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. Numerical experiments demonstrate that by simultaneously exploiting the harmonic structure in all dimensions and possibly also 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 [8] and references therein. In the companion paper [12] we have presented a connection between MHR and the recently introduced coupled Canonical Polyadic Decomposition (CPD) modeling framework [11], [13].

In this paper we explain that the coupled CPD model also provides an algebraic framework from which algorithms for MHR can be derived. In particular, we develop a linear algebra based Simultaneous matrix Diagonalization (SD) method for MHR that under mild conditions is guaranteed to find the decomposition in the noiseless case. In the noisy case, the SD method can be used to initialize an optimization-based method.

In the context of MHR, a reason to consider coupled CPD as opposed to the conventional VDM-CPD methods is that the former approach can support multirate sampling. This makes the approach interesting for large scale MHR problems where sub-Nyquist sampling is often a desired feature.

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 connection between MHR and coupled CPD. Section III develops a new SD method for MHR. In Section IV we show that, unlike the existing VDM-CPD based methods for MHR, the coupled CPD framework for MHR can algebraically support 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:

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

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

Matlab index notation will be used for submatrices of a given matrix. For example, $\mathbf{A}(1{:}k,:)$ represents the submatrix of $\mathbf{A}$ consisting of the rows from 1 to k of $\mathbf{A}$. $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$ denotes the diagonal matrix holding row

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.

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.

k of $\mathbf{A} \in \mathbb{C}^{I\times J}$ on its diagonal. Similarly, $\mathrm{Vecd}(\mathbf{A})$ denotes the column vector defined by $(\mathrm{Vecd}(\mathbf{A}))_i = (\mathbf{A})_{ii}$.

The k-th compound matrix of $\mathbf{A} \in \mathbb{C}^{I\times R}$ is denoted by $C_k(\mathbf{A}) \in \mathbb{C}^{C_I^k \times C_R^k}$, where $C_m^k = \frac{m!}{k!(m-k)!}$. It is the matrix containing the determinants of all $k\times k$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order. For more elementary notation, we refer to [12].

Tensor decompositions: Let us briefly present the tensor decompositions that are used in this paper. See the references in [12] 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:

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

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 $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$ 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:

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

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

II. MHR via coupled CPD

A. Link between MHR and coupled CPD

In this section we briefly explain how MHR and coupled CPD are connected. More details can be found in the companion paper [12]. Consider 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

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)} \circ \mathbf{c}_r\,, \tag{3}$$

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 unstructured $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{M\times R}$, where M is the number of snapshots. Note that in contrast to (1), the rank-1 terms in (3) are Vandermonde constrained. The minimal number of Vandermonde constrained rank-1 terms $\mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)} \circ \mathbf{c}_r$ that yield $\mathcal{X}$ in (3) in a linear combination will be denoted by $r_{\mathrm{VDM}}(\mathcal{X})$.

The PD of $\mathcal{X}$ in (3) has the following matrix representation

$$\mathbf{X} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{C}^T \in \mathbb{C}^{(\prod_{n=1}^{N} I_n)\times M}. \tag{4}$$

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

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

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

$$\mathbf{X}_{\mathrm{SS}}^{(K_1,\ldots,K_N)} = \left(\mathbf{A}^{(K_1,1)} \odot \cdots \odot \mathbf{A}^{(K_N,N)}\right) \mathbf{C}_{\mathrm{SS}}^T, \tag{5}$$

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

$$\mathbf{C}_{\mathrm{SS}} = \mathbf{A}^{(L_1,1)} \odot \cdots \odot \mathbf{A}^{(L_N,N)} \odot \mathbf{C}\,, \tag{6}$$

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 $\mathbf{C}_{\mathrm{SS}}$, replacing $\mathbf{C}$, has full column rank.

Define the row-selection matrices

$$\underline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{p=1}^{n-1} I_p} \otimes \underline{\mathbf{I}}_{I_n} \otimes \mathbf{I}_{\prod_{q=n+1}^{N} I_q}\,, \tag{7}$$

$$\overline{\mathbf{S}}_{(n)}^{(I_1,\ldots,I_N)} = \mathbf{I}_{\prod_{p=1}^{n-1} I_p} \otimes \overline{\mathbf{I}}_{I_n} \otimes \mathbf{I}_{\prod_{q=n+1}^{N} I_q}\,, \tag{8}$$

where $\underline{\mathbf{I}}_{I_n}$ and $\overline{\mathbf{I}}_{I_n}$ are the row-selection matrices given by

$$\begin{aligned} \underline{\mathbf{I}}_{I_n} &= \mathbf{I}_{I_n}(1:I_n-1, 1:I_n) \in \mathbb{C}^{(I_n-1)\times I_n}, \\ \overline{\mathbf{I}}_{I_n} &= \mathbf{I}_{I_n}(2:I_n, 1:I_n) \in \mathbb{C}^{(I_n-1)\times I_n}. \end{aligned}$$

A second spatial smoothing in the n-th dimension yields the tensor $\mathcal{X}_{\mathrm{SS}}^{(n)} \in \mathbb{C}^{2\times(\times_{p=1}^{n-1} K_p)\times(K_n-1)\times(\times_{q=n+1}^{N} K_q)\times(\times_{s=1}^{N} L_s)\times M}$ with matrix representation $\mathbf{X}_{\mathrm{SS}}^{(n)} \in \mathbb{C}^{2(\prod_{p=1,p\neq n}^{N} K_p)(K_n-1)\times(\prod_{q=1}^{N} L_q)M}$ admitting the factorization

$$\mathbf{X}_{\mathrm{SS}}^{(n)} = \begin{bmatrix} \underline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{X}_{\mathrm{SS}}^{(K_1,\ldots,K_N)} \\ \overline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{X}_{\mathrm{SS}}^{(K_1,\ldots,K_N)} \end{bmatrix} = \left(\mathbf{A}^{(2,n)} \odot \mathbf{B}_{\mathrm{SS}}^{(n)}\right)\mathbf{C}_{\mathrm{SS}}^T, \tag{9}$$

where $\mathbf{B}_{\mathrm{SS}}^{(n)} \in \mathbb{C}^{J_n\times R}$ with $J_n = (\prod_{p=1,p\neq n}^{N} K_p)(K_n-1)$ is given by

$$\mathbf{B}_{\mathrm{SS}}^{(n)} = \bigodot_{p=1}^{n-1} \mathbf{A}^{(K_p,p)} \odot \mathbf{A}^{(K_n-1,n)} \odot \bigodot_{p=n+1}^{N} \mathbf{A}^{(K_p,p)}. \tag{10}$$

Overall, we obtain

$$\begin{bmatrix} \mathbf{X}_{\mathrm{SS}}^{(1)} \\ \vdots \\ \mathbf{X}_{\mathrm{SS}}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(2,1)} \odot \mathbf{B}_{\mathrm{SS}}^{(1)} \\ \vdots \\ \mathbf{A}^{(2,N)} \odot \mathbf{B}_{\mathrm{SS}}^{(N)} \end{bmatrix} \mathbf{C}_{\mathrm{SS}}^T. \tag{11}$$

The structure of (11) is such that $\mathcal{X}$ must have the N-dimensional harmonic structure (3), even when the Khatri-Rao structure in (10) is ignored [12]. Clearly, (11) corresponds to a matrix representation of a coupled CPD. Hence, any numerical method for coupled CPD can be used to compute the MHR decomposition (3) via (11). See [7] and references therein for optimization-based methods for coupled CPD. In this paper we will focus on linear algebra based methods.

B. Uniqueness conditions for MHR

In this section we formulate the uniqueness conditions obtained in [12] for MHR. Based on these uniqueness conditions, we will in Sections III–IV discuss linear algebra based methods for MHR relying on the coupled CPD.

For the cases where $\mathbf{C}$ in (4) has full column rank, we have the MHR uniqueness condition stated in Theorem II.1. It makes use of the matrix

$$\mathbf{G}^{(N)} = \begin{bmatrix} C_2\left(\mathbf{A}^{(2,1)}\right) \odot C_2\left(\mathbf{B}^{(1)}\right) \\ \vdots \\ C_2\left(\mathbf{A}^{(2,N)}\right) \odot C_2\left(\mathbf{B}^{(N)}\right) \end{bmatrix}. \tag{12}$$

Theorem II.1. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times M}$ in (3) where the factor matrices $\{\boldsymbol{A}^{(n)}\}$ are Vandermonde. If*

$$\begin{cases} \boldsymbol{C} \text{ in (4) has full column rank,} \\ \boldsymbol{G}^{(N)} \text{ in (12) has full column rank,} \end{cases} \tag{13}$$

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

Theorem II.2 below extends Theorem II.1 to cases where $\mathbf{C}$ does not have full column rank. It makes use of a variant of the matrix in (12) obtained via spatial smoothing:

$$\mathbf{G}_{\text{SS}}^{(N)} = \begin{bmatrix} C_2\left(\mathbf{A}^{(2,1)}\right) \odot C_2\left(\mathbf{B}_{\text{SS}}^{(1)}\right) \\ \vdots \\ C_2\left(\mathbf{A}^{(2,N)}\right) \odot C_2\left(\mathbf{B}_{\text{SS}}^{(N)}\right) \end{bmatrix}. \tag{14}$$

Theorem II.2. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times M}$ in (3) where the factor matrices $\{\boldsymbol{A}^{(n)}\}$ are Vandermonde. If there exist pairs $\{(K_n, L_n)\}$ subject to $K_n + L_n = I_n + 1$ such that*

$$\begin{cases} \boldsymbol{C}_{SS} \text{ in (6) has full column rank,} \\ \boldsymbol{G}_{SS}^{(N)} \text{ in (14) has full column rank,} \end{cases} \tag{15}$$

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

III. MHR via SD

In this section we explain that the MHR problem, like classical ESPRIT for one-dimensional HR [6], can be reduced to a Generalized EigenValue Decomposition (GEVD) problem.

A. Step 0: Dimensionality reduction

We assume that condition (15) in Theorem II.2 is satisfied. This implies that $\bigodot_{n=1}^{N} \mathbf{A}^{(K_n,n)}$ has full column rank. Indeed, if $\bigodot_{n=1}^{N} \mathbf{A}^{(K_n,n)}$ does not have full column rank, then there exists a vector $\mathbf{y} \in \ker\left(\bigodot_{n=1}^{N} \mathbf{A}^{(K_n,n)}\right)$ and a vector $\mathbf{x} \in \mathbb{C}^{(\prod_{n=1}^{N} L_n)M)}$ such that we obtain an alternative decomposition $\mathbf{X}_{\text{SS}}^{(K_1,\ldots,K_N)} = \left(\bigodot_{n=1}^{N} \mathbf{A}^{(K_n,n)}\right)\left(\mathbf{C}_{\text{SS}} + \mathbf{x}\mathbf{y}^T\right)^T$, contradicting Theorem II.2. Thus, condition (15) implies that both $\bigodot_{n=1}^{N} \mathbf{A}^{(K_n,n)}$ and $\mathbf{C}_{\text{SS}}$ have full column rank. Let $\mathbf{X}_{\text{SS}}^{(K_1,\ldots,K_N)} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$ denote the compact SVD of $\mathbf{X}_{\text{SS}}^{(K_1,\ldots,K_N)}$ in (5), where $\mathbf{U} \in \mathbb{C}^{(\prod_{n=1}^{N} K_n) \times R}$ and $\mathbf{V} \in \mathbb{C}^{(\prod_{n=1}^{N} L_n)M) \times R}$ are columnwise orthonormal and $\boldsymbol{\Sigma} \in \mathbb{C}^{R \times R}$ is diagonal. If R is unknown, then it can be estimated as the rank of $\mathbf{X}_{\text{SS}}^{(K_1,\ldots,K_N)}$. We know that $\text{range}\left(\mathbf{X}_{\text{SS}}^{(K_1,\ldots,K_N)}\right) = \text{range}\,(\mathbf{U})$ and that there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R \times R}$ such that from relation (5) we obtain

$$\mathbf{U}_{\text{SS}}^{(K_1,\ldots,K_N)} := \mathbf{U}\boldsymbol{\Sigma} = \left(\mathbf{A}^{(K_1,1)} \odot \cdots \odot \mathbf{A}^{(K_N,N)}\right)\mathbf{F}^T, \tag{16}$$

$$\mathbf{V}^*\mathbf{F}^{-1} = \mathbf{C}_{\text{SS}} = \mathbf{A}^{(L_1,1)} \odot \cdots \odot \mathbf{A}^{(L_N,N)} \odot \mathbf{C}. \tag{17}$$

B. Step 1: From MHR to coupled CPD

From (16) we build the tensors $\mathcal{U}_{\text{SS,red}}^{(n)} \in \mathbb{C}^{2 \times (\times_{p=1}^{n-1} K_p) \times (K_n - 1) \times (\times_{q=n+1}^{N} K_q) \times R}$, $n \in \{1, \ldots, N\}$, each admitting a factorization

$$\mathbf{U}_{\text{SS,red}}^{(n)} := \begin{bmatrix} \mathbf{U}^{(1,n)} \\ \mathbf{U}^{(2,n)} \end{bmatrix} = (\mathbf{A}^{(2,n)} \odot \mathbf{B}_{\text{SS}}^{(n)})\mathbf{F}^T, \tag{18}$$

where 'red' stands for reduced, and

$$\mathbf{U}^{(1,n)} = \underline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\text{SS}}^{(K_1,\ldots,K_N)} \in \mathbb{C}^{M_n(K_n-1) \times R}, \tag{19}$$

$$\mathbf{U}^{(2,n)} = \overline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\text{SS}}^{(K_1,\ldots,K_N)} \in \mathbb{C}^{M_n(K_n-1) \times R}, \tag{20}$$

in which $M_n = (\prod_{m=1, m \neq n}^{N} K_m)$. Note that $J_n = M_n(K_n - 1)$. Overall, we obtain the compressed variant of (11):

$$\begin{bmatrix} \mathbf{U}_{\text{SS,red}}^{(1)} \\ \vdots \\ \mathbf{U}_{\text{SS,red}}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(2,1)} \odot \mathbf{B}_{\text{SS}}^{(1)} \\ \vdots \\ \mathbf{A}^{(2,N)} \odot \mathbf{B}_{\text{SS}}^{(N)} \end{bmatrix} \mathbf{F}^T. \tag{21}$$

Since (21) corresponds to a matrix representation of a coupled CPD of a set of third-order tensors, the MHR factorization (3) can be computed via any algorithm for coupled CPD. In the next subsection we explain that the SD method in [9] is also suited for the MHR problem.

C. Step 2: From coupled CPD to SD

From (21) we build R symmetric matrices $\{\mathbf{M}^{(r)}\}$, admitting the factorizations

$$\mathbf{M}^{(r)} = \mathbf{G}\boldsymbol{\Lambda}^{(r)}\mathbf{G}^T, \quad r \in \{1, \ldots, R\}, \tag{22}$$

where $\mathbf{G} = \mathbf{F}^{-1}$ and $\boldsymbol{\Lambda}^{(r)} \in \mathbb{C}^{R \times R}$ are diagonal matrices. The construction will be explained below. In the exact case, the SD problem (22) can be solved by means of a GEVD, with the columns of $\mathbf{G}^{-T}$ as generalized

eigenvectors. The GEVD can be computed using the QZ method [1]. The robustness of the computation can be increased via SD techniques that simultaneously exploit all R equations in (22). For instance, the extended QZ method [16] is an attempt to generalize the QZ method to the SD case (22). A further refinement by means of an optimization-based method (e.g., [15], [7]) can lead to a more accurate estimate in the inexact case.

Following the steps in [9] we will now explain how to construct the matrices $\{\mathbf{M}^{(r)}\}$ in (22). Thereafter, we explain that the complexity of the construction can be reduced by exploiting the MHR structure.

Construction of matrices $\{\boldsymbol{M}^{(r)}\}$: 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 $f_R : \mathbb{N}^2 \rightarrow \{1, 2, \ldots, C_R^2\}$ that returns the position of its argument in the set S_R, e.g., $f_R((1,3)) = 3$. Define also $\mathbf{Q}^{(1,n)} \in \mathbb{C}^{R\times R}$, $\mathbf{Q}^{(2,n)} \in \mathbb{C}^{R\times C_R^2}$ and $\mathbf{Q}^{(3,n)} \in \mathbb{C}^{C_R^2\times C_R^2}$ as follows

$$(\mathbf{Q}^{(1,n)})_{r,s} = 8\left(w_{rs}^{(n)} w_{rs}^{(n)} - \mathbf{b}_{rs}^{(n)H}\mathbf{b}_{rs}^{(n)}\right), \tag{23}$$

$$(\mathbf{Q}^{(2,n)})_{r,f_R(s_1,s_2)} = 16(w_{rs_1}^{(n)} w_{rs_2}^{(n)} - \mathbf{b}_{rs_1}^{(n)H}\mathbf{b}_{rs_2}^{(n)}), \tag{24}$$

$$(\mathbf{Q}^{(3,n)})_{f_R(r_1,r_2),f_R(s_1,s_2)} = 4(w_{r_1s_1}^{(n)} w_{r_2s_2}^{(n)} + w_{r_2s_1}^{(n)} w_{r_1s_2}^{(n)} - \mathbf{b}_{r_1s_1}^{(n)H}\mathbf{b}_{r_2s_2}^{(n)} - \mathbf{b}_{r_2s_1}^{(n)H}\mathbf{b}_{r_1s_2}^{(n)}), \tag{25}$$

where

$$\begin{aligned} w_{rs}^{(n)} &:= w_{rs}^{(1,n)} + w_{rs}^{(2,n)}, \\ w_{rs}^{(1,n)} &= \mathbf{u}_r^{(1,n)H}\mathbf{u}_s^{(1,n)}, \qquad (26)\\ w_{rs}^{(2,n)} &= \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(2,n)}, \qquad (27)\\ \mathbf{b}_{rs}^{(n)} &:= [w_{rs}^{(1,n)}, z_{rs}^{(n)}, z_{sr}^{(n)*}, w_{rs}^{(2,n)}]^T, \\ z_{rs}^{(n)} &= \mathbf{u}_r^{(2,n)H}\mathbf{u}_s^{(1,n)}. \qquad (28) \end{aligned}$$

It can be verified that if Theorem II.2 is satisfied for the coupled CPD of $\{\boldsymbol{\mathcal{U}}_{\mathrm{SS,red}}^{(n)}\}$, then

$$\mathbf{Q}^H\mathbf{Q} = \sum_{n=1}^{N}\begin{bmatrix} \mathbf{Q}^{(1,n)} & 2\cdot\mathbf{Q}^{(3,n)} \\ 2\cdot\mathbf{Q}^{(3,n)H} & 4\cdot\mathbf{Q}^{(2,n)} \end{bmatrix} \tag{29}$$

has an R-dimensional kernel and the matrices $\{\mathbf{M}^{(r)}\}$ can be obtained from the kernel of $\mathbf{Q}^H\mathbf{Q}$ [13]. Namely, $\mathbf{M}^{(r)}$ satisfies

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

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

Exploiting MHR structure: Consider the following complementary row-selection matrices of $\underline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}$ and $\overline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}$:

$$\begin{aligned} \underline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)} &= \mathbf{I}_{\prod_{p=1}^{n-1}K_p} \otimes \mathbf{I}_{K_n}(K_n,:) \otimes \mathbf{I}_{\prod_{q=n+1}^{N}I_q}\,, \\ \overline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)} &= \mathbf{I}_{\prod_{p=1}^{n-1}K_p} \otimes \mathbf{I}_{K_n}(1,:) \otimes \mathbf{I}_{\prod_{q=n+1}^{N}I_q}\,. \end{aligned}$$

The selection is complementary in the sense that $\underline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)} + \underline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)} = \overline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)} + \overline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)} = \mathbf{I}_{\prod_{n=1}^{N}K_n}$.

For $n \in \{1, \ldots, N\}$, we define the matrices

$$\widetilde{\mathbf{W}}^{(1,n)} = (\underline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)})^H(\underline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}), \tag{30}$$

$$\widetilde{\mathbf{W}}^{(2,n)} = (\overline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)})^H(\overline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}), \tag{31}$$

where $(\underline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}), (\overline{\mathbf{T}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}) \in \mathbb{C}^{M_n\times R}$. Since the matrices in (30) and (31) are Hermitian, the construction of $\{\widetilde{\mathbf{W}}^{(1,n)}, \widetilde{\mathbf{W}}^{(2,n)}\}$ requires $2C_{R+1}^2(2(\sum_{n=1}^{N}M_n) - N)$ flops[1].

We conclude from (19), (20), (23), (24), (25), (26), (27), (28), (30) and (31) that the computational cost of $\mathbf{Q}^H\mathbf{Q}$ is dominated by the construction of the matrices

$$\mathbf{Y} = \mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)H}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)} \in \mathbb{C}^{R\times R}, \tag{32}$$

$$\mathbf{W}^{(1,n)} = \mathbf{U}^{(1,n)H}\mathbf{U}^{(1,n)} = \mathbf{Y} - \widetilde{\mathbf{W}}^{(1,n)} \in \mathbb{C}^{R\times R}, \tag{33}$$

$$\mathbf{W}^{(2,n)} = \mathbf{U}^{(2,n)H}\mathbf{U}^{(2,n)} = \mathbf{Y} - \widetilde{\mathbf{W}}^{(2,n)} \in \mathbb{C}^{R\times R}, \tag{34}$$

$$\mathbf{Z}^{(n)} = \mathbf{U}^{(2,n)H}\mathbf{U}^{(1,n)} \in \mathbb{C}^{R\times R}, \tag{35}$$

where $n \in \{1, \ldots, N\}$. Observe that the matrices in (32)–(34) are Hermitian, implying that only the upper C_{R+1}^2 entries of $\{\mathbf{Y}, \mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}\}$ need to be computed. Thus, the computation of $\mathbf{Y}$ requires $C_{R+1}^2(2\prod_{n=1}^{N}K_n - 1) = C_{R+1}^2(2M_nK_n - 1)$ flops while the computation of $\{\mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}\}$ requires $2NC_{R+1}^2$ flops. The computation of $\{\mathbf{Z}^{(n)}\}$ requires $R^2(2(\sum_{n=1}^{N}M_n(K_n - 1)) - N)$ flops. Overall, the complexity of the computation of $\{\mathbf{Y}, \mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}$ is of the order of $O((N + 1)(\prod_{n=1}^{N}K_n)R^2)$ flops.

D. Step 3: From SD to single-tone MHR problems

After $\mathbf{G}$ has been found from solving the SD problem (22) we obtain $\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}\mathbf{G}^T$ and $\mathbf{C}_{\mathrm{SS}}$ from relations (16) and (17), respectively. The generators $\{z_{r,n}\}_{n=1}^{N}$ associated with the rth source can now be deduced from the rth column of $\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}\mathbf{G}^T$. In other words, the coupled CPD framework has turned the MHR problem (3) into a set of decoupled single-tone MHR problems. In this subsection we briefly explain that the single-source Direction Of Arrival (DOA) estimation method in [9] can also be used for single-tone MHR. However, any single-tone MHR method can be used.

1) Extraction from $\boldsymbol{U}_{SS}^{(K_1,\ldots,K_N)}\boldsymbol{G}^T$: Denote

$$\underline{\boldsymbol{\Gamma}}^{(n)} = \underline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}\mathbf{G}^T\,, \tag{36}$$

$$\overline{\boldsymbol{\Gamma}}^{(n)} = \overline{\mathbf{S}}_{(n)}^{(K_1,\ldots,K_N)}\mathbf{U}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}\mathbf{G}^T\,. \tag{37}$$

Due to the shift-invariance property $\mathbf{A}(2{:}K_n,:) = \mathbf{A}^{(K_n,n)}(1 : K_n - 1,:)D_1([z_{r,1},\ldots,z_{R,n}])$, we have $\underline{\boldsymbol{\gamma}}_r^{(n)} = \overline{\boldsymbol{\gamma}}_r^{(n)}z_{r,n}$ so that the generator follows from $z_{r,n} = (\overline{\boldsymbol{\gamma}}_r^{(n)H}\underline{\boldsymbol{\gamma}}_r^{(n)})/(\overline{\boldsymbol{\gamma}}_r^{(n)H}\overline{\boldsymbol{\gamma}}_r^{(n)})$. In several applications, such as DOA estimation, $z_{r,n}$ is located on the unit circle. In

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

that case we set $z_{r,n} = (x_1^{(r,n)} + i \cdot x_2^{(r,n)})^2$, in which $\left[\begin{smallmatrix} x_1^{(r,n)} \\ x_2^{(r,n)} \end{smallmatrix}\right] \in \mathbb{R}^2$ corresponds to the eigenvector associated with the dominant eigenvalue of $\left[\begin{smallmatrix} -\mathrm{Re}\{\beta_{r,n}\} & \mathrm{Im}\{\beta_{r,n}\} \\ \mathrm{Im}\{\beta_{r,n}\} & \mathrm{Re}\{\beta_{r,n}\} \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$ with $\beta_{r,n} = \underline{\boldsymbol{\gamma}}_r^{(n)H} \overline{\boldsymbol{\gamma}}_r^{(n)}$ [9].

2) Extraction from $\boldsymbol{\mathcal{U}}_{SS}^{(K_1,\ldots,K_N)} \boldsymbol{G}^T$ and $\boldsymbol{C}_{SS}$: If spatial smoothing is used such that $L_n \geqslant 2$, then we may additionally use $\mathbf{C}_{\mathrm{SS}}$ to extract the generator $z_{r,n}$. Define $\underline{\boldsymbol{\Upsilon}}^{(n)} = (\underline{\mathbf{S}}_{(n)}^{(L_1,\ldots,L_N)} \otimes \mathbf{I}_M)\mathbf{C}_{\mathrm{SS}}$ and $\overline{\boldsymbol{\Upsilon}}^{(n)} = (\overline{\mathbf{S}}_{(n)}^{(L_1,\ldots,L_N)} \otimes \mathbf{I}_M)\mathbf{C}_{\mathrm{SS}}$. Due to the shift-invariance property $\mathbf{A}^{(L_n,n)}(2:L_n,:) = \mathbf{A}^{(L_n,n)}(1:L_n-1,:)D_1([z_{r,1},\ldots,z_{R,n}])$, we have $\underline{\boldsymbol{v}}_r^{(n)} = \overline{\boldsymbol{v}}_r^{(n)} z_{r,n}$. Hence, by replacing $\underline{\boldsymbol{\Gamma}}^{(n)}$ in (36) with $\left[\begin{smallmatrix} \underline{\boldsymbol{\Gamma}}^{(n)} \\ \underline{\boldsymbol{\Upsilon}}^{(n)} \end{smallmatrix}\right]$ and by replacing $\overline{\boldsymbol{\Gamma}}^{(n)}$ in (37) with $\left[\begin{smallmatrix} \overline{\boldsymbol{\Gamma}}^{(n)} \\ \overline{\boldsymbol{\Upsilon}}^{(n)} \end{smallmatrix}\right]$, the generator $z_{r,n}$ can be determined as explained in the preceding paragraph.

E. Summary and discussion of SD method for MHR

The SD procedure is summarized as Algorithm 1. If condition (15) in Theorem II.2 is satisfied, then Algorithm 1 is guaranteed to find the MHR decomposition in the exact case. On the other hand, if condition (15) is not satisfied, then Algorithm 1 does not apply. The reason is that in the latter case we cannot find the R symmetric matrices $\{\mathbf{M}^{(r)}\}$ in (22) from $\mathbf{Q}^H\mathbf{Q}$.

If M is large (compared to R), then the choice of spatial smoothing parameters is not so critical. Thus, for the case where M is large (and $\mathbf{C}$ has full column rank), we may simply set $L_1 = \cdots = L_N = 1$. If M is small, then we may resort to any heuristic used for 1D HR. For instance, in line with the experiments for ESPRIT in [17], [5], one may choose the pairs $\{K_n, L_n\}$ such that the dimensions of the matrices $(\bigodot_{n=1}^N \mathbf{A}^{(K_n,n)})$ and $(\bigodot_{n=1}^N \mathbf{A}^{(L_n,n)} \odot \mathbf{C})$ in $\mathbf{X}^{(K_1,\ldots,K_N)} = (\bigodot_{n=1}^N \mathbf{A}^{(K_n,n)})(\bigodot_{n=1}^N \mathbf{A}^{(L_n,n)} \odot \mathbf{C})^T$ are roughly square.

Note that the difference between Algorithm 1 and the existing VDM-CPD methods is that the former approach tries to take all the tensors in $\{\boldsymbol{\mathcal{U}}_{\mathrm{SS,red}}^{(n)}\}$ into account while the latter approach only considers one of them. In more detail, the VDM-CPD approach solves the MHR problem by first computing the VDM-CPD of $\boldsymbol{\mathcal{U}}_{\mathrm{SS,red}}^{(n)}$ for some $n \in \{1,\ldots,N\}$ (e.g., via [8]). This yields $\mathbf{F}^{-1} = \mathbf{G}$. Once $\mathbf{G}$ is known, the generators $\{z_{r,n}\}$ can easily be determined, e.g., as explained in Subsection III-D.

If the MHR decomposition is computed via an iterative optimization method, then Algorithm 1 can also be used as an initialization procedure. As an example, Algorithm 1 can be used to initialize the Alternating Least Squares (ALS) method for coupled CPD, which alternates between the updates of $\{\mathbf{A}^{(2,n)}\}$, $\{\mathbf{B}_{\mathrm{SS}}^{(n)}\}$ and $\mathbf{F}$, see [13]. See also [7] and references therein for other optimization-based methods.

F. Compressed version of SD method for MHR

If some of the matrices $\mathbf{B}_{\mathrm{SS}}^{(n)}$, say for $n \in S \subseteq \{1,\ldots,N\}$, have full column rank, then it is possible to reduce the

Algorithm 1 SD method for MHR (Theorem II.2).

Input: $\mathbf{X} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{C}^T$ of the form (4).

1. Choose pairs $\{(K_n, L_n)\}_{n=1}^N$ subject to $K_n + L_n = I_n + 1$.
2. Build $\mathbf{X}_{\mathrm{SS}}^{(K_1,\ldots,K_N)}$ in (5) using spatial smoothing.
3. Compute SVD $\mathbf{X}_{\mathrm{SS}}^{(K_1,\ldots,K_N)} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$ (and find R from $\boldsymbol{\Sigma}$).
4. Build $\mathbf{U}_{\mathrm{SS,red}}^{(n)}$ in (18) for $n \in \{1,\ldots,N\}$.
5. Construct $\{\mathbf{Y}, \mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^N$ given by (32)–(35).
6. Construct $\mathbf{Q}^H\mathbf{Q}$ in (29) via $\{\mathbf{Y}, \mathbf{W}^{(1,n)}, \mathbf{W}^{(2,n)}, \mathbf{Z}^{(n)}\}_{n=1}^N$.
7. Obtain matrices $\{\mathbf{M}^{(r)}\}$ from the kernel of $\mathbf{Q}^H\mathbf{Q}$.
8. Solve SD problem $\mathbf{M}^{(r)} = \mathbf{G}\boldsymbol{\Lambda}^{(r)}\mathbf{G}^T$, $r \in \{1,\ldots,R\}$.
9. Compute $\mathbf{N} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{G}^T$.
10. Obtain $\{z_{r,n}\}$ from $\mathbf{N}$ via single-tone MHR.

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

the dimension of the corresponding matrices $\mathbf{U}^{(m,n)}$ in (19)–(20), yielding a smaller matrix in (21). To see this, consider the corresponding tensors $\boldsymbol{\mathcal{U}}_{\mathrm{SS,red}}^{(n)} \in \mathbb{C}^{2\times J_n \times R}$ and the matrix factorization in (18). We compress the tensors $\{\boldsymbol{\mathcal{U}}_{\mathrm{SS,red}}^{(n)}\}_{n\in S}$ to $\boldsymbol{\mathcal{W}}_{\mathrm{SS,red}}^{(n)} \in \mathbb{C}^{2\times R\times R}$, $n \in S$. In detail, let $[\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)}] = \mathbf{B}_{\mathrm{SS}}^{(n)}(\mathbf{A}^{(2,n)} \odot \mathbf{F})^T = \widetilde{\mathbf{U}}^{(n)}\widetilde{\boldsymbol{\Sigma}}^{(n)}\widetilde{\mathbf{V}}^{(n)H}$ denote the compact SVD of $[\mathbf{U}^{(1,n)}, \mathbf{U}^{(2,n)}]$, in which $\widetilde{\mathbf{U}}^{(n)} \in \mathbb{C}^{J_n\times R}$. Then

$$\mathbf{W}^{(m,n)} = \widetilde{\mathbf{U}}^{(n)H}\mathbf{U}^{(m,n)} = \mathbf{H}^{(n)} D_m(\mathbf{A}^{(2,n)})\mathbf{F}^T, \quad m \in \{1,2\}, \tag{38}$$

where $\mathbf{H}^{(n)} = \widetilde{\mathbf{U}}^{(n)H}\mathbf{B}_{\mathrm{SS}}^{(n)} \in \mathbb{C}^{R\times R}$ is nonsingular. Stacking yields

$$\mathbf{W}_{\mathrm{SS,red}}^{(n)} := \left[\begin{array}{c} \mathbf{W}^{(1,n)} \\ \mathbf{W}^{(2,n)} \end{array}\right] = (\mathbf{A}^{(2,n)} \odot \mathbf{H}^{(n)})\mathbf{F}^T, \quad n \in S, \tag{39}$$

which can be used in a reduced-size variant of (21). Note that the compression has destroyed the MHR structures in $\mathbf{U}^{(m,n)}$ and consequently we cannot build $\mathbf{Q}^H\mathbf{Q}$ in (29) by steps 5-6 in Algorithm 1. Instead the compressed SD method constructs $\mathbf{Q}^H\mathbf{Q}$ by an explicit construction via (23)–(25). A benefit of compression is the reduction of the computational complexity of the construction of $\mathbf{Q}^H\mathbf{Q}$. Note that this compression step may also be relevant for the SD method developed in [9].

IV. Application in multirate MHR

So far we have only discussed how to recover the generators $z_{r,n}$ from $\mathcal{X}$. However, in many signal processing applications, $z_{r,n}$ is complex and we are interested in the argument of $z_{r,n}$, denoted by $\arg(z_{r,n})$. Because of the periodicity of the complex exponential, we have $z_{r,n}^{\beta_n} = |z_{r,n}|^{\beta_n} e^{i(\arg(z_{r,n})\beta_n + 2\pi k)}$ for every $k \in \mathbb{Z}$ and $\beta_n \in \mathbb{N}$. The values $\arg(z_{r,n})$ and β_n can be interpreted as the radial frequency of the complex exponential and the interval between two consecutive sampling points, respectively. This leads to the following recovery condition of $\{\arg(z_{r,n})\}$ from $\{z_{r,n}^{\beta_n}\}$:

$$\arg(z_{r,n})\beta_n < 2\pi\ ,\ \forall r \in \{1,\ldots,R\},\ \forall n \in \{1,\ldots,N\}. \tag{40}$$

In other words, the multivariate complex exponentials in $\sum_{r=1}^R \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)} \circ \mathbf{c}_r$ can only be recovered from the

uniformly spaced sampling points $\{x_{i_1,\ldots,i_N}\}$ in (3) if the radial frequencies $\{\arg(z_{r,n})\}$ satisfy condition (40). Due to the uniform sampling in (3), (40) corresponds in fact to the Nyquist sampling condition for multivariate complex exponentials. A way to relax this condition on $\{\arg(z_{r,n})\}$, and thereby extend the MHR problem to sub-Nuquist sampling cases where $\arg(z_{r,n})\beta_n \geqslant 2\pi$, is to employ nonuniform sampling (see for instance the references in [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 ($\arg(z_{r,n})\beta_n \geqslant 2\pi$). This is in contrast to the conventional VDM-CPD approach for MHR (e.g., [2], [3], [4], [14]), which can only support sampling at or above the Nyquist rate.

A. Multirate sampling in cases where $\mathbf{C}$ *has full column rank*

Consider a 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 $\mathbf{a}_r^{(p,n)} = z_{r,n}^{k_{p,n}}[1, z_{r,n}^{\beta_{p,n}}, z_{r,n}^{2\beta_{p,n}}, \ldots, z_{r,n}^{(I_{p,n}-1)\beta_{p,n}}]^T \in \mathbb{C}^{I_{p,n}}$, 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

$$\mathbf{a}_r^{(n)} = [\mathbf{a}_r^{(1,n)T}, \mathbf{a}_r^{(2,n)T}, \ldots, \mathbf{a}_r^{(P_n,n)T}]^T \in \mathbb{C}^{I_n}, \tag{41}$$

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 $\mathbf{a}_r^{(n)}$ in (3) now takes the form (41). This leads to a multirate version of $\mathcal{X}$ in (3):

$$\mathcal{X}_{\text{MR}} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)} \circ \mathbf{c}_r \in \mathbb{C}^{I_1 \times \cdots \times I_N \times M}, \tag{42}$$

with matrix representation (cf. (4)):

$$\mathbf{X}_{\text{MR}} = (\mathbf{A}^{(1)} \odot \cdots \odot \mathbf{A}^{(N)})\mathbf{C}^T \in \mathbb{C}^{(\prod_{n=1}^{N} I_n) \times M}, \tag{43}$$

where 'MR' stands for multirate and the columns of $\mathbf{A}^{(n)}$ in (43) are given by (41). Note that the samples stored in $\mathbf{a}_r^{(n)}$ may not be distinct, i.e., the subsampling and stacking in (41) may cause that $a_r^{(p,n)}(s) = a_r^{(q,n)}(t)$ for some $s \in \{1, \ldots, I_{p,n}\}$ and $t \in \{1, \ldots, I_{q,n}\}$. As an example, if $P_n = 4$, $k_{1,n} = 1$, $k_{2,n} = 13$, $k_{3,n} = 1$, $k_{4,n} = 1$, $\beta_{1,n} = 5$, $\beta_{2,n} = 5$, $\beta_{3,n} = 10$, $\beta_{4,n} = 13$, $\mathbf{a}_r^{(1,n)} = [1\ z_{r,n}^5\ z_{r,n}^{10}\ z_{r,n}^{15}\ z_{r,n}^{20}]^T$, $\mathbf{a}_r^{(2,n)} = [z_{r,n}^{13}\ z_{r,n}^{18}\ z_{r,n}^{23}\ z_{r,n}^{28}]^T$, $\mathbf{a}_r^{(3,n)} = [1\ z_{r,n}^{10}\ z_{r,n}^{20}\ z_{r,n}^{30}]^T$ and $\mathbf{a}_r^{(4,n)} = [1\ z_{r,n}^{13}]^T$, then

$$\mathbf{a}_r^{(n)} = [1\ z_{r,n}^5\ z_{r,n}^{10}\ z_{r,n}^{15}\ z_{r,n}^{20} \mid z_{r,n}^{13}\ z_{r,n}^{18}\ z_{r,n}^{23}\ z_{r,n}^{28} \mid 1\ z_{r,n}^{10}\ z_{r,n}^{20}\ z_{r,n}^{30} \mid 1\ z_{r,n}^{13}]^T \tag{44}$$

with $\mathbf{a}_r^{(n)}(1) = \mathbf{a}_r^{(n)}(10) = \mathbf{a}_r^{(n)}(14) = 1$, $\mathbf{a}_r^{(n)}(3) = \mathbf{a}_r^{(n)}(11) = z_{r,n}^{10}$, $\mathbf{a}_r^{(n)}(5) = \mathbf{a}_r^{(n)}(12) = z_{r,n}^{20}$ and $\mathbf{a}_r^{(n)}(6) = \mathbf{a}_r^{(n)}(15) = z_{r,n}^{13}$.

For every N-tuple $(p_1, \ldots, p_N)$ of sampling rates in the N modes, we can now extract a subtensor $\mathcal{X}_{\text{MR}}^{(p_1,\ldots,p_N)} \in \mathbb{C}^{I_{p_1,1} \times \cdots \times I_{p_N,N} \times M}$, $1 \leqslant p_n \leqslant P_n$ and $1 \leqslant n \leqslant N$, from $\mathcal{X}_{\text{MR}}$ with MHR factorization

$$\mathcal{X}_{\text{MR}}^{(p_1,\ldots,p_N)} = \sum_{r=1}^{R} \mathbf{a}_r^{(p_1,1)} \circ \cdots \circ \mathbf{a}_r^{(p_N,N)} \circ \mathbf{c}_r\,. \tag{45}$$

This leads to the following instances of (4):

$$\mathbf{X}_{\text{MR}}^{(p_1,\ldots,p_N)} = (\mathbf{A}^{(p_1,1)} \odot \cdots \odot \mathbf{A}^{(p_N,N)})\mathbf{C}^T, \tag{46}$$

where $\mathbf{A}^{(p_n,n)} = [\mathbf{a}_1^{(p_n,n)}, \ldots, \mathbf{a}_R^{(p_n,n)}] \in \mathbb{C}^{I_{p_n,n} \times R}$, $1 \leqslant p_n \leqslant P_n$ and $1 \leqslant n \leqslant N$. We can obtain $\mathbf{X}_{\text{MR}}^{(p_1,\ldots,p_N)}$ in (46) from $\mathbf{X}_{\text{MR}}$ in (43) via a row-selection matrix $\mathbf{Z}^{(p_1,\ldots,p_N)} \in \mathbb{C}^{(\prod_{n=1}^{N} I_{p,n}) \times (\prod_{n=1}^{N} I_n)}$:

$$\mathbf{X}_{\text{MR}}^{(p_1,\ldots,p_N)} = \mathbf{Z}^{(p_1,\ldots,p_N)}\mathbf{X}_{\text{MR}}. \tag{47}$$

To summarize, by utilizing several sampling rates, we have gone from working with the single MHR tensor $\mathcal{X}$ in (3) to a set of MHR tensors $\{\mathcal{X}_{\text{MR}}^{(p_1,\ldots,p_N)}\}$ of the form (45). Note that the matrix factorizations (46) of the latter tensors are coupled via $\mathbf{C}$. Hence, the multirate sampled MHR problem corresponds to a coupled CPD problem. For simplicity, we limit the discussion to cases where $\mathbf{C}$ in (46) has full column rank ($M \geqslant R$). The following paragraphs explain how to solve this multirate MHR problem via Algorithm 1.

1) Dimensionality reduction: As in Subsection III-A, we first reduce the $(N+1)$th dimension of the tensor $\mathcal{X}_{\text{MR}}$ in (42). Additionally, we only consider the distinct samples contained in $\mathcal{X}_{\text{MR}}$. Assume that there are $J_n \leqslant I_n$ distinct samples along the n-th mode of $\mathcal{X}_{\text{MR}}$. As an example, if the sampling pattern of the n-th mode is of the form (44), then the tenth, eleventh, twelfth, fourteenth and fifteenth samples along the n-th mode are omitted such that $J_n = I_n - 5$. Let $\widetilde{\mathcal{X}}_{\text{MR}} \in \mathbb{C}^{J_1 \times \cdots \times J_N \times M}$ denote the tensor built from all the distinct samples contained in $\mathcal{X}_{\text{MR}}$ and let $\widetilde{\mathbf{X}}_{\text{MR}} \in \mathbb{C}^{(\prod_{n=1}^{N} J_n) \times M}$ denote its matrix representation. Consider the compact SVD $\widetilde{\mathbf{X}}_{\text{MR}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^H$ and define $\widetilde{\mathbf{U}}_{\text{MR}} := \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}} \in \mathbb{C}^{(\prod_{n=1}^{N} J_n) \times R}$. Then there exists a nonsingular matrix $\mathbf{F}$ such that we obtain the compressed variant of (46):

$$\mathbf{U}_{\text{MR,red}}^{(p_1,\ldots,p_N)} := \widetilde{\mathbf{Z}}^{(p_1,\ldots,p_N)}\widetilde{\mathbf{U}}_{\text{MR}} = (\mathbf{A}^{(p_1,1)} \odot \cdots \odot \mathbf{A}^{(p_N,N)})\mathbf{F}^T, \tag{48}$$

where $\widetilde{\mathbf{Z}}^{(p_1,\ldots,p_N)} \in \mathbb{C}^{(\prod_{n=1}^{N} I_{p,n}) \times (\prod_{n=1}^{N} J_n)}$ is the compressed variant of the row-selection matrix in (47) that takes into account that repeated samples in $\mathcal{X}_{\text{MR}}$ have been omitted.

2) Reduction to SD: A priori, by processing each MHR factorization of the form (48) individually, the multirate MHR problem via coupled CPD can immediately be reduced to a SD problem. In short, applying Steps 4–6 in Algorithm 1 on each $\mathbf{U}_{\text{MR,red}}^{(p_1,\ldots,p_N)}$ we obtain the following multirate MHR version of (29) in Step 6:

$$\mathbf{Q}^H\mathbf{Q} = \sum_{p_1,\ldots,p_N} \begin{bmatrix} \mathbf{Q}^{(1,p_1,\ldots,p_N)} & 2 \cdot \mathbf{Q}^{(3,p_1,\ldots,p_N)} \\ 2 \cdot \mathbf{Q}^{(3,p_1,\ldots,p_N)H} & 4 \cdot \mathbf{Q}^{(2,p_1,\ldots,p_N)} \end{bmatrix}, \tag{49}$$

where $\mathbf{Q}^{(1,p_1,\ldots,p_N)}$, $\mathbf{Q}^{(2,p_1,\ldots,p_N)}$ and $\mathbf{Q}^{(3,p_1,\ldots,p_N)}$ are of the forms (23), (24) and (25), respectively, and built from $\mathbf{U}_{\text{MR,red}}^{(p_1,\ldots,p_N)}$ for every $1 \leqslant p_n \leqslant P_n$ and $1 \leqslant n \leqslant N$. Working with $\mathbf{Q}^H\mathbf{Q}$ given by (49), Steps 7 and 8 in Algorithm 1 now leads to the SD factorization problem (22).

A more elaborate approach is to combine the different sampling rates, as explained next. We will obtain a multirate version of (18) in Subsection III-B. From $\widetilde{\mathbf{U}}_{\text{MR}}$ we extract the matrices

$$\mathbf{T}^{(p_n,n)}\widetilde{\mathbf{U}}_{\text{MR}} = \left(\bigodot_{s=1}^{n-1} \widetilde{\mathbf{A}}^{(s)} \odot \mathbf{A}^{(p_n,n)} \odot \bigodot_{t=n+1}^{N} \widetilde{\mathbf{A}}^{(t)}\right)\mathbf{F}^T, \quad (50)$$

where $1 \leqslant p_n \leqslant P_n$, $1 \leqslant n \leqslant N$, $\mathbf{T}^{(p_n,n)} \in \mathbb{C}^{(\prod_{p=1,p\neq n}^{N} J_n)I_{p_n,n} \times (\prod_{n=1}^{N} J_n)}$ is a row-selection matrix that selects the rows of $\widetilde{\mathbf{U}}_{\text{MR}}$ associated with sampling rate p_n of the n-th mode, and $\widetilde{\mathbf{A}}^{(s)} \in \mathbb{C}^{J_n \times R}$ is a row-reduced version of the factor matrix $\mathbf{A}^{(s)} \in \mathbb{C}^{I_n \times R}$ in (43) consisting only of its distinct row vectors. As in the uniformly sampled case, spatial smoothing produces the tensor $\mathcal{U}_{\text{MR,red}}^{(p_n,n)} \in \mathbb{C}^{2\times(\times_{p=1,p\neq n}^{N} J_p)\times(I_{p,n}-1)\times R}$, the PD of which admits a matrix representation that corresponds to a compressed multirate version of (18):

$$\mathbf{U}_{\text{MR,red}}^{(p_n,n)} := \begin{bmatrix} \mathbf{U}^{(1,p_n,n)} \\ \mathbf{U}^{(2,p_n,n)} \end{bmatrix} = (\mathbf{A}^{(2,p_n,n)} \odot \mathbf{B}^{(p_n,n)})\mathbf{F}^T, \quad (51)$$

where $\mathbf{A}^{(2,p_n,n)} = \mathbf{A}^{(p_n,n)}(1:2,:)$, and

$$\mathbf{U}^{(1,p_n,n)} = \underline{\mathbf{S}}_{(n)}^{(J_1,\ldots,J_{n-1},I_{p_n,n},J_{n+1},\ldots,J_N)}\mathbf{T}^{(p_n,n)}\widetilde{\mathbf{U}}_{\text{MR}}, \quad (52)$$

$$\mathbf{U}^{(2,p_n,n)} = \overline{\mathbf{S}}_{(n)}^{(J_1,\ldots,J_{n-1},I_{p_n,n},J_{n+1},\ldots,J_N)}\mathbf{T}^{(p_n,n)}\widetilde{\mathbf{U}}_{\text{MR}}, \quad (53)$$

$$\mathbf{B}^{(p_n,n)} = \bigodot_{s=1}^{n-1} \widetilde{\mathbf{A}}^{(s)} \odot \mathbf{A}^{(I_{p_n,n}-1,p_n,n)} \odot \bigodot_{t=n+1}^{N} \widetilde{\mathbf{A}}^{(t)}, \quad (54)$$

where $\mathbf{A}^{(I_{p_n,n}-1,p_n,n)} = \mathbf{A}^{(p_n,n)}(1:I_{p_n,n}-1,:)$ and where $\underline{\mathbf{S}}_{(n)}^{(J_1,\ldots,J_{n-1},I_{p_n,n},J_{n+1},\ldots,J_N)}$ and $\overline{\mathbf{S}}_{(n)}^{(J_1,\ldots,J_{n-1},I_{p_n,n},J_{n+1},\ldots,J_N)}$ are row-selection matrices defined according to (7) and (8), respectively. Hence, if condition (13) in Theorem II.1 is satisfied for the coupled CPD of the $\sum_{n=1}^{N} P_n$ tensors in the set $\{\mathcal{U}_{\text{MR,red}}^{(p_n,n)}\}_{1\leqslant n\leqslant N}^{1\leqslant p_n\leqslant P_n}$, then the multirate MHR terms $\{\mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(N)} \circ \mathbf{c}_r\}_{r=1}^{R}$ in (42) are unique. Furthermore, the multirate MHR decomposition in (45) can be computed by following Steps 5–8 in Algorithm 1, using a multirate version of $\mathbf{Q}^H\mathbf{Q}$ in Step 6, i.e.,

$$\mathbf{Q}^H\mathbf{Q} = \sum_{\substack{1\leqslant p_n\leqslant P_n \\ 1\leqslant n\leqslant N}} \begin{bmatrix} \mathbf{Q}^{(1,p_n,n)} & 2\cdot\mathbf{Q}^{(3,p_n,n)} \\ 2\cdot\mathbf{Q}^{(3,p_n,n)H} & 4\cdot\mathbf{Q}^{(2,p_n,n)} \end{bmatrix}, \quad (55)$$

where $\mathbf{Q}^{(1,p_n,n)}$, $\mathbf{Q}^{(2,p_n,n)}$ and $\mathbf{Q}^{(3,p_n,n)}$ are of the form (23), (24) and (25), respectively, and built from $\mathbf{U}_{\text{MR,red}}^{(p_n,n)}$ in (51) for $1 \leqslant p_n \leqslant P_n$ and $1 \leqslant n \leqslant N$.

3) Recovery of generators from multirate sampled single-tone MHR: Note that in step 10 in Algorithm 1 we now have to solve multirate sampled single-tone MHR problems. This can for instance be accomplished by the polynomial rooting procedure presented in [10]. Briefly, if one of the following two conditions

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

is satisfied, then $z_{r,n}$ can be recovered from the minimizer of the univariate polynomial $f(z_{r,n}) = \sum_{m=1}^{P_n} \|\mathbf{U}^{(1,m,n)}\mathbf{g}_r z_{r,n}^{\beta_{m,n}} - \mathbf{U}^{(2,m,n)}\mathbf{g}_r\|_F^2$, where $\mathbf{G} = [\mathbf{g}_1,\ldots,\mathbf{g}_R]$ denotes the nonsingular matrix obtained after solving the SD problem in step 8 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 $f(\psi_{r,n}) = \sum_{m=1}^{P_n} \|\mathbf{U}^{(1,m,n)}\mathbf{g}_r e^{i\cdot\psi_{r,n}\cdot\beta_{m,n}} - \mathbf{U}^{(2,m,n)}\mathbf{g}_r\|_F^2$, which in turn can be found by rooting a polynomial of degree $2\max_{1\leqslant m\leqslant P_n}\beta_{m,n}$. See [10] for details about the polynomial rooting procedure.

4) Compressed version of SD method for multirate MHR: As in the uniformly sampled MHR case, if some of the matrices $\mathbf{B}^{(p_n,n)}$ in (54) have full column rank, then it is possible to reduce the dimension of the corresponding matrices in (52)–(53). Briefly, as in Subsection III-F, let $[\mathbf{U}^{(1,p_n,n)}, \mathbf{U}^{(2,p_n,n)}] = \mathbf{B}^{(p_n,n)}(\mathbf{A}^{(2,p_n,n)} \odot \mathbf{F})^T = \widetilde{\mathbf{U}}^{(p_n,n)}\widetilde{\mathbf{\Sigma}}^{(p_n,n)}\widetilde{\mathbf{V}}^{(p_n,n)H}$ denote the compact SVD of $[\mathbf{U}^{(1,p_n,n)}, \mathbf{U}^{(2,p_n,n)}]$, then

$$\mathbf{W}^{(m,p_n,n)} = \widetilde{\mathbf{U}}^{(p_n,n)H}\mathbf{U}^{(m,p_n,n)} = \mathbf{H}^{(p_n,n)}D_m(\mathbf{A}^{(2,p_n,n)})\mathbf{F}^T, \quad (57)$$

where $\mathbf{H}^{(p_n,n)} = \widetilde{\mathbf{U}}^{(p_n,n)H}\mathbf{B}^{(p_n,n)} \in \mathbb{C}^{R\times R}$ is nonsingular, with $m \in \{1,2\}$. Stacking yields the multirate version of (39):

$$\mathbf{W}_{\text{MR,red}}^{(p_n,n)} := \begin{bmatrix} \mathbf{W}^{(1,p_n,n)} \\ \mathbf{W}^{(2,p_n,n)} \end{bmatrix} = (\mathbf{A}^{(2,p_n,n)} \odot \mathbf{H}^{(p_n,n)})\mathbf{F}^T, \quad (58)$$

which can be used in a reduced-size variant of (51). In the same way as in the uniformly sampled MHR case, the compressed SD method for multirate MHR constructs $\mathbf{Q}^H\mathbf{Q}$ in (55) by an explicit construction via relations (23)–(25) based on the compressed matrices $\{\mathbf{W}_{\text{MR,red}}^{(p_n,n)}\}$.

B. Illustrative example in 2D DOA estimation

Let us demonstrate how 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 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 (46) takes the form

$$\mathbf{X}^{(p_1,p_2)} = (\mathbf{A}^{(p_1,1)} \odot \mathbf{A}^{(p_2,2)})\mathbf{C}^T, \quad (59)$$

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. See [9], [10] and references therein for background. The DOAs $\{\theta_r, \phi_r\}$ can be computed from (59) 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 (4). 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] $= 10\log(\|\mathbf{X}\|_F^2/\|\beta\mathbf{N}\|_F^2)$.

We consider 2D DOA estimation based on URAs. In this case we have $N = 2$, $z_{r,1} = e^{-i\omega_c d_1 \sin(\phi_r)\cos(\theta_r)/c}$ and $z_{r,2} = e^{-i\omega_c d_2 \sin(\phi_r)\sin(\theta_r)/c}$. In the experiments below we fix $\omega_c = 2\pi \cdot 10^9$, $c = 3 \cdot 10^8$ and $d_1 = d_2 = \frac{\pi c}{\omega_c}$ while the azimuth angles $\{\theta_r\}$ and elevation angles $\{\phi_r\}$ are drawn from a uniform distribution over $[0, \pi]$ and $[-\pi/2, \pi/2]$, respectively.

The performance evaluation will be based on the distance between the true DOAs $\{\theta_r, \phi_r\}$ 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 r-th element of the corresponding permuted set.

A. 2D DOA estimation via multiple snapshot 2D HR

We 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 $R = 7$. We compare Algorithm 1 and its compressed version with the basic VDM-CPD approach in [8] (which basically ignores the Vandermonde structure in $\mathbf{A}^{(2)}$), as discussed in Subsection III-E. The methods will be referred to as SD, compressed SD and VDM-CPD, respectively. We also consider the multirate sampled URA presented in Subsection IV-B in which $P_1 = P_2 = 2$ and the columns of $\mathbf{A}^{(p,1)}$ and $\mathbf{A}^{(p,2)}$ in (59) 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 $n \in \{1, 2\}$. Algorithm 1 applied on multirate sampled data 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. All methods except MR-SD and compressed MR-SD apply Forward-Backward Averaging (FBA) in a preprocessing step.[2]

[2]FBA means that $\mathbf{X}$ in (4) is replaced by $[\mathbf{X}, \mathbf{J}\mathbf{X}^*]$, where $\mathbf{J}$ is an antidiagonal matrix with unit entries on its antidiagonal. See [12] for further details and references.

The mean and median $\mathrm{E_{RMS}}$ values over 100 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. The compressed SD method performs better than VDM-CPD. We also observe that the pure SD method for MHR on average performs worse than VDM-CPD. This is due to a few runs in which the matrix $\mathbf{Q}^H\mathbf{Q}$ in (29) is ill-conditioned. The median performance of the SD method is clearly better than that of the VDM-CPD method, as it is not affected by the worst trials. The extra orthonormalization step in Section III-F seems to make the compressed version more robust. Overall, the compressed and multirate versions of SD seem to be the methods of choice for this experiment. 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. This is of course due to the fact that VDM-CPD works with one tensor while the SD methods work with several tensors.

B. 2D DOA estimation via single snapshot 2D HR

We also consider the more difficult single snapshot 2D DOA estimation problem. We assume that the URA model parameters are $I_1 = I_2 = 8$, $M = 1$ and $R = 5$. We compare the SD and compressed SD methods with the MultiDimensional Folding (MDF) method developed in [4] for single-snapshot MHR. The latter method represents a basic VDM-CPD approach for MHR. In accordance with the heuristic mentioned in Subsection III-E, the smoothing parameters for the SD and compressed SD methods are set to $(K_1, K_2) = (5, 5)$. All methods apply FBA in a preprocessing step.

The mean and median $\mathrm{E_{RMS}}$ values over 100 trials as a function of SNR can be seen in Figure 2. We first notice that the compressed SD method performs better than the SD and MDF methods. Again, we observe that the transformation step from coupled CPD in the pure SD method is sensitive to noise. In terms of performance, the compressed SD method seems to be the method of choice for this experiment. Again, since MDF is a basic VDM-CPD method it has a lower complexity than the SD methods.

VI. Conclusion

Using a link between the coupled CPD framework and MHR we have derived an algebraic SD method 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 proposed SD method, can simultaneously exploit the harmonic structure in all dimensions. This implies that coupled CPD based methods work under more relaxed conditions compared to existing VDM-CPD based methods. Furthermore, since the SD method reduces the MHR problem into a GEVD problem, the algorithm can be interpreted as an MHR extension of the ESPRIT method.

We also presented a compressed variant that is more robust against perturbation noise. The SD method for MHR (and its compressed variant) can be used to initialize computationally more expensive optimization-based methods for MHR.

We explained that the coupled CPD framework for MHR 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 SD procedures 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] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed. John Hopkins university press, 2013.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] ——, "Coupled tensor decompositions in array processing," ESAT-STADIUS, KU Leuven, Belgium, Tech. Rep. 13-241, 2014.

[10] ——, "Multiple invariance ESPRIT: A coupled matrix-tensor factorization approach," ESAT-STADIUS, KU Leuven, Belgium, Tech. Rep. 13-242, 2014.

[11] ——, "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.

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

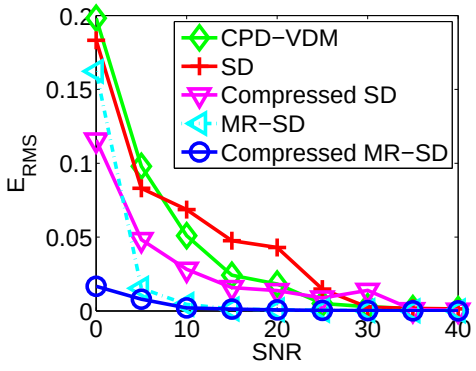
[13] M. Sørensen, I. Domanov, and L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part II: Algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 3, pp. 1015–1045, 2015.

[14] 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.

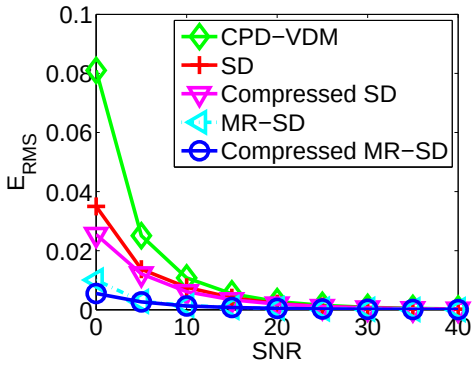
[15] A.-J. van der Veen, "Joint diagonalization via subspace fitting techniques," in *Proc. ICASSP*, May 07-11, Salt Lake City, Utah, USA, 2001.

[16] 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.

[17] 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.



(a) Mean E_{RMS}.



(b) Median E_{RMS}.

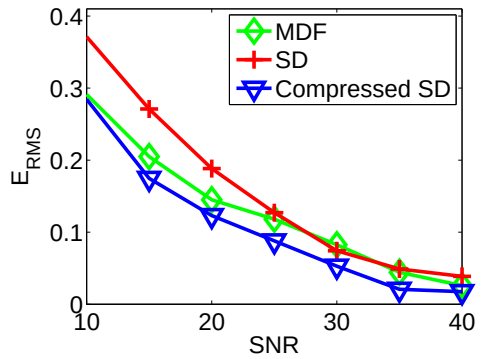
Fig. 1. Mean and median E_{RMS} over 100 trials, multiple snapshot case.



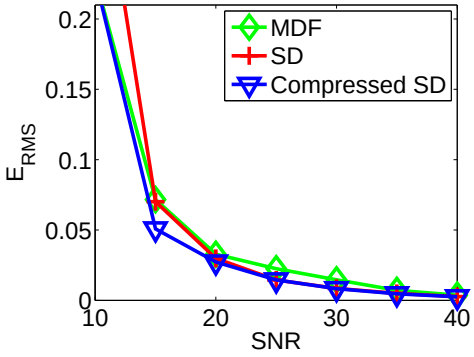
Mikael Sørensen received the Master's degree from Aalborg University, Denmark, and the Ph.D. degree from University of Nice, France, in 2006 and 2010, respectively, both in electrical engineering. Since 2010 he has been a Postdoctoral Fellow with the KU Leuven, Belgium. His research interests include applied linear algebra, tensor decompositions, and tensor-based signal and array processing.



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.



(a) Mean E_{RMS}.



(b) Median E_{RMS}.

Fig. 2. Mean and median E_{RMS} over 100 trials, single snapshot case.