Coupled Tensor Decompositions 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 first contribution, we present a paradigm shift from the conventional matrix/tensor decomposition approach to coupled matrix/tensor based array processing. More precisely, we propose the coupled CPD model as a new algebraic framework for array processing. In particular, we demonstrate that the coupled CPD can handle array processing problems involving multiple shift-invariance structures. This leads to relaxed uniqueness conditions and algebraic multidimensional ESPRIT methods for many antenna array configurations found in the literature. Second, we also present a link between wideband vector-sensor array processing and the coupled CPD. This leads to a new coherent data processing framework for VSA that takes the wideband structure of the signals into account.**

***Index Terms*—array processing, direction-of-arrival estimation, tensor, coupled canonical polyadic decomposition, coupled matrix-pencil, multidimensional array.**

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 geometries and tensors was recognized in [24]. More precisely, a link between the Canonical Polyadic Decomposition (CPD) and arrays composed of several displaced but identical subarrays [23], [33], [32], [34], [15] was presented in [24].

In practice, the array geometries may be irregular, the subarrays may be sparse, different or some of the sensors (or equivalently some sensor outputs) may be missing. 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.

To overcome some of the limitations of traditional array signal processing, we suggest a shift from the conventional matrix/tensor decomposition approach to coupled matrix/tensor based array processing. We argue that the coupled CPD framework presented in [29], [30] is a good framework for array processing. Part of this work was presented in the conference paper [26].

We first demonstrate that the coupled CPD provides a unified algebraic framework for one-, two-, three- or N-dimensional arrays with shift-invariance structures. In particular, the coupled CPD framework yields new dedicated uniqueness conditions and algorithms for many array geometries proposed in the literature, such as rectangular shaped, hexagonal shaped, cross-shaped, frame-shaped, L-shaped, V-shaped, Y-shaped, triangular shaped, octagon shaped and multi-scale invariance arrays (e.g. [12], [11], [38], [35], [8], [14], [10], [9], [18]).

Second, we briefly explain that the coupled CPD appears in Vector-Sensor Array (VSA) processing [20], [19]. In particular, the coupled CPD provides a coherent data processing framework for VSA in the case of wideband signals, which is relevant in the context of acoustic (e.g. [2], [17], [37]) or seismic signal processing (e.g. [7], [21]).

Even though we only focus on applications in array processing, multiple shift-invariance structured problems appear in a wide range of areas such as multi-sensor data fusion and multi-modal biomedical engineering. See [29], [30], [25] for references to concrete applications. As such, the concepts presented in this paper are relevant beyond array processing.

The paper is organized as follows. The rest of the introduction will present the notation. Section II reviews the necessary algebraic prerequisites. Section III briefly discusses CPD based array processing. Section IV introduces the coupled CPD framework for array processing problems involving array geometries with multiple shift-invariance structures. Section V briefly explains that the coupled CPD also provides a framework for VSA processing, even in the case of random array geometries. Many coupled CPD problems in array processing can be interpreted as coupled Matrix-Pencil (MP) problems; these involve a coupled CPD of two-slice tensors. Section VI develops a Simultaneous matrix Diagonalization (SD) method suited for this problem. Numerical experiments are reported in section VII. Section VIII 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, range, kernel, Frobenius norm, real and imaginary part of a matrix $\mathbf{A}$ are denoted by $\mathbf{a}_r$, $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\text{range}(\mathbf{A})$, $\ker(\mathbf{A})$, $\|\mathbf{A}\|_F$, $\text{Re}\{\mathbf{A}\}$ and $\text{Im}\{\mathbf{A}\}$, respectively.

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-kulak.be.

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 exchange matrix with unit entries on the anti-diagonal and zeros elsewhere is denoted by $\mathbf{J}_I \in \mathbb{C}^{I\times I}$. $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$ 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. Let $\mathbf{X} \in \mathbb{C}^{IJ\times K}$. Then we will also make use of the linear operator $\mathbf{S}_{I,J}(\mathbf{X}) = \begin{bmatrix} (\underline{\mathbf{I}}_J\otimes\mathbf{I}_I)\mathbf{X} \\ (\overline{\mathbf{I}}_I\otimes\mathbf{I}_I)\mathbf{X} \end{bmatrix} \in \mathbb{C}^{2(I-1)J\times K}$, where $\mathbf{I}_m \in \mathbb{C}^{m\times m}$ denotes the identity matrix.

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. Algebraic Prerequisites

A. Vandermonde matrices

The matrix $\mathbf{A} \in \mathbb{C}^{I\times R}$ is called Vandermonde if

$$\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_R], \quad \mathbf{a}_r = \left[1, z_r, z_r^2, \dots, z_r^{I-1}\right]^T, \tag{1}$$

where the scalars $\{z_r\}$ are called the generators of $\mathbf{A}$.

If the generators of the Vandermonde matrix $\mathbf{A}$ are located on the unit circle ($|z_r| = 1$), then we can make use of the well-known Forward-Backward Averaging (FBA) procedure (e.g. [22]). More precisely, if $z_r = e^{i\alpha_r}$ where $\alpha_r \in \mathbb{R}$, $\forall r \in \{1,\dots,R\}$, then $\mathbf{J}_I\mathbf{A}^* = \mathbf{A}\mathbf{D}_\mathbf{A}$ in which $\mathbf{D}_\mathbf{A} = D_1\left(\left[z_1^{-(I-1)}, z_2^{-(I-1)}, \dots, z_R^{-(I-1)}\right]\right)$. FBA basically doubles the amount of observed data. As an example, consider the factorization $\mathbf{X} = \mathbf{A}\mathbf{S}^T$ in which $\mathbf{A}$ is Vandermonde with unit-modulus generators. FBA now provides the augmented factorization $[\mathbf{X}, \mathbf{J}_I\mathbf{X}^*] = \mathbf{A}[\mathbf{S}^T, \mathbf{D}_\mathbf{A}\mathbf{S}^H]$.

B. Canonical Polyadic Decomposition (CPD)

Consider the third-order tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$. We say that $\mathcal{X}$ is a rank-1 tensor if it is equal to the outer product of some non-zero vectors $\mathbf{a} \in \mathbb{C}^I$, $\mathbf{b} \in \mathbb{C}^J$ and $\mathbf{c} \in \mathbb{C}^K$ such that $x_{ijk} = a_i b_j c_k$. A Polyadic Decomposition (PD) is a decomposition of $\mathcal{X}$ into a sum of rank-1 terms:

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

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 (2) is called the CPD of $\mathcal{X}$. Let us stack the vectors $\{\mathbf{a}_r\}$, $\{\mathbf{b}_r\}$ and $\{\mathbf{c}_r\}$ into the matrices $\mathbf{A} = [\mathbf{a}_1,\dots,\mathbf{a}_R]$, $\mathbf{B} = [\mathbf{b}_1,\dots,\mathbf{b}_R]$ and $\mathbf{C} = [\mathbf{c}_1,\dots,\mathbf{c}_R]$. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ will be referred to as the factor matrices of the PD or CPD of $\mathcal{X}$.

1) Matrix Representation: Let $\mathbf{X}^{(i\cdot\cdot)} \in \mathbb{C}^{J\times K}$ denote the matrix such that $\left(\mathbf{X}^{(i\cdot\cdot)}\right)_{jk} = x_{ijk}$, then $\mathbf{X}^{(i\cdot\cdot)} = \mathbf{B}D_i(\mathbf{A})\mathbf{C}^T$ and we obtain the following matrix representation of (2):

$$\mathbb{C}^{IJ\times K} \ni \mathbf{X}_{(1)} := \left[\mathbf{X}^{(1\cdot\cdot)T}, \dots, \mathbf{X}^{(I\cdot\cdot)T}\right]^T = (\mathbf{A}\odot\mathbf{B})\mathbf{C}^T. \tag{3}$$

2) Uniqueness: The uniqueness properties of the CPD have been intensely studied in recent years, see [4], [5], [27] and references therein. In this paper we consider the case where at least one factor has full column rank. In wireless communication and radar processing this typically corresponds to the factor matrix in which transmitted symbols or data snapshots are stacked, respectively. For this case the following sufficient uniqueness condition has been obtained.

Theorem II.1. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in* (2)*. If*

$$\begin{cases} \boldsymbol{C} \text{ has full column rank,} \\ C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B}) \text{ has full column rank,} \end{cases} \tag{4}$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique [13], [3], [4]. Generically[1]*, condition* (4) *is satisfied if $2R(R-1) \le I(I-1)J(J-1)$ and $R \le K$ [3].*

The uniqueness condition (4) is relatively easy to check and yet very powerful. Moreover, under condition (4) the CPD can be computed by algebraic means [3]. More precisely, the CPD problem was reduced to a Simultaneous matrix Diagonalization (SD) problem. In the case of an exact CPD, the SD problem can be solved via a standard Generalized EigenValue Decomposition (GEVD). The approach has been further generalized to cases where $\mathbf{C}$ does not have full column rank [6]. For these reasons all CPD based uniqueness conditions stated in this paper will be derived from condition (4).

C. Coupled CPD

We say that a collection of tensors $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n\times J_n\times K}$, $n \in \{1,\dots,N\}$, admits an R-term coupled PD if each tensor $\mathcal{X}^{(n)}$ can be written as [28]:

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r\,, \quad n \in \{1,\dots,N\}, \tag{5}$$

with factor matrices $\mathbf{A}^{(n)} = \left[\mathbf{a}_1^{(n)},\dots,\mathbf{a}_R^{(n)}\right]$, $\mathbf{B}^{(n)} = \left[\mathbf{b}_1^{(n)},\dots,\mathbf{b}_R^{(n)}\right]$ and $\mathbf{C} = [\mathbf{c}_1,\dots,\mathbf{c}_R]$. We define the coupled rank of $\{\mathcal{X}^{(n)}\}$ as the minimal number of coupled rank-1 tensors $\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r$ that yield $\{\mathcal{X}^{(n)}\}$ in a linear

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

combination. Assume that the coupled rank of $\{\mathcal{X}^{(n)}\}$ is R, then (5) will be called the coupled CPD of $\{\mathcal{X}^{(n)}\}$.

The coupled PD or CPD of $\{\mathcal{X}^{(n)}\}$ has the following matrix representation

$$\mathbf{X} = \left[\mathbf{X}_{(1)}^{(1)T}, \ldots, \mathbf{X}_{(1)}^{(N)T}\right]^T = \mathbf{F}\mathbf{C}^T \in \mathbb{C}^{(\sum_{n=1}^{N} I_n J_n)\times K}, \tag{6}$$

where $\mathbf{F} = [(\mathbf{A}^{(1)} \odot \mathbf{B}^{(1)})^T, \ldots, (\mathbf{A}^{(N)} \odot \mathbf{B}^{(N)})^T]^T \in \mathbb{C}^{(\sum_{n=1}^{N} I_n J_n)\times R}$.

The coupled rank-1 tensors in (5) 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. Uniqueness conditions for the coupled CPD have been developed in [29]. Theorem II.2 extends Theorem II.1 to coupled CPD.

Theorem II.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 (5). Define*

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

If

$$\begin{cases} \boldsymbol{C} \text{ has full column rank}, \\ \boldsymbol{G} \text{ has full column rank}, \end{cases} \tag{7}$$

then the coupled rank of $\{\mathcal{X}^{(n)}\}$ is R and the coupled CPD of $\{\mathcal{X}^{(n)}\}$ is unique [29].

As in ordinary CPD, the uniqueness condition (7) is relatively easy to check and yet very powerful. It is also clear that more relaxed uniqueness conditions are obtained by taking the coupling into account. (Compare for instance condition (4) with condition (7)).

Furthermore, an extension of the SD method in [3] to the coupled CPD has been developed in [30]. We emphasize that in the noise-free case the SD problem for coupled CPD can be solved exactly via a GEVD.

III. 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

$$x_{ik} = \sum_{r=1}^{R} s_r\left(k - \tau_{ri}\right), \tag{8}$$

where τ_{ri} denotes the delay between the ith sensor and the rth source. The azimuth angle associated with the rth source is denoted by θ_r and is measured counter-clockwise (see figure 1). Similarly, the elevation angle associated with the rth source is denoted by ϕ_r and is measured clockwise. 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 (e.g. [31]). Under these assumptions the array response vector $\mathbf{a}_r \in \mathbb{C}^I$ associated with the rth source can (approximately) be expressed as

$$\mathbf{a}_r = \left[e^{-i\omega_c \tau_{r1}} \ \ldots \ e^{-i\omega_c \tau_{rI}}\right]^T = \left[e^{-i\omega_c \mathbf{b}_r^T \mathbf{p}_1/c} \ \ldots \ e^{-i\omega_c \mathbf{b}_r^T \mathbf{p}_I/c}\right]^T, \tag{9}$$

where ω_c is the carrier frequency, c is the speed of propagation, $\mathbf{b}_r = [\sin(\phi_r)\cos(\theta_r), \ \sin(\phi_r)\sin(\theta_r), \ \cos(\phi_r)]^T$ is the bearing vector depicted in figure 1, and the product $\tau_{ri} = \mathbf{b}_r^T \mathbf{p}_i / c$ corresponds to the propagation delay associated with the ith sensor and the rth source. 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

$$\mathbf{X} = \mathbf{A}\mathbf{S}^T \in \mathbb{C}^{I\times K}, \quad \mathbf{A} = [\mathbf{a}_1 \ \ldots \ \mathbf{a}_R], \ \mathbf{S} = [\mathbf{s}_1 \ \ldots \ \mathbf{s}_R]. \tag{10}$$

In direction-of-arrival 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. In this paper we focus on the former approach.

Because of their simplicity, arrays composed of several displaced but identical subarrays are often considered in the literature (e.g. [23], [33], [32], [24], [34], [15]). 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 data tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ admits the factorization

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{d}_r \circ \mathbf{s}_r, \tag{11}$$

where $\mathbf{a}_r = \left[e^{-i\omega_c \mathbf{b}_r^T \mathbf{p}_1/c} \ \ldots \ e^{-i\omega_c \mathbf{b}_r^T \mathbf{p}_I/c}\right]^T \in \mathbb{C}^I$, $\mathbf{d}_r = \left[e^{-i\omega_c \mathbf{b}_r^T \mathbf{t}_1/c} \ \ldots \ e^{-i\omega_c \mathbf{b}_r^T \mathbf{t}_J/c}\right]^T \in \mathbb{C}^J$, 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 $\mathbf{d}_r$:

$$\mathbf{a}_r \mathbf{d}_r^T = [\mathbf{a}_r(\mathbf{d}_r)_1 \ \ldots \ \mathbf{a}_r(\mathbf{d}_r)_J]. \tag{12}$$

The rank-1 structure in (12), reflecting the overall array geometry, explains why CPD is the proper model. We say that an array with a rank-1 structure of the form (12) enjoys a shift-invariance property if $\mathbf{a}_r$ and/or $\mathbf{d}_r$ are Vandermonde vectors.

A notable limitation of the CPD is that it can only model array configurations in which all subarrays are identical. An array may more generally be composed of subarrays that have different configurations, leading to multiple shift-invariance substructures. Such arrays will be discussed in sections IV and V.

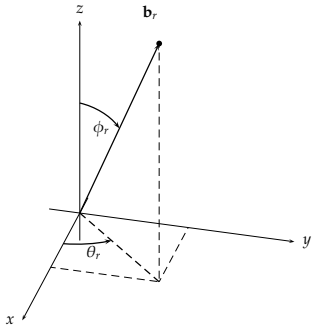


Fig. 1. Bearing vector $\mathbf{b}_r$ associated with the rth source in which θ_r and ϕ_r denote the azimuth and elevation angle, respectively.

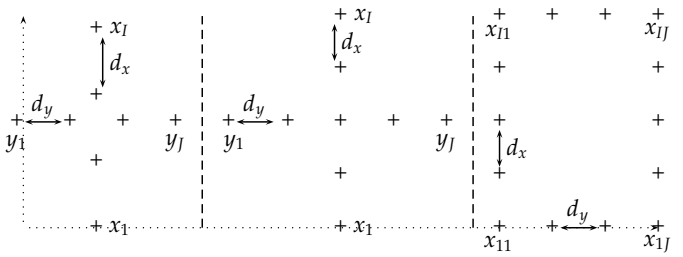


Fig. 2. Cross-shaped arrays (left and middle) and frame-shaped array (right) where '+' represents an antenna element.

IV. New results for arrays with multiple shift-invariance structures based on the coupled CPD

The main limitation of the standard tensor-based methods for array processing is that the observed data must admit a (constrained) CPD interpretation of the form (11). To alleviate this problem, we suggest to model data received by an array that has N shift-invariance structures by a coupled CPD of the form

$$\mathbb{C}^{I_n \times J_n \times K} \ni \mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{d}_r^{(n)} \circ \mathbf{s}_r, \quad n \in \{1, \ldots, N\}, \quad (13)$$

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.

For convenience of presentation, we mainly focus on very regular array geometries with multiple shift-invariance structures. We only consider 2D planar arrays ($z_r = 0$ in (9)). The extension to three-dimensional (3D) arrays ($z_r \neq 0$ in (9)) is completely analogous, i.e., the coupled CPD framework can handle any N-dimensional array as long as it contains shift-invariance structures. We stress that our approach can handle quite irregular array geometries provided they contain shift-invariance structures.

A. Centro-symmetric regular arrays

Let us first explain how the coupled CPD model (13) can be used in the case of centro-symmetric arrays (i.e., arrays that are symmetric around their "epicenter") with multiple shift-invariance structures.

1) Cross-shaped Array (CA): Consider the CA in figure 2 (middle), composed of a horizontal and vertical array (e.g. [11]). It is assumed that I and J are odd numbers such that the horizontal and vertical arrays share one sensor. Since the horizontal and vertical arrays are not shifted versions of each other, it is crucial in our framework that the individual arrays have shift-invariance properties. Here we consider the simple case where both the horizontal and vertical arrays are ULAs. Let us choose the sensor located at the "epicenter" of the CA as the reference sensor. The observed data can be stored into the matrices $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, where $x_{ik}^{(1)}$ and $x_{ik}^{(2)}$ denote the outputs of the vertical and horizontal arrays at snapshot k, respectively. The narrowband and far-field assumptions imply that

$$\mathbf{X}^{(1)} = \mathbf{B}^{(1)}\mathbf{S}^T, \quad \mathbf{X}^{(2)} = \mathbf{B}^{(2)}\mathbf{S}^T, \quad (14)$$

where $\mathbf{b}_r^{(1)} = [x_r^{(1-I)/2}, \ldots, x_r^{(I-1)/2}]$ in which $x_r = e^{-i\omega_c d_x \sin(\phi_r)\cos(\theta_r)/c}$, $\mathbf{b}_r^{(2)} = [y_r^{(1-J)/2}, \ldots, y_r^{(J-1)/2}]$ in which $y_r = e^{-i\omega_c d_y \sin(\phi_r)\sin(\theta_r)/c}$, and $\mathbf{S} \in \mathbb{C}^{K \times R}$ is the collected signal matrix. FBA yields

$$\begin{bmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} & \mathbf{J}_I\mathbf{X}^{(1)*} \\ \mathbf{X}^{(2)} & \mathbf{J}_J\mathbf{X}^{(2)*} \end{bmatrix} = \begin{bmatrix} \mathbf{B}^{(1)} \\ \mathbf{B}^{(2)} \end{bmatrix} \left[\mathbf{S}^T, \mathbf{S}^H\right], \quad (15)$$

where we exploited the centro-symmetry in $\mathbf{J}_I\mathbf{B}^{(1)*} = \mathbf{B}^{(1)}$ and $\mathbf{J}_J\mathbf{B}^{(2)*} = \mathbf{B}^{(2)}$. From (15) we obtain tensors $\{\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)}\}$ of which the matrix representation admits the coupled decomposition

$$\begin{bmatrix} \mathbf{Z}^{(1)} \\ \mathbf{Z}^{(2)} \end{bmatrix} = \begin{bmatrix} S_{I,1}(\mathbf{Y}^{(1)}) \\ S_{J,1}(\mathbf{Y}^{(2)}) \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{(1)} \odot \mathbf{A}^{(1)} \\ \mathbf{D}^{(2)} \odot \mathbf{A}^{(2)} \end{bmatrix} \left[\mathbf{S}^T, \mathbf{S}^H\right], \quad (16)$$

where

$$\mathbf{a}_r^{(1)} = \left[x_r^{(2-I)/2}, \ldots, x_r^{(I-2)/2}\right]^T, \quad \mathbf{d}_r^{(1)} = \left[x_r^{-\frac{1}{2}}, x_r^{\frac{1}{2}}\right]^T,$$
$$\mathbf{a}_r^{(2)} = \left[y_r^{(2-J)/2}, \ldots, y_r^{(J-2)/2}\right]^T, \quad \mathbf{d}_r^{(2)} = \left[y_r^{-\frac{1}{2}}, y_r^{\frac{1}{2}}\right]^T.$$

From (16) it is clear that CA problems involving shift-invariant subarrays can be addressed as coupled CPD problems. This leads to improved uniqueness conditions. As an example, consider the case where $I = J = 5$ and $K \geq R$. If we only exploit the CPD structure of, say $\mathcal{Z}^{(1)}$, then from Theorem II.1 we obtain the bound $R \leq 4$. Thanks to Theorem II.2 the coupled CPD structure of $\{\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)}\}$ relaxes the bound to $R \leq 5$. The coupled CPD of $\{\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)}\}$ can also be computed by means of the SD algorithm developed in Section VI.

Let us now briefly consider the case where I and J are even numbers. Let us choose the left-middle sensor in the horizontal ULA as the reference sensor. By exploiting the phase indeterminacy of the problem (due to the delay between the source and the reference sensor) we can choose the centre of the "x-y plane" as the phase reference center. Consider the following phase-centered variant of (14):

$$\mathbf{X}^{(1)} = \mathbf{B}^{(1)}\mathbf{D}\mathbf{S}^T, \quad \mathbf{X}^{(2)} = \mathbf{B}^{(2)}\mathbf{D}\mathbf{S}^T,$$

where again $\mathbf{b}_r^{(1)} = [x_r^{(1-I)/2}, \ldots, x_r^{(I-1)/2}]$ and $\mathbf{b}_r^{(2)} = [y_r^{(1-I)/2}, \ldots, y_r^{(I-1)/2}]$ while now $\mathbf{D} = D_1([y_1^{1/2}, \ldots, y_R^{1/2}])$. Following the same procedure as before we obtain (16) with $[\mathbf{S}^T, \mathbf{S}^H]$ replaced by $[\mathbf{D}\mathbf{S}^T, \mathbf{D}^*\mathbf{S}^H]$.

2) Frame-shaped Array (FA): Consider a FA composed of two horizontal and vertical arrays as depicted in figure 2 (right). The difference between the CA and the FA is that in the latter configuration there is a shift-invariance structure between the involved horizontal and vertical arrays. Consequently, in a FA the individual subarrays are not required to be shift-invariant. Here we consider the simple case where the horizontal and vertical arrays are ULAs and explain how this additional ULA structure can be taken into account.

Build the FA observation matrices $\mathbf{X}^{(1)} \in \mathbb{C}^{2I \times K}$ with columns $\mathbf{x}_k^{(1)} = [x_{11k}, \ldots, x_{I1k}, x_{1Jk}, \ldots, x_{IJk}]^T$ and $\mathbf{X}^{(2)} \in \mathbb{C}^{2J \times K}$ with columns $\mathbf{x}_k^{(2)} = [x_{11k}, \ldots, x_{1Jk}, x_{I1k}, \ldots, x_{IJk}]^T$, where x_{ijk} denotes the output of sensor x_{ij} at the kth snapshot and with indexing as indicated in figure 2 (right). The matrices $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ admit the factorizations

$$\mathbf{X}^{(1)} = \left(\mathbf{B}^{(1)} \odot \mathbf{C}^{(1)}\right)\mathbf{S}^T, \quad \mathbf{X}^{(2)} = \left(\mathbf{B}^{(2)} \odot \mathbf{C}^{(2)}\right)\mathbf{S}^T, \qquad (17)$$

where

$$\begin{aligned} \mathbf{b}_r^{(1)} &= [1 \; x_r \; \ldots \; x_r^{(I-1)}]^T, \quad \mathbf{c}_r^{(1)} = [1 \; y_r^{(J-1)}]^T, \\ \mathbf{b}_r^{(2)} &= [1 \; y_r \; \ldots \; y_r^{(J-1)}]^T, \quad \mathbf{c}_r^{(2)} = [1 \; x_r^{(I-1)}]^T, \end{aligned}$$

in which $x_r = e^{-i\omega_c d_x \sin(\phi_r)\cos(\theta_r)/c}$ where d_x denotes the inter-element spacing along the x-axis, and $y_r = e^{-i\omega_c d_y \sin(\phi_r)\sin(\theta_r)/c}$ where d_y denotes the inter-element spacing along the y-axis. FBA yields

$$\begin{bmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} & \mathbf{J}_{I2}\mathbf{X}^{(1)*} \\ \mathbf{X}^{(2)} & \mathbf{J}_{J2}\mathbf{X}^{(2)*} \end{bmatrix} = \begin{bmatrix} \mathbf{B}^{(1)} \odot \mathbf{C}^{(1)} \\ \mathbf{B}^{(2)} \odot \mathbf{C}^{(2)} \end{bmatrix}\left[\mathbf{S}^T, \mathbf{D}\mathbf{S}^H\right],$$

where $\mathbf{D} = D_1([x_1^{-(I-1)}y_1^{-(J-1)}, \ldots, x_R^{-(I-1)}y_R^{-(J-1)}])$. Overall, we obtain

$$\begin{bmatrix} \mathbf{Z}^{(1)} \\ \mathbf{Z}^{(2)} \end{bmatrix} = \begin{bmatrix} S_{I,2}(\mathbf{Y}^{(1)}) \\ S_{J,2}(\mathbf{Y}^{(2)}) \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{(1)} \odot \mathbf{A}^{(1)} \\ \mathbf{D}^{(2)} \odot \mathbf{A}^{(2)} \end{bmatrix}\left[\mathbf{S}^T, \mathbf{D}\mathbf{S}^H\right], \quad (18)$$

where $\mathbf{A}^{(1)} = \underline{\mathbf{B}}^{(1)} \odot \mathbf{C}^{(1)}$, $\mathbf{A}^{(2)} = \underline{\mathbf{B}}^{(2)} \odot \mathbf{C}^{(2)}$ and

$$\mathbf{D}^{(1)} = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_R \end{bmatrix}, \quad \mathbf{D}^{(2)} = \begin{bmatrix} 1 & \cdots & 1 \\ y_1 & \cdots & y_R \end{bmatrix}.$$

Note that in (18) we have exploited both the shift-invariance structures within and between the horizontal and vertical arrays. The structure in (18) leads to improved uniqueness conditions. As an example, consider the case where $I = J = 5$ and $K \geq R$. If we only exploit the CPD structure of, say $\mathcal{Z}^{(1)}$, then from Theorem II.1 we obtain the bound $R \leq 8$. Thanks to Theorem II.2 the coupled CPD structure of $\{\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)}\}$ relaxes the bound to $R \leq 10$. The coupled CPD of $\{\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)}\}$ can also be computed by means of the SD algorithm developed in Section VI. We limited the exposition to a FA composed of a single frame. However, the result can be extended to FAs composed of several frames.

Other centro-symmetric arrays that fit into this framework include hexagonal shaped arrays [35].

B. Asymmetric arrays

Many array configurations are not symmetric around their "epicenter". We call them asymmetric. For simplicity, we limit the exposition to arrays composed of ULAs.

1) Arrays composed of several linear arrays: Several proposed array configurations are based on the combination of ULAs. Some well-known examples are V-Shaped Arrays (VAs) [10], [9] and L-Shaped Arrays (LAs) [12]. Here we consider the Y-shaped Array (YA) in figure 3 (middle). The YA is clearly composed of three ULAs. Assume that K snapshots are available. Let us choose the sensor denoted by $x_1 = y_1 = z_1$ in figure 3 (middle) as the reference sensor. The outputs of the three ULAs can be stacked into the matrices $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$ and $\mathbf{X}^{(3)}$ according to $(\mathbf{X}^{(1)})_{ik} = x_{ik}$, $(\mathbf{X}^{(2)})_{ik} = y_{ik}$ and $(\mathbf{X}^{(3)})_{ik} = z_{ik}$, where the index k denotes the snapshot number. This leads to the matrix factorizations

$$\mathbf{X}^{(1)} = \mathbf{B}^{(1)}\mathbf{S}^T, \quad \mathbf{X}^{(2)} = \mathbf{B}^{(2)}\mathbf{S}^T, \quad \mathbf{X}^{(3)} = \mathbf{B}^{(3)}\mathbf{S}^T, \qquad (19)$$

where $\mathbf{S} \in \mathbb{C}^{K \times R}$ is the collected signal matrix and

$$\begin{aligned} \mathbf{b}_r^{(1)} &= [1, x_r, \ldots, x_r^{(I_1-1)}], \quad x_r = e^{-i\omega_c d_x \sin(\phi_r)\cos(\theta_r-\alpha)/c}, \\ \mathbf{b}_r^{(2)} &= [1, y_r, \ldots, y_r^{(I_2-1)}], \quad y_r = e^{-i\omega_c d_y \sin(\phi_r)\cos(\theta_r-\beta)/c}, \\ \mathbf{b}_r^{(3)} &= [1, z_r, \ldots, z_r^{(I_3-1)}], \quad z_r = e^{-i\omega_c d_z \sin(\phi_r)\cos(\theta_r-\gamma)/c}, \end{aligned}$$

in which α, β and γ denote the angles between the "x axis" and the individual ULA "axes". Using spatial smoothing on the matrices $\{\mathbf{X}^{(n)}\}$ we obtain tensors $\{\mathcal{Y}^{(n)}\}$ of which the matrix representation admits the coupled decomposition

$$\begin{bmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \\ \mathbf{Y}^{(3)} \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{(1)} \odot \mathbf{A}^{(1)} \\ \mathbf{D}^{(2)} \odot \mathbf{A}^{(2)} \\ \mathbf{D}^{(3)} \odot \mathbf{A}^{(3)} \end{bmatrix}\mathbf{S}^T, \qquad (20)$$

where $\mathbf{A}^{(n)} = \underline{\mathbf{B}}^{(n)}$ and $\mathbf{D}^{(n)} = \mathbf{B}^{(n)}(1:2,:)$. Clearly (20) corresponds to a coupled CPD problem of the form (13). The CPD structure of $\{\mathcal{Y}^{(n)}\}$ is thanks to the ULA geometry. On the other hand, the coupling is due to the fact that $\mathbf{S}$ is the same for the three subarrays. The coupled CPD structure leads to improved uniqueness conditions for the YA. Furthermore, the SD algorithm developed in Section VI can be used for the actual computation.

The VA in [10], [9] is just a YA with $I_3 = 1$. Similarly, the LA in [12] is also just a special case of the YA. For instance, by setting $I_3 = 1$, $\alpha = \frac{\pi}{2}$ and $\beta = \pi$, we obtain the L-shaped array depicted in figure 3 (left).

2) Triangular-shaped Array (TA): Consider the nine-element TA depicted in figure 3 (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 (10) are given by

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

where $y_r = e^{-i\frac{\omega_c}{c} d_y \sin(\theta_r)\sin(\phi_r)}$ and $x_r = e^{-i\frac{\omega_c}{c} d_u \cos(\theta_r-\alpha)\sin(\phi_r)}$ in which α denotes the angle between the "x axis"

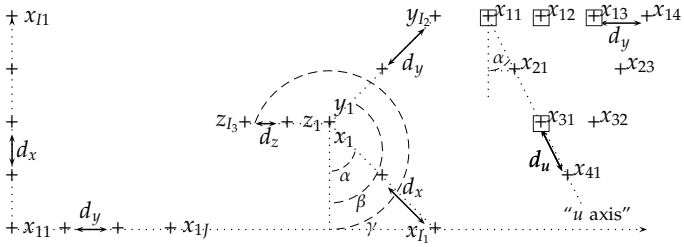


Fig. 3. L-shaped array (left), Y-shaped array (middle) and triangular-shaped array (right) where '+' represents an antenna element.

and the "u axis", see figure 3 (right). At first sight one may view the TA as three superimposed ULAs each composed of three sensors. However, we observe that this TA contains 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 TA has the following three shift-invariance structures:

$$\mathbf{a}_r^{(1)} := \mathbf{a}_r([1,2,3,\ 7]) = \mathbf{a}_r([2,3,4,\ 8])\, y_r^{-1}, \quad (21)$$

$$\mathbf{a}_r^{(2)} := \mathbf{a}_r([1,5,7,\ 3]) = \mathbf{a}_r([5,7,9,\ 7])\, x_r^{-1}, \quad (22)$$

$$\mathbf{a}_r^{(3)} := \mathbf{a}_r([4,6,8,\ 2]) = \mathbf{a}_r([6,8,9,\ 5])\,(x_r/y_r)^{-1}. \quad (23)$$

This motivates the construction of the tensors $\mathcal{X}^{(1)}$, $\mathcal{X}^{(2)}$ and $\mathcal{X}^{(3)}$ 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

$$\mathbf{Y} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \\ \mathbf{X}^{(3)} \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{(1)} \odot \mathbf{A}^{(1)} \\ \mathbf{D}^{(2)} \odot \mathbf{A}^{(2)} \\ \mathbf{D}^{(3)} \odot \mathbf{A}^{(3)} \end{bmatrix} \mathbf{S}^T, \quad (24)$$

where the columns of $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$, $\mathbf{A}^{(3)}$ are of the form (21)–(23) and the columns of $\mathbf{D}^{(1)}$, $\mathbf{D}^{(2)}$, $\mathbf{D}^{(3)}$ are given by $\mathbf{d}_r^{(1)} = \begin{bmatrix} 1 \\ y_r \end{bmatrix}$, $\mathbf{d}_r^{(2)} = \begin{bmatrix} 1 \\ x_r \end{bmatrix}$ and $\mathbf{d}_r^{(3)} = \begin{bmatrix} 1 \\ x_r/y_r \end{bmatrix}$, respectively. Note that if we only exploit the CPD structure of, say $\mathcal{X}^{(1)}$, then from Theorem II.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 II.2 relaxes the bound to $R \leq 6$. Note also that the SD method developed in Section VI can be used for the computation.

V. Vector-sensor array processing via coupled CPD

Consider R far-field wideband signals impinging on an array composed of I vector-sensors (e.g. [20], [19]). The delay between emitter r and the ith vector-sensor is denoted by $\tau_{i,r} \in \mathbb{R}$. The output of the ith vector-sensor at the kth observation is

$$y_{il}(k) = \sum_{r=1}^{R} d_{lr} s_r(k - \tau_{i,r}), \quad (25)$$

where $\mathbf{d}_r \in \mathbb{C}^L$ is a diversity vector. The diversity can for instance be related to the use of polarization [7], [21] or velocity [2], [17], [37] sensitive sensors. Let us choose the first sensor indexed by $i = 1$ as the reference sensor. Since the impinging signals are not narrowband, the observation model in (10) does not apply anymore. Instead we first partition the observed data into M frames each lasting F samples such that the mth frame is given by $\mathbf{y}^{(i,m,l)} = [y_{il}((m-1)F+1), \ldots, y_{il}(mF)]^T$. Next, we transform the data into the frequency domain by means of the DFT matrix $\mathbf{W} \in \mathbb{C}^{F \times F}$ with $(\mathbf{W})_{gh} = e^{-i \cdot 2\pi \cdot g \cdot h/F}$ as follows $\mathbf{z}^{(i,m,l)} = \mathbf{W}\mathbf{y}^{(i,m,l)}$. Finally, we stack the vectors $\{\mathbf{z}^{(i,m,l)}\}$ into $\mathcal{Z}^{(f)} \in \mathbb{C}^{I \times M \times L}$ as follows $z_{i,m,l}^{(f)} = z_f^{(i,m,l)}$. The tensors $\mathcal{Z}^{(f)}$ admit the factorizations

$$\mathcal{Z}^{(f)} = \sum_{r=1}^{R} \mathbf{a}_r^{(f)} \circ \mathbf{s}_r^{(f)} \circ \mathbf{d}_r, \quad f \in \{1, \ldots, F\}, \quad (26)$$

where $\mathbf{s}_r^{(f)} \in \mathbb{C}^M$ denotes the rth signal vector at frequency f and

$$\mathbf{a}_r^{(f)} = [1, e^{-i2\pi f \mathbf{b}_r^T \mathbf{p}_2/c}, \ldots, e^{-i2\pi f \mathbf{b}_r^T \mathbf{p}_I/c}]^T \in \mathbb{C}^I \quad (27)$$

is now the frequency dependent steering vector associated with the rth signal at frequency f. As in the narrowband model (10), $\mathbf{p}_i$ denotes the position of the ith vector-sensor and $\mathbf{b}_r$ denotes the bearing vector associated with the rth source.

The set of tensors $\{\mathcal{Z}^{(f)}\}$ given by (26) clearly admits a coupled CPD interpretation. Hence, the uniqueness conditions and algorithms developed in [29], [30] for coupled CPD can directly be applied to the wideband VSA problem in (26). Note that we did not impose any particular geometry on the sensor array. Wideband VSA processing in the case of random array geometries is for instance relevant for sonobouys with arbitrary displacements due to the drifting in the water (e.g. [17]).

In VSA one of the involved factor matrices may be real. For instance, in acoustic VSA processing $\mathbf{d}_r$ in (26) takes the form $\mathbf{d}_r = [1,\ \sin(\phi_r)\cos(\theta_r),\ \sin(\phi_r)\sin(\theta_r),\ \cos(\phi_r)]^T$ [19]. This property can be exploited. In short, define $\mathcal{X}^{(f)} \in \mathbb{C}^{I \times M \times L}$ as follows $x_{iml}^{(f)} = z_{iml}^{(f)*}$. We now obtain the tensors $\{\mathcal{Z}^{(f)}, \mathcal{X}^{(f)}\}$ with coupled CPD of the form

$$\begin{cases} \mathcal{Z}^{(f)} = \sum_{r=1}^{R} \mathbf{a}_r^{(f)} \circ \mathbf{s}_r^{(f)} \circ \mathbf{d}_r, & f \in \{1, \ldots, F\}, \\ \mathcal{X}^{(f)} = \sum_{r=1}^{R} \mathbf{a}_r^{(f)*} \circ \mathbf{s}_r^{(f)*} \circ \mathbf{d}_r, & f \in \{1, \ldots, F\}. \end{cases}$$

In practice, the VSA is often regularly structured. For instance, in [2], [7], [21] the VSA is ULA structured. We stress that the result presented here is valid for any structured VSA enjoying a shift-invariance property. However, let us for simplicity limit the discussion to the case of a ULA along the x-axis such that the array response vector (27) takes the form $\mathbf{a}_r^{(f)} = [1, x_r(f), \ldots, x_r(f)^{I-1}]^T \in \mathbb{C}^I$, in which $x_r(f) = e^{-i2\pi f d_x \sin(\phi_r)\cos(\theta_r)/c}$ and d_x denotes the inter-element spacing along the x-axis. Using spatial smoothing we obtain the fourth-order tensors $\mathcal{V}^{(f)} \in \mathbb{C}^{K_{1,f} \times L_{1,f} \times M \times L}$ with factorizations

$$\mathcal{V}^{(f)} = \sum_{r=1}^{R} \mathbf{a}_r^{(K_1,f)} \circ \mathbf{a}_r^{(L_1,f)} \circ \mathbf{s}_r^{(f)} \circ \mathbf{d}_r, \quad f \in \{1, \ldots, F\}, \quad (28)$$

where $\mathbf{a}_r^{(K_1,f)} = [1\ x_r(f)\ \dots\ x_r(f)^{(K_{1,f}-1)}]^T$ and $\mathbf{a}_r^{(L_1,f)} = [1\ x_r(f)\ \dots\ x_r(f)^{(L_{1,f}-1)}]^T$ are subject to $K_{1,f} + L_{1,f} = I + 1$. Clearly the factorization of the fourth-order tensors $\{\mathcal{V}^{(f)}\}$ in (28) admits a coupled CPD interpretation. The combination of coupled and higher-order structures leads to improved uniqueness conditions and algorithms, see [29], [30] for details.

VI. SD method for Coupled MPs

In Sections III–V we demonstrated that several problems in array processing admit a coupled CPD interpretation. This in turn implies that these problems can be solved by means of the SD procedure for coupled CPD [30]. From Section IV it was also clear that a large subset of these problems are of the form $\mathbf{X}^{(n)} = \mathbf{A}^{(n)}\mathbf{S}^T \in \mathbb{C}^{I_n \times K}$, $n \in \{1,\dots,N\}$, in which $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ enjoys a shift-invariance property. Subsection VI-A provides a coupled MP interpretation for this problem. Next, in Subsection VI-B we present a SD method suited for this coupled MP problem. We pay special attention to the shift-invariance property, $\overline{\mathbf{A}}^{(n)} = \underline{\mathbf{A}}^{(n)} D_1([x_{n,1},\dots,x_{n,R}])$ with $x_{n,r} \in \mathbb{C}$, characterizing the Vandermonde structure associated with the uniform linear subarrays.

A. Connection between coupled MP and coupled CPD

Consider the decomposition (covering all cases in Section IV):[2] [3]

$$\mathbf{X} = \mathbf{A}\mathbf{S}^T \in \mathbb{C}^{I \times K}, \tag{29}$$

where the matrix $\mathbf{A} \in \mathbb{C}^{I \times R}$ is shift-invariance structured while $\mathbf{S} \in \mathbb{C}^{K \times R}$ is unstructured. More precisely, let $\mathbf{C}^{(n)} \in \mathbb{C}^{I_n \times I}$ denote a row-selection matrix such that $\mathbf{A}^{(n)} = \mathbf{C}^{(n)}\mathbf{A} \in \mathbb{C}^{I_n \times R}$ consists of I_n rows of $\mathbf{A}$. Then it is assumed that there exist matrices $\hat{\mathbf{P}}^{(n)}, \check{\mathbf{P}}^{(n)} \in \mathbb{C}^{J_n \times I_n}$ with $J_n < I_n$ and scalars $x_{n,r} \in \mathbb{C}$ such that

$$\hat{\mathbf{P}}^{(n)}\mathbf{A}^{(n)}\mathbf{S}^T = \check{\mathbf{P}}^{(n)}\mathbf{A}^{(n)} D_1([x_{n,1},\dots,x_{n,R}])\mathbf{S}^T, \quad n \in \{1,\dots,N\}. \tag{30}$$

The simplest example is the exponential case in which $a_{ir}^{(n)} = x_{n,r}^i$ such that $a_{ir}^{(n)} = a_{i-1,r}^{(n)} x_{n,r}$, $\hat{\mathbf{P}}^{(n)} = \overline{\mathbf{I}}_{I_n}$ and $\check{\mathbf{P}}^{(n)} = \underline{\mathbf{I}}_{I_n}$. Another very basic example is sinusoidal functions $a_{ir}^{(n)} = \alpha_{r,n}\cos(\omega_{r,n}\cdot i + \beta_{r,n})$ with $\alpha_{r,n} \in \mathbb{C}$, $\omega_{r,n}, \beta_{r,n} \in \mathbb{R}$ and $a_{i+1,r}^{(n)} + a_{i-1,r}^{(n)} = 2\cdot a_{i,r}^{(n)} x_{n,r}$ in which $x_{n,r} = \cos(\omega_{r,n})$ (e.g. [36], [16]). In the sinusoidal case we have $\hat{\mathbf{P}}^{(n)} = \mathbf{I}_{I_n}(3 : I_n, :) + \mathbf{I}_{I_n}(1 : I_n - 2, :)$ and $\check{\mathbf{P}}^{(n)} = 2 \cdot I_{I_n}(2 : I_n - 1, :)$.

[2] In the case of centro-symmetric arrays we first apply FBA such that $\mathbf{X} = \mathbf{A}\mathbf{S}^T$ in (29) is replaced by $[\mathbf{X}, \mathbf{J}_I\mathbf{X}^*] = \mathbf{A}[\mathbf{S}^T, \mathbf{S}^H]$.

[3] Note that model (29) can also be used for 3D arrays (e.g. [8]). Model (29) can also be used for M widely separated antenna arrays provided that the impinging signals are narrowband for the global array (e.g. [1]). In short, by taking $\mathbf{A}$ in (29) equal to $\mathbf{A} = [(\mathbf{B}^{(1)}\Lambda^{(1)})^T, \dots, (\mathbf{B}^{(M)}\Lambda^{(M)})^T]^T$, where $\mathbf{B}^{(m)}$ denotes a shift-invariance structured antenna response matrix associated with the mth antenna array and where $\Lambda^{(m)}$ denotes the diagonal matrix that holds on its diagonal the channel attenuation factors between the mth antenna array and the emitters, the widely separated case is indeed also described by (29).

Clearly (30) is related to coupled CPD. Indeed, for every $n \in \{1,\dots,N\}$ we can build a tensor $\mathcal{Y}^{(n)} \in \mathbb{C}^{2 \times J_n \times K}$ for which the matrix representation admits the decomposition

$$\mathbf{Y}^{(n)} = \begin{bmatrix} \hat{\mathbf{P}}^{(n)}\mathbf{C}^{(n)}\mathbf{X} \\ \check{\mathbf{P}}^{(n)}\mathbf{C}^{(n)}\mathbf{X} \end{bmatrix} = \left(\mathbf{D}^{(n)} \odot (\hat{\mathbf{P}}^{(n)}\mathbf{A}^{(n)})\right)\mathbf{S}^T, \ n \in \{1,\dots,N\}, \tag{31}$$

where $\mathbf{D}^{(n)} = \begin{bmatrix} 1 & \cdots & 1 \\ x_{n,1} & \cdots & x_{n,R} \end{bmatrix} \in \mathbb{C}^{2 \times R}$ and the relation $\hat{\mathbf{P}}^{(n)}\mathbf{A}^{(n)} = \check{\mathbf{P}}^{(n)}\mathbf{A}^{(n)} D_1([x_{n,1},\dots,x_{n,R}]) \in \mathbb{C}^{J_n \times R}$ has been used.

Since each tensor $\mathcal{Y}^{(n)}$, $n \in \{1,\dots,N\}$ only consists of two matrix slices, the coupled CPD of $\{\mathcal{Y}^{(n)}\}$ can be interpreted as a coupled MP problem. More precisely, define $\mathbf{\Gamma} = [\boldsymbol{\gamma}_1,\dots,\boldsymbol{\gamma}_R] := (\mathbf{S}^T)^\dagger$ and $\mathbf{\Upsilon}^{(n)} = D_1([\lambda_{n,1},\dots,\lambda_{n,R}]) := (D_1([x_{n,1},\dots,x_{n,R}]))^{-1}$, then we obtain the coupled MP problem

$$\hat{\mathbf{P}}^{(n)}\mathbf{X}^{(n)}\mathbf{\Gamma} = \check{\mathbf{P}}^{(n)}\mathbf{X}^{(n)}\mathbf{\Gamma}\mathbf{\Upsilon}^{(n)} \Leftrightarrow (\hat{\mathbf{P}}^{(n)}\mathbf{X}^{(n)} - \lambda_{n,r}\check{\mathbf{P}}^{(n)}\mathbf{X}^{(n)})\boldsymbol{\gamma}_r = \mathbf{0}_{J_n},$$

where $\mathbf{X}^{(n)} = \mathbf{C}^{(n)}\mathbf{X}$, $r \in \{1,\dots,R\}$ and $n \in \{1,\dots,N\}$. In other words, the factorization problem (29) amounts to finding the common generalized eigenvectors of a set of matrix-pencils $\{(\hat{\mathbf{P}}^{(n)}\mathbf{X}^{(n)}, \check{\mathbf{P}}^{(n)}\mathbf{X}^{(n)})\}_{n=1}^N$. In contrast to the matrix-pencil case ($N = 1$), the generalized eigenvectors $(\boldsymbol{\gamma}_r)$ associated with the coupled MP can be unique despite $\ker\left(\hat{\mathbf{P}}^{(n)}\mathbf{X}^{(n)}\right) \cap \ker\left(\check{\mathbf{P}}^{(n)}\mathbf{X}^{(n)}\right) \neq \emptyset$ or $\lambda_{n,r} = \lambda_{n,s}$ for some $r \neq s$.

B. SD method for coupled MP

We now develop an efficient SD procedure for the shift-invariant structured matrix factorization problem (29) that admits a coupled MP interpretation. We pay special attention to the shift-invariance structure $\overline{\mathbf{A}}^{(n)} = \underline{\mathbf{A}}^{(n)} D_1([x_{n,1},\dots,x_{n,R}])$ characterizing the exponential structure in for example DOA estimation.

a) Step 0: Dimensionality reduction: Recall that we assume that the full column rank conditions (7) in Theorem II.2 are satisfied for the coupled CPD of $\{\mathcal{Y}^{(n)}\}$ with matrix representation (31). Then $\mathbf{A}$ in (29) has full column rank. Indeed, if $\mathbf{A}$ does not have full column rank, then there exists a vector $\mathbf{y} \in \ker(\mathbf{A})$ and a vector $\mathbf{x} \in \mathbb{C}^K$ such that $\mathbf{X} = \mathbf{A}(\mathbf{S} + \mathbf{x}\mathbf{y}^T)^T$ has rank $R - 1$. This in turn contradicts Theorem II.2. Let $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ denote the compact SVD of $\mathbf{X}$, where $\mathbf{U} \in \mathbb{C}^{I \times R}$ and $\mathbf{V} \in \mathbb{C}^{K \times R}$ are columnwise orthonormal and $\mathbf{\Sigma} \in \mathbb{C}^{R \times R}$ is diagonal. We know that $\text{range}(\mathbf{X}) = \text{range}(\mathbf{U})$ and that there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\mathbf{\Sigma} = \mathbf{A}\mathbf{F}^T \text{ and } \mathbf{V}^*\mathbf{F}^{-1} = \mathbf{S}. \tag{32}$$

b) Step 1: From shift-invariant structured matrix factorization to coupled MP: From (32) we build tensors $\mathcal{Y}_{\text{red}}^{(n)} \in \mathbb{C}^{2 \times J_n \times R}$, $n \in \{1,\dots,N\}$, each admitting the decomposition

$$\mathbf{Y}_{\text{red}}^{(n)} = \begin{bmatrix} \mathbf{U}^{(1,n)} \\ \mathbf{U}^{(2,n)} \end{bmatrix} := \begin{bmatrix} \hat{\mathbf{P}}^{(n)}\mathbf{U}^{(n)} \\ \check{\mathbf{P}}^{(n)}\mathbf{U}^{(n)} \end{bmatrix} = (\mathbf{D}^{(n)} \odot (\hat{\mathbf{P}}^{(n)}\mathbf{A}^{(n)}))\mathbf{F}^T, \tag{33}$$

where 'red' stands for reduced and $\mathbf{U}^{(n)} = \mathbf{C}^{(n)}\mathbf{U}\mathbf{\Sigma} \in \mathbb{C}^{I_n \times R}$. In other words, by working with $\mathbf{U}\mathbf{\Sigma}$ in (32) instead of $\mathbf{X}$ in (29), we obtain the compressed variant $\mathbf{Y}_{\text{red}}^{(n)}$ of $\mathbf{Y}^{(n)}$ in (31). Stacking for all $n \in \{1, \ldots, N\}$ yields:

$$\begin{bmatrix} \mathbf{Y}_{\text{red}}^{(1)} \\ \vdots \\ \mathbf{Y}_{\text{red}}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{(1)} \odot (\hat{\mathbf{P}}^{(1)}\mathbf{A}^{(1)}) \\ \vdots \\ \mathbf{D}^{(N)} \odot (\hat{\mathbf{P}}^{(N)}\mathbf{A}^{(N)}) \end{bmatrix} \mathbf{F}^T. \tag{34}$$

c) Step 2: From coupled MP to SD: From (34) we build R symmetric matrices $\{\mathbf{M}^{(r)}\}$, admitting the factorizations

$$\mathbf{M}^{(r)} = \mathbf{G}\mathbf{\Lambda}^{(r)}\mathbf{G}^T, \quad r \in \{1, \ldots, R\}, \tag{35}$$

where $\mathbf{G} = \mathbf{F}^{-1}$ and $\mathbf{\Lambda}^{(r)} \in \mathbb{C}^{R \times R}$ are diagonal matrices. The construction will be explained below. Once the SD factorization (35) has been computed, we obtain $\mathbf{A}$ and $\mathbf{S}$ via relation (32). Note that, in the exact case, the SD factorization (35) can be computed via a GEVD, with the columns of $\mathbf{G}^{-T}$ as generalized eigenvectors.

We will now first explain how to construct the matrices $\{\mathbf{M}^{(r)}\}$ in (35). Thereafter, we explain that the complexity of the construction can be reduced by exploiting the shift-invariance structures in $\mathbf{A}$.

Construction of matrices $\{\boldsymbol{M}^{(r)}\}$: Consider the bilinear mappings $\Phi^{(n)} : \mathbb{C}^{J_n \times 2} \times \mathbb{C}^{J_n \times 2} \to \mathbb{C}^{4J_n^2}$ defined by [3]:

$$\Phi^{(n)}(\mathbf{X}, \mathbf{Y})_{(i-1)4J_n+4(j-1)+2(k-1)+l} = x_{ik}y_{jl} + y_{ik}x_{jl} - x_{il}y_{jk} - y_{il}x_{jk}.$$

Denote $\mathbf{Y}_{\text{red}}^{(n)} = [\widetilde{\mathbf{y}}_1^{(n)}, \ldots, \widetilde{\mathbf{y}}_R^{(n)}]$ and define $\Phi_{r,s}^{(n)} = \Phi^{(n)}\left(\mathbf{Y}_{\text{red}}^{(n,r)}, \mathbf{Y}_{\text{red}}^{(n,s)}\right)$ in which

$$\begin{aligned} \mathbf{Y}_{\text{red}}^{(n,r)} &= \text{Unvec}\left(\widetilde{\mathbf{y}}_r^{(n)}\right) = [\hat{\mathbf{P}}^{(n)}\mathbf{u}_r^{(n)}, \check{\mathbf{P}}^{(n)}\mathbf{u}_r^{(n)}] \\ &= [\mathbf{u}_r^{(1,n)}, \mathbf{u}_r^{(2,n)}], \end{aligned} \tag{36}$$

$$\widetilde{\mathbf{y}}_r^{(n)} = \text{Vec}\left(\mathbf{Y}_{\text{red}}^{(n,r)}\right) = \begin{bmatrix} \mathbf{u}_r^{(1,n)} \\ \mathbf{u}_r^{(2,n)} \end{bmatrix}. \tag{37}$$

Construct $\mathbf{Q} \in \mathbb{C}^{4(\sum_{n=1}^N J_n^2) \times \frac{R(R+1)}{2}}$ as follows

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}^{(1)} \\ \vdots \\ \mathbf{Q}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}^{(1,1)} & 2 \cdot \mathbf{Q}^{(2,1)} \\ \vdots & \vdots \\ \mathbf{Q}^{(1,N)} & 2 \cdot \mathbf{Q}^{(2,N)} \end{bmatrix}, \tag{38}$$

where $\mathbf{Q}^{(1,n)} = \left[\Phi_{1,1}^{(n)}, \Phi_{2,2}^{(n)}, \ldots, \Phi_{R,R}^{(n)}\right] \in \mathbb{C}^{4J_n^2 \times R}$ and $\mathbf{Q}^{(2,n)} = \left[\Phi_{1,2}^{(n)}, \Phi_{1,3}^{(n)}, \Phi_{2,3}^{(n)}, \ldots, \Phi_{R-1,R}^{(n)}\right] \in \mathbb{C}^{4J_n^2 \times \frac{R(R-1)}{2}}$. It can be verified that if Theorem II.2 is satisfied for the coupled CPD of $\{\mathcal{Y}^{(n)}\}$, then

$$\mathbf{Q}^H\mathbf{Q} = \sum_{n=1}^N \begin{bmatrix} \mathbf{Q}^{(1,n)H}\mathbf{Q}^{(1,n)} & 2 \cdot \mathbf{Q}^{(1,n)H}\mathbf{Q}^{(2,n)} \\ 2 \cdot \mathbf{Q}^{(2,n)H}\mathbf{Q}^{(1,n)} & 4 \cdot \mathbf{Q}^{(2,n)H}\mathbf{Q}^{(2,n)} \end{bmatrix} \tag{39}$$

has an R-dimensional kernel and the matrices $\{\mathbf{M}^{(r)}\}$ can be obtained from the kernel of $\mathbf{Q}^H\mathbf{Q}$ [30]. Namely, $\mathbf{M}^{(r)}$ satisfies

$$\mathbf{Q}^H\mathbf{Q}\mathbf{m}^{(r)} = \mathbf{0}, \ \forall r \in \{1, \ldots, R\},$$

where $\mathbf{m}^{(r)} = \left[m_{1,1}^{(r)}, m_{2,2}^{(r)}, \ldots, m_{R,R}^{(r)}, m_{1,2}^{(r)}, m_{1,3}^{(r)}, \ldots, m_{R-1,R}^{(r)}\right]^T$.

Define the set $S_R = \{(s_1, s_2) : 1 \leq s_1 < s_2 \leq R\}$ in which the C_R^2 elements are ordered lexicographically. Consider the mapping $f_R : \mathbb{N}^2 \to \{1, 2, \ldots, C_R^2\}$ that returns the position of its argument in the set S_R, e.g., $f_R((1,3)) = 3$. Then a technical derivation yields

$$\begin{aligned} &(\mathbf{Q}^{(1,n)H}\mathbf{Q}^{(1,n)})_{r,s} = \Phi_{r,r}^{(n)H}\Phi_{s,s}^{(n)} = 8\left(w_{rs}^{(n)}w_{rs}^{(n)} - \mathbf{b}_{rs}^{(n)H}\mathbf{b}_{rs}^{(n)}\right), \\ &(\mathbf{Q}^{(1,n)H}\mathbf{Q}^{(2,n)})_{r,f_R(s_1,s_2)} = \Phi_{r,r}^{(n)H}\Phi_{s_1,s_2}^{(n)} = 16(w_{rs_1}^{(n)}w_{rs_2}^{(n)} - \mathbf{b}_{rs_1}^{(n)H}\mathbf{b}_{rs_2}^{(n)}), \\ &(\mathbf{Q}^{(2,n)H}\mathbf{Q}^{(2,n)})_{f_R(r_1,r_2),f_R(s_1,s_2)} = \Phi_{r_1,r_2}^{(n)H}\Phi_{s_1,s_2}^{(n)} = \\ &\quad 4(w_{r_1s_1}^{(n)}w_{r_2s_2}^{(n)} + w_{r_2s_1}^{(n)}w_{r_1s_2}^{(n)} - \mathbf{b}_{r_1s_1}^{(n)H}\mathbf{b}_{r_2s_2}^{(n)} - \mathbf{b}_{r_2s_1}^{(n)H}\mathbf{b}_{r_1s_2}^{(n)}), \end{aligned}$$

where

$$\begin{aligned} w_{rs}^{(n)} &:= \text{Vec}\left(\mathbf{Y}_{\text{red}}^{(n,r)}\right)^H \text{Vec}\left(\mathbf{Y}_{\text{red}}^{(n,s)}\right) = w_{rs}^{(1,n)} + w_{rs}^{(2,n)}, \\ w_{rs}^{(1,n)} &= \mathbf{u}_r^{(1,n)H}\mathbf{u}_s^{(1,n)}, \\ w_{rs}^{(2,n)} &= \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(2,n)}, \end{aligned} \tag{40}$$

and

$$\begin{aligned} \mathbf{b}_{rs}^{(n)} &:= \text{Vec}\left(\mathbf{Y}_{\text{red}}^{(n,r)H}\mathbf{Y}_{\text{red}}^{(n,s)}\right) = [w_{rs}^{(1,n)}, z_{rs}^{(n)}, z_{rs}^{(n)*}, w_{rs}^{(2,n)}]^T, \\ z_{rs}^{(n)} &= \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(1,n)}, \\ \mathbf{b}_{sr}^{(n)} &= \begin{bmatrix} 1&0&0&0 \\ 0&0&1&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{bmatrix}\mathbf{b}_{rs}^{(n)*} = [w_{rs}^{(1,n)*}, z_{rs}^{(n)*}, z_{rs}^{(n)}, w_{rs}^{(2,n)*}]^T. \end{aligned} \tag{41}$$

We can now conclude that the computational cost of $\mathbf{Q}^H\mathbf{Q}$ is dominated by the construction of the matrices

$$\mathbf{W}^{(1,n)} = \mathbf{U}^{(1,n)H}\mathbf{U}^{(1,n)} \in \mathbb{C}^{R \times R}, \quad n \in \{1, \ldots, N\}, \tag{42}$$

$$\mathbf{W}^{(2,n)} = \mathbf{U}^{(2,n)H}\mathbf{U}^{(2,n)} \in \mathbb{C}^{R \times R}, \quad n \in \{1, \ldots, N\}, \tag{43}$$

$$\mathbf{Z}^{(n)} = \mathbf{U}^{(2,n)H}\mathbf{U}^{(1,n)} \in \mathbb{C}^{R \times R}, \quad n \in \{1, \ldots, N\}, \tag{44}$$

where $\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)} \in \mathbb{C}^{J_n \times R}$. Observe that the matrices in (42) and in (43) are Hermitian, implying that only the upper C_{R+1}^2 entries of $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}\}$ need to be computed. The computation of $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^N$ requires $(2C_{R+1}^2 + R^2)\sum_{n=1}^N(2J_n - 1) = (4R^2 + 2R)(\sum_{n=1}^N J_n) - (2R^2 + R)N = O((4\sum_{n=1}^N J_n) - 2N)R^2)$ flops[4].

Exploiting shift-invariance structures in $\boldsymbol{A}$: The complexity of the construction of $\{\mathbf{M}^{(r)}\}$ can be reduced by utilizing the specific shift-invariance structure in $\mathbf{A}$. In the case of exponential structures we have $\overline{\mathbf{A}}^{(n)} = \underline{\mathbf{A}}^{(n)}D_1([x_{n,1}, \ldots, x_{n,R}])$. This property implies that $J_n = I_n - 1$ and that $\begin{bmatrix} \mathbf{U}^{(1,n)} \\ \mathbf{U}^{(2,n)} \end{bmatrix} \in \mathbb{C}^{2J_n \times R}$ in (33) has the structure

$$\mathbf{U}^{(1,n)} = \begin{bmatrix} \mathbf{U}^{(n)}(1,:) \\ \mathbf{U}^{(2,n)}(1:I_n-2,:) \end{bmatrix}, \ \mathbf{U}^{(2,n)} = \begin{bmatrix} \mathbf{U}^{(1,n)}(2:I_n-1,:) \\ \mathbf{U}^{(n)}(I_n,:) \end{bmatrix}.$$

The property $\overline{\mathbf{A}}^{(n)} = \underline{\mathbf{A}}^{(n)}D_1([x_{n,1}, \ldots, x_{n,R}])$ also implies that the matrix in (36) and the vector in (37) are equal to

$$\begin{aligned} \mathbf{Y}^{(n,r)} &= [\mathbf{u}_r^{(1,n)}, \mathbf{u}_r^{(2,n)}] = [\mathbf{u}_r^{(n)}(1:I_n-1), \mathbf{u}_r^{(n)}(2:I_n)], \\ \widetilde{\mathbf{y}}_r^{(n)} &= \text{Vec}\left(\mathbf{Y}^{(n,r)}\right) = \begin{bmatrix} \mathbf{u}_r^{(1,n)} \\ \mathbf{u}_r^{(2,n)} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_r^{(n)}(1:I_n-1) \\ \mathbf{u}_r^{(n)}(2:I_n) \end{bmatrix}. \end{aligned}$$

The computation of $\{\mathbf{W}^{(2,n)}\}$ in (43) now simplifies to

$$\mathbf{W}^{(2,n)} = \mathbf{W}^{(1,n)} - \hat{\mathbf{u}}^{(n)}\hat{\mathbf{u}}^{(n)H} + \check{\mathbf{u}}^{(n)}\check{\mathbf{u}}^{(n)H}, \quad n \in \{1, \ldots, N\}, \tag{45}$$

[4]A flop is here understood as a complex scalar addition, subtraction or multiplication.

where $\hat{\mathbf{u}}^{(n)H} = \mathbf{U}^{(n)}(1,:)$ and $\check{\mathbf{u}}^{(n)H} = \mathbf{U}^{(n)}(I_n,:)$. Note that (45) only requires two rank-1 updates, requiring $4C_{R+1}^2$ flops in total. Hence, in the case of exponential structures with property $\overline{\mathbf{A}}^{(n)} = \underline{\mathbf{A}}^{(n)} D_1([x_{n,1},\ldots,x_{n,R}])$, the computation of $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \overline{\mathbf{Z}}^{(n)}\}_{n=1}^N$ only requires $(C_{R+1}^2 + R^2)\sum_{n=1}^N (2J_n - 1) + 4NC_{R+1}^2 = (3R^2 + R)(\sum_{n=1}^N J_n) + \frac{1}{2}R^2N + \frac{3}{2}RN = O((3\sum_{n=1}^N J_n + \frac{N}{2})R^2)$ flops.

d) Step 3: From SD to single-source DOA problems: After $\mathbf{G}$ has been found from solving the SD problem (35) we obtain $\mathbf{A}$ and $\mathbf{S}$ from relation (32). The shift-invariance parameters associated with the rth source can now be deduced from $\mathbf{a}_r$. In other words, the coupled CPD framework has turned the factorization problem (29) into a set of decoupled single-source DOA problems.

Let $x_{r,n}$ denote the generator associated with the rth source and the nth shift-invariance structure. Define

$$\mathbf{c}_r^{(1,n)} = \hat{\mathbf{P}}^{(n)}\mathbf{a}_r^{(n)} \quad \text{and} \quad \mathbf{c}_r^{(2,n)} = \check{\mathbf{P}}^{(n)}\mathbf{a}_r^{(n)}. \tag{46}$$

If $x_{r,n} \in \mathbb{C}$, then since $\mathbf{c}_r^{(1,n)} = \mathbf{c}_r^{(2,n)} x_{r,n}$, the generator follows from $x_{r,n} = (\mathbf{c}_r^{(2,n)H}\mathbf{c}_r^{(1,n)})/(\mathbf{c}_r^{(2,n)H}\mathbf{c}_r^{(2,n)})$. However, in DOA estimation $x_{r,n}$ is located on the unit circle, i.e., $x_{r,n} = e^{i\cdot\psi_{r,n}}$ for some $\psi_{r,n} \in \mathbb{R}$. Using the identities $\cos(\psi_{r,n}) = \cos^2(\psi_{r,n}/2) - \sin^2(\psi_{r,n}/2)$ and $\sin(\psi_{r,n}) = 2\sin(\psi_{r,n}/2)\cos(\psi_{r,n}/2)$, the unit-norm generator can be determined from

$$\begin{aligned} f(\psi_{r,n}) &= \left\|\mathbf{c}_r^{(1,n)} x_{r,n} - \mathbf{c}_r^{(2,n)}\right\|_F^2 = \left\|\mathbf{c}_r^{(1,n)} e^{i\cdot\psi_{r,n}} - \mathbf{c}_r^{(2,n)}\right\|_F^2 \\ &= \alpha_{r,n} - 2\mathrm{Re}\{\beta_{r,n}\}\cos(\psi_{r,n}) - 2\mathrm{Im}\{\beta_n\}\sin(\psi_{r,n}) \\ &= \alpha_{r,n} - 2\mathrm{Re}\{\beta_{r,n}\}\left(\cos^2(\psi_{r,n}/2) - \sin^2(\psi_{r,n}/2)\right) \\ &\quad - 4\mathrm{Im}\{\beta_{r,n}\}\sin(\psi_{r,n}/2)\cos(\psi_{r,n}/2) \\ &= \alpha_{r,n} - 2\begin{bmatrix} \cos(\psi_{r,n}/2) \\ \sin(\psi_{r,n}/2) \end{bmatrix}^T \mathbf{G}^{(r,n)} \begin{bmatrix} \cos(\psi_{r,n}/2) \\ \sin(\psi_{r,n}/2) \end{bmatrix}, \end{aligned} \tag{47}$$

where $\alpha_{r,n} = \mathbf{c}_r^{(1,n)H}\mathbf{c}_r^{(1,n)} + \mathbf{c}_r^{(2,n)H}\mathbf{c}_r^{(2,n)}$, $\beta_{r,n} = \mathbf{c}_r^{(1,n)H}\mathbf{c}_r^{(2,n)}$ and $\mathbf{G}^{(r,n)} = \begin{bmatrix} -\mathrm{Re}\{\beta_{r,n}\} & \mathrm{Im}\{\beta_{r,n}\} \\ \mathrm{Im}\{\beta_{r,n}\} & \mathrm{Re}\{\beta_{r,n}\} \end{bmatrix}$. Let $\mathbf{x}^{(r,n)} \in \mathbb{R}^2$ denote the eigenvector associated with the dominant eigenvalue of $\mathbf{G}^{(r,n)}$, then $x_{r,n} = (x_1^{(r,n)} + i\cdot x_2^{(r,n)})^2$.

e) Summary of algorithm: The proposed SD procedure tailored to the shift-invariance structured matrix factorization problem (29) is summarized as Algorithm 1. Note that if Theorem II.2 is satisfied for the coupled CPD of $\{\mathcal{Y}_{\text{red}}^{(n)}\}$ with matrix decomposition (34), then Algorithm 1 is guaranteed to find the decomposition in the exact case. Finally, from $\{x_{r,n}\}$ we can deduce the DOA parameters $\{\theta_r, \phi_r\}$. The exact procedure for extracting $\{\theta_r, \phi_r\}$ from $\{x_{r,n}\}$ depends on the given antenna array configuration.

VII. Numerical experiments

Consider the factorization $\mathbf{X} = \mathbf{A}\mathbf{S}^T$ in (10). The goal is to estimate the angles $\{\theta_r, \phi_r\}$ from $\mathbf{T} = \mathbf{X} + \beta\mathbf{N}$, where $\mathbf{N}$ is an unstructured perturbation matrix and $\beta \in \mathbb{R}$ controls the noise level. The azimuth angles $\{\theta_r\}$ and elevation angles $\{\phi_r\}$ are randomly chosen on the intervals $[0, \pi]$ and $[-\pi/2, \pi/2]$, respectively. The real and imaginary entries of the perturbation matrices are randomly drawn from a Gaussian distribution with zero mean and unit variance. The following Signal-to-Noise Ratio (SNR) measure will be used: $\text{SNR [dB]} = 10\log(\|\mathbf{X}\|_F^2/\|\beta\mathbf{N}\|_F^2)$. The performance evaluation will be based on the distance between the true angles $\{\theta_r, \phi_r\}$ and their estimates $\{\hat{\theta}_r, \hat{\phi}_r\}$. The distance is measured according to

$$\mathrm{E_{RMS}} = \min_{\sigma\in\Lambda_R} \sqrt{\frac{1}{R}\sum_{r=1}^R \left(\theta_r - \hat{\theta}_{\sigma(r)}\right)^2 + \left(\phi_r - \hat{\phi}_{\sigma(r)}\right)^2},$$

where Λ_R denotes the set of $R!$ possible permutations of the elements $\{1,\ldots,R\}$ and $\sigma(r)$ denotes the rth element of the corresponding permuted set.

Algorithm 1 SD procedure for matrix factorization involving N shift-invariance structures (Theorem II.2).

Input: $\mathbf{X} = \mathbf{A}\mathbf{S}^T$ of the form (29).
1. Compute SVD $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ (and find R from $\mathbf{\Sigma}$).
2. Choose matrices $\{(\mathbf{C}^{(n)}, \hat{\mathbf{P}}^{(n)}, \check{\mathbf{P}}^{(n)})\}_{n=1}^N$.
3. Build $\mathbf{Y}_{\text{red}}^{(n)}$ in (33) for every $n \in \{1,\ldots,N\}$.
4. Construct $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^N$ given by (42)–(44).
5. Construct $\mathbf{Q}^H\mathbf{Q}$ in (39) via $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^N$.
6. Obtain matrices $\{\mathbf{M}^{(r)}\}$ from the kernel of $\mathbf{Q}^H\mathbf{Q}$.
7. Solve SD problem $\mathbf{M}^{(r)} = \mathbf{G}\mathbf{\Lambda}^{(r)}\mathbf{G}^T$, $r \in \{1,\ldots,R\}$.
8. Compute $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{G}^T$ and $\mathbf{S} = \mathbf{V}^*\mathbf{G}^{-1}$.
9. Obtain $\{x_{r,n}\}$ from $\mathbf{A}$ (e.g. via $f(\psi_{r,n})$ given by (47)).

Output: $\{x_{r,n}\}$, $\mathbf{A}$ and $\mathbf{S}$.

A. Centro-symmetric Cross-shaped Array (CA)

Consider the CA in subsection IV-A with $I = J = 5$, $K = 50$ and R varying. We compare Algorithm 1 with an adapted version of the 2D unitary ESPRIT method for centro-symmetric arrays, of which all subarrays yield unique DOAs [38]. This is in contrast to Algorithm 1 which only requires that the global array yields unique DOAs.

Case $R = 4$: Note that $R = 4$ is the maximal bound on R for the 2D unitary ESPRIT method while Algorithm 1 can also handle the case $R = 5$. The mean $\mathrm{E_{RMS}}$ value over 100 trials as a function of SNR can be seen in figure 4. We observe that Algorithm 1 performs better than the 2D unitary ESPRIT method. The reason is that the former method takes the whole CA antenna array into account when estimating the DOAs while the latter method only exploits the horizontal (vertical) subarray when estimating the azimuth (elevation) angles.

Case $R = 3$: We now set $R = 3$ and, to demonstrate that the DOAs are not required to be distinct, we also set $\theta_1 = \theta_2$ and $\phi_2 = \phi_3$. The mean $\mathrm{E_{RMS}}$ value over 100 trials as a function of SNR can also be seen in figure 4. For the same reason as before, Algorithm 1 performs better than the 2D unitary ESPRIT approach. The repeated azimuth and elevation angles make the estimation problem more difficult, which explains the degradation in performance despite R has been reduced to three.

B. Asymmetric L-shaped Array (LA)

Consider the LA in subsection IV-B with $I = J = 5$, $K = 50$ and R varying.

Note that the 2D unitary ESPRIT method is not valid for the asymmetric LA. If $R < I$, we compare Algorithm 1 with an ESPRIT method [23] that first computes the azimuth angles $\{\theta_r\}$ and $\mathbf{S}$ via the horizontal ULA ($\mathbf{X}^{(1)} = \mathbf{B}^{(1)}\mathbf{S}^T$) and thereafter finds the elevation angles $\{\phi_r\}$ via the vertical ULA ($\mathbf{X}^{(2)} = \mathbf{B}^{(2)}\mathbf{S}^T$) with fixed $\mathbf{S}$.

In order to illustrate the use of Algorithm 1 for initializing an optimization-based method, we also refine the estimate obtained from Algorithm 1 by the Alternating Least Squares (ALS) method for coupled CPD [30]. In short, ALS alternates between conditional updates of the factor matrices of the coupled CPD. The overall method will be referred to as Algorithm 1-ALS. See [25] for other optimization-based methods.

Case $R = 4$: The mean $\mathrm{E_{RMS}}$ value over 100 trials as a function of SNR can be seen in figure 5. We observe that the basic ESPRIT approach suffers from the same problem as 2D unitary ESPRIT in the CA case. The difference in performance is quite significant. We also observe that below 30 dB, ALS slightly improves the estimate obtained by Algorithm 1.

Case $R = 5$: Since $\max(I, J) = R$ the individual horizontal and vertical subarrays of the LA do not yield unique DOAs. For this reason the previous ESPRIT algorithm cannot be used. The mean $\mathrm{E_{RMS}}$ value over 100 trials as a function of SNR can also be seen in figure 5. We observe that Algorithm 1-ALS yields more accurate results than Algorithm 1. Hence, Algorithm 1 can be used as a computationally cheap initialization procedure for a more costly optimization method, such as ALS.

VIII. Conclusion

Despite its importance, sensor array processing still presents important algebraic challenges. We have first made a link between several narrowband sensor array processing problems and the coupled CPD. This has led to a new uniqueness condition for sensor array processing problems involving multiple shift-invariance structures. To the the authors' knowledge, this is the first uniqueness condition reported in the signal processing literature that formally demonstrates that by simultaneously exploiting several subarray structures, improved identifiability results for the global array can be obtained.

Second, the coupled CPD also provides an algebraic SD method that under mild conditions is guaranteed to find the decomposition in the exact case. Since the complexity of the proposed SD method is basically linear in the array dimensions, it can be used as an initialization method for a more computationally demanding optimization based method. The low cost of the SD method also makes it interesting for shift-invariant structured large scale arrays composed of hundreds and possibly even thousands of sensors.

Numerical experiments showed that a higher accuracy of the DOA estimates can be obtained by simultaneously exploiting several subarray structures via coupled CPD compared to ordinary CPD based methods.

Third, we made a link between wideband VSA processing and the coupled CPD. This has led to an algebraic coherent data processing framework for wideband VSA, in which uniqueness conditions and algorithms can be derived from results for coupled CPD.

References

[1] J. Bosse, A. Ferréol, C. Germond, and P. Larzabel, "Passive geolocalization of radio transmitters: Algorithm and performance in narrowband context," *Signal Processing*, vol. 92, pp. 841–852, 2012.

[2] H.-W. Chen and J.-W. Zhao, "Wideband MVDR beamforming for acoustic sensor linear array," *IEE Proc.-Radar, Sonar Navig.*, vol. 151, no. 3, pp. 158–162, Jun. 2004.

[3] L. De 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.

[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] D. Donno, A. Nehorai, and U. Spagnolini, "Seismic velocity and polarization estimation for wavefield separation," *IEEE Trans. Signal Process.*, vol. 56, no. 10, Oct. 2008.

[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] T. Filik and E. Tuncer, "Uniform and nonuniform V-shaped isotropic planar arrays," in *Proc. of the 5th IEEE SAM Workshop*, july 21-23, Darmstadt, Germany, 2008.

[10] H. Gazzah and S. Marcos, "Cramer-Rao bounds for antenna array design," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 336–345, Jan. 2006.

[11] Y. Hua and T. K. Sarkar, "A note on the Cramer-Rao bound for 2-D direction finding based on 2-D array," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1215–1218, May 1991.

[12] Y. Hua, T. K. Sarkar, and D. D. Weiner, "An L-shaped array for estimating 2-D directions of wave arrival," *IEEE Trans. Antennas. Propag.*, vol. 39, no. 2, pp. 143–146, Feb. 1991.

[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] T. Koruda, N. Kikuma, and N. Inagaki, "DOA estimation and pairing method in 2D-ESPRIT using triangular antenna array," *Electronics and Communications in Japan*, vol. 86, no. 6, pp. 1505–1513, 2003.

[15] L.-H. Lim and P. Comon, "Blind multilinear identification," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1260–1280, Feb. 2014.

[16] K. Mahata and T. Söderström, "ESPRIT-like estimation of real-valued sinusoidal frequencies," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1161–1170, May 2004.

[17] A. Massoud, A. Osman, and A. Noureldin, "DOA estimation based on fourth order cumulant beamforming for nonunifom linear array of DIFAR sonobuoys," *Comput. Electr. Eng.*, vol. 38, no. 4, pp. 986–993, Jul. 2012.

[18] S. Miron, Y. Song, D. Brie, and K. T. Wong, "A multilinear approach of direction finding using a sensor-array with multiple scales of spatial invariance,," *Accepted for publication in IEEE Trans. Aerosp. Electron. Syst.*

[19] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2481–2491, Sep. 1994.

[20] ——, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 376–398, Feb. 1994.
[21] C. Paulus and J. I. Mars, "Vector-sensor array processing for polarization parameters and DOA estimation," *EURASIP Journal on Advances in Signal Processing*, 2010, article ID 850265.
[22] 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.
[23] R. Roy and T. Kailath, "Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, no. 7, pp. 984–995, Jul. 1989.
[24] 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.
[25] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured data fusion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 586–600, Jun. 2015.
[26] M. Sørensen and L. De Lathauwer, "Coupled tensor decompositions for applications in array signal processing," in *Proc. CAMSAP 2013*, Dec. 15-18, Saint Martin, Dutch Antilles, 2013.
[27] ——, "New uniqueness conditions for the canonical polyadic decomposition of third-order tensors," *Accepted for publication in SIAM J. Matrix Anal. Appl.*
[28] ——, "Blind signal separation via tensor decompositions with a Vandermonde factor: Canonical polyadic decomposition," *IEEE Trans. Signal Processing*, vol. 61, no. 22, pp. 5507–5519, Nov. 2013.
[29] ——, "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.
[30] 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.
[31] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Prentice Hall, 2005.
[32] A. Swindlehurst and T. Kailath, "Azimuth/elevation direction finding using regular array geometries," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 145–156, Jan. 1993.
[33] A. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, "Multiple invariance ESPRIT," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 868–881, Apr. 1992.
[34] A. Swindlehurst, P. Stoica, and M. Jansson, "Exploiting arrays with multiple invariances using MUSIC and MODE," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2511–2521, Nov. 2001.
[35] Z. Tian and H. L. Van Trees, "DOA estimation with hexagonal arrays," in *Proc. ICASSP*, May 12-15, Seattle, WA, USA, 1998.
[36] D. W. Tufts and P. D. Fiore, "Simple, effective estimation of frequency based on Prony's method," in *Proc. ICASSP 1996*, May 7-10, Atlanta, GA, USA, 1996.
[37] S. Zhao, T. Saluev, and D. L. Jones, "Underdetermined direction of arrival estimation using acoustic vector sensor," *Signal Processing*, vol. 100, no. 1, pp. 160–168, 2014.
[38] M. D. Zoltowski, M. Haardt, and C. P. Matthews, "Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 316–328, Feb. 1996.

Acknowledgment

Research supported by: (1) Research Council KU Leuven: CoE EF/05/006 Optimization in Engineering (OPTEC), C1 project c16/15/059-nD (2) F.W.O.: project G.0830.14N, G.0881.14N, (3) the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012–2017), (4) 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.

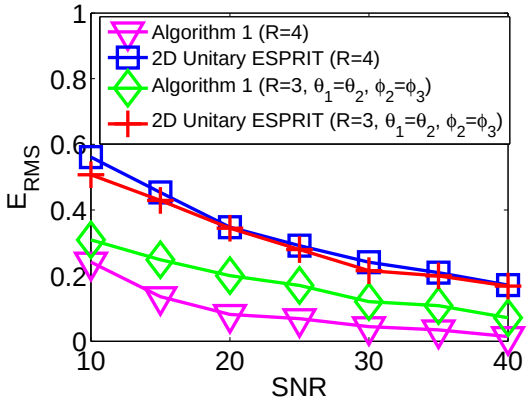


Fig. 4. Mean E_{RMS} values over 100 trials while SNR is varying from 10 to 40 dB, cross-shaped array.

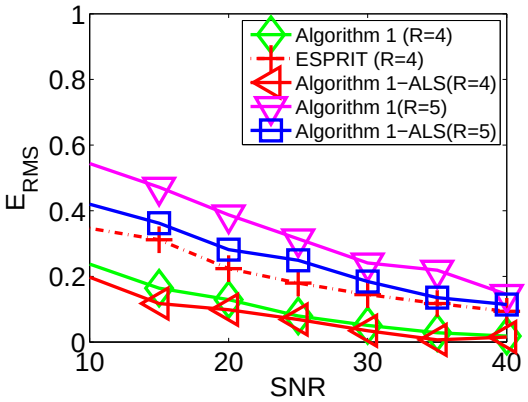


Fig. 5. Mean E_{RMS} values over 100 trials while SNR is varying from 10 to 40 dB, L-shaped array.



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 the KU Leuven, Belgium. His research interests include applied linear algebra, tensor decompositions, and tensor-based (array) signal processing.



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.