Joint Low-Rank Factorizations with Shared and Unshared Components: Identifiability and Algorithms

Mikael Sorensen
University of Virginia
ms8tz@virginia.edu

Nikolaos D. Sidiropoulos
University of Virginia
nikos@virginia.edu

***Abstract*—We study the joint low-rank factorization of the matrices X=[A B]G and Y=[A C]H, in which the columns of the shared factor matrix A correspond to vectorized rank-one matrices, the unshared factors B and C have full column rank, and the matrices G and H have full row rank. The objective is to find the shared factor A, given only X and Y. We first explain that if the matrix [A B C] has full column rank, then a basis for the column space of the shared factor matrix A can be obtained from the null space of the matrix [X Y]. This in turn implies that the problem of finding the shared factor matrix A boils down to a basic Canonical Polyadic Decomposition (CPD) problem that in many cases can directly be solved by means of an eigenvalue decomposition. Next, we explain that by taking the rank-one constraint of the columns of the shared factor matrix A into account when computing the null space of the matrix [X Y], more relaxed identifiability conditions can be obtained that do not require that [A B C] has full column rank. The benefit of the unconstrained null space approach is that it leads to simple algorithms while the benefit of the rank-one constrained null space approach is that it leads to relaxed identifiability conditions. Finally, a joint unbalanced orthogonal Procrustes and CPD fitting approach for computing the shared factor matrix A from noisy observation matrices X and Y will briefly be discussed.**

***Index Terms*—Coupled decompositions, canonical polyadic decomposition (CPD), joint low-rank tensor factorizations, joint unbalanced orthogonal Procrustes and CPD fitting, joint dimensionality reduction and CPD fitting.**

I. Introduction

The decomposition of a tensor into a sum of rank-one components, known as the Canonical Polyadic Decomposition (CPD), has found many applications in signal processing and machine learning; see [1] and references therein. We have now entered into an era where data is everywhere. This fact combined with advances in sensor and network technologies have made data fusion and multimodal data analytics important research areas across science and engineering. Coupled CPD models for a set of tensors, in which the rank-one components of the individual tensors are coupled, have been considered for data fusion and multimodal data analytics in signal processing and machine learning. We mention applications in multistatic MIMO radar [2], in joint blind source separation [3], and more recently in the context of high-dimensional statistical modeling [4]. A limitation of the coupled CPD approach is that the coupled rank-one components of the individual tensors are assumed to be *shared* across the collection of decomposed tensors (the meaning of *shared* and *unshared* components will be made clear in Section II). In practice, not all the rank-one components of the individual tensors may be shared. Joint tensor factorization models with shared *and* unshared rank-one components have received little attention in the literature. To the best of our knowledge, only an identifiability result for the basic coupled matrix-tensor factorization model [5] has been reported in [6]. In contrast to the tensor case, joint factorization of matrices with *shared and unshared* components have a long history. A classical example is Canonical Correlation Analysis (CCA) and its variants; see [7] and references therein.

In this paper we first provide in Section II a model for joint low-rank tensor factorizations with shared and unshared components. Based on this modeling framework, in Section III we will present identifiability conditions and algorithms for the recovery of the shared components. The results will explain that by taking into account (i) that the shared components lie in a common subspace and (ii) that the shared components are rank-one structured, more relaxed recovery conditions can be obtained compared to results that do not consider both of the mentioned properties. Before diving into the details, a brief review of the CPD is provided next.

A. Canonical Polyadic Decomposition (CPD)

Consider the CPD of the tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$:

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

where R denotes the rank of $\mathcal{X}$, $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_R] \in \mathbb{C}^{I\times R}$, $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_R] \in \mathbb{C}^{J\times R}$, $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_R] \in \mathbb{C}^{K\times R}$ are the CPD factor matrices of $\mathcal{X}$ and '$\circ$' denotes the outer product, e.g., $(\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r)_{ijk} = a_{ir}b_{jr}c_{kr}$. Note that $\mathbf{H}^{(r)} = \mathbf{a}_r\mathbf{b}_r^T$ is a rank-1 matrix. This fact will be exploited in Section III-B. We will consider the following matrix representation of (1):

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

where '$\odot$' denotes the Khatri–Rao (columnwise Kronecker) product and '$(\cdot)^T$' denotes the transpose. The rows of $\mathbf{X}$ correspond to the mode-3 fibers $\{\mathbf{x}_{ij\bullet}\}$ of $\mathcal{X}$, defined as $(\mathbf{x}_{ij\bullet})_{ijk} = x_{ijk}$.

A key feature of the CPD that will be used in Section III is that it is unique under mild conditions, i.e., **A**, **B** and **C** are unique (up to intrinsic column scaling and permutation ambiguities); see [1], [8]–[12] and references therein. In this paper we will make use of the relatively easy to check sufficient uniqueness condition stated in Theorem 1 below. It makes use of the matrix $C_2(\mathbf{A}) \odot C_2(\mathbf{B})$, in which $C_2(\mathbf{A}) \in \mathbb{C}^{C_I^2 \times C_R^2}$ denotes the second-order compound matrix containing the determinants of all 2×2 submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order [10], and where $C_I^2 = \frac{1}{2}I(I-1)$ and $C_R^2 = \frac{1}{2}R(R-1)$ denote binomial coefficients. (Similarly for $C_2(\mathbf{B}) \in \mathbb{C}^{C_J^2 \times C_R^2}$.)

Theorem 1: Consider the CPD of the tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in (1). If

$$\begin{cases} \mathbf{C} \ \text{has full column rank,} \\ C_2(\mathbf{A}) \odot C_2(\mathbf{B}) \ \text{has full column rank,} \end{cases} \tag{3}$$

then the CPD of $\mathcal{X}$ is unique [8]–[11]. Generically, condition (3) is satisfied if $C_R^2 \le C_I^2 C_J^2$ and $R \le K$ [9], [13].

II. Joint low-rank factorizations with shared and unshared components

In this paper we are interested in joint low-rank factorizations involving R_1 shared components and $R_{2,1} + R_{2,2}$ unshared components:

$$\begin{cases} \mathbf{X}^{(1)} = \mathbf{M}^{(1,1)}\mathbf{C}^{(1,1)T} + \mathbf{M}^{(2,1)}\mathbf{C}^{(2,1)T} \in \mathbb{C}^{IJ \times K_1}, \\ \mathbf{X}^{(2)} = \mathbf{M}^{(1,1)}\mathbf{C}^{(1,2)T} + \mathbf{M}^{(2,2)}\mathbf{C}^{(2,2)T} \in \mathbb{C}^{IJ \times K_2}, \end{cases} \tag{4}$$

where the columns of the shared factor matrix $\mathbf{M}^{(1,1)} \in \mathbb{C}^{IJ \times R_1}$ are vectorized rank-1 matrices:

$$\mathbf{M}^{(1,1)} = \mathbf{A} \odot \mathbf{B} \in \mathbb{C}^{IJ \times R_1}, \ \mathbf{A} \in \mathbb{C}^{I \times R_1}, \ \mathbf{B} \in \mathbb{C}^{J \times R_1}, \tag{5}$$

and where the unshared factor matrices $\mathbf{M}^{(2,1)} \in \mathbb{C}^{IJ \times R_{2,1}}$ and $\mathbf{M}^{(2,2)} \in \mathbb{C}^{IJ \times R_{2,2}}$ are not necessarily structured. We also assume that the matrices

$$\mathbf{C}^{(1)} = [\mathbf{C}^{(1,1)}, \mathbf{C}^{(2,1)}] \in \mathbb{C}^{K_1 \times (R_1 + R_{2,1})},$$
$$\mathbf{C}^{(2)} = [\mathbf{C}^{(1,2)}, \mathbf{C}^{(2,2)}] \in \mathbb{C}^{K_2 \times (R_1 + R_{2,2})},$$

have full column rank, where $\mathbf{C}^{(1,1)} \in \mathbb{C}^{K_1 \times R_1}$, $\mathbf{C}^{(2,1)} \in \mathbb{C}^{K_1 \times R_{2,1}}$, $\mathbf{C}^{(1,2)} \in \mathbb{C}^{K_1 \times R_1}$ and $\mathbf{C}^{(2,2)} \in \mathbb{C}^{K_2 \times R_{2,2}}$. W.l.o.g. we can now assume that $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$ are nonsingular ($K_1 = R_1 + R_{2,1}$ and $K_2 = R_1 + R_{2,2}$). We say that the components associated with $\mathbf{M}^{(1,1)}$ are shared between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ while the components associated with $\mathbf{M}^{(2,1)}$ and $\mathbf{M}^{(2,2)}$ are unshared. Note that when there are no unshared components ($R_{2,1} = R_{2,2} = 0$), then the joint low-rank factorization (4) reduces to a CPD of $[\mathbf{X}^{(1)}, \ \mathbf{X}^{(2)}] = (\mathbf{A} \odot \mathbf{B})[\mathbf{C}^{(1,1)T}, \mathbf{C}^{(1,2)T}]$ (cf. Eq. (2)). The goal is now to find the shared factor $\mathbf{M}^{(1,1)} = \mathbf{A} \odot \mathbf{B}$, given only $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. Solutions to this problem will be discussed in Section III.

III. Identifiability conditions and computation of shared components $\mathbf{M}^{(1,1)}$

In this section we will demonstrate that by exploiting both the common subspace structure range$(\mathbf{M}^{(1,1)})$ between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ and the rank-one structures in $\mathbf{M}^{(1,1)}$, relaxed identifiability conditions for the shared components $\mathbf{M}^{(1,1)} = \mathbf{A} \odot \mathbf{B}$ can be obtained. We also briefly discuss algorithms for the computation of the shared factors $\mathbf{A}$ and $\mathbf{B}$.

A. Exploiting common subspace structure

Two-channel factor analysis for detection of unknown cell-edge users in communication systems via CCA have been proposed in [14]. A related approach for finding range$(\mathbf{M}^{(1,1)})$ will now be presented. More precisely, we first obtain an identifiability condition for the shared components $\mathbf{M}^{(1,1)}$ that exploits the shared subspace range$(\mathbf{M}^{(1,1)}\mathbf{C}^{(1,1)T})$ = range$(\mathbf{M}^{(1,1)}\mathbf{C}^{(1,2)T})$ between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, where $'$range$(\cdot)'$ denotes the range of a matrix. More precisely, we will consider the subspace $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$, where $'\ker(\cdot)'$ denotes the kernel of a matrix. Note that the minimal dimension of $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$ is R_1. Indeed, let $\mathbf{e}_i^{(I)} \in \mathbb{C}^I$ denote a unit vector with unit entry at the ith position and zeros elsewhere. Then for any column $\mathbf{m}_r^{(1,1)}$ of $\mathbf{M}^{(1,1)} \in \mathbb{C}^{IJ \times R_1}$ we can obviously find a vector $\mathbf{n}_r = [((\mathbf{C}^{(1)T})^{-1}\mathbf{e}_r^{(K_1)})^T, -((\mathbf{C}^{(2)T})^{-1}\mathbf{e}_r^{(K_2)})^T]^T \in \mathbb{C}^{(2R_1 + R_{2,1} + R_{2,2})}$, $1 \le r \le R_1$, such that $[\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]\mathbf{n}_r = \mathbf{m}_r^{(1,1)} - \mathbf{m}_r^{(1,1)} = \mathbf{0}$. The assumption that the dimension of $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$ is minimal (i.e., R_1) implies that

$$\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]) = \text{range}\left(\begin{bmatrix} \mathbf{Q}^{(1)} \\ \mathbf{Q}^{(2)} \end{bmatrix}\right), \tag{6}$$

where $\mathbf{Q}^{(1)} \in \mathbb{C}^{(R_1 + R_{2,1}) \times R_1}$ and $\mathbf{Q}^{(2)} \in \mathbb{C}^{(R_1 + R_{2,2}) \times R_1}$ are matrices with property range$(\mathbf{X}^{(1)}\mathbf{Q}^{(1)})$ = range$(\mathbf{X}^{(2)}\mathbf{Q}^{(2)})$ = range$(\mathbf{M}^{(1,1)})$. In words, the column space of $\mathbf{M}^{(1,1)}$ can be obtained via the null space of $[\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]$. Note that in general, we only have that range$\left(\begin{bmatrix} \mathbf{Q}^{(1)} \\ \mathbf{Q}^{(2)} \end{bmatrix}\right) \subseteq \ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$. However, it is not hard to see that if the matrix $[\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}] \in \mathbb{C}^{IJ \times (R_1 + R_{2,1} + R_{2,2})}$ has full column rank, then R_1 is also the maximal dimension of $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$, i.e., if $[\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}]$ has full column rank, then relation (6) holds. In detail, consider a vector $\mathbf{n} \in \ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$ with partitioning $\mathbf{n} = [\mathbf{n}_1, -\mathbf{n}_2]^T \in \mathbb{C}^{(2R_1 + R_{2,1} + R_{2,2})}$ in which $\mathbf{n}_1 \in \mathbb{C}^{(R_1 + R_{2,1})}$ and $\mathbf{n}_2 \in \mathbb{C}^{(R_1 + R_{2,2})}$. Using this partitioning, relation $[\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]\mathbf{n} = \mathbf{0}$ can be written as

$$[\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}] \begin{bmatrix} \mathbf{C}^{(1,1)T}\mathbf{n}_1 - \mathbf{C}^{(1,2)T}\mathbf{n}_2 \\ \mathbf{C}^{(2,1)T}\mathbf{n}_1 \\ -\mathbf{C}^{(2,2)T}\mathbf{n}_2 \end{bmatrix} = \mathbf{0}. \tag{7}$$

The full column rank assumption on $[\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}]$ implies that (7) reduces to

$$\begin{bmatrix} \mathbf{C}^{(1,1)T}\mathbf{n}_1 - \mathbf{C}^{(1,2)T}\mathbf{n}_2 \\ \mathbf{C}^{(2,1)T}\mathbf{n}_1 \\ -\mathbf{C}^{(2,2)T}\mathbf{n}_2 \end{bmatrix} = \mathbf{0}. \tag{8}$$

Note that range$(\mathbf{M}^{(1)})$ = range$(\mathbf{M}^{(1)}\mathbf{C}^{(1)T})$ and range$(\mathbf{M}^{(2)})$ = range$(\mathbf{M}^{(2)}\mathbf{C}^{(2)T})$. In other words, we are only interested in the involved subspaces. Consequently, we can w.l.o.g. make use of change-of-basis transforms such that $\mathbf{M}^{(1)}\mathbf{C}^{(1)T} = (\mathbf{M}^{(1)}\mathbf{C}^{(1)T}) \cdot ((\mathbf{C}^{(1)T})^{-1}\mathbf{C}^{(1)T})$ and $\mathbf{M}^{(2)}\mathbf{C}^{(2)T} = (\mathbf{M}^{(2)}\mathbf{C}^{(2)T}) \cdot ((\mathbf{C}^{(2)T})^{-1}\mathbf{C}^{(2)T})$. Let $\mathbf{I}_m \in \mathbb{C}^{m \times m}$ denote the identity matrix. Then this implies that we can set $\mathbf{C}^{(1)} = [\mathbf{C}^{(1,1)}, \mathbf{C}^{(2,1)}] = \mathbf{I}_{R_1 + R_{2,1}}$ and $\mathbf{C}^{(2)} = [\mathbf{C}^{(1,2)}, \mathbf{C}^{(2,2)}] = \mathbf{I}_{R_1 + R_{2,2}}$ in relation (8):

$$\begin{bmatrix} \mathbf{n}_{11} - \mathbf{n}_{22} \\ \mathbf{n}_{12} \\ -\mathbf{n}_{22} \end{bmatrix} = \mathbf{0}, \quad (9)$$

where $\mathbf{n}_1 = [\mathbf{n}_{11}^T, \mathbf{n}_{21}^T]^T \in \mathbb{C}^{(R_1+R_{2,1})}$ with $\mathbf{n}_{11} \in \mathbb{C}^{R_1}$ and $\mathbf{n}_{21} \in \mathbb{C}^{R_{2,1}}$ and $\mathbf{n}_2 = [\mathbf{n}_{12}^T, \mathbf{n}_{22}^T]^T \in \mathbb{C}^{(R_1+R_{2,2})}$ with $\mathbf{n}_{12} \in \mathbb{C}^{R_1}$ and $\mathbf{n}_{22} \in \mathbb{C}^{R_{2,2}}$. Clearly, condition (9) is only satisfied when $\mathbf{n}_{12} = \mathbf{0}$, $\mathbf{n}_{22} = \mathbf{0}$ and for any $\mathbf{n}_{11} = \mathbf{n}_{22} \in \mathbb{C}^{R_1}$. The latter property implies that there exist at most R_1 linearly independent vectors in $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$. We can now conclude that if $[\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}]$ has full column rank, then the subspace $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$ is indeed R_1-dimensional and that relation (6) holds. Let us summarize the common subspace identifiability condition in Lemma 1 below.

Lemma 1: Consider the matrix $[\mathbf{X}^{(1)}, \mathbf{X}^{(2)}] \in \mathbb{C}^{IJ \times (K_1+K_2)}$, in which $\mathbf{X}^{(1)} \in \mathbb{C}^{IJ \times K_1}$ and $\mathbf{X}^{(2)} \in \mathbb{C}^{IJ \times K_2}$ are given by (4). If

$$\begin{cases} [\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}] \text{ has full column rank}, \\ \mathbf{C}^{(1)} \text{ and } \mathbf{C}^{(2)} \text{ have full column rank}, \end{cases} \quad (10)$$

then $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}]) = \text{range}\left(\begin{bmatrix} \mathbf{Q}^{(1)} \\ \mathbf{Q}^{(2)} \end{bmatrix}\right)$, where $\mathbf{Q}^{(1)}$ and $\mathbf{Q}^{(2)}$ are matrices with property $\text{range}(\mathbf{X}^{(1)}\mathbf{Q}^{(1)}) = \text{range}(\mathbf{X}^{(2)}\mathbf{Q}^{(2)}) = \text{range}(\mathbf{M}^{(1,1)})$. Generically, the conditions in (10) hold if $R_1 + R_{2,1} + R_{2,2} \leq IJ$, $R_1 + R_{2,1} \leq K_1$ and $R_1 + R_{2,2} \leq K_2$.

Note that when the Khatri–Rao structure of $\mathbf{M}^{(1,1)} = \mathbf{A} \odot \mathbf{B}$ is ignored, then it is clear from (7) that the full column rank condition (10) is also necessary. In Section III-B we will demonstrate that when the Khatri–Rao structure $\mathbf{M}^{(1,1)} = \mathbf{A} \odot \mathbf{B}$ is taken into account, the full column rank condition on $[\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}]$ can be relaxed without sacrificing the uniqueness of the shared components $\mathbf{M}^{(1,1)} = \mathbf{A} \odot \mathbf{B}$. An important practical implication of (6) is that a basis for $\text{range}(\mathbf{M}^{(1,1)})$ can be obtained from it. More precisely, since $\text{range}(\mathbf{X}^{(1)}\mathbf{Q}^{(1)}) = \text{range}(\mathbf{X}^{(2)}\mathbf{Q}^{(2)}) = \text{range}(\mathbf{M}^{(1,1)})$, there exists a nonsingular change-of-basis matrix $\mathbf{F} \in \mathbb{C}^{R_1 \times R_1}$ such that we obtain $\mathbf{Y} \in \mathbb{C}^{IJ \times R_1}$ with factorizations

$$\mathbf{Y} := \mathbf{X}^{(1)}\mathbf{Q}^{(1)} = -\mathbf{X}^{(2)}\mathbf{Q}^{(2)} = \mathbf{M}^{(1,1)}\mathbf{F}^T = (\mathbf{A} \odot \mathbf{B})\mathbf{F}^T. \quad (11)$$

Clearly, relation (11) corresponds to a matrix representation of the CPD of a tensor

$$\mathcal{Y} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{f}_r \in \mathbb{C}^{I \times J \times R_1}. \quad (12)$$

Since the factor matrix $\mathbf{F}$ has full column rank, the combination of Lemma 1 and Theorem 1 yields a uniqueness condition for the shared CPD factors $\mathbf{A}$ and $\mathbf{B}$. We summarize the result as Theorem 2 below.

Theorem 2: Consider the joint low-rank factorization of $\mathbf{X}^{(1)} \in \mathbb{C}^{IJ \times K_1}$ and $\mathbf{X}^{(2)} \in \mathbb{C}^{IJ \times K_2}$ in (4) with R_1 shared components and $R_{2,1} + R_{2,2}$ unshared components. If

$$\begin{cases} [\mathbf{M}^{(1,1)}, \mathbf{M}^{(2,1)}, \mathbf{M}^{(2,2)}] \text{ has full column rank}, \\ \mathbf{C}^{(1)} \text{ and } \mathbf{C}^{(2)} \text{ have full column rank}, \\ C_2(\mathbf{A}) \odot C_2(\mathbf{B}) \text{ has full column rank}, \end{cases} \quad (13)$$

then the shared factor matrices $\mathbf{A}$ and $\mathbf{B}$ are unique. Generically, condition (13) holds if $R_1 + R_{2,1} + R_{2,2} \leq IJ$, $R_1 + R_{2,1} \leq K_1$, $R_1 + R_{2,2} \leq K_2$ and $C_{R_1}^2 \leq C_I^2 C_J^2$.

An interesting property of Theorem 2 is that it admits a constructive interpretation. In short, if condition (13) is satisfied, then the CPD of $\mathcal{Y}$ given by (12), and obtained via $\ker([\mathbf{X}^{(1)}, \mathbf{X}^{(2)}])$ as earlier explained, can be computed by means of an EigenValue Decomposition (EVD) [1], [9], [12].

B. Exploiting both common subspace and rank-one structures

We will now exploit both the common subspace structure $\text{range}(\mathbf{M}^{(1,1)})$ and the rank-one structures between and within $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. This means that we are now looking for a pair of vectors $\mathbf{v} \in \mathbb{C}^{(R_1+R_{2,1})}$ and $\mathbf{w} \in \mathbb{C}^{(R_1+R_{2,2})}$ with properties

$$\mathbf{X}^{(1)}\mathbf{v} - \mathbf{X}^{(2)}\mathbf{w} = \mathbf{0}, \quad (14)$$

$$\mathbf{X}^{(1)}\mathbf{v} = \mathbf{a} \otimes \mathbf{b}, \quad (15)$$

$$\mathbf{X}^{(2)}\mathbf{w} = \mathbf{a} \otimes \mathbf{b}, \quad (16)$$

where '$\otimes$' denotes the Kronecker product. Note that relation (14) takes the common subspace structure of $\text{range}(\mathbf{M}^{(1,1)})$ and the rank-one structure between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ into account. Likewise, relations (15) and (16) take the rank-one structures within $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ into account. In the next sections we explain how to combine all of the mentioned structures.

1) Exploiting common subspace structure range($\boldsymbol{M}^{(1,1)}$): Using (14), the common subspace structure $\text{range}(\mathbf{M}^{(1,1)})$ can be exploited. In Section III-B5 it will become clear that it is more convenient to work with the equivalent expressions

$$(\mathbf{I}_{R_1+R_{2,1}} \otimes \mathbf{X}^{(1)})(\mathbf{v} \otimes \mathbf{v}) - (\mathbf{I}_{R_1+R_{2,1}} \otimes \mathbf{X}^{(2)})(\mathbf{v} \otimes \mathbf{w}) = \mathbf{0}, \quad (17)$$

$$(\mathbf{I}_{R_1+R_{2,2}} \otimes \mathbf{X}^{(1)})(\mathbf{w} \otimes \mathbf{v}) - (\mathbf{I}_{R_1+R_{2,2}} \otimes \mathbf{X}^{(2)})(\mathbf{w} \otimes \mathbf{w}) = \mathbf{0}. \quad (18)$$

2) Exploiting rank-one structure between $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$: Using (14), we can also exploit the rank-one structure between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. For ease of presentation, we use the following notation for the rank-one terms in (4):

$$\mathbf{H}^{(r)} = \mathbf{a}_r \circ \mathbf{b}_r = \mathbf{a}_r \mathbf{b}_r^T \in \mathbb{C}^{I \times J}, \quad r \in \{1, \ldots, R_1\}. \quad (19)$$

The rank-one constraint implies that

$$\begin{vmatrix} h_{i_1 j_1}^{(r)} & h_{i_1 j_2}^{(r)} \\ h_{i_2 j_1}^{(r)} & h_{i_2 j_2}^{(r)} \end{vmatrix} = h_{i_1 j_1}^{(r)} h_{i_2 j_2}^{(r)} - h_{i_2 j_1}^{(r)} h_{i_1 j_2}^{(r)} = 0, \quad (20)$$

where $1 \leq i_1 < i_2 \leq I$ and $1 \leq j_1 < j_2 \leq J$. The combination of (14) and (20) yields

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_1}^{(J)})^T \mathbf{X}^{(1)}\mathbf{v} & (\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_2}^{(J)})^T \mathbf{X}^{(2)}\mathbf{w} \\ (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_1}^{(J)})^T \mathbf{X}^{(1)}\mathbf{v} & (\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_2}^{(J)})^T \mathbf{X}^{(2)}\mathbf{w} \end{vmatrix} = \mathbf{q}^{(n,1,2)} (\mathbf{v} \otimes \mathbf{w}) = 0, \quad (21)$$

where $\mathbf{q}^{(n,1,2)} = ((\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_1}^{(J)})^T \mathbf{X}^{(1)}) \otimes ((\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_2}^{(J)})^T \mathbf{X}^{(2)}) - ((\mathbf{e}_{i_2}^{(I)} \otimes \mathbf{e}_{j_1}^{(J)})^T \mathbf{X}^{(1)}) \otimes ((\mathbf{e}_{i_1}^{(I)} \otimes \mathbf{e}_{j_2}^{(J)})^T \mathbf{X}^{(2)})$, in which the superscript 'n' in the row-vector $\mathbf{q}^{(n,1,2)} \in \mathbb{C}^{1 \times (2R_1+R_{2,1}+R_{2,2})^2}$ takes all the subscripts i_1, i_2, j_1 and j_2 into account. Throughout this section the variables $L = 2R_1 + R_{2,1} + R_{2,2}$ and $N = C_I^2 C_J^2$ will be used. Stacking yields

$$\mathbf{Q}^{(N,1,2)}\left(\mathbf{v}\otimes\mathbf{w}\right)=\mathbf{0}, \tag{22}$$

where $\mathbf{Q}^{(N,1,2)} = [\mathbf{q}^{(1,1,2)T},\ldots,\mathbf{q}^{(N,1,2)T}]^T \in \mathbb{C}^{N\times L^2}$. In Section III-B5 it will become clear that it is more convenient to also consider the following combination of (14) and (20):

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(2)}\mathbf{w} & (\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(1)}\mathbf{v} \\ (\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(2)}\mathbf{w} & (\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(1)}\mathbf{v} \end{vmatrix} = \mathbf{q}^{(n,2,1)}\left(\mathbf{w}\otimes\mathbf{v}\right)=0, \tag{23}$$

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

$$\mathbf{Q}^{(N,2,1)}\left(\mathbf{w}\otimes\mathbf{v}\right)=\mathbf{0}, \tag{24}$$

where $\mathbf{Q}^{(N,2,1)} = [\mathbf{q}^{(1,2,1)T},\ldots,\mathbf{q}^{(N,2,1)T}]^T \in \mathbb{C}^{N\times L^2}$.

3) Exploiting rank-one structure within $\mathbf{X}^{(1)}$: The combination of (15) and (20) yields

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(1)}\mathbf{v} & (\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(1)}\mathbf{v} \\ (\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(1)}\mathbf{v} & (\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(1)}\mathbf{v} \end{vmatrix} = \mathbf{q}^{(n,1,1)}\left(\mathbf{v}\otimes\mathbf{v}\right)=0, \tag{25}$$

where $\mathbf{q}^{(n,1,1)} = ((\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(1)})\otimes((\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(1)}) - ((\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(1)})\otimes((\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(1)})$, and the superscript '$n$' in the row-vector $\mathbf{q}^{(n,1,1)}\in\mathbb{C}^{1\times(R_1+R_{2,1})^2}$ takes all the subscripts i_1, i_2, j_1 and j_2 into account. Stacking yields

$$\mathbf{Q}^{(N,1,1)}\left(\mathbf{v}\otimes\mathbf{v}\right)=\mathbf{0}, \tag{26}$$

where $\mathbf{Q}^{(N,1,1)} = [\mathbf{q}^{(1,1,1)T},\ldots,\mathbf{q}^{(N,1,1)T}]^T \in \mathbb{C}^{N\times(R_1+R_{2,1})^2}$.

4) Exploiting rank-one structure within $\mathbf{X}^{(2)}$: As before, the combination of (16) and (20) yields

$$\begin{vmatrix} (\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(2)}\mathbf{w} & (\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(2)}\mathbf{w} \\ (\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(2)}\mathbf{w} & (\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(2)}\mathbf{w} \end{vmatrix} = \mathbf{q}^{(n,2,2)}\left(\mathbf{w}\otimes\mathbf{w}\right)=0, \tag{27}$$

where $\mathbf{q}^{(n,2,2)} = ((\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(2)})\otimes((\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(2)}) - ((\mathbf{e}_{i_2}^{(I)}\otimes\mathbf{e}_{j_1}^{(J)})^T\mathbf{X}^{(2)})\otimes((\mathbf{e}_{i_1}^{(I)}\otimes\mathbf{e}_{j_2}^{(J)})^T\mathbf{X}^{(2)})$, and the superscript '$n$' in the row-vector $\mathbf{q}^{(n,2,2)}\in\mathbb{C}^{1\times(R_1+R_{2,2})^2}$ takes all the subscripts i_1, i_2, j_1 and j_2 into account. Stacking yields

$$\mathbf{Q}^{(N,2,2)}\left(\mathbf{w}\otimes\mathbf{w}\right)=\mathbf{0}, \tag{28}$$

where $\mathbf{Q}^{(N,2,2)} = [\mathbf{q}^{(1,2,2)T},\ldots,\mathbf{q}^{(N,2,2)T}]^T \in \mathbb{C}^{N\times(R_1+R_{2,2})^2}$.

5) Combination of common subspace and rank-one structures: The combination of the common subspace structures (17) and (18) and the rank-one structures (22), (24), (26) and (28) yields

$$\mathbf{G}^{(\text{tot})}\left(\left[\begin{smallmatrix}\mathbf{v}\\ \mathbf{w}\end{smallmatrix}\right]\otimes\left[\begin{smallmatrix}\mathbf{v}\\ \mathbf{w}\end{smallmatrix}\right]\right)=\mathbf{0}, \tag{29}$$

where $\mathbf{G}^{(\text{tot})}\in\mathbb{C}^{N_{\text{tot}}\times L^2}$ is given by

$$\mathbf{G}^{(\text{tot})} = \begin{bmatrix} [(\mathbf{I}_{R_1+R_{2,1}}\otimes\mathbf{X}^{(1)}),-(\mathbf{I}_{R_1+R_{2,1}}\otimes\mathbf{X}^{(2)}),\mathbf{0}]\mathbf{\Pi}_1 \\ [(\mathbf{I}_{R_1+R_{2,2}}\otimes\mathbf{X}^{(1)}),-(\mathbf{I}_{R_1+R_{2,2}}\otimes\mathbf{X}^{(2)}),\mathbf{0}]\mathbf{\Pi}_2 \\ [\mathbf{Q}^{(N,1,1)},\mathbf{0}]\mathbf{\Pi}_3 \\ [\mathbf{Q}^{(N,1,2)},\mathbf{0}]\mathbf{\Pi}_4 \\ [\mathbf{Q}^{(N,2,2)},\mathbf{0}]\mathbf{\Pi}_5 \\ [\mathbf{Q}^{(N,2,1)},\mathbf{0}]\mathbf{\Pi}_6 \end{bmatrix}, \tag{30}$$

in which $N_{\text{tot}} = LIJ+4N$, $\{\mathbf{\Pi}_n\}$ denote appropriate column permutation matrices and $\{\mathbf{0}\}$ denote zero matrices of conformable sizes. Let S_L denote the C_L^2-dimensional subspace of vectorized $(L\times L)$ symmetric matrices. From (29) it is clear that if the dimension of the subspace $\ker(\mathbf{G}^{(\text{tot})})\cap S_L$ is minimal (i.e., R_1), then $\mathbf{V}$ and $\mathbf{W}$ can be obtained from it. In more detail, let the columns of $\mathbf{R}\in\mathbb{C}^{L^2\times R_1}$ form a basis for $\ker(\mathbf{G}^{(\text{tot})})\cap S_L$, then there exists a nonsingular change-of-basis matrix $\mathbf{F}\in\mathbb{C}^{R_1\times R_1}$ such that

$$\mathbf{R} = (\left[\begin{smallmatrix}\mathbf{V}\\ \mathbf{W}\end{smallmatrix}\right]\odot\left[\begin{smallmatrix}\mathbf{V}\\ \mathbf{W}\end{smallmatrix}\right])\mathbf{F}^T. \tag{31}$$

Clearly, (31) corresponds to a third-order tensor

$$\mathcal{R} = \sum_{r=1}^{R_1}\left[\begin{smallmatrix}\mathbf{v}_r\\ \mathbf{w}_r\end{smallmatrix}\right]\circ\left[\begin{smallmatrix}\mathbf{v}_r\\ \mathbf{w}_r\end{smallmatrix}\right]\circ\mathbf{f}_r\in\mathbb{C}^{L\times L\times R_1}, \tag{32}$$

whose CPD is unique. We summarize the result as Theorem 3 below.

Theorem 3: Consider the joint low-rank factorization of $\mathbf{X}^{(1)}\in\mathbb{C}^{IJ\times K_1}$ and $\mathbf{X}^{(2)}\in\mathbb{C}^{IJ\times K_2}$ in (4) with R_1 shared components and $R_{2,1}+R_{2,2}$ unshared components. If

$$\begin{cases} \mathbf{C}^{(1)} \text{ and } \mathbf{C}^{(2)} \text{ have full column rank,} \\ \ker(\mathbf{G}^{(\text{tot})})\cap S_L \text{ is an } R_1\text{-dimensional subspace,} \end{cases} \tag{33}$$

then the shared factor matrices $\mathbf{A}$ and $\mathbf{B}$ are unique.

Theorem 3 can be understood as a version of Theorem 1 for joint low-rank factorizations with shared and unshared components in which we are interested in the shared components. Note that condition (33) in Theorem 3 is only sufficient and it can be improved upon by the use of tensorization methods [12].

A nice property of Theorem 3 is that it leads to relaxed identifiability conditions. We will now consider an example that demonstrates that improved identifiability conditions can be obtained by simultaneously exploiting the common subspace structure range$(\mathbf{M}^{(1,1)})$ and the rank-one structures between and within $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. In Table I we report upper bounds on $R_1+R_{2,1}+R_{2,2}$ as a function of $I=J$ when $R_1=10$ and conditions (13) and (33) are used. By inspection of the table it is clear that improved identifiability conditions can be obtained by simultaneously exploiting (i) the common subspace structure range$(\mathbf{M}^{(1,1)})$, (ii) the rank-one structure between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ and (iii) the rank-one structures within $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. Note also that when the Khatri–Rao structure of $\mathbf{M}^{(1,1)}$ is taken into account, the uniqueness of $\mathbf{A}$ and $\mathbf{B}$ can be guaranteed, despite the fact that $[\mathbf{M}^{(1,1)},\mathbf{M}^{(2,1)},\mathbf{M}^{(2,2)}]$ does not have full column rank.

$I=J$	5	6	7	8
condition (13)	25	36	49	64
condition (33)	29	47	68	93

TABLE I

AN UPPER BOUND ON $R_1+R_{2,1}+R_{2,2}$ AS A FUNCTION OF $I=J$ WHEN $R_1=10$ AND CONDITIONS (13) AND (33) ARE USED.

Another interesting property of Theorem 3 is that it admits a constructive interpretation. In short, if condition (32) is satisfied, then $\mathbf{V}$ and $\mathbf{W}$ can be obtained via the CPD of $\mathcal{R}$

given by (32). which in turn can be computed by means of an EVD [1], [9], [12]. Once $\mathbf{V}$ and $\mathbf{W}$ have been computed, $\mathbf{A}$ and $\mathbf{B}$ can be obtained via (15) and (16). (Due to space limitations, further details are deferred to a journal version.)

C. A joint unbalanced orthogonal Procrustes and CPD fitting approach (joint dimensionality reduction and CPD fitting)

In practice, the decomposition (4) is rarely exact. For this reason, we will now briefly discuss an optimization-based approach for computing the shared factors $\mathbf{A}$ and $\mathbf{B}$ from noisy observation matrices $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$. Consider the matrix $\left[\begin{smallmatrix}\mathbf{V}\\ \mathbf{W}\end{smallmatrix}\right]$ in (31) and whose columns have the properties (14)–(16). Let $\mathbf{Z}\mathbf{F}^{-T} = \left[\begin{smallmatrix}\mathbf{V}\\ \mathbf{W}\end{smallmatrix}\right]$ denote the factorization in which $\mathbf{Z} \in \mathbb{C}^{(2R_1+R_{2,1}+R_{2,2})\times R_1}$ is columnwise orthonormal ($\mathbf{Z}^H\mathbf{Z} = \mathbf{I}_{R_1}$) and $\mathbf{F} \in \mathbb{C}^{R_1\times R_1}$ is nonsingular. As an alternative to (4) we propose to compute $\mathbf{A}$ and $\mathbf{B}$ via

$$f(\mathbf{Z},\mathbf{A},\mathbf{B},\mathbf{F}) = \left\| \begin{bmatrix} \mathbf{X}^{(1)} & \mathbf{0} \\ \mathbf{0} & -\mathbf{X}^{(2)} \end{bmatrix} \mathbf{Z} - \begin{bmatrix} \mathbf{A}\odot\mathbf{B} \\ \mathbf{A}\odot\mathbf{B} \end{bmatrix} \mathbf{F}^T \right\|_F^2, \quad (34)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Observe that if we fix $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{F}$, then the problem of minimizing (34) corresponds to an Unbalanced Orthogonal Procrustes (UOP) problem [15]. Likewise, if we fix $\mathbf{Z}$, then (34) basically corresponds to a CPD fitting problem. Hence, the problem of minimizing (34) can be interpreted as a joint UOP and CPD fitting problem. This can also be understood as a joint dimensionality reduction and CPD fitting approach for computing $\mathbf{A}$ and $\mathbf{B}$.

Let us end the section with an illustrative example. Consider (4) with $I = J = 7$, $K_1 = R_1 + R_{2,1}$, $K_2 = R_1 + R_{2,2}$, $R_{2,1} = R_{2,2} = 5$ and varying R_1. The goal is to estimate $\mathbf{A}$ and $\mathbf{B}$ from $\mathbf{Y}^{(n)} = \mathbf{X}^{(n)} + \mathbf{N}^{(n)}$, $n \in \{1,2\}$, where $\mathbf{N}^{(n)}$ is an unstructured perturbation matrix. In each trial of the Monte Carlo experiment, the involved factor and noise perturbation matrices are randomly drawn from a Gaussian distribution with zero mean and unit variance. As a performance measure we use the distance between $\mathbf{A}$ and its estimate, $\widehat{\mathbf{A}}$. The distance is measured according to the criterion: $\mathrm{P}(\mathbf{A}) = \min_{\mathbf{\Pi}\mathbf{\Lambda}} \left\|\mathbf{A} - \widehat{\mathbf{A}}\mathbf{\Pi}\mathbf{\Lambda}\right\|_F / \|\mathbf{A}\|_F$, where $\mathbf{\Pi}$ and $\mathbf{\Lambda}$ denote a permutation matrix and a diagonal matrix, respectively. We compare the joint UOP and CPD fitting approach[1] with the two algebraic EVD-based methods associated with Theorems 2 and 3. The mean $\mathrm{P}(\mathbf{A})$ over 50 Monte Carlo runs in which $R_1 = 5$ or $R_1 = 10$ are shown in Figure 1. A gain in performance is observed when the perturbation noise is taken into account.

IV. Conclusion

We first presented identifiability conditions for joint low-rank factorizations with shared and unshared components. In particular, we showed that by jointly exploiting the common subspace and low-rank structures of the shared components, more relaxed identifiability conditions can be obtained. Next, we also briefly explained that in the noiseless case, the shared components can be computed by means of an EVD. Finally, a joint UOP and CPD fitting approach for computing the shared components in the noisy case was proposed.

[1]Numerically, we minimize (34) by means of an alternating optimization method that alternates between the updates of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{F}$ and $\mathbf{Z}$. For the updates of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{F}$ standard conditional least squares updates are used. For the update of $\mathbf{Z}$ a column-by-column based updating approach is used in which the spherical constraint $\|\mathbf{z}_r\|_F = 1$ is first relaxed and then imposed in a subsequent projection step.

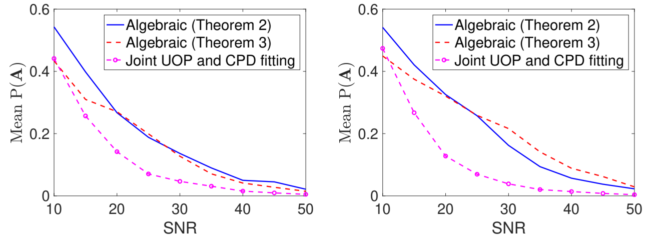

Fig. 1. (Left) Mean $\mathrm{P}(\mathbf{A})$ values over 50 Monte Carlo runs when $R_1 = 5$. (Right) Mean $\mathrm{P}(\mathbf{A})$ values over 50 Monte Carlo runs when $R_1 = 10$.

Acknowledgment

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

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] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Processing*, vol. 58, no. 11, pp. 5693–5705, Nov. 2010.

[3] X.-F. Gong, Q.-H. Lin, F.-Y. Cong, and L. De Lathauwer, "Double coupled canonical polyadic decomposition for joint blind source separation," *IEEE Trans. Signal Processing*, vol. 66, pp. 3475–3490, 2018.

[4] N. Kargas, N. D. Sidiropoulos, and X. Fu, "Tensors, learning, and 'Kolmogorov extension' for finite-alphabet random vectors," *IEEE Trans. Signal Processing*, vol. 66, no. 18, pp. 4854–4868, Sep. 2018.

[5] E. Acar, R. Bro, and A. Smilde, "Data fusion in metabolomics using coupled matrix and tensor factorizations," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1602–1620, 2015.

[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] X. Fu, K. Huang, M. Hong, N. D. Sidiropoulos, and A. M. So, "Scalable and flexible multiview max-var canonical correlation analysis," *IEEE Trans. Signal Processing*, vol. 16, pp. 4150–4165, 2017.

[8] 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.

[9] 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.

[10] 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.

[11] ——, "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.

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

[13] A. Stegeman, J.M.F. ten Berge, and L. De Lathauwer, "Sufficient conditions for uniqueness in Candecomp/Parafac and Indscal with random component matrices," *Psychometrika*, vol. 71, pp. 219–229, 2006.

[14] M. Ibrahim and N. D. Sidiropoulos, "Cell-edge interferometry: Reliable detection of unknown cell-edge users via canonical correlation analysis," submitted to *IEEE SPAWC*, July 2-5, 2019, Cannes, France.

[15] L. Elden and H. Park, "A Procrustes problem on Stiefel manifolds," *Numer. Math.*, vol. 82, pp. 599–619, 1999.