Multidimensional Harmonic Retrieval via Coupled Canonical Polyadic Decomposition — Part II: Algorithm and Multirate Sampling

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

***Abstract*—In Part I, we have presented a link between Multidimensional Harmonic Retrieval (MHR) and the recently proposed coupled Canonical Polyadic Decomposition (CPD), which implies new uniqueness conditions for MHR that are more relaxed than the existing results based on a Vandermonde constrained CPD. In Part II, we explain that the coupled CPD also provides a computational framework for MHR. In particular, we present an algebraic method for MHR based on simultaneous matrix diagonalization that is guaranteed to find the exact solution in the noiseless case, under conditions discussed in Part I. Since the simultaneous matrix diagonalization method reduces the MHR problem into an eigenvalue problem, the proposed algorithm can be interpreted as a MHR generalization of the classical ESPRIT method for one-dimensional harmonic retrieval. We also demonstrate that the presented coupled CPD framework for MHR can algebraically support multirate sampling. We develop an efficient implementation which has about the same computational complexity for single-rate and multirate sampling. Numerical experiments demonstrate that by simultaneously exploiting the harmonic structure in all dimensions and making use of multirate sampling, the coupled CPD framework for MHR can lead to an improved performance compared to the conventional Vandermonde constrained CPD approaches.**

***Index Terms*—coupled canonical polyadic decomposition, simultaneous matrix diagonalization, multidimensional harmonic retrieval, multirate sampling, direction of arrival estimation.**

I. Introduction

In the signal processing literature, the uniqueness properties of Multidimensional Harmonic Retrieval (MHR) have mainly been studied using arguments that implicitly or explicitly relate to properties of a Vandermonde constrained Canonical Polyadic Decomposition (VDM-CPD), see [21] and references therein. Based on this connection several algorithms have been developed (e.g. [4], [7], [9], [10], [5]) that solve the MHR problem via a Generalized EigenValue Decomposition (GEVD).

In the companion paper [23] we presented a link between MHR and the recently introduced coupled Canonical Polyadic Decomposition (CPD) modeling framework [22], [25], which led to improved identifiability results. In this paper we explain that the coupled CPD model also provides an algebraic framework from which algorithms for MHR can be derived. We develop an algorithm based on coupled CPD that, similar to ESPRIT [16] for one-dimensional (1D) Harmonic Retrieval (HR), reduces the MHR problem into a set of decoupled single-tone MHR problems via a GEVD. A closed-form solution to the single-tone MHR problem will be provided. In the inexact case, the algorithm can be used to initialize an optimization-based method for MHR (e.g. [2], [20]).

We argue that an additional reason to consider coupled CPD as opposed to the conventional VDM-CPD is that the former approach offers a more flexible framework. In particular, we show that coupled CPD can support multirate sampling. This makes it interesting for large-scale MHR problems where sub-Nyquist sampling is often a desired feature in order to reduce hardware, data acquisition or computational costs.

The paper is organized as follows. The rest of the introduction will present the notation and the tensor decompositions used throughout the paper. Section II briefly reviews the connections between MHR, VDM-CPD and coupled CPD. From the coupled CPD interpretation of MHR, an algebraic ESPRIT-type algorithm is derived in Section III. In Section IV we extend the algorithm to multirate sampling. 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 vector of $\mathbf{A}$ is denoted by $\mathbf{a}_r$. The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri-Rao product:

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

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.

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.

in which $(\mathbf{A})_{mn} = a_{mn}$. The outer product of N vectors $\mathbf{a}^{(n)} \in \mathbb{C}^{I_n}$ is denoted by $\mathbf{a}^{(1)} \otimes \cdots \otimes \mathbf{a}^{(N)} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, such that $(\mathbf{a}^{(1)} \otimes \cdots \otimes \mathbf{a}^{(N)})_{i_1,i_2,\ldots,i_N} = a^{(1)}_{i_1} a^{(2)}_{i_2} \cdots a^{(N)}_{i_N}$.

The conjugate, transpose, conjugate-transpose, Frobenius norm, Moore-Penrose pseudo-inverse, range and kernel of the matrix $\mathbf{A}$ are denoted by $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\|\mathbf{A}\|_F$, $\mathbf{A}^\dagger$, range($\mathbf{A}$) and ker($\mathbf{A}$) respectively. The real and imaginary parts of a number a are denoted by $\mathrm{Re}\{a\}$ and $\mathrm{Im}\{a\}$, respectively. From the context it should be clear when i denotes the imaginary unit number, i.e., $i = \sqrt{-1}$. Kronecker's delta function, denoted by δ_{ij}, is equal to one when $i = j$ and zero otherwise. Matlab index notation will occasionally be used for submatrices of a given matrix. We write $\mathbf{A}(1{:}k,:)$ to denote the submatrix of $\mathbf{A}$ consisting of the rows from 1 to k of $\mathbf{A}$. Let $\mathbf{A} \in \mathbb{C}^{I \times J}$, then $\underline{\mathbf{A}} = \mathbf{A}(1 : I-1,:) \in \mathbb{C}^{(I-1)\times J}$ and $\overline{\mathbf{A}} = \mathbf{A}(2 : I,:) \in \mathbb{C}^{(I-1)\times J}$, i.e., $\underline{\mathbf{A}}$ and $\overline{\mathbf{A}}$ are obtained by deleting the bottom and top row of $\mathbf{A}$, respectively. $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$ denotes the diagonal matrix holding row k of $\mathbf{A} \in \mathbb{C}^{I\times J}$ on its diagonal.

The k-th compound matrix of $\mathbf{A} \in \mathbb{C}^{I\times R}$ is denoted by $C_k(\mathbf{A}) \in \mathbb{C}^{C^k_I \times C^k_R}$, where $C^k_m = \frac{m!}{k!(m-k)!}$. It is the matrix containing the determinants of all $k \times k$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order. For more elementary notation, we refer to [23].

Tensor decompositions: Let us briefly present the tensor decompositions that are used in this paper. See the references in [23] for further details.

1) Polyadic Decomposition (PD) and CPD: A PD is a decomposition of a tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ into rank-1 terms:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r\,, \tag{1}$$

with factor matrices $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R] \in \mathbb{C}^{I\times R}$, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R] \in \mathbb{C}^{J\times R}$ and $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{K\times R}$. The rank of a tensor $\mathcal{X}$ is equal to the minimal number of rank-1 tensors $\mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$ that yield $\mathcal{X}$ in a linear combination. Assume that the rank of $\mathcal{X}$ is R, then (1) is the CPD of $\mathcal{X}$.

2) Coupled PD and coupled CPD: More generally, we say that a collection of tensors $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$, $n \in \{1, \ldots, N\}$, admits an R-term coupled PD if each tensor $\mathcal{X}^{(n)}$ can be written as:

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}^{(n)}_r \otimes \mathbf{b}^{(n)}_r \otimes \mathbf{c}_r\,, \quad n \in \{1, \ldots, N\}, \tag{2}$$

with factor matrices $\mathbf{A}^{(n)} = [\mathbf{a}^{(n)}_1, \ldots, \mathbf{a}^{(n)}_R] \in \mathbb{C}^{I_n\times R}$, $\mathbf{B}^{(n)} = [\mathbf{b}^{(n)}_1, \ldots, \mathbf{b}^{(n)}_R] \in \mathbb{C}^{J_n\times R}$ and $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{K\times R}$. We define the coupled rank of $\{\mathcal{X}^{(n)}\}$ as the minimal number of coupled rank-1 tensors $\mathbf{a}^{(n)}_r \otimes \mathbf{b}^{(n)}_r \otimes \mathbf{c}_r$ that yield $\{\mathcal{X}^{(n)}\}$ in a linear combination. Assume that the coupled rank of $\{\mathcal{X}^{(n)}\}$ is R, then (2) is the coupled CPD of $\{\mathcal{X}^{(n)}\}$.

II. Review of links between multidimensional harmonic retrieval and (coupled) CPD

The connections between MHR, CPD and coupled CPD will be reviewed in this section. More details can be found in the companion paper [23]. In order to understand why coupled CPD is a natural framework for MHR, we will first review the link between 1D HR and CPD.

A. 1D HR and CPD

1) CPD interpretation of 1D HR: Consider the 1D HR factorization

$$\mathbf{X} = \mathbf{A}\mathbf{C}^T \in \mathbb{C}^{I\times M}, \tag{3}$$

where $\mathbf{A} \in \mathbb{C}^{I\times R}$ is a Vandermonde matrix with columns $\mathbf{a}_r = [1, z_r, z_r^2, \ldots, z_r^{I-1}]^T$ and where $\mathbf{C} \in \mathbb{C}^{M\times R}$ is an unconstrained matrix that has full column rank. Let us stack $\underline{\mathbf{X}}$ and $\overline{\mathbf{X}}$ as slices of $\mathcal{Y} \in \mathbb{C}^{2\times J\times M}$. Because of the shift-invariance $\overline{\mathbf{A}} = \underline{\mathbf{A}} D_2(\mathbf{A})$, we have

$$\mathbf{Y} = \begin{bmatrix} \underline{\mathbf{X}} \\ \overline{\mathbf{X}} \end{bmatrix} = \begin{bmatrix} \mathbf{B}\mathbf{C}^T \\ \mathbf{B}D_2(\mathbf{A}^{(2)})\mathbf{C}^T \end{bmatrix} = (\mathbf{A}^{(2)} \odot \mathbf{B})\mathbf{C}^T \in \mathbb{C}^{2J\times M}, \tag{4}$$

where $\mathbf{A}^{(2)} = \begin{bmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_R \end{bmatrix} \in \mathbb{C}^{2\times R}$ and $\mathbf{B} = \underline{\mathbf{A}} \in \mathbb{C}^{J\times R}$ with $J = I - 1$. Equation (4) can be seen as the matrix representation of the Vandermonde constrained CPD of $\mathcal{Y}$:

$$\mathcal{Y} = \sum_{r=1}^{R} \mathbf{a}^{(2)}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r \in \mathbb{C}^{2\times J\times M}. \tag{5}$$

It was explained in [23] that the Vandermonde structure in (5) can be relaxed without affecting the uniqueness of the decomposition. In other words, the 1D HR factorization of $\mathbf{X}$ is unique if and only if the unconstrained CPD of $\mathcal{Y}$ is unique. The latter decomposition is known to be unique if $\mathbf{B}$ and $\mathbf{C}$ have full column rank and if $z_r \neq z_s$, $\forall r \neq s$. It is known that under this condition, the CPD of $\mathcal{Y}$ can be determined via the GEVD of the matrix pencil $(\overline{\mathbf{X}}, \underline{\mathbf{X}})$ (e.g., [17], [16], [6], [8]). This is in fact precisely what happens in ESPRIT, see the next subsection.

2) ESPRIT: A particular method for computing the 1D HR factorization of $\mathbf{X}$ via the matrix pencil $(\overline{\mathbf{X}}, \underline{\mathbf{X}})$ in (4) is ESPRIT [16]. It was originally formulated in terms of the correlation matrix $\mathbf{X}\mathbf{X}^H = \mathbf{A}\mathbf{S}^T\mathbf{S}^*\mathbf{A}^H$, but it can also be formulated in terms of an SVD. Briefly, let the columns of $\mathbf{U} \in \mathbb{C}^{I\times R}$ constitute a basis for range($\mathbf{X}$), numerically obtained from the SVD of $\mathbf{X}$. Since $\mathbf{B}$ and $\mathbf{C}$ have full column rank, there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that $\mathbf{U} = \mathbf{A}\,\mathbf{F}^T$. The shift-invariance property $\underline{\mathbf{A}} = \overline{\mathbf{A}} D_2(\mathbf{A}^{(2)})$ now results in the EVD

$$\underline{\mathbf{U}}^\dagger \overline{\mathbf{U}} = \mathbf{F}^{-1} D_2(\mathbf{A}^{(2)}) \mathbf{F}. \tag{6}$$

As the generators $\{z_r\}_{r=1}^R$ form the second row of $\mathbf{A}^{(2)}$, they can be found as the eigenvalues of $\underline{\mathbf{U}}^\dagger\overline{\mathbf{U}}$, i.e., the generalized eigenvalues of the pencil $(\overline{\mathbf{U}}, \underline{\mathbf{U}})$. We have now found the Vandermonde column vectors of $\mathbf{A}$ that form a basis for range($\mathbf{X}$). Finally, the linear coefficients of the columns of $\mathbf{X}$ in this Vandermonde basis can be determined as $\mathbf{C} = (\mathbf{A}^\dagger\mathbf{X})^T$.

B. MHR and Vandermonde constrained CPD

1) VDM-CPD interpretation of MHR: Consider now an N-dimensional HR problem that involves R complex exponentials. The generator of the r-th exponential in the n-th mode is denoted by $z_{r,n}$, $1 \leqslant r \leqslant R$, $1 \leqslant n \leqslant N$. The data tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times M}$ admits a constrained PD

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \otimes \cdots \otimes \mathbf{a}_r^{(N)} \otimes \mathbf{c}_r \,, \tag{7}$$

with Vandermonde factor matrices $\mathbf{A}^{(n)} = [\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}] \in \mathbb{C}^{I_n \times R}$ with $\mathbf{a}_r^{(n)} = [1, z_{r,n}, z_{r,n}^2, \ldots, z_{r,n}^{I_n-1}]^T$ and unconstrained $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{M \times R}$, where M is the number of snapshots. The PD of $\mathcal{X}$ in (7) has the following matrix representation

$$\mathbf{X} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{C}^T \in \mathbb{C}^{(\prod_{n=1}^{N} I_n) \times M}. \tag{8}$$

Note that the Vandermonde constraints in (7) can increase the number of rank-1 terms that are required to fit $\mathcal{X}$, compared to unconstrained CPD. The minimal number of Vandermonde constrained rank-1 terms $\mathbf{a}_r^{(1)} \otimes \cdots \otimes \mathbf{a}_r^{(N)} \otimes \mathbf{c}_r$ that yield $\mathcal{X}$ will be denoted by $r_{\text{MHR}}(\mathcal{X})$.

2) Exploiting harmonic structure in one mode: If the Vandermonde structure of only one of the factor matrices, say $\mathbf{A}^{(n)}$, is taken into account, then (7) corresponds to a VDM-CPD problem that can be reduced to a CPD as in Subsection II-A. In detail, using the row-selection matrices

$$\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{p=1}^{n-1} I_p} \otimes \underline{\mathbf{I}}_{I_n} \otimes \mathbf{I}_{\prod_{q=n+1}^{N} I_q} \,, \tag{9}$$

$$\overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{p=1}^{n-1} I_p} \otimes \bar{\mathbf{I}}_{I_n} \otimes \mathbf{I}_{\prod_{q=n+1}^{N} I_q} \,, \tag{10}$$

we build the tensor $\mathcal{Y}^{(n)} \in \mathbb{C}^{2 \times (\times_{p=1}^{n-1} I_p) \times (I_n - 1) \times (\times_{q=n+1}^{N} I_q) \times M}$:

$$\mathcal{Y}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(2,n)} \otimes \mathbf{b}_r^{(n)} \otimes \mathbf{c}_r \,, \tag{11}$$

with matrix representation

$$\mathbf{Y}^{(n)} = \begin{bmatrix} \underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} \mathbf{X} \\ \overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} \mathbf{X} \end{bmatrix} = (\mathbf{A}^{(2,n)} \odot \mathbf{B}^{(n)})\mathbf{C}^T \in \mathbb{C}^{2J_n \times M}, \tag{12}$$

where $J_n = (\prod_{p=1, p \neq n}^{N} I_p)(I_n - 1)$ and

$$\mathbf{A}^{(2,n)} = \begin{bmatrix} 1 & \cdots & 1 \\ z_{1,n} & \cdots & z_{R,n} \end{bmatrix} \in \mathbb{C}^{2 \times R}, \tag{13}$$

$$\mathbf{B}^{(n)} = \bigodot_{p=1}^{n-1} \mathbf{A}^{(p)} \odot \underline{\mathbf{A}}^{(n)} \odot \bigodot_{q=n+1}^{N} \mathbf{A}^{(q)}. \tag{14}$$

Ignoring the structure of $\mathbf{B}^{(n)}$, (11) can be interpreted as an unconstrained two-slice CPD. Hence, if $\mathbf{B}^{(n)}$ and $\mathbf{C}$ have full column rank and if $z_{r,n} \neq z_{s,n}$, $\forall r \neq s$, then the generators $\{z_{r,n}\}_{r=1}^{R}$ can be found via a (G)EVD. In a similar way as in ESPRIT for 1D HR, the remaining unknown factor matrices $\{\mathbf{A}^{(n)}\}$ and $\mathbf{C}$ in (8) can now be deduced. Details and variants of this procedure can be found in [4], [7], [9], [10], [27], [21], [5] and references therein.

C. MHR and coupled CPD

1) Coupled CPD interpretation of MHR: In contrast to the existing algebraic VDM-CPD approaches to MHR, coupled CPD simultaneously considers all tensors in the set $\{\mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(N)}\}$. Overall, we obtain

$$\begin{bmatrix} \mathbf{Y}^{(1)} \\ \vdots \\ \mathbf{Y}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(2,1)} \odot \mathbf{B}^{(1)} \\ \vdots \\ \mathbf{A}^{(2,N)} \odot \mathbf{B}^{(N)} \end{bmatrix} \mathbf{C}^T, \tag{15}$$

where $\mathbf{Y}^{(n)}$ is given by (12). The structure of (15) is such that $\mathcal{X}$ must have the N-dimensional harmonic structure (7), even when the structure of $\{\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(N)}\}$ is ignored [23]. In fact, we make the assumption that $\mathbf{A}^{(2,n)}$, $n \in \{1, \ldots, N\}$ do not have zero entries, i.e., we assume that $\mathbf{A}^{(2,n)}$ have a (trivial) Vandermonde structure. Hence, in principle any numerical method for coupled CPD can be used to compute the MHR decomposition (7) via (15).

2) Exploiting several harmonic structures: Both necessary and sufficient MHR uniqueness conditions have been obtained in [23]. In this paper we develop a linear algebra based method. Similar to ESPRIT, it finds the pure harmonics from range($\mathbf{X}$), and it does so via (G)EVD. Like ESPRIT, it is guaranteed to find the solution in the noise-free case under some algebraic assumptions. It relies on the sufficient and "almost necessary" MHR uniqueness condition stated in Theorem II.1, which jointly takes the Vandermonde structures of $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}$ into account. Consequently, it works under more relaxed conditions than the methods that essentially rely on the harmonic structure in one mode. Theorem II.1 makes use of the matrix

$$\mathbf{G}^{(N)} = \begin{bmatrix} C_2\left(\mathbf{A}^{(2,1)}\right) \odot C_2\left(\mathbf{B}^{(1)}\right) \\ \vdots \\ C_2\left(\mathbf{A}^{(2,N)}\right) \odot C_2\left(\mathbf{B}^{(N)}\right) \end{bmatrix}. \tag{16}$$

Theorem II.1. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times M}$ in (7) where the factor matrices $\{\boldsymbol{A}^{(n)}\}$ are Vandermonde. If*

$$\begin{cases} \boldsymbol{C} \text{ in (8) has full column rank,} \\ \boldsymbol{G}^{(N)} \text{ in (16) has full column rank,} \end{cases} \tag{17}$$

then $r_{MHR}(\mathcal{X}) = R$ and the VDM-CPD of $\mathcal{X}$ is unique.

III. Coupled CPD algorithms for MHR

In this section we use coupled CPD to obtain ESPRIT-type algorithms for MHR that work under the assumptions in Theorem II.1.

A. Find a basis for range($\boldsymbol{A}^{(1)} \odot \cdots \odot \boldsymbol{A}^{(N)}$)

Similar to ESPRIT [14], [16] we will find the individual multidimensional harmonics from range($\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)}$). We first explain that under condition (17), the latter is readily available as range($\mathbf{X}$). First we note that $\mathbf{G}^{(N)}$ having full column rank implies that $\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)}$

has full column rank. Indeed, if $\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)}$ does not have full column rank, then there exists a vector $\mathbf{y} \in \ker(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})$ and a vector $\mathbf{x} \in \mathbb{C}^M$ such that we obtain an alternative decomposition $\mathbf{X} = \left(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)}\right)\left(\mathbf{C} + \mathbf{x}\mathbf{y}^T\right)^T$, contradicting Theorem II.1. Since $\mathbf{C}$ is assumed to have full column rank, we have that $\text{range}(\mathbf{X}) = \text{range}(\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})$. Numerically, we obtain $\text{range}(\mathbf{X})$ from the compact SVD $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, where $\mathbf{U} \in \mathbb{C}^{(\prod_{n=1}^N I_n) \times R}$ and $\mathbf{V} \in \mathbb{C}^{M \times R}$ are columnwise orthonormal and $\mathbf{\Sigma} \in \mathbb{C}^{R \times R}$ is diagonal. We know that $\text{range}(\mathbf{X}) = \text{range}(\mathbf{U}\mathbf{\Sigma})$ and that there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\mathbf{\Sigma} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{F}^T. \tag{18}$$

Compared to (8), (18) involves the smaller $(R \times R)$ matrix $\mathbf{F}$. Note that this dimensionality reduction step is similar to the one in classical ESPRIT for 1D HR, see Subsection II-A. Note also that if R is unknown, then it can be estimated as the number of significant singular values of $\mathbf{X}$.

B. From MHR to coupled CPD

From (18) we build the tensors $\mathcal{U}_{\text{red}}^{(n)} \in \mathbb{C}^{2 \times (\times_{p=1}^{n-1} I_p) \times (I_n - 1) \times (\times_{q=n+1}^{N} I_q) \times R}$, $n \in \{1, \ldots, N\}$, each admitting a factorization

$$\mathbf{U}_{\text{red}}^{(n)} := \begin{bmatrix} \mathbf{U}^{(1,n)} \\ \mathbf{U}^{(2,n)} \end{bmatrix} = (\mathbf{A}^{(2,n)} \odot \mathbf{B}^{(n)})\mathbf{F}^T, \tag{19}$$

where 'red' stands for reduced, and

$$\mathbf{U}^{(1,n)} = \underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} \mathbf{U}\mathbf{\Sigma} \in \mathbb{C}^{M_n(I_n-1) \times R}, \tag{20}$$

$$\mathbf{U}^{(2,n)} = \overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} \mathbf{U}\mathbf{\Sigma} \in \mathbb{C}^{M_n(I_n-1) \times R}, \tag{21}$$

in which $M_n = (\prod_{m=1, m \neq n}^{N} I_m)$ such that $J_n = M_n(I_n - 1)$. Note that (19), in combination with the way $\mathbf{U}_{\text{red}}^{(n)}$ is constructed, implements the harmonic structure in the nth mode. Overall, we obtain the compressed variant of (15):

$$\begin{bmatrix} \mathbf{U}_{\text{red}}^{(1)} \\ \vdots \\ \mathbf{U}_{\text{red}}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(2,1)} \odot \mathbf{B}^{(1)} \\ \vdots \\ \mathbf{A}^{(2,N)} \odot \mathbf{B}^{(N)} \end{bmatrix} \mathbf{F}^T, \tag{22}$$

in which each submatrix implements the harmonic structure in the corresponding mode. In the next subsection we adapt the algebraic SD method in [1] to our MHR problem. In particular, a more efficient implementation will be developed that takes the harmonic structure of the MHR problem into account.

C. From coupled CPD to SD

From (22) we build R symmetric matrices $\{\mathbf{M}^{(r)}\}$, admitting the factorizations

$$\mathbf{M}^{(r)} = \mathbf{G}\mathbf{\Lambda}^{(r)}\mathbf{G}^T, \quad r \in \{1, \ldots, R\}, \tag{23}$$

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. Note that (23) can be seen as the slice-wise representation of the CPD of the partially symmetric tensor $\mathcal{M} = \sum_{r=1}^{R} \mathbf{g}_r \otimes \mathbf{g}_r \otimes \boldsymbol{\zeta}_r$ with factor matrices $\mathbf{G}$ and $\mathbf{\Delta} = [\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_R]$ in which $\boldsymbol{\zeta}_r = [\lambda_{11}^{(r)}, \ldots, \lambda_{RR}^{(r)}]^T$. It can verified that both $\mathbf{G}$ and $\mathbf{\Gamma}$ have full column rank [1]. This implies that in the exact case, it suffices to exploit the structure of only two of the matrix slices, denoted by $\mathbf{M}^{(r_1)}$ and $\mathbf{M}^{(r_2)}$. More precisely, as in the ESPRIT method, the problem can be solved by means of a GEVD of the matrix pencil $(\mathbf{M}^{(r_1)}, \mathbf{M}^{(r_2)})$ in which the columns of $\mathbf{G}^{-T}$ are the generalized eigenvectors. In the inexact case, the robustness can be increased by simultaneously exploiting all R equations in (23), using for instance optimization-based methods (e.g., [28], [20]). The CPD of $\mathcal{M}$ can equivalently be seen as the simultaneous diagonalization of its matrix slices, which is why (23) is referred to as SD.

Following the steps in [25] we will now explain how to construct the matrices $\{\mathbf{M}^{(r)}\}$ in (23). Thereafter, we explain that the complexity of the construction can be reduced by exploiting the MHR structure.

Construction of matrices $\{\mathbf{M}^{(r)}\}$: 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 [1]:

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

Denote $\Phi_{r,s}^{(n)} = \Phi^{(n)}([\mathbf{u}_r^{(1,n)}, \mathbf{u}_r^{(2,n)}], [\mathbf{u}_s^{(1,n)}, \mathbf{u}_s^{(2,n)}])$, in which $\mathbf{u}_t^{(1,n)}$ and $\mathbf{u}_t^{(2,n)}$ correspond to the t-th column of $\mathbf{U}^{(1,n)}$ and $\mathbf{U}^{(2,n)}$, respectively. Define $\mathbf{\Psi}^{(n)} = [\mathbf{\Psi}_1^{(n)}, 2 \cdot \mathbf{\Psi}_2^{(n)}] \in \mathbb{C}^{4J_n^2 \times C_{R+1}^2}$ with $\mathbf{\Psi}_1^{(n)} = [\Phi_{1,1}^{(n)}, \Phi_{2,2}^{(n)}, \ldots, \Phi_{R,R}^{(n)}] \in \mathbb{C}^{4J_n^2 \times R}$ and $\mathbf{\Psi}_2^{(n)} = [\Phi_{1,2}^{(n)}, \Phi_{1,3}^{(n)}, \Phi_{2,3}^{(n)}, \ldots, \Phi_{R-1,R}^{(n)}] \in \mathbb{C}^{4J_n^2 \times C_R^2}$. It can be verified that if condition (17) in Theorem II.1 is satisfied, then

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{\Psi}^{(1)} \\ \vdots \\ \mathbf{\Psi}^{(N)} \end{bmatrix} \in \mathbb{C}^{4(\sum_{n=1}^{N} J_n^2) \times C_{R+1}^2} \tag{24}$$

has an R-dimensional kernel [25]. Furthermore, the kernel vectors can be associated with the matrices $\{\mathbf{M}^{(r)}\}$. Namely, $\mathbf{M}^{(r)}$ satisfies

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

where $\mathbf{m}^{(r)} = [m_{1,1}^{(r)}, m_{2,2}^{(r)}, \ldots, m_{R,R}^{(r)}, m_{1,2}^{(r)}, m_{1,3}^{(r)}, \ldots, m_{R-1,R}^{(r)}]^T$. Since $\ker(\mathbf{\Psi}) = \ker(\mathbf{\Psi}^H\mathbf{\Psi})$, a basis for the kernel of (24) can be found from the smaller $(C_{R+1}^2 \times C_{R+1}^2)$ positive semi-definite matrix

$$\mathbf{\Psi}^H\mathbf{\Psi} = \sum_{n=1}^{N} \mathbf{\Psi}^{(n)H}\mathbf{\Psi}^{(n)}. \tag{25}$$

Numerically, we can take $\{\mathbf{m}^{(r)}\}$ equal to the eigenvectors associated with the R smallest eigenvalues of $\mathbf{\Psi}^H\mathbf{\Psi}$. In the computation of the smaller matrix (25), the structure of $\{\mathbf{\Psi}^{(n)H}\mathbf{\Psi}^{(n)}\}$ can be exploited. Concretely, define the set $S_R = \{(s_1, s_2) \; : \; 1 \leqslant s_1 < s_2 \leqslant R\}$ 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., $f_R((1,3)) = 3$. Define also $\mathbf{Y}^{(1,n)} \in \mathbb{C}^{R\times R}$, $\mathbf{Y}^{(2,n)} \in \mathbb{C}^{R\times C_R^2}$ and $\mathbf{Y}^{(3,n)} \in \mathbb{C}^{C_R^2\times C_R^2}$ as follows

$$(\mathbf{Y}^{(1,n)})_{r,s} = 8(w_{rs}^{(n)} - \mathbf{b}_{rs}^{(n)H}\mathbf{b}_{sr}^{(n)}), \tag{26}$$

$$(\mathbf{Y}^{(2,n)})_{r,f_R(s_1,s_2)} = 16(w_{rs_1}^{(n)}w_{rs_2}^{(n)} - \mathbf{b}_{rs_1}^{(n)H}\mathbf{b}_{rs_2}^{(n)}), \tag{27}$$

$$(\mathbf{Y}^{(3,n)})_{f_R(r_1,r_2),f_R(s_1,s_2)} = 4(w_{r_1s_1}^{(n)}w_{r_2s_2}^{(n)} + w_{r_2s_1}^{(n)}w_{r_1s_2}^{(n)} - \mathbf{b}_{r_1s_1}^{(n)H}\mathbf{b}_{r_2s_2}^{(n)} - \mathbf{b}_{r_2s_1}^{(n)H}\mathbf{b}_{r_1s_2}^{(n)}), \tag{28}$$

where

$$w_{rs}^{(n)} := w_{rs}^{(1,n)} + w_{rs}^{(2,n)},$$

$$w_{rs}^{(1,n)} = \mathbf{u}_r^{(1,n)H}\mathbf{u}_s^{(1,n)}, \tag{29}$$

$$w_{rs}^{(2,n)} = \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(2,n)}, \tag{30}$$

$$\mathbf{b}_{rs}^{(n)} := [w_{rs}^{(1,n)}, z_{rs}^{(n)}, z_{sr}^{(n)*}, w_{rs}^{(2,n)}]^T,$$

$$z_{rs}^{(n)} = \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(1,n)}. \tag{31}$$

It can be verified that $\mathbf{\Psi}^{(n)H}\mathbf{\Psi}^{(n)} = \left[\begin{smallmatrix} \mathbf{Y}^{(1,n)} & 2\cdot\mathbf{Y}^{(2,n)} \\ 2\cdot\mathbf{Y}^{(2,n)H} & 4\cdot\mathbf{Y}^{(3,n)} \end{smallmatrix}\right]$, so that (25) can be expressed in terms of (26)–(28):

$$\mathbf{\Psi}^H\mathbf{\Psi} = \sum_{n=1}^{N}\mathbf{\Psi}^{(n)H}\mathbf{\Psi}^{(n)} = \sum_{n=1}^{N}\left[\begin{array}{cc} \mathbf{Y}^{(1,n)} & 2\cdot\mathbf{Y}^{(2,n)} \\ 2\cdot\mathbf{Y}^{(2,n)H} & 4\cdot\mathbf{Y}^{(3,n)} \end{array}\right]. \tag{32}$$

Note that the matrices $\mathbf{Y}^{(1,n)}$ and $\mathbf{Y}^{(3,n)}$ are Hermitian, implying that it suffices to compute their upper triangular part. Note also that the vector products $\mathbf{b}_{r_1s_1}^{(n)H}\mathbf{b}_{r_2s_2}^{(n)}$ that appear in (26)–(28) can be computed prior to the construction of $\mathbf{Y}^{(1,n)}$, $\mathbf{Y}^{(2,n)}$ and $\mathbf{Y}^{(3,n)}$. In addition, due to the MHR structure, the involved scalars $\{w_{rs}^{(1,n)}, w_{rs}^{(2,n)}\}$ can be computed efficiently, as explained next.

Exploiting MHR structure: By construction the matrix slices $\{\mathbf{U}^{(1,n)}\}$ of the tensors $\{\mathcal{U}_{\text{red}}^{(n)}\}$ are largely overlapping. (Similarly for $\{\mathbf{U}^{(2,n)}\}$.) More concretely, observe that for every $n \in \{1, \ldots, N\}$, the vector $\mathbf{u}_r^{(1,n)} = \underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{u}_r$ consists of all except M_n entries of $\mathbf{u}_r$. To put it differently, the vectors in the set $\{\mathbf{u}_r^{(1,1)}, \ldots, \mathbf{u}_r^{(1,N)}\}$ are almost identical. (Similarly for the vectors in the set $\{\mathbf{u}_r^{(2,1)}, \ldots, \mathbf{u}_r^{(2,N)}\}$, which all are related via $\mathbf{u}_r^{(2,n)} = \overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{u}_r$.) Moreover, $\mathbf{U}^{(1,n)}$ and $\mathbf{U}^{(2,n)}$ in (19) are taken from the matrix $\mathbf{U\Sigma}$, the columns of which are mutually orthogonal, see (20)–(21). These properties enable a faster computation of $w_{rs}^{(1,n)}$ and $w_{rs}^{(2,n)}$ in (29)–(30). We work as follows. Consider the following row-selection matrices complementary to $\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}$ and $\overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}$:

$$\underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{p=1}^{n-1}I_p} \otimes \mathbf{I}_{I_n}(I_n, :) \otimes \mathbf{I}_{\prod_{q=n+1}^{N}I_q} \in \mathbb{C}^{M_n\times(\prod_{n=1}^{N}I_n)},$$

$$\overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{p=1}^{n-1}I_p} \otimes \mathbf{I}_{I_n}(1, :) \otimes \mathbf{I}_{\prod_{q=n+1}^{N}I_q} \in \mathbb{C}^{M_n\times(\prod_{n=1}^{N}I_n)}.$$

The selection is complementary in the sense that $\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} + \underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)} = \overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} + \overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{n=1}^{N}I_n}$, i.e., $\underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}$ and $\overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}$ select the M_n rows not selected by $\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}$ and $\overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}$, respectively. This property together with relations (20)–(21) allows us to express (29)–(30) as follows:

$$\begin{aligned} w_{rs}^{(1,n)} &= \mathbf{u}_r^{(1,n)H}\mathbf{u}_s^{(1,n)} = \sigma_r\sigma_s\mathbf{u}_r^H\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{u}_s \\ &= \sigma_r\sigma_s(\delta_{rs} - \mathbf{u}_r^H\underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{u}_s), \end{aligned} \tag{33}$$

$$\begin{aligned} w_{rs}^{(2,n)} &= \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(2,n)} = \sigma_r\sigma_s\mathbf{u}_r^H\overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{u}_s \\ &= \sigma_r\sigma_s(\delta_{rs} - \mathbf{u}_r^H\overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{u}_s), \end{aligned} \tag{34}$$

where the orthogonality $\mathbf{u}_r^H\mathbf{u}_s = \delta_{rs}$ has been taken into account and where σ_r denotes the r-th singular value of $\mathbf{X}$.

Denoting the Hermitian matrices

$$\widetilde{\mathbf{W}}^{(1,n)} = (\underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{U\Sigma})^H \cdot (\underline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{U\Sigma}) \in \mathbb{C}^{R\times R}, \tag{35}$$

$$\widetilde{\mathbf{W}}^{(2,n)} = (\overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{U\Sigma})^H \cdot (\overline{\mathbf{T}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{U\Sigma}) \in \mathbb{C}^{R\times R}, \tag{36}$$

the products in (29)–(31) can be computed as the entries of

$$\mathbf{W}^{(1,n)} = \mathbf{U}^{(1,n)H}\mathbf{U}^{(1,n)} = \mathbf{\Sigma}^2 - \widetilde{\mathbf{W}}^{(1,n)} \in \mathbb{C}^{R\times R}, \tag{37}$$

$$\mathbf{W}^{(2,n)} = \mathbf{U}^{(2,n)H}\mathbf{U}^{(2,n)} = \mathbf{\Sigma}^2 - \widetilde{\mathbf{W}}^{(2,n)} \in \mathbb{C}^{R\times R}, \tag{38}$$

$$\mathbf{Z}^{(n)} = \mathbf{U}^{(2,n)H}\mathbf{U}^{(1,n)} \in \mathbb{C}^{R\times R}, \tag{39}$$

where $n \in \{1, \ldots, N\}$. Note that by exploiting the structure of $\mathbf{u}_r^{(1,n)H}\mathbf{u}_s^{(1,n)}$ and $\mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(2,n)}$, the computational cost of $w_{rs}^{(1,n)}$ and $w_{rs}^{(2,n)}$ have both been reduced from $2M_n(I_n - 1) - 1$ flops to $2M_n$ flops.

The different steps in the transformation of the original MHR decomposition (8) into the SD factorization (23) are listed in Table I, together with an estimate of their complexity (in flops). t is assumed that the R-SVD and QZ methods are used for the computation of the SVD and GEVD, respectively [3]. The most expensive steps are the initial dimensionality reduction step, which has a complexity of the size of $6(\prod_{n=1}^{N}I_n)M^2$, and the computation of the matrices $\{\mathbf{Z}^{(n)}\}$ used in the construction of $\mathbf{\Psi}^H\mathbf{\Psi}$, with a complexity of the order of $2N(\prod_{n=1}^{N}I_n)R^2$. The total complexity is of the order of $(6M + 2N\frac{R^2}{M})(\prod_{n=1}^{N}I_n)M$. If there are many snapshots ($M >> R$), then the cost is dominated by the conventional SVD reduction in Step 1. If $M \backsimeq R$, then the cost is about $(6 + 2N)R$ flops per tensor entry.

D. From SD to single-tone MHR problems

After $\mathbf{G} = \mathbf{F}^{-1}$ has been found from (23) we obtain $\mathbf{U\Sigma G}^T = \mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)}$ from relation (18). The generators $\{z_{r,n}\}_{n=1}^{N}$ associated with the rth source can now be found from the rth column of $\mathbf{U\Sigma G}^T$. In other words, coupled CPD has turned the MHR problem (7) into a set of decoupled single-tone MHR problems. Denote

$$\underline{\mathbf{\Gamma}}^{(n)} = \underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{U\Sigma G}^T, \tag{40}$$

$$\overline{\mathbf{\Gamma}}^{(n)} = \overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)}\mathbf{U\Sigma G}^T. \tag{41}$$

Algorithm operations	Eq.	Complexity order (flops)
SVD of $\mathbf{X}$	(8)	$6(\prod_{n=1}^{N} I_n)M^2 + 11M^3$
Compute $\{\widetilde{\mathbf{W}}^{(1,n)}, \widetilde{\mathbf{W}}^{(2,n)}\}_{n=1}^{N}$	(35)–(36)	$2(2(\sum_{n=1}^{N} M_n) - N)C_{R+1}^2$
Compute $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}\}_{n=1}^{N}$	(37)–(38)	$2NR$
Compute $\{\mathbf{Z}^{(n)}\}_{n=1}^{N}$	(39)	$(2(\sum_{n=1}^{N} M_n(I_n-1)) - N)R^2$
Compute $\{\mathbf{b}_{r_1 s_1}^{(n)H}\mathbf{b}_{r_2 s_2}^{(n)}\}_{1\leqslant n\leqslant N}^{1\leqslant r_1, r_2, s_1, s_2\leqslant R}$	(26)–(28)	$7NR^4$
Compute $\{\mathbf{Y}^{(1,n)}, \mathbf{Y}^{(2,n)}, \mathbf{Y}^{(3,n)}\}_{n=1}^{N}$	(26)–(28)	$N(2C_{R+1}^2 + 3RC_R^2 + 6C_{R+1}^2)$
Compute $\mathbf{\Psi}^H\mathbf{\Psi}$	(32)	$N(C_{R+1}^2 + 2RC_R^2 + 2C_{R+1}^2)$
EVD of $\mathbf{\Psi}^H\mathbf{\Psi}$	(32)	$\frac{4}{3}(C_{R+1}^2)^3$
GEVD of pair $(\mathbf{M}^{(r_1)}, \mathbf{M}^{(r_2)})$	(23)	$30R^2$

TABLE I
AN ESTIMATE OF THE COMPLEXITY OF THE DIFFERENT STEPS IN THE TRANSFORMATION FROM THE ORIGINAL MHR DECOMPOSITION (8) INTO THE SD FACTORIZATION (23). WE DENOTE $M_n = (\prod_{m=1,m\neq n} I_m)$.

Due to the shift-invariance property $\overline{\mathbf{A}}^{(n)} = \underline{\mathbf{A}}^{(n)} D_1([z_{r,1},\ldots,z_{R,n}])$, we have $\underline{\boldsymbol{\gamma}}_r^{(n)} = \overline{\boldsymbol{\gamma}}_r^{(n)} z_{r,n}$ so that the generator can be obtained as $z_{r,n} = (\overline{\boldsymbol{\gamma}}_r^{(n)H}\underline{\boldsymbol{\gamma}}_r^{(n)})/(\overline{\boldsymbol{\gamma}}_r^{(n)H}\overline{\boldsymbol{\gamma}}_r^{(n)})$.

In several applications, such as DOA estimation, $z_{r,n}$ is located on the unit circle, i.e., $z_{r,n} = e^{i\cdot\psi_{r,n}} = \cos(\psi_{r,n}) + i\sin(\psi_{r,n})$ for some $\psi_{r,n} \in [0, 2\pi)$. In that case, we can find the generator by minimizing

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

where $\alpha_{r,n} = \underline{\boldsymbol{\gamma}}_r^{(n)H}\underline{\boldsymbol{\gamma}}_r^{(n)} + \overline{\boldsymbol{\gamma}}_r^{(n)H}\overline{\boldsymbol{\gamma}}_r^{(n)}$, $\beta_{r,n} = \underline{\boldsymbol{\gamma}}_r^{(n)H}\overline{\boldsymbol{\gamma}}_r^{(n)}$ 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 normalized eigenvector associated with the dominant eigenvalue of $\mathbf{G}^{(r,n)}$, then the unit-norm generator is obtained as $z_{r,n} = (x_1^{(r,n)} + i\cdot x_2^{(r,n)})^2$.

E. Summary and discussion of overall method

The overall procedure is summarized as Algorithm 1. If condition (17) in Theorem II.1 is satisfied, then Algorithm 1 is guaranteed to find the MHR decomposition in the exact case. On the other hand, if condition (17) is not satisfied, then Algorithm 1 does not apply. In the latter case, what goes wrong is that we cannot find the R symmetric matrices $\{\mathbf{M}^{(r)}\}$ in (23) from $\mathbf{\Psi}^H\mathbf{\Psi}$, as the dimension of its kernel is larger than R [25]. It is also clear that Algorithm 1 works under more relaxed conditions than the approaches in Subsection II-B since it takes all harmonic structures into account at once. See [23] for a thorough discussion of the conditions.

In the inexact case, Algorithm 1 does not necessarily provide the Least Squares (LS) estimate of the factors of the original MHR decomposition (7) as they would be obtained from direct fitting. However, some robustness is built into the algorithm. More precisely, in Steps 1 and 5, LS estimates of range$(\mathbf{A}^{(1)}\odot\cdots\odot\mathbf{A}^{(N)})$ and the kernel of $\mathbf{\Psi}^H\mathbf{\Psi}$ can be obtained via SVD and EVD, respectively. Likewise, in Step 8 a LS estimate of the generators $\{z_{r,n}\}$ can be obtained via the closed-form solutions discussed in Subsection III-D. As mentioned in Subsection III-C, an algebraic solution to the SD problem (23) in Step 6 can be obtained via a GEVD. The result can serve as a good initialization for an SD method that in turn takes all the matrix slices in (23) into account in LS sense. Obviously, Algorithm 1 as a whole can be used to initialize a LS fitting method for MHR (e.g., [20]), i.e., the optimization algorithm is only used to refine the estimates.

Algorithm 1 SD method for MHR (Theorem II.1).

Input: $\mathbf{X} = (\mathbf{A}^{(1)}\odot\cdots\odot\mathbf{A}^{(N)})\mathbf{C}^T$ of the form (8).
1. Compute SVD $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ (and find R from $\mathbf{\Sigma}$).
2. Build $\mathbf{U}_{\text{red}}^{(n)}$ in (19) for $n \in \{1,\ldots,N\}$.
3. Construct $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^{N}$ given by (37)–(39).
4. Construct $\mathbf{\Psi}^H\mathbf{\Psi}$ in (32) via $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^{N}$.
5. Obtain matrices $\{\mathbf{M}^{(r)}\}$ from the kernel of $\mathbf{\Psi}^H\mathbf{\Psi}$.
6. Solve SD problem $\mathbf{M}^{(r)} = \mathbf{G}\mathbf{\Lambda}^{(r)}\mathbf{G}^T$, $r \in \{1,\ldots,R\}$.
7. Compute $\mathbf{N} = \mathbf{U}\mathbf{\Sigma}\mathbf{G}^T$.
8. Obtain $\{z_{r,n}\}$ from $\mathbf{N}$ via single-tone MHR.

Output: $\{z_{r,n}\}$

F. Variant 1: Compression

Note that the submatrices in (22) have the size $(J_n \times R)$, with $J_n \backsimeq \prod_{n=1}^{N} I_n$. We may consider compressing the tensors $\mathcal{U}_{\text{red}}^{(n)}$. Ideally, $\mathbf{U}^{(1,n)}$ and $\mathbf{U}^{(2,n)}$ in (20)–(21) will be replaced by matrices of size $(R \times R)$, as there are only R harmonics. It makes sense to project on the R-dimensional dominant subspace of $[\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)}]$, which may be computed by means of an SVD. We may equivalently compute the R-dimensional dominant eigenspace of

$$[\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)}]^H \cdot [\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)}] \tag{42}$$

$$= \begin{bmatrix} \mathbf{W}^{(1,n)} & \mathbf{Z}^{(n)H} \\ \mathbf{Z}^{(n)} & \mathbf{W}^{(2,n)} \end{bmatrix} \in \mathbb{C}^{2R\times 2R}, \tag{43}$$

where $\mathbf{W}^{(1,n)}$, $\mathbf{W}^{(2,n)}$ and $\mathbf{Z}^{(n)}$ are the matrices that already appeared in (37)–(39). While the diagonal blocks can be computed cheaply thanks to the orthogonality of $\mathbf{U}$, the computation of $\mathbf{Z}^{(n)}$ is expensive. Recall from Subsection III-C that the computation of $\mathbf{Z}^{(n)}$ was also the most expensive part of the construction of $\mathbf{\Psi}^H\mathbf{\Psi}$. Consequently, although replacing $\mathbf{U}^{(1,n)}$ and $\mathbf{U}^{(2,n)}$ by $(R\times R)$ matrices makes the computational cost of Steps 3-4 in Algorithm 1 negligible, the savings are compensated by the cost of the compression itself, i.e., the complexity is about the same as that of Algorithm 1. However, the "compressed variant" of Algorithm 1 has another interesting feature. Recall from Subsection III-B that we use $\mathcal{U}_{\text{red}}^{(n)}$ to impose the harmonic structure in the nth

mode. In the noiseless case, the matrix $[\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)}]$ will exactly have rank R. Under less ideal circumstances however, the projection on the dominant subspace will implement a signal enhancement, increasing the robustness and accuracy of the parameter estimates. This will be confirmed by numerical experiments in Section V.

G. Variant 2: Spatial smoothing

As the ESPRIT method for 1D HR, Algorithm 1 requires that $\mathbf{C}$ in (8) has full column rank. A classical way of extending ESPRIT-type methods to cases where $\mathbf{C}$ does not have full column rank is spatial smoothing [18]. In detail, by capitalizing on the Vandermonde structure of the factor matrices $\{\mathbf{A}^{(n)}\}$, spatial smoothing yields a $(2N+1)$th-order tensor $\mathcal{X}^{(K_1,\ldots,K_N)} \in \mathbb{C}^{K_1\times\cdots\times K_N\times L_1\times\cdots\times L_N\times M}$ with entries

$$\begin{aligned} x^{(K_1,\ldots,K_N)}_{k_1,\ldots,k_N,l_1,\ldots,l_N,m} &= x_{l_1+k_1-1,\ldots,l_N+k_N-1,m} \\ &= \sum_{r=1}^{R}\prod_{p=1}^{N} z_{r,p}^{k_p-1}\prod_{q=1}^{N} z_{r,q}^{l_q-1} c_{m,r}, \end{aligned}$$

where $K_n+L_n = I_n+1$. Combining the first N modes and the last $N+1$ modes in a matrix representation $\mathbf{X}^{(K_1,\ldots,K_N)}_{\text{SS}} \in \mathbb{C}^{(\prod_{n=1}^{N} K_n)\times(\prod_{n=1}^{N} L_n)M}$, we obtain

$$\mathbf{X}^{(K_1,\ldots,K_N)}_{\text{SS}} = \left(\mathbf{A}^{(K_1,1)} \odot \cdots \odot \mathbf{A}^{(K_N,N)}\right) \mathbf{C}^T_{\text{SS}}, \quad (44)$$

where 'SS' stands for spatial smoothing, $\mathbf{A}^{(K_n,n)} = \mathbf{A}^{(n)}(1:K_n,:) \in \mathbb{C}^{K_n\times R}$, and

$$\mathbf{C}_{\text{SS}} = \mathbf{A}^{(L_1,1)} \odot \cdots \odot \mathbf{A}^{(L_N,N)} \odot \mathbf{C}, \quad (45)$$

in which $\mathbf{A}^{(L_n,n)} = \mathbf{A}^{(n)}(1:L_n,:) \in \mathbb{C}^{L_n\times R}$. This spatial smoothing is meant to make sure that $\mathbf{C}_{\text{SS}}$, replacing $\mathbf{C}$, has full column rank. Consequently, if condition (17) is satisfied after replacing the matrices $\mathbf{A}^{(1)},\ldots,\mathbf{A}^{(N)}$ and $\mathbf{C}$ by $\mathbf{A}^{(K_1,1)},\ldots,\mathbf{A}^{(K_N,N)}$ and the full column rank matrix $\mathbf{C}_{\text{SS}}$, then Algorithm 1 can still be used to solve the MHR problem via $\mathbf{X}^{(K_1,\ldots,K_N)}_{\text{SS}}$.

In practice, the smoothing parameters $\{K_n, L_n\}$ are most often chosen in a heuristic manner. One may for instance, in line with the experiments for ESPRIT in [30], [13], choose the pairs $\{K_n, L_n\}$ in such a way that the dimension of the matrix $\mathbf{X}^{(K_1,\ldots,K_N)} = (\mathbf{A}^{(K_1,1)} \odot \cdots \odot \mathbf{A}^{(K_N,N)})\mathbf{C}^T_{\text{SS}}$ is roughly square.

IV. Application in multirate MHR

So far we have implicitly assumed that the MHR tensor $\mathcal{X}$ is built from uniformly spaced sampling points (single-rate sampling) with normalized unit inter-element spacing along all directions, i.e., $\mathbf{a}^{(n)}_r = [1\ z_{r,n}\ z^2_{r,n}\ \ldots\ z^{I_n-1}_{r,n}]^T$, $n \in \{1,\ldots,N\}$. In this section we explain that Algorithm 1 can be extended to multirate sampling, which implies that the interval between consecutive samplings points can be larger than a single (normalized) unit. This is not possible for the single-rate case, as can easily be understood from the properties of the complex exponential. Let $\beta_n \in \mathbb{N}$ denote the interval between two consecutive sampling points along the n-th direction so that $\mathbf{a}^{(n)}_r = [1\ z^{\beta_n}_{r,n}\ z^{2\beta_n}_{r,n}\ \ldots\ z^{(I_n-1)\beta_n}_{r,n}]^T$. The question is whether $z_{r,n}$ can be uniquely recovered from $\mathbf{a}^{(n)}_r$. Because of the periodicity of the complex exponential, we have $z^{\beta_n}_{r,n} = |z_{r,n}|^{\beta_n} e^{i(\arg(z_{r,n})\beta_n + 2\pi\alpha)}$ for every $\alpha \in \mathbb{Z}$ and $\beta_n \in \mathbb{N}$. For a unique recovery of $z_{r,n}$ from $z^{\beta_n}_{r,n}$, the following is required:

$$\arg(z_{r,n})\beta_n < 2\pi\ ,\ \forall r \in \{1,\ldots,R\},\ \forall n \in \{1,\ldots,N\}. \quad (46)$$

Condition (46) corresponds in fact to the Nyquist sampling condition for multivariate complex exponentials. A way to relax this condition and thereby extend the MHR problem to sub-Nyquist sampling cases where $\arg(z_{r,n})\beta_n \geqslant 2\pi$, is to employ nonuniform sampling (e.g., [12]). In this section we explain how the proposed coupled CPD framework for MHR can be used to develop multirate sampling schemes (i.e., nonuniform sampling with regularities) for MHR that can algebraically support sub-Nyquist sampling ($\arg(z_{r,n})\beta_n \geqslant 2\pi$). Note that this is an advantage over the existing VDM-CPD based approaches for MHR, which only support sampling at or above the Nyquist rate.

A. Multirate sampling via coupled CPD

Consider an N-dimensional system in which multirate sampling is used in the various modes. More precisely, we assume that P_n different sampling rates are used along the n-th mode, $1 \leqslant n \leqslant N$. Let us for the r-th multidimensional component stack $I_{p,n}$ samples associated with the p-th sampling rate along the n-th mode in the column vector $\mathbf{a}^{(p,n)}_r = z^{k_{p,n}}_{r,n}[1, z^{\beta_{p,n}}_{r,n}, z^{2\beta_{p,n}}_{r,n}, \ldots, z^{(I_{p,n}-1)\beta_{p,n}}_{r,n}]^T \in \mathbb{C}^{I_{p,n}}$, where $\beta_{p,n}$ is the interleaving factor associated with the p-th sample rate of the n-th mode, and $k_{p,n}$ is an offset. Overall, we collect all n-mode samples in the column vector

$$\mathbf{a}^{(n)}_r = [\mathbf{a}^{(1,n)T}_r, \mathbf{a}^{(2,n)T}_r, \ldots, \mathbf{a}^{(P_n,n)T}_r]^T \in \mathbb{C}^{I_n}, \quad (47)$$

where $I_n = \sum_{p=1}^{P_n} I_{p,n}$. The connection with the preceding discussion that we are making is that the column vector $\mathbf{a}^{(n)}_r$ in (7) now takes the form (47). This leads to a multirate version of $\mathcal{X}$ in (7):

$$\mathcal{X}_{\text{MR}} = \sum_{r=1}^{R} \mathbf{a}^{(1)}_r \otimes \cdots \otimes \mathbf{a}^{(N)}_r \otimes \mathbf{c}_r \in \mathbb{C}^{I_1\times\cdots\times I_N\times M}, \quad (48)$$

with matrix representation (cf. (8)):

$$\mathbf{X}_{\text{MR}} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{C}^T \in \mathbb{C}^{(\prod_{n=1}^{N} I_n)\times M}, \quad (49)$$

where 'MR' stands for multirate and the columns of $\mathbf{A}^{(n)}$ in (49) are given by (47). Note that the samples stored in $\mathbf{a}^{(n)}_r$ may not be distinct, i.e., the subsampling and stacking in (47) may cause that $a^{(p,n)}_r(s) = a^{(q,n)}_r(t)$ for some $s \in \{1,\ldots,I_{p,n}\}$ and $t \in \{1,\ldots,I_{q,n}\}$. As an example,

if $P_n = 4$ and

$$\begin{aligned} &k_{1,n} = 1, \quad \beta_{1,n} = 5, \quad \mathbf{a}_r^{(1,n)} = [1\ z_{r,n}^{5}\ z_{r,n}^{10}\ z_{r,n}^{15}\ z_{r,n}^{20}]^T, \\ &k_{2,n} = 13, \quad \beta_{2,n} = 5, \quad \mathbf{a}_r^{(2,n)} = [z_{r,n}^{13}\ z_{r,n}^{18}\ z_{r,n}^{23}\ z_{r,n}^{28}]^T, \\ &k_{3,n} = 1, \quad \beta_{3,n} = 10, \quad \mathbf{a}_r^{(3,n)} = [1\ z_{r,n}^{10}\ z_{r,n}^{20}\ z_{r,n}^{30}]^T, \\ &k_{4,n} = 1, \quad \beta_{4,n} = 13, \quad \mathbf{a}_r^{(4,n)} = [1\ z_{r,n}^{13}]^T, \end{aligned}$$

then

$$\mathbf{a}_r^{(n)} = [1\ z_{r,n}^{5}\ z_{r,n}^{10}\ z_{r,n}^{15}\ z_{r,n}^{20} \mid z_{r,n}^{13}\ z_{r,n}^{18}\ z_{r,n}^{23}\ z_{r,n}^{28} \mid 1\ z_{r,n}^{10}\ z_{r,n}^{20}\ z_{r,n}^{30} \mid 1\ z_{r,n}^{13}]^T \tag{50}$$

with the equalities

$$\begin{aligned} &\mathbf{a}_r^{(n)}(1) = \mathbf{a}_r^{(n)}(10) = \mathbf{a}_r^{(n)}(14) = 1, \quad \mathbf{a}_r^{(n)}(3) = \mathbf{a}_r^{(n)}(11) = z_{r,n}^{10}, \\ &\mathbf{a}_r^{(n)}(5) = \mathbf{a}_r^{(n)}(12) = z_{r,n}^{20}, \qquad\qquad \mathbf{a}_r^{(n)}(6) = \mathbf{a}_r^{(n)}(15) = z_{r,n}^{13}. \end{aligned}$$

For the pth sampling rate along the nth mode we can extract a subtensor $\mathcal{X}_{\text{MR}}^{(p,n)} \in \mathbb{C}^{I_1 \times \cdots I_{n-1} \times I_{p,n} \times I_{n+1} \times \cdots \times I_N \times M}$ from $\mathcal{X}_{\text{MR}}$, admitting the PD

$$\mathcal{X}_{\text{MR}}^{(p,n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \otimes \cdots \otimes \mathbf{a}_r^{(n-1)} \otimes \mathbf{a}_r^{(p,n)} \otimes \mathbf{a}_r^{(n+1)} \otimes \cdots \mathbf{a}_r^{(N)} \otimes \mathbf{c}_r, \tag{51}$$

with matrix representation

$$\mathbf{X}_{\text{MR}}^{(p,n)} = \left(\bigodot_{s=1}^{n-1} \mathbf{A}^{(s)} \odot \mathbf{A}^{(p,n)} \odot \bigodot_{t=n+1}^{N} \mathbf{A}^{(t)} \right) \mathbf{C}^T, \tag{52}$$

where $\mathbf{A}^{(p,n)} = [\mathbf{a}_1^{(p,n)}, \ldots, \mathbf{a}_R^{(p,n)}] \in \mathbb{C}^{I_{p,n} \times R}$ is Vandermonde. In other words, only the samples associated with the pth single rate in the nth mode of $\mathcal{X}_{\text{MR}}$ are considered. Similar to single-rate MHR, the factorization of $\mathbf{X}_{\text{MR}}^{(p,n)}$ can be interpreted as a VDM-CPD when the structure of the factors $\{\mathbf{A}^{(m)}\}_{m \neq n}$ is ignored. Note also that the factorizations (52) are coupled via $\mathbf{C}$. Hence, the multirate sampled MHR problem already corresponds to a coupled CPD problem before invoking shift-invariance.

B. ESPRIT-type algorithm for multirate sampling

A nice feature of Algorithm 1, is that in cases where $\mathbf{C}$ in (52) has full column rank ($M \geqslant R$), it can easily be adapted to multirate MHR problems. We briefly discuss the modifications of the different steps in Algorithm 1. In particular, it will be explained that multirate sampling does not increase the computational complexity.

1) Step 1: As in the case of uniform sampling, the dimension of $\mathcal{X}_{\text{MR}}$ in (48) is first reduced in a preprocessing step. Consequently, $\mathbf{C}$ in (49) is replaced by a nonsingular matrix $\mathbf{F}$, which leads to the following equivalent of (18):

$$\mathbf{U}_{\text{MR}} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{F}^T. \tag{53}$$

The difference with (18) is that $\mathbf{A}^{(n)}$, $n \in \{1, \ldots, N\}$ are not Vandermonde but multirate extensions.

2) Step 2: Similar to the way (51)–(52) has been obtained from $\mathcal{X}_{\text{MR}}$, we extract for each sampling rate $1 \leqslant p \leqslant P_n$ along each mode $1 \leqslant n \leqslant N$ a tensor $\mathcal{U}_{\text{red}}^{(p,n)} \in \mathbb{C}^{2 \times J_{p,n} \times R}$ with matrix representation (cf. (12) and (19)):

$$\mathbf{U}_{\text{red}}^{(p,n)} = \begin{bmatrix} \mathbf{U}^{(1,p,n)} \\ \mathbf{U}^{(2,p,n)} \end{bmatrix} = (\mathbf{A}^{(2,p,n)} \odot \mathbf{B}^{(p,n)})\mathbf{F}^T, \tag{54}$$

where $\mathbf{A}^{(2,p,n)} \in \mathbb{C}^{2 \times R}$ with $\mathbf{a}_r^{(2,p,n)} = [1, z_{r,n}^{\beta_{p,n}}]^T$ and $\mathbf{B}^{(p,n)} = (\bigodot_{s=1}^{n-1} \mathbf{A}^{(s)}) \odot \underline{\mathbf{A}}^{(p,n)} \odot (\bigodot_{t=n+1}^{N} \mathbf{A}^{(t)}) \in \mathbb{C}^{J_{p,n} \times R}$ with $J_{p,n} = (I_{p,n} - 1)(\prod_{m=1, m \neq n}^{N} I_m)$. The decompositions (54) can be jointly handled as in Section III. The difference is that we have a PD per mode *and* per rate. The matrix $\mathbf{G}^{(N)}$ in Theorem II.1 is replaced by

$$\mathbf{G}_{\text{MR}}^{(N)} = \begin{bmatrix} \mathbf{G}_{\text{MR}}^{(P_1,1)} \\ \vdots \\ \mathbf{G}_{\text{MR}}^{(P_N,N)} \end{bmatrix},$$

where

$$\mathbf{G}_{\text{MR}}^{(P_n,n)} = \begin{bmatrix} C_2\left(\mathbf{A}^{(2,1,n)}\right) \odot C_2\left(\mathbf{B}^{(1,n)}\right) \\ \vdots \\ C_2\left(\mathbf{A}^{(2,P_n,n)}\right) \odot C_2\left(\mathbf{B}^{(P_n,n)}\right) \end{bmatrix}.$$

Hence, if $\mathbf{G}_{\text{MR}}^{(N)}$ and $\mathbf{C}$ have full column rank, then the multirate MHR terms $\{\mathbf{a}_r^{(1)} \otimes \cdots \otimes \mathbf{a}_r^{(N)} \otimes \mathbf{c}_r\}_{r=1}^{R}$ in (48) are unique and can be obtained via the multirate variant of Algorithm 1. Note that, in ensuring uniqueness, it is possible to trade off the number of sampling rates for the number of samples per rate.

3) Steps 3 and 4: The multirate version of $\mathbf{\Psi}^H\mathbf{\Psi}$ is given by

$$\mathbf{\Psi}^H\mathbf{\Psi} = \sum_{n=1}^{N} \sum_{p=1}^{P_n} \mathbf{\Psi}^{(p,n)H}\mathbf{\Psi}^{(p,n)}, \tag{55}$$

where $\mathbf{\Psi}^{(p,n)H}\mathbf{\Psi}^{(p,n)}$ is the matrix obtained from $\mathbf{U}_{\text{red}}^{(p,n)}$ via Steps 3–4 in Algorithm 1. More specifically, the products (33)–(34) associated with $\mathbf{\Psi}^{(p,n)H}\mathbf{\Psi}^{(p,n)}$ are built from the columns of the $((I_{p,n} - 1)(\prod_{m=1, m \neq n}^{N} I_m) \times R)$ matrices $\mathbf{U}^{(1,p,n)}$ and $\mathbf{U}^{(2,p,n)}$ in (54). From Subsection III-C we know that the cost of the construction of a "single-rate" matrix in (55), say $\mathbf{\Psi}^{(p,n)H}\mathbf{\Psi}^{(p,n)}$ is of the order $I_{p,n}(\prod_{m=1, m \neq n}^{N} I_m)R^2$. Hence, the complexity of the construction of the multirate extension of $\mathbf{\Psi}^H\mathbf{\Psi}$ is of the order $\sum_{n=1}^{N} \sum_{p=1}^{P_n} I_{p,n}(\prod_{m=1, m \neq n}^{N} I_m)R^2$. In case no samples are used for more than one sampling rate (cf. discussion in Subsection IV-A), we have $I_n = \sum_{p=1}^{P_n} I_{p,n}$ and we obtain exactly the same expression for the complexity as in single-rate MHR. In general we can say that multirate sampling does not substantially affect the computational complexity.

4) Steps 5, 6 and 7: Identical to single-rate MHR.

5) Step 8: A difference with uniform sampling is that we now have to solve multirate sampled single-tone MHR problems in Step 8 in Algorithm 1. This can for instance be accomplished by the polynomial rooting procedure presented in [24]. Briefly, if one of the following two conditions

$$\begin{cases} \beta_{p,n} \text{ and } \beta_{q,n} \text{ are coprime for some } p \neq q, & \text{(56a)} \\ \exists\ \beta_{p,n} \ :\ \arg(z_{r,n})\beta_{p,n} < 2\pi\ ,\ \forall r \in \{1,\ldots,R\}, & \text{(56b)} \end{cases}$$

is satisfied, then $z_{r,n}$ can be recovered from the minimizer of the univariate polynomial $f(z_{r,n}) = \sum_{p=1}^{P_n} \|\mathbf{U}^{(1,p,n)}\mathbf{g}_r z_{r,n}^{\beta_{p,n}} - \mathbf{U}^{(2,p,n)}\mathbf{g}_r\|_F^2$, where $\mathbf{G} = [\mathbf{g}_1,\ldots,\mathbf{g}_R]$ denotes the nonsingular matrix obtained after solving the SD problem in Step 6 in Algorithm 1. Note that condition (56a) does not prevent sub-Nyquist sampling. In DOA estimation, the generator $z_{r,n}$ is known to be located on the unit circle such that $z_{r,n} = e^{i\cdot\psi_{r,n}}$ for some $\psi_{r,n} \in \mathbb{R}$. In this special case, the real number $\psi_{r,n}$ associated with $z_{r,n}$ can be found from the minimizer of $f(\psi_{r,n}) = \sum_{m=1}^{P_n} \|\mathbf{U}^{(1,p,n)}\mathbf{g}_r e^{i\cdot\psi_{r,n}\cdot\beta_{p,n}} - \mathbf{U}^{(2,p,n)}\mathbf{g}_r\|_F^2$, which in turn can be found by rooting a polynomial of degree $2\max_{1\leqslant p\leqslant P_n}\beta_{p,n}$. See [24] for details about the polynomial rooting procedure.

C. Illustrative example in DOA estimation

By way of example, let us mention that the presented multirate MHR sampling scheme can be used in DOA estimation based on Uniform Rectangular Arrays (URAs). A distinct advantage of the regularly sampled URA is its computational simplicity (e.g., computation via VDM-CPD). Compared to URA, irregularly sampled arrays can however have a higher aperture. Multirate MHR sampling based on superimposed sub-URAs can provide a good compromise between the regularly sampled URA and the irregularly sampled sparse array. In short, we assume that (52) takes the form

$$\mathbf{X}^{(p_1,p_2)} = (\mathbf{A}^{(p_1,1)} \odot \mathbf{A}^{(p_2,2)})\mathbf{C}^T, \qquad (57)$$

where $\mathbf{C}$ has full column rank, $z_{r,1} = e^{-i\omega_c d_1 \sin(\phi_r)\cos(\theta_r)/c}$ and $z_{r,2} = e^{-i\omega_c d_2 \sin(\phi_r)\sin(\theta_r)/c}$ in which ω_c is the carrier frequency, c is the speed of propagation, d_n denotes the nominal inter-element spacing of the sensors along the nth dimension (typically $d_n = \lambda/2$ in which $\lambda = \frac{2\pi c}{\omega_c}$ denotes the signal wavelength), θ_r is the azimuth angle associated with the r-th source, and ϕ_r is the elevation angle associated with the r-th source. The DOAs $\{\theta_r, \phi_r\}$ can be computed from (57) via Algorithm 1, despite sub-Nyquist sampling $\min_{1\leqslant m\leqslant P_n} d_n\beta_{m,n} > \lambda/2$.

V. Numerical experiments

Consider the MHR factorization of $\mathbf{X}$ in (8). The goal is to estimate the generators $\{z_{r,n}\}$ 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 Signal-to-Noise Ratio (SNR) is measured as: SNR [dB] $= 10\log(\|\mathbf{X}\|_F^2/\|\beta\mathbf{N}\|_F^2)$.

A. DOA estimation via multirate 2D HR sampling

a) Case 1: We first consider multiple snapshot 2D DOA estimation in which the model parameters for the uniformly sampled URA are $I_1 = I_2 = 9$, $M = 50$ and $R = 5$. Furthermore, half-wavelength sensor inter-element spacing will be used, implying that the generators are given by $z_{r,1} = e^{-i\pi\sin(\phi_r)\cos(\theta_r)}$ and $z_{r,2} = e^{-i\pi\sin(\phi_r)\sin(\theta_r)}$. The DOA parameters are set to $\boldsymbol{\theta} = [\theta_1,\theta_2,\theta_3,\theta_4,\theta_5]^T = \frac{\pi}{180}[10,\ 30,\ 50,\ 70,\ 80]^T$ and $\boldsymbol{\phi} = [\phi_1,\phi_2,\phi_3,\phi_4,\phi_5]^T = \frac{\pi}{180}[-70,\ -50,\ 50,\ 70,\ 80]^T$.

We compare Algorithm 1 and its compressed version with the ESPRIT-type method in [21] which represents the VDM-CPD approach discussed in Subsection II-B. The methods will be referred to as SD, compressed SD and VDM-CPD, respectively. We also consider a multirate sampled URA constructed from two superimposed regular URAs, which allows us to increase the DOA resolution without increasing the number of spatial sampling points. The spatial sampling parameters are chosen such that $P_1 = P_2 = 2$ and the columns of $\mathbf{A}^{(p_1,1)}$ and $\mathbf{A}^{(p_2,2)}$ in (57) are of the form $\mathbf{a}_r^{(1,n)} = [1\ z_{r,n}^5\ z_{r,n}^{10}\ z_{r,n}^{15}\ z_{r,n}^{20}]^T$ and $\mathbf{a}_r^{(2,n)} = [1\ z_{r,n}^{11}\ z_{r,n}^{22}\ z_{r,n}^{33}\ z_{r,n}^{44}]^T$ with $n \in \{1,2\}$. The multirate variant of Algorithm 1 will be referred to as MR-SD. The compressed version of MR-SD will be referred to as compressed MR-SD. Note that the uniformly sampled and the multirate sampled methods make use of the same number of spatial sampling points. The SD problem in Step 6 of Algorithm 1 will be solved by means of the extended QZ method [29], which generalizes the QZ method for the computation of a GEVD [3] to more than two matrices. The centro-symmetry of the regular URA allows us to apply Forward-Backward Averaging (FBA) in a preprocessing step of the SD, compressed SD and VDM-CPD methods.[1] Due to the asymmetric structure of the multirate sampled URA, FBA is not possible for MR-SD and compressed MR-SD.

We compute the root mean square error $\mathrm{E}_{\mathrm{RMS}} = \sqrt{\frac{1}{2RT_{\mathrm{MC}}}\sum_{t=1}^{T_{\mathrm{MC}}} \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_t\|_F^2 + \|\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}_t\|_F^2}$, where T_{MC}, $\hat{\boldsymbol{\theta}}_t$ and $\hat{\boldsymbol{\phi}}_t$ denote the number of Monte Carlo trials and the trial estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, respectively. The root mean square of the deterministic Cramér-Rao Bound (CRB) [26], [32] are also computed. The $\mathrm{E}_{\mathrm{RMS}}$ and CRB values over $T_{\mathrm{MC}} = 200$ trials as a function of SNR can be seen in Figure 1. We first notice that due to the increased aperture, the MR-SD and compressed MR-SD methods perform much better than the other methods. We also observe that the pure SD method for MHR on average performs worse than compressed SD and VDM-CPD. This is due to a few runs in which the matrix $\mathbf{Q}^{(N)}$ in (32) is ill-conditioned. The extra orthonormalization step in Section III-F seems to make the compressed version more robust. Overall, the multirate versions of SD seem to be the methods of choice for this experiment despite

[1] FBA means that $\mathbf{X}$ in (8) is replaced by $[\mathbf{X}, \mathbf{J}\mathbf{X}^*]$, where $\mathbf{J}$ is an antidiagonal matrix with unit entries on its antidiagonal [15]. In short, FBA virtually doubles the number of snapshots from M to $2M$. See also [23] for further details and discussion.

the fact that they were not supported by FBA. However, since VDM-CPD essentially amounts to the GEVD of a $((I_1-1)I_2 \times R)$ matrix pencil, it has a lower complexity than the SD methods.

b) Case 2: In our second multirate sampling experiment, we set $R = 3$ while in each Monte Carlo experiment the azimuth angles $\{\theta_r\}$ and elevation angles $\{\phi_r\}$ are randomly drawn from a uniform distribution over $[0, \pi]$ and $[-\pi/2, \pi/2]$, respectively. The $\mathrm{E_{RMS}}$ values over $T_{\mathrm{MC}} = 200$ trials as a function of SNR can be seen in Figure 2. The randomness of $\{\theta_r, \phi_r\}$ can yield difficult cases, explaining the degradation on performance despite R has been reduced to three. We also observe that below 20 dB SNR, the compressed SD method performed better than the CPD-VDM and SD methods.

B. DOA estimation via sparse 2D HR sampling

We now consider a multiple snapshot 2D DOA estimation problem that involves a larger array where the model parameters for the uniformly sampled URA are $I_1 = I_2 = 21$, $M = 50$ and $R = 5$. Multirate sampling can also be used to implement a thinned version of this URA, which leads to a reduction in the number of spatial sampling points. Let $P_1 = P_2 = 2$ and let the columns of $\mathbf{A}^{(p_1,1)}$ and $\mathbf{A}^{(p_2,2)}$ in (57) be of the form $\mathbf{a}_r^{(1,n)} = [1\ z_{r,n}^3\ z_{r,n}^6\ z_{r,n}^9\ z_{r,n}^{12}\ z_{r,n}^{15}\ z_{r,n}^{18}\ z_{r,n}^{21}]^T$ and $\mathbf{a}_r^{(2,n)} = [1\ z_{r,n}^7\ z_{r,n}^{14}\ z_{r,n}^{21}]^T$ with $n \in \{1, 2\}$. The DOA parameters $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are the same as in the previous experiment. The $\mathrm{E_{RMS}}$ and CRB values over $T_{\mathrm{MC}} = 200$ trials as a function of SNR can be seen in Figure 3. As expected, sparse sampling leads to a loss of performance (about 5 dB SNR). On the other hand, a significant reduction of the number of spatial sampling points has been accomplished (less than 20% of the fully sampled URA).

C. DOA estimation via single snapshot 2D HR

We also consider the single snapshot DOA estimation problem. We assume that the URA model parameters are $I_1 = I_2 = 8$, $M = 1$ and $R = 4$. The DOA parameters are set to $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4]^T = \frac{\pi}{180}[10,\ 30,\ 50,\ 50]^T$ and $\boldsymbol{\phi} = [\phi_1, \phi_2, \phi_3, \phi_4]^T = \frac{\pi}{180}[-70,\ -50,\ 50,\ 70]^T$. We compare the SD and compressed SD methods with the MultiDimensional Folding (MDF) method developed in [11] for single-snapshot MHR and the unitary ESPRIT method in [4], which takes the unit norm property of the generators into account. In accordance with the heuristic mentioned in Subsection III-E, the smoothing parameters for the unitary ESPRIT, SD and compressed SD methods are set to $(K_1, K_2) = (5, 5)$. In order to illustrate the effect of the choice of smoothing parameters, compressed SD and unitary ESPRIT with configuration $(K_1, K_2) = (6, 6)$ are also considered. All methods apply FBA in a preprocessing step. The $\mathrm{E_{RMS}}$ values over 500 trials as a function of SNR can be seen in Figure 4. We observe that below 30 dB SNR, SD performs worse than the other methods. In this experiment, the choice between $(K_1, K_2) = (5, 5)$ or $(K_1, K_2) = (6, 6)$ did not significantly affect the performance of the compressed SD method while, on the other hand, it did seem to affect the performance of the unitary ESPRIT method.

D. Small-scale 2D HR

As our final example, we consider a small-scale 2D HR problem in which $I_1 = I_2 = 3$, $M = 50$ and $R = 4$. The goal of this experiment is to show that optimization-based methods for MHR can benefit from a proper initialization. We compare the algebraic SD and compressed SD methods with the optimization-based Nonlinear Least Squares (NLS) algorithm in Tensorlab [31]. More precisely, we consider both an unconstrained and a Vandermonde constrained CPD interpretation of (8). In the latter case the Vandermonde structure of the factor matrices is taken into account while in the former they are ignored. Randomly initialized versions of these optimization-based methods are referred to as CPD and VDM-CPD, respectively. Note that the complexity of NLS can be quite high[2], though it can be significantly reduced by exploiting the MHR structure (e.g. [2]). In order to illustrate that SD techniques can provide a good initialization for optimization-based methods, compressed SD will be applied in combination with VDM-CPD. The overall procedure is referred to as SD-VDM-CPD. In this experiment the real and imaginary parts of the generators $\{z_{r,n}\}$ are randomly drawn from a uniform distribution with support $[-\frac{1}{2}, \frac{1}{2}]$. We compute the root mean square error $\mathrm{E_{RMS}} = \sqrt{\frac{1}{2RT_{\mathrm{MC}}} \sum_{t=1}^{T_{\mathrm{MC}}} \sum_{r=1}^{R} |z_{r,1}^{(t)} - \hat{z}_{r,1}^{(t)}|^2 + |z_{r,2}^{(t)} - \hat{z}_{r,2}^{(t)}|^2}$, where T_{MC}, $z_{r,n}^{(t)}$ and $\hat{z}_{r,n}^{(t)}$ denote the number of trials, the true generator in trial t and the associated estimate, respectively. The $\mathrm{E_{RMS}}$ values over 200 trials as a function of SNR can be seen in Figure 5. SD-VDM-CPD clearly performed best. The figure also shows that the optimization-based methods can be sensitive to their initialization.

VI. Conclusion

Using a link between the coupled CPD framework and MHR we have derived an algebraic algorithm for MHR that under mild conditions is guaranteed to find the decomposition in the exact case. Methods based on the coupled CPD framework, such as the presented algorithm, can simultaneously exploit the harmonic structure in all dimensions. As such, the method works under more relaxed conditions compared to existing algebraic VDM-CPD based methods. Furthermore, since the algorithm reduces the MHR problem to the computation of a GEVD, it can be interpreted as an MHR extension of ESPRIT. We also presented a compressed variant that is more robust against perturbation noise. The algorithm

[2]As an example, the Gauss-Newton algorithm with dogleg trust region for a Nth order cubic tensor with dimensions I has a complexity of the order $2(N + \mathrm{it}_{gn})RI^N + \frac{8}{3}N^2R^3I^3$, where it_{gn} denotes the number of iterations [19].

(and its compressed variant) can be used to initialize computationally more expensive optimization-based methods for MHR.

Coupled CPD is a flexible tool that can be used for various tasks in array processing [24]. This point was illustrated by showing that the coupled CPD approach can algebraically support multirate sampling, i.e., even the multirate sampling MHR problem can be reduced to a GEVD problem. This makes it interesting for large scale MHR problems in which sub-Nyquist sampling is often a desired feature.

Numerical experiments showed that for 2D DOA estimation based on URAs, the presented techniques can lead to improved DOA estimation compared to the conventional VDM-CPD approaches. In particular, the benefits of compression and multirate sampling were confirmed.

References

[1] 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.

[2] W. Ding, L. Qi, and Y. Wei, "Fast Hankel tensor-vector product and its application to exponential data fitting," *Numer. Linear Algebra Appl.*, vol. 22, pp. 814–832, 2015.

[3] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed. John Hopkins university press, 2013.

[4] M. Haardt and J. A. Nossek, "Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 161–169, Jan. 1998.

[5] M. Haardt, M. Pesavento, F. Roemer, and M. E. Korso, "Subspace methods and exploitation of special array structures," in *Academic Press Library in Signal Processing: Volume 3 — Array and Statistical Signal Processing*, A. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, ch. 15, pp. 651–717.

[6] Y. Hua and T. K. Sarkar, "On SVD for estimating generalized eigenvalues of singular matrix pencil in noise," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 892–900, Apr. 1991.

[7] T. Jiang, N. D. Sidiropoulos, and J. M. F. ten Berge, "Almost-sure identifiability of multidimensional harmonic retrieval," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1849–1859, Sep. 2001.

[8] 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.

[9] J. Liu and X. Liu, "An eigenvector-based approach for multidimensional frequency estimation with improved identifiability," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4543–4557, Dec. 2006.

[10] J. Liu, X. Liu, and W. Ma, "Multidimensional frequency estimation with finite snapshots in the presence of identical frequencies," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5179–5194, Nov. 2007.

[11] X. Liu and N. D. Sidiropoulos, "Almost sure identifiability of multidimensional constant modulus harmonic retrieval," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2366–2368, Sep. 2002.

[12] M. Misha and Y. C. Eldar, "Sub-Nyquist sampling: Bridging theory and practice," *IEEE Signal Processing Magazine*, vol. 29, pp. 98–124, Nov. 2011.

[13] J. M. Papy, L. De Lathauwer, and S. Van Huffel, "Exponential data fitting using multilinear algebra: The single-channel and multi-channel case," *Numerical Linear Algebra with Applications*, vol. 12, no. 8, pp. 809–826, 2005.

[14] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation to signal parameter estimation," *Proc. of the IEEE*, vol. 74, no. 7, pp. 1044–1045, Jul. 1986.

[15] 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.

[16] R. Roy and T. Kailath, "ESPRIT – Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[17] E. Sanchez and B. R. Kowalski, "Generalized rank annihilation factor analysis," *Anal. Chem.*, vol. 58, no. 2, p. 496, Feb. 1986.

[18] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-on-arrival of coherent signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 8, pp. 806–811, Aug. 1985.

[19] L. Sorber, M. Van Barel, and L. De Lathauwer, "Optimization-based algorithms for tensor decompositions : Canonical polyadic decomposition, decomposition in rank-$(L_r, L_r, 1)$ terms and a new generalization," *SIAM J. Opt.*, vol. 23, pp. 695–720, 2013.

[20] ——, "Structured data fusion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 586–600, Jun. 2015.

[21] M. Sørensen and L. De Lathauwer, "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.

[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] ——, "Multidimensional harmonic retrieval via coupled canonical polyadic decompositions — Part I: Model and identifiability," ESAT-STADIUS, KU Leuven, Belgium, Tech. Rep. 15-149, 2015.

[24] ——, "Multiple invariance ESPRIT for nonuniform linear arrays: A coupled canonical polyadic decomposition approach," *IEEE Trans. Signal Processing*, vol. 64, no. 14, pp. 3693–3704, 2016.

[25] 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.

[26] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on ASSP*, vol. 37, no. 5, pp. 720–741, 1989.

[27] A. Thakre, M. Haardt, F. Roemer, and K. Giridhar, "Tensor-based spatial smoothing (TB-SS) using multiple snapshots," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2715–2728, May 2010.

[28] A.-J. van der Veen, "Joint diagonalization via subspace fitting techniques," in *Proc. ICASSP*, May 07-11, Salt Lake City, Utah, USA, 2001.

[29] 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.

[30] S. Van Huffel, H. Chen, C. Decanniere, and P. Van Hecke, "Algorithm for time-domain NMR data fitting based on total least squares," *J. Magn. Reson. Ser. A*, vol. 110, pp. 238–237, 1994.

[31] 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/.

[32] S. F. Yau and Y. Bresler, "A compact Cramér-Rao bound expression for parametric estimation of superimposed signals," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1226–1230, May 1992.



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 signal and array processing.

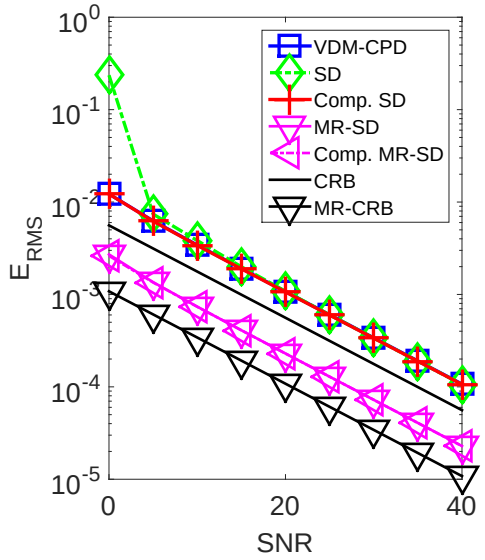

Fig. 1. E_{RMS} over 200 trials, multirate sampling case 1.

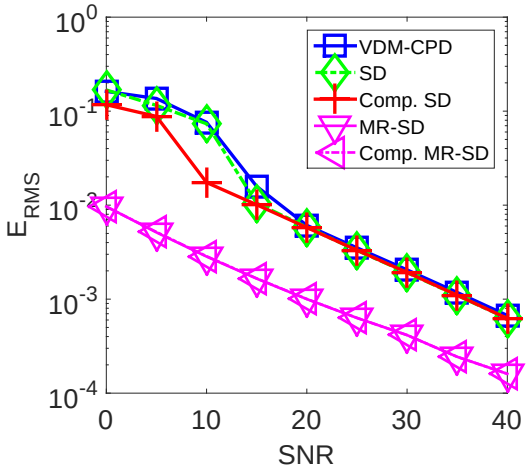

Fig. 2. E_{RMS} over 200 trials, multirate sampling case 2.

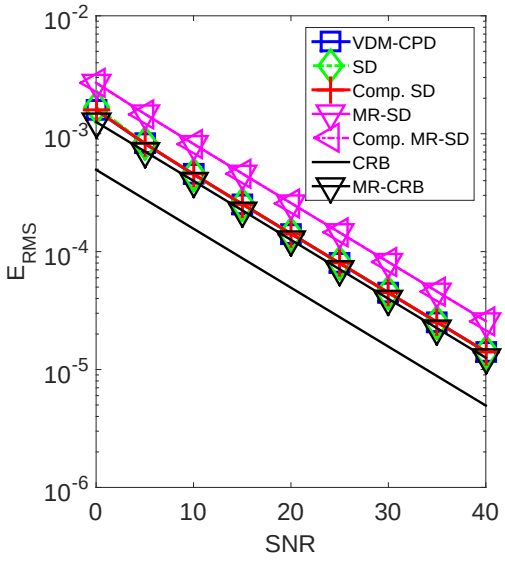

Fig. 3. E_{RMS} over 200 trials, sparse sampling case.

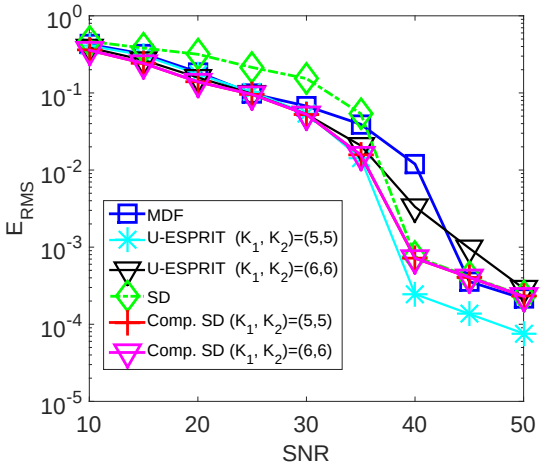

Fig. 4. E_{RMS} over 500 trials, single snapshot case.

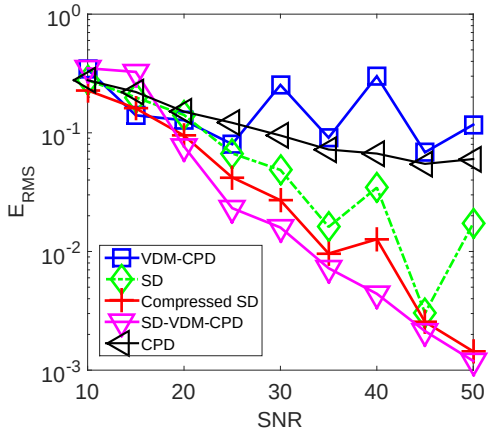

Fig. 5. E_{RMS} over 200 trials, small-scale case.



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.