Coupled tensor decompositions and monomial structure in array processing

Mikael Sørensen and Lieven De Lathauwer, *Fellow, IEEE*

***Abstract*—The Canonical Polyadic Decomposition (CPD) plays an important role in array processing in the case of arrays composed of several displaced but identical subarrays. However, the CPD model is less appropriate for more complex array geometries. As our contribution, we first present a coupled tensor decomposition approach. We demonstrate that the coupled CPD can handle array processing problems involving multiple shift-invariance structures. Next, we make a connection between bilinear factorizations involving monomial structures and the coupled Block Term Decomposition (BTD). This leads to an extension of the coupled decomposition framework to more general array processing problems involving monomial structures. This includes nonuniform and near-field array processing and the separation of M-PSK signals using an antenna array. We also discuss a variant that can handle monomials with conjugate variables. The latter can be used to impose constant modulus constraints on steering vectors or impinging signals. We obtain a deterministic condition for monomial factorizations that is both necessary and sufficient in the multiple snapshot case but which may be difficult to check in practice. We derive a deterministic relaxation that admits a constructive interpretation and is also more easy to verify. Finally, we explain that the link between monomials and tensor decompositions allows a further unification and yields a framework that encompasses classical tensor methods for array processing, such as ESPRIT, CPD and ACMA.**

***Index Terms*—array processing, tensor, canonical polyadic decomposition, block term decomposition, coupled decomposition, monomials.**

I. Introduction

Sensor arrays have found important applications in radar, sonar, wireless communication and medical diagnostics. The connection between sensor array processing problems with regular array geometries composed of several displaced but identical subarrays and the Canonical Polyadic Decomposition (CPD) was made in [18]. In practice, array geometries may also be irregular or sparse. A notable limitation of the CPD is that it can only handle (very) regular array geometries. Hence, the need for a flexible and algebraic framework for array processing that can improve and break the confinements of the standard CPD approach is apparent.

M. Sørensen and L. De Lathauwer are with KU Leuven - E.E. Dept. (ESAT) - STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, the Group Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium, and iMinds Medical IT, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, {Mikael.Sorensen, Lieven.DeLathauwer}@kuleuven.be.

To overcome some of the limitations of traditional tensor based array signal processing, we have already suggested a shift from conventional CPD-based approach to coupled matrix/tensor based array processing [24], [26], [27]. In this paper we further broaden the approach. First, we argue that the coupled CPD framework presented in [22], [28] provides a unified algebraic framework for handling various types of one-, two-, three- or N-dimensional arrays with shift-invariance properties. Part of this work was presented in the conference paper [21].

Next, we extend the coupled tensor decomposition framework to array processing problems that involve monomial structures. This includes Direction-Of-Arrival (DOA) estimation using sparse rectangular arrays (e.g. [35], [12]), near-field extensions of DOA and range estimation using a linear array (e.g. [30], [10]), and blind separation of digital communication signals using an antenna array (e.g. [31], [9]). Using the coupled Block Term Decomposition (BTD), we present deterministic uniqueness conditions and an algebraic algorithm that in the exact case reduces a factorization problem involving monomial structure to a classical EigenValue Decomposition (EVD). An algorithmic variant for monomials with conjugate variables is also presented. This is for instance relevant in the case of constant modulus constraints.

Finally, we explain that the link between monomially constrained factorization problems and coupled tensor decompositions leads to a very general framework in which the classical methods ESPRIT [17], CPD [18] and ACMA [33] are special cases.

The paper is organized as follows. The rest of the introduction will present the notation. Section II reviews classical tensor based array processing. Section III introduces the coupled CPD framework for array processing problems involving array geometries with multiple shift-invariance structures. Section IV extends the presented framework to monomial structures and briefly discusses applications. We also make the connections with ESPRIT, CPD and ACMA. Numerical experiments are reported in Section V. Section VI concludes the paper.

Notation: Vectors, matrices and tensors are denoted by lower case boldface, upper case boldface and upper case calligraphic letters, respectively. The rth column, conjugate, transpose, conjugate-transpose, determinant, rank, Moore-Penrose pseudoinverse, Frobenius norm, range and kernel of a matrix $\mathbf{A}$ are denoted by $\mathbf{a}_r$, $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $|\mathbf{A}|$, $\mathbf{A}^\dagger$, $\|\mathbf{A}\|_F$, $r(\mathbf{A})$, range $(\mathbf{A})$, ker $(\mathbf{A})$, respectively.

The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri-Rao product, defined as

$$\mathbf{A}\otimes\mathbf{B} := \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}, \mathbf{A}\odot\mathbf{B} := [\mathbf{a}_1 \otimes \mathbf{b}_1 \;\; \mathbf{a}_2 \otimes \mathbf{b}_2 \;\dots],$$

in which $(\mathbf{A})_{mn} = a_{mn}$. The outer product of, say, three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ is denoted by $\mathbf{a}\circ\mathbf{b}\circ\mathbf{c}$, such that $(\mathbf{a}\circ\mathbf{b}\circ\mathbf{c})_{ijk} = a_i b_j c_k$. From the context it should be clear when i denotes the imaginary unit number, i.e., $i = \sqrt{-1}$.

The all-ones vector is denoted by $\mathbf{1}_R = [1,\dots,1]^T \in \mathbb{C}^R$. The I-element unit vector with unit entry at position i and zero elsewhere is denoted $\mathbf{e}_i^{(I)} \in \mathbb{C}^I$. $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$ denotes the diagonal matrix holding row k of $\mathbf{A} \in \mathbb{C}^{I\times J}$ on its diagonal. 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 $\mathbf{A} \in \mathbb{C}^{I\times R}$, then $\underline{\mathbf{A}} = \mathbf{A}(1:I-1,:) \in \mathbb{C}^{(I-1)\times R}$ and $\overline{\mathbf{A}} = \mathbf{A}(2:I,:) \in \mathbb{C}^{(I-1)\times R}$, i.e., $\underline{\mathbf{A}}$ and $\overline{\mathbf{A}}$ are obtained by deleting the bottom and upper row of $\mathbf{A}$, respectively.

The binomial coefficient is denoted by $C_m^k = \frac{m!}{k!(m-k)!}$. The k-th compound matrix of $\mathbf{A} \in \mathbb{C}^{I\times R}$ is denoted by $C_k(\mathbf{A}) \in \mathbb{C}^{C_I^k \times C_R^k}$. 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 [4] and references therein for a discussion.

II. Tensor based array processing

Consider R signals impinging on an array composed of I sensors such that the output of the ith sensor at the kth observation is

$$x_{ik} = \sum_{r=1}^{R} s_r(k - \tau_{ri}), \tag{1}$$

where τ_{ri} denotes the delay between the ith sensor and the rth source. The position of the ith sensor in cartesian coordinates is given by $\mathbf{p}_i = [x_i \; y_i \; z_i]^T$. Let us additionally assume that the sources are located in the far-field and that the narrowband assumption holds. Under these assumptions the array response vector $\mathbf{a}_r \in \mathbb{C}^I$ associated with the rth source can (approximately) be expressed as

$$\mathbf{a}_r = [e^{-\sqrt{-1}\omega_c \mathbf{b}_r^T\mathbf{p}_1/c} \;\dots\; e^{-\sqrt{-1}\omega_c \mathbf{b}_r^T\mathbf{p}_I/c}]^T, \tag{2}$$

where ω_c is the carrier frequency, $\mathbf{b}_r = [\sin(\phi_r)\cos(\theta_r),\; \sin(\phi_r)\sin(\theta_r),\; \cos(\phi_r)]^T$ is the bearing vector in which θ_r and ϕ_r denote the azimuth and elevation angle, respectively, c is the speed of propagation, and the product $\mathbf{b}_r^T\mathbf{p}_i/c$ corresponds to the propagation delay in (1) associated with the ith sensor and the rth source, that is $\tau_{ri} = \mathbf{b}_r^T\mathbf{p}_i/c$. Assume that R sources are impinging on the sensor array and that K snapshots are available such that $\mathbf{s}_r \in \mathbb{C}^K$ denotes the signal vector associated with the rth source. Then the observed data matrix admits the factorization

$$\mathbf{X} = \mathbf{A}\mathbf{S}^T \in \mathbb{C}^{I\times K}, \quad \mathbf{A} = [\mathbf{a}_1 \;\dots\; \mathbf{a}_R],\; \mathbf{S} = [\mathbf{s}_1 \;\dots\; \mathbf{s}_R]. \tag{3}$$

In Direction-Of-Arrival (DOA) applications the goal is to estimate $\{\theta_r, \phi_r\}$ via $\mathbf{A}$, observing $\mathbf{X}$. On the other hand, in source separation the goal is to estimate $\mathbf{S}$, observing $\mathbf{X}$. For both problems, uniqueness conditions and algorithms need to be developed. To this end, either arrays with simple geometries or signals with certain known properties (statistical independence, finite alphabet, sparsity, etc.) are used.

A. CPD model for array processing

Because of their simplicity, arrays composed of several displaced but identical subarrays are often considered in the literature (see [18] and references therein). More precisely, let x_{ijk} denote the output of the ith sensor of the jth subarray at the kth snapshot. Assume that J subarrays each composed of I sensors are used. Then the observed third-order data tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ admits a so-called Canonical Polyadic Decomposition (CPD) [1], [11], [14], as pointed out in [18]:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{d}_r \circ \mathbf{s}_r, \tag{4}$$

where $\mathbf{a}_r = [e^{-\sqrt{-1}\omega_c \mathbf{b}_r^T\mathbf{p}_1/c} \;\dots\; e^{-\sqrt{-1}\omega_c \mathbf{b}_r^T\mathbf{p}_I/c}]^T \in \mathbb{C}^I$, $\mathbf{d}_r = [e^{-\sqrt{-1}\omega_c \mathbf{b}_r^T\mathbf{t}_1/c} \;\dots\; e^{-\sqrt{-1}\omega_c \mathbf{b}_r^T\mathbf{t}_J/c}]^T \in \mathbb{C}^J$, and $\mathbf{s}_r \in \mathbb{C}^K$ denote the reference subarray, the displacement and signal vector associated with the rth source, respectively. In this case $\mathbf{p}_i$ denotes the position of the ith sensor in the reference subarray while $\mathbf{t}_j$ denotes the translation vector associated with the jth subarray. Note that the I element subarray vector $\mathbf{a}_r$ is repeated J times and scaled according to the entries of $\mathbf{d}_r$. This *polyadic structure* is the algebraic manifestation of the *shift-invariance* property of the array. Uniqueness conditions and algebraic algorithms for the CPD can be found in [13], [3], [4], [5], [6], [7], [19] and references therein.

Throughout the paper matrix representations of tensors will be used. Consider the horizontal matrix slice $\mathbf{X}^{(i\cdot\cdot)} \in \mathbb{C}^{J\times K}$ of $\mathcal{X}$, defined by $(\mathbf{X}^{(i\cdot\cdot)})_{jk} = x_{ijk} = \sum_{r=1}^R a_{ir}d_{jr}s_{kr}$. The tensor $\mathcal{X}$ can be interpreted as a collection of matrix slices $\mathbf{X}^{(1\cdot\cdot)},\dots,\mathbf{X}^{(I\cdot\cdot)}$, each admitting the factorization $\mathbf{X}^{(i\cdot\cdot)} = \sum_{r=1}^R a_{ir}\mathbf{d}_r\mathbf{s}_r^T = \mathbf{D}D_i(\mathbf{A})\mathbf{S}^T$. Stacking yields the matrix representation of (4):

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}^{(1\cdot\cdot)} \\ \vdots \\ \mathbf{X}^{(I\cdot\cdot)} \end{bmatrix} = \begin{bmatrix} \mathbf{D}D_1(\mathbf{A}) \\ \vdots \\ \mathbf{D}D_I(\mathbf{A}) \end{bmatrix} \mathbf{S}^T = (\mathbf{A}\odot\mathbf{D})\mathbf{S}^T \in \mathbb{C}^{IJ\times K}. \tag{5}$$

B. ESPRIT method for array processing

One of the first proposed array processing methods that can be interpreted in terms of the CPD model (4) is ESPRIT [17]. Recall that ESPRIT considers a bilinear factorization of the form (3) in which $\mathbf{S}$ has full column rank and the columns of $\mathbf{A}$ are Vandermonde vectors of the form

$$\mathbf{a} = [a_0, a_1, \dots, a_{I-1}]^T = [1, z, \dots, z^{I-1}]^T \in \mathbb{C}^I. \tag{6}$$

Using the shift-property $z^{m+1} = z \cdot z^m$, the Vandermonde vector (6) can be related to the rank-1 Hankel matrix

$$\mathbf{H} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{I-2} \\ a_1 & a_2 & \cdots & a_{I-1} \end{bmatrix} = \begin{bmatrix} 1 \\ z \end{bmatrix} \underline{\mathbf{a}}^T = \begin{bmatrix} 1 \\ z \end{bmatrix} \circ \underline{\mathbf{a}}. \quad (7)$$

The shift-invariance property also implies that $\overline{\mathbf{X}} = \overline{\mathbf{A}}\mathbf{S}^T = \underline{\mathbf{A}}D_1([z_1, \ldots, z_R])\mathbf{S}^T = \underline{\mathbf{A}}D_2(\mathbf{A})\mathbf{S}^T$. Let us stack $\underline{\mathbf{X}}$ and $\overline{\mathbf{X}}$ in a two-slice tensor $\mathcal{X} \in \mathbb{C}^{2\times(I-1)\times K}$. According to (5), $\underline{\mathbf{X}} = \underline{\mathbf{A}}\mathbf{S}^T$ and $\overline{\mathbf{X}} = \underline{\mathbf{A}}D_2(\mathbf{A})\mathbf{S}^T$ can be seen as a CPD $\mathcal{X} = \sum_{r=1}^R \begin{bmatrix} 1 \\ z_r \end{bmatrix} \circ \underline{\mathbf{a}}_r \circ \mathbf{s}_r$, with matrix representation

$$\mathbf{X} = \begin{bmatrix} \underline{\mathbf{X}} \\ \overline{\mathbf{X}} \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{A}} \\ \underline{\mathbf{A}}D_2(\mathbf{A}) \end{bmatrix} = \left(\begin{bmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_R \end{bmatrix} \odot \underline{\mathbf{A}} \right) \mathbf{S}^T. \quad (8)$$

The first explicit mentioning of the link between CPD and ESPRIT in the context of array processing was probably in [18]. It is well-known that if $\underline{\mathbf{A}}$ and $\mathbf{S}$ have full column rank and the Vandermonde generators of $\mathbf{A}$ are distinct (i.e., $z_r \neq z_s$ for all $r \neq s$), then the CPD of $\mathcal{X}$ is unique and can be computed via an EVD [15]. In the next section we will see that the coupled CPD model provides multidimensional extensions of ESPRIT.

III. Coupled tensor decompositions and arrays with multiple shift-invariance structures

A. Coupled CPD model for array processing

Standard tensor-based methods for array processing are essentially limited to situations where the observed data admit a CPD interpretation of the form (4). In this section we explain that the coupled *CPD model* (9) below enables a more general modelling. It extends the CPD model to arrays enjoying *multiple shift-invariance* structures:

$$\mathbb{C}^{I_n \times J_n \times K} \ni \mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{d}_r^{(n)} \circ \mathbf{s}_r, \quad n \in \{1, \ldots, N\}, \quad (9)$$

where N denotes the number of shift-invariance structures that have been taken into account, and where $\mathbf{a}_r^{(n)} \in \mathbb{C}^{I_n}$ and $\mathbf{d}_r^{(n)} \in \mathbb{C}^{J_n}$ denote the nth reference subarray response vector and its associated displacement vector, respectively. Obviously, if $N = 1$, then (9) reduces to (4). Similar to the CPD matrix representation (5), the coupled CPD of $\{\mathcal{X}^{(n)}\}$ has the following matrix representation:

$$\begin{bmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(1)} \odot \mathbf{D}^{(1)} \\ \vdots \\ \mathbf{A}^{(N)} \odot \mathbf{D}^{(N)} \end{bmatrix} \mathbf{S}^T \in \mathbb{C}^{\left(\sum_{n=1}^N I_n J_n\right) \times K}. \quad (10)$$

As for CPD, uniqueness conditions and algebraic algorithms for coupled CPD (9) have been developed [22], [28]. In this paper we limit the discussion to the multiple snapshot case, meaning that we assume that the factor matrix $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_R] \in \mathbb{C}^{K \times R}$ has full column rank. However, this does not mean that the factor matrices $\mathbf{A}^{(n)} = [\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}] \in \mathbb{C}^{I_n \times R}$ and $\mathbf{D}^{(n)} = [\mathbf{d}_1^{(n)}, \ldots, \mathbf{d}_R^{(n)}] \in \mathbb{C}^{J_n \times R}$ are required to have full column rank. (In fact, in array processing it is not uncommon that some of matrices in the set $\{\mathbf{A}^{(n)}, \mathbf{D}^{(n)}\}_{n=1}^N$ do not have full column rank.) Before presenting uniqueness conditions for coupled CPD, a few concepts need to be explained.

a) Definitions of coupled rank, coupled CPD and coupled PD: We define the coupled rank of $\{\mathcal{X}^{(n)}\}$ as the minimal number of coupled rank-1 tensors $\mathbf{a}_r^{(n)} \circ \mathbf{d}_r^{(n)} \circ \mathbf{c}_r$ that yield $\{\mathcal{X}^{(n)}\}$ in a linear combination. Assume that the coupled rank of $\{\mathcal{X}^{(n)}\}$ is R, then (9) will be called the coupled CPD of $\{\mathcal{X}^{(n)}\}$. On the other hand, if R in (9) is not necessarily minimal, then (9) will simply be called a coupled Polyadic Decomposition (PD) of $\{\mathcal{X}^{(n)}\}$.

b) Definition of uniqueness of coupled CPD: The coupled rank-1 tensors in (9) can be arbitrarily permuted and the vectors within the same coupled rank-1 tensor can be arbitrarily scaled provided the overall coupled rank-1 term remains the same. We say that the coupled CPD is unique when it is only subject to these trivial indeterminacies.

c) Uniqueness condition for coupled CPD: In this paper we consider the case where the signal matrix $\mathbf{S}$ has full column rank. A necessary and sufficient uniqueness condition for coupled CPD with full column rank $\mathbf{S}$ can be found in [22], and is summarized in the following theorem.

Theorem III.1. *Consider the coupled PD of $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$, $n \in \{1, \ldots, N\}$ in (9). Define $\boldsymbol{E}^{(n)}(\boldsymbol{w}) = \sum_{r=1}^R w_r \boldsymbol{a}_r^{(n)} \boldsymbol{d}_r^{(n)T}$ and $\Omega = \left\{ \boldsymbol{x} \in \mathbb{C}^R \middle| \omega(\boldsymbol{x}) \geq 2 \right\}$. Assume that $\boldsymbol{S}$ has full column rank. The coupled rank of $\{\mathcal{X}^{(n)}\}$ is R and the coupled CPD of $\{\mathcal{X}^{(n)}\}$ is unique if and only if*

$$\forall \boldsymbol{w} \in \Omega\,,\ \exists\, n \in \{1, \ldots, N\} \,:\, r\left(\boldsymbol{E}^{(n)}(\boldsymbol{w})\right) \geq 2\,. \quad (11)$$

In practice, condition (11) can be hard to check. For this reason we will in this section resort to the following more easy to check and yet very powerful uniqueness condition.

Theorem III.2. *Consider the coupled PD of $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$, $n \in \{1, \ldots, N\}$ in (9). Define*

$$\boldsymbol{G} = \begin{bmatrix} C_2\left(\boldsymbol{A}^{(1)}\right) \odot C_2\left(\boldsymbol{D}^{(1)}\right) \\ \vdots \\ C_2\left(\boldsymbol{A}^{(N)}\right) \odot C_2\left(\boldsymbol{D}^{(N)}\right) \end{bmatrix} \in \mathbb{C}^{\left(\sum_{n=1}^N C_{I_n}^2 C_{J_n}^2\right) \times C_R^2}. \quad (12)$$

If

$$\begin{cases} \boldsymbol{S} \text{ in (9) has full column rank,} & (13a) \\ \boldsymbol{G} \text{ in (12) has full column rank,} & (13b) \end{cases}$$

then the coupled rank of $\{\mathcal{X}^{(n)}\}$ is R, the coupled CPD of $\{\mathcal{X}^{(n)}\}$ is unique, and the coupled CPD can be computed via an EVD [22], [28].

From the definition of $\mathbf{G}$ it is clear that more relaxed uniqueness conditions are obtained by taking the coupling into account; indeed, every PD in the set contributes a submatrix to $\mathbf{G}$. Subsections III-B and III-C demonstrate that the coupled CPD model (9) plays an important role in array processing problems involving shift-invariance structures.

B. Arrays composed of linear or rectangular arrays

Several practical array configurations are based on the combination of uniform linear or rectangular arrays (ULA, URA). Such configurations can be handled in the coupled CPD modeling framework. Despite the generality of configurations that can be described, coupled CPD yields explicit uniqueness conditions. Moreover, the EVD in Theorem III.2 can be interpreted as an extension of ESPRIT to such geometries. We discuss the increasingly complex cases of L-shaped, triangular-shaped and frame-shaped arrays.

1) L-shaped array: A simple first example is the L-shaped array depicted in Figure 1 (left), which is clearly composed of two ULAs. Assume that K snapshots are available. The outputs of the two ULAs can be stacked into the matrices $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ according to $(\mathbf{Y}^{(1)})_{ik} = y_{i1k}$ and $(\mathbf{Y}^{(2)})_{ik} = y_{1ik}$ where the index k denotes the snapshot number. This leads to the matrix factorizations $\mathbf{Y}^{(1)} = \mathbf{B}^{(1)}\mathbf{S}^T$ and $\mathbf{Y}^{(2)} = \mathbf{B}^{(2)}\mathbf{S}^T$, where $\mathbf{S} \in \mathbb{C}^{K\times R}$ is the collected signal matrix and

$$\begin{aligned}\mathbf{b}_r^{(1)} &= [1, z_{1,r}, \ldots, z_{1,r}^{(I_1-1)}], \quad z_{1,r} = e^{-\sqrt{-1}\omega_c d_x \sin(\phi_r)\cos(\theta_r)/c},\\ \mathbf{b}_r^{(2)} &= [1, z_{2,r}, \ldots, z_{2,r}^{(I_2-1)}], \quad z_{2,r} = e^{-\sqrt{-1}\omega_c d_y \sin(\phi_r)\cos(\theta_r)/c}.\end{aligned}$$

As in ESPRIT, one can from each ULA subarray data matrix $\mathbf{Y}^{(n)}$ build a tensor $\mathcal{X}^{(n)} = \sum_{r=1}^{R}\begin{bmatrix}1\\ z_{n,r}\end{bmatrix} \circ \underline{\mathbf{b}}_r^{(n)} \circ \mathbf{s}_r \in \mathbb{C}^{2\times(I_n-1)\times K}$ with matrix representation

$$\mathbf{X}^{(n)} = \begin{bmatrix}\underline{\mathbf{Y}}^{(n)}\\ \overline{\mathbf{Y}}^{(n)}\end{bmatrix} = \begin{bmatrix}\underline{\mathbf{B}}^{(n)}\mathbf{S}^T\\ \overline{\mathbf{B}}^{(n)}\mathbf{S}^T\end{bmatrix} = (\mathbf{A}^{(n)} \odot \mathbf{D}^{(n)})\mathbf{S}^T, \qquad (14)$$

where $\mathbf{A}^{(n)} = \begin{bmatrix}1 & \cdots & 1\\ z_{n,1} & \cdots & z_{n,R}\end{bmatrix}$ and $\mathbf{D}^{(n)} = \underline{\mathbf{B}}^{(n)}$. Considering the overall array, we obtain

$$\begin{bmatrix}\mathbf{X}^{(1)}\\ \mathbf{X}^{(2)}\end{bmatrix} = \begin{bmatrix}\mathbf{A}^{(1)} \odot \mathbf{D}^{(1)}\\ \mathbf{A}^{(2)} \odot \mathbf{D}^{(2)}\end{bmatrix}\mathbf{S}^T. \qquad (15)$$

From (10) it is clear that (15) corresponds to a matrix representation of a coupled CPD of the form (9) with $N = 2$. The CPD structure of the individual data tensors $\mathcal{X}^{(1)}, \mathcal{X}^{(2)}$ is thanks to the fact that each of them corresponds to a ULA. On the other hand, the coupling is due to the fact that $\mathbf{S}$ is the same for the two subarrays. Comparing (8) with (15) it is clear that this coupled CPD approach can be interpreted as an L-shaped array variant of ESPRIT.

2) Triangular-shaped array: Associating a coupled CPD with a given array configuration can be less straightforward than in the case of the L-shaped array. Let us consider the nine-element triangular array depicted in Figure 1 (right). Let x_{mnk} denote the output of sensor x_{mn} at snapshot k. Then the kth column of $\mathbf{X}$ and the rth column of $\mathbf{A}$ in (3) are given by

$$\begin{aligned}\mathbf{x}_k &= [x_{11k}, x_{12k}, x_{13k}, x_{14k}, x_{21k}, x_{23k}, x_{31k}, x_{32k}, x_{41k}]^T,\\ \mathbf{a}_r &= \left[1, y_r, y_r^2, y_r^3, z_r, z_r y_r^2, z_r^2, z_r^2 y_r, z_r^3\right]^T,\end{aligned}$$

where $y_r = e^{-\sqrt{-1}\frac{\omega_c}{c} d_y \sin(\theta_r)\sin(\phi_r)}$ and $z_r = e^{-\sqrt{-1}\frac{\omega_c}{c} d_u \cos(\theta_r - \alpha)\sin(\phi_r)}$ in which α denotes the angle between the vertical axis and the diagonal axis of the triangular array, see Figure 1 (right). At first sight one may view the triangular array as three superimposed ULAs each composed of three sensors. However, we can also see the triangular array as a collection of three subarrays, each composed of *four* sensors that also enjoy a shift-invariance property (one of the four-element subarrays is highlighted in the figure by squared boxes). More precisely, the triangular array has the following three shift-invariance structures:

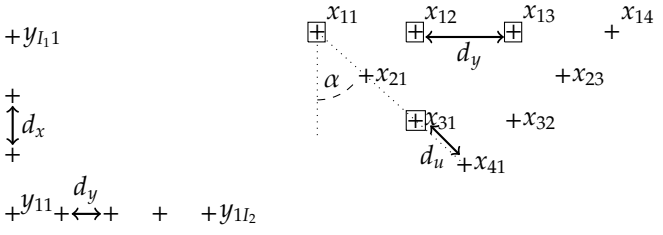


Fig. 1. L-shaped array (left) and triangular-shaped array (right) where '+' represents an antenna element.

$$\mathbf{d}_r^{(1)} := \mathbf{a}_r([1,2,3,7]) = \mathbf{a}_r([2,3,4,8])\, y_r^{-1}, \qquad (16)$$

$$\mathbf{d}_r^{(2)} := \mathbf{a}_r([1,5,7,3]) = \mathbf{a}_r([5,7,9,7])\, z_r^{-1}, \qquad (17)$$

$$\mathbf{d}_r^{(3)} := \mathbf{a}_r([4,6,8,2]) = \mathbf{a}_r([6,8,9,5])\,(z_r/y_r)^{-1}. \qquad (18)$$

This motivates the construction of the tensors $\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, \mathcal{X}^{(3)} \in \mathbb{C}^{2\times 4\times K}$ with matrix representation

$$\begin{aligned}\mathbf{X}^{(1)} &= [\mathbf{X}([1,2,3,7],:)^T, \mathbf{X}([2,3,4,8],:)^T]^T,\\ \mathbf{X}^{(2)} &= [\mathbf{X}([1,5,7,3],:)^T, \mathbf{X}([5,7,9,7],:)^T]^T,\\ \mathbf{X}^{(3)} &= [\mathbf{X}([4,6,8,2],:)^T, \mathbf{X}([6,8,9,5],:)^T]^T.\end{aligned}$$

The matrix representation of the tensors $\{\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, \mathcal{X}^{(3)}\}$ admits the coupled decomposition

$$\begin{bmatrix}\mathbf{X}^{(1)}\\ \mathbf{X}^{(2)}\\ \mathbf{X}^{(3)}\end{bmatrix} = \begin{bmatrix}\mathbf{A}^{(1)} \odot \mathbf{D}^{(1)}\\ \mathbf{A}^{(2)} \odot \mathbf{D}^{(2)}\\ \mathbf{A}^{(3)} \odot \mathbf{D}^{(3)}\end{bmatrix}\mathbf{S}^T, \qquad (19)$$

where the columns of $\mathbf{D}^{(1)}$, $\mathbf{D}^{(2)}$, $\mathbf{D}^{(3)}$ are of the form (16)–(18) and the columns of $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$, $\mathbf{A}^{(3)}$ are given by $\mathbf{a}_r^{(1)} = \begin{bmatrix}1\\ y_r\end{bmatrix}$, $\mathbf{a}_r^{(2)} = \begin{bmatrix}1\\ z_r\end{bmatrix}$ and $\mathbf{a}_r^{(3)} = \begin{bmatrix}1\\ z_r/y_r\end{bmatrix}$, respectively. Note that if we only exploit the CPD structure of, say $\mathcal{X}^{(1)}$, then from Theorem III.2 with $N = 1$ we obtain the upper bound $R \leq 4$. By exploiting the coupled CPD structure of $\{\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, \mathcal{X}^{(3)}\}$, Theorem III.2 with $N = 3$ relaxes the bound to $R \leq 6$.

3) Frame-shaped array: Consider the frame-shaped array depicted in Figure 2 (left). Compared to the L-shaped array, there is an additional shift-invariance structure between the two vertical subarrays and between the two horizontal subarrays. We explain that this additional shift-invariance structure can be integrated in the coupled CPD formulation. We build the observation matrices $\mathbf{Y}^{(1)} \in \mathbb{C}^{2I_1\times K}$ with columns $\mathbf{y}_k^{(1)} = [y_{11k}, \ldots, y_{I_1 1k}, y_{1I_2k}, \ldots, y_{I_1I_2k}]^T$ and $\mathbf{Y}^{(2)} \in \mathbb{C}^{2I_2\times K}$ with columns $\mathbf{y}_k^{(2)} = [y_{11k}, \ldots, y_{1I_2k}, y_{I_1 1k}, \ldots, y_{I_1I_2k}]^T$, where y_{ijk}

denotes the output of sensor y_{ij} at the kth snapshot and with indexing as indicated in Figure 2 (left). The matrices $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ admit the factorizations

$$\mathbf{Y}^{(1)} = \left(\mathbf{B}^{(1)} \odot \mathbf{C}^{(1)}\right)\mathbf{S}^T, \quad \mathbf{Y}^{(2)} = \left(\mathbf{B}^{(2)} \odot \mathbf{C}^{(2)}\right)\mathbf{S}^T, \tag{20}$$

where

$$\mathbf{b}_r^{(1)} = [1 \; z_{1,r} \; \dots \; z_{1,r}^{(I_1-1)}]^T, \quad \mathbf{c}_r^{(1)} = [1 \; z_{2,r}^{(I_2-1)}]^T,$$
$$\mathbf{b}_r^{(2)} = [1 \; z_{2,r} \; \dots \; z_{2,r}^{(I_2-1)}]^T, \quad \mathbf{c}_r^{(2)} = [1 \; z_{1,r}^{(I_1-1)}]^T,$$

in which $z_{1,r} = e^{-\sqrt{-1}\omega_c d_x \sin(\phi_r)\cos(\theta_r)/c}$ where d_x denotes the inter-element spacing along the x-axis, and $z_{2,r} = e^{-\sqrt{-1}\omega_c d_y \sin(\phi_r)\sin(\theta_r)/c}$ where d_y denotes the inter-element spacing along the y-axis. Clearly (20) corresponds to a matrix representation of the decomposition of the two tensors $\mathcal{Y}^{(1)} \in \mathbb{C}^{2\times(I_1-1)I_2\times K}$ and $\mathcal{Y}^{(2)} \in \mathbb{C}^{2\times I_1(I_2-1)\times K}$ that take the shift-invariance between the two vertical subarrays and between the two horizontal subarrays into account, respectively. The next step is to build tensors that exploit the shift-invariance of the vertical and horizontal ULA. Similar to the L-shaped case, this structure can be taken into account by building a set of tensors $\{\mathcal{X}^{(1)}, \mathcal{X}^{(2)}\}$ whose matrix representation

$$\mathbf{X}^{(1)} = \begin{bmatrix} (\underline{\mathbf{I}}_{I_1} \otimes \mathbf{I}_{I_2})\mathbf{Y}^{(1)} \\ (\bar{\mathbf{I}}_{I_1} \otimes \mathbf{I}_{I_2})\mathbf{Y}^{(1)} \end{bmatrix}, \quad \mathbf{X}^{(2)} = \begin{bmatrix} (\underline{\mathbf{I}}_{I_2} \otimes \mathbf{I}_{I_1})\mathbf{Y}^{(2)} \\ (\bar{\mathbf{I}}_{I_2} \otimes \mathbf{I}_{I_1})\mathbf{Y}^{(2)} \end{bmatrix}$$

admits the coupled CPD interpretation:

$$\begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(1)} \odot \mathbf{D}^{(1)} \\ \mathbf{A}^{(2)} \odot \mathbf{D}^{(2)} \end{bmatrix}\mathbf{S}^T, \tag{21}$$

where $\mathbf{A}^{(n)} = \begin{bmatrix} 1 & \cdots & 1 \\ z_{n,1} & \cdots & z_{n,R} \end{bmatrix}$ and $\mathbf{D}^{(n)} = \underline{\mathbf{B}}^{(n)} \odot \mathbf{C}^{(n)}$. Note that the latter is just an augmented version of the factor matrix ($\mathbf{D}^{(n)} = \underline{\mathbf{B}}^{(n)}$) in (14) associated with the L-shaped array. The coupling of $\{\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, \mathcal{Y}^{(1)}, \mathcal{Y}^{(2)}\}$, as expressed by (20)–(21), leads to improved uniqueness conditions. By way of example, consider the case where $I_1 = I_2 = 5$ and $K \geq R$. If we only exploit the CPD structure of $\mathcal{X}^{(1)}$, which exploits the shift-invariance structures of the vertical subarrays, then from Theorem III.2 with $N = 1$ (i.e., ordinary CPD) we obtain the bound $R \leq 8$. Thanks to Theorem III.2 with $N = 2$ the coupled CPD structure of $\{\mathcal{X}^{(1)}, \mathcal{X}^{(2)}, \mathcal{Y}^{(1)}, \mathcal{Y}^{(2)}\}$ relaxes the bound to $R \leq 10$.

Other antenna arrays that consist of superimposed linear subarrays, possibly also with identical translated subarrays, include cross-shaped [8] and hexagonal arrays [34]. From the derivation in this section it follows that the coupled CPD based results in [24], [26], [27] can directly be used for such array configurations.

C. Centro-symmetric arrays

In this section we consider centro-symmetric arrays, i.e., arrays that are symmetric around their "epicenter". In particular, we will argue that the centro-symmetry of an array by itself can be sufficient to ensure identifiability. We will explain that the coupled CPD model, albeit not optimal, can actually be used for *any* centro-symmetric array. The key observation is that for a centro-symmetric array, we know that for any two sensors there exist two "mirrored" sensors so that the four sensors together form a 2-by-2 rectangular array. (For example, in Figure 2(right), sensor x_1 and x_2 together with the two corresponding "mirrored" sensors x_5 and x_6 form a 2-by-2 rectangular array. More details will be provided at the end of this subsection.) Formally speaking, let $\mathbf{\Xi}^{(n)} \in \mathbb{C}^{2\times R}$ denote a row selection matrix that selects two rows of $\mathbf{A}$ in (3), denoted by $\mathbf{D}^{(n)} = \mathbf{\Xi}^{(n)}\mathbf{A} \in \mathbb{C}^{2\times R}$. Similarly, let $\mathbf{\Xi}_{(n)} \in \mathbb{C}^{2\times R}$ denote a row selection matrix that selects the two corresponding "mirrored" rows of $\mathbf{A}$ so that $\mathbf{D}^{(n)}\mathbf{\Lambda}^{(n)} = \mathbf{\Xi}_{(n)}\mathbf{A}$, where $\mathbf{\Lambda}^{(n)} \in \mathbb{C}^{R\times R}$ is a diagonal matrix. We emphasize that due to the centro-symmetry, there exists a row-selection pair $(\mathbf{\Xi}^{(n)}, \mathbf{\Xi}_{(n)})$ for all choices of two sensors, except for the trivial case in which $\mathbf{\Xi}^{(n)}$ selects two sensors that already mirror each other (e.g., sensors x_1 and x_5 in Figure 2(right)). In other words, no matter which two rows the selection matrix $\mathbf{\Xi}^{(n)}$ has chosen, there always exists a selection matrix $\mathbf{\Xi}_{(n)}$ that picks two rows from $\mathbf{A}$ so that we can build a tensor $\mathcal{X}^{(n)} \in \mathbb{C}^{2\times 2\times R}$ whose matrix representation admits the decomposition

$$\mathbf{X}^{(n)} = \begin{bmatrix} \mathbf{\Xi}^{(n)}\mathbf{X} \\ \mathbf{\Xi}_{(n)}\mathbf{X} \end{bmatrix} = (\mathbf{A}^{(n)} \odot \mathbf{D}^{(n)})\mathbf{S}^T \in \mathbb{C}^{4\times K}, \tag{22}$$

where $\mathbf{A}^{(n)} = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_{11}^{(n)} & \cdots & \lambda_{RR}^{(n)} \end{bmatrix}$ in which $\lambda_{rr}^{(n)} \in \mathbb{C}$ corresponds to the rth diagonal element of $\mathbf{\Lambda}^{(n)}$. Assume that for N choices of sensor pairs, we construct tensors $\mathcal{X}^{(n)} \in \mathbb{C}^{2\times 2\times R}$, $n \in \{1, \dots, N\}$, each with a matrix factorization of the form (22). Theorem III.2 now tells us that if

$$\begin{cases} \mathbf{G} \text{ in (12) has full column rank,} & \text{(23a)} \\ \mathbf{S} \text{ in (3) has full column rank,} & \text{(23b)} \end{cases}$$

where $\mathbf{G}$ is built from $\{\mathbf{A}^{(n)}, \mathbf{D}^{(n)}\}_{n=1}^N$ in (22), then the coupled CPD of $\{\mathcal{X}^{(n)}\}$ is unique and can actually be computed via a matrix EVD. Consequently, if conditions (23a)–(23b) are satisfied, then $\mathbf{S}$ is essentially unique[1]. This in turn implies that $\mathbf{A} = \mathbf{X}(\mathbf{S}^T)^\dagger$ in the original matrix factorization (3) is also essentially unique.

By way of example, let us consider the specific Uniform Circular Array (UCA) depicted in Figure 2(right). For a UCA composed of I sensors the entries of $\mathbf{A}$ in (3) are given by $a_{i,r} = e^{\sqrt{-1}\cdot 2\pi\cdot d\cdot\cos(\theta_r-\gamma_i)\cdot\sin(\phi_r)/\lambda}$, in which d is the radius of the UCA, λ is the signal wavelength, $\gamma_i = 2\pi(i-1)/I$, θ_r is the azimuth angle of the rth source and ϕ_r is the elevation angle of the rth source. Note that any UCA with an even number of sensors is centro-symmetric. For the particular example in Figure 2(right) with $I = 8$, the two sensors indexed by the pair (x_1, x_2) together with the two "mirrored" sensors indexed by the pair (x_5, x_6) constitute a 2-by-2 rectangular subarray. Denote this correspondence by $(x_1, x_2) \leftrightarrow (x_5, x_6)$. We

[1] A matrix, say $\mathbf{S}$, is said to be essentially unique if it is unique up to column scaling and permutation ambiguities.

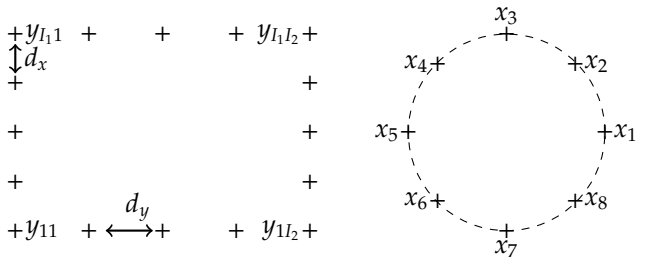


Fig. 2. Centro-symmetric frame-shaped array (left) and circular array (right) where '+' represents an antenna element.

then have the following six couples $(x_1, x_2) \leftrightarrow (x_5, x_6)$, $(x_1, x_3) \leftrightarrow (x_5, x_7)$, $(x_1, x_4) \leftrightarrow (x_5, x_8)$ $(x_2, x_3) \leftrightarrow (x_6, x_7)$, $(x_2, x_4) \leftrightarrow (x_6, x_8)$, $(x_2, x_5) \leftrightarrow (x_7, x_8)$, each corresponding to a 2-by-2 rectangular subarray. More generally, we can extract $C^2_{I/2}$ 2-by-2 rectangular arrays from $\mathbf{X}$, each with a matrix representation:

$$\mathbf{X}^{(m,n)} = \begin{bmatrix} \mathbf{\Xi}^{(m,n)}\mathbf{X} \\ \mathbf{\Xi}^{(n+I/2,m+I/2)}\mathbf{X} \end{bmatrix} = (\mathbf{A}^{(m,n)} \odot \mathbf{D}^{(m,n)})\mathbf{S}^T \in \mathbb{C}^{4\times K}, \tag{24}$$

where $1 \leq m \leq I/2$ and $m < n \leq I/2$, $\mathbf{\Xi}^{(p,q)} \in \mathbb{C}^{2\times I}$ is a row-selection matrix that selects rows p and q, $\mathbf{A}^{(m,n)} = \begin{bmatrix} 1 & \cdots & 1 \\ d_1^{(m,n)} & \cdots & d_R^{(m,n)} \end{bmatrix}$ in which $d_r^{(m,n)} \in \mathbb{C}$ has the property $\mathbf{a}_r^{(n+I/2,m+I/2)} = \mathbf{a}_r^{(m,n)} d_r^{(m,n)}$, and $\mathbf{D}^{(m,n)} = \mathbf{A}([m,n],:) \in \mathbb{C}^{2\times R}$. Clearly the set of matrices $\{\mathbf{X}^{(m,n)}\}$ with $1 \leq m \leq I/2$ and $m < n \leq I/2$ corresponds to a matrix representation of a coupled CPD. Note that this coupled CPD approach does not fully exploit the *rotational* invariance of the UCA and, hence, does not provide the most relaxed bound on the number of signals that can be separated in a unique way. On the other hand, it provides a uniqueness condition that is easy-to-check and valid for *any* centro-symmetric array. Note that each row of $\mathbf{G}$ in condition (13b) is associated with a 2-by-2 rectangular subarray. Since we consider $C^2_4 = 6$ subarrays, $\mathbf{G}$ will be an $(6\times C^2_R)$ matrix. It can be verified that this matrix generically has full column rank if $R \leq 3$. Consequently, Theorem III.2 generically guarantees uniqueness for this UCA if $R \leq 3$.

Note that due to the centro-symmetry property, forward-backward averaging [16] can be used in a pre-processing step to relax condition (23b). To make things more tangible, let us focus again on the above UCA. Observe that any UCA with an even number of I sensors has the property $a_{(n+I/2),r} = a^*_{n,r}$. This property implies that

$$\mathbf{Y} = [\mathbf{X}, (\mathbf{J}_2 \otimes \mathbf{I}_{I/2})\mathbf{X}^*] = \mathbf{A}\widetilde{\mathbf{S}}^T, \quad \widetilde{\mathbf{S}} = \begin{bmatrix} \mathbf{S} \\ \mathbf{S}^* \end{bmatrix},$$

where $\mathbf{J}_2 = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. We can now proceed with $\mathbf{S}$ in (3) and in (23b) replaced by the augmented matrix $\widetilde{\mathbf{S}} \in \mathbb{C}^{2K\times R}$, which generically has full column rank if $2K \geq R$ [23].

IV. Monomial structure in array processing

In this section we will extend our tensor-based framework for array processing problems involving multiple shift-invariance structures to more general monomial structures. (A monomial is a product of variables, possibly with repetitions.) More precisely, we now consider bilinear factorizations of the form

$$\mathbf{X} = \mathbf{A}\mathbf{S}^T \in \mathbb{C}^{I\times K}, \tag{25}$$

in which the columns of $\mathbf{A} \in \mathbb{C}^{I\times R}$ (or similarly the columns of $\mathbf{S} \in \mathbb{C}^{K\times R}$) enjoy monomial relations of the form $a_0^{\alpha_0} a_{p_1}^{\beta_1} \cdots a_{p_q}^{\beta_q} - a_0^{\gamma_0} a_{s_1}^{\omega_1} \cdots a_{s_t}^{\omega_t} = 0$. It is not obvious to see that this will indeed generalize ESPRIT and coupled CPD. At this point, let us just mention that in Subsection IV-C1 we explain that ESPRIT exploits monomial relations of the form $x^{n+1}x^{n-1} - x^{2n} = 0$. Similarly, in Subsection IV-C2, we explain that coupled CPD exploits monomial relations of the form $ab - cd = 0$.

The section is organised as follows. By way of motivation, we will first in Subsection IV-A discuss some applications in array processing. Next, in Subsection IV-B we establish a link between coupled tensor decompositions and low-rank matrix factorizations with monomial structures. This link yields BTD-based uniqueness conditions and EVD-based algorithms. In Subsection IV-C we explain that this "monomial" approach is an extension of the ESPRIT [17], CPD [18] and coupled CPD [22], [28] approaches for array processing problems. Finally, in Subsection IV-D we make a connection with ACMA [33] and develop a CPD-based formulation.

A. Motivating array processing applications

1) Separation of M-PSK signals using an antenna array: The separation of digital communication signals is probably one of the earliest array processing examples involving bilinear factorizations with monomial structures. For instance, blind separation of M-PSK signals in which the entries of $\mathbf{S}$ in (25) take the form

$$s_{kr} = e^{\sqrt{-1}u_{kr}} \quad \text{with} \quad u_{kr} \in \left\{0, \frac{2\pi}{M}, \ldots, \frac{2\pi(M-1)}{M}\right\} \tag{26}$$

has been considered (e.g. [31], [9]). From (26) it is clear that $s^M_{k_1 r} = s^M_{k_2 r}$ for all $k_1, k_2 \in \{1, \ldots, K\}$. In other words, for every pair (k_1, k_2) with property $k_1 < k_2$, we can exploit the monomial relation

$$s^M_{k_1 r} - s^M_{k_2 r} = 0\,. \tag{27}$$

Overall, we obtain a set of C^2_K relations of the form (27). In Subsections IV-B and IV-D we will explain how to translate this problem into a tensor decomposition.

2) DOA estimation using a non-uniform array with grid structures: Another classical example is narrowband DOA estimation using an antenna array configuration with sensors located on points of one or several rectangular grids. By way of example, let us consider the one-dimensional case in which the columns of $\mathbf{A}$ in (25) take the form

$$\mathbf{a}_r = [1 \; x_r^3 \; x_r^5 \; x_r^8 \; x_r^{13} \; x_r^{21}]^T, \quad r \in \{1, \ldots, R\}, \tag{28}$$

where $x_r = e^{-\sqrt{-1}\omega_c d_x \cos(\theta_r)/c}$ in which $0 \leq \theta_r < \pi$ denotes the azimuth angle associated with the rth source, d_x is the

nominal unit measure along the sensor axis ("x-axis"), ω_c is the carrier frequency, and c is the propagation speed. It is clear that the vector given by (28) enjoys several monomial relations, such as

$$a_{1r}^1 a_{4r}^1 = a_{2r}^1 a_{3r}^1 \Leftrightarrow 1^1 (x_r^8)^1 = (x_r^3)^1 (x_r^5)^1, \tag{29}$$
$$a_{1r}^1 a_{5r}^1 = a_{3r}^1 a_{4r}^1 \Leftrightarrow 1^1 (x_r^{13})^1 = (x_r^5)^1 (x_r^8)^1, \tag{30}$$
$$a_{1r}^1 a_{6r}^1 = a_{4r}^1 a_{5r}^1 \Leftrightarrow 1^1 (x_r^{21})^1 = (x_r^8)^1 (x_r^{13})^1. \tag{31}$$

This "monomial" approach also provides a better understanding of nonredundant arrays [35], such as

$$\mathbf{a}_r = [1 \; x_r \; x_r^4 \; x_r^9 \; x_r^{11}]^T. \tag{32}$$

Note that in [24] we only considered cases in which the array response vectors enjoy several shift-invariance structures. This is clearly not the case for array response vectors associated with nonredundant arrays. However, they do enjoy monomial structure. For example, the array response vector (32) enjoys the following relations

$$a_{2r}^2 a_{4r}^1 = a_{1r}^2 a_{5r} \Leftrightarrow (x_r)^2 (x_r^9)^1 = 1^2 (x_r^{11})^1, \tag{33}$$
$$a_{2r}^1 a_{3r}^2 = a_{1r}^2 a_{4r}^1 \Leftrightarrow (x_r)^1 (x_r^4)^2 = 1^2 (x_r^9)^1, \tag{34}$$
$$a_{1r}^1 a_{3r}^1 a_{4r}^1 = a_{2r}^2 a_{5r}^1 \Leftrightarrow 1^1 (x_r^4)^1 (x_r^9)^1 = (x_r^1)^2 (x_r^{11})^1, \tag{35}$$
$$a_{3r}^3 = a_{1r}^1 a_{2r}^1 a_{5r}^1 \Leftrightarrow (x_r^4)^3 = 1^1 (x_r^1)^1 (x_r^{11})^1, \tag{36}$$
$$a_{3r}^1 a_{4r}^2 = a_{1r}^1 a_{5r}^2 \Leftrightarrow (x_r^4)^1 (x_r^9)^2 = 1^1 (x_r^{11})^2. \tag{37}$$

3) Near-field array processing: Bilinear factorization problems of the form (25) involving monomial relations also occur in near-field array processing problems in which the columns of $\mathbf{A}$ in (25) are approximated by polynomials [30]. As an example, consider a 9-element linear array in which a second-order Taylor series approximation is used [10]:

$$\mathbf{a}_r^{(n)} = [a_{-4r}, a_{-3r}, a_{-2r}, a_{-1r}, a_{0r}, a_{1r}, a_{2r}, a_{3r}, a_{4r}]^T, \tag{38}$$

where $a_{mr} = e^{\sqrt{-1}(\omega_r \cdot m + \gamma_r \cdot m^2)}$ with $\omega_r = -\frac{2\pi d_x}{\lambda}\sin(\theta_r)$ and $\gamma_r = \frac{\pi d_x^2}{\lambda \Delta_r}\cos^2(\theta_r)$ in which the new variable Δ_r denotes the range associated with the rth source. The vector (38) enjoys several monomial relations, such as

$$a_{-2} a_1^2 = a_2 a_{-1}^2 = \gamma^6, \tag{39}$$
$$a_{-3} a_1 a_2 = a_3 a_{-1} a_{-2} = \gamma^{14}, \tag{40}$$
$$a_2^2 a_{-4} = a_{-2}^2 a_4 = \gamma^{24}, \tag{41}$$
$$a_1 a_3 a_{-4} = a_{-1} a_{-3} a_4 = \gamma^{25}. \tag{42}$$

B. Link between low-rank factorization with monomial constraints and coupled block term decompositions

1) Coupled BTD: In this section we explain that a low-rank factorization with monomial constraints can be translated into a coupled BTD. Recall first that the multilinear rank-$(P,P,1)$ term decomposition [2] of a tensor is an extension of the CPD (4), where each term in the decomposition now consists of the outer product of a vector and a matrix that is low-rank. More formally, $\mathbf{a}_r \circ \mathbf{d}_r \circ \mathbf{s}_r$ in (4) is replaced by $\mathbf{G}_r \circ \mathbf{s}_r$, in which $\mathbf{G}_r \in \mathbb{C}^{I \times J}$ is a rank-P matrix with $\min(I,J) > P$. We will consider the extension of the coupled CPD (9), in which the set of tensors $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$, $n \in \{1,\ldots,N\}$ is now decomposed into a sum of coupled multilinear rank-$(P,P,1)$ terms [22], or coupled BTD for short:

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{G}_r^{(n)} \circ \mathbf{s}_r \;, \quad n \in \{1,\ldots,N\}, \tag{43}$$

where $\mathbf{G}_r^{(n)} \in \mathbb{C}^{I_n \times J_n}$ is a rank-P matrix with $\min(I_n, J_n) > P$ and $\mathbf{s}_r \in \mathbb{C}^K$. Note that if $P = 1$, then (43) indeed reduces to (9) with $\mathbf{G}_r^{(n)} = \mathbf{a}_r^{(n)} \mathbf{d}_r^{(n)T} = \mathbf{a}_r^{(n)} \circ \mathbf{d}_r^{(n)}$.

2) Matrix implementation of monomial structure: Consider a vector of the form $\mathbf{a} = [a_0 \; a_1 \; \ldots \; a_{I-1}]^T \in \mathbb{C}^I$ with $a_0 = 1$ and monomial relations[2]

$$a_0^{\alpha_0} a_{p_1}^{\beta_1} \cdots a_{p_q}^{\beta_q} = a_0^{\gamma_0} a_{s_1}^{\omega_1} \cdots a_{s_t}^{\omega_t} \Leftrightarrow a_0^{\alpha_0} a_{p_1}^{\beta_1} \cdots a_{p_q}^{\beta_q} - a_0^{\gamma_0} a_{s_1}^{\omega_1} \cdots a_{s_t}^{\omega_t} = 0 \tag{44}$$

subject to $L := \alpha_0 + \beta_1 + \cdots + \beta_q = \gamma_0 + \omega_1 + \cdots + \omega_t$. Define the vectors

$$\mathbf{b} = \begin{bmatrix} \mathbf{1}_{\alpha_1} \\ a_{p_1}\mathbf{1}_{\beta_1} \\ \vdots \\ a_{p_q}\mathbf{1}_{\beta_q} \end{bmatrix} \in \mathbb{C}^L \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} \mathbf{1}_{\gamma_0} \\ a_{s_1}\mathbf{1}_{\omega_1} \\ \vdots \\ a_{s_t}\mathbf{1}_{\omega_t} \end{bmatrix} \in \mathbb{C}^L. \tag{45}$$

Then relation (44) can be related to the matrix

$$\mathbf{A}_L(\mathbf{b},\mathbf{c}) := \begin{bmatrix} b_1 & 0 & \cdots & 0 & (-1)^L \cdot c_1 \\ c_2 & b_2 & \ddots & & 0 \\ 0 & c_3 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_L & b_L \end{bmatrix} \in \mathbb{C}^{L \times L}. \tag{46}$$

From the cofactor expansion of $|\mathbf{A}_L(\mathbf{b},\mathbf{c})|$ along the first row, the connection between (44) and (46) becomes clear:

$$\begin{aligned} |\mathbf{A}_L(\mathbf{b},\mathbf{c})| &= b_1 \cdot \begin{vmatrix} b_2 & 0 & \cdots & 0 \\ c_3 & \ddots & \ddots & \vdots \\ & \ddots & \ddots & 0 \\ & & c_L & b_L \end{vmatrix} \\ &\quad + (-1)^L \cdot c_1 \cdot (-1)^{L+1} \cdot \begin{vmatrix} c_2 & b_2 & & \\ 0 & c_3 & \ddots & \\ \vdots & \ddots & \ddots & b_{L-1} \\ 0 & \cdots & 0 & c_L \end{vmatrix} \\ &= \prod_{n=1}^{L} b_n - \prod_{m=1}^{L} c_m = 0, \end{aligned} \tag{47}$$

where we exploited that the two involved $(L-1)\times(L-1)$ minors in (47) are triangular. The determinant property (47) also explains that $\mathbf{A}_L(\mathbf{b},\mathbf{c})$ is low-rank under the condition (44). In fact, since the minors in (47) do not vanish under condition (44), $\mathbf{A}_L(\mathbf{b},\mathbf{c})$ will be a rank-$(L-1)$ matrix (the only possible exception is the trivial case where $\prod_{m=1}^{q} a_{p_m} = 0$ and $\prod_{n=1}^{t} a_{s_n} = 0$).

[2]For ease of presentation and explanation we limit the discussion to monomials of the same degree L. However, the results presented in the paper can be extended to monomials of varying degree or to the sum of monomials, as explained in [20].

3) Uniqueness and computation: Consider the bilinear factorization (25) in which the columns of $\mathbf{A}$ enjoy N monomial relations of the form (44). The mapping (46) enables us to transform (25) into a coupled BTD. In more detail, for every monomial relation $n \in \{1,\dots,N\}$, we build a tensor $\mathcal{X}^{(n)} \in \mathbb{C}^{L\times L\times K}$ with matrix slices $\mathcal{X}^{(n)}(:,:,k) = \mathbf{A}_L(\mathbf{y}_k^{(n)}, \mathbf{z}_k^{(n)})$ in which the vectors $\mathbf{y}_k^{(n)}, \mathbf{z}_k^{(n)} \in \mathbb{C}^L$ are constructed from the entries of the kth column of $\mathbf{X}$ in accordance to the nth monomial relation. Overall, this yields the coupled BTD

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{A}_L(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)}) \circ \mathbf{s}_r, \quad n \in \{1,\dots,N\}, \tag{48}$$

where the vectors $\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)} \in \mathbb{C}^L$ are constructed from the entries of the kth column of $\mathbf{A}$ in (25) in accordance to the nth monomial relation. Theorem IV.1 below is an adaptation of the necessary and sufficient uniqueness condition for coupled BTD [22, Theorem 5.5] to the monomial case.

Theorem IV.1. *Consider the coupled BTD of $\mathcal{X}^{(n)} \in \mathbb{C}^{L\times L\times K}$, $n \in \{1,\dots,N\}$ in (48). Define $\boldsymbol{E}^{(n)}(\boldsymbol{w}) = \sum_{r=1}^{R} w_r \boldsymbol{A}_L(\boldsymbol{b}_r^{(n)}, \boldsymbol{c}_r^{(n)})$ and $\Omega = \left\{\boldsymbol{x} \in \mathbb{C}^R \middle| \omega(\boldsymbol{x}) \geq 2\right\}$. Assume that $\boldsymbol{S}$ has full column rank. The decomposition of $\{\mathcal{X}^{(n)}\}$ into coupled multilinear rank-$(L-1, L-1, 1)$ terms is unique if and only if*

$$\forall \boldsymbol{w} \in \Omega\,,\ \exists n \in \{1,\dots,N\} \,:\, r_{\boldsymbol{E}^{(n)}(\boldsymbol{w})} \geq \max_{r|w_r\neq 0} L\,. \tag{49}$$

Theorem IV.1 extends Theorem III.1 to the monomial case. Like condition (11), condition (49) can be hard to check. For this reason, we also derive a monomial variant of an easier-to-check but only sufficient uniqueness condition for coupled BTD that was derived in [28]. The condition assumes that the dimension of the third mode has been reduced to R. In cases where $\mathbf{S}$ is tall and has full column rank, this can always be achieved via for instance an SVD of $\mathbf{X}$ (full column rank of $\mathbf{A}$ is necessary for uniqueness). Briefly, let $\widetilde{\mathbf{X}} \in \mathbb{C}^{I\times R}$ and $\widetilde{\mathbf{S}} \in \mathbb{C}^{R\times R}$ denote the compressed versions of $\mathbf{X}$ and $\mathbf{S}$, i.e., $\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{V} = \mathbf{A}\widetilde{\mathbf{S}}$ and $\widetilde{\mathbf{S}} = \mathbf{S}\mathbf{V}$ in which the columns of $\mathbf{V} \in \mathbb{C}^{K\times R}$ form an orthonormal basis for range$\left(\mathbf{X}^T\right)$. Similarly, let $\widetilde{\mathcal{X}}^{(n)} \in \mathbb{C}^{L\times L\times R}$ denote the compressed version of $\mathcal{X}^{(n)}$ and let $\widetilde{\mathbf{y}}_k^{(n)}$ and $\widetilde{\mathbf{z}}_k^{(n)}$ be the associated compressed versions of $\mathbf{y}_k^{(n)}$ and $\mathbf{z}_k^{(n)}$, respectively. Consider the $(1\times R^L)$ row-vector $\mathbf{r}_L(\widetilde{\mathcal{X}}^{(n)})$ in which the

$$\left((r_L-1)R^{L-1} + (r_{L-1}-1)R^{L-2} + \cdots + (r_2-1)R + r_1\right)\text{–th entry}$$

is equal to

$$\sum_{k=1}^{L}(-1)^{L-k} \sum_{1\leq p_1<p_2<\cdots<p_k\leq L} \left|\left(\widetilde{\mathcal{X}}^{(n)}(:,:,r_{p_1}) + \cdots + \widetilde{\mathcal{X}}^{(n)}(:,:,r_{p_k})\right)\right|$$
$$= \sum_{k=1}^{L}(-1)^{L-k} \sum_{1\leq p_1<p_2<\cdots<p_k\leq L} \left(\prod_{l=1}^{L} \alpha_l^{(p_1,\dots,p_k,n)} - \prod_{l=1}^{L} \beta_l^{(p_1,\dots,p_k,n)}\right), \tag{50}$$

where the identity is due to (47) and where $\boldsymbol{\alpha}^{(p_1,\dots,p_k,n)} = \widetilde{\mathbf{y}}_{r_{p_1}}^{(n)} + \cdots + \widetilde{\mathbf{y}}_{r_{p_k}}^{(n)}$ and $\boldsymbol{\beta}^{(p_1,\dots,p_k,n)} = \widetilde{\mathbf{z}}_{r_{p_1}}^{(n)} + \cdots + \widetilde{\mathbf{z}}_{r_{p_k}}^{(n)}$. Stacking yields

$$\mathbf{R}_L^{(N)} = \begin{bmatrix} \mathbf{r}_L(\widetilde{\mathcal{X}}^{(1)}) \\ \vdots \\ \mathbf{r}_L(\widetilde{\mathcal{X}}^{(N)}) \end{bmatrix} \in \mathbb{C}^{N\times R^L}. \tag{51}$$

It can be verified that $\mathbf{r}_L(\widetilde{\mathcal{X}}^{(n)})$ has C_{R+L-1}^{L} distinct entries; actually the rows of $\mathbf{R}_L^{(N)}$ represent vectorized symmetric tensors. Denote $\mathbf{W} = [\mathbf{w}_1,\dots,\mathbf{w}_R] = \widetilde{\mathbf{S}}^{-1}$ with $\mathbf{w}_r \in \mathbb{C}^R$. Furthermore, let $\pi_S^{(L)}$ denote the subspace of vectorized $\underbrace{R\times\cdots\times R}_{L}$ symmetric tensors. [22, Lemma 5.3] implies that $\underbrace{\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r}_{L} \in \ker\left(\mathbf{R}_L^{(N)}\right) \cap \pi_S^{(L)}$. Since the R vectors in the set $\{\mathbf{w}_r\}$ are linearly independent, the R vectors in the set $\{\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r\}$ are also linearly independent. Consequently, we know that the dimension of the subspace of $\ker\left(\mathbf{R}_L^{(N)}\right) \cap \pi_S^{(L)}$ is at least R. Let the columns of $\mathbf{M}$ constitute a basis for $\ker\left(\mathbf{R}_L^{(N)}\right) \cap \pi_S^{(L)}$ and assume that the dimension of this subspace is minimal (i.e., R), then there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{M} = (\underbrace{\mathbf{W} \odot \cdots \odot \mathbf{W}}_{L})\mathbf{F}^T. \tag{52}$$

From (5) it is clear that relation (52) corresponds to a matrix representation of the CPD of a tensor $\mathcal{M} \in \mathbb{C}^{R\times\cdots\times R}$ of order $L+1$:

$$\mathcal{M} = \sum_{r=1}^{R} \mathbf{w}_1 \circ \cdots \circ \mathbf{w}_r \circ \mathbf{f}_r\,. \tag{53}$$

Note that we have simplified the coupled BTD (43) in two ways. Eq. (53) involves just a CPD, and it is not coupled. Since all the factor matrices of the CPD in (53) have full column rank, we know that in the exact case $\mathbf{W}$ can be recovered from (53) via a EVD (e.g., [11], [15]). This also means that we obtain $\mathbf{A} = \widetilde{\mathbf{X}} \cdot \mathbf{W} = (\mathbf{A}\widetilde{\mathbf{S}}) \cdot (\widetilde{\mathbf{S}}^{-1})$ and $\mathbf{S} = \mathbf{W}\mathbf{V}^H = \widetilde{\mathbf{S}}^{-1}\mathbf{V}^H$. In the inexact case, the solution obtained via (53) can be used to initialize an optimization-based method (e.g., alternating or nonlinear least squares, see [36] and references therein). Theorem IV.2 below summarizes the above uniqueness result for the coupled BTD (48). As we have just explained, the result is moreover constructive in the sense that it converts the coupled BTD problem into a CPD problem that can be solved via EVD.

Theorem IV.2. *Consider the coupled BTD of $\mathcal{X}^{(n)} \in \mathbb{C}^{L\times L\times K}$, $n \in \{1,\dots,N\}$ in (48). Let $\pi_S^{(L)}$ denote the subspace of vectorized $\underbrace{R\times\cdots\times R}_{L}$ symmetric tensors. If*

$$\begin{cases} \boldsymbol{S} \text{ has full column rank,} \\ \boldsymbol{ker}\left(\boldsymbol{R}_L^{(N)}\right) \cap \pi_S^{(L)} \text{ is an } \boldsymbol{R}\text{-dimensional subspace,} \end{cases} \tag{54}$$

then the decomposition of $\{\mathcal{X}^{(n)}\}$ into coupled multilinear rank-$(L-1, L-1, 1)$ terms is unique and it can be reduced to the

CPD of $\mathcal{M}$ in (53). *The latter decomposition can be computed algebraically via a EVD.*

Theorem IV.2 is in fact a variant of Theorem III.2 for the implementation of the monomial structure. This will be made more clear in Subsection IV-C2. We also mention that an extension of Theorem III.2 in terms of factor matrices is provided in [20].

4) Illustrations: To make things more concrete, let us briefly explain how to apply the approach to the M-PSK signal separation and DOA estimation problems discussed in Subsection IV-A.

M-PSK signals: Using the mapping (46), we can translate the identity (27) into a BTD of a tensor $\mathcal{X}^{(k_1,k_2)} \in \mathbb{C}^{M\times M\times I}$ with matrix slices

$$\mathcal{X}^{(k_1,k_2)}(:,:,i) = \mathbf{A}_M(\mathbf{b}_{ik_1}, \mathbf{c}_{ik_2}),$$

with $\mathbf{b}_{ik_1} = x_{ik_1}\mathbf{1}_M$ and $\mathbf{c}_{ik_2} = x_{ik_2}\mathbf{1}_M$, in which $\mathbf{A}_M(\mathbf{b}_{ik_1}, \mathbf{c}_{ik_2})$ is a rank-$(M-1)$ matrix. Overall, we obtain a set of C_K^2 coupled decompositions:

$$\mathcal{X}^{(k_1,k_2)} = \sum_{r=1}^{R} \mathbf{A}_M(\mathbf{b}_r^{(k_1)}, \mathbf{c}_r^{(k_2)}) \circ \mathbf{a}_r\,, \quad 1 \le k_1 < k_2 \le K\,,$$

where $\mathbf{b}_r^{(k_1)} = s_{k_1 r}\mathbf{1}_M$ and $\mathbf{c}_r^{(k_2)} = s_{k_2 r}\mathbf{1}_M$.

Nonuniform arrays: Based on the relations (29)–(31) we can build the tensors $\mathcal{X}^{(n)} \in \mathbb{C}^{2\times 2\times K}$, $n \in \{1,2,3\}$ with matrix slices

$$\mathcal{X}^{(n)}(:,:,k) = \mathbf{A}_2(\mathbf{b}_k^{(n)}, \mathbf{c}_k^{(n)}), \quad 1 \le k \le K,$$

in which $\mathbf{A}_2(\mathbf{b}_k^{(n)}, \mathbf{c}_k^{(n)})$ is a rank-1 matrix with

$$\mathbf{b}_k^{(n)} = \begin{bmatrix} x_{1,k} \\ x_{n+3,k} \end{bmatrix} \text{ and } \mathbf{c}_k^{(n)} = \begin{bmatrix} x_{n+1,k} \\ x_{n+2,k} \end{bmatrix}.$$

This leads to a set of three coupled decompositions:

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{A}_2(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)}) \circ \mathbf{s}_r\,, \quad 1 \le n \le 3\,,$$

where $\mathbf{b}_r^{(1)} = [a_{1r}\ a_{4r}]^T$, $\mathbf{c}_r^{(1)} = [a_{2r}\ a_{3r}]^T$, $\mathbf{b}_r^{(2)} = [a_{1r}\ a_{5r}]^T$, $\mathbf{c}_r^{(2)} = [a_{3r}\ a_{4r}]^T$, $\mathbf{b}_r^{(3)} = [a_{1r}\ a_{6r}]^T$ and $\mathbf{c}_r^{(3)} = [a_{4r}\ a_{5r}]^T$.

Similarly, based on the monomial relations (33)–(37) and the mapping (46) with $L = 3$, we can translate narrowband DOA estimation problems involving array response vectors of the form (32) into a coupled tensor decomposition of the form (48) with $N = 5$.

Near-field sources: Exploiting the monomial relations (39)–(42), the near-field array processing problem in Subsection IV-A can be translated into a coupled BTD of the form

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{A}_3(\mathbf{b}_r^{(n)}, \mathbf{c}_r^{(n)}) \circ \mathbf{s}_r\,, \quad 1 \le n \le 4\,,$$

where $\mathbf{b}_r^{(1)} = [a_{-2r}\ a_{1r}\ a_{1r}]^T$, $\mathbf{c}_r^{(1)} = [a_{2r}\ a_{-1r}\ a_{-1r}]^T$, $\mathbf{b}_r^{(2)} = [a_{-3r}\ a_{1r}\ a_{2r}]^T$, $\mathbf{c}_r^{(2)} = [a_{3r}\ a_{-1r}\ a_{-2r}]^T$, $\mathbf{b}_r^{(3)} = [a_{2r}\ a_{2r}\ a_{-4r}]^T$, $\mathbf{c}_r^{(3)} = [a_{-2r}\ a_{-2r}\ a_{4r}]^T$, $\mathbf{b}_r^{(4)} = [a_{1r}\ a_{3r}\ a_{-4r}]^T$ and $\mathbf{c}_r^{(3)} = [a_{-1r}\ a_{-3r}\ a_{4r}]^T$. More details on the connection between coupled tensor decompositions and near-field array processing are given in the supplementary material [25].

C. Connections to ESPRIT and coupled CPD

1) Connection to ESPRIT: At first sight it may not be obvious to see that the presented coupled BTD approach is an extension of the ESPRIT method discussed in Subsection II-B. We now explain the connection between the coupled BTD approach discussed in the previous subsection for the case of monomial structures and the CPD approach used in the ESPRIT method for the special case of exponential structure. The I-element Vandermonde vector (6) can be partitioned into $I-2$ overlapping 3-element vectors:

$$\mathbf{a}^{(1)} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}, \mathbf{a}^{(2)} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \ldots, \mathbf{a}^{(I-2)} = \begin{bmatrix} a_{I-3} \\ a_{I-2} \\ a_{I-1} \end{bmatrix}. \tag{55}$$

Each vector in (55) can be associated with a rank-one (2×2) Hankel matrix:

$$\mathbf{H}^{(1)} = \begin{bmatrix} a_0 & a_1 \\ a_1 & a_2 \end{bmatrix} = \begin{bmatrix} 1 & x \\ x & x^2 \end{bmatrix}, \tag{56}$$

$$\vdots$$

$$\mathbf{H}^{(I-2)} = \begin{bmatrix} a_{I-3} & a_{I-2} \\ a_{I-2} & a_{I-1} \end{bmatrix} = \begin{bmatrix} x^{I-3} & x^{I-2} \\ x^{I-2} & x^{I-1} \end{bmatrix}. \tag{57}$$

Note that (56)–(57) is equivalent to (7). From the relation $\left|\mathbf{H}^{(n)}\right| = x^{n+1}x^{n-1} - x^n x^n = x^{2n} - x^{2n} = 0$, it is clear that $\mathbf{H}^{(n)}$ corresponds to a matrix of the form (46) with $L = 2$, $\mathbf{b} = \left[\begin{smallmatrix} a_{n+1} \\ a_{n-1} \end{smallmatrix}\right] = \left[\begin{smallmatrix} x^{n+1} \\ x^{n-1} \end{smallmatrix}\right]$ and $\mathbf{c} = \left[\begin{smallmatrix} a_n \\ a_n \end{smallmatrix}\right] = \left[\begin{smallmatrix} x^n \\ x^n \end{smallmatrix}\right]$. This shows that ESPRIT can be interpreted as a special case of the proposed coupled BTD approach.

2) Connection to coupled CPD: In this section we explain that the presented coupled BTD approach for monomial structure generalizes the coupled CPD method in [22], [28]. This is significant, as coupled CPD forms the basis of a whole class of array processing techniques [24], [26], [27], [29], as has briefly been illustrated in Section III. Consider the kth matrix slice of the coupled CPD of $\mathcal{X}^{(n)}$ in (9) with factorization

$$\mathbf{X}_k^{(n)} := \mathcal{X}^{(n)}(:,:,k) = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{d}_r^{(n)} s_{kr} = \sum_{r=1}^{R} \mathbf{G}_r^{(n)} s_{kr},$$

in which $\mathbf{G}_r^{(n)} = \mathbf{a}_r^{(n)} \circ \mathbf{d}_r^{(n)} = \mathbf{a}_r^{(n)}\mathbf{d}_r^{(n)T}$ is a rank-1 matrix. The core idea of coupled CPD is to exploit this rank-1 structure. Similar to the monomial interpretation of ESPRIT for exponential structure. Obviously, $\mathbf{G}_r^{(n)}$ is rank-1 if and only if all its (2×2) submatrices are rank-1. There are $C_{I_n}^2$ ways of selecting two rows of $\mathbf{G}_r^{(n)}$ and $C_{J_n}^2$ ways of selecting two columns of $\mathbf{G}_r^{(n)}$. We consider $C_{I_n}^2 C_{J_n}^2$ (overlapping) rank-1 submatrices:

$$\mathbf{G}_r^{(1,n)} = \begin{bmatrix} \mathbf{e}_1^{(I_n)} & \mathbf{e}_2^{(I_n)} \end{bmatrix}^T \mathbf{G}_r^{(n)} \begin{bmatrix} \mathbf{e}_1^{(J_n)} & \mathbf{e}_2^{(J_n)} \end{bmatrix} = \mathbf{a}_r^{(1,n)}\mathbf{d}_r^{(1,n)T},$$
$$\mathbf{G}_r^{(2,n)} = \begin{bmatrix} \mathbf{e}_1^{(I_n)} & \mathbf{e}_2^{(I_n)} \end{bmatrix}^T \mathbf{G}_r^{(n)} \begin{bmatrix} \mathbf{e}_1^{(J_n)} & \mathbf{e}_3^{(J_n)} \end{bmatrix} = \mathbf{a}_r^{(2,n)}\mathbf{d}_r^{(2,n)T},$$

$$\vdots$$

$$\mathbf{G}_r^{(C_{I_n}^2 C_{J_n}^2, n)} = \begin{bmatrix} \mathbf{e}_{I_n-1}^{(I_n)} & \mathbf{e}_{I_n}^{(I_n)} \end{bmatrix}^T \mathbf{G}_r^{(n)} \begin{bmatrix} \mathbf{e}_{J_n-1}^{(J_n)} & \mathbf{e}_{J_n}^{(J_n)} \end{bmatrix} = \mathbf{a}_r^{(C_{I_n}^2 C_{J_n}^2, n)}\mathbf{d}_r^{(C_{I_n}^2 C_{J_n}^2, n)T},$$

where $\mathbf{a}_r^{(\gamma,n)} = \left[\mathbf{e}_{i_1}^{(I_n)}\ \mathbf{e}_{i_2}^{(I_n)}\right]^T \cdot \mathbf{a}_r^{(n)} = \mathbf{A}^{(n)}([i_1,i_2],r)$ and $\mathbf{d}_r^{(\gamma,n)} = \left[\mathbf{e}_{j_1}^{(J_n)}\ \mathbf{e}_{j_2}^{(J_n)}\right]^T \cdot \mathbf{d}_r^{(n)} = \mathbf{D}^{(n)}([j_1,j_2],r)$, in which the pairs (i_1,i_2) and (j_1,j_2) with $i_1 < i_2$ and $j_1 < j_2$ are indexed by $\gamma \in \{1,\ldots,C_{I_n}^2 C_{J_n}^2\}$. In words, if γ involves the pair (i_1,i_2), then $\mathbf{a}_r^{(\gamma,n)}$ consists of the i_1-th and the i_2-th entry of the r-th column of $\mathbf{A}^{(n)}$ (similarly for $\mathbf{d}_r^{(\gamma,n)}$). The associated 2-by-2 submatrix of $\mathbf{X}_k^{(n)}$ can be written as

$$\mathbf{X}_k^{(\gamma,n)} = \sum_{r=1}^{R} \mathbf{G}_r^{(\gamma,n)} s_{kr}. \tag{58}$$

For every $\gamma \in \{1,\ldots,C_{I_n}^2 C_{J_n}^2\}$ and $n \in \{1,\ldots,N\}$, we have that

$$\begin{aligned}\left|\mathbf{G}_r^{(\gamma,n)}\right| &= (\mathbf{G}_r^{(\gamma,n)})_{11}(\mathbf{G}_r^{(\gamma,n)})_{22} - (\mathbf{G}_r^{(\gamma,n)})_{12}(\mathbf{G}_r^{(\gamma,n)})_{21} \\ &= a_{i_1r}^{(n)} a_{i_2r}^{(n)} d_{j_1r}^{(n)} d_{j_2r}^{(n)} - a_{i_1r}^{(n)} a_{i_2r}^{(n)} d_{j_1r}^{(n)} d_{j_2r}^{(n)} = 0\,. \end{aligned} \tag{59}$$

It is now clear that by exploiting this monomial relation, the coupled CPD (9) can be can be turned into a coupled BTD model (48) as follows:

$$\mathcal{X}^{(\gamma,n)} = \sum_{r=1}^{R} \mathbf{G}_r^{(\gamma,n)} \circ \mathbf{s}_r, \quad \gamma \in \{1,\ldots,C_{I_n}^2 C_{J_n}^2\},\ n \in \{1,\ldots,N\},$$

where $\mathcal{X}^{(\gamma,n)}(:,:,k) = \mathbf{X}_k^{(\gamma,n)}$. Here, relation (59) is expressed as the rank-deficiency of a matrix $\mathbf{G}_r^{(\gamma,n)}$ of the form (46) with $L = 2$, $\mathbf{b} = \begin{bmatrix}(\mathbf{G}_r^{(\gamma,n)})_{11}\\(\mathbf{G}_r^{(\gamma,n)})_{22}\end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix}(\mathbf{G}_r^{(\gamma,n)})_{12}\\(\mathbf{G}_r^{(\gamma,n)})_{21}\end{bmatrix}$. Although the reasoning is different, this is precisely what is done in the coupled CPD method in [22], [28].

To summarize, the coupled CPD model (9) extends the ESPRIT and CPD approaches for array processing to the case of multiple shift-invariance structures. The presented coupled BTD model (48) in turn extends the coupled CPD approach to more general cases involving monomial structures.

D. Constant modulus constraint and alternative formulation

Consider again the bilinear factorization (25) in which the entries of either $\mathbf{A}$ or $\mathbf{S}$ are now subject to a Constant Modulus (CM) constraint. The Analytical Constant Modulus Algorithm (ACMA) developed in [33] translates such factorizations into a tensor decomposition. We note in passing that the original ACMA paper [33] focused on the case where the source matrix $\mathbf{S}$ is subject to a CM constraint. This includes the M-PSK source separation problem discussed in Subsection IV-A. However, mixing matrices $\mathbf{A}$ with a CM constraint are also common. This includes the nonuniform and near-field DOA estimation problems discussed in Subsection IV-A, where the CM constraint is due to the fact that the generators of the array response vectors are located on the unit circle.

1) ACMA revisited: Let us focus on the case where the mixing matrix $\mathbf{A} \in \mathbb{C}^{I\times R}$ is CM constrained and has full column rank. This implies that

$$a_{i_1r}a_{i_1r}^* - a_{i_2r}a_{i_2r}^* = 0, \quad 1 \le i_1 < i_2 \le I. \tag{60}$$

Note that, in contrast to the coupled BTD approach discussed in Subsection IV-B, the monomial in (60) involves conjugate variables. This is just a technical variation that does not pose conceptual problems. Recall that ACMA-type procedures attempt to find all linearly independent "beamforming" vectors $\mathbf{w} \in \mathbb{C}^K$ for which [33], [32]:

$$\mathbf{X}\mathbf{w} = \mathbf{a}_r, \quad r \in \{1,\ldots,R\}. \tag{61}$$

Inserting (61) into (60) yields

$$\begin{aligned} &a_{i_1r}a_{i_1r}^* - a_{i_2r}a_{i_2r}^* = 0 \Leftrightarrow \\ &\left(\mathbf{e}_{i_1}^{(I)T}\mathbf{X}\mathbf{w}\right)\left(\mathbf{e}_{i_1}^{(I)T}\mathbf{X}^*\mathbf{w}^*\right) - \left(\mathbf{e}_{i_2}^{(I)T}\mathbf{X}\mathbf{w}\right)\left(\mathbf{e}_{i_2}^{(I)T}\mathbf{X}^*\mathbf{w}^*\right) = 0 \Leftrightarrow \\ &\left\{\mathbf{e}_{i_1}^{(I)T}\mathbf{X}\otimes\mathbf{e}_{i_1}^{(I)T}\mathbf{X}^* - \mathbf{e}_{i_2}^{(I)T}\mathbf{X}\otimes\mathbf{e}_{i_2}^{(I)T}\mathbf{X}^*\right\}(\mathbf{w}\otimes\mathbf{w}^*) = 0 \Leftrightarrow \\ &\mathbf{p}_{\mathrm{CM}}^{(i_1,i_2)}\mathbf{y}_{\mathrm{CM}} = 0, \end{aligned} \tag{62}$$

where

$$\begin{aligned}\mathbf{p}_{\mathrm{CM}}^{(i_1,i_2)} &:= \mathbf{e}_{i_1}^{(I)T}\mathbf{X}\otimes\mathbf{e}_{i_1}^{(I)T}\mathbf{X}^* - \mathbf{e}_{i_2}^{(I)T}\mathbf{X}\otimes\mathbf{e}_{i_2}^{(I)T}\mathbf{X}^*, \\ \mathbf{y}_{\mathrm{CM}} &:= \mathbf{w}\otimes\mathbf{w}^* \in \mathbb{C}^{K^2}.\end{aligned}$$

Stacking yields

$$\mathbf{p}_{\mathrm{CM}}^{(C_I^2)} \cdot \mathbf{y}_{\mathrm{CM}} = \mathbf{0}_{C_I^2}\,, \tag{63}$$

in which

$$\mathbf{P}_{\mathrm{CM}}^{(C_I^2)} = \begin{bmatrix}\mathbf{p}_{\mathrm{CM}}^{(1,2)}\\ \vdots \\ \mathbf{p}_{\mathrm{CM}}^{(I-1,I)}\end{bmatrix} \in \mathbb{C}^{C_I^2\times K^2}. \tag{64}$$

Note that the vectors in the set $\{\mathbf{w}_r\otimes\mathbf{w}_r^*\}$ represent vectorized Hermitian matrices. Since $\mathbf{S}$ has full column rank, we also know that if $\mathbf{w}_r$ is a column of $(\mathbf{S}^T)^\dagger$, then $\mathbf{w}_r\otimes\mathbf{w}_r^* \in \ker(\mathbf{P}_{\mathrm{CM}}^{(C_I^2)})$ and the R vectors in the set $\{\mathbf{w}_r\}$ are linearly independent. Consequently, the R vectors in the set $\{\mathbf{w}_r\otimes\mathbf{w}_r^*\}$ are also linearly independent and the dimension of the subspace $\ker(\mathbf{P}_{\mathrm{CM}}^{(C_I^2)})$ is at least R. Assume that the dimension of $\ker(\mathbf{P}_{\mathrm{CM}}^{(C_I^2)})$ is minimal (i.e. R), then the problem of finding $\{\mathbf{w}_r\}$ from (63) can be reduced to a CPD problem. Namely, if the columns of $\mathbf{N}$ denote a basis for the subspace $\ker(\mathbf{P}_{\mathrm{CM}}^{(C_I^2)})$, then there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that

$$(\mathbf{W}\odot\mathbf{W}^*)\mathbf{F}^T = \mathbf{N}\,. \tag{65}$$

From (5), it is clear that (65) corresponds to a matrix version of a CPD $\mathcal{N} = \sum_{r=1}^{R}\mathbf{w}_r\circ\mathbf{w}_r^*\circ\mathbf{f}_r \in \mathbb{C}^{R\times R\times R}$ that can be computed via a EVD, implying uniqueness of $\mathbf{A} = \mathbf{X}\mathbf{W}$ and $\mathbf{S} = (\mathbf{A}^\dagger\mathbf{X})^T$.[3] This transformation of the CM constrained bilinear factorization (25) into the CPD (65) is the core of ACMA [33]. Here we have given a short derivation in terms of monomials with conjugate variables.

[3]It is explained in [20] that if the dimension of $\ker(\mathbf{P}_{\mathrm{CM}}^{(C_I^2)})$ is minimal (i.e., R), then $\mathbf{A}$ has full column rank.

2) Alternative ACMA-type formulation for monomial structures: In the literature ACMA variants for M-PSK source separation [31], [9] have been presented that exploit the property $s_{k_1 r}^M - s_{k_2 r}^M = 0$. Recall that we have discussed this type of monomial constraint in Subsections IV-A and IV-B4. Let us now work in the opposite direction, and extend ACMA to the general monomial structure in (44). In this way we will obtain an algorithm that is an alternative to the coupled BTD method in Subsection IV-B. Plugging (61) into (44) yields

$$\begin{aligned}
&a_0^{\alpha_0} a_{p_1}^{\beta_1} \cdots a_{p_q}^{\beta_q} - a_0^{\gamma_0} a_{s_1}^{\omega_1} \cdots a_{s_t}^{\omega_t} = 0 \Leftrightarrow \\
&\left(\mathbf{e}_1^{(I)T}\mathbf{X}\mathbf{w}\right)^{\alpha_0}\left(\mathbf{e}_{p_1+1}^{(I)T}\mathbf{X}\mathbf{w}\right)^{\beta_1}\cdots\left(\mathbf{e}_{p_q+1}^{(I)T}\mathbf{X}\mathbf{w}\right)^{\beta_q} \\
&-\left(\mathbf{e}_1^{(I)T}\mathbf{X}\mathbf{w}\right)^{\gamma_0}\left(\mathbf{e}_{s_1+1}^{(I)T}\mathbf{X}\mathbf{w}\right)^{\omega_1}\cdots\left(\mathbf{e}_{s_t+1}^{(I)T}\mathbf{X}\mathbf{w}\right)^{\omega_t} = 0 \Leftrightarrow \\
&\Big(\underbrace{\mathbf{w}\otimes\cdots\otimes\mathbf{w}}_{L \text{ times}}\Big)^T\Big\{\otimes_{\alpha_0}\mathbf{X}^T\mathbf{e}_1^{(I)}\otimes_{\beta_1}\mathbf{X}^T\mathbf{e}_{p_1+1}^{(I)}\otimes\cdots\otimes_{\beta_q}\mathbf{X}^T\mathbf{e}_{p_q+1}^{(I)} \\
&-\otimes_{\gamma_0}\mathbf{X}^T\mathbf{e}_1^{(I)}\otimes_{\omega_1}\mathbf{X}^T\mathbf{e}_{s_1+1}^{(I)}\otimes\cdots\otimes_{\omega_t}\mathbf{X}^T\mathbf{e}_{s_t+1}^{(I)}\Big\} = 0, \qquad (66)
\end{aligned}$$

where $\mathbf{e}_i^{(I)} \in \mathbb{C}^I$ denotes the unit vector with unit entry at the ith position and zero elsewhere, $\otimes_{\gamma_0}\mathbf{X}^T\mathbf{e}_1^{(I)} = \underbrace{\mathbf{X}^T\mathbf{e}_1^{(I)}\otimes\cdots\otimes\mathbf{X}^T\mathbf{e}_1^{(I)}}_{\gamma_0 \text{ times}}$ and similarly for the other Kronecker products in (66). Define

$$\begin{aligned}
\mathbf{p}_L &:= \otimes_{\alpha_0}\mathbf{X}^T\mathbf{e}_1^{(I)}\otimes_{\beta_1}\mathbf{X}^T\mathbf{e}_{p_1+1}^{(I)}\otimes\cdots\otimes_{\beta_q}\mathbf{X}^T\mathbf{e}_{p_q+1}^{(I)} \\
&\quad -\otimes_{\gamma_0}\mathbf{X}^T\mathbf{e}_1^{(I)}\otimes_{\omega_1}\mathbf{X}^T\mathbf{e}_{s_1+1}^{(I)}\otimes\cdots\otimes_{\omega_t}\mathbf{X}^T\mathbf{e}_{s_t+1}^{(I)} \in \mathbb{C}^{K^L}, \qquad (67) \\
\mathbf{y} &:= \mathbf{w}\otimes\cdots\otimes\mathbf{w} \in \mathbb{C}^{K^L}. \qquad (68)
\end{aligned}$$

Then (66) can be written as

$$\mathbf{p}_L^T \cdot \mathbf{y} = 0\,. \qquad (69)$$

Let $\mathbf{p}_L^{(n)}$ denote the vector in (69) associated with the nth monomial relation. Stacking yields:

$$\mathbf{P}_L^{(N)} = \left[\mathbf{p}_L^{(1)}, \ldots, \mathbf{p}_L^{(N)}\right]^T \in \mathbb{C}^{N\times K^L} \qquad (70)$$

and

$$\mathbf{P}_L^{(N)} \cdot \mathbf{y} = \mathbf{0}_N\,. \qquad (71)$$

As in Subsection IV-D1, we know that there exist at least R linearly independent vectors $\{\mathbf{w}_r\otimes\cdots\otimes\mathbf{w}_r\}$, each with property $\mathbf{w}_r\otimes\cdots\otimes\mathbf{w}_r \in \ker\left(\mathbf{P}_L^{(N)}\right)\cap\pi_s^{(L)}$. Consequently, if the dimension of the subspace $\ker\left(\mathbf{P}_L^{(N)}\right)\cap\pi_s^{(L)}$ is minimal (i.e., R), the problem of finding $\{\mathbf{w}_r\}$ from (71) can be reduced to a CPD problem. Namely, if the columns of $\mathbf{M}$ denote a basis for the subspace $\ker\left(\mathbf{P}_L^{(N)}\right)\cap\pi_s^{(L)}$, then there exists a nonsingular matrix $\mathbf{F}\in\mathbb{C}^{R\times R}$ such that

$$\underbrace{(\mathbf{W}\odot\cdots\odot\mathbf{W})}_{L}\mathbf{F}^T = \mathbf{M}\,. \qquad (72)$$

Like (65), it is clear that (72) corresponds to a matrix representation of a CPD $\mathcal{M} = \sum_{r=1}^R \mathbf{w}_r\circ\cdots\circ\mathbf{w}_r\circ\mathbf{f}_r \in \mathbb{C}^{R\times\cdots\times R}$ that can be computed via a EVD, which again implies the uniqueness of $\mathbf{A} = \mathbf{X}\mathbf{W}$ and $\mathbf{S} = (\mathbf{A}^\dagger\mathbf{X})^T$.[4]

V. Numerical experiments

Let us consider the factorization $\mathbf{X} = \mathbf{A}\mathbf{S}^T$ in (3) where the columns $\mathbf{A}$ are of the form (28) or (32). The goal is to estimate the angles $\{\theta_r\}$ from $\mathbf{T} = \mathbf{X} + \beta\mathbf{N}$, where $\mathbf{N}$ is an unstructured perturbation matrix and $\beta\in\mathbb{R}$ controls the signal-to-noise level. We will consider Monte Carlo simulations, where for each trial the real and imaginary entries of $\mathbf{S}$ and $\mathbf{N}$ are randomly drawn from a Gaussian distribution with zero mean and unit variance. The following Signal-to-Noise Ratio (SNR) measure will be used: SNR [dB] $= 10\log(\|\mathbf{X}\|_F^2/\beta^2\|\mathbf{N}\|_F^2)$.

A. Case 1

Consider the scenario in which $K = 50$, $R = 2$ and the columns of $\mathbf{A}$ are of the form (32). We will retrieve $\boldsymbol{\theta} = [\theta_1, \theta_2]$ from $\mathbf{X}$ using the coupled BTD and CPD based methods discussed in Subsections IV-B2 and IV-D2, respectively. The monomial relations (33)–(37) will be exploited and the associated CPD problems (52) and (72) will numerically be solved by QZ iteration [33]. Finally, the decoupled single-source DOA estimation problems (i.e., recover θ_r from the estimated array response vector $\mathbf{a}_r$) will be solved by spectral peak picking. The overall methods will be referred to as 'coupled BTD-QZ' and 'CPD-QZ', respectively. We compare them with the dedicated spectral MUSIC method, referred to simply as 'MUSIC'. We compute the root mean square error $\mathrm{E_{RMS}} = \sqrt{\frac{1}{RT_{\mathrm{MC}}}\sum_{t=1}^{T_{\mathrm{MC}}}\|\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_t\|_F^2}$, where T_{MC} and $\hat{\boldsymbol{\theta}}_t$ denote the number of Monte Carlo trials and the associated trial estimate of $\boldsymbol{\theta}$, respectively. The root mean square of the Cramér-Rao Bound (CRB) will also be computed. The $\mathrm{E_{RMS}}$ and CRB values over $T_{\mathrm{MC}} = 50$ trials as a function of SNR when $\boldsymbol{\theta} = \frac{\pi}{180}[10,\ 40]^T$ can be seen in figure 3(a). We observe that both the algebraic coupled BTD-QZ and CPD-QZ performed about the same as the dedicated spectral search method MUSIC.

Let us now consider a more difficult case in which $\boldsymbol{\theta} = \frac{\pi}{180}[40,\ 45]^T$. The $\mathrm{E_{RMS}}$ and CRB values over $T_{\mathrm{MC}} = 50$ trials as a function of SNR can be seen in figure 3(b). Above 15 dB SNR, the algebraic coupled BTD-QZ and CPD-QZ methods performed about the same and almost as good as the dedicated MUSIC method. On the other hand, the former methods do no require the solution to a more complicated multi-peak spectral search problem.

B. Case 2

Consider the same scenario as before, but now $R = 3$ and the columns of $\mathbf{A}$ are of the form (28) with $\boldsymbol{\theta} = \frac{\pi}{180}[10,\ 40,\ 45]^T$. The monomial relations (29)–(31) will be exploited in the coupled BTD-QZ and CPD-QZ methods.

[4] It is explained in [20] that if the dimension of $\ker\left(\mathbf{P}_L^{(N)}\right)\cap\pi_s^{(L)}$ is minimal (i.e. R), then $\mathbf{A}$ has full column rank.

The E_{RMS} and CRB values over $T_{MC} = 50$ trials as a function of SNR can be seen in figure 4. We make the same observations as the previous case.

VI. Conclusion

Despite its importance, sensor array processing beyond the CPD model still presents important algebraic challenges. We first explained that the coupled CPD model provides multidimensional extensions of the CPD/ESPRIT framework to arrays with multiple shift-invariance structures.

Second, we explained that monomial structures play an important role in array processing. To accommodate the use of monomial structures in array processing, we extended the coupled CPD model to bilinear factorizations involving a monomial structure. This resulted in a new uniqueness condition, a new coupled BTD based algorithm and a new CPD based algorithm.

Third, we explained that the link between monomials and tensor decompositions led to a very general framework that encompasses many classical tensor-based methods. In particular, new monomial interpretations of the ESPRIT and ACMA methods for Vandermonde and CM structured bilinear factorizations were provided.

For ease of exposition we limited the presentation to monomials of the same degree. However, the results can be extended monomials of varying degree or to the sum of monomials, as demonstrated in [20].

References

[1] J. D. Carroll and J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[2] L. De Lathauwer, "Decomposition of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.

[3] ——, "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.

[4] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, 2013.

[5] ——, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Overall uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, 2013.

[6] ——, "Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 2, pp. 636–660, 2014.

[7] ——, "Canonical polyadic decomposition of third-order tensors: Relaxed uniqueness conditions and algebraic algorithm," *Linear Algebra Appl.*, vol. 513, pp. 342–375, Jan. 2017.

[8] R. Eckhoff, "Direction-of-arrival determination using 3-axis crossed array and ESPRIT," in *Proc. of the 9th IEEE International Symposium PIMRC 1998*, september 8-11, Boston, MA, USA, 1998.

[9] O. Grellier and P. Comon, "Analytical blind discrete source separation," in *Proc. EUSIPCO*, May 5-8 , Tampere, Finland, 2000.

[10] E. Grosicki, K. Abed-Meraim, and Y. Hua, "A weighted linear prediction method for near-field source localization," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3651–3660, Oct. 2005.

[11] 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.

[12] S. Holm, A. Austeng, K. Iranpour, and J. F. Hopperstad, "Sparse sampling in array processing," in *Nonuniform sampling: Theory and Practice*, F. Marvasti, Ed. Springer Science and Business Media, 2001, ch. 19, pp. 787–833.

[13] 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.

[14] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics," *Linear Algebra and its Applications*, vol. 18, pp. 95–138, 1977.

[15] S. E. Leurgans, R. T. Ross, and R. B. Abel, "A decomposition of three-way arrays," *SIAM J. Matrix Anal. Appl.*, vol. 14, pp. 1064–1083, 1993.

[16] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 1, pp. 8–15, Jan. 1989.

[17] R. Roy and T. Kailath, "Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. ASSP*, vol. 32, no. 7, pp. 984–995, Jul. 1989.

[18] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Processing*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

[19] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Processing, To appear*.

[20] M. Sørensen and L. De Lathauwer, "Coupled tensor decompositions and monomial structure in array processing — Supplementary Material," Tech. Rep.

[21] ——, "Coupled tensor decompositions for applications in array signal processing," in *Proc. CAMSAP 2013*, Dec. 15-18, Saint Martin, Dutch Antilles, 2013.

[22] ——, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part I: Uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 2, pp. 496–522, 2015.

[23] ——, "New uniqueness conditions for the canonical polyadic decomposition of third-order tensors," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 4, pp. 1381–1403, 2015.

[24] ——, "Multiple invariance ESPRIT for nonuniform linear arrays: A coupled canonical polyadic decomposition approach," *IEEE Trans. Signal Processing*, vol. 64, no. 4, pp. 3693–3704, Jul. 2016.

[25] ——, "Coupled tensor decompositions and monomial structure in array processing," ESAT-STADIUS, KU Leuven, Belgium, Tech. Rep., 2017.

[26] ——, "Multidimensional harmonic retrieval via coupled canonical polyadic decomposition — Part I: Model and identifiability," *IEEE Trans. Signal Processing*, vol. 65, no. 2, pp. 517–527, Jan. 2017.

[27] ——, "Multidimensional harmonic retrieval via coupled canonical polyadic decomposition — Part II: Algorithm and multirate sampling," *IEEE Trans. Signal Processing*, vol. 65, no. 2, pp. 528–539, Jan. 2017.

[28] M. Sørensen, I. Domanov, and L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part II: Algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 3, pp. 1015–1045, 2015.

[29] M. Sørensen, F. Van Eeghem, and L. De Lathauwer, "Blind multichannel deconvolution and convolutive extensions of canonical polyadic and block term decompositions," ESAT-STADIUS, KU Leuven, Belgium, Tech. Rep. 16-73, 2016.

[30] A. T. Swindlehurst and T. Kailath, "Passive direction-of-arrival and range estimation for near-field sources," in *Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, Minneapolis, Minnesota, USA, Aug. 3-5 1988.

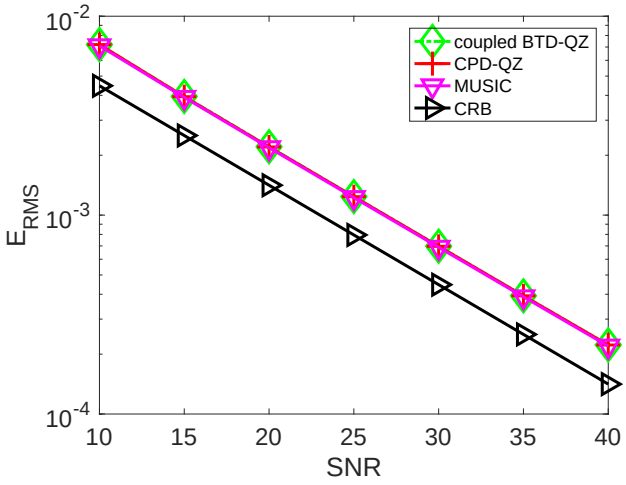
[31] A.-J. Van Der Veen, "Analytical method for blind binary signal separation," *IEEE Trans. Signal Process.*, vol. 45, no. 4, pp. 1078–1082, Apr. 1997.

[32] ——, "Analytical constant modulus algorithms," in *Signal Processing Advances in Wireless and Mobile Communications*, G. Giannakis et al., Ed. Prentice Hall, 2000.

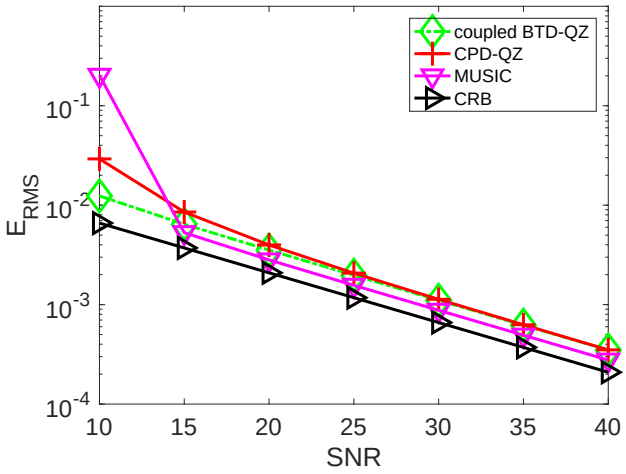
[33] A.-J. Van Der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[34] H. L. Van Trees, *Detection, estimation, and modulation theory, optimum array processing (Part IV)*. Wiley-Interscience, 2002.

[35] E. Vertatschitsch and S. Haykin, "Nonredundant arrays," *Proc. of the IEEE*, vol. 74, no. 1, p. 217, Jan. 1986.
[36] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab v3.0*, Available online, March 2016,. [Online]. Available: http://www.tensorlab.net/.



(a) $\theta_1 = \pi\frac{10}{180}$ and $\theta_2 = \pi\frac{40}{180}$.



(b) $\theta_1 = \pi\frac{40}{180}$ and $\theta_2 = \pi\frac{45}{180}$.

Fig. 3. E_{RMS} over 50 trials, case $\mathbf{a}_r = [1\ x_r\ x_r^4\ x_r^9\ x_r^{11}]^T$.

Acknowledgment

Research supported by: (1) Research Council KU Leuven: C1 project C16/15/059-nD, (2) F.W.O.: project G.0830.14N, G.0881.14N, (3) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

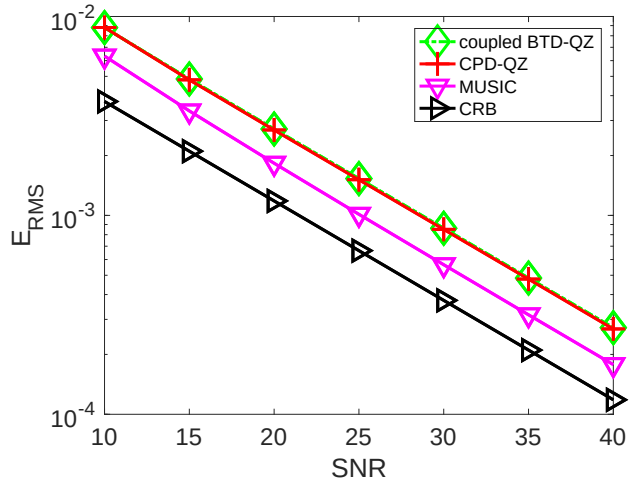


Fig. 4. E_{RMS} over 50 trials, case $\mathbf{a}_r = [1\ x_r^3\ x_r^5\ x_r^8\ x_r^{13}\ x_r^{21}]^T$.



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. From 2010 to 2017 he was a Postdoctoral Fellow with the KU Leuven, Belgium. Past research interests include applied linear algebra, tensor decompositions, and tensor-based array and signal processing. His current research activities are focused on tensor methods for machine learning and big data analytics.



Lieven De Lathauwer 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 Professor with KU Leuven, Belgium. Dr. De Lathauwer is an Associate Editor of the SIAM Journal on Matrix Analysis and Applications and has served as an Associate Editor for the IEEE Transactions on Signal Processing. His research concerns the development of tensor tools for engineering applications.