Blind Signal Separation via Tensor Decomposition With Vandermonde Factor: Canonical Polyadic Decomposition

Mikael Sørensen and Lieven De Lathauwer, *Senior Member, IEEE*

***Abstract*—Several problems in signal processing have been formulated in terms of the Canonical Polyadic Decomposition of a higher-order tensor with one or more Vandermonde constrained factor matrices. We first propose new, relaxed uniqueness conditions. We explain that, under these conditions, the number of components may simply be estimated as the rank of a matrix. We propose an efficient algorithm for the computation of the factors that only resorts to basic linear algebra. We demonstrate the use of the results for various applications in wireless communication and array processing.**

***Index Terms*—Array processing, Polyadic Decomposition, tensor, Vandermonde matrix, wireless communication.**

I. Introduction

CANONICAL Polyadic Decomposition (CPD) of a higher-order tensor is its decomposition in a minimal number of rank-1 terms. Compared to decomposition of a matrix in rank-1 terms, CPD is unique under mild conditions [7]. This makes CPD a natural tool for signal separation [4]. In particular, orthogonality constraints like in Singular Value Decomposition (SVD) are not *per se* required for uniqueness. Nevertheless, imposing such constraints may increase the number of rank-1 terms for which uniqueness can be established, increase the accuracy of estimates, reduce the computational cost, facilitate interpretability, and so on [38]. In this paper we consider CPD with a different type of constraint, namely Vandermonde structure. We mean that the columns of at least one factor matrix take the form of exponentials. This is relevant for a broad class of signal processing problems, as is clear from the extensive literature on one-dimensional [27], [28] and multidimensional harmonic retrieval [18], [23], [25], [32]. We also mention that the Vandermonde structure may result from the use of a sinusoidal carrier in telecommunication applications, or from the use of a Uniform Linear Array (ULA) in array processing [33], [42], [45] and in wireless communication [8], [24]. The Vandermonde structure may also be caused by Doppler effects, carrier frequency offset or small delay spread in array processing and in wireless communication [16], [26], [35].

The paper is organized as follows. The rest of the introduction presents our notation, followed by a brief review of the CPD. In Section II some properties of Vandermonde constrained CPD and some techniques for handling exponential signals are briefly reviewed. Section III consists of five parts. In Section III-A we present new uniqueness results for CPD with a Vandermonde factor matrix. The derivation is constructive in the sense that it also provides us with an algorithm for the actual computation of the factors. The method can be understood as a generalization of well-known ESPRIT [29] from the matrix to the tensor case. The derivation also shows that in the noise-free case (or for a sufficiently high Signal-to-Noise Ratio (SNR)), the number of rank-1 terms can be estimated as the rank of a matrix. This is to be contrasted with the general CPD setting, in which the estimation of the number of components is NP-hard [14], [15]. The rank result is presented in Section III-B and the algorithm in Section III-C. In Section III-D we discuss situations in which more than one factor matrix is Vandermonde constrained. Section III-E pays special attention to the case of Vandermonde vectors with multiplicity larger than one. In Section IV–Section VII we discuss several applications in wireless communication and array processing. Section IV concerns MIMO radar. Section V concerns the blind separation of signals impinging on a polarization sensitive array. In Section VI we explain the use in wireless communication systems employing a ULA at the receiver. Section VII explains that the results lead to improved identifiability results in MIMO OFDM with carrier frequency offset. Simulations are reported in Section VIII. We conclude in Section IX.

Part of this work was presented in the conference paper [37].

Notation

Vectors, matrices and tensors are denoted by lower case boldface, upper case boldface and upper case calligraphic letters, respectively. The $r\text{th}$ column vector of $\mathbf{A}$ is denoted by $\mathbf{a}_r$. The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri-Rao products, defined as

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix},$$

$$\mathbf{A} \odot \mathbf{B} \triangleq [\, \mathbf{a}_1 \otimes \mathbf{b}_1 \quad \mathbf{a}_2 \otimes \mathbf{b}_2 \quad \dots \,],$$

Manuscript received November 16, 2012; revised April 01, 2013 and July 08, 2013; accepted July 19, 2013. Date of publication August 01, 2013; date of current version October 08, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chong-Yung Chi. This work was supported by the Research Council KU Leuven: GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT1/08/23, F.W.O.: project G.0427.10N, the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012–2017).

The authors are with the Electrical Engineering Department (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, and iMinds Future Health Department, KU Leuven, B-3001 Leuven-Heverlee, Belgium, and also with The Group Science, Engineering and Technology, KU Leuven, Kulak 8500 Kortrijk, Belgium (e-mail: Mikael.Sorensen@kuleuven-kulak.be; Lieven.DeLathauwer@kuleuven-kulak.be).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2276416

1053-587X © 2013 IEEE

in which $(\mathbf{A})_{mn} = a_{mn}$. Sometimes we denote the Khatri-Rao product of N matrices $\{\mathbf{A}^{(n)}\}_{n=1}^{N}$ by

$$\bigodot_{n=1}^{N} \mathbf{A}^{(n)} = \mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)}.$$

The Hadamard product is denoted by $*$ and satisfies $(\mathbf{A} * \mathbf{B})_{mn} = a_{mn} b_{mn}$. The outer product is denoted by $\circ$, such that $\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ satisfies

$$\left(\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)}\right)_{i_1, i_2, \ldots, i_N} = a_{i_1}^{(1)} a_{i_2}^{(2)} \cdots a_{i_N}^{(N)}.$$

We sometimes denote the Cartesian product of N sets $\{\mathbb{C}^{I_n}\}_{n=1}^{N}$ by

$$\mathbb{C}^{\overset{N}{\underset{n=1}{\times}} I_n} = \mathbb{C}^{I_1 \times \cdots \times I_N}.$$

Matlab index notation will be used for submatrices of a given matrix. For example, $\mathbf{A}(1:k,:)$ represents the submatrix of $\mathbf{A}$ consisting of the rows from 1 to k of $\mathbf{A}$. Let $\mathbf{T} \in \mathbb{C}^{I \times R}$, then $\overline{\mathbf{T}} = \mathbf{T}(2:I,:) \in \mathbb{C}^{(I-1) \times R}$ and $\underline{\mathbf{T}} = \mathbf{T}(1:I-1,:) \in \mathbb{C}^{(I-1) \times R}$, i.e., $\overline{\mathbf{T}}$ and $\underline{\mathbf{T}}$ are obtained by deleting the top and bottom row of $\mathbf{T}$, respectively.

The transpose, conjugate, conjugate-transpose, inverse, Moore-Penrose pseudo-inverse, Frobenius norm and column space of a matrix are denoted by $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, $\|\cdot\|_F$ and $\mathrm{Col}(\cdot)$, respectively. The identity matrix and all-ones vector are denoted by $\mathbf{I}_R \in \mathbb{C}^{R \times R}$ and $\mathbf{1}_R = [1, \ldots, 1]^T \in \mathbb{C}^R$, respectively.

Given $\mathbf{A} \in \mathbb{C}^{I \times J}$, then $\mathrm{Vec}(\mathbf{A}) \in \mathbb{C}^{IJ}$ will denote the column vector defined by $(\mathrm{Vec}(\mathbf{A}))_{i+(j-1)I} = (\mathbf{A})_{ij}$. Given $\mathbf{a} \in \mathbb{C}^{IJ}$, the reverse operation is $\mathrm{Unvec}(\mathbf{a}) = \mathbf{A} \in \mathbb{C}^{I \times J}$, such that $(\mathbf{a})_{i+(j-1)I} = (\mathbf{A})_{ij}$. $D_k(\mathbf{A}) \in \mathbb{C}^{J \times J}$ denotes the diagonal matrix holding row k of $\mathbf{A} \in \mathbb{C}^{I \times J}$ on its diagonal.

The rank of $\mathbf{A}$ is denoted by $r(\mathbf{A})$. The k-rank of a matrix $\mathbf{A}$ is denoted by $k(\mathbf{A})$. It is equal to the largest integer $k(\mathbf{A})$ such that every subset of $k(\mathbf{A})$ columns of $\mathbf{A}$ is linearly independent.

The k-th compound matrix of $\mathbf{A} \in \mathbb{C}^{I \times R}$ is denoted by $C_k(\mathbf{A}) \in \mathbb{C}^{C_I^k \times C_R^k}$, where $C_m^k = \frac{m!}{k!(m-k)!}$. It is the matrix containing the determinants of all $k \times k$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order, see [6] for a detailed explanation.

A. Canonical Polyadic Decomposition (CPD)

Given is an Nth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$. We say that $\mathcal{X}$ is a rank-1 tensor if it equals the outer product of some non-zero vectors $\mathbf{a}^{(n)} \in \mathbb{C}^{I_n}$, $n \in \{1, \ldots, N\}$, such that $\mathcal{X}_{i_1 \ldots i_N} = \prod_{n=1}^{N} a_{i_n}^{(n)}$. Decompositions into rank-1 terms are called Polyadic Decompositions (PD):

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)}. \tag{1}$$

The rank of a tensor $\mathcal{X}$ is equal to the minimal number of rank-1 tensors that yield $\mathcal{X}$ in a linear combination. Assume that the rank of $\mathcal{X}$ is R, then (1) is called the Canonical PD (CPD) of $\mathcal{X}$. Let us stack the vectors $\{\mathbf{a}_r^{(n)}\}$ into the matrices

$$\mathbf{A}^{(n)} = \left[\mathbf{a}_1^{(n)}, \quad \ldots \quad, \mathbf{a}_R^{(n)}\right] \in \mathbb{C}^{I_n \times R}, \quad n \in \{1, \ldots, N\}.$$

The matrices $\mathbf{A}^{(n)}$ will be referred to as factor matrices.

The following matrix representation of a PD or CPD of $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ will be used throughout the paper. Let $\mathbf{X}^{(i_1 \cdot \cdot)} \in \mathbb{C}^{I_2 \times I_3}$ denote the matrix such that $\mathbf{X}_{i_2 i_3}^{(i_1 \cdot \cdot)} = \mathcal{X}_{i_1 i_2 i_3}$, then $\mathbf{X}_{(1)} \in \mathbb{C}^{I_1 I_2 \times I_3}$ is given by

$$\begin{aligned} \mathbf{X}_{(1)} &\triangleq \left[\mathbf{X}^{(1 \cdot \cdot)T}, \ldots, \mathbf{X}^{(I_1 \cdot \cdot)T}\right]^T \\ &= \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right) \mathbf{A}^{(3)T}. \end{aligned} \tag{2}$$

Similarly, let the matrix $\mathbf{X}^{(\cdot i_2 \cdot)} \in \mathbb{C}^{I_3 \times I_1}$ be constructed such that $\mathbf{X}_{i_3 i_1}^{(\cdot i_2 \cdot)} = \mathcal{X}_{i_1 i_2 i_3}$, then $\mathbf{X}_{(2)} \in \mathbb{C}^{I_2 I_3 \times I_1}$ is given by

$$\begin{aligned} \mathbf{X}_{(2)} &\triangleq \left[\mathbf{X}^{(\cdot 1 \cdot)T}, \ldots, \mathbf{X}^{(\cdot I_2 \cdot)T}\right]^T \\ &= \left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(3)}\right) \mathbf{A}^{(1)T}. \end{aligned} \tag{3}$$

For Nth-order tensors, the following matrix representation will also be used throughout the paper. Consider $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R, then

$$\begin{aligned} &\mathbf{X}^{[P]} \\ &\triangleq \begin{bmatrix} \mathcal{X}_{1,\ldots,1,1,\ldots,1} & \mathcal{X}_{1,\ldots,1,1,\ldots,2} & \cdots & \mathcal{X}_{1,\ldots,1,I_{P+1},\ldots,I_N} \\ \mathcal{X}_{1,\ldots,2,1,\ldots,1} & \mathcal{X}_{1,\ldots,2,1,\ldots,2} & \cdots & \mathcal{X}_{1,\ldots,2,I_{P+1},\ldots,I_N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}_{I_1,\ldots,I_P,1,\ldots,1} & \mathcal{X}_{I_1,\ldots,I_P,1,\ldots,2} & \cdots & \mathcal{X}_{I_1,\ldots,I_P,I_{P+1},\ldots,I_N} \end{bmatrix} \\ &= \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(P)}\right) \left(\mathbf{A}^{(P+1)} \odot \cdots \odot \mathbf{A}^{(N)}\right)^T. \end{aligned} \tag{4}$$

A CPD of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is said to be unique if the rank-1 terms are unique up to permutation of the terms and scaling/counterscaling of vectors within the same term. Formally, the CPD is unique if all N-tuplets $\left(\widehat{\mathbf{A}}^{(1)}, \ldots, \widehat{\mathbf{A}}^{(N)}\right)$ satisfying (1) are related via

$$\widehat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \mathbf{P} \Delta_{\mathbf{A}^{(n)}}, \quad \forall n \in \{1, \ldots, N\},$$

where $\mathbf{P}$ is a permutation matrix and $\{\Delta_{\mathbf{A}^{(n)}}\}$ are diagonal matrices satisfying $\prod_{n=1}^{N} \Delta_{\mathbf{A}^{(n)}} = \mathbf{I}_R$. Unlike the decomposition of a matrix in rank-1 terms, CPD is unique under mild conditions, which makes it an important tool for blind signal separation. A well-known uniqueness result is the following.

Theorem I.1: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation (2). If

$$k\left(\mathbf{A}^{(1)}\right) + k\left(\mathbf{A}^{(2)}\right) + k\left(\mathbf{A}^{(3)}\right) \geq 2R + 2 \tag{5}$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique [20], [41].

In the generic case[1], condition (5) becomes

$$\min(I_1, R) + \min(I_2, R) + \min(I_3, R) \geq 2R + 2. \tag{6}$$

Condition (5) is sufficient but not necessary. For instance, for the case where one of the factor matrices of the CPD has full column rank, say $\mathbf{A}^{(3)}$, the following more relaxed uniqueness condition has been obtained [5], [7], [17], [39].

[1]A CPD property is called generic if it holds with probability one when the entries of the factor matrices are drawn from absolutely continuous probability density functions.

Theorem I.2: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation (2). If

$$\begin{cases} \mathbf{A}^{(3)} \text{ has full column rank,} \\ C_2\left(\mathbf{A}^{(1)}\right) \odot C_2\left(\mathbf{A}^{(2)}\right) \text{ has full column rank,} \end{cases} \tag{7}$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique. Generically, condition (7) is satisfied if

$$2R(R-1) \leq I_1(I_1-1)I_2(I_2-1) \quad \text{and} \quad R \leq I_3\,. \tag{8}$$

An extension of Theorem I.2 to higher-order tensors ($N > 3$) can be found in [40]. Under the conditions in Theorem I.2 the CPD can be computed by means of linear algebra [5]. In general, one resorts to optimization-based methods [36].

One has recently obtained a relaxation of Theorem I.2 to the case where none of the involved factor matrices needs to have full column rank [6], [7].

Theorem I.3: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation (2). If the condition in the equation at the bottom of the page is satisfied, then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique.

II. CPD With Vandermonde Factor Matrix

The factor matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ is said to be Vandermonde if

$$\begin{aligned} \mathbf{A}^{(n)} &= \left[\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}\right] \\ \mathbf{a}_r^{(n)} &= \left[1, z_{n,r}, z_{n,r}^2, \ldots, z_{n,r}^{I_n-1}\right]^T, \end{aligned} \tag{9}$$

where the scalars $\{z_{n,r}\}$ are called the generators of $\mathbf{A}^{(n)}$. If some of the factor matrices are Vandermonde, then this may obviously increase the minimal R in (1). This minimal value will be denoted by $r_{\mathrm{VDM}}(\mathcal{X})$. If some of the factor matrices are Vandermonde and $R = r_{\mathrm{VDM}}(\mathcal{X})$, then (1) will be called a Vandermonde constrained CPD of $\mathcal{X}$.

In a Vandermonde constrained CPD the scaling/counter-scaling indeterminacies do not involve the Vandermonde factors. Consider $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with rank R and constructed from the factor matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$, $n \in \{1, \ldots, N\}$. Assume that the factors $\mathbf{A}^{(p)}$, $1 \leq p \leq P \leq N$ are Vandermonde matrices. Let the N-tuplet $\left(\widehat{\mathbf{A}}^{(1)}, \ldots, \widehat{\mathbf{A}}^{(N)}\right)$ yield an alternative decomposition of $\mathcal{X}$. Formally, the Vandermonde constrained CPD of $\mathcal{X}$ is said to be unique if

$$\begin{aligned} \widehat{\mathbf{A}}^{(p)} &= \mathbf{A}^{(p)}\mathbf{P}\,, \quad \forall p \in \{1, \ldots, P\} \\ \widehat{\mathbf{A}}^{(q)} &= \mathbf{A}^{(q)}\mathbf{P}\Delta_q\,, \quad \forall q \in \{P+1, \ldots, N\}\,, \end{aligned}$$

where $\mathbf{P} \in \mathbb{C}^{R \times R}$ is a permutation matrix and $\Delta_q \in \mathbb{C}^{R \times R}$ are diagonal matrices with property $\prod_{q=P+1}^{N} \Delta_q = \mathbf{I}_R$.

Section II-A reviews existing uniqueness conditions for a CPD with a Vandermonde factor matrix. Next, Section II-B presents a direct variant of ESPRIT [29]. In Section II-C we discuss the common technique of spatial smoothing.

A. Existing Uniqueness Results

By combining Vandermonde structure with Theorem I.1, the following uniqueness result for a CPD with a Vandermonde factor matrix was obtained in [35]:

Theorem II.1: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation (2). Assume that $\mathbf{A}^{(1)}$ is a Vandermonde matrix with distinct generators. If

$$\min\left(2R, I_1 + k\big(\mathbf{A}^{(3)}\big)\right) + k\left(\mathbf{A}^{(2)}\right) \geq 2(R+1), \tag{10}$$

then $R = r_{\mathrm{VDM}}(\mathcal{X})$ and the Vandermonde constrained CPD of $\mathcal{X}$ is unique. In the generic case, condition (10) becomes

$$\min\Big(2R, I_1 + \min\left(I_3, R\right)\Big) + \min\left(I_2, R\right) \geq 2(R+1).$$

For the case where $\mathbf{A}^{(3)}$ additionally has full column rank, the following result was obtained in [11]:

Theorem II.2: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation (2). Assume that $\mathbf{A}^{(1)}$ is a Vandermonde matrix with distinct generators. If

$$\mathrm{r}\left(\mathbf{A}^{(3)}\right) = R \quad \text{and} \quad I_1 + r\big(\mathbf{A}^{(2)}\big) - 2 \geq R\,, \tag{11}$$

then $R = r_{\mathrm{VDM}}(\mathcal{X})$ and the Vandermonde constrained CPD of $\mathcal{X}$ is unique. In the generic case, condition (11) becomes

$$\min\Big(I_3, I_1 + \min\left(I_2, R\right) - 2\Big) \geq R.$$

B. ESPRIT

If $\mathbf{A}^{(1)}$ is a tall ($I_1 > R$) Vandermonde matrix and $\mathbf{A}^{(2)} \odot \mathbf{A}^{(3)}$ has full column rank, then we may just ignore the Khatri-Rao structure, compute $\mathbf{A}^{(1)}$ by means of classical ESPRIT [29] and then compute $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$, as explained in the following theorem.

Theorem II.3: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation (3). Assume that $\mathbf{A}^{(1)}$ is a Vandermonde matrix with distinct generators. If

$$I_1 - 1 \geq R \quad \text{and} \quad r\left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(3)}\right) = R\,, \tag{12}$$

$$\begin{cases} \min\left(k\left(\mathbf{A}^{(1)}\right) + k\left(\mathbf{A}^{(3)}\right), k\left(\mathbf{A}^{(2)}\right) + k\left(\mathbf{A}^{(3)}\right)\right) \geq R+1, \\ \max\left(k\left(\mathbf{A}^{(1)}\right) + k\left(\mathbf{A}^{(3)}\right), k\left(\mathbf{A}^{(2)}\right) + k\left(\mathbf{A}^{(3)}\right)\right) \geq R+2, \\ C_{R-\mathrm{r}\left(\mathbf{A}^{(3)}\right)+2}\left(\mathbf{A}^{(1)}\right) \odot C_{R-\mathrm{r}\left(\mathbf{A}^{(3)}\right)+2}\left(\mathbf{A}^{(2)}\right) \text{ has full column rank,} \end{cases}$$

then $R = r_{\mathrm{VDM}}(\mathcal{X})$ and the Vandermonde constrained CPD of $\mathcal{X}$ is unique. In the generic case, condition (12) becomes

$$\min(I_1 - 1, I_2 I_3) \geq R. \tag{13}$$

Proof: As explained in [29], we can under condition (12) find R as $r\left(\mathbf{X}_{(2)}\right)$ and determine $\mathbf{A}^{(1)}$ from $\mathrm{Col}\left(\mathbf{X}_{(2)}^T\right)$. Once $\mathbf{A}^{(1)}$ has been found, we compute $\mathbf{Y} = \mathbf{X}_{(2)}\left(\mathbf{A}^{(1)T}\right)^{\dagger}$. The remaining factor matrices now trivially follow from rank-1 approximation problems:

$$\min_{\mathbf{a}_r^{(2)}, \mathbf{a}_r^{(3)}} \left\| \mathbf{y}_r - \mathbf{a}_r^{(2)} \otimes \mathbf{a}_r^{(3)} \right\|_F^2, \quad r \in \{1, \ldots, R\}.$$

The generic condition (13) follows from the same arguments, taking into account the Khatri-Rao product of unstructured matrices generically has full rank [5], [18]. ∎

C. Spatial Smoothing

Spatial smoothing is a technique commonly applied in sensor array processing to overcome the problems that may occur when some of the involved matrices are rank deficient [31]. The technique has been extended to higher-order tensors in for instance [32], [35], [42]. Consider a tensor $\mathcal{X} \in \mathbb{C}^{\overset{N}{\underset{n=1}{\times}} I_n}$ with PD (1), where we assume that the factor matrices $\mathbf{A}^{(p)} \in \mathbb{C}^{I_p \times R}$, $1 \leq p \leq P < N$, are Vandermonde matrices. From (4) we obtain the following matrix representation of (1):

$$\mathbb{C}^{\prod_{p=1}^{P} I_p \times \prod_{q=P+1}^{N} I_q} \ni \mathbf{X}^{[P]} = \left(\bigodot_{p=1}^{P} \mathbf{A}^{(p)} \right) \left(\bigodot_{q=P+1}^{N} \mathbf{A}^{(q)} \right)^T. \tag{14}$$

If the matrix $\bigodot_{q=P+1}^{N} \mathbf{A}^{(q)}$ does not have full column rank, i.e., its rank is less than R, then the matrix $\mathbf{X}^{[P]}$ also has rank less than R. The goal of spatial smoothing is to obtain a matrix that has rank R.

Using $z_{p,r}^{l_p + k_p - 2} = z_{p,r}^{l_p - 1} z_{p,r}^{k_p - 1}$, spatial smoothing maps $\mathbf{A}^{(p)} \in \mathbb{C}^{I_p \times R}$ to $\mathbf{A}^{(K_p,p)} \mathbf{A}^{(L_p,p)T} \in \mathbb{C}^{K_p \times L_p}$, where $\mathbf{A}^{(K_p,p)} \in \mathbb{C}^{K_p \times R}$, $\mathbf{A}^{(L_p,p)} \in \mathbb{C}^{L_p \times R}$, increasing the order by one for each $p \in \{1, \ldots, P\}$. Thus, the total increase of order is P. Choose $Q \in \{P+1, \ldots, N\}$ and construct a tensor

$$\mathcal{Y} \in \mathbb{C}^{\left(\overset{P}{\underset{p=1}{\times}} K_p\right) \times \left(\overset{Q}{\underset{q=P+1}{\times}} I_q\right) \times \left(\overset{P}{\underset{p=1}{\times}} L_p\right) \times \left(\overset{N}{\underset{q=Q+1}{\times}} I_q\right)}$$

as follows:

$$\begin{aligned} \mathcal{Y}_{k_1,\ldots,k_P,i_{P+1},\ldots,i_Q,l_1,\ldots,l_P,i_{Q+1},\ldots,i_N} &= \mathcal{X}_{l_1+k_1-1,\ldots,l_P+k_P-1,i_{P+1},\ldots,i_N} \\ &= \sum_{r=1}^{R} \prod_{p=1}^{P} z_{p,r}^{l_p+k_p-2} \prod_{q=P+1}^{N} a_{i_q,r}^{(q)} \\ &= \sum_{r=1}^{R} \prod_{s=1}^{P} z_{s,r}^{l_s-1} \prod_{t=1}^{P} z_{t,r}^{k_t-1} \prod_{q=P+1}^{N} a_{i_q,r}^{(q)}, \end{aligned}$$

where $l_p \in \{1, \ldots, L_p\}$, $k_p \in \{1, \ldots, K_p\}$ and the integers K_p and L_p must be chosen such that $L_p + K_p = I_p + 1$, $p \in \{1, \ldots, P\}$. From matrix representation (4) we obtain

$$\mathbf{Y}^{[P]} = \left(\bigodot_{p=1}^{P} \mathbf{A}^{(K_p,p)} \odot \bigodot_{q=P+1}^{Q} \mathbf{A}^{(q)} \right) \times \left(\bigodot_{p=1}^{P} \mathbf{A}^{(L_p,p)} \odot \bigodot_{q=Q+1}^{N} \mathbf{A}^{(q)} \right)^T \tag{15}$$

with $\mathbf{A}^{(K_p,p)} \in \mathbb{C}^{K_p \times R}$ and $\mathbf{A}^{(L_p,p)} \in \mathbb{C}^{L_p \times R}$ given by

$$\begin{aligned} \mathbf{A}^{(K_p,p)} &= \left[\mathbf{a}_1^{(K_p,p)}, \ldots, \mathbf{a}_R^{(K_p,p)} \right], \\ \mathbf{a}_r^{(K_p,p)} &= \left[1, z_{p,r}, \ldots, z_{p,r}^{K_p-1} \right]^T, \\ \mathbf{A}^{(L_p,p)} &= \left[\mathbf{a}_1^{(L_p,p)}, \ldots, \mathbf{a}_R^{(L_p,p)} \right], \\ \mathbf{a}_r^{(L_p,p)} &= \left[1, z_{p,r}, \ldots, z_{p,r}^{L_p-1} \right]^T. \end{aligned}$$

In this way, a matrix representation (14) of the tensor $\mathcal{X}$ that involves a rank deficient matrix $\bigodot_{q=P+1}^{N} \mathbf{A}^{(q)}$, may possibly be replaced by a matrix representation (15) that only involves full column rank matrices $\bigodot_{p=1}^{P} \mathbf{A}^{(K_p,p)} \odot \bigodot_{q=P+1}^{Q} \mathbf{A}^{(q)}$ and $\bigodot_{p=1}^{P} \mathbf{A}^{(L_p,p)} \odot \bigodot_{q=Q+1}^{N} \mathbf{A}^{(q)}$.

III. New Results For Vandermonde Structured CPD

Section III-A presents new uniqueness conditions for a CPD with a Vandermonde factor matrix. Section III-B briefly discusses the estimation of the rank. Section III-C provides an algorithm for the actual computation. Section III-D discusses extensions to cases where more than one factor matrix is Vandermonde. In Section III-E we discuss the special case where some of the generators of the Vandermonde matrix are repeated.

A. Uniqueness

For the case where $\mathbf{A}^{(1)}$ is Vandermonde and $\mathbf{A}^{(3)}$ has full column rank, we obtain a uniqueness result that is more relaxed than Theorems II.1 and II.3. This result is given as Proposition III.2 below. The proof of the deterministic part is a Khatri-Rao variant of the derivation of ESPRIT [29]. The proof of the generic version is based on the following lemma.

Lemma III.1: Let $\mathbf{A}^{(1)} \in \mathbb{C}^{I_1 \times R}$ be a Vandermonde matrix and let $\mathbf{A}^{(2)} \in \mathbb{C}^{I_2 \times R}$, then the matrix $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}$ generically has rank $\min(I_1 I_2, R)$.

Proof: The result follows from Theorem 3 and Corollary 1 in [18]. ∎

Proposition III.2: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ admit a PD with matrix representation (2). Let $\mathbf{A}^{(1)}$ be a Vandermonde matrix with distinct generators $\{z_r\}_{r=1}^{R}$. If

$$\begin{cases} r\left(\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}\right) = R \\ r\left(\mathbf{A}^{(3)}\right) = R, \end{cases} \tag{16}$$

then $R = r_{\mathrm{VDM}}(\mathcal{X})$ and the Vandermonde constrained CPD of $\mathcal{X}$ is unique. Generically, condition (16) is satisfied if and only if

$$\min\left(\left(I_1 - 1\right) I_2, I_3\right) \geq R\,. \tag{17}$$

Proof: Consider the matrix representation $\mathbf{X}_{(1)} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{A}^{(3)T}$ of $\mathcal{X}$ and let $\mathbf{X}_{(1)} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact Singular Value Decomposition (SVD) of $\mathbf{X}_{(1)}$, in which $\mathbf{U} \in \mathbb{C}^{I_1 I_2 \times R}$, $\Sigma \in \mathbb{C}^{R\times R}$ and $\mathbf{V} \in \mathbb{C}^{I_3 \times R}$ with $R = r\left(\mathbf{X}_{(1)}\right)$. Since we assume that the matrices $\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$ have full column rank, we know that there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{U}\mathbf{M} = \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}. \tag{18}$$

Due to the Vandermonde structure of $\mathbf{A}^{(1)}$ we have

$$\left(\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{Z} = \overline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}, \tag{19}$$

where $\mathbf{Z} = D_1\left([z_1, \ldots, z_R]\right)$. The Vandermonde structure of $\mathbf{A}^{(1)}$ and relation (18) imply that

$$\mathbf{U}_1\mathbf{M} = \underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)} \tag{20}$$

$$\mathbf{U}_2\mathbf{M} = \overline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}, \tag{21}$$

where

$$\mathbf{U}_1 = \mathbf{U}\left(1 : (I_1 - 1)I_2, :\right) \in \mathbb{C}^{(I_1-1)I_2 \times R}$$
$$\mathbf{U}_2 = \mathbf{U}\left(I_2 + 1 : I_1 I_2, :\right) \in \mathbb{C}^{(I_1-1)I_2 \times R}.$$

By combining (19), (20) and (21) we obtain the following equalities:

$$\mathbf{U}_2\mathbf{M} = \overline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)} = \left(\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{Z} = \mathbf{U}_1\mathbf{M}\mathbf{Z}. \tag{22}$$

Thus, from (22) we have $\mathbf{U}_2 = \mathbf{U}_1\widehat{\mathbf{Z}}$, where $\widehat{\mathbf{Z}} = \mathbf{M}\mathbf{Z}\mathbf{M}^{-1}$. Assuming that $\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}$ has full column rank, $\mathbf{U}_1$ and $\mathbf{U}_2$ have full column rank and $\widehat{\mathbf{Z}} = \mathbf{U}_1^{\dagger}\mathbf{U}_2$. From the EigenValue Decomposition (EVD) $\widehat{\mathbf{Z}} = \mathbf{M}\mathbf{Z}\mathbf{M}^{-1}$ the generators $\{z_r\}_{r=1}^{R}$, $\mathbf{A}^{(1)}$ and $\mathbf{M}$ are obtained, up to permutation ambiguities.

Given $\mathbf{A}^{(1)}$ and $\mathbf{M}$, the next step is to find $\mathbf{A}^{(2)}$. Note that $\left(\frac{\mathbf{a}_r^{(1)H}}{\mathbf{a}_x^{(1)H}\mathbf{a}_y^{(1)}} \otimes \mathbf{I}_{I_2}\right)\left(\mathbf{a}_r^{(1)} \otimes \mathbf{a}_r^{(2)}\right) = \mathbf{a}_r^{(2)}$. Due to the scaling ambiguity, we can without loss of generality assume that the column vectors of $\mathbf{A}^{(1)}$ have unit norm. Thus, $\mathbf{a}_r^{(2)}$ follows from (18) as follows:

$$\mathbf{a}_r^{(2)} = \left(\mathbf{a}_r^{(1)H} \otimes \mathbf{I}_{I_2}\right)\mathbf{U}\mathbf{m}_r\,, \quad r \in \{1, \ldots, R\}\,. \tag{23}$$

Given $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$, $\mathbf{A}^{(3)}$ follows directly from [19]:

$$\mathbf{A}^{(3)T} = \left(\left(\mathbf{A}^{(1)H}\mathbf{A}^{(1)}\right) * \left(\mathbf{A}^{(2)H}\mathbf{A}^{(2)}\right)\right)^{-1} \times \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)^{H}\mathbf{X}_{(1)}\,.$$

TABLE I
MAXIMUM VALUE R FOR WHICH THEOREMS 1.1, 1.2, II.1, II.2 AND II.3 AND PROPOSITION III.2 GUARANTEE UNIQUENESS IN THE GENERIC CASE OF A CPD WITH $I_1 = I_2, I_3 \geq R$ AND $\mathbf{A}^{(1)}$ VANDERMONDE

$I_1 = I_2$	2	3	4	5	6	7	8	9	10
Theorem I.1	**2**	4	6	8	10	12	14	16	18
Theorem I.2	**2**	4	9	14	21	30	40	51	64
Theorem II.1	**2**	4	6	8	10	12	14	16	18
Theorem II.2	**2**	4	6	8	10	12	14	16	18
Theorem II.3	1	2	3	4	5	6	7	8	9
Proposition III.2	**2**	**6**	**12**	**20**	**30**	**42**	**56**	**72**	**90**

The generic condition (17) follows along the same lines, taking Lemma III.1 into account. ■

Note that Proposition III.2 does not prevent $k\left(\mathbf{A}^{(2)}\right) = 1$. This is in contrast with the unconstrained CPD, where $k\left(\mathbf{A}^{(n)}\right) \geq 2$, $n \in \{1, 2, 3\}$, is necessary for uniqueness (e.g., [41]). Uniqueness condition (16) in Proposition III.2 may also be derived via spatial smoothing. Indeed, consider the matrix representation $\mathbf{X} = \left(\mathbf{A}^{(I_1-1,1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(2,1)} \odot \mathbf{A}^{(3)}\right)^T = \left(\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(2,1)} \odot \mathbf{A}^{(3)}\right)^T$, obtained by spatial smoothing of $\mathcal{X}$. The matrix $\mathbf{X}$ may be interpreted as a matrix representation of a third-order tensor with two full column rank factor matrices, $\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$, and third factor matrix $\mathbf{A}^{(2,1)}$, with property $k\left(\mathbf{A}^{(2,1)}\right) \geq 2$. Theorem I.1 now tells us that $R = r_{\mathrm{VDM}}(\mathcal{X})$ and that the Vandermonde constrained CPD of $\mathcal{X}$ is unique.

When $\mathbf{A}^{(3)}$ has full column rank, then Theorem II.2 is more relaxed than Theorem II.1, which in turn is more relaxed than Theorem I.1. We now explain that the uniqueness condition provided by Proposition III.2 is more relaxed than the one stated in Theorem II.2. Theorem II.2 requires that the generators of $\mathbf{A}^{(1)}$ are distinct and that the inequality $I_1 - 1 + r\left(\mathbf{A}^{(2)}\right) \geq R + 1$ is satisfied. Recall that if $k\left(\mathbf{A}^{(1)}\right) + r\left(\mathbf{A}^{(2)}\right) \geq R + 1$, then $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}$ has full column rank [11]. Consequently the condition $I_1 - 1 + r\left(\mathbf{A}^{(2)}\right) \geq R + 1$ implies that $\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}$ has full column rank. Hence, the conditions in Theorem II.2 imply the conditions in Proposition III.2. However, the converse is not true.

Table I shows a few values for the case where $\mathbf{A}^{(3)}$ has full column rank and $\mathbf{A}^{(1)}$ is Vandermonde. It is clear that Theorem I.1 is not appropriate in this case. Theorems II.1 and II.2 yield the same result as Theorem I.1. Theorem I.2 yields a more relaxed uniqueness condition than Theorems 1.1, II.1 and II.2. The reason is that Theorem I.2 takes into account that $\mathbf{A}^{(3)}$ has full column rank. Finally, we notice that the new Proposition III.2 yields an even more relaxed uniqueness condition than Theorem I.2. The reason is that Proposition III.2 simultaneously takes into account that $\mathbf{A}^{(1)}$ is Vandermonde and that $\mathbf{A}^{(3)}$ has full column rank. In Theorem II.3 R is bounded by the dimension of the Vandermonde factor. Note also that (17) is more relaxed than (8).

We may also combine Proposition III.2 with spatial smoothing. In this way we can obtain uniqueness results in cases where $\mathbf{A}^{(3)}$ does not have full column rank. We have the following uniqueness result.

Theorem III.3: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ be a tensor with matrix representation $\mathbf{X}_{(1)} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{A}^{(3)T}$, where $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ and $\mathbf{A}^{(1)}$ is a Vandermonde matrix with

distinct generators $\{z_r\}_{r=1}^{R}$. Consider the matrix representation $\mathbf{X} = \left(\mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}\right)^T$ with $K_1 + L_1 = I_1 + 1$. If

$$\begin{cases} r\left(\underline{\mathbf{A}}^{(K_1,1)} \odot \mathbf{A}^{(2)}\right) = R \\ r\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}\right) = R \end{cases} \tag{24}$$

for some $K_1 + L_1 = I_1 + 1$, then $R = r_{\mathrm{VDM}}(\mathcal{X})$ and the Vandermonde constrained CPD of $\mathcal{X}$ is unique. Generically, condition (24) is satisfied if and only if

$$\left\lceil \frac{R}{I_2} \right\rceil + \left\lceil \frac{R}{I_3} \right\rceil \leq I_1 . \tag{25}$$

Proof: Let us first prove condition (24). Let $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{X}$, in which $\mathbf{U} \in \mathbb{C}^{K_1 I_2 \times R}$, $\Sigma \in \mathbb{C}^{R \times R}$ and $\mathbf{V} \in \mathbb{C}^{L_1 I_3 \times R}$ with $R = r(\mathbf{X})$, then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\mathbf{M} = \mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(2)} \tag{26}$$

$$\mathbf{V}^*\Sigma\mathbf{N} = \mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}, \quad \mathbf{N} = \mathbf{M}^{-T}. \tag{27}$$

From Proposition III.2 we know that $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ and $\mathbf{M}$ follow from (26). From (27) we obtain $\mathbf{A}^{(3)}$ as follows:

$$\mathbf{a}_r^{(3)} = \left(\frac{\mathbf{a}_r^{(L_1,1)H}}{\mathbf{a}_r^{(L_1,1)H}\mathbf{a}_r^{(L_1,1)}} \otimes \mathbf{I}_{I_3} \right) \times \mathbf{V}^*\Sigma\mathbf{n}_r , \quad r \in \{1, \ldots, R\} .$$

Let us now prove that (25) is generically equivalent to (24). From Lemma III.1 we know that condition (24) is generically true if

$$\min\left((K_1 - 1)I_2, L_1 I_3\right) \geq R , \tag{28}$$

i.e., $K_1 - 1 \geq \left\lceil \frac{R}{I_2} \right\rceil$ and $L_1 \geq \left\lceil \frac{R}{I_3} \right\rceil$ in which K_1 and L_1 can be chosen such that $K_1 - 1 + L_1 = I_1$. This is equivalent to (25). ∎

Uniqueness condition (24) in Theorem III.3 may also be derived via an additional spatial smoothing of $\mathbf{X} = \left(\mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}\right)^T$ so that the matrix $\mathbf{Y} = \left(\mathbf{A}^{(K_1-1,1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(2,1)} \odot \mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}\right)^T$ is obtained. The matrix $\mathbf{Y}$ can be interpreted as a matrix representation of a third-order tensor with two full column rank factor matrices, $\mathbf{A}^{(K_1-1,1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}$, and third factor matrix $\mathbf{A}^{(2,1)}$, with property $k\left(\mathbf{A}^{(2,1)}\right) \geq 2$. Theorem I.1 now tells us that $R = r_{\mathrm{VDM}}(\mathcal{X})$ and that the Vandermonde constrained CPD of $\mathcal{X}$ is unique.

Theorem III.3 allows us to deal with Vandermonde constrained CPD problems where none of the factor matrices have full column rank. Let us compare with Theorems 1.1, 1.3 and II.1 in a few cases. Assume that $\mathbf{A}^{(1)}$ is Vandermonde and let $I_1 = I_2 = I_3$. The maximal value R for which the conditions guarantee uniqueness in the generic case are given in Table II for various I_1 values. It is clear that Theorem III.3 yields a significantly more relaxed bound on R than Theorems 1.1, 1.3 and II.1.

TABLE II
MAXIMUM VALUE R FOR WHICH THEOREMS 1.1, 1.3, II.1 AND III.3 GUARANTEE UNIQUENESS IN THE GENERIC CASE OF A CPD WITH $I_1 = I_2 = I_3$ AND $\mathbf{A}^{(1)}$ VANDERMONDE.

$I_1 = I_2 = I_3$	2	3	4	5	6	7	8	9	10
Theorem I.1	**2**	**3**	5	6	8	9	11	12	14
Theorem I.3	**2**	**3**	5	6	8	9	11	13	15
Theorem II.1	**2**	**3**	5	6	8	9	11	12	14
Theorem III.3	**2**	**3**	**8**	**10**	**18**	**21**	**32**	**36**	**50**

Note that Theorem III.3 does not prevent $k\left(\mathbf{A}^{(2)}\right) = 1$ and/or $k\left(\mathbf{A}^{(3)}\right) = 1$. It even allows $I_2 = 1$ or $I_3 = 1$. Theorem III.3 also allows us to deal with Vandermonde constrained CPD problems where $\min_{m \neq n} I_m I_n < R$. This is in contrast with the unconstrained CPD where $\min_{m \neq n} I_m I_n \geq R$ is necessary for uniqueness [24].

B. Rank

In general the determination of tensor rank is NP-hard [14], [15]. Vandermonde structure of a factor matrix simplifies the determination of tensor rank considerably. We assume that $\mathbf{A}^{(1)}$ is Vandermonde.

Let us first assume that $\mathbf{A}^{(3)}$ has full column rank. If $\underline{\mathbf{A}}^{(1)} \odot \mathbf{A}^{(2)}$ also has full column rank, then the (tensor) rank of $\mathcal{X}$ is equal to the (matrix) rank of its matrix representation $\mathbf{X}_{(1)} = (\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)})\mathbf{A}^{(3)T}$. Lemma III.1 implies that this is generically true for R bounded as $\min\left((I_1 - 1)I_2, I_3\right) \geq R$.

If $\mathbf{A}^{(3)}$ does not have full column rank, then we first apply spatial smoothing as in Theorem III.3 and work with $\mathbf{X} = \left(\mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}\right)^T$, where $K_1 + L_1 = I_1 + 1$. If $\underline{\mathbf{A}}^{(K_1,1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}$ have full column rank, then the (tensor) rank of $\mathcal{X}$ is equal to the (matrix) rank of $\mathbf{X}$. Lemma III.1 implies that this is generically true if $\min\left((K_1 - 1)I_2, L_1 I_3\right) \geq R$.

C. Algorithm

It is also important to take the Vandermonde structure into account in the actual computation of the tensor decomposition. Unconstrained optimization methods such as the popular Alternating Least Squares (ALS) algorithm [1], [13] fail when the rank is high, even in the exact case. In this respect, an important observation is that the proofs of Proposition III.2 and Theorem III.3 are constructive, i.e., they provide us with an algorithm to compute the decomposition when the uniqueness conditions are satisfied. This algorithm only resorts to standard linear algebra and is guaranteed to return the exact solution in the noise-free case. If $\mathbf{A}^{(3)}$ has full column rank, then we work as in Proposition III.2. If $\mathbf{A}^{(3)}$ does not have full column rank, then we work as in Theorem III.3.

An outline of the proposed method for computing a CPD with a Vandermonde factor matrix is given as Algorithm 1. Note that the algorithm is computationally very cheap. The most expensive steps in Algorithm 1 are the computation of the SVD of

the matrix $\mathbf{X} \in \mathbb{C}^{K_1 I_2 \times L_1 I_3}$ of rank R and the determination of the EVD of $\mathbf{U}_1^{\dagger}\mathbf{U}_2$. The latter problem may by computed as a Generalized EVD (GEVD) of the rank-R matrices $\mathbf{U}_1 \in \mathbb{C}^{(K_1-1)I_2 \times R}$ and $\mathbf{U}_2 \in \mathbb{C}^{(K_1-1)I_2 \times R}$. For more details on SVD and GEVD we refer to the numerical linear algebra literature (e.g., [9]). Algorithm 1 basically has the same cost as one iteration of an optimization method (e.g., [36]).

Algorithm 1: Computation of CPD With Vandermonde Factor Matrix via Proposition III.2 and Theorem III.3.

Input: $\mathbf{X}_{(1)} = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)\mathbf{A}^{(3)T}$

Choose pair (K_1, L_1) subject to $K_1 + L_1 = I_1 + 1$:

$$\mathbf{X} = \left(\mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(2)}\right)\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}\right)^T$$

Compute SVD $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^H$.

Determine R from Σ.

Build $\mathbf{U}_1 = \mathbf{U}(1 : (K_1 - 1)I_2, :)$ and $\mathbf{U}_2 = \mathbf{U}(I_2 + 1 : K_1 I_2, :)$.

Determine $\{z_r\}_{r=1}^R$ and $\mathbf{M}$ from EVD $\mathbf{U}_1^{\dagger}\mathbf{U}_2 = \mathbf{M}\mathbf{Z}\mathbf{M}^{-1}$.

Build $\mathbf{a}_r^{(1)} = \left[1, z_r, z_r^2, \ldots, z_r^{I_1-1}\right]^T$, $r \in \{1, \ldots, R\}$.

Compute $\mathbf{a}_r^{(2)} = \left(\mathbf{a}_r^{(K_1,1)H} \otimes \mathbf{I}_{I_2}\right)\mathbf{U}\mathbf{m}_r$, $r \in \{1, \ldots, R\}$.

if $\left(\left(\mathbf{A}^{(1)H}\mathbf{A}^{(1)}\right) * \left(\mathbf{A}^{(2)H}\mathbf{A}^{(2)}\right)\right)^{-1}$ exists **then**

Compute $\mathbf{F} = \left(\left(\mathbf{A}^{(1)H}\mathbf{A}^{(1)}\right) * \left(\mathbf{A}^{(2)H}\mathbf{A}^{(2)}\right)\right)^{-1}$.

Compute $\mathbf{A}^{(3)T} = \mathbf{F}\left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right)^H \mathbf{X}_{(1)}$.

else

Compute $\mathbf{N} = \mathbf{M}^{-T}$.

Compute $\mathbf{a}_r^{(3)} = \left(\frac{\mathbf{a}_r^{(L_1,1)H}}{\mathbf{a}_r^{(L_1,1)H}\mathbf{a}_r^{(L_1,1)}} \otimes \mathbf{I}_{I_3}\right)\mathbf{V}^*\Sigma\mathbf{n}_r$, $r \in \{1, \ldots, R\}$.

end if

Output: $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$

In line with the experiments for ESPRIT in [27], [43], one may choose the pair (K_1, L_1) such that the dimensions of the matrices $\mathbf{A}^{(K_1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(L_1)} \odot \mathbf{A}^{(3)}$ are close, with the inequality $\lceil \frac{I_2}{R} \rceil + \lceil \frac{I_3}{R} \rceil \leq I_1$ satisfied. If I_1 is of moderate size, one may even try all I_1 possible pairs, i.e., let K_1 vary from 1 to I_1, and keep the pair (K_1, L_1) that yields the best fit to $\mathcal{X}$. Note that such a direct approach is still significantly more efficient than an optimization-based method such as ALS.

D. More Than One Vandermonde Factor

If two of the factor matrices of the CPD, say $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$, are Vandermonde, a variant of Algorithm 1 may be used.

Corollary III.4: Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ have matrix representation (2), where $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are Vandermonde matrices with generators $\{z_{1,r}\}_{r=1}^R$ and $\{z_{2,r}\}_{r=1}^R$, respectively. Consider the matrix representation $\mathbf{X} = \left(\mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(K_2,2)}\right)\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(L_2,2)} \odot \mathbf{A}^{(3)}\right)^T$ with $K_1 + L_1 = I_1 + 1$ and $K_2 + L_2 = I_2 + 1$. If

$$\begin{cases} z_{1,s} \neq z_{1,t}\,, \forall s \neq t \\ r\left(\underline{\mathbf{A}}^{(K_1,1)} \odot \mathbf{A}^{(K_2,2)}\right) = R \\ r\left(\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(L_2,2)} \odot \mathbf{A}^{(3)}\right) = R, \end{cases} \tag{29}$$

then $R = r_{\mathrm{VDM}}(\mathcal{X})$ and the Vandermonde constrained CPD of $\mathcal{X}$ is unique. Generically, condition (29) is satisfied if

$$\min_{\substack{L_m = I_m - K_m + 1 \\ m \in \{1,2\}}} \left((K_1 - 1)K_2, L_1 L_2 I_3\right) \geq R. \tag{30}$$

Proof: Let $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{X}$, then we know that under condition (29) there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\mathbf{M} = \mathbf{A}^{(K_1,1)} \odot \mathbf{A}^{(K_2,2)}. \tag{31}$$

Ignoring the Vandermonde structure of $\mathbf{A}^{(K_2,2)}$ and the Khatri-Rao structure of $\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(L_2,2)} \odot \mathbf{A}^{(3)}$, Algorithm 1 allows us to find $\mathbf{A}^{(K_1,1)}$ (and thereby also $\mathbf{A}^{(1)}$) and $\mathbf{M}$ from $\mathrm{Col}(\mathbf{X})$. Next, we compute $\mathbf{A}^{(K_2,2)}$ as follows:

$$\mathbf{a}_r^{(K_2,2)} = \left(\mathbf{a}_r^{(K_1,1)H} \otimes \mathbf{I}_{I_{K_2}}\right)\mathbf{U}\mathbf{m}_r\,.$$

Taking the Vandermonde structure of $\mathbf{A}^{(K_2,2)}$ into account, the generators of $\mathbf{A}^{(2)}$ can be found as $z_{2,r} = \left(\underline{\mathbf{a}}_r^{(K_2,2)}\right)^{\dagger}\overline{\mathbf{a}}_r^{(K_2,2)}$, $r \in \{1, \ldots, R\}$. Let $\mathbf{N} = \mathbf{M}^{-T}$, then $\mathbf{a}_r^{(3)}$ can be determined as follows:

$$\mathbf{a}_r^{(3)} = \left(\frac{\mathbf{a}_r^{(L_1,1)H}}{\mathbf{a}_r^{(L_1,1)H}\mathbf{a}_r^{(L_1,1)}} \otimes \frac{\mathbf{a}_r^{(L_2,2)H}}{\mathbf{a}_r^{(L_2,2)H}\mathbf{a}_r^{(L_2,2)}} \otimes \mathbf{I}_{I_3}\right)\mathbf{V}^*\Sigma\mathbf{n}_r. \tag{32}$$

Lemma III.1 implies that the rank assumptions in (29) hold generically if (30) is satisfied. ■

Bound (30) is to be compared with (28). Obviously, the fact that spatial smoothing may be applied in more than one mode offers more flexibility, implying a relaxed bound. Note also that Corollary III.4 does not prevent $k\left(\mathbf{A}^{(1)}\right) = 1$ and/or $k\left(\mathbf{A}^{(2)}\right) = 1$ and/or $k\left(\mathbf{A}^{(3)}\right) = 1$.

The extension of Corollary III.4 to the case of N th-order tensors in which several of the CPD factor matrices are Vandermonde is straightforward.

E. Vandermonde Factor Matrix With Repeated Generators

So far we have only considered the case where all the generators of the Vandermonde matrix $\mathbf{A}^{(1)}$ are distinct, which is the common assumption in the matrix setting. However, in the tensor setting partial computation is possible in the case where not all generators are distinct. Assume that the generators appear in Q groups of size $P_1, P_2, \ldots, P_Q$, i.e., $\mathbf{A}^{(1)}$ takes the form

$$\mathbf{A}^{(1)} = \left[\mathbf{1}_{P_1}^T \otimes \mathbf{b}_1, \ldots, \mathbf{1}_{P_Q}^T \otimes \mathbf{b}_Q\right], \tag{33}$$

where $\{\mathbf{b}_q\}_{q=1}^{Q}$ are distinct Vandermonde vectors. Partition $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$ in accordance with (33) as follows:

$$\mathbf{A}^{(2)} = \left[\mathbf{A}_1^{(2)}, \ldots, \mathbf{A}_Q^{(2)}\right], \quad \mathbf{A}_q^{(2)} \in \mathbb{C}^{I_2 \times P_q}$$
$$\mathbf{A}^{(3)} = \left[\mathbf{A}_1^{(3)}, \ldots, \mathbf{A}_Q^{(3)}\right], \quad \mathbf{A}_q^{(3)} \in \mathbb{C}^{I_3 \times P_q}.$$

Assume, despite the multiplicities of the vectors in (33), that $\underline{\mathbf{A}}^{(K_1,1)} \odot \mathbf{A}^{(2)}$ and $\mathbf{A}^{(L_1,1)} \odot \mathbf{A}^{(3)}$ have full column rank, as in Theorem III.3. It is easy to verify that $\mathbf{A}^{(1)}$ can still be obtained up to block-permutation via Algorithm 1. The eigenvalues of $\mathbf{U}_1^{\dagger}\mathbf{U}_2$ will be recovered with their multiplicities. For isolated eigenvalues also the corresponding eigenvectors, and consequently the corresponding column vectors of $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$ will be recovered. For repeated eigenvalues, only the eigenspace can be recovered, and consequently only the column spaces of the corresponding blocks of $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(3)}$. In wireless communication, $\mathbf{A}^{(3)}$ often contains the transmitted signals and is the matrix of interest (see next sections). After having separated the signals corresponding to different Vandermonde vectors, the recovery of the individual column vectors of $\mathbf{A}_q^{(3)}$ from $\mathrm{Col}\left(\mathbf{A}_q^{(3)}\right)$ may be based on properties such as statistical independence, constant modulus, finite alphabet, etc.

IV. MIMO Radar

MIMO radar has received quite some attention in recent years due to its superior identifiability properties over standard single-phased radar [12], [21], [22], [44]. In [26] a tensor-based approach to MIMO radar was proposed for the following configuration.

Assume that the receive and transmit antenna arrays consist of N and M antennas, respectively. Furthermore, assume that there are R targets in the range of interest, that the transmitted pulse signals last L samples per pulse period and that Q pulse periods are considered. Ignoring noise terms, the observed data for the qth pulse period are given by [26]:

$$\mathbf{X}^{(q)} = \mathbf{B} D_q(\mathbf{C}) \mathbf{A}^T \mathbf{S} \in \mathbb{C}^{N \times L}, \tag{34}$$

where $\mathbf{A} \in \mathbb{C}^{M \times R}$ and $\mathbf{B} \in \mathbb{C}^{N \times R}$ are the transmit and receive antenna steering matrices, respectively, $\mathbf{S} \in \mathbb{C}^{M \times L}$ is the collected transmitted pulse waveform matrix and

$$\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{Q \times R}, \quad \mathbf{c}_r = [\beta_{1r}, \ldots, \beta_{Qr}]^T,$$

are the collected Radar Cross Section (RCS) fading coefficients. The underlying assumptions in [26] are that both the transmit and receive arrays are equipped with a ULA, and that $\mathbf{S}$ is known and has full row rank. In this case the problem of finding $\mathbf{A}$ and $\mathbf{B}$ from (34) reduces to a two-dimensional harmonic retrieval problem, which may be solved by various methods, see [26] for references and details. Another interesting ULA-type configuration can be found in [2], where it is again assumed that both the transmitter and receiver are equipped with a ULA. However, if the transmit array is not ULA structured or if the transmitted waveforms are unknown, such as in passive radar, the approaches in [2], [26] are not valid anymore. As a contribution, we propose a method for finding $\mathbf{B}$ and possibly also $\mathbf{A}$. We first rewrite (34) as

$$\mathbf{X}_{(1)} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(Q)} \end{bmatrix} = (\mathbf{C} \odot \mathbf{B}) \mathbf{G} \in \mathbb{C}^{QN \times L}, \tag{35}$$

where $\mathbf{G} = \mathbf{A}^T \mathbf{S} \in \mathbb{C}^{R \times L}$.

A. Arbitrary Transmit Array, Arbitrary Receive Array and Constant RCS Coefficients

If the RCS coefficients are time invariant, then $(\mathbf{C})_{qr} = \beta_{qr} e^{(q-1)\kappa_r \sqrt{-1}}$, where $\beta_{qr} = \beta_{pr}, \forall p, q$, are the constant RCS coefficients and κ_r is the Doppler frequency for the rth target. Thus, if the generators $\{\kappa_r\}$ of the Vandermonde matrix $\mathbf{C}$ are distinct, then from Theorem III.3 we know that the Vandermonde constrained CPD (35) is generically unique if $\lceil \frac{R}{N} \rceil + \lceil \frac{R}{L} \rceil \leq Q$. Furthermore, we may use Algorithm 1 for the computation. Note that in contrast to [26] we did not impose any particular structure on the steering matrices, i.e., the transmit and receive antenna arrays may be chosen freely.

B. Arbitrary Transmit Array, ULA Receive Array and Varying RCS Coefficients

Consider now the case where the RCS coefficients change over time, such that $(\mathbf{C})_{qr} = \beta_{qr} e^{(q-1)\kappa_r \sqrt{-1}}$. Due to the CPD structure of (35) it is possible to find the factor matrices of $\mathbf{X}_{(1)}$ under mild conditions, as stated in Theorems 1.2 and 1.3. Such CPDs are usually computed by means of optimization algorithms that are not guaranteed to find the global optimum. We now explain that ULA structure of the receiver yields a more relaxed uniqueness condition and an interesting computational alternative. Assume that the receiver is equipped with a ULA such that

$$\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R], \quad \mathbf{b}_r = \left[1, z_r, z_R^2, \ldots, z_r^{N-1}\right]^T,$$

where $z_r = e^{\frac{2\pi}{\lambda} \sin(\theta_r) \Delta_x \sqrt{-1}}$ in which Δ_x is the displacement along the x-axis between each sensor, θ_r is the direction of arrival angle with respect to the receive array and λ is the carrier wavelength. Then Theorem III.3 implies that it is possible to recover the factor matrices from (35) if $\left\lceil \frac{R}{Q} \right\rceil + \lceil \frac{R}{L} \rceil \leq N$. Again, Algorithm 1 provides a numerical method.

V. Polarization Sensitive Array Processing

In this section we develop an identifiability result and an algorithm for blind signal separation in the context of polarization sensitive array processing, using the results obtained in Section III. The contribution amounts to an extension of the results in [45] to the underdetermined case, i.e., to the case where the number of impinging signals R exceeds the number of sensors J in the receive antenna array.

Starting from the signal model in [3], the following blind receiver was proposed in [45]. Assume that R signals $\{s_r(t)\}$ are

impinging on a crossed-dipole array. The output of a polarization sensitive array employing J pairs of crossed-dipoles is

$$\mathbb{C}^{2J\times K} \ni \mathbf{Y}_{(1)} = \sum_{r=1}^{R} (\mathbf{a}_r \otimes \mathbf{b}_r)\, \mathbf{s}_r^T, \tag{36}$$

where $\mathbf{s}_r \in \mathbb{C}^K$ is the signal vector lasting K sample periods, $\mathbf{b}_r \in \mathbb{C}^J$ is the steering vector associated with signal r, and

$$\mathbb{C}^2 \ni \mathbf{a}_r = \begin{bmatrix} \cos(\theta_r)\cos(\phi_r), & -\sin(\phi_r) \\ \cos(\theta_r)\sin(\phi_r), & \cos(\phi_r) \end{bmatrix} \begin{bmatrix} \sin(\gamma_r) e^{\eta_r \sqrt{-1}} \\ \cos(\gamma_r) \end{bmatrix}$$

is the polarization vector, where θ_r is the azimuth angle, ϕ_r is the elevation angle and the pair (γ_r, η_r) describes the polarization state of the rth impinging signal, respectively. For a detailed discussion of this data model we refer to [3]. The observed data can be considered as a rank-R tensor

$$\mathbb{C}^{2\times J\times K} \ni \mathcal{Y} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{s}_r\,. \tag{37}$$

In [45] it was assumed that the receiver is equipped with a ULA and that the sources are in the far-field, such that

$$\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R]\,, \quad \mathbf{b}_r = \left[1, z_r, z_r^2, \ldots, z_r^{J-1}\right]^T, \tag{38}$$

where $z_r = e^{\frac{2\pi}{\lambda}\sin(\theta_r)\sin(\phi_r)\Delta_x\sqrt{-1}}$ in which Δ_x is the displacement along the x-axis between each sensor and λ is the wavelength.

Let $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R]$ and $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_R]$. We expect to have $\mathrm{k}(\mathbf{A}) = \min(2, R)$ and $\mathrm{k}(\mathbf{S}) = \min(K, R)$. In practice, we have $K \geq R \geq 2$ and thus we expect that $\mathrm{k}(\mathbf{A}) = 2$ and $\mathrm{k}(\mathbf{S}) = R$. Equating $\mathbf{A}^{(1)} = \mathbf{B}$ and substitution into (10), with $\mathbf{A}^{(2)} = \mathbf{A}$ and $\mathbf{A}^{(3)} = \mathbf{S}$, and vice-versa, yields the following inequalities

$$\min(2R, J+R) + 2 \geq 2(R+1) \Leftrightarrow J \geq R\,, \tag{39}$$

$$\min(2R, J+2) + R \geq 2(R+1) \Leftrightarrow J \geq R\,. \tag{40}$$

In order to satisfy (39) or (40) we must have that $J \geq R$. Hence, the identifiability result (10) is only applicable in the overdetermined case, i.e., in order to guaranty uniqueness of the CPD (37) the number of impinging signals R is not allowed to exceed the number of antennas J of the ULA.

On the other hand, Theorem III.3 and consequently also Algorithm 1 are still valid in some underdetermined cases. Let us first apply spatial smoothing. Consider the fourth-order tensor $\mathcal{X} \in \mathbb{C}^{2\times L_2\times K_2\times K}$ with $L_2 + K_2 = J + 1$, constructed as follows

$$\mathcal{X}(:, l_2, k_2, :) = \mathcal{Y}(:, l_2 + k_2 - 1, :)\,.$$

We have

$$\mathbb{C}^{K_2K\times 2L_2} \ni \mathbf{X} = \begin{bmatrix} \mathcal{X}(:,1,1,:) & \cdots & \mathcal{X}(:,1,K_2,:) \\ \vdots & \ddots & \vdots \\ \mathcal{X}(:,L_2,1,:) & \cdots & \mathcal{X}(:,L_2,K_2,:) \end{bmatrix}^T = \left(\mathbf{B}^{(K_2)} \odot \mathbf{S}\right)\left(\mathbf{B}^{(L_2)} \odot \mathbf{A}\right)^T, \tag{41}$$

in which $\mathbf{B}^{(K_2)} \in \mathbb{C}^{K_2\times R}$ and $\mathbf{B}^{(L_2)} \in \mathbb{C}^{L_2\times R}$ are Vandermonde matrices. From Proposition III.2 and using the structure of the matrix (41), we may find the factor matrices of the Vandermonde constrained CPD of $\mathcal{Y}$ up to the inherent ambiguities if the condition $\min((K_2-1)K, 2L_2) \geq R$, or equivalently, $\lceil \frac{R}{K} \rceil + \lceil \frac{R}{2} \rceil \leq J$ is satisfied. As an example, let $R = 6$ and $J = 4$, then (17) becomes $\min((K_2-1)K, 2L_2) \geq 6$. By letting $L_2 = 3$, and thereby also $K_2 = 2$, we get $\min(K, 6) \geq 6 \Leftrightarrow K \geq 6$, which is generally true in practice.

In [10] another polarization sensitive array processing problem of the form (37) was considered. It was again assumed that the receiver is equipped with a ULA. The difference with the above is that the polarization vector is 6-dimensional. Again, the authors of [10] resorted to Kruskal's uniqueness condition (5). Also for this problem an improved identifiability result may be obtained from Theorem III.3. Furthermore, Algorithm 1 may be used to compute the Vandermonde constrained CPD.

VI. Blind Separation in Wireless Communication With ULA Receiver

In this section we show how the results derived in this paper can be used for blind separation in the context of wireless communication. Again, the Vandermonde structure is due to the use of a ULA at the receiver.

Consider a DS-CDMA uplink system, where the base-station is equipped with I antennas and R mobile stations are transmitting synchronously over flat fading channels. In DS-CDMA systems the transmitted encoded data y_{jkr} from user r at the jth chip period and kth symbol period are given by $y_{jkr} = h_{jr}s_{kr}$, where h_{jr} denotes the jth chip of the spreading code of user r, and s_{kr} the kth symbol transmitted by user r. Assume that a spreading code of length J is employed and that the users are transmitting over a flat fading channel with gain b_{ir} between the ith receive antenna and the rth transmitter. Then the output of the ith antenna is

$$x_{ijk} = \sum_{r=1}^{R} b_{ir}y_{jkr} = \sum_{r=1}^{R} b_{ir}h_{jr}s_{kr}.$$

Let us stack the channel gains, spreading codes and the transmitted symbols in the matrices $\mathbf{B} \in \mathbb{C}^{I\times R}$, $\mathbf{H} \in \mathbb{C}^{J\times R}$ and $\mathbf{S} \in \mathbb{C}^{K\times R}$ such that $(\mathbf{B})_{ir} = b_{ir}$, $(\mathbf{H})_{jr} = h_{jr}$ and $(\mathbf{S})_{kr} = s_{kr}$. The received data may then be stored in a tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ such that [34]

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{b}_r \circ \mathbf{h}_r \circ \mathbf{s}_r. \tag{42}$$

Using guard chips, the model still holds when reflections take place in the far field of the array; $\mathbf{h}_r$ then consists of the convolution of the rth spreading sequence and the rth channel impulse response [34].

In [24] the authors additionally considered the use of a ULA at the receiver. It was explained that the ULA structure may improve the condition of the blind signal separation problem. As a contribution, we explain that the ULA additionally improves

the uniqueness properties and that it also leads to a simple algorithm. If a ULA is employed and the sources can be considered to be in the far-field, then

$$\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R], \quad \mathbf{b}_r = \left[1, z_r, z_R^2, \ldots, z_r^{I-1}\right]^T,$$

where $z_r = e^{\frac{2\pi}{\lambda}\sin(\theta_r)\sin(\phi_r)\Delta_x\sqrt{-1}}$ in which Δ_x is the displacement along the x-axis between each sensor, θ_r is the azimuth angle, ϕ_r is the elevation angle and λ is the wavelength. We now know from Theorem III.3 that the Vandermonde constrained CPD is generically unique if $\lceil \frac{R}{J} \rceil + \lceil \frac{R}{K} \rceil \leq I$. Moreover, we may compute the factor matrices $\mathbf{B}$, $\mathbf{H}$ and $\mathbf{S}$ by means of Algorithm 1. Note also that Theorem III.3 and Algorithm 1 can be applied even when $k(\mathbf{H}) = 1$ and/or $k(\mathbf{S}) = 1$.

The results presented in this section may also be used to obtain improved uniqueness conditions and an efficient algorithm for blind signal separation in multiuser uplink cooperative systems in which the receiver is equipped with a ULA, such as those developed in [8].

VII. Blind Separation of MIMO OFDM Signals in the Case of Carrier Frequency Offset and Multipath

To the authors' knowledge the first tensor-based blind receiver for an OFDM system was presented in [16]. We first review the result in [16]. As a contribution, we explain that the receiver can be used under more relaxed conditions than has been suggested.

We consider a MIMO OFDM system where the receiver is equipped with I antennas and each user employs N subcarriers. In the kth transmission frame the symbol vector

$$\mathbf{s}^{(r)}(k) = \left[s_1^{(r)}(k), \ldots, s_N^{(r)}(k)\right]^T \in \mathbb{C}^N$$

is transmitted by user r. The symbol vector $\mathbf{s}^{(r)}(k)$ is loaded onto each subcarrier and a Cyclic Prefix (CP) is added so that the transmitted data block becomes

$$\mathbf{x}^{(r)}(k) = \mathbf{T}_{\mathrm{CP}}\mathbf{F}^H\mathbf{s}^{(r)}(k) \in \mathbb{C}^{N+D}$$

where $\mathbf{F} \in \mathbb{C}^{N\times N}$ is the DFT matrix, defined by $(\mathbf{F})_{mn} = \omega_N^{(m-1)(n-1)}$ with $\omega_N = \frac{1}{\sqrt{N}}e^{-\frac{2\pi\sqrt{-1}}{N}}$, and where $\mathbf{T}_{\mathrm{CP}} = \left[\mathbf{I}_N(1:D,:)^T, \mathbf{I}_N^T\right]^T \in \mathbb{C}^{(N+D)\times N}$ is the CP encoding matrix, with D exceeding the maximum delay spread of the communication channel. Before transmission, a parallel to serial conversion takes place such that the output of the ith receive antenna at transmission frame k and at time instant t is [16]

$$u_i^{(k)}(t) \triangleq \mathbf{u}_i(k(N+D)+t) = \sum_{r=1}^{R} e^{\phi_r t\sqrt{-1}} \sum_{l_{ir}=0}^{L_{ir}} h_i^{(r)}(l_{ir})x^{(r)}(t-l_{ir})$$

where R is the total number of users, ϕ_r is the normalized carrier frequency offset associated with user r, caused by Doppler effects or synchronization errors, and additive noise has been ignored for notational convenience. For identifiability purposes we assume that $|\phi_r| < \frac{\pi}{N}$. Stack the observed data in the vector

$$\mathbf{u}_i^{(k)} = \left[u_i^{(k)}(1), \ldots, u_i^{(k)}(N+D)\right]^T \in \mathbb{C}^{N+D}.$$

After reception, the transmitted data are decoded as

$$\mathbf{y}_i^{(k)} = \mathbf{T}_{\mathrm{CP}^{-1}}\mathbf{u}_i^{(k)} \in \mathbb{C}^N, \tag{43}$$

in which $\mathbf{T}_{\mathrm{CP}^{-1}} = [\mathbf{0}_{N,D}, \mathbf{I}_N] \in \mathbb{C}^{N\times(N+D)}$ is the CP decoding matrix. Due to the CP it can be shown that (43) is equal to [16]:

$$\mathbb{C}^N \ni \mathbf{y}_i^{(k)} = \mathbf{P}^{(r)}\mathbf{F}^H D_i\left(\mathbf{H}^{(r)}\right)\mathbf{s}^{(r)}(k)e^{\phi_r((k+1)(N+D)-N)\sqrt{-1}} \tag{44}$$

where

$$\mathbf{P}^{(r)} = D_1\left(\left[1, e^{\phi_r\sqrt{-1}}, \ldots, e^{(N-1)\phi_r\sqrt{-1}}\right]^T\right) \in \mathbb{C}^{N\times N},$$
$$\mathbf{H}^{(r)} = \left[\mathbf{h}_1^{(r)}, \ldots, \mathbf{h}_N^{(r)}\right] \in \mathbb{C}^{I\times N},$$
$$\mathbf{h}_n^{(r)} = \left[h_{1n}^{(r)}, h_{2n}^{(r)}, \ldots, h_{In}^{(r)}\right]^T,$$

in which $h_{in}^{(r)} = \sum_{l_{ir}=0}^{L_{ir}} h_i^{(r)}(l_{ir})e^{\frac{\pi\cdot l_{ir}\cdot n\cdot\sqrt{-1}}{N}}$ is the frequency response of the nth subcarrier associated with the ith receive antenna. In [16] it was assumed that $\mathbf{P}^{(r)}$ and $\mathbf{H}^{(r)}$ remain constant over K transmission blocks. Denoting

$$\mathbf{Y}_i = [\mathbf{y}_i(1), \ldots, \mathbf{y}_i(K)] \in \mathbb{C}^{N\times K},$$
$$\mathbf{S}^{(r)} = \left[\mathbf{s}^{(r)}(1), \ldots, \mathbf{s}^{(r)}(K)\right] \in \mathbb{C}^{N\times K},$$
$$\mathbf{Q}^{(r)} = D_1\Big(\left[e^{\phi_r(2D+N)\sqrt{-1}}, e^{\phi_r(3D+2N)\sqrt{-1}}, \ldots, e^{\phi_r((K+1)(N+D)-N)\sqrt{-1}}\right]^T\Big) \in \mathbb{C}^{K\times K},$$

we obtain the CPD

$$\mathbf{Y} = \begin{bmatrix}\mathbf{Y}_1\\ \vdots \\ \mathbf{Y}_I\end{bmatrix} = \sum_{r=1}^{R}\left(\mathbf{H}^{(r)} \odot \mathbf{P}^{(r)}\mathbf{F}^H\right)\left(\mathbf{Q}^{(r)}\mathbf{S}^{(r)}\right)^T = \left(\mathbf{H} \odot \widetilde{\mathbf{P}}\right)\widetilde{\mathbf{S}}^T, \tag{45}$$

where

$$\mathbf{H} = \left[\mathbf{H}^{(1)}, \ldots, \mathbf{H}^{(R)}\right] \in \mathbb{C}^{I\times NR},$$
$$\widetilde{\mathbf{P}} = \left[\mathbf{P}^{(1)}\mathbf{F}^H, \ldots, \mathbf{P}^{(R)}\mathbf{F}^H\right] \in \mathbb{C}^{N\times NR},$$
$$\widetilde{\mathbf{S}} = \left[\mathbf{Q}^{(1)T}\mathbf{S}^{(1)T}, \ldots, \mathbf{Q}^{(R)T}\mathbf{S}^{(R)T}\right] \in \mathbb{C}^{K\times NR}.$$

In [16] the authors resorted to Kruskal's identifiability condition(5), which states that the CPD (45) is unique if

$$k(\mathbf{H}) + k\left(\widetilde{\mathbf{P}}\right) + k\left(\widetilde{\mathbf{S}}\right) \geq 2NR + 2. \tag{46}$$

In practice we expect that (46) reduces to

$$\min(I, NR) + N + \min(K, NR) \geq 2NR + 2. \tag{47}$$

In this derivation the structure of $\widetilde{\mathbf{P}}$ has been ignored and consequently the restrictive identifiability condition (46) has been obtained. Note that condition(47) is only useful when N is small, which may not be the case in practice. More recently, in [30] it has been observed that $\widetilde{\mathbf{P}}$ is in fact a Vandermonde matrix. Indeed, let $z_r = e^{\phi_r\sqrt{-1}}$ and $\rho_N = \omega_N^*$, then

$$\mathbf{P}^{(r)}\mathbf{F}^H = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_r & z_r\rho_N & \cdots & z_r\rho_N^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ z_r^{N-1} & (z_r\rho_N)^{N-1} & \cdots & (z_r\rho_N^{N-1})^{N-1} \end{bmatrix}. \tag{48}$$

The matrix in (48) is a Vandermonde matrix with generators $\{z_r\rho_N^n\}_{n=0}^{N-1}$. By simultaneously exploiting the Vandermonde structure of $\widetilde{\mathbf{P}}$ and the Khatri-Rao product structure of $\mathbf{H} \odot \widetilde{\mathbf{P}}$, Theorem III.3 guarantees that the CPD (45) is unique if

$$\underline{\widetilde{\mathbf{P}}}^{(K_1)} \odot \mathbf{H} \text{ and } \widetilde{\mathbf{P}}^{(L_1)} \odot \widetilde{\mathbf{S}} \text{ have full column rank,} \tag{49}$$

subject to $K_1 + L_1 = N+1$, in which the Vandermonde matrices $\widetilde{\mathbf{P}}^{(K_1)}$ and $\widetilde{\mathbf{P}}^{(L_1)}$ are obtained by spatial smoothing. In practice, we expect that (49) yields the condition

$$\min\left((K_1 - 1)\, I, L_1 K\right) \geq NR \text{ subject to } K_1 + L_1 = N + 1\,. \tag{50}$$

Comparison of (47) and (50) shows that the identifiability condition has been significantly relaxed.

Algorithm 1 does not take into account the relationship between the columns in (48). An optimization method that does, is presented in [16].

VIII. Simulations

The applications discussed in this paper essentially lead to the same computation. We illustrate the performance of the algorithms for the polarization sensitive array processing problem described in Section V and the problem of separating oversampled signals described in Section VI. Let $\mathbf{Y}_{(1)} \in \mathbb{C}^{IJ\times K}$ denote the noise-free tensor and let $\mathbf{X}_{(1)} = \mathbf{Y}_{(1)} + \beta\mathbf{N}_{(1)}$, where $\beta \in \mathbb{R}$ and $\mathbf{N}_{(1)} \in \mathbb{C}^{IJ\times K}$ is a noise term with elements randomly drawn from a Gaussian distribution with zero mean and unit variance.

The Signal to Noise Ratio (SNR) is measured as

$$\text{SNR [dB]} = 10\log\left(\left\|\mathbf{Y}_{(1)}\right\|_F^2 / \left\|\beta\mathbf{N}_{(1)}\right\|_F^2\right).$$

In the case $I_n \geq R$, the distance between the factor matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n\times R}$ and its estimate $\widehat{\mathbf{A}}^{(n)}$ is measured as

$$\mathrm{P}\left(\mathbf{A}^{(n)}\right) = \min_{\Pi\Lambda}\left\|\mathbf{A}^{(n)} - \widehat{\mathbf{A}}^{(n)}\Pi\Lambda\right\|_F / \left\|\mathbf{A}^{(n)}\right\|_F,$$

where Π denotes a permutation matrix and Λ denotes a diagonal matrix. In order to find Π and Λ we use the greedy least squares column matching algorithm proposed in [34].

We compare Algorithm 1 with the classical ALS method, which ignores the Vandermonde structure. Let $f\left(\widehat{\mathbf{X}}_{(1)}^{(k)}\right) = \left\|\mathbf{X}_{(1)} - \widehat{\mathbf{X}}_{(1)}^{(k)}\right\|_F$, where $\widehat{\mathbf{X}}_{(1)}^{(k)}$ denotes the estimated tensor at iteration k. We decide that the ALS method has converged when $\left|f\left(\widehat{\mathbf{X}}_{(1)}^{(k)}\right) - f\left(\widehat{\mathbf{X}}_{(1)}^{(k+1)}\right)\right| < 1e-8$ or when the number of iterations exceeds 2000. In the ALS method the factor matrices are initialized randomly or by means of the GEVD of $\mathbf{X}^{(\cdot\cdot 1)}$ and $\mathbf{X}^{(\cdot\cdot 2)}$ if possible. When only one random initialization is used, the ALS method will be referred to as ALS-CP. When the best out of ten random initializations is used, the ALS method will be referred to as ALS-CP-10. When the ALS method is initialized by means of the GEVD, then it will be referred to as ALS-CP-GEVD. Algorithm 1 will be referred to as CP-VDM. For the spatial smoothing parameter K_1, the value that yields the best fit of $\mathcal{X}$ is retained. More precisely, when the factor matrix $\mathbf{A}^{(1)}$ is Vandermonde structured, the spatial smoothing parameter K_1 is chosen as the minimizer of

$$g\left(\widehat{\mathbf{A}}_{K_1}^{(1)}, \widehat{\mathbf{A}}_{K_1}^{(2)}, \widehat{\mathbf{A}}_{K_1}^{(3)}\right) = \left\|\mathbf{X}_{(1)} - \left(\widehat{\mathbf{A}}_{K_1}^{(1)} \odot \widehat{\mathbf{A}}_{K_1}^{(2)}\right)\widehat{\mathbf{A}}_{K_1}^{(3)T}\right\|_F,$$

where $\{\widehat{\mathbf{A}}_{K_1}^{(n)}\}$ denote the estimates of the factor matrices of $\mathbf{Y}_{(1)}$ obtained by CP-VDM with the given K_1. If the Vandermonde factor matrix is tall, then we will also compare Algorithm 1 with the ESPRIT method followed by rank-1 approximation, as explained in Theorem II.3. This approach will simply be referred to as ESPRIT.

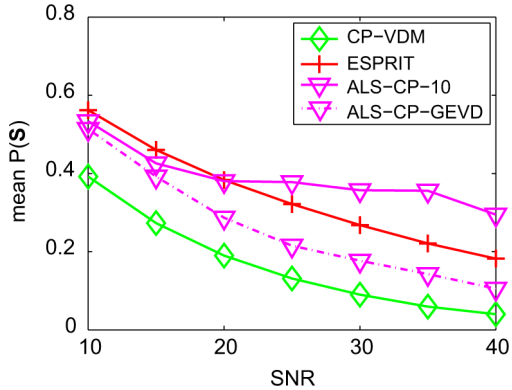

Fig. 1. Mean P (**S**) over 100 trials for varying SNR, case 1.

A. Polarization Sensitive Array Processing

The polarization parameters γ_r and η_r are randomly drawn from a Gaussian distribution with zero mean and unit variance while the azimuth and elevation angles θ_r and ϕ_r are randomly drawn from a uniform distribution with support $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The elements of the symbol matrix are randomly drawn from a QPSK constellation.

Case 1: $I < R$ *and* $J, K > R$*:* The model parameters are $I = 2$, $J = 6$, $K = 100$ and $R = 5$. For this configuration Kruskal's condition (6) is satisfied. Hence, the CPD is unique even when the Vandermonde structure is not taken into account. The mean P (**S**) values over 100 trials for varying SNR are shown in Fig. 1. We notice that CP-VDM yields the best performance. We also notice that the ALS method is sensitive w.r.t. its initialization.

Case 2: $I, J < R$ *and* $K > R$*:* The model parameters are $I = 2$, $J = 4$, $K = 100$ and $R = 5$. In this configuration CPD is not unique by itself. Also, initialization by GEVD or the use

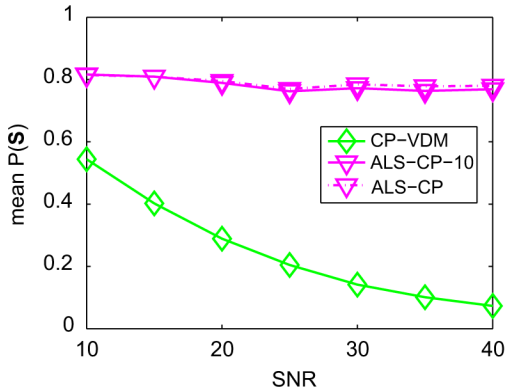

Fig. 2. Mean P $(\mathbf{S})$ over 100 trials for varying SNR, case 2.

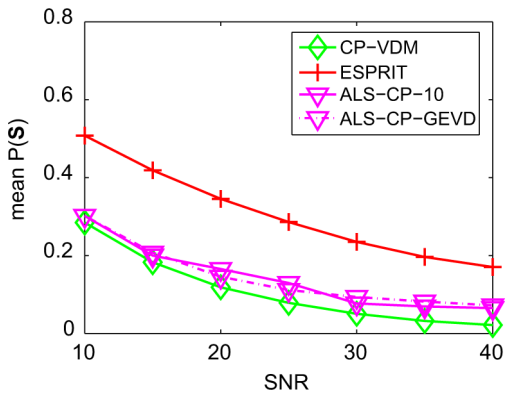

Fig. 3. Mean P $(\mathbf{S})$ over 100 trials for varying SNR, case 3.

of ESPRIT is not possible. On the other hand, Theorem III.3 tells us that the Vandermonde constrained CPD is unique. The mean P $(\mathbf{S})$ values over 100 trials for varying SNR are shown in Fig. 2. We notice that CP-VDM works well while ALS-CP and ALS-CP-10 indeed fail.

B. Blind Separation of DS-CDMA Signals

The azimuth and elevation angles θ_r and ϕ_r are randomly drawn from a uniform distribution with support $[-\frac{\pi}{2}, \frac{\pi}{2}]$. To take into account possible interchip interference the entries of the spreading code matrix $\mathbf{H}$ are randomly drawn from a Gaussian distribution with zero mean and unit variance. Finally, the elements of the symbol matrix are randomly drawn from a QPSK constellation.

Case 3: $J < R$ and $I, K > R$: The model parameters are $I = 6$, $J = 3$, $K = 100$ and $R = 5$. For this configuration Kruskal's condition (6) is satisfied. Hence, the CPD is unique even when the Vandermonde structure is not taken into account. The mean P $(\mathbf{S})$ values over 100 trials for varying SNR are shown in Fig. 3. We notice that the CP-VDM method performs better than the ALS methods, which in turn perform better than ESPRIT. The pure CPD algorithms perform better than in the simulations in Section VIII -A because the CPD structure is stronger, due to the fact that the tensor has more slices.

Case 4: $I, J < R$ and $K > R$: The model parameters are $I = 4$, $J = 3$, $K = 100$ and $R = 8$. In this configuration CPD is not unique by itself. Also, initialization by GEVD or the use of ESPRIT is not possible. On the other hand, Theorem III.3 tells us that this Vandermonde constrained CPD is unique. The mean P $(\mathbf{S})$ values over 100 trials for varying SNR are shown in Fig. 4. We notice that CP-VDM works well, while ALS-CP and ALS-CP-10 fail as expected.

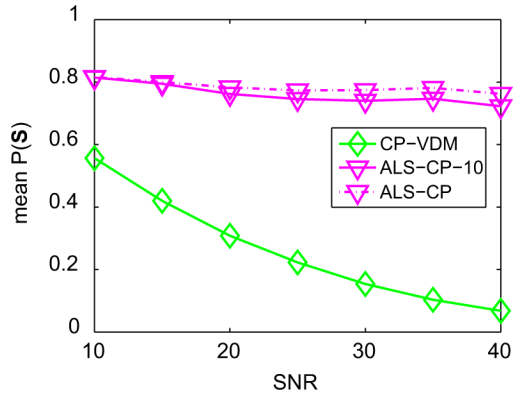

Fig. 4. Mean P $(\mathbf{S})$ over 100 trials for varying SNR, case 4.

IX. Conclusion

Many problems in signal processing can be formulated as the computation of a CPD with a Vandermonde factor matrix. We first derived a new uniqueness condition that is more relaxed than the uniqueness conditions that are available for the unconstrained CPD. We explained that under this condition, the tensor rank can be estimated as the rank of a matrix, while for the unconstrained CPD the determination of tensor rank is in general NP-hard. Second, we presented an algorithm for computing the factor matrices under the new uniqueness condition. This algorithm essentially relies on a matrix EVD and is guaranteed to find the decomposition if it is exact. For optimization-based algorithms for the unconstrained CPD there is no such guarantee.

In the second part of the paper we illustrated how the results may be used in sensor array processing. More precisely, we showed how they can be used in MIMO radar. We also illustrated the relevance in underdetermined blind signal separation problems that have arisen in the context of polarization sensitive arrays. We also illustrated the usefulness of the results in blind signal separation in wireless communication. In particular, we considered multiuser uplink cooperative systems and systems using oversampling such as DS-CDMA. We also showed how the results can be used to obtain identifiability results for MIMO OFDM systems using cyclic prefixing in the case of carrier frequency offset.

In the third part of the paper simulations confirmed the practical interest of the proposed algorithm. Especially in the highly underdetermined case $(I, J < R)$ the method showed superior performance.

In this paper we have limited ourselves to problems where blind signal separation can be achieved through CPD. In a follow-up paper we will extend the results to array processing problems in the case of coherent or incoherent channels with small delay spread.

References

[1] 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, no. 3, pp. 283–319, 1970.

[2] C. Chen and P. P. Vaidyanathan, "Mimo radar space-time adaptive processing using prolate spheroidal wave functions," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 623–635, Feb. 2008.
[3] R. T. Compton, "The tripole antenna: An adaptive array with full polarization flexibility," *IEEE Trans. Antennas Propagat.*, no. 6, pp. 944–952, Nov. 1981.
[4] L. D. Lathauwer, "A short introduction to tensor-based methods for factor analysis and blind source separation," in *Proc. 7th Int. Symp. on Image and Signal Process. Anal. (ISPA 2011)*, Dubrovnik, Croatia, Sep. 4–6, 2011, pp. 558–563.
[5] L. D. Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.
[6] I. Domanov and L. D. Lathauwer, "On the uniqueness of the Canonical Polyadic Decomposition—part I: basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, 2013.
[7] I. Domanov and L. D. Lathauwer, "On the uniqueness of the Canonical Polyadic Decomposition—Part II: Overall uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, 2013.
[8] C. A. R. Fernandes, A. L. F. de Almeida, and D. B. da Costa, "Unified tensor modeling for blind receivers in multiuser uplink cooperative systems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 247–250, May 2012.
[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The John Hopkins Univ. Press, 1996.
[10] X. Guo, S. Miron, and D. Brie, "Three-way array analysis on polarization signals for direction-finding and blind source separation," in *Proc. IAR 2007*, Grenoble, France, Nov. 15–16, 2007.
[11] X. Guo, S. Miron, D. Brie, S. Zhu, and X. Liao, "A candecomp/parafac perspective on uniqueness of doa estimation using a vector sensor array," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3475–3481, Jul. 2011.
[12] A. Haimovich, R. S. Blum, and L. J. Cimini, Jr, "Mimo radar with widely separated antennas," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 116–129, Jan. 2008.
[13] R. A. Harshman, "Foundations of the parafac procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonet.*, vol. 16, pp. 1–84, 1970.
[14] J. Håstad, "Tensor rank is NP-complete," *J. Algorithms*, vol. 11, no. 4, pp. 644–654, 1990.
[15] C. J. Hillar and L.-H. Lim, Most Tensor Problems are NP-Hard Tech. Rep., Eprint arXiv:0911.1393v4.
[16] T. Jiang and N. D. Sidiropoulos, "A direct blind receiver for SIMO and MIMO OFDM systems subject to unknown frequency offset and multipath," in *Proc. IEEE 4th Workshop on Signal Process. Adv. Wireless Commun.*, Rome, Italy, Jun. 15–18, 2003, pp. 358–362.
[17] T. Jiang and N. D. Sidiropoulos, "Kruskal's permutation lemma and the identification of candecomp/parafac and bilinear model with constant modulus constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, Sep. 2004.
[18] T. Jiang, N. D. Sidiropoulos, and J. M. F. T. Berge, "Almost-sure identifiability of multidimensional harmonic retrieval," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1849–1859, Sep. 2001.
[19] H. A. L. Kiers, "Towards a standardized notation and terminology in multiway analysis," *J. Chemometr.*, vol. 14, no. 3, pp. 105–122, 2000.
[20] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, pp. 95–138, 1977.
[21] J. Li and P. Stoica, "Mimo radar with colocated antennas," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 106–114, Sep. 2007.
[22] J. Li, P. Stoica, L. Xu, and W. Roberts, "On parameter identifiability of MIMO radar," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 968–971, Dec. 2007.
[23] J. Liu and X. Liu, "An eigenvector-based approach for multidimensional frequency estimation with improved identifiability," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 2074–2086, Dec. 2006.
[24] X. Liu and N. D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2074–2086, Sep. 2001.
[25] X. Liu and N. D. Sidiropoulos, "Almost sure identifiability of multidimensional constant modulus harmonic retrieval," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2366–2368, Sep. 2002.
[26] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5693–5705, Nov. 2010.
[27] J. M. Papy, L. D. Lathauwer, and S. V. Huffel, "Exponential data fitting using multilinear algebra: the single-channel and multi-channel case," *Numer. Linear Algebra Appl.*, vol. 12, pp. 809–826, 2005.
[28] J. M. Papy, L. D. Lathauwer, and S. V. Huffel, "Exponential data fitting using multilinear algebra: The decimative case," *J. Chemometr.*, vol. 23, no. 7-8, pp. 341–351, 2009.
[29] R. Roy and T. Kailath, "Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Syst. Signal Process.*, vol. 32, no. 7, pp. 984–995, Jul. 1989.
[30] A. Santra and K. Hari, "Novel subspace methods for blind estimation of multiple CFOS in MIMO-OFDM systems," in *Proc. ICSP 2010*, Beijing, China, Oct. 28, 2010.
[31] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-on-arrival of coherent signals," *IEEE Trans. Acoust. Syst. Signal Process.*, vol. 33, no. 8, pp. 806–811, Aug. 1985.
[32] N. D. Sidiropoulos, "Generalizing carathéodory's uniqueness of harmonic parameterization to N dimensions," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1687–1690, May 2001.
[33] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.
[34] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind parafac receivers for DS-CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 810–823, Mar. 2000.
[35] N. D. Sidiropoulos and X. Liu, "Identifiability results for blind beamforming in incoherent multipath with small delay spread," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 228–236, Jan. 2001.
[36] L. Sorber, M. V. Barel, and L. D. Lathauwer, "Optimization-based algorithms for tensor decompositions: Canonical Polyadic Decomposition, decomposition in rank- $(L_r, L_r, 1)$ terms and a new generalization," *SIAM J. Opt.*, vol. 32, no. 2, pp. 695–702, 2013.
[37] M. Sørensen and L. D. Lathauwer, "Tensor decompositions with a Vandermonde factor and applications in signal processing," in *Proc. Asilomar*, Pacific Grove, California, USA, Nov. 4–7, 2012.
[38] M. Sørensen, L. D. Lathauwer, P. Comon, S. Icart, and L. Deneire, "Canonical Polyadic Decomposition with a columnwise orthonormal factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 33, no. 4, pp. 1190–1213, 2012.
[39] A. Stegeman, "On uniqueness conditions for candecomp/parafac and indscal with full column rank in one mode," *Linear Algebra Appl.*, vol. 431, pp. 211–227, 2009.
[40] A. Stegeman, "On uniqueness of the nth order tensor decomposition into rank-1 terms with linear independence in one mode," *SIAM J. Matrix Anal. Appl.*, vol. 31, pp. 2498–2516, 2010.
[41] A. Stegeman and N. D. Sidiropoulos, "On Kruskal's uniqueness condition for the candecomp/parafac decomposition," *Linear Algebra Appl.*, vol. 420, pp. 540–552, 2007.
[42] A.-J. V. D. Veen, "Algebraic methods for deterministic blind beamforming," *Proc. IEEE*, vol. 10, pp. 1987–2008, 1998.
[43] S. V. Huffel, H. Chen, C. Decanniere, and P. V. Hecke, "Algorithm for time-domain nmr data fitting based on total least squares," *J. Magn. Reson. Ser. A*, vol. 110, pp. 238–237, 1994.
[44] H. Yan, J. Li, and G. Liao, "Multitarget identification and localization using bistatic radar systems," *EURASIP J. Adv. Signal Process.*, 2008, Article ID 283483.
[45] X. Zhang and D. Xu, "Blind parafac signal detection for polarization sensitive array," *EURASIP J. Adv. Signal Process.*, 2007, Article ID 12025.



Mikael Sørensen received the Master's degree from Aalborg University, Denmark, and the Ph.D. degree from University of Nice, France, in 2006 and 2010, respectively, both in electrical engineering.

Since 2010, he has been a Postdoctoral Fellow with KU Leuven, Belgium. His research interests include applied linear and multilinear algebra, blind source separation and system identification, and signal processing for wireless communication and sensor array processing.



Lieven De Lathauwer (SM'06) received the Master's degree in electromechanical engineering and the Ph.D. degree in applied sciences from KU Leuven, Belgium, in 1992 and 1997, respectively.

He is currently an ssociate Professor with KU Leuven, Belgium. His research interests include linear and multilinear algebra, statistical signal and array processing, higher order statistics, independent component analysis, identification, blind identification, and equalization.

Dr. De Lathauwer is an Associate Editor of the SIAM *Journal on Matrix Analysis and Applications* and has been an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.