Multi-set low-rank factorizations with shared and unshared components

Mikael Sørensen and Nicholas D. Sidiropoulos, *Fellow, IEEE*

***Abstract*—Low-rank matrix/tensor factorizations play a significant role in science and engineering. An important example is the canonical polyadic decomposition (CPD). There is also a growing interest in multi-set extensions of low-rank matrix/tensor factorizations in which the associated factor matrices are partially shared. In this paper we propose a more unified framework for multi-set matrix/tensor factorizations. In particular, we propose a multi-set extension of bilinear factorizations subject to monomial equality constraints to the case of shared and unshared factors. The presented framework encompasses (generalized) canonical correlation analysis (CCA) and (coupled) CPD models as special cases. CPD, CCA and hybrid models between them feature interesting uniqueness properties. We derive uniqueness conditions for CCA and multi-set low-rank factorization with partially shared entities. Computationally, we reduce multi-set low-rank factorizations with shared and unshared components into a special CPD problem, which can be solved via a matrix eigenvalue decomposition. Finally, numerical experiments demonstrate the importance of taking the coupling between multi-set low-rank factorizations into account in the actual computation.**

***Index Terms*—tensor, canonical polyadic decomposition, canonical correlation analysis, coupled decomposition, monomial, uniqueness, eigenvalue decomposition.**

I. Introduction

Low-rank factorizations of the form

$$\mathbf{X} = \mathbf{M}\mathbf{S}^T \in \mathbb{C}^{I\times K}, \tag{1}$$

where $\mathbf{M} \in \mathbb{C}^{I\times R}$ and $\mathbf{S} \in \mathbb{C}^{K\times R}$, are ubiquitous in science and engineering. In many signal processing and machine learning applications the factor matrix $\mathbf{M} \in \mathbb{C}^{I\times R}$ or $\mathbf{S} \in \mathbb{C}^{K\times R}$ is structured [1]. A classical example is blind separation of wireless communication signals, in which the columns of $\mathbf{M}$ or $\mathbf{S}$ are subject to for instance a constant modulus (CM) constraint (e.g., [2]). Another classical signal processing example that will be considered throughout the paper is the case where $\mathbf{M}$ is Khatri–Rao structured (e.g., [3]):

$$\mathbf{X} = \mathbf{M}\mathbf{S}^T = (\mathbf{A}\odot\mathbf{B})\mathbf{S}^T \in \mathbb{C}^{IJ\times K}, \tag{2}$$

where $\mathbf{M} = \mathbf{A}\odot\mathbf{B} \in \mathbb{C}^{IJ\times R}$ in which $\mathbf{A} \in \mathbb{C}^{I\times R}$ and $\mathbf{B} \in \mathbb{C}^{J\times R}$, and '$\odot$' denotes the Khatri–Rao (columnwise Kronecker) product. When R is minimal, the decomposition (2) is known as the canonical polyadic decomposition (CPD) and will be reviewed in Section I-D. In practice, interfering signals can lead to heterogenous mixtures of the form

$$\mathbf{X} = [\mathbf{M}, \mathbf{C}]\mathbf{S}^T \in \mathbb{C}^{I\times K}, \tag{3}$$

Mikael Sørensen and Nicholas D. Sidiropoulos are with the Dept. of Electrical and Computer Engineering, University of Virginia, 351 McCormick Road, Charlottesville, VA 22904, USA, {ms8tz, nikos}@virginia.edu.

where now $\mathbf{S} \in \mathbb{C}^{K\times(R+Q)}$ and the interference matrix $\mathbf{C} \in \mathbb{C}^{I\times Q}$ has a different structure than $\mathbf{M}$. In the context of blind source separation only a few studies exist for the case where the factorization of $\mathbf{X}$ involves a low-rank structured interference term. We mention that for the special case where $\mathbf{M}$ is Khatri–Rao structured (i.e., the columns of $\mathbf{M}$ are vectorized rank-one matrices), identifiability conditions and algorithms have been presented in [4].

In this paper we will consider multi-set extensions of "single-set" low-rank factorizations of the form (1)–(3), as will be reviewed next.

A. Overview of multi-set low-rank factorizations

In recent years, data fusion and multimodal data analytics based on multi-set low-rank factorizations with shared and unshared components have received considerable attention (e.g., [5], [6], [7], [8], [9]). These works have identified a wide range of promising applications, and convincing experimental / numerical results for several of these applications, thus motivating a deeper look into the more mathematical aspects of multi-set low-rank modeling with shared and unshared components. The one aspect that stands out as deserving a closer look is model identifiability – i.e., conditions under which one can guarantee that the shared and the unshared components can be uniquely identified. Such an analysis is missing from the literature (with the exception of limited results, such as [6]), and it is our goal in this paper to make progress in this direction, with an eye towards a more unified identifiability analysis encompassing several known models and results. The main objective of this paper is to provide a more unified identification framework for multi-set low-rank factorizations with shared and unshared components. In short, we combine and unify the canonical correlation analysis (CCA) and CPD models, leading to a hybrid CCA and CPD model with interesting uniqueness properties. However, before we dive into the details, let us briefly provide an overview and motivate the multi-set low-rank factorization models discussed in this paper.

1) Coupled low-rank factorization with a common factor matrix: The multi-set extension of the basic single-set low-rank model (3) is given by

$$\mathbf{X}^{(n)} = [\mathbf{M}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{I\times K_n}, \quad n \in \{1,\ldots,N\}, \tag{4}$$

where $\mathbf{M} \in \mathbb{C}^{I\times R}$, $\mathbf{C}^{(n)} \in \mathbb{C}^{I\times Q_n}$ and $\mathbf{S}^{(n)} \in \mathbb{C}^{K_n\times(R+Q_n)}$, $n \in \{1,\ldots,N\}$. Note that $\mathbf{M}$ is shared. Data modeling concepts that fit within this framework include partial least squares regression (e.g., [10]) and Joint and Individual Variation Explained (JIVE) analysis [11]. The model (4) is also amenable to canonical correlation analysis (CCA) [12] and Generalized CCA (GCCA) [13].

When only the matrix $\mathbf{M}$ in (4) is of interest, then an alternative model that does not explicitly involve $\mathbf{C}^{(n)}$ can be useful. Namely, by assuming that relation (4) holds, then there exist columnwise orthonormal dimensionality reduction matrices $\mathbf{V}^{(n)} \in \mathbb{C}^{K_n \times R}$, $n \in \{1,\ldots,N\}$ and nonsingular change-of-basis matrices $\mathbf{F}^{(n)} \in \mathbb{C}^{R\times R}$, $n \in \{1,\ldots,N\}$, such that

$$\mathbf{X}^{(n)}\mathbf{V}^{(n)} = \mathbf{M}\mathbf{F}^{(n)T} \in \mathbb{C}^{I\times R}, \quad n \in \{1,\ldots,N\}. \tag{5}$$

In other words, there exists an R-dimensional common subspace between $\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(N)}$ that is spanned by the columns of $\mathbf{M}$. Let range($\mathbf{M}$) denote the range of $\mathbf{M}$. An identifiability condition that guarantees the recovery of range($\mathbf{M}$) from (4) is discussed in [14]. In this paper we are interested in cases where $\mathbf{M}$ is structured, which will lead to a different factorization problem discussed next.

2) Coupled low-rank factorization with a common structured factor matrix: In certain applications the shared factor matrix $\mathbf{M}$ in (4) is structured. One example is cell-edge user detection in wireless communication [15] in which $\mathbf{M}$ is subject to a finite alphabet constraint. Another related signal processing example is blind separation of partially overlapping data packets [16]. To the best of our knowledge, dedicated identifiability conditions and algorithms that *jointly* exploit the coupled low-rank structure between $\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(N)}$ and the structure of the shared factor $\mathbf{M}$ have not been discussed in the literature. In this paper we will provide a more unified framework that can combine the constrained bilinear factorization models (1) and (3) with the multi-set low-rank factorization model (4) in which the columns of the shared matrix $\mathbf{M}$ satisfy M monomial equality constraints, such as

$$m_{\alpha_1,r}\cdots m_{\alpha_L,r} - m_{\beta_1,r}\cdots m_{\beta_L,r} = 0, \tag{6}$$

where $m_{\alpha_l,r}$ denotes the α_l-th entry of the r-th column of $\mathbf{M}$, $m_{\beta_l,r}$ denotes the β_l-th entry of the r-th column of $\mathbf{M}$ and L denotes the degree of the monomials in (6). Compared to the results in [14] that only exploit the coupled low-rank structure, we will show that when the structure of $\mathbf{M}$ is also taken into account, improved identifiability conditions can be obtained.

a) Example 1: To make things more concrete, let us consider the cell-edge user detection problem in [15], in which R transmitted BPSK signals of length I are impinging on N widely separated antenna arrays, each equipped with K_n antennas, so that the observation data matrices $\mathbf{X}^{(n)}$ admit the factorizations (4), where $\mathbf{M} \in \{-1,1\}^{I\times R}$ is the BPSK signal matrix of interest, $\mathbf{C}^{(n)} \in \mathbb{C}^{I\times Q_n}$ is an interference term associated with the n-th antenna array, and $\mathbf{S}^{(n)} \in \mathbb{C}^{K_n\times(R+Q_n)}$ is the channel response matrix associated with the n-th antenna array. In [15] an identifiability condition based on CCA for the two-view case ($N = 2$) that exploits the coupled low-rank structure was proposed. A necessary condition for the CCA based approach in [15] is that $R + Q_1 \leq K_1$, $R + Q_2 \leq K_2$ and $R + Q_1 + Q_2 \leq I$. In this paper we will show that by also exploiting the $\{-1,1\}$-binary structure of $\mathbf{M}$, a more relaxed identifiability condition can be obtained, which allows tolerating (being immune to) more interference signals, or unraveling more cell-edge users. Furthermore, as shown in section V-C, our approach leads to a better algorithm that works well in practice when the prior art fails. More precisely, the finite alphabet property of $\mathbf{M}$ implies that $m_{ir}^2 = 1$. This can also be expressed as a monomial equality constraint:

$$m_{i_1,r}^2 - m_{i_2,r}^2 = 0, \quad 1 \leq i_1 < i_2 \leq I. \tag{7}$$

b) Example 2: In this paper we will focus on bilinear factorizations in which $\mathbf{M}$ is Khatri–Rao structured. The coupled low-rank model (4) with $\mathbf{M} = \mathbf{A} \odot \mathbf{B}$ corresponds to a multi-set extension of the CPD model (2) with shared and unshared components:

$$\mathbf{X}^{(n)} = [\mathbf{A}\odot\mathbf{B}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{IJ\times K_n},\ n \in \{1,\ldots,N\}. \tag{8}$$

The model (8) can be understood as an extension of the coupled low-rank factorization model (4) to the tensorial case, in which $\mathbf{X}^{(n)} \in \mathbb{C}^{IJ\times K_n}$ is a matrix representation of a tensor $\mathcal{X}^{(n)} \in \mathbb{C}^{I\times J\times K_n}$ with rank R. (More details about tensor rank and matrix representations of tensors will be provided in Section I-D). In the context of blind separation of DS-CDMA signals [3], (8) can model a communication system in which several receive antenna arrays are used so that $\mathbf{A}$ is the transmitted symbol matrix of interest, $\mathbf{B}$ is the spreading code matrix used for spectral diversity, and $\mathbf{S}^{(n)}$ and $\mathbf{C}^{(n)}$ are the antenna response matrix and interference matrix associated with the n-th receive antenna array, respectively. In addition to the coupled low-rank structure, the Khatri–Rao structure of $\mathbf{M} = \mathbf{A}\odot\mathbf{B}$ can also be exploited. In Section II-B2 it will be made clear that the latter structure implies that the following monomial equality constraints have to be satisfied:

$$m_{(i_1-1)J+j_1,r}m_{(i_2-1)J+j_2,r} - m_{(i_2-1)J+j_1,r}m_{(i_1-1)J+j_2,r} = 0, \tag{9}$$

where $1 \leq i_1 < i_2 \leq I$ and $1 \leq j_1 < j_2 \leq J$. It is worth noticing that the model (8) encompasses the single-set low-rank CPD model (2) and the coupled low-rank factorization model (4) associated with GCCA. More precisely, when $\mathbf{B} = \mathbf{1}_R^T$ with $\mathbf{1}_R^T = [1,\ldots,1]$, then (8) reduces to (4). Similarly, when $\mathbf{C}^{(n)} = \mathbf{0}$ for all $n \in \{1,\ldots,N\}$, then (8) reduces to (2). In terms of (5), relation (8) can be expressed as

$$\mathbf{X}^{(n)}\mathbf{V}^{(n)} = (\mathbf{A}\odot\mathbf{B})\mathbf{F}^{(n)T} \in \mathbb{C}^{IJ\times R},\ n \in \{1,\ldots,N\}. \tag{10}$$

3) Coupled low-rank factorization with partially shared entities: In some applications, rows of $\mathbf{X}^{(n)}$ in (4) are missing. For example, if individuals evaluate a set of objects using only a subset of available attributes then the involved matrices can have missing rows [17]. Formally, we consider the following extension of (4):

$$\mathbf{X}^{(n)} = [\mathbf{M}^{(n)}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{I_n\times K_n},\ n \in \{1,\ldots,N\}, \tag{11}$$

where the matrices $\mathbf{M}^{(1)} \in \mathbb{C}^{I_1\times R},\ldots,\mathbf{M}^{(N)} \in \mathbb{C}^{I_N\times R}$ are all assumed to be submatrices of the matrix $\mathbf{M} \in \mathbb{C}^{I\times R}$ while the entries of the matrices $\mathbf{C}^{(1)} \in \mathbb{C}^{I_1\times Q_1},\ldots,\mathbf{C}^{(N)} \in \mathbb{C}^{I_N\times Q_N}$ do not necessarily depend on each other. For ease of explanation, we limit the discussion to the tensorial extension of (11) where $\mathbf{M}$ is Khatri–Rao structured and there exist row selection matrices $\mathbf{D}^{(n)} \in \mathbb{C}^{I_n J_n \times IJ}$ with property $\mathbf{D}^{(n)}(\mathbf{A} \odot \mathbf{B}) = \mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}$ such that (cf. Eq. (8)):

$$\mathbf{X}^{(n)} = [\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{I_n J_n \times K_n}, \tag{12}$$

where $n \in \{1, \ldots, N\}$, $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ and $\mathbf{B}^{(n)} \in \mathbb{C}^{J_n \times R}$, and $\mathbf{A}^{(1)} \odot \mathbf{B}^{(1)} \in \mathbb{C}^{I_1 J_1 \times R}, \ldots, \mathbf{A}^{(N)} \odot \mathbf{B}^{(N)} \in \mathbb{C}^{I_N J_N \times R}$ are submatrices of the matrix $\mathbf{A} \odot \mathbf{B} \in \mathbb{C}^{IJ \times R}$. Again, in terms of (5), relation (12) can be expressed as

$$\mathbf{X}^{(n)}\mathbf{V}^{(n)} = (\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)})\mathbf{F}^{(n)T} \in \mathbb{C}^{I_n J_n \times R}. \tag{13}$$

B. Organization and contributions of the paper

The rest of the introduction will present the notation used throughout the paper followed by a brief review of the CPD. As our first contribution, we will in Section II propose a two-step range subspace intersection approach for multi-set low-rank factorizations of the form (4) in which the columns of $\mathbf{M}$ satisfy monomial equality constraints of the form (6). This two-step approach first exploits the common subspace structure of $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ and thereafter exploits the monomial structure of $\mathbf{M}$ in a subsequent step. In particular, we explain that when both the common subspace structure of $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ and the monomial structure of $\mathbf{M}$ are simultaneously exploited, improved identifiability conditions are obtained. It will also be clear that this approach generalizes existing tensor-based methods for bilinear factorizations, such as ACMA [2] and CPD [3], to the multi-set case. As our second contribution, we will in Section III propose a kernel subspace intersection framework that will lead to dedicated identifiability conditions for multi-set monomial factorizations with partially shared entities of the form (11). Part of this work appeared in the conference paper [18]. As our third and final contribution, we will in Section IV explain that the discussed range and kernel subspace intersection approaches lead to algebraic algorithms for the computation of multi-set low-rank factorizations. An optimization based method will also be proposed. In Section V numerical experiments will be reported that demonstrate that for multi-set low-rank factorizations with shared and unshared components, improved performance can be obtained when both the coupled and the monomial equality constrained structure of the decompositions are jointly exploited.

C. Notation

Vectors, matrices and tensors are denoted by lower case boldface, upper case boldface and upper case calligraphic letters, respectively. The r-th column of a matrix $\mathbf{A}$ is denoted by $\mathbf{a}_r$, i.e., $(\mathbf{A})_{ir} = (\mathbf{a}_r)_i$. The conjugate, transpose, conjugate-transpose, determinant, Frobenius norm, inverse, left-inverse, range and kernel of a matrix $\mathbf{A}$ are denoted by $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $|\mathbf{A}|$, $\|\mathbf{A}\|_F$, $\mathbf{A}^{-1}$, $\mathbf{A}^{\dagger}$, range$(\mathbf{A})$, ker$(\mathbf{A})$, respectively. The dimension and orthogonal complement of a subspace S are denoted by $\dim(S)$ and $S^{\perp}$, respectively. The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri–Rao (columnwise Kronecker) product. The Kronecker product of K matrices $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(K)}$ will sometimes be denoted by $\bigotimes_{k=1}^{K} \mathbf{A}^{(k)} = \mathbf{A}^{(1)} \otimes \cdots \otimes \mathbf{A}^{(K)}$. The outer product of, say, three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ is denoted by $\mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$, such that $(\mathbf{a} \circ \mathbf{b} \circ \mathbf{c})_{ijk} = a_i b_j c_k$. The number of non-zero entries of a vector $\mathbf{x}$ is denoted by $\omega(\mathbf{x})$. Let diag$(\mathbf{a}) \in \mathbb{C}^{J \times J}$ denote a diagonal matrix that holds a column vector $\mathbf{a} \in \mathbb{C}^{J \times 1}$ or a row vector $\mathbf{a} \in \mathbb{C}^{1 \times J}$ on its diagonal. Furthermore, let vec$(\mathbf{A})$ denote the vector obtained by stacking the columns of $\mathbf{A} \in \mathbb{C}^{I \times J}$ into a column vector vec$(\mathbf{A}) = [\mathbf{a}_1^T, \ldots, \mathbf{a}_J^T]^T \in \mathbb{C}^{IJ}$. Let $\mathbf{e}_n^{(N)} \in \mathbb{C}^N$ denote the unit vector with unit entry at position n and zeros elsewhere. The identity matrix and all-ones vector are denoted by $\mathbf{I}_m \in \mathbb{C}^{m \times m}$ and $\mathbf{1}_m = [1, \ldots, 1]^T \in \mathbb{C}^m$, respectively. 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 $\mathbf{A}$. The binomial coefficient is denoted by $C_m^k = \frac{m!}{k!(m-k)!}$. The k-th compound matrix of $\mathbf{A} \in \mathbb{C}^{I \times R}$ is denoted by $\mathcal{C}_k(\mathbf{A}) \in \mathbb{C}^{C_I^k \times C_R^k}$. It is the matrix containing the determinants of all $k \times k$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order. Finally, let Sym$^L(\mathbb{C}^R)$ denote the vector space of all symmetric L-th order tensors defined on $\mathbb{C}^R$. The associated set of vectorized ("flattened") versions of the symmetric tensors in Sym$^L(\mathbb{C}^R)$ will be denoted by $\pi_S^{(L,R)}$, i.e., a symmetric tensor $\mathcal{X} \in \mathbb{C}^{R \times \cdots \times R}$ in Sym$^L(\mathbb{C}^R)$ is associated with a vector $\mathbf{x} \in \mathbb{C}^{R^L}$ in $\pi_S^{(L,R)}$.

D. Canonical Polyadic Decomposition (CPD)

Consider the tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$. We say that $\mathcal{X}$ is a rank-one tensor if it is equal to the outer product of nonzero vectors $\mathbf{a} \in \mathbb{C}^I$, $\mathbf{b} \in \mathbb{C}^J$ and $\mathbf{s} \in \mathbb{C}^K$ such that $x_{ijk} = a_i b_j s_k$. A Polyadic Decomposition (PD) is a decomposition of $\mathcal{X}$ into a sum of rank-one terms [19], [20]:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{G}_r \circ \mathbf{s}_r = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{s}_r \in \mathbb{C}^{I \times J \times K}, \tag{14}$$

where $\mathbf{G}_r = \mathbf{a}_r \mathbf{b}_r^T = \mathbf{a}_r \circ \mathbf{b}_r \in \mathbb{C}^{I \times J}$ is a rank-one matrix. The rank of a tensor $\mathcal{X}$ is equal to the minimal number of rank-one tensors that yield $\mathcal{X}$ in a linear combination. Assume that the rank of $\mathcal{X}$ is R, then (14) is called the CPD of $\mathcal{X}$. Let us stack the vectors $\{\mathbf{a}_r\}$, $\{\mathbf{b}_r\}$ and $\{\mathbf{s}_r\}$ into the 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 $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_R] \in \mathbb{C}^{K \times R}$. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{S}$ will be referred to as the factor matrices of the (C)PD of $\mathcal{X}$ in (14).

1) Matrix representation: We will consider the following matrix representation of (14):

$$\mathbf{X} = \left[\text{vec}(\mathbf{G}_1^T), \ldots, \text{vec}(\mathbf{G}_R^T)\right] \mathbf{S}^T = (\mathbf{A} \odot \mathbf{B})\, \mathbf{S}^T \in \mathbb{C}^{IJ \times K}. \tag{15}$$

The rows of $\mathbf{X}$ correspond to the mode-3 fibers $\{\mathbf{x}_{ij\bullet}\}$ of $\mathcal{X}$, defined as $(\mathbf{x}_{ij\bullet})_k = x_{ijk}$. Note that vec$(\mathbf{G}_r^T) = \mathbf{a}_r \otimes \mathbf{b}_r$ in (15) corresponds to a vectorized rank-one matrix.

2) Uniqueness conditions for CPD: The rank-one tensors in (14) can be arbitrarily permuted and the vectors within the same rank-one tensor can be arbitrarily scaled provided the overall rank-one term remains the same. We say that the CPD is unique when it is only subject to these trivial indeterminacies. In this paper we will make use of the relatively easy to check *deterministic* uniqueness condition stated in Theorem 1.1 below.

Theorem 1.1: [21], [22], [23], [24] Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in (14). If

$$\begin{cases} \mathbf{S} \text{ has full column rank,} \\ \mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}) \text{ has full column rank,} \end{cases} \quad (16)$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique.

II. Multi-set low-rank factorizations with shared and unshared components

In this section we will explain how to generalize bilinear factorizations of the form (1) or (3) to the multi-set case

$$\mathbf{X}^{(n)} = [\mathbf{M}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{I\times K_n}, \quad n \in \{1,\ldots,N\}, \quad (17)$$

where the columns of the matrix of interest $\mathbf{M} \in \mathbb{C}^{I\times R}$ can exhibit monomial structure, $\mathbf{C}^{(n)} \in \mathbb{C}^{I\times Q_n}$ are individual interference matrices, and $\mathbf{S}^{(n)} \in \mathbb{C}^{K_n\times(R+Q_n)}$. The goal is to recover the shared factor matrix $\mathbf{M}$, given only the observed matrices $\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(N)}$. Throughout the paper we assume that $\mathbf{M}$ has full column rank, which is a necessary recovery condition. Since we are only interested in the recovery of the full column rank matrix $\mathbf{M}$ it can be verified that a necessary condition is that the matrices $\{[\mathbf{M}, \mathbf{C}^{(n)}]\}$ in (17) have full column rank. For simplicity, we will throughout the paper also assume that the matrices $\{\mathbf{S}^{(n)}\}$ in (17) have full column rank, even though the latter is not necessary. Since we assume that both $[\mathbf{M}, \mathbf{C}^{(n)}]$ and $\mathbf{S}^{(n)}$ have full column rank, we can without loss of generality (w.l.o.g.) also assume that $\mathbf{S}^{(n)}$ is nonsingular (implying that $K_n = R + Q_n$). Compared to the multi-set factorization (4), in which $\mathbf{M}$ is only assumed to be a rank-R matrix, the monomially constrained multi-set factorization (17) allows dependencies between the interference terms $\mathbf{C}^{(m)}$ and $\mathbf{C}^{(n)}$, e.g., $\mathbf{c}_r^{(m)} = \mathbf{c}_s^{(n)}$ for some $m \neq n$ and $r \neq s$ is permitted. In addition, by exploiting the monomial structure of $\mathbf{M}$ more relaxed bounds on R and $Q_1 \ldots, Q_N$ can be obtained, as will be illustrated in Sections II-B and III-B. Compared to the single-set monomially constrained bilinear factorization model (3), the multi-set extension (17) does not prevent the columns of the individual interference terms $\mathbf{C}^{(n)}$ to have the same monomial structure as the columns of $\mathbf{M}$ (e.g., if $\mathbf{M}$ is Khatri–Rao structured, then $\mathbf{C}^{(n)}$ can also be Khatri–Rao structured). Specifically, in Section II-A we propose a basic two-step range subspace intersection method that first exploits the common subspace structure of $\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(N)}$ and thereafter exploits the monomial structure of $\mathbf{M}$.

A. A range subspace intersection approach

In this section we present a two-step range subspace intersection approach for finding $\mathbf{M}$ via the multi-set monomial factorization model (17). The two key assumptions made in this section is that there exist nonzero vectors $\mathbf{w}_r^{(1)},\ldots,\mathbf{w}_r^{(N)}$ such that

$$\mathbf{X}^{(n_1)}\mathbf{w}_r^{(n_1)} = \mathbf{X}^{(n_2)}\mathbf{w}_r^{(n_2)}, \quad 1 \le n_1 < n_2 \le N, \quad (18)$$

$$\mathbf{X}^{(n)}\mathbf{w}_r^{(n)} = \mathbf{m}_r \in \mathcal{M}, \quad 1 \le n \le N, \quad (19)$$

where $\mathcal{M}$ denotes a set of vectors that satisfy M monomial equality constraints of the form (6). Assumption (18) implies that $\mathbf{M} \subseteq \bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})$. Thus, the overall idea is to use assumption (18) to first find the common column subspace $\bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})$ and thereafter use assumption (19) to recover $\mathbf{M}$ from it.

1) Reduction to bilinear factorization subject to monomial equality constraints, possibly with a low-rank interference term, by exploiting common subspace: Observe that when the matrices in the sets $\{[\mathbf{M}, \mathbf{C}^{(n)}]\}$ and $\{\mathbf{S}^{(n)}\}$ have full column rank, then

$$Y := \bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)}) = \text{range}(\mathbf{M}) \oplus C, \quad (20)$$

where '$\oplus$' denotes the direct sum, and

$$C := \bigcap_{n=1}^{N} \text{range}(\mathbf{C}^{(n)}) \cap \text{range}(\mathbf{M})^{\perp}. \quad (21)$$

The dimension of C will be denoted by Q.

a) Basic case $Q = 0$: Since $\text{range}(\mathbf{M}) \subseteq Y$, the minimal dimension of Y is R. This also means that if $\dim(Y) = R$, then $Q = 0$, $C = \{\mathbf{0}\}$ and $Y = \text{range}(\mathbf{M})$. Consequently, if the columns of $\mathbf{Y} \in \mathbb{C}^{I\times R}$ form a basis for Y, then the original multi-set factorization problem (17) can be reduced to a bilinear factorization problem studied in [25]:

$$\mathbf{Y} = \mathbf{M}\mathbf{F}^T, \quad (22)$$

where $\mathbf{F} \in \mathbb{C}^{R\times R}$ is a nonsingular change-of-basis matrix.

b) More general case $Q \ge 0$: Consider now the more challenging case where the subspace Y given by (20) is $(R+Q)$-dimensional, i.e., the subspace C given by (21) is Q-dimensional with $Q \ge 0$. Let the columns of $\mathbf{C} \in \mathbb{C}^{I\times Q}$ form a basis for C. Similarly, let the columns of $\mathbf{Y} \in \mathbb{C}^{I\times(R+Q)}$ form a basis for Y. It is now clear that relation (20) can be expressed in terms of a matrix factorization:

$$\mathbf{Y} = [\mathbf{M}, \mathbf{C}]\mathbf{F}^T, \quad (23)$$

where $\mathbf{F} \in \mathbb{C}^{(R+Q)\times(R+Q)}$ is a nonsingular change-of-basis matrix. Comparing (22) with (23), it is clear that the former is just a special case of the latter in which the low-rank interference term $\mathbf{C}$ is omitted. For this reason, we only consider the matrix factorization (23) in the subsequent discussion where it will be made clear that when the monomial equality constraints (6) are exploited, $\mathbf{M}$ can be recovered from Y, even if $Q \ge 1$.

2) Reduction to CPD by exploiting monomial structure: Assume that the columns of $\mathbf{M}$ satisfy N monomial equality constraints of the form (6). By exploiting this structure, the shared factor $\mathbf{M}$ can be recovered from $\mathbf{Y}$, despite $Q \ge 1$. In more detail, we are now looking for a dimensionality reduction matrix $\mathbf{W} \in \mathbb{C}^{(R+Q)\times R}$ with full column rank and with property

$$\mathbf{Y}\mathbf{W} = [\mathbf{M}, \mathbf{C}]\mathbf{F}^T\mathbf{W} = \mathbf{M} \Leftrightarrow \mathbf{Y}\mathbf{w}_r = \mathbf{m}_r \in \mathcal{M}, \quad 1 \le r \le R, \quad (24)$$

where $\mathcal{M}$ denotes the set of vectors that satisfy the M monomial equality constraints of the form (6). Note that this is only possible if any nontrivial linear combination of the columns of $\mathbf{Y}$ yields a vector that is not contained in $\mathcal{M}$, i.e., $\omega(\mathbf{w}) \geq 2 \Rightarrow \mathbf{Y}\mathbf{w} \notin \mathcal{M}$.[1] More specifically, we are looking for a condition that ensures that any linear combination of the columns of $\mathbf{M}$ and $\mathbf{C}$ has the property

$$[\mathbf{M}, \mathbf{C}]\mathbf{w} \in \mathcal{M} \Rightarrow \omega(\mathbf{w}) \leq 1. \tag{25}$$

By exploiting the monomial equality constraints of the form (6), condition (25) can, under certain conditions, be satisfied. In more detail, from (24) we conclude that $m_{ir} = \mathbf{e}_i^{(I)T}\mathbf{Y}\mathbf{w}_r$. Hence, the combination of (6) and (24) yields

$$\prod_{l=1}^{L} m_{\alpha_{l,m}} - \prod_{l=1}^{L} m_{\beta_{l,m}} = \prod_{l=1}^{L}(\mathbf{e}_{\alpha_{l,m}}^{(I)T}\mathbf{Y}\mathbf{w}_r) - \prod_{l=1}^{L}(\mathbf{e}_{\beta_{l,m}}^{(I)T}\mathbf{Y}\mathbf{w}_r) = \mathbf{p}_L^{(n)T} \cdot (\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r) = 0, \quad r \in \{1, \ldots, R\}, \tag{26}$$

where $\mathbf{p}_L^{(m)} = \bigotimes_{l=1}^{L}(\mathbf{Y}^T\mathbf{e}_{\alpha_{l,m}}^{(I)}) - \bigotimes_{l=1}^{L}(\mathbf{Y}^T\mathbf{e}_{\beta_{l,m}}^{(I)}) \in \mathbb{C}^{(R+Q)^L}$ and subscript 'm' in $\alpha_{l,m}$ and $\beta_{l,m}$ denotes the m-th monomial equality constraint of the form (6). Stacking yields

$$\mathbf{P}^{(M,L)} \cdot (\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r) = \mathbf{0}, \quad r \in \{1, \ldots, R\}, \tag{27}$$

where

$$\mathbf{P}^{(M,L)} = [\mathbf{p}_L^{(1)}, \ldots, \mathbf{p}_L^{(M)}]^T \in \mathbb{C}^{M \times (R+Q)^L}. \tag{28}$$

From (27) we know that there exist at least R linearly independent vectors $\{\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r\}$, each with property $\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r \in \ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}$. Thus, if the dimension of $\ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}$ is minimal (i.e., R) and the columns of $\mathbf{R} \in \mathbb{C}^{(R+Q)^L \times R}$ form a basis for $\ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}$, then there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{R} = (\mathbf{W} \odot \cdots \odot \mathbf{W})\mathbf{F}^T, \tag{29}$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_R] \in \mathbb{C}^{(R+Q) \times R}$ appears L times in (29). Clearly, (29) corresponds to a matrix representation of an $(L+1)$-th order tensor $\mathcal{R} = \sum_{r=1}^{R} \mathbf{w}_r \circ \cdots \circ \mathbf{w}_r \circ \mathbf{f}_r \in \mathbb{C}^{(R+Q) \times \cdots \times (R+Q) \times R}$. Since $\mathbf{W}$ and $\mathbf{F}$ have full column rank, the CPD of $\mathcal{R}$ is unique (see Theorem 1.1 with $\mathbf{A} = \mathbf{W} \odot \cdots \odot \mathbf{W}$, $\mathbf{B} = \mathbf{W}$ and $\mathbf{S} = \mathbf{F}$). This in turn implies the uniqueness of $\mathbf{W}$ and $\mathbf{M}$.

Theorem 2.1 summarizes the obtained uniqueness condition for multi-set low-rank factorizations of the form (17). In words, the first two conditions $\text{rank}(\mathbf{S}^{(n)}) = \text{rank}([\mathbf{M}, \mathbf{C}^{(n)}]) = R + Q_n$ in (30) ensure that $\mathbf{M} \subseteq \bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})$. The third condition tells us that $Q = \dim(\bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})) - R$. Finally, since $\mathbf{w}_r \otimes \cdots \otimes \mathbf{w}_r \in \ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}$ for all $r \in \{1, \ldots, R\}$, the fourth condition $\dim(\ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}) = R$ tells us that $\mathbf{W}$ and $\mathbf{M}$ are unique, as explained above.

[1] If there exists a vector $\mathbf{w}$ with property $\omega(\mathbf{w}) \geq 2$ such that $[\mathbf{M}, \mathbf{C}]\mathbf{w} \in \mathcal{M}$, then a nontrivial linear combination of the columns of $[\mathbf{M}, \mathbf{C}]$ yields an alternative solution (24), i.e., $\mathbf{Y}\mathbf{w}$ does not correspond to a (scaled version) of a column of $\mathbf{M}$, implying that $\mathbf{M}$ in (24) is not unique.

Theorem 2.1: Consider the multi-set low-rank factorization of $\mathbf{X}^{(n)} \in \mathbb{C}^{I \times K_n}$, $n \in \{1, \ldots, N\}$ in (17). If

$$\begin{cases} \mathbf{S}^{(1)}, \ldots, \mathbf{S}^{(N)} \text{ have full column rank}, & \text{(30a)} \\ [\mathbf{M}, \mathbf{C}^{(1)}], \ldots, [\mathbf{M}, \mathbf{C}^{(N)}] \text{ have full column rank}, & \text{(30b)} \\ \bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)}) \text{ is } (R+Q)\text{-dimensional}, & \text{(30c)} \\ \ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)} \text{ is } R\text{-dimensional}, & \text{(30d)} \end{cases}$$

then the shared factor matrix $\mathbf{M}$ is unique.

Similar to Theorem 1.1, conditions (30c) and (30d) in Theorem 2.1 could alternatively have been formulated in terms of the involved factor matrices $\mathbf{M}, \mathbf{C}^{(1)}, \ldots, \mathbf{C}^{(N)}$. A reason to formulate them in terms of the observed data matrices $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ is that Theorem 2.1 now admits a constructive interpretation that allows us to develop algorithms for computing $\mathbf{M}$, as will be explained in Section IV-B.

When the model variables in (17) for a given tuple $\{R, Q_1, \ldots, Q_N\}$ can be considered to have been drawn from an absolutely continuous probability distribution, then Theorem 2.1 can be used to obtain a generic identifiability condition for $\mathbf{M}$. More precisely, recall that an $m \times n$ matrix has rank $p \leq \min(m, n)$ if and only if it has a nonvanishing $p \times p$ minor and all higher-degree minors, if any, vanish. Since a minor is an analytic function, if it is nonzero at one point (one constructive example) then it is nonzero generically (at almost every point except for a set of measure zero). This fact can be used to check the conditions in (30), as briefly explained next.

a) Checking conditions (30a) *and* (30b)*:* From the above discussion we know that $\mathbf{S}^{(n)}$ has full column rank if and only if it contains a nonvanishing $(R+Q_n) \times (R+Q_n)$ minor. This is generically true when $K_n \geq R + Q_n$. By a similar reason we know that $[\mathbf{M}, \mathbf{C}^{(n)}]$ generically has full column rank if $I \geq R + Q_n$.

b) Checking condition (30c)*:* Let the columns of $\mathbf{U}^{(n)}$ form a basis for $\text{range}(\mathbf{X}^{(n)})$. Then the problem of checking if $\dim(\bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})) = R+Q$ amounts to checking if the kernel of the $(NR + \sum_{n=1}^{N} Q_n) \times (NR + \sum_{n=1}^{N} Q_n)$ matrix[2]

$$\mathbf{\Xi} = \begin{bmatrix} (N-1)\mathbf{U}^{(1)H}\mathbf{U}^{(1)} & \cdots & -\mathbf{U}^{(1)H}\mathbf{U}^{(N)} \\ -\mathbf{U}^{(2)H}\mathbf{U}^{(1)} & \ddots & \vdots \\ \vdots & & -\mathbf{U}^{(N-1)H}\mathbf{U}^{(N)} \\ -\mathbf{U}^{(N)H}\mathbf{U}^{(1)} & \cdots & (N-1)\mathbf{U}^{(N)H}\mathbf{U}^{(N)} \end{bmatrix} \tag{31}$$

is $(R+Q)$-dimensional. This is true if and only if the largest nonvanishing minor of $\mathbf{\Xi}$ is of size $(NR + \sum_{n=1}^{N} Q_n - R - Q)$-by-$(NR + \sum_{n=1}^{N} Q_n - R - Q)$.

c) Checking condition (30d)*:* Finally, note that

$$\mathbf{x} \in \ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)} \Leftrightarrow f^{(L)}(\mathbf{x}) \in \ker(\mathbf{P}^{(M,L)}\mathbf{D}^{(L)}),$$

[2] The reason is that $\dim(\ker(\mathbf{\Xi})) = \dim(\bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)}))$. Alternatively, it can be computed via relation $\ker(\sum_{n=1}^{N} \mathbf{P}_{\mathbf{U}^{(n)}}^{\perp}) = \bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})$, where $\mathbf{P}_{\mathbf{U}^{(n)}}^{\perp}$ denotes the projector onto $\text{range}(\mathbf{U}^{(n)})^{\perp}$ [26], i.e., $\dim(\ker(\sum_{n=1}^{N} \mathbf{P}_{\mathbf{U}^{(n)}}^{\perp})) = \dim(\bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)}))$.

where the vector $f^{(L)}(\mathbf{x}) \in \mathbb{C}^{C^L_{R+Q+L-1}}$ consists of all the $C^L_{R+Q+L-1}$ distinct entries of $\bigotimes_{l=1}^L \mathbf{d} = \mathbf{d} \otimes \cdots \otimes \mathbf{d}$ and $\mathbf{D}^{(L)} \in \{0,1\}^{(R+Q)^L \times C^L_{R+Q+L-1}}$ is the "compression" matrix that takes the structure of $\mathbf{x} \in \pi_S^{(L,R+Q)}$ into account, so that $\mathbf{P}^{(M,L)}\mathbf{x} = \mathbf{P}^{(M,L)}\mathbf{D}^{(L)} f^{(L)}(\mathbf{x})$ when $\mathbf{x} \in \pi_S^{(L,R+Q)}$. Hence, the problem of checking if $\dim(\ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}) = R$ amounts to checking if $\dim(\ker(\mathbf{P}^{(M,L)}\mathbf{D}^{(L)})) = R$ for a single instance. In the next section we demonstrate how, for a given tuple $\{R, Q_1, \ldots, Q_N\}$, Theorem 2.1 can be used to obtain a generic identifiability condition for $\mathbf{M}$.

B. Illustrative examples

1) Exploiting coupled low-rank structure leads to improved identifiability conditions: Using Theorem 2.1, we will first demonstrate that improved identifiability conditions can be obtained by taking into account that $\text{range}(\mathbf{M}) \subseteq \cap_{n=1}^N \text{range}(\mathbf{X}^{(n)})$. Consider the multi-set factorization (17) in which $Q_n = 2$, $N = 1$ or $N = 2$ and $\mathbf{M}$ is CM constrained, i.e., the columns of $\mathbf{M}$ satisfies the $M = C_I^2$ monomial equality constraints $m_{i_1 r}m^*_{i_1 r} - m_{i_2 r}m^*_{i_2 r} = 0$, $1 \le i_1 < i_2 \le I$, of degree $L = 2$. This can be understood as a double extension of ACMA for blind separation of CM signals [2], i.e., from $N = 1$ and $Q_n = 0$ to $N \ge 1$ and $Q_n \ge 0$. In Appendix A-1 we explain how to construct $\mathbf{P}^{(M,L)}$ in (30d). (Note that due to the complex conjugation associated with the CM constraint, the requirement that $\ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}$ is R-dimensional in condition (30) now becomes that $\ker(\mathbf{P}^{(M,2)})$ has to be R-dimensional; see Appendix A-1 for details.) In Table I we report upper bounds on R as a function of I when condition (30) is used. By inspection of the table it is clear that a relaxed condition on R can be obtained by exploiting that $\text{range}(\mathbf{M}) \subseteq \text{range}(\mathbf{X}^{(1)}) \cap \text{range}(\mathbf{X}^{(2)})$. In practice, this means that more sources $\mathbf{m}_1, \ldots, \mathbf{m}_R$ from a mixture of CM signals of the form (17) can be unraveled when the common subspace structure $\text{range}(\mathbf{M})$ is exploited.

I	10	20	30	40	50	60
$N = 1$	1	2	3	4	5	6
$N = 2$	3	4	5	6	7	8

TABLE I
AN UPPER BOUND ON R AS A FUNCTION OF I WHEN $\mathbf{M}$ IS CM CONSTRAINED, $N = 1$ AND $Q_1 = 2$ OR $N = 2$ AND $Q_1 = Q_2 = 2$, AND CONDITION (30) IS USED.

*2) Exploiting monomial structure of **M** leads to improved identifiability conditions:* Let us now demonstrate that the bound on the variables $Q_1, \ldots, Q_N$ can be relaxed when the monomial structure of $\mathbf{M}$ is taken into account, i.e., we allow the dimension of the subspace $\bigcap_{n=1}^N \text{range}(\mathbf{X}^{(n)})$ to be greater than R. Consider the multi-set factorization (17) in which $R = 10$, $N = 2$, and $\mathbf{M} = \mathbf{A} \odot \mathbf{B} \in \mathbb{C}^{IJ \times R}$ is Khatri–Rao structured with $\mathbf{A} \in \mathbb{C}^{I \times R}$ and $\mathbf{B} \in \mathbb{C}^{J \times R}$. This implies that the columns of $\mathbf{M}$ satisfy the $M = C_I^2 C_J^2$ monomial equality constraints $(\mathbf{G}_r)_{i_1 j_1}(\mathbf{G}_r)_{i_2 j_2} - (\mathbf{G}_r)_{i_1 j_2}(\mathbf{G}_r)_{i_2 j_1} = 0$, $1 \le i_1 < i_2 \le I$, $1 \le j_1 < j_2 \le J$ of degree $L = 2$, where we recall that $\mathbf{m}_r = \mathbf{a}_r \otimes \mathbf{b}_r = \text{vec}(\mathbf{G}_r)$. This can be understood as a double extension of CPD, i.e., from $N = 1$ and $Q_n = 0$ to $N \ge 1$ and $Q_n \ge 0$. In Appendix A-2 we explain how to construct $\mathbf{P}^{(M,L)}$ in (30d). In Table II we report upper bounds on Q_1 and Q_2 as a function of I and J when condition (30) is used with $Q = 0$ and $Q \ge 0$. More precisely, for a given Q, the triplets $(Q_1, Q_2; Q)$ in Table II indicate the maximal value for the pair (Q_1, Q_2) with property $Q_1 = Q_2$ or $Q_1 = Q_2 - 1$. In cases where $Q \ge 0$, we have $Q = \dim(\bigcap_{n=1}^N \text{range}(\mathbf{X}^{(n)})) - R$. By inspection of the table it is observed that the bound on (Q_1, Q_2) can be improved when $Q \ge 0$ is permitted. In practice, this means that when both the common subspace structure $\text{range}(\mathbf{M})$ and the monomial structure associated with $\mathbf{M}$ are exploited, then more interference signals $\mathbf{c}_1^{(n)}, \ldots, \mathbf{c}_{Q_n}^{(n)}$ are allowed, without affecting the ability of recovering the Kronecker structured sources $\mathbf{a}_1 \otimes \mathbf{b}_1, \ldots, \mathbf{a}_R \otimes \mathbf{b}_R$.

III. EXTENSION TO PARTIALLY SHARED ENTITIES

Note that in Section II it was assumed that the observed matrices $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ have the same row dimension, I, and more importantly that $\text{range}(\mathbf{M}) \subseteq \bigcap_{n=1}^N \text{range}(\mathbf{X}^{(n)})$. In this section we relax both these conditions. In more detail, we explain that the low-rank multi-set factorization framework can be extended to the case of partially shared and unshared components, where the row dimensions of $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ vary, i.e.,

$$\mathbf{X}^{(n)} = [\mathbf{M}^{(n)}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{I_n \times K_n}, \quad 1 \le n \le N, \quad (32)$$

in which $I_m \ne I_n$ for some $m \ne n$ is permitted. In addition, we only require that a subset of the rows of $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$ are shared, i.e., $\mathbf{D}^{(m,n)}\mathbf{M}^{(m)} = \mathbf{D}^{(n,m)}\mathbf{M}^{(n)} \in \mathbb{C}^{F_{m,n} \times R}$ for some a priori known row selection matrices $\mathbf{D}^{(m,n)} \in \mathbb{C}^{F_{m,n} \times I_m}$ and $\mathbf{D}^{(n,m)} \in \mathbb{C}^{F_{m,n} \times I_n}$ in which $F_{m,n}$ denotes the number of shared rows between $\mathbf{M}^{(m)}$ and $\mathbf{M}^{(n)}$. In the context of digital communication, this could happen if different temporal sampling patterns are used for the different views, e.g., because they are sampled at different sampling rates or sampling different and partially overlapping blocks of data. Note that if there is a common "band" of rows in $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$, then we could in principle make use of the range subspace intersection approach discussed in Section II-A, in which the structure associated with the common rows among $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$ is exploited. However, there are two reasons why considering an alternative approach makes good sense. First, since the range subspace intersection approach discussed in Section II-A only exploits the structure associated with the common rows among $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$, it may not be suitable in cases where one pair of matrices $(\mathbf{M}^{(p)}, \mathbf{M}^{(q)})$ with $p \ne q$ share a different subset of rows than another pair of matrices $(\mathbf{M}^{(s)}, \mathbf{M}^{(t)})$ with $s \ne t$. Second, in cases where a subset of rows in a matrix $\mathbf{M}^{(p)}$ is uncommon with any of the matrices in $\{\mathbf{M}^{(n)}\}_{n=1, n\ne p}^N$, this information is not exploited by the range subspace intersection approach discussed in Section II-A. For these two reasons, a kernel subspace intersection approach will be discussed next that can also take the structures associated with both the partially shared rows between $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$ and the uncommon rows of the matrices $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$.

I	6				7			8		9
J	6	7	8	9	7	8	9	8	9	9
$Q = 0$	(12,13;0)	(16,16;0)	(19,19;0)	(22,22;0)	(19,20;0)	(23,23;0)	(26,27;0)	(27,27;0)	(31,31;0)	(35,36;0)
$Q \geq 0$	(18,19;11)	(23,24;15)	(28,28;17)	(33,33;22)	(27,27;15)	(35,35;24)	(40,41;28)	(41,42;29)	(48,48;34)	(55,56; 40)

TABLE II
AN UPPER BOUND ON (Q_1, Q_2) FOR A FIXED Q AS A FUNCTION OF I AND J WHEN $\mathbf{M}$ IS KHATRI–RAO STRUCTURED, $N = 2$, $R = 10$, AND CONDITION (30) IS USED WITH $Q = 0$ AND $Q \geq 0$. THE ENTRIES IN THE TABLE ARE TRIPLETS OF THE FORM $(Q_1, Q_2; Q)$.

A. A kernel subspace intersection approach

We will now propose a kernel subspace intersection approach for multi-set factorizations of the form (32) in which $I_m \neq I_n$ for some $m \neq n$ is permitted. The overall idea is to look for full column rank matrices $\mathbf{W}^{(1)} \in \mathbb{C}^{(R+Q_1)\times R}, \ldots, \mathbf{W}^{(N)} \in \mathbb{C}^{(R+Q_N)\times R}$ whose columns have properties

$$\mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)}\mathbf{w}_r^{(n_1)} - \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)}\mathbf{w}_r^{(n_2)} = \mathbf{0}, \quad (33)$$

$$\mathbf{X}^{(n)}\mathbf{w}_r^{(n)} = \mathbf{m}_r \in \mathcal{M}, \quad (34)$$

where $1 \leq n_1 < n_2 \leq N$ and $1 \leq n \leq N$. Due to the increased complexity of the kernel subspace intersection approach, we will limit the discussion to the case where $\mathbf{M}^{(n)}$ in (32) is Khatri–Rao structured:

$$\mathbf{X}^{(n)} = [\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{I_n J_n \times K_n}, \quad 1 \leq n \leq N, \quad (35)$$

where the matrices in the sets $\{[\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}, \mathbf{C}^{(n)}]\}$ and $\{\mathbf{S}^{(n)}\}$ are assumed to have full column rank. Note that there exist row selection matrices $\mathbf{D}^{(m,n)} \in \mathbb{C}^{F_{m,n}\times I_m J_m}$ and $\mathbf{D}^{(n,m)} \in \mathbb{C}^{F_{m,n}\times I_n J_n}$ with property $\mathbf{D}^{(m,n)}(\mathbf{A}^{(m)} \odot \mathbf{B}^{(m)}) = \mathbf{D}^{(n,m)}(\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)})$. Based on this property we derive a condition stated in Theorem B.1 in Appendix B that ensures the recovery of $\mathbf{A}$ and $\mathbf{B}$ from the multi-set factorization (35). In the next section we demonstrate that this kernel subspace intersection approach can lead to improved identifiability conditions compared to the range subspace intersection approach discussed in Section II-A.

B. Illustrative examples

1) CPD with shared and unshared factors: Let us demonstrate that the bound on the variables $Q_1, \ldots, Q_N$ can be relaxed when the rank-one structures that are not shared between $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$ are also taken into account, i.e., when relation (34) is also taken into account. Consider the same experiment as in Section II-B2 in which $N = 2$, but now we investigate cases where $J_1 < J_2$. More precisely, we consider the case where $\mathbf{X}^{(1)} = [\mathbf{A}^{(1)} \odot \mathbf{B}^{(1)}, \mathbf{C}^{(1)}]\mathbf{S}^{(1)T}$ and $\mathbf{X}^{(2)} = [\mathbf{A}^{(2)} \odot \mathbf{B}^{(2)}, \mathbf{C}^{(2)}]\mathbf{S}^{(2)T}$, in which $\mathbf{A}^{(1)} \odot \mathbf{B}^{(1)} = (\mathbf{I}_{I_2} \otimes \mathbf{S_B})(\mathbf{A}^{(2)} \odot \mathbf{B}^{(2)})$ is a submatrix of $\mathbf{A}^{(2)} \odot \mathbf{B}^{(2)} \in \mathbb{C}^{I_2 J_2 \times R}$, where $\mathbf{A}^{(1)} = \mathbf{A}^{(2)}$ and $\mathbf{S_B} \in \mathbb{C}^{J_1 \times J_2}$ is a row selection matrix that selects the top J_1 rows of $\mathbf{B}^{(2)}$. Note that in this experiment, $\mathbf{M} := \mathbf{M}^{(2)} = \mathbf{A}^{(2)} \odot \mathbf{B}^{(2)}$ is the partially shared factor matrix. In Table III we report upper bounds on Q_1 and Q_2 as a function of (I_1, J_1) and (I_2, J_2) when $N = 2$, $R = 10$ and condition (92) in Theorem B.1 in Appendix B is used. By comparing the entries in Table II where $J_1 = J_2$ with the entries in Table III where $J_1 < J_2$, it is observed that the bound on (Q_1, Q_2) can be improved when also exploiting rank-one structures that are not shared between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$.

(I_1, J_1)	(5,5)	(5,5)	(6,6)	(6,6)	(7,7)
(I_2, J_2)	(5,6)	(5,7)	(6,7)	(6,8)	(7,8)
(Q_1, Q_2)	(11,11)	(12,13)	(20,20)	(21,22)	(30,31)

TABLE III
AN UPPER BOUND ON (Q_1, Q_2) AS A FUNCTION OF (I_1, J_1) AND (I_2, J_2) WHEN $\mathbf{M} := \mathbf{M}^{(2)}$ IS KHATRI–RAO STRUCTURED AND $\mathbf{M}^{(1)}$ IS A SUBMATRIX OF $\mathbf{M}^{(2)}$ (SEE TEXT FOR DETAILS), $N = 2$, $R = 10$, AND CONDITION (92) IN THEOREM B.1 IN APPENDIX B IS USED.

2) Coupled CPD with shared and unshared factors: Let us now show that the kernel subspace intersection approach can exploit the coupling between the matrix factorizations in (35) in cases where the range subspace intersection approach discussed in Section II-A cannot. Consider the following special case of (35) where $\mathbf{A} := \mathbf{A}^{(1)} = \mathbf{A}^{(2)}$ with $N = 2$:

$$\mathbf{X}^{(n)} = [\mathbf{A} \odot \mathbf{B}^{(n)}, \mathbf{C}^{(n)}]\mathbf{S}^{(n)T} \in \mathbb{C}^{IJ_n \times K_n}, \quad n \in \{1, 2\}. \quad (36)$$

Note that (36) corresponds to a coupled CPD [27], [28] with shared and unshared components. Observe that range$(\mathbf{A} \odot \mathbf{B}^{(1)}) \neq$ range$(\mathbf{A} \odot \mathbf{B}^{(2)})$ and consequently the subspace intersection approach discussed in Section II-A cannot exploit the coupling between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. In contrast, using property (34), the kernel subspace intersection approach can exploit the coupling

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_1}^{(J_1)})^T \mathbf{X}^{(1)}\mathbf{w}_r^{(1)} & (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_2}^{(J_2)})^T \mathbf{X}^{(2)}\mathbf{w}_r^{(2)} \\ (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_1}^{(J_1)})^T \mathbf{X}^{(1)}\mathbf{w}_r^{(1)} & (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_2}^{(J_2)})^T \mathbf{X}^{(2)}\mathbf{w}_r^{(2)} \end{vmatrix} = 0, \quad (37)$$

where $1 \leq i_1 < i_2 \leq I$, $1 \leq j_1 \leq J_1$ and $1 \leq j_2 \leq J_2$. The approach discussed here is a special case of the more general kernel subspace intersection approach outlined in Appendix B in which only the rank-one structure between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ is exploited. Assume that the matrices $[\mathbf{A} \odot \mathbf{B}^{(1)}, \mathbf{C}^{(1)}]$ and $[\mathbf{A} \odot \mathbf{B}^{(2)}, \mathbf{C}^{(2)}]$ have full column rank and that the matrices $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ are nonsingular. From (36) we observe that

$$\mathbf{X}^{(n)}\mathbf{w}_r^{(n)} = \mathbf{a}_r \otimes \mathbf{b}_r^{(n)} \in \mathbb{C}^{IJ_n}, \quad r \in \{1, \ldots, R\}. \quad (38)$$

By exploiting the monomial relations

$$\begin{vmatrix} a_{i_1 r} b_{j_1 r}^{(1)} & a_{i_1 r} b_{j_2 r}^{(2)} \\ a_{i_2 r} b_{j_1 r}^{(1)} & a_{i_2 r} b_{j_2 r}^{(2)} \end{vmatrix} = a_{i_1 r} a_{i_2 r}(b_{j_1 r}^{(1)} b_{j_2 r}^{(2)} - b_{j_1 r}^{(1)} b_{j_2 r}^{(2)}) = 0,$$

where $1 \leq i_1 < i_2 \leq I$, $1 \leq j_1 \leq J_1$ and $1 \leq j_2 \leq J_2$, we obtain

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_1}^{(J_1)})^T \mathbf{X}^{(1)}\mathbf{w}_r^{(1)} & (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_2}^{(J_2)})^T \mathbf{X}^{(2)}\mathbf{w}_r^{(2)} \\ (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_1}^{(J_1)})^T \mathbf{X}^{(1)}\mathbf{w}_r^{(1)} & (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_2}^{(J_2)})^T \mathbf{X}^{(2)}\mathbf{w}_r^{(2)} \end{vmatrix}$$
$$= \mathbf{q}^{(s,1,2)}(\mathbf{w}_r^{(1)} \otimes \mathbf{w}_r^{(2)}) = 0, \quad (39)$$

where $\mathbf{q}^{(s,1,2)} = ((\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_1}^{(J_1)})^T \mathbf{X}^{(1)}) \otimes ((\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_2}^{(J_2)})^T \mathbf{X}^{(2)}) - ((\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_1}^{(J_1)})^T \mathbf{X}^{(1)}) \otimes ((\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_2}^{(J_2)})^T \mathbf{X}^{(2)})$, and the superscript '$s$' in $\mathbf{q}^{(s,1,2)} \in \mathbb{C}^{1\times(R+Q_1)(R+Q_2)}$ takes all the subscripts i_1, i_2, j_1 and j_2 into account. Stacking yields

$$\mathbf{Q}^{(M)}(\mathbf{w}_r^{(1)} \otimes \mathbf{w}_r^{(2)}) = \mathbf{0}, \tag{40}$$

where $\mathbf{Q}^{(M)} = [\mathbf{q}^{(1,1,2)T}, \ldots, \mathbf{q}^{(M,1,2)T}]^T \in \mathbb{C}^{M\times(R+Q_1)(R+Q_2)}$ with $M = C_I^2 J_1 J_2$. From (40) it is clear that if $\dim(\ker(\mathbf{Q}^{(M)})) = R$ (which is minimal since $\mathbf{Q}^{(M)}(\mathbf{W}^{(1)} \odot \mathbf{W}^{(2)}) = \mathbf{0}$), then $\mathbf{A} \odot \mathbf{B}^{(n)} = \mathbf{X}^{(n)}\mathbf{W}^{(n)}$, $n \in \{1,2\}$. More precisely, let the columns of the matrix $\mathbf{R} \in \mathbb{C}^{(R+Q_1)(R+Q_2)\times R}$ constitute a basis for $\ker(\mathbf{Q}^{(M)})$. Then there exists a nonsingular change-of-basis matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that we obtain the CPD:

$$\mathbf{R} = (\mathbf{W}^{(1)} \odot \mathbf{W}^{(2)})\mathbf{F}^T. \tag{41}$$

To summarize, if $\ker(\mathbf{Q}^{(M)})$ is R-dimensional, then $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ follows from the unique CPD of (41), which in turn implies that $\mathbf{A}$, $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ can be obtained from $\mathbf{X}^{(1)}\mathbf{W}^{(1)}$ and $\mathbf{X}^{(2)}\mathbf{W}^{(2)}$. As an example, let $R = 6$, $Q_1 = Q_2 = 10$, $I = J_1 = J_2 = 5$ and $K_1 = K_2 = 16$. Then generically $\dim(\ker(\mathbf{Q}^{(M)})) = R$, implying that $\mathbf{A}$, $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ in (36) are generically unique, despite the presence of $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$.

IV. Algorithms

In this section we discuss algorithms for computing the shared components of multi-set low-rank factorizations of the form (17) or (32). For simplicity, we limit the discussion to the special case where $\mathbf{M} = \mathbf{A} \odot \mathbf{B}$. The extension to other monomial structures is analogous. We also mainly limit the discussion to the basic multi-set low-rank factorization of the form (17). In Appendix C an algebraic algorithm for the more general multi-set low-rank factorizations of the form (35) that allows for partially shared entities will be outlined.

A. A least squares fitting approach

Consider the multi-set low-rank factorization of the form (17) with $\mathbf{M} = \mathbf{A}\odot\mathbf{B}$. Let the columns of $\mathbf{U}^{(n)} \in \mathbb{C}^{IJ\times(R+Q_n)}$ form an orthonormal basis for range$(\mathbf{X}^{(n)})$. The identifiability condition (30) in Theorem 2.1 ensures that there exist columnwise orthonormal matrices $\mathbf{V}^{(n)} \in \mathbb{C}^{(R+Q_n)\times R}$ and nonsingular matrices $\mathbf{F}^{(n)} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{U}^{(n)}\mathbf{V}^{(n)} = (\mathbf{A} \odot \mathbf{B})\mathbf{F}^{(n)T}, \quad n \in \{1,\ldots,N\}. \tag{42}$$

In practice, relation (42) is rarely exact and consequently we consider the least squares fitting criterion

$$f(\mathbf{A},\mathbf{B},\{\mathbf{V}^{(n)}\},\{\mathbf{F}^{(n)}\}) = \sum_{n=1}^{N} \|\mathbf{U}^{(n)}\mathbf{V}^{(n)} - (\mathbf{A}\odot\mathbf{B})\mathbf{F}^{(n)T}\|_F^2. \tag{43}$$

Since $\mathbf{U}^{(n)}$ and $\mathbf{V}^{(n)}$ are columnwise orthonormal, minimizing (43) is equivalent to minimizing

$$g(\mathbf{A},\mathbf{B},\{\mathbf{\Gamma}^{(n)}\}) = \sum_{n=1}^{N} \|\mathbf{U}^{(n)} - (\mathbf{A}\odot\mathbf{B})\mathbf{\Gamma}^{(n)T}\|_F^2, \tag{44}$$

where $\mathbf{\Gamma}^{(n)} = \mathbf{V}^{(n)*}\mathbf{F}^{(n)}$. Note that the latter corresponds to a standard least squares CPD fitting problem, implying that if condition (30) is satisfied and the columns of $\mathbf{U}^{(n)}$ form an orthonormal basis for range$(\mathbf{X}^{(n)})$, then $\mathbf{A}$ and $\mathbf{B}$ can be obtained via a standard least squares fitting method for CPD, despite the presence of the interfering terms $\mathbf{C}^{(1)}, \ldots, \mathbf{C}^{(N)}$.

By similar reasoning, if the columns of $\mathbf{U}^{(n)} \in \mathbb{C}^{I_n J_n\times(R+Q_n)}$ form an orthonormal basis for range$(\mathbf{X}^{(n)})$, where $\mathbf{X}^{(n)}$ is given by (35), then in the partially shared factor case, $\mathbf{A}$ and $\mathbf{B}$ can be computed by minimizing the least-squares cost function

$$h(\mathbf{A},\mathbf{B},\{\mathbf{\Gamma}^{(n)}\}) = \sum_{n=1}^{N} \|\mathbf{U}^{(n)} - \mathbf{P}_{\text{sel}}^{(n)}(\mathbf{A}\odot\mathbf{B})\mathbf{\Gamma}^{(n)T}\|_F^2, \tag{45}$$

where $\mathbf{P}_{\text{sel}}^{(n)} \in \mathbb{C}^{I_n J_n\times IJ}$ is a known row selection matrix with property $\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)} = \mathbf{P}_{\text{sel}}^{(n)}(\mathbf{A}\odot\mathbf{B})$.

In the next section an algebraic method for computing $\mathbf{A}$ and $\mathbf{B}$ will be outlined that can be used to initialize an optimization-based method.

B. Range subspace intersection approaches

1) A basic range subspace intersection approach: In Section II-A we alluded that Theorem 2.1 admits a constructive interpretation that can be used to compute the shared factors $\mathbf{A}$ and $\mathbf{B}$, given only $\{\mathbf{X}^{(n)}\}$. Details will now be provided.

Step 1: Using SVD, the first step is to find matrices $\mathbf{U}^{(1)} \in \mathbb{C}^{IJ\times(R+Q_1)}, \ldots, \mathbf{U}^{(N)} \in \mathbb{C}^{IJ\times(R+Q_N)}$ whose columns form orthonormal bases for $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$, respectively.

Step 2: The next step is to find a columnwise orthonormal matrix $\mathbf{U} \in \mathbb{C}^{IJ\times(R+Q)}$ whose columns span $\bigcap_{n=1}^{N}$ range$(\mathbf{U}^{(n)})$. This can be accomplished as follows. Let $Q = \dim(\bigcap_{n=1}^{N} \text{range}(\mathbf{X}^{(n)})) - R$ and let $\mathbf{F} \in \mathbb{C}^{(NR+\sum_{n=1}^{N} Q_n)\times(R+Q)}$ be a columnwise orthonormal matrix with partitioning $\mathbf{F} = [\mathbf{F}^{(1)T}, \ldots, \mathbf{F}^{(N)T}]^T$ and whose submatrices $\mathbf{F}^{(n)} \in \mathbb{C}^{(R+Q_n)\times(R+Q)}$, $n \in \{1,\ldots,N\}$ satisfy the relation

$$[\mathbf{U}^{(1)}\mathbf{F}^{(1)}, \ldots, \mathbf{U}^{(N)}\mathbf{F}^{(N)}] = [\mathbf{A}\odot\mathbf{B}, \mathbf{C}](\mathbf{1}_N^T \otimes \mathbf{G}^T), \tag{46}$$

where $\mathbf{U}^{(n)}\mathbf{F}^{(n)} \in \mathbb{C}^{IJ\times(R+Q)}$ and $\mathbf{G} \in \mathbb{C}^{(R+Q)\times(R+Q)}$ is a nonsingular matrix. Using SVD, the matrix $\mathbf{F}$ with property (46) can be obtained via the subspace $\bigcap_{1\le n_1<n_2\le N} \ker([\mathbf{0}_{I\times\alpha_m}, \mathbf{U}^{(n_1)}, \mathbf{0}_{I\times\beta_m}, -\mathbf{U}^{(n_2)}, \mathbf{0}_{I\times\gamma_m}])$, where $\alpha_m = (n-1)R + \sum_{m=1}^{n-1} Q_m$, $\beta_m = (n-1)R + \sum_{m=1}^{n-1} Q_m$ and $\gamma_m = (n-1)R + \sum_{m=1}^{n-1} Q_m$. For example, the $R+Q$ columns of $\mathbf{F}$ can be chosen to be the $R+Q$ right singular vectors associated with the $R+Q$ smallest singular values of the matrix $\mathbf{\Xi}$ given by (31). Observe that $\bigcap_{n=1}^{N}$ range$(\mathbf{U}^{(n)})$ = range$([\mathbf{U}^{(1)}\mathbf{F}^{(1)}, \ldots, \mathbf{U}^{(N)}\mathbf{F}^{(N)}])$. Thus $\mathbf{U}$ can be obtained from the SVD of $[\mathbf{U}^{(1)}\mathbf{F}^{(1)}, \ldots, \mathbf{U}^{(N)}\mathbf{F}^{(N)}]$. We note in passing that finding $\mathbf{U}$ and $\mathbf{F}$ via $\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(N)}$ is related to generalized CCA [13] and generalized Procrustes analysis [29]. Hence, numerical methods developed to solve these problems can also be used to find $\mathbf{U}$ and $\mathbf{F}$. For example, $\mathbf{U}$ and $\mathbf{F}$ can be computed via the so-called maxvar criterion:

$$\min_{\mathbf{U},\{\mathbf{F}^{(n)}\}} \sum_{n=1}^{N} \left\|\mathbf{U} - \mathbf{U}^{(n)}\mathbf{F}^{(n)}\right\|_F^2. \tag{47}$$

It is well-known that the solution to (47) can be obtained from an eigenvalue decomposition (EVD), i.e., the columns of $\mathbf{U}$ correspond to the $(R+Q)$ eigenvectors associated with the $(R+Q)$ dominant eigenvalues of $\sum_{n=1}^{N}\mathbf{U}^{(n)}\mathbf{U}^{(n)H}$ and that $\mathbf{F}^{(n)}=\mathbf{U}^{(n)H}\mathbf{U}$.

Step 3: The final step is to compute $\mathbf{A}$ and $\mathbf{B}$, given $\mathbf{U}$.

a) Basic case $Q=0$: Let us first consider the basic case where $Q=0$. Since range$(\mathbf{U})=$ range$(\mathbf{A}\odot\mathbf{B})$, there exists a nonsingular change-of-basis matrix $\mathbf{\Gamma}\in\mathbb{C}^{R\times R}$ such that

$$\mathbf{U}=(\mathbf{A}\odot\mathbf{B})\mathbf{\Gamma}^{T}\in\mathbb{C}^{IJ\times R}. \tag{48}$$

Clearly, relation (48) corresponds to a CPD (cf. Eq. (15)). Hence, $\mathbf{A}$ and $\mathbf{B}$ can be obtained from CPD of $\mathbf{U}$. See [1], [30] and references therein for CPD algorithms.

b) More general case $Q\geq 0$: Consider now the case where dim(range$(\mathbf{U})$) $=R+Q$ with $Q\geq 0$. This implies that there exists a nonsingular change-of-basis matrix $\mathbf{\Gamma}\in\mathbb{C}^{(R+Q)\times(R+Q)}$ and a full column matrix $\mathbf{C}\in\mathbb{C}^{IJ\times Q}$ such that

$$\mathbf{U}=[\mathbf{A}\odot\mathbf{B},\mathbf{C}]\mathbf{\Gamma}^{T}\in\mathbb{C}^{IJ\times(R+Q)}. \tag{49}$$

Assume that condition (30) in Theorem 2.1 is satisfied, then there exists a columnwise orthonormal matrix $\mathbf{V}\in\mathbb{C}^{(R+Q)\times R}$ and a nonsingular matrix $\mathbf{F}\in\mathbb{C}^{R\times R}$ such that

$$\mathbf{U}\mathbf{V}=(\mathbf{A}\odot\mathbf{B})\mathbf{F}^{T}\in\mathbb{C}^{IJ\times R}. \tag{50}$$

Since $\mathbf{U}$ and $\mathbf{V}$ are columnwise orthonormal, then, similar to (44), $\mathbf{A}$ and $\mathbf{B}$ can be obtained by minimizing the least squares cost function

$$g(\mathbf{A},\mathbf{B},\mathbf{\Gamma})=\|\mathbf{U}-(\mathbf{A}\odot\mathbf{B})\mathbf{\Gamma}^{T}\|_F^2, \tag{51}$$

where $\mathbf{\Gamma}=\mathbf{V}^{*}\mathbf{F}\in\mathbb{C}^{(R+Q)\times R}$. In Section IV-B2 an algebraic method for computing $\mathbf{A}$ and $\mathbf{B}$ via (49) will be discussed.

An outline of the basic range subspace intersection approach for multi-set low-rank factorizations with shared and unshared factors is given as Algorithm 1. The algorithm is based on a constructive use of Theorem 2.1 when $\mathbf{M}=\mathbf{A}\odot\mathbf{B}$. Note that if $Q=0$, then the problem of finding $\mathbf{A}$ and $\mathbf{B}$ reduces to computing the CPD given by (48). Hence, when $Q=0$ and condition (30) in Theorem 2.1 is satisfied, then Algorithm 1 is guaranteed to find $\mathbf{A}$ and $\mathbf{B}$ in the exact case. The dominant cost of Algorithm 1 is the computation of $\mathbf{U}$, which can be obtained via the EVD of $\mathbf{\Xi}$ given by (31) with computational cost $\mathcal{O}((NR+\sum_{n=1}^{N}Q_n)^3)$ or via the EVD of $\sum_{n=1}^{N}\mathbf{U}^{(n)}\mathbf{U}^{(n)H}$ with computational cost $\mathcal{O}((IJ)^3)$. As mentioned earlier, any method, including a cheaper one, for computing the common subspace of $\{\mathbf{X}^{(n)}\}$ can be used.

2) An algebraic range subspace intersection approach: An algebraic version of Algorithm 1 can be extended to cases where dim$(\bigcap_{n=1}^{N}$ range$(\mathbf{X}^{(n)}))$ $=R+Q$ with $Q\geq 0$. In short, let $\mathbf{U}\in\mathbb{C}^{IJ\times(R+Q)}$ be the same matrix as in Algorithm 1. Using the $M=C_I^2C_J^2$ monomial equality constraints $m_{(i_1-1)J_n+j_1,r}m_{(i_2-1)J_n+j_2,r}-m_{(i_2-1)J_n+j_1,r}m_{(i_1-1)J_n+j_2,r}=0$, $1\leq i_1<i_2\leq I$, $1\leq j_1<j_2\leq J$ of degree $L=2$ that the columns of $\mathbf{M}=\mathbf{A}\odot\mathbf{B}$ exhibit, we build the matrix $\mathbf{P}^{(M,2)}\in\mathbb{C}^{M\times(R+Q)^2}$ given by (28) with $L=2$ and $M=C_I^2C_J^2$. From $\ker(\mathbf{P}^{(M,2)})\cap\pi_S^{(2)}$, $\mathbf{W}$ can be obtained. In more details, the next step is to find a matrix $\mathbf{R}$ whose columns form an orthonormal basis for $\ker(\mathbf{P}^{(M,2)})\cap\pi_S^{(2)}$. Let the columns of $\mathbf{P}^{(M,2)}$ be indexed by the pairs $\{(r_1,r_2)\}_{1\leq r_1,r_2\leq R+Q}$, and ordered lexicographically:

$$\mathbf{P}^{(M,2)}=[\mathbf{p}_{(1,1)}^{(M,2)}\mathbf{p}_{(1,2)}^{(M,2)},\ldots,\mathbf{p}_{(R+Q,R+Q)}^{(M,2)}]. \tag{52}$$

Let the matrix $\mathbf{P}^{(\text{sym})}\in\mathbb{C}^{M\times C_{R+Q+1}^2}$ be constructed from the columns $\{\mathbf{p}_{(r_1,r_2)}^{(M,2)}\}_{r_1\leq r_2}$ as follows

$$\mathbf{P}^{(\text{sym})}=[\mathbf{p}_{(1,1)}^{(M,2)},2\mathbf{p}_{(1,2)}^{(M,2)},\ldots,2\mathbf{p}_{(1,R+Q)}^{(M,2)},\mathbf{p}_{(2,2)}^{(M,2)},2\mathbf{p}_{(2,3)}^{(M,2)},\ldots,2\mathbf{p}_{(2,R+Q)}^{(M,2)},\ldots,\mathbf{p}_{(R+Q,R+Q)}^{(M,2)}]. \tag{53}$$

Note that the columns $\{\mathbf{p}_{(r_1,r_2)}^{(M,2)}\}_{r_1<r_2}$ are scaled by a factor two. We now have that

$$\mathbf{w}\otimes\mathbf{w}\in\ker(\mathbf{P}^{(M,2)})\cap\pi_S^{(2)}\Leftrightarrow\mathbf{f}^{(2)}(\mathbf{w})\in\ker(\mathbf{P}^{(\text{sym})}), \tag{54}$$

where $\mathbf{f}^{(2)}(\mathbf{w})\in\mathbb{C}^{C_{R+Q+1}^2}$ is a structured vector of the form

$$\mathbf{f}^{(2)}(\mathbf{w})=[w_1w_1,w_1w_2,\ldots,w_{R+Q}w_{R+Q}]^T. \tag{55}$$

In words, $\mathbf{f}^{(2)}(\mathbf{w})$ consists of all distinct entries of $\mathbf{w}\otimes\mathbf{w}$. Hence, if condition (30) in Theorem 2.1 is satisfied, then dim$(\ker(\mathbf{P}^{(\text{sym})}))=R$. Let the columns of $\mathbf{V}^{(\text{sym})}$ form a basis for $\ker(\mathbf{P}^{(\text{sym})}$ and obtained via the SVD of $\mathbf{P}^{(\text{sym})}$. Then

$$\mathbf{R}=\mathbf{D}_{R+Q}\cdot\mathbf{V}^{(\text{sym})}=(\mathbf{W}\odot\mathbf{W})\mathbf{\Theta}^T\in\mathbb{C}^{(R+Q)^2\times R}, \tag{56}$$

where $\mathbf{D}_{R+Q}\in\mathbb{C}^{(R+Q)^2\times C_{R+Q+1}^2}$ is the duplication matrix with property $\mathbf{w}\otimes\mathbf{w}=\mathbf{D}_{R+Q}\cdot\mathbf{f}^{(2)}(\mathbf{w})$, $\mathbf{\Theta}\in\mathbb{C}^{R\times R}$ is a nonsingular change-of-basis matrix, and $\mathbf{W}=[\mathbf{w}_1,\ldots,\mathbf{w}_R]\in\mathbb{C}^{(R+Q)\times R}$ is the full column rank matrix of interest. Clearly, relation (56) corresponds to a CPD (cf. Eq. (15)). Hence, $\mathbf{W}$ can be obtained from CPD of $\mathbf{R}$. Now that $\mathbf{W}$ has been obtained, $\mathbf{A}$ and $\mathbf{B}$ follow from rank-one factorizations of the columns of $\mathbf{U}\mathbf{W}=\mathbf{A}\odot\mathbf{B}$.

Algorithm 1 A range subspace intersection approach for multi-set low-rank factorizations of the form (17) with $\mathbf{M}=\mathbf{A}\odot\mathbf{B}$ and based on a constructive use of Theorem 2.1.

Input: $\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(N)}$, R, Q, $Q_1,\ldots,Q_N$.

1. Use SVD to obtain matrix $\mathbf{U}^{(n)}\in\mathbb{C}^{IJ\times(R+Q_n)}$ whose columns form an orthonormal basis for range$(\mathbf{X}^{(n)})$, $\forall n$.
2. Obtain columnwise orthonormal matrix $\mathbf{U}\in\mathbb{C}^{IJ\times(R+Q)}$ whose columns form an orthonormal basis for $\bigcap_{n=1}^{N}$ range$(\mathbf{U}^{(n)})$ via an EVD of $\mathbf{\Xi}$ given by (31) or via an EVD of $\sum_{n=1}^{N}\mathbf{U}^{(n)}\mathbf{U}^{(n)H}$.
3. Obtain $\mathbf{A}$ and $\mathbf{B}$ from the CPD of $\mathbf{U}$ given by (48) when $Q=0$ and (50) otherwise.

Output: $\mathbf{A}$ and $\mathbf{B}$.

An outline of the more advanced range subspace intersection approach for multi-set low-rank factorizations with shared and unshared factors is given as Algorithm 2. The only difference between Algorithm 1 and Algorithm 2 is that the latter is still guaranteed to find $\mathbf{A}$ and $\mathbf{B}$ in the exact case when $Q>0$ and condition (30) in Theorem 2.1 is satisfied. The complexity of Algorithm 2 is dominated by the computation of $\mathbf{U}$, as in Algorithm 1, and the computation of the SVD of $\mathbf{P}^{(\text{sym})}$ given by (53). The latter can be obtained by means of an EVD with computational cost $\mathcal{O}((C^2_{R+Q+1})^3)$.

Algorithm 2 A range subspace intersection approach for multi-set low-rank factorizations of the form (17) with $\mathbf{M} = \mathbf{A} \odot \mathbf{B}$ and based on a constructive use of Theorem 2.1.

Input: $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$, R, Q, $Q_1, \dots, Q_N$.

1. Use SVD to obtain matrix $\mathbf{U}^{(n)}$ whose columns form an orthonormal basis for range$(\mathbf{X}^{(n)})$, $n \in \{1, \dots, N\}$.
2. Obtain columnwise orthonormal matrix $\mathbf{U} \in \mathbb{C}^{IJ \times (R+Q)}$ whose columns form an orthonormal basis for $\bigcap_{n=1}^{N}$ range$(\mathbf{U}^{(n)})$ via an EVD of $\boldsymbol{\Xi}$ given by (31) or via an EVD of $\sum_{n=1}^{N} \mathbf{U}^{(n)}\mathbf{U}^{(n)H}$.
3. Build matrix $\mathbf{P}^{(M,2)}$ given by (52) from $\mathbf{U}$.
4. Use SVD to obtain matrix $\mathbf{R}$ given by (56) whose columns form an orthonormal basis for $\ker(\mathbf{P}^{(M,2)}) \cap \pi_S^{(2)}$.
5. Obtain $\mathbf{W}$ from CPD of $\mathbf{R}$.
6. Obtain $\mathbf{A}$ and $\mathbf{B}$ from rank-one factorizations of the columns of $\mathbf{UW}$.

Output: $\mathbf{A}$ and $\mathbf{B}$.

V. Numerical experiments

A. Exploiting both monomial and common subspace structures can lead to improved performance

Consider (8) with $N = 2$, $I = J = 10$, $K = 50$, $R = 10$ and $Q_1 = Q_2 = 30$. The matrices $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are perturbed by additive noise so that observed matrices are of the form $\mathbf{Y}^{(n)} = \mathbf{X}^{(n)} + \mathbf{N}^{(n)}$, where $\mathbf{N}^{(n)}$ is an unstructured perturbation matrix. In each trial of the Monte Carlo experiment, the entries of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}^{(n)}$, $\mathbf{S}^{(n)}$ and $\mathbf{N}^{(n)}$ are randomly drawn from a Gaussian distribution with zero mean and unit variance. The following Signal-to-Noise Ratio (SNR) measure will be used: SNR $= 10\log_{10}(\sum_{n=1}^{N} \|\mathbf{X}^{(n)}\|_F^2 / \sum_{n=1}^{N} \|\mathbf{N}^{(n)}\|_F^2)$. As a performance measure we use the distance between $\mathbf{A}$ and its estimate, $\widehat{\mathbf{A}}$. The distance is measured according to

$$\mathrm{P}(\mathbf{A}) = \min_{\boldsymbol{\Pi}\boldsymbol{\Lambda}} \|\mathbf{A} - \widehat{\mathbf{A}}\boldsymbol{\Pi}\boldsymbol{\Lambda}\|_F / \|\mathbf{A}\|_F \,, \tag{57}$$

where $\boldsymbol{\Pi}$ and $\boldsymbol{\Lambda}$ denote a permutation matrix and a diagonal matrix, respectively. Note that when $\boldsymbol{\Lambda}$ in (57) is fixed, then the problem of finding $\mathrm{P}(\mathbf{A})$ amounts to solving a linear assignment problem, which can for instance be solved using the Hungarian method. However, since $\boldsymbol{\Lambda}$ is unknown, the function `cpderr.m` in Tensorlab [30] is used for the computation of $\mathrm{P}(\mathbf{A})$, where $\boldsymbol{\Pi}$ and $\boldsymbol{\Lambda}$ are numerically found using a greedy least squares matching algorithm [3].

We compare a standard least squares fitting CPD method that ignores $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$ when computing $\mathbf{A}$ via the CPD of $[\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]$ with the least squares fitting approach described in Section IV-A in which $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are replaced with the columnwise orthonormal matrices $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$. The former method will be referred to as 'CPD' while the latter method will be referred to as 'ONB-CPD', where ONB stands for orthonormal basis. In both cases the function `cpd.m` in Tensorlab [30] is used for the CPD computation. We will also consider Algorithms 1 and 2 associated with Theorem 2.1. Finally, we consider ONB-CPD initialized by Algorithm 1, which will be referred to as 'Algorithm 1+ONB-CPD'.

In this experiment we know that in the noiseless case we have $Q = 0$, implying that in the exact case both Algorithms 1 and 2 are guaranteed to perfectly recover $\mathbf{A}$. The former algorithm assumes that $Q = 0$ while the latter algorithm only assumes that $Q \geq 0$. To demonstrate that a good performance can be obtained despite overestimated Q values, we will consider Algorithm 2 with $Q = 0$ and $Q = 3$. The mean $\mathrm{P}(\mathbf{A})$ values over 100 Monte Carlo runs can be seen in Figure 1. We observe that Algorithms 1 and 2 perform about the same. We also observe that CPD does not perform well and that 'ONB-CPD' is sensitive w.r.t. initialization. Indeed, 'Algorithm 1+ONB-CPD' performed slightly better than Algorithms 1 and 2.

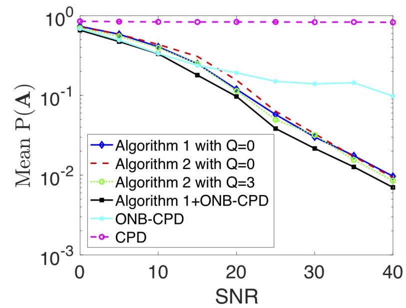

Fig. 1. Mean $\mathrm{P}(\mathbf{A})$ values over 100 Monte Carlo runs.

B. Exploiting both shared and individual rows can lead to improved performance

Consider (12) with $N = 3$, $I_n = J_n = 5$, $n \in \{1, 2, 3\}$, $K = 50$, $R = 3$ and $Q_1 = Q_2 = Q_2 = 3$. We set $\mathbf{A}^{(1)}(1:3,:) = \mathbf{A}^{(2)}(1:3,:) = \mathbf{A}^{(3)}(1:3,:)$ and $\mathbf{B}^{(1)}(1:3,:) = \mathbf{B}^{(2)}(1:3,:) = \mathbf{B}^{(3)}(1:3,:)$. In words, $\mathbf{A}^{(1)} \odot \mathbf{B}^{(1)}$, $\mathbf{A}^{(2)} \odot \mathbf{B}^{(2)}$ and $\mathbf{A}^{(3)} \odot \mathbf{B}^{(3)}$ only have 9 rows (out of 57) in common. To demonstrate that improved performance can be obtained when all 57 rows in the noisy observation matrices $\{\mathbf{Y}^{(n)}\}$ are exploited, we compare Algorithm 1, which only exploits the 9 shared rows, with Algorithm 3, which takes both the shared and individual rows into account. The mean $\mathrm{P}(\mathbf{A})$ values over 100 Monte Carlo runs can be seen in Figure 2. By inspection of the figure it is clear that Algorithm 3 performs much better than Algorithm 1. We also observe that the refinement step by 'ONB-CPD' did not improve the performance.

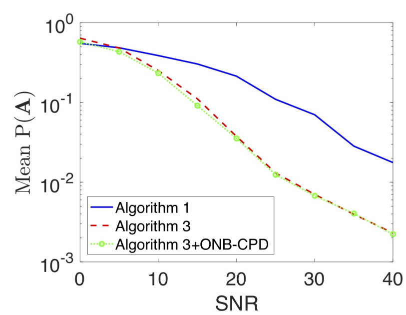

Fig. 2. Mean $\mathrm{P}(\mathbf{A})$ values over 100 Monte Carlo runs.

C. BPSK cell-edge signal separation

Consider the BPSK signal separation problem discussed in Section I-A2 in which $N = 2$, $R = 5$, $Q_1 = Q_2 = 24$, $I = 50$, $K_1 = K_2 = 60$ and the columns of $\mathbf{M}$ satisfy monomial equality constraints of the form (7). We fix the signal-to-interference noise ratio to -20 dB, i.e, $20\log_{10}(\|\mathbf{m}_r\|_F/\|\mathbf{c}_r^{(n)}\|_F) = -20$, $\forall r, s$. We compare the combined CCA-RACMA method proposed in [15] with an adapted version of Algorithm 2 with $Q = \dim(\cap_{n=1}^{2}\text{range}(\mathbf{X}^{(n)})) - R = 3$ or $Q = 4$ for BPSK signals. The goal is to show that BPSK signal separation is possible even if the CCA identifiability conditions are not satisfied. In the CCA-RACMA method in [15] it is assumed that $R + Q_1 + Q_2 \leq I$, which is not the case in this experiment. The Matlab built-in function `canoncorr.m` is used for CCA computation while the RACMA method in [31] is used for BPSK signal separation. For the BPSK variant of Algorithm 2 we first compute columnwise orthonormal matrices $\mathbf{U}^{(n)} \in \mathbb{C}^{I\times(R+Q_n)}$, $n \in \{1, 2\}$ via SVD such that in the noiseless case we have $\text{range}(\mathbf{U}^{(n)}) = \text{range}(\mathbf{X}^{(n)})$, $n \in \{1, 2\}$. Next, we compute the SVD of $\mathbf{U}^{(1,2)} = [\mathbf{U}^{(1)}, \mathbf{U}^{(2)}] \in \mathbb{C}^{I\times(2R+Q_1+Q_1)}$.[3] Let the columns of $\mathbf{U} \in \mathbb{C}^{I\times(R+Q)}$ form an orthonormal basis for $\text{range}(\mathbf{U}^{(1,2)})$. By exploiting the monomial equality constraints (7), $\mathbf{M}$ is now computed via $\mathbf{U}$. The mean bit error rate (BER) over 100 Monte Carlo runs can be seen in Figure 3. We observe that Algorithm 2 with $Q = 3$ or $Q = 4$ works (even in the presence of noise) while CCA-RACMA does not, even without noise, as expected.

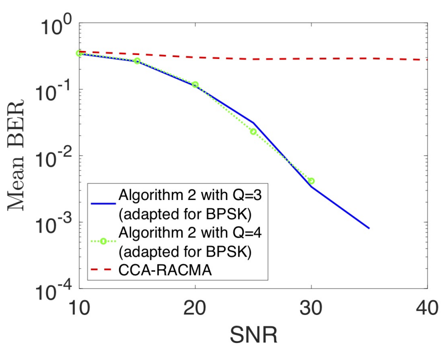

Fig. 3. Mean BER values over 100 Monte Carlo runs.

VI. Conclusion

In this paper we studied multi-set low-rank factorizations with shared and unshared components of which one of the involved factor matrices can be subject to monomial equality constraints. We have explained that such multi-set low-rank factorizations generalize both the well-known CPD for higher-order tensors and the generalized CCA for multi-set / multi-view arrays. More precisely, based on subspace intersection, we presented a link between multi-set low-rank factorizations with shared and unshared components and monomial factorizations, possibly with a low-rank interference term. This led to a uniqueness condition which demonstrated that improved identifiability conditions can be obtained by simultaneously exploiting the common subspace structure and the monomial structure of the involved factorizations. By taking into account the specific structure that captures the monomial constraint, we have even reduced the monomial factorization, possibly with a low-rank interference term, to a CPD. This in turn led to algebraic algorithms for multi-set low-rank factorizations with shared and unshared components. We extended the results to the case of partially shared entities, where both shared and unshared structural information is taken into account.

[3]Note that similar to (43)–(44), $\mathbf{U}$ corresponds to the minimizer of $\|[\mathbf{U}^{(1)}\mathbf{V}^{(1)}, \mathbf{U}^{(2)}\mathbf{V}^{(2)}] - \mathbf{U}[\mathbf{F}^{(1)T}, \mathbf{F}^{(2)T}]\|_F^2$, or equivalently the minimizer of $\|[\mathbf{U}^{(1)}, \mathbf{U}^{(2)}] - \mathbf{U}\mathbf{\Gamma}^T\|_F^2$, where $\mathbf{\Gamma} = [\mathbf{F}^{(1)T}\mathbf{V}^{(1)H}, \mathbf{F}^{(2)T}\mathbf{V}^{(2)H}]$. Alternatively, $\mathbf{U}$ could have been obtained via relation (31) or (47).

Appendix A
Bilinear factorizations subject to monomial equality constraints, possibly with low-rank interference terms

In this appendix we explain that the blind source separation problems in [2] and [3] can be interpreted as bilinear factorizations of the form (1) in which the columns of $\mathbf{M}$ satisfy M monomial equality constraints of the form (6). More importantly, we explain that the factorization approach in [25] allows us in a straightforward way to consider bilinear factorizations of the form (3) that also involve a low-rank interference term. This section will also demonstrate how to construct $\mathbf{P}^{(M,L)}$ in Theorem 2.1 for two concrete cases.

1) Example 1: Constant modulus constraint: Consider the bilinear factorizations of the form (1) in which the columns of $\mathbf{M}$ are subject to a CM constraint and $\mathbf{S}$ has full column rank [2]. Without loss of generality we assume that $\mathbf{S}$ is nonsingular so that $K = R + Q$. Compared to [2] a slightly different but more concise formulation will be used that is in line with the bilinear factorization framework in [25]. We will also consider the more general bilinear matrix factorization of the form (3) in which the columns of $\mathbf{C}$ are *not* CM constrained. By exploiting the monomial relation property $m_{i_1}m_{i_1}^* - m_{i_2}m_{i_2}^* = 0$ of a CM constrained column $\mathbf{m}$, we will derive a condition that ensures the recovery of $\mathbf{M}$, given only $\mathbf{X}$. Since $\mathbf{m} \in \text{range}(\mathbf{X})$ there exists a vector $\mathbf{w} \in \mathbb{C}^{(R+Q)}$ such that $\mathbf{X}\mathbf{w} = \mathbf{m}$. In more detail, we are looking for R vectors $\mathbf{w}_1, \dots, \mathbf{w}_R \in \mathbb{C}^{(R+Q)}$, each with property

$$\mathbf{X}\mathbf{w} = [\mathbf{M}, \mathbf{C}]\mathbf{S}^T\mathbf{w} = \mathbf{m} \in \mathcal{M}_{\text{CM}}, \tag{58}$$

where the subscript 'r' has been dropped for clarity and where M_{CM} denotes the set of CM constrained vectors $\mathbf{m}$ with property $m_1^*m_1 = \cdots = m_I^*m_I$. The CM constraint implies that

$$m_{i_1}m_{i_1}^* - m_{i_2}m_{i_2}^* = 0, \quad 1 \leq i_1 < i_2 \leq I. \tag{59}$$

Compared to (6), (59) involves complex conjugate variables m_i^*. This is just a technical variant of the former. The combination of (58) and (59) yields

$$\begin{aligned}
&m_{i_1}m_{i_1}^* - m_{i_2}m_{i_2}^* = 0 \Leftrightarrow \\
&(\mathbf{e}_{i_1}^{(I)T}\mathbf{X}\mathbf{w})(\mathbf{e}_{i_1}^{(I)T}\mathbf{X}^*\mathbf{w}^*) - (\mathbf{e}_{i_2}^{(I)T}\mathbf{X}\mathbf{w})(\mathbf{e}_{i_2}^{(I)T}\mathbf{X}^*\mathbf{w}^*) = 0 \Leftrightarrow \\
&\left\{\mathbf{e}_{i_1}^{(I)T}\mathbf{X} \otimes \mathbf{e}_{i_1}^{(I)T}\mathbf{X}^* - \mathbf{e}_{i_2}^{(I)T}\mathbf{X} \otimes \mathbf{e}_{i_2}^{(I)T}\mathbf{X}^*\right\}(\mathbf{w} \otimes \mathbf{w}^*) = 0 \Leftrightarrow \\
&\mathbf{p}_{\text{CM}}^{(i_1,i_2)}(\mathbf{w} \otimes \mathbf{w}^*) = 0,
\end{aligned} \tag{60}$$

where $\mathbf{p}_{\text{CM}}^{(i_1,i_2)} := \mathbf{e}_{i_1}^{(I)T}\mathbf{X} \otimes \mathbf{e}_{i_1}^{(I)T}\mathbf{X}^* - \mathbf{e}_{i_2}^{(I)T}\mathbf{X} \otimes \mathbf{e}_{i_2}^{(I)T}\mathbf{X}^* \in \mathbb{C}^{1\times(R+Q)^2}$. Stacking yields

$$\mathbf{P}_{\text{CM}}^{(C_I^2)} \cdot (\mathbf{w} \otimes \mathbf{w}^*) = \mathbf{0}\,, \tag{61}$$

in which

$$\mathbf{P}_{\text{CM}}^{(C_I^2)} = \left[\mathbf{p}_{\text{CM}}^{(1,2)T}, \ldots, \mathbf{p}_{\text{CM}}^{(I-1,I)T}\right]^T \in \mathbb{C}^{C_I^2\times(R+Q)^2}. \tag{62}$$

Note that Theorem 2.1 can be adapted to the case of complex conjugate variables, simply by replacing the condition $\dim(\ker(\mathbf{P}^{(M,L)}) \cap \pi_S^{(L,R+Q)}) = R$ with $\dim(\ker(\mathbf{P}^{(M,L)})) = R$, where $\mathbf{P}^{(M,L)} = \mathbf{P}_{\text{CM}}^{(C_I^2)}$ with $M = C_I^2$ and $L = 2$. Assume that $\dim(\ker(\mathbf{P}_{\text{CM}}^{(C_I^2)})) = R$, which is minimal. Let the columns of $\mathbf{Q} \in \mathbb{C}^{(R+Q)^2\times R}$ form a basis for $\ker(\mathbf{P}_{\text{CM}}^{(C_I^2)})$. Then there exists a nonsingular change-of-basis matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{Q} = (\mathbf{W} \odot \mathbf{W}^*)\mathbf{F}^T, \tag{63}$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_R]$. It is clear that (63) corresponds to a matrix version of the tensor $\mathcal{Q} = \sum_{r=1}^R \mathbf{w}_r \circ \mathbf{w}_r^* \circ \mathbf{f}_r \in \mathbb{C}^{(R+Q)\times(R+Q)\times R}$, whose CPD is unique (see Theorem 1.1 with $\mathbf{A} = \mathbf{W}$, $\mathbf{B} = \mathbf{W}^*$ and $\mathbf{S} = \mathbf{F}$). To summarize, if the dimension of $\ker(\mathbf{P}_{\text{CM}}^{(C_I^2)})$ is minimal (i.e., R), then $\mathbf{W}$ and consequently also $\mathbf{M} = \mathbf{X}\mathbf{W}$ are unique.

2) Example 2: Low-rank constraint: In Sections II and III it was made clear that it can be convenient to interpret the CPD of a tensor with matrix representation (15) as a bilinear factorization subject to monomial equality constraints of the form $(\mathbf{G}_r)_{i_1j_1}(\mathbf{G}_r)_{i_2j_2} - (\mathbf{G}_r)_{i_1j_2}(\mathbf{G}_r)_{i_2j_1} = 0$. In this section we will consider the more general case in which the columns of $\mathbf{M}$ of (3) correspond to vectorized rank-P matrices of the form $\mathbf{m}_r = \text{vec}(\mathbf{G}_r)$ with $\mathbf{G}_r \in \mathbb{C}^{I\times J}$ being a rank-P matrix. We note in passing that a factorization of the form (3) in which the columns of $\mathbf{M}$ are vectorized rank-P matrices but the low-rank interference term $\mathbf{C}$ is absent (i.e., $\mathbf{C} = \mathbf{0}$) is known as a Block Term Decomposition (BTD) [32] and it has been thoroughly studied in [33]. Compared to [33] a slightly different formulation will be used that is more in line with the bilinear factorization framework in [25]. Since $\mathbf{S}$ is assumed to have full column rank, we can w.l.o.g. assume that it is nonsingular, implying that $K = R + Q$. Similar to (58), we are looking for R vectors $\mathbf{w}_1, \ldots, \mathbf{w}_R \in \mathbb{C}^{(R+Q)}$, each with property

$$\mathbf{X}\mathbf{w} = [\mathbf{M}, \mathbf{C}]\mathbf{S}^T\mathbf{w} = \mathbf{m} \in \mathcal{M}_{\text{BTD}}, \tag{64}$$

where again the subscript 'r' has been dropped for clarity and M_{BTD} denotes the set of vectors of the form $\mathbf{m} = \text{vec}(\mathbf{G})$ in which $\mathbf{G} \in \mathbb{C}^{I\times J}$ is a matrix with at most rank $P < \min(I, J)$. The low-rank constraint implies that

$$\mathcal{C}_{P+1}(\mathbf{G}) = \mathbf{0}. \tag{65}$$

Let the $M = C_I^{P+1}C_J^{P+1}$ minors of $\mathbf{G}$ of size $(P+1)$-by-$(P+1)$ be indexed by a superscript 'm' , i.e., $\mathbf{G}^{(m)} \in \mathbb{C}^{(P+1)\times(P+1)}$ is a submatrix formed by $P+1$ rows and $P+1$ columns of $\mathbf{G}$. Relation (65) implies that any $(P+1)$-by-$(P+1)$ minor of $\mathbf{G}$ has the property

$$\left|\mathbf{G}^{(m)}\right| = \sum_{\sigma\in S_{P+1}} \text{sgn}(\sigma) \prod_{i=1}^{P+1} g_{i,\sigma(i)}^{(m)} = 0, \tag{66}$$

where S_{P+1} denotes the set of all permutations of $1, 2, \ldots, P+1$ and $\text{sgn}(\sigma)$ denotes the sign of the permutation σ. Since $P > 0$ there are $\frac{(P+1)!}{2}$ even permutations and $\frac{(P+1)!}{2}$ odd permutations, meaning that (66) can be decomposed as follows

$$\left|\mathbf{G}^{(m)}\right| = \left(\sum_{\sigma\in A_{P+1}} \prod_{i=1}^{P+1} g_{i,\sigma(i)}^{(m)}\right) - \left(\sum_{\tau\in B_{P+1}} \prod_{i=1}^{P+1} g_{i,\tau(i)}^{(m)}\right) = 0, \tag{67}$$

where A_{P+1} and B_{P+1} denote the sets of even and odd permutations, respectively. Note that (67) involves a sum of $\frac{(P+1)!}{2}$ relations of the form (6). Again, the former is just a technical variant of the latter. Note also that $P = 1$ corresponds to the CPD case in which (67) reduces to $g_{1,1}^{(m)}g_{2,2}^{(m)} - g_{1,2}^{(m)}g_{2,1}^{(m)} = 0$. Observe that $g_{i,\sigma(i)}^{(m)} = (\mathbf{e}_{i_m}^{(I)} \otimes \mathbf{e}_{\sigma(i_n)}^{(J)})^T\mathbf{X}\mathbf{w}$, where the subscript '$m$' of i_m takes into account that $\mathbf{G}^{(m)}$ is a submatrix of $\mathbf{G}$.[4] Similar to (60), the combination of (64) and (67) yields

$$\left|\mathbf{G}^{(m)}\right| = 0 \Leftrightarrow \sum_{q=1}^{\frac{(P+1)!}{2}} \mathbf{P}_{P+1}^{(q,m)} \cdot (\underbrace{\mathbf{w} \otimes \cdots \otimes \mathbf{w}}_{P+1 \text{ times}}) = 0,$$

where

$$\mathbf{p}_{P+1}^{(q,m)} = \bigotimes_{p=1}^{P+1}((\mathbf{e}_{i_p}^{(I)} \otimes \mathbf{e}_{\sigma_q(i_p)}^{(J)})^T\mathbf{X}) - \bigotimes_{p=1}^{P+1}((\mathbf{e}_{i_p}^{(I)} \otimes \mathbf{e}_{\tau_q(i_p)}^{(J)})^T\mathbf{X}),$$

in which $(i_1, i_2, \ldots, i_{P+1})$ is determined by 'm', σ_q denotes the qth element of A_{P+1} and τ_q denotes the qth element of B_{P+1}. Stacking yields

$$\mathbf{P}_{\text{BTD}}^{(M,P)} \cdot (\mathbf{w} \otimes \cdots \otimes \mathbf{w}) = \mathbf{0}\,, \tag{68}$$

in which

$$\mathbf{P}_{\text{BTD}}^{(M,P)} = \begin{bmatrix} \sum_{q=1}^{\frac{(P+1)!}{2}} \mathbf{P}_{P+1}^{(q,1)} \\ \vdots \\ \sum_{q=1}^{\frac{(P+1)!}{2}} \mathbf{P}_{P+1}^{(q,M)} \end{bmatrix} \in \mathbb{C}^{M\times(R+Q)^{P+1}}. \tag{69}$$

Note that Theorem 2.1 can also be used in the BTD case, i.e., simply set $\mathbf{P}^{(M,L)} = \mathbf{P}_{\text{BTD}}^{(M,P)}$ with $M = C_I^{P+1}C_J^{P+1}$ and $L = P+1$. Assume that $\dim(\ker(\mathbf{P}_{\text{BTD}}^{(M,P)}) \cap \pi_S^{(P+1,R+Q)}) = R$, which is minimal. Let the columns of $\mathbf{Q} \in \mathbb{C}^{(R+Q)^{P+1}\times R}$ form a basis for $\ker(\mathbf{P}_{\text{BTD}}^{(M,P)}) \cap \pi_S^{(P+1,R+Q)}$. Then there exists a nonsingular change-of-basis matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{Q} = (\mathbf{W} \odot \cdots \odot \mathbf{W})\mathbf{F}^T, \tag{70}$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_R] \in \mathbb{C}^{(R+Q)\times R}$. It is clear that (70) corresponds to a matrix version of the tensor $\mathcal{Q} = \sum_{r=1}^R \mathbf{w}_r \circ \cdots \circ \mathbf{w}_r \circ \mathbf{f}_r \in \mathbb{C}^{(R+Q)\times\cdots\times(R+Q)\times R}$, whose CPD is unique (see Theorem 1.1 with $\mathbf{A} = \mathbf{W}\odot\cdots\odot\mathbf{W}$, $\mathbf{B} = \mathbf{W}$ and $\mathbf{S} = \mathbf{F}$). To summarize, if the dimension of $\dim(\ker(\mathbf{P}_{\text{BTD}}^{(N,P)}) \cap \pi_S^{(P+1,R+Q)}$ is minimal (i.e., R), then $\mathbf{W}$ and $\mathbf{M} = \mathbf{X}\mathbf{W}$ are unique.

We note in passing that it can be verified that in the CPD case where $P = 1$ and $Q = 0$ (i.e., $\mathbf{C} = \mathbf{0}$), the condition

[4] Alternatively, this could have been expressed as $g_{i,\sigma(i)}^{(m)} = (\mathbf{e}_i^{(I)} \otimes \mathbf{e}_{\sigma(i)}^{(J)})^T\mathbf{X}^{(m)}\mathbf{w}$, where the superscript 'm' now extracts the appropriate $(P+1) \times (P+1)$ submatrix of $\mathbf{X}$.

that $\ker(\mathbf{P}_{\text{BTD}}^{(M,1)}) \cap \pi_S^{(2,R)}$ is minimal (i.e., R) and $\mathbf{S}$ has full column rank is equivalent to the CPD uniqueness condition (16) stated in Theorem 1.1, i.e., we can replace the condition that $\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ has full column rank with the equivalent condition that $\dim(\ker(\mathbf{P}_{\text{BTD}}^{(C_I^2 C_J^2,1)}) \cap \pi_S^{(2,R)}) = R$.

To summarize, the problem of computing $\mathbf{M}$ via a bilinear factorization of the form (3) in which the columns of $\mathbf{M}$ exhibit a monomial structure can be transformed into a low-rank CPD problem, even in cases where a low-rank interference term $\mathbf{C}$ is present.

Appendix B
Identifiability condition based on the kernel subspace intersection formulation

We will derive an identifiability condition for the multi-set factorization (35) based on relations (33) and (34). First, we explain how to exploit the rank-one structure associated with $\mathbf{M}^{(n)} = \mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}$ and the common subspace structure between $\text{range}(\mathbf{X}^{(m)})$ and $\text{range}(\mathbf{X}^{(n)})$. Next, we explain how to merge these two structures, which will lead to a uniqueness condition for $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$. Finally, we derive a condition that ensures the uniqueness of $\mathbf{A}$ and $\mathbf{B}$, given that $\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(N)}$ are unique (up to intrinsic scaling ambiguities).

1) Exploiting rank-one structure within $\boldsymbol{X}^{(n)}$: Since the columns of the Khatri–Rao structured matrix $\mathbf{M}^{(n)} = \mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}$ correspond to vectorized rank-one matrices, we already know from the discussion in Section A-2 that $\mathbf{M}^{(n)}$ satisfies $M_n = C_{I_n}^2 C_{J_n}^2$ monomial relations of the form

$$m^{(n)}_{(i_1-1)J_n+j_1,r} m^{(n)}_{(i_2-1)J_n+j_2,r} - m^{(n)}_{(i_2-1)J_n+j_1,r} m^{(n)}_{(i_1-1)J_n+j_2,r} = \begin{vmatrix} a^{(n)}_{i_1,r} b^{(n)}_{j_1,r} & a^{(n)}_{i_1,r} b^{(n)}_{j_2,r} \\ a^{(n)}_{i_1,r} b^{(n)}_{j_1,r} & a^{(n)}_{i_1,r} b^{(n)}_{j_2,r} \end{vmatrix} = 0, \tag{71}$$

where $1 \leq i_1 < i_2 \leq I_n$ and $1 \leq j_1 < j_2 \leq J_n$. The combination of (33) and (71) yields

$$\begin{vmatrix} (\mathbf{e}^{(I_n)}_{i_1} \otimes \mathbf{e}^{(J_n)}_{j_1})^T \mathbf{X}^{(n)} \mathbf{w}^{(n)}_r & (\mathbf{e}^{(I_n)}_{i_1} \otimes \mathbf{e}^{(J_n)}_{j_2})^T \mathbf{X}^{(n)} \mathbf{w}^{(n)}_r \\ (\mathbf{e}^{(I_n)}_{i_2} \otimes \mathbf{e}^{(J_n)}_{j_1})^T \mathbf{X}^{(n)} \mathbf{w}^{(n)}_r & (\mathbf{e}^{(I_n)}_{i_2} \otimes \mathbf{e}^{(J_n)}_{j_2})^T \mathbf{X}^{(n)} \mathbf{w}^{(n)}_r \end{vmatrix} = \mathbf{q}^{(m,n,n)} (\mathbf{w}^{(n)}_r \otimes \mathbf{w}^{(n)}_r) = 0, \tag{72}$$

where $\mathbf{q}^{(m,n,n)} = ((\mathbf{e}^{(I_n)}_{i_1} \otimes \mathbf{e}^{(J_n)}_{j_1})^T \mathbf{X}^{(n)}) \otimes ((\mathbf{e}^{(I_n)}_{i_2} \otimes \mathbf{e}^{(J_n)}_{j_2})^T \mathbf{X}^{(n)}) - ((\mathbf{e}^{(I_n)}_{i_2} \otimes \mathbf{e}^{(J_n)}_{j_1})^T \mathbf{X}^{(n)}) \otimes ((\mathbf{e}^{(I_n)}_{i_1} \otimes \mathbf{e}^{(J_n)}_{j_2})^T \mathbf{X}^{(n)})$, and the superscript '$m$' in the row-vector $\mathbf{q}^{(m,n,n)} \in \mathbb{C}^{1 \times (R+Q_n)^2}$ takes all the subscripts i_1, i_2, j_1 and j_2 into account. Stacking yields

$$\mathbf{Q}^{(M_n,n,n)} (\mathbf{w}^{(n)} \otimes \mathbf{w}^{(n)}) = \mathbf{0}, \tag{73}$$

where $\mathbf{Q}^{(M_n,n,n)} = [\mathbf{q}^{(1,n,n)T}, \ldots, \mathbf{q}^{(M_n,n,n)T}]^T \in \mathbb{C}^{M_n \times (R+Q_n)^2}$ with $M_n = C_{I_n}^2 C_{J_n}^2$. Note that all rows in $\mathbf{X}^{(n)}$ are involved in the construction of $\mathbf{Q}^{(M_n,n,n)}$.

2) Exploiting rank-one structure between $\boldsymbol{X}^{(n_1)}$ and $\boldsymbol{X}^{(n_2)}$: Using (34), we can also exploit the rank-one structure between $\mathbf{X}^{(n_1)}$ and $\mathbf{X}^{(n_2)}$, as explained next. Consider two integers n_1 and n_2 with $1 \leq n_1 < n_2 \leq N$. Let

$$\mathbf{S}_{\mathbf{A}}^{(n_1,n_2)} \in \mathbb{C}^{I_{n_1,n_2} \times I_{n_1}}, \qquad \mathbf{S}_{\mathbf{B}}^{(n_1,n_2)} \in \mathbb{C}^{J_{n_1,n_2} \times J_{n_1}},$$
$$\mathbf{S}_{\mathbf{A}}^{(n_2,n_1)} \in \mathbb{C}^{I_{n_1,n_2} \times I_{n_2}}, \qquad \mathbf{S}_{\mathbf{B}}^{(n_2,n_1)} \in \mathbb{C}^{J_{n_1,n_2} \times J_{n_2}},$$

denote row-selection matrices such that

$$(\mathbf{S}_{\mathbf{A}}^{(n_1,n_2)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_1,n_2)}) \mathbf{M}^{(n_1)} = (\mathbf{S}_{\mathbf{A}}^{(n_2,n_1)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_2,n_1)}) \mathbf{M}^{(n_2)} = \mathbf{A}^{(n_1,n_2)} \odot \mathbf{B}^{(n_1,n_2)},$$

where $\mathbf{A}^{(n_1,n_2)} = \mathbf{S}_{\mathbf{A}}^{(n_1,n_2)} \mathbf{A} = \mathbf{S}_{\mathbf{A}}^{(n_2,n_1)} \mathbf{A}$ and $\mathbf{B}^{(n_1,n_2)} = \mathbf{S}_{\mathbf{B}}^{(n_1,n_2)} \mathbf{B} = \mathbf{S}_{\mathbf{B}}^{(n_2,n_1)} \mathbf{B}$. In words, $\mathbf{S}_{\mathbf{A}}^{(n_1,n_2)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_1,n_2)}$ and $\mathbf{S}_{\mathbf{A}}^{(n_2,n_1)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_2,n_1)}$ select the shared rows between $\mathbf{A}^{(n_1)} \odot \mathbf{B}^{(n_1)}$ and $\mathbf{A}^{(n_2)} \odot \mathbf{B}^{(n_2)}$. Note that the row selection matrices $\mathbf{S}_{\mathbf{A}}^{(n_1,n_2)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_1,n_2)}$ and $\mathbf{S}_{\mathbf{A}}^{(n_2,n_1)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_2,n_1)}$ can be more restrictive (in the sense of selecting rows in a structured fashion) than the row selection matrices $\mathbf{D}^{(n_1,n_2)}$ and $\mathbf{D}^{(n_2,n_1)}$ in (33). Define

$$\mathbf{Y}^{(n_1,n_2)} = (\mathbf{S}_{\mathbf{A}}^{(n_1,n_2)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_1,n_2)}) \mathbf{X}^{(n_1)}, \tag{74}$$
$$\mathbf{Y}^{(n_2,n_1)} = (\mathbf{S}_{\mathbf{A}}^{(n_2,n_1)} \otimes \mathbf{S}_{\mathbf{B}}^{(n_2,n_1)}) \mathbf{X}^{(n_2)}. \tag{75}$$

The combination of (34), (71), (74) and (75) yields

$$\begin{vmatrix} \boldsymbol{\gamma}^{(I_{n_1},J_{n_1})T}_{i_1,j_1} \mathbf{Y}^{(n_1,n_2)} \mathbf{v}^{(n_1)}_r & \boldsymbol{\gamma}^{(I_{n_2},J_{n_2})T}_{i_1,j_1} \mathbf{Y}^{(n_2,n_1)} \mathbf{v}^{(n_2)}_r \\ \boldsymbol{\gamma}^{(I_{n_1},J_{n_1})T}_{i_2,j_1} \mathbf{Y}^{(n_1,n_2)} \mathbf{v}^{(n_1)}_r & \boldsymbol{\gamma}^{(I_{n_2},J_{n_2})T}_{i_2,j_2} \mathbf{Y}^{(n_2,n_1)} \mathbf{v}^{(n_2)}_r \end{vmatrix} = \mathbf{q}^{(m,n_1,n_2)} (\mathbf{w}^{(n_1)}_r \otimes \mathbf{w}^{(n_2)}_r) = 0,$$

where $\boldsymbol{\gamma}^{(I_n,J_n)}_{i_p,j_q} = \mathbf{e}^{(I_n)}_{i_p} \otimes \mathbf{e}^{(J_n)}_{j_q} \in \mathbb{C}^{I_n J_n}$ and

$$\mathbf{q}^{(m,n_1,n_2)} = (\boldsymbol{\gamma}^{(I_{n_1},J_{n_1})T}_{i_1,j_1} \mathbf{Y}^{(n_1,n_2)}) \otimes (\boldsymbol{\gamma}^{(I_{n_2},J_{n_2})T}_{i_2,j_2} \mathbf{Y}^{(n_2,n_1)}) - (\boldsymbol{\gamma}^{(I_{n_2},J_{n_2})T}_{i_1,j_1} \mathbf{Y}^{(n_1,n_2)}) \otimes (\boldsymbol{\gamma}^{(I_{n_1},J_{n_1})T}_{i_2,j_1} \mathbf{Y}^{(n_2,n_1)}),$$

in which the superscript 'm' in the row-vector $\mathbf{q}^{(m,n_1,n_2)} \in \mathbb{C}^{1 \times (R+Q_{n_1})(R+Q_{n_2})}$ takes all the subscripts i_1, i_2, j_1 and j_2 into account. Stacking yields

$$\mathbf{Q}^{(M_{n_1,n_2},n_1,n_2)} (\mathbf{w}^{(n_1)}_r \otimes \mathbf{w}^{(n_2)}_r) = \mathbf{0}, \tag{76}$$

where $\mathbf{Q}^{(M_{n_1,n_2},n_1,n_2)} = [\mathbf{q}^{(1,n_1,n_2)T}, \ldots, \mathbf{q}^{(M_{n_1,n_2},n_1,n_2)T}]^T \in \mathbb{C}^{M_{n_1,n_2} \times (R+Q_{n_1})(R+Q_{n_2})}$ with $M_{n_1,n_2} = C^2_{I_{n_1,n_2}} C^2_{J_{n_1,n_2}}$.

3) Exploiting common subspace structure $\text{range}(\boldsymbol{X}^{(m)})$ and $\text{range}(\boldsymbol{X}^{(n)})$: Note that by exploiting the rank-one structures within and between the matrices $\mathbf{X}^{(m)}$ and $\mathbf{X}^{(n)}$, we obtained several systems of equations of the form (73) and (76). We wish to combine them into a single system of equations that can be more conveniently studied using the CPD. This will be accomplished by exploiting the common subspace structure between $\text{range}(\mathbf{X}^{(m)})$ and $\text{range}(\mathbf{X}^{(n)})$. In detail, since $\mathbf{w}^{(n_1)}_r \neq \mathbf{0}$, $\mathbf{D}^{(n_1,n_2)} \mathbf{X}^{(n_1)} \mathbf{w}^{(n_1)}_r - \mathbf{D}^{(n_2,n_1)} \mathbf{X}^{(n_2)} \mathbf{w}^{(n_2)}_r = \mathbf{0}$ if and only if $\mathbf{w}^{(n_1)}_r \otimes (\mathbf{D}^{(n_1,n_2)} \mathbf{X}^{(n_1)} \mathbf{w}^{(n_1)}_r - \mathbf{D}^{(n_2,n_1)} \mathbf{X}^{(n_2)} \mathbf{w}^{(n_2)}_r) = \mathbf{0}$, which in turn is equivalent to

$$(\mathbf{I}_{R+Q_{n_1}} \otimes \mathbf{D}^{(n_1,n_2)} \mathbf{X}^{(n_1)}) (\mathbf{w}^{(n_1)}_r \otimes \mathbf{w}^{(n_1)}_r) - (\mathbf{I}_{R+Q_{n_1}} \otimes \mathbf{D}^{(n_2,n_1)} \mathbf{X}^{(n_2)}) (\mathbf{w}^{(n_1)}_r \otimes \mathbf{w}^{(n_2)}_r) = \mathbf{0}. \tag{77}$$

By a similar reasoning, we obtain

$$(\mathbf{I}_{R+Q_{n_2}} \otimes \mathbf{D}^{(n_1,n_2)} \mathbf{X}^{(n_1)}) (\mathbf{w}^{(n_2)}_r \otimes \mathbf{w}^{(n_1)}_r) - (\mathbf{I}_{R+Q_{n_2}} \otimes \mathbf{D}^{(n_2,n_1)} \mathbf{X}^{(n_2)}) (\mathbf{w}^{(n_2)}_r \otimes \mathbf{w}^{(n_2)}_r) = \mathbf{0}. \tag{78}$$

Alternatively, we can also work with the relations

$$(\mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)})(\mathbf{w}_r^{(n_1)} \otimes \mathbf{w}_r^{(n_1)}) - (\mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)})(\mathbf{w}_r^{(n_2)} \otimes \mathbf{w}_r^{(n_2)}) = \mathbf{0}, \quad (79)$$

$$(\mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)})(\mathbf{w}_r^{(n_2)} \otimes \mathbf{w}_r^{(n_1)}) - (\mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)})(\mathbf{w}_r^{(n_2)} \otimes \mathbf{w}_r^{(n_2)}) = \mathbf{0}. \quad (80)$$

In addition, we can also make use of the relation

$$(\mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)})(\mathbf{w}_r^{(n_1)} \otimes \mathbf{w}_r^{(n_1)}) - (\mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)})(\mathbf{w}_r^{(n_2)} \otimes \mathbf{w}_r^{(n_2)}) = \mathbf{0}. \quad (81)$$

We observe that every rank-one structured vector of the form $\mathbf{w}_r^{(m)} \otimes \mathbf{w}_r^{(n)}$, $1 \le m,n \le N$ appears in (73) and (76)–(81). This property can be used to conveniently combine the involved system of equations, as will be explained next.

4) Combination of rank-one and common subspace structures: The combination of the common subspace structures (77)–(80) and the rank-one structures (73) and (76) yields

$$\mathbf{G}^{(n_1,n_2)}\left(\begin{bmatrix}\mathbf{w}_r^{(n_1)}\\ \mathbf{w}_r^{(n_2)}\end{bmatrix} \otimes \begin{bmatrix}\mathbf{w}_r^{(n_1)}\\ \mathbf{w}_r^{(n_2)}\end{bmatrix}\right) = \mathbf{0}, 1 \le n_1 < n_2 \le N, \quad (82)$$

where $\mathbf{G}^{(n_1,n_2)} \in \mathbb{C}^{M_{n_1,n_2}\times R^2_{n_1,n_2}}$ is given by (91) in which $M_{n_1,n_2} = F_{n_1,n_2}(2R+Q_{n_1}+Q_{n_2}+3F_{n_1,n_2}) + C^2_{I_{n_1}}C^2_{J_{n_1}} + C^2_{I_{n_2}}C^2_{J_{n_2}} + C^2_{I_{n_1}}C^2_{J_{n_2}} + C^2_{I_{n_2}}C^2_{J_{n_1}}$, $R_{n_1,n_2} = 2R + Q_{n_1} + Q_{n_2}$, $\{\mathbf{\Pi}_n\}$ denote appropriate column permutation matrices associated with the pair (n_1,n_2) and $\{\mathbf{0}\}$ denote zero matrices of conformable sizes.

5) Reduction to CPD: The Kronecker structured system of equations (82) can be combined as follows:

$$\mathbf{G}^{(\text{tot})}\left(\begin{bmatrix}\mathbf{w}_r^{(1)}\\ \vdots\\ \mathbf{w}_r^{(N)}\end{bmatrix} \otimes \begin{bmatrix}\mathbf{w}_r^{(1)}\\ \vdots\\ \mathbf{w}_r^{(N)}\end{bmatrix}\right) = \mathbf{0}, \quad (83)$$

in which

$$\mathbf{G}^{(\text{tot})} = \begin{bmatrix}[\mathbf{G}^{(1,2)},\mathbf{0}]\mathbf{\Pi}_{1,2}\\ [\mathbf{G}^{(1,3)},\mathbf{0}]\mathbf{\Pi}_{1,3}\\ \vdots\\ [\mathbf{G}^{(N-1,N)},\mathbf{0}]\mathbf{\Pi}_{N-1,N}\end{bmatrix} \in \mathbb{C}^{M_{\text{tot}}\times R^2_{\text{tot}}}, \quad (84)$$

where $\mathbf{G}^{(n_1,n_2)}$ is given by (91), $R_{\text{tot}} = NR + \sum_{n=1}^{N} Q_n$, $M_{\text{tot}} = \sum_{1\le n_1<n_2\le N} M_{n_1,n_2}$, $\{\mathbf{\Pi}_{m,n}\}$ denote appropriate column permutation matrices and $\{\mathbf{0}\}$ denote zero matrices of conformable sizes. Let $S_{R_{\text{tot}}}$ denote the $C^2_{R_{\text{tot}}}$-dimensional subspace of vectorized ($R_{\text{tot}} \times R_{\text{tot}}$) symmetric matrices. From (83) it is clear that if the dimension of the subspace $\ker(\mathbf{G}^{(\text{tot})}) \cap S_{R_{\text{tot}}}$ is minimal (i.e., R), then the matrices $\mathbf{W}^{(1)},\dots,\mathbf{W}^{(N)}$ can be obtained from it. In more detail, let the columns of $\mathbf{R} \in \mathbb{C}^{R^2_{\text{tot}}\times R}$ form a basis for $\ker(\mathbf{G}^{(\text{tot})}) \cap S_{R_{\text{tot}}}$, then there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{R} = \left(\begin{bmatrix}\mathbf{W}^{(1)}\\ \vdots\\ \mathbf{W}^{(N)}\end{bmatrix} \odot \begin{bmatrix}\mathbf{W}^{(1)}\\ \vdots\\ \mathbf{W}^{(N)}\end{bmatrix}\right)\mathbf{F}^T. \quad (85)$$

Clearly, (85) corresponds to a third-order tensor whose CPD is unique. This property together with relations (33) implies the uniqueness of $\mathbf{M}^{(1)},\dots,\mathbf{M}^{(N)}$, which in turn implies the uniqueness of $\{\mathbf{A}^{(n)}\}$ and $\{\mathbf{B}^{(n)}\}$ up to individual column scaling ambiguities.

A. Uniqueness of A *and* B *based on graph connectivity*

The goal is now to ensure the uniqueness of the global rank-one matrices $\mathbf{m}_r = \mathbf{a}_r\mathbf{b}_r^T$, given the local rank-one matrices $\mathbf{m}_r^{(n)} = \mathbf{a}_r^{(n)}\mathbf{b}_r^{(n)T}$, $n \in \{1,\dots,N\}$. Since each subvector $\mathbf{a}_r^{(n)}$ of $\mathbf{a}_r$ is subject to an individual scaling, we need to make sure that we can get rid of this ambiguity. Intuitively speaking, this is only possible if the different subvectors are sufficiently "overlapping". Specifically, let $G^{(\mathbf{a}_r)}$ denote an intersection graph formed from a family of sets $\{E_r^{(n)}\}$, in which the elements of the set $E_r^{(n)}$ correspond to the entries of $\mathbf{a}_r^{(n)}$, i.e., $E_r^{(n)} = \{a_{1r}^{(n)},\dots,a_{I_n r}^{(n)}\}$. The edge between $E_r^{(n_1)}$ and $E_r^{(n_2)}$ is denoted by $\mathbf{e}_{(n_1)}^{(n_2)}$ and it is equal to

$$\mathbf{e}_{(n_1)}^{(n_2)} = \begin{cases}1, & \text{if } \omega\left(\boldsymbol{\tau}^{(n_1)} * \boldsymbol{\tau}^{(n_2)}\right) \ge 1,\\ 0, & \text{if } \omega\left(\boldsymbol{\tau}^{(n_1)} * \boldsymbol{\tau}^{(n_2)}\right) = 0,\end{cases} \quad (86)$$

where '$*$' denotes the Hadamard (elementwise) product and $\boldsymbol{\tau}^{(n)} = \mathbf{\Pi}^{(n)T}\begin{bmatrix}\mathbf{1}_{I_n}\\ \mathbf{0}\end{bmatrix} \in \mathbb{C}^I$ in which $\mathbf{\Pi}^{(n)} \in \{0,1\}^{I_n\times I}$ is a row-selection matrix with property $\mathbf{a}_r^{(n)} = \mathbf{\Pi}^{(n)}\mathbf{a}_r$. In words, $E_r^{(n_1)}$ and $E_r^{(n_2)}$ are connected if $\mathbf{e}_{(n_1)}^{(n_2)} = 1$, i.e., the selection matrices $\mathbf{\Pi}^{(n_1)}$ and $\mathbf{\Pi}^{(n_2)}$ select a common entry in $\mathbf{a}_r$. If the simple intersection graph $G^{(\mathbf{a}_r)}$ is connected and

$$\forall i \in \{1,\dots,I\},\ \exists n \in \{1,\dots,N\} \text{ such that } a_{ir} \in E_r^{(n)}, \quad (87)$$

then $\mathbf{a}_r$ is unique, i.e., we can get rid of the individual scaling ambiguities associated with the subvectors $\{\mathbf{a}_r^{(n)}\}$.

Now that $\mathbf{A}$ has been obtained, $\mathbf{A}^{(1)},\dots,\mathbf{A}^{(N)}$ are known. This in turn implies that $\mathbf{B}$ can be obtained by solving a system of linear equations. In more detail, observe that

$$\mathbf{y}_r^{(n)} = \mathbf{a}_r^{(n)} \otimes \mathbf{b}_r^{(n)} = \mathbf{a}_r^{(n)} \otimes (\mathbf{S}_r^{(n)}\mathbf{b}_r) = (\mathbf{a}_r^{(n)} \otimes \mathbf{S}_r^{(n)})\mathbf{b}_r, \quad (88)$$

where $n \in \{1,\dots,N\}$ and $\mathbf{S}_r^{(n)} \in \{0,1\}^{J_n\times J}$ is a row-selection matrix such that $\mathbf{S}_r^{(n)}\mathbf{b}_r = \mathbf{b}_r^{(n)}$. The r-th column of $\mathbf{B}$ can be obtained from the solution to the following system of linear equations

$$\begin{bmatrix}\mathbf{y}_r^{(1)}\\ \vdots\\ \mathbf{y}_r^{(N)}\end{bmatrix} = \begin{bmatrix}\mathbf{a}_r^{(1)} \otimes \mathbf{S}_r^{(1)}\\ \vdots\\ \mathbf{a}_r^{(N)} \otimes \mathbf{S}_r^{(N)}\end{bmatrix}\mathbf{b}_r = \mathbf{G}^{(\mathbf{b}_r)}\mathbf{b}_r,\ r \in \{1,\dots,R\}, \quad (89)$$

where

$$\mathbf{G}^{(\mathbf{b}_r)} = \left[(\mathbf{a}_r^{(1)} \otimes \mathbf{S}_r^{(1)})^T,\dots,(\mathbf{a}_r^{(N)} \otimes \mathbf{S}_r^{(N)})^T\right]^T. \quad (90)$$

B. Summary

Theorem B.1 below summarizes the obtained uniqueness condition for the multi-set low-rank factorization of the form (35) in which $\mathbf{M} = \mathbf{A} \odot \mathbf{B}$ is partially shared.

$$\mathbf{G}^{(n_1,n_2)} = \begin{bmatrix} [\mathbf{I}_{R+Q_{n_1}} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} - \mathbf{I}_{R+Q_{n_1}} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)}, \mathbf{0}]\mathbf{\Pi}_1 \\ [\mathbf{I}_{R+Q_{n_2}} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} - \mathbf{I}_{R+Q_{n_2}} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)}, \mathbf{0}]\mathbf{\Pi}_2 \\ [\mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} - \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)}, \mathbf{0}]\mathbf{\Pi}_3 \\ [\mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} - \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)}, \mathbf{0}]\mathbf{\Pi}_4 \\ [\mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} \otimes \mathbf{D}^{(n_1,n_2)}\mathbf{X}^{(n_1)} - \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)} \otimes \mathbf{D}^{(n_2,n_1)}\mathbf{X}^{(n_2)}, \mathbf{0}]\mathbf{\Pi}_5 \\ [\mathbf{Q}^{(M_{n_1,n_1},n_1,n_1)}, \mathbf{0}]\mathbf{\Pi}_6 \\ [\mathbf{Q}^{(M_{n_1,n_2},n_1,n_2)}, \mathbf{0}]\mathbf{\Pi}_7 \\ [\mathbf{Q}^{(M_{n_2,n_2},n_2,n_2)}, \mathbf{0}]\mathbf{\Pi}_8 \\ [\mathbf{Q}^{(M_{n_2,n_1},n_2,n_1)}, \mathbf{0}]\mathbf{\Pi}_9 \end{bmatrix}. \tag{91}$$

Theorem B.1: Consider the multi-set low-rank factorization of $\mathbf{X}^{(n)} \in \mathbb{C}^{I_n J_n \times K_n}$, $n \in \{1,\dots,N\}$ in (35). If

$$\begin{cases} \ker(\mathbf{G}^{(\text{tot})}) \cap S_{R_{\text{tot}}} \text{ is } R\text{-dimensional}, \\ \mathbf{S}^{(1)},\dots,\mathbf{S}^{(N)} \text{ have full column rank}, \\ G^{(\mathbf{a}_r)} \text{ is connected and has property (87)}, \forall r \in \{1,\dots,R\}, \\ \mathbf{G}^{(\mathbf{b}_r)} \text{ has full column rank}, \forall r \in \{1,\dots,R\}, \end{cases} \tag{92}$$

then the partially shared Khatri–Rao structured factor matrix $\mathbf{M} = \mathbf{A} \odot \mathbf{B}$ is unique.

Appendix C
An Algebraic Algorithm Based on the Kernel Subspace Intersection Approach

In Appendix B we alluded that Theorem B.1 admits a constructive interpretation that can be used to compute the shared factors $\mathbf{A}$ and $\mathbf{B}$, given only $\mathbf{X}^{(1)} \in \mathbb{C}^{I_1 J_1 \times K_1},\dots,\mathbf{X}^{(N)} \in \mathbb{C}^{I_N J_N \times K_N}$. We will now outline the steps of this algorithm. Let the columns of the matrix $\mathbf{U}^{(n)} \in \mathbb{C}^{I_n J_n \times (R+Q_n)}$ form an orthonormal basis for the subspace range$(\mathbf{X}^{(n)})$. Then there exist full column rank matrices $\mathbf{W}^{(n)} \in \mathbb{C}^{(R+Q_n)\times R}$, $n \in \{1,\dots,N\}$, such that

$$\mathbf{U}^{(n)}\mathbf{W}^{(n)} = \mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}, \quad n \in \{1,\dots,N\}. \tag{93}$$

Using $\mathbf{U}^{(1)},\dots,\mathbf{U}^{(N)}$, we build $\mathbf{G}^{(\text{tot})}$ given by (84). The next step is to find a matrix $\mathbf{R}$ whose columns form an orthonormal basis for $\ker(\mathbf{G}^{(\text{tot})}) \cap \pi_S^{(2)}$. Let the columns of $\mathbf{G}^{(\text{tot})} \in \mathbb{C}^{N_{\text{tot}} \times R_{\text{tot}}^2}$ be indexed by the pairs $\{(r_1,r_2)\}_{1\le r_1,r_2 \le R_{\text{tot}}}$, and ordered lexicographically (recall from Section B-5 that $R_{\text{tot}} = NR + \sum_{n=1}^N Q_n$):

$$\mathbf{G}^{(\text{tot})} = [\mathbf{g}^{(\text{tot})}_{(1,1)}, \mathbf{g}^{(\text{tot})}_{(1,2)}, \dots, \mathbf{g}^{(\text{tot})}_{(R_{\text{tot}},R_{\text{tot}})}].$$

Similar to (53), $\mathbf{G}^{(\text{sym})} \in \mathbb{C}^{N_{\text{tot}} \times C^2_{R_{\text{tot}}+1}}$ is constructed from the columns $\{\mathbf{g}^{(\text{tot})}_{(r_1,r_2)}\}_{r_1 \le r_2}$ as follows

$$\mathbf{G}^{(\text{sym})} = [\mathbf{g}^{(\text{tot})}_{(1,1)}, 2\mathbf{g}^{(\text{tot})}_{(1,2)}, \dots, 2\mathbf{g}^{(\text{tot})}_{(1,R_{\text{tot}})}, \mathbf{g}^{(\text{tot})}_{(2,2)}, 2\mathbf{g}^{(\text{tot})}_{(2,3)}, \dots, 2\mathbf{g}^{(\text{tot})}_{(2,R_{\text{tot}})}, \dots, \mathbf{g}^{(\text{tot})}_{(R_{\text{tot}},R_{\text{tot}})}]. \tag{94}$$

We now have that

$$[\mathbf{w}_r^{(1)T},\dots,\mathbf{w}_r^{(N)T}]^T \otimes [\mathbf{w}_r^{(1)T},\dots,\mathbf{w}_r^{(N)T}]^T \in \ker(\mathbf{G}^{(\text{tot})}) \cap \pi_S^{(2)} \Leftrightarrow \mathbf{f}^{(2)}([\mathbf{w}_r^{(1)T},\dots,\mathbf{w}_r^{(N)T}]^T) \in \ker(\mathbf{G}^{(\text{sym})}), \tag{95}$$

where $\mathbf{f}^{(2)}([\mathbf{w}_r^{(1)T},\dots,\mathbf{w}_r^{(N)T}]^T) \in \mathbb{C}^{C^2_{L_{\text{tot}}}}$ is a structured vector of the form (55), but built from the distinct entries of the vector $[\mathbf{w}_r^{(1)T},\dots,\mathbf{w}_r^{(N)T}]^T \otimes [\mathbf{w}_r^{(1)T},\dots,\mathbf{w}_r^{(N)T}]^T$. Hence, if condition (92) in Theorem B.1 is satisfied, $\ker(\mathbf{G}^{(\text{sym})})$ is an R-dimensional subspace. Let the columns of $\mathbf{V}^{(\text{sym})}$ form a basis for $\ker(\mathbf{G}^{(\text{sym})})$ and obtained via the SVD of $\mathbf{G}^{(\text{sym})}$. Then

$$\mathbf{R} = \mathbf{D}_{R_{\text{tot}}} \cdot \mathbf{V}^{(\text{sym})} \tag{96}$$

corresponds to (85) and $\mathbf{D}_{R_{\text{tot}}} \in \mathbb{C}^{R_{\text{tot}}^2 \times C^2_{R+1}}$ is the duplication matrix with property $[\mathbf{w}^{(1)T},\dots,\mathbf{w}^{(N)T}]^T \otimes [\mathbf{w}^{(1)T},\dots,\mathbf{w}^{(N)T}]^T = \mathbf{D}_{R_{\text{tot}}} \cdot \mathbf{f}^{(2)}([\mathbf{w}^{(1)T},\dots,\mathbf{w}^{(N)T}]^T)$. The matrices $\mathbf{W}^{(1)},\dots,\mathbf{W}^{(N)}$ can be obtained from the CPD of $\mathbf{R}$ given by (85). The matrices $\{\mathbf{A}^{(n)}\}$ and $\{\mathbf{B}^{(n)}\}$ can be obtained from the rank-one factorizations of the columns of

$$\mathbf{Y}^{(n)} := \mathbf{U}^{(n)}\mathbf{W}^{(n)} = \mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}, \quad n \in \{1,\dots,N\}. \tag{97}$$

From $\{\mathbf{A}^{(n)}\}$ we can find $\mathbf{A}$. Let us here solve this problem via rank-one completion. Define the matrix

$$\mathbf{P}_r^{(\mathbf{A})} = \left[\mathbf{\Pi}^{(1)T} \cdot \begin{bmatrix} \mathbf{a}_r^{(1)} \\ \star \end{bmatrix}, \dots, \mathbf{\Pi}^{(N)T} \cdot \begin{bmatrix} \mathbf{a}_r^{(N)} \\ \star \end{bmatrix}\right], \tag{98}$$

where '$\star$' denotes indeterminate entries and $\mathbf{\Pi}^{(n)} \in \{0,1\}^{I_n \times I}$ is the row-selection matrix used in (86). If the intersection graph $G^{(\mathbf{a}_r)}$ with edges defined by (86) is connected and has the property (87), then $\mathbf{a}_r$ can be obtained from $\mathbf{P}_r^{(\mathbf{A})}$ via a rank-one subspace identification procedure [34]. Finally, $\mathbf{b}_r$ can be obtained by solving the system of linear equations (89).

We summarize the kernel subspace approach for multi-set low-rank factorizations with shared and unshared factors as Algorithm 3, which is based on a use of Theorem B.1. Consequently, if condition (92) in Theorem B.1 is satisfied, then Algorithm 3 is guaranteed to find $\mathbf{A}$ and $\mathbf{B}$ in the exact case. The computational cost of Algorithm 3 is dominated by the computation of $\mathbf{R}$, which can be obtained by means of an EVD of $\mathbf{G}^{(\text{sym})}$ given by (94) with computational cost $\mathcal{O}(R_{\text{tot}}^3)$ with $R_{\text{tot}} = NR + \sum_{n=1}^N Q_n$.

References

[1] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Processing*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.
[2] 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.

Algorithm 3 A kernel subspace intersection approach for multi-set low-rank factorizations of the form (32) with $\mathbf{M}^{(n)} = \mathbf{A}^{(n)} \odot \mathbf{B}^{(n)}$ and based on a constructive use of Theorem B.1.

Input: $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N)}$, R, $Q_1, \ldots, Q_N$.

1. Use SVD to obtain matrix $\mathbf{U}^{(n)}$ whose columns form an orthonormal basis for range($\mathbf{X}^{(n)}$), $n \in \{1, \ldots, N\}$.
3. Build $\mathbf{G}^{(\text{tot})}$ given by (84) from $\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(N)}$.
4. Use SVD to obtain matrix $\mathbf{R}$ given by (85) and whose columns form an orthonormal basis for $\ker(\mathbf{G}^{(\text{tot})}) \cap \pi_S^{(2)}$; see equations (94)–(96) for details.
5. Obtain $\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(N)}$ from CPD of $\mathbf{R}$.
6. Compute $\mathbf{Y}^{(n)} = \mathbf{U}^{(n)}\mathbf{W}^{(n)}$, $n \in \{1, \ldots, N\}$.
7. Obtain $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ from rank-one factorizations of the columns of $\mathbf{Y}^{(n)}$, $n \in \{1, \ldots, N\}$.
8. Obtain $\mathbf{a}_r$ from the rank-one completion of $\mathbf{P}_r^{(\mathbf{A})}$ given by (98), $r \in \{1, \ldots, R\}$.
9. Obtain $\mathbf{b}_r$ by solving the system of linear equations (89), $r \in \{1, \ldots, R\}$.

Output: $\mathbf{A}$ and $\mathbf{B}$.

[3] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 810–823, Mar. 2000.
[4] O. Debals, M. Sohail, and L. De Lathauwer, "Analytical multi-modulus algorithms based on coupled canonical polyadic decompositions," Tech. Rep., 16-150, ESAT-STADIUS, KU Leuven (Leuven, Belgium), 2016.
[5] A. K. Smilde, I. Måge, T. Næs, T. Hankemeier, M. Lips, H. A. L. Kiers, E. Acar, and R. Bro, "Common and distinct components in data fusion," *J. Chemometrics*, vol. 31, no. 7, 2017.
[6] L. De Lathauwer and E. Kofidis, "Coupled matrix-tensor factorizations — the case of partially shared factors," in *Proc. of the 51st Asilomar Conference on Signals, Systems, and Computers*, Oct. 29–Nov. 2, Pacific Grove, CA, USA, 2017.
[7] E. Acar, E. Papalexakis, G. Gurdeniz, M. Rasmussen, A. Lawaetz, M. Nilsson, and R. Bro, "Structure-revealing data fusion," *BMC Bioinformatics*, vol. 15, 2014.
[8] E. Acar, Y. Levin-Schwartz, V. Calhoun, and T. Adali, "ACMTF for fusion of multi-modal neuroimaging data and identification of biomarkers," in *Proc. EUSIPCO 2017*, Aug. 28 – Sep. 2, Kos, Greece, 2017.
[9] M. Cocchi, Ed., *Data fusion methodology and applications*. Elsevier, 2019.
[10] P. Geladi and B. R. Kowalski, "Partial least-squares regression: a tutorial," *Analytica Chimica Acta*, vol. 185, pp. 1–17, 1986.
[11] E. F. Lock, K. A. Hoadley, J. Marron, and A. B. Nobel, "Joint and Individual Variation Explained (JIVE) for integrated analysis of multiple data types," *Ann Appl Stat.*, vol. 7, no. 1, pp. 523–542, 2013.
[12] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3–4, pp. 321–377, 1936.
[13] J. R. Kettenring, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, 1971.
[14] M. Sørensen, C. Kanatsoulis, and N. D. Sidiropoulos, "Generalized canonical correlation analysis: a subspace intersection approach," Tech. Rep., 2020, arXiv:2003.11205.
[15] M. Ibrahim and N. D. Sidiropoulos, "Reliable detection of unknown cell-edge users via canonical correlation analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4170–4182, Jun. 2020.
[16] M. Zhou and A.-J. van der Veen, "Blind separation of partially overlapping data packets," *Digit. Signal Process.*, vol. 68, pp. 154–166, Sep. 2017.
[17] M. van de Velden and T. H. A. Bijmolt, "Generalized canonical correlation analysis of matrices with missing rows: a simulation study," *Psychometrika*, vol. 71, no. 2, pp. 323–331, Jun 2006.
[18] M. Sørensen and N. D. Sidiropoulos, "Joint low-rank factorizations with shared and unshared components: identifiability and algorithms," in *Proc. EUSIPCO 2019*, sep. 2–6, A Coruña, Spain, 2019.
[19] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.
[20] J. D. Carroll and J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.
[21] T. Jiang and N. D. Sidiropoulos, "Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear model with constant modulus constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2625–2636, Sep. 2004.
[22] 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.
[23] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 855–875, 2013.
[24] ——, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Overall uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, 2013.
[25] M. Sørensen, L. De Lathauwer, and N. D. Sidiropoulos, "Bilinear factorizations subject monomial equality constraints via tensor decompositions," *Submitted for publication*.
[26] A. Ben-Israel, "Projectors on intersection of subspaces," in *Contemporary Mathematics*, 2015, vol. 636, pp. 41–50.
[27] M. Sørensen and L. De Lathauwer, "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.
[28] M. Sørensen, I. Domanov, and L. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part II: Algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 3, pp. 1015–1045, 2015.
[29] J. C. Gower and G. B. Dijksterhuis, *Procrustes problems*. Oxford University Press, 2004.
[30] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab v3.0.* [Online]. Available: http://www.tensorlab.net/.
[31] A.-J. van der Veen, "Analytical method for blind binary signal separation," *IEEE Trans. Signal Process.*, vol. 45, no. 4, pp. 1078–1082, Apr. 1997.
[32] L. De Lathauwer, "Decomposition of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.
[33] I. Domanov and L. De Lathauwer, "On uniqueness and computation of the decomposition of a tensor into multilinear rank-$(1, L_r, L_r)$ terms," *SIAM J. Matrix Anal. Appl.*, vol. 41, no. 2, pp. 747–803, 2020.
[34] M. Sørensen and L. De Lathauwer, "Fiber sampling approach to canonical polyadic decomposition and tensor completion," *SIAM J. Matrix Anal. Appl.*, vol. 40, no. 3, pp. 888–917, 2019.

ACKNOWLEDGMENT

This work was supported by NSF ECCS-1608961, NSF ECCS-1807660, and IIS-1704074.

Mikael Sørensen received the Master's degree from Aalborg University, Denmark, and the Ph.D. degree from University of Nice, France, in 2006 and 2010, respectively, both in electrical engineering. From 2010 to 2017 he was a Post-doctoral Fellow with the KU Leuven, Belgium. Since 2018 he has been a research scientist with the University of Virginia.

Nicholas D. Sidiropoulos (F'09) earned the Diploma in Electrical Engineering from Aristotle University of Thessaloniki, Greece, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland at College Park, in 1988, 1990 and 1992, respectively. He has served on the faculty of the University of Virginia, University of Minnesota, and the Technical University of Crete, Greece, prior to his current appointment as Louis T. Rader Professor and Chair of ECE at UVA. From 2015 to 2017 he was an ADC Chair Professor at the University of Minnesota. His research interests are in signal processing, communications, optimization, tensor decomposition, and factor analysis, with applications in machine learning and communications. He received the NSF/CAREER award in 1998, the IEEE Signal Processing Society (SPS) Best Paper Award in 2001, 2007, and 2011, served as IEEE SPS Distinguished Lecturer (2008-2009), and as Vice President - Membership of IEEE SPS (2017-2019). He received the 2010 IEEE Signal Processing Society Meritorious Service Award, and the 2013 Distinguished Alumni Award from the University of Maryland, Dept. of ECE. He is a Fellow of IEEE (2009) and a Fellow of EURASIP (2014).