FIBER SAMPLING APPROACH TO CANONICAL POLYADIC DECOMPOSITION AND TENSOR COMPLETION

MIKAEL SØRENSEN * AND LIEVEN DE LATHAUWER †

Abstract. Tensor decompositions play an important role in a variety of applications, such as signal processing and machine learning. In practice, the tensor can be incomplete or very large, making it more difficult to analyse and process using conventional tensor decomposition results. In this paper we focus on the basic Canonical Polyadic Decomposition (CPD). We propose an algebraic framework for CPD of tensors that have missing fibers. This includes fiber sampled extensions of a uniqueness condition and an algebraic algorithm originally developed for the ordinary CPD. In particular, we explain that tensor decomposition problems involving missing fibers can be reduced to relatively simple matrix completion problems via a matrix eigenvalue decomposition. By capitalizing on results from matrix completion, we briefly explain that the fiber sampling approach can be extended to cases where the tensor entries are missing at random. Fiber sampling also leads to a new algebraic multi-grid subsampling scheme, making it interesting for large scale problems. In particular, we explain that the CPD of a large-scale tensor can be found via a smaller-scaled coupled CPD.

Key words. tensors, polyadic decomposition, fiber sampling, missing data.

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. Tensor methods have found many applications in psychometrics, chemometrics, signal processing, machine learning, data compression, and so on; see [29] and references therein. In practice, the low-rank properties of a tensor may not be easy to directly exploit. An example is the Canonical Polyadic Decomposition (CPD) model with missing entries, which is commonly used in the context of tensors with incomplete, corrupt and large scale data (e.g., [19, 38, 1, 43, 29]). In this paper we focus on the problem of determining the CPD of an incomplete tensor in which fibers are missing. The problem of CPD with missing fibers appears in chemometrics due to scattering [38], in EEG processing based on time-frequency-measurement analysis where some of the electrodes are malfunctioning [1], and in NMR spectroscopy due to nonuniform sampling [25, 40, 46, 24]. A more recent application of fiber sampled CPD is tensor-based graph clustering in which higher-order network structures are exploited [41]. However, no uniqueness conditions have been proposed in the mentioned applications. To the best of our knowledge, the first formal study of fiber sampled CPD was carried out by the authors in the conference paper [32] in the context of sensor array processing. We note in passing that a problem related to CPD with missing fibers is tensor completion (e.g., [15]). The difference is that the latter problem is concerned about finding the missing entries of the incomplete tensor and not neces-

*mr.mikael.sorensen@gmail.com.

†KU Leuven - E.E. Dept. (ESAT) - STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, the Group Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium, and iMinds Medical IT, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, Lieven.DeLathauwer@kuleuven.be.

Research supported by: (1) Research Council KU Leuven: C1 project C16/15/059-nD, (2) F.W.O.: project G.0830.14N, G.0881.14N, (3) the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012–2017), (4) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

sarily about finding the factors of an underlying CPD. Another notable difference is that many of the existing matrix/tensor completion methods (e.g., [4, 27, 28, 15, 45]) usually rely on probabilistic sampling conditions while the proposed fiber sampling method relies on an easy-to-check deterministic condition.

As our first contribution, we present a new algebraic framework for CPD with missing fibers. In particular, we provide a coupled CPD [30, 31] interpretation that reduces the problem into a set of relatively simple matrix completion problems. The approach leads to a new easy-to-check and yet powerful uniqueness condition. We also briefly explain that by capitalizing on results from matrix completion, the fiber sampling approach can be extended to cases where the tensor entries are missing at random.

As our second contribution, we explain that if the presented uniqueness condition is satisfied, then the CPD of a tensor with missing fibers can be determined via a matrix EigenValue Decomposition (EVD), despite missing fibers. Fiber sampling also leads to a new algebraic multi-grid subsampling scheme for tensors that enjoy a low-rank CPD structure, which makes it interesting for large-scale problems.

The paper is organized as follows. The rest of the introduction presents the notation. Section 2 reviews the necessary algebraic prerequisites. In Section 3 we present the algebraic framework for CPD with missing fibers, which leads to a new deterministic uniqueness condition that is relatively easy to check. It also leads to the earlier mentioned multi-grid sampling scheme that computes the CPD of a large-scale tensor via the coupled CPD of a set of smaller-scaled tensors. In Section 4 we develop an algorithm that transforms the CPD problem with missing fibers into a simultaneous matrix diagonalization problem, which in turn is transformed into simple matrix completion problems via a matrix EVD. Numerical experiments are reported in Section 5 where it is demonstrated that the multi-grid fiber sampling scheme can be used to find the CPD of a large-scale low-rank tensor without a significant loss in terms of performance, even when less than one percent of tensor entries are used in the computation.

Notation. Vectors, matrices and tensors are denoted by lower case boldface, upper case boldface and upper case calligraphic letters, respectively. The rth column, conjugate, transpose, conjugate-transpose, inverse, Moore-Penrose pseudoinverse, Frobenius norm, rank, range and kernel of a matrix $\mathbf{A}$ are denoted by $\mathbf{a}_r$, $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^{-1}$, $\mathbf{A}^\dagger$, $\|\mathbf{A}\|_F$, $r_{\mathbf{A}}$, $\text{range}\,(\mathbf{A})$ and $\ker(\mathbf{A})$, respectively.

The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri-Rao product, defined as

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

in which $(\mathbf{A})_{mn} = a_{mn}$. The symbol $*$ denotes the Hadamard product, e.g., $(\mathcal{A} * \mathcal{B})_{ijk} = a_{ijk}b_{ijk}$ in the case of third-order tensors. 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 $\omega(\mathbf{x})$ in the tensor decomposition literature, dating back to the work of Kruskal [22]. $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$ denotes the diagonal matrix holding row k of $\mathbf{A} \in \mathbb{C}^{I\times J}$ on its diagonal. $\mathbf{I}_N \in \mathbb{C}^{N\times N}$ and $\mathbf{e}_n^{(N)} \in \mathbb{C}^N$ denote the identity matrix and the unit vector with unit entry at position n and zeros elsewhere, 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 $C_k(\mathbf{A}) \in \mathbb{C}^{C_I^k \times C_R^k}$. 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.

2. Algebraic Prerequisites.

2.1. Canonical Polyadic Decomposition (CPD). Consider the third-order tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$. We say that $\mathcal{X}$ is a rank-1 tensor if it is equal to the outer product of non-zero vectors $\mathbf{a} \in \mathbb{C}^I$, $\mathbf{b} \in \mathbb{C}^J$ and $\mathbf{c} \in \mathbb{C}^K$ such that $x_{ijk} = a_i b_j c_k$. A Polyadic Decomposition (PD) is a decomposition of $\mathcal{X}$ into a sum of rank-1 terms [18, 3]:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \,. \tag{2.1}$$

The rank of a tensor $\mathcal{X}$ is equal to the minimal number of rank-1 tensors that yield $\mathcal{X}$ in a linear combination. Assume that the rank of $\mathcal{X}$ is R, then (2.1) is called the CPD of $\mathcal{X}$. Let us stack the vectors $\{\mathbf{a}_r\}$, $\{\mathbf{b}_r\}$ and $\{\mathbf{c}_r\}$ into the matrices $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R]$, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R]$ and $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ will be referred to as the factor matrices of the PD or CPD of $\mathcal{X}$ in (2.1).

2.1.1. Matrix Representation. Let $\mathbf{X}^{(i\cdot\cdot)} \in \mathbb{C}^{J\times K}$ denote the matrix such that $(\mathbf{X}^{(i\cdot\cdot)})_{jk} = x_{ijk}$, then $\mathbf{X}^{(i\cdot\cdot)} = \mathbf{B} D_i(\mathbf{A}) \mathbf{C}^T$ and we obtain the following matrix representation of (2.1):

$$\mathbb{C}^{IJ\times K} \ni \mathbf{X} := \left[\mathbf{X}^{(1\cdot\cdot)T}, \ldots, \mathbf{X}^{(I\cdot\cdot)T}\right]^T = (\mathbf{A} \odot \mathbf{B})\, \mathbf{C}^T. \tag{2.2}$$

2.1.2. Uniqueness. The rank-1 tensors in (2.1) can be arbitrarily permuted and scaled provided that the overall rank-1 term remains the same. We say that the CPD is unique when it is only subject to these trivial indeterminacies. CPD uniqueness has been intensively studied in recent years, see [9, 10, 13, 29] and references therein. In many practical problems at least one factor has full column rank. For this case the following necessary and sufficient uniqueness condition was obtained in [21] and later reformulated in terms of compound matrices in [9]. It makes use of the vector

$$\mathbf{d}^{(2)} = [d_1 d_2, d_1 d_3, \ldots, d_{R-1} d_R]^T \in \mathbb{C}^{C_R^2}. \tag{2.3}$$

Note that $\mathbf{d}^{(2)}$ consists of all products of entries (d_r, d_s) from the vector $\mathbf{d} = [d_1, \ldots, d_R]^T$ minus the R products $d_1 d_1, d_2 d_2, \ldots, d_R d_R$.

THEOREM 2.1. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in (2.1). Let $\boldsymbol{d}^{(2)} \in \mathbb{C}^{C_R^2}$ be built from $\boldsymbol{d} \in \mathbb{C}^R$ as in (2.3). Assume that $\boldsymbol{C}$ has full column rank. The rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique if and only if*

$$(C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B}))\boldsymbol{d}^{(2)} = \boldsymbol{0} \Rightarrow \omega(\boldsymbol{d}) \leq 1. \tag{2.4}$$

Generically[1]*, condition (2.4) is satisfied if and only if $R \leq (I-1)(J-1)$ [37, 5, 12].*

[1] A generic property is a property that holds everywhere except for a set of Lebesgue measure zero. In particular, a generic tensor decomposition property is a property that holds with probability one if the entries of the factor matrices are randomly drawn from continuous distributions.

In practice, condition (2.4) can be hard to check. Observe that if $C_2(\mathbf{A}) \odot C_2(\mathbf{B})$ in (2.4) has full column rank, then $\mathbf{d}^{(2)} = \mathbf{0}$ and the condition is immediately satisfied. This fact leads to the following easy-to-check uniqueness condition, which is only sufficient.

THEOREM 2.2. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in (2.1). If*

$$\begin{cases} \boldsymbol{C} \text{ has full column rank,} \\ C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B}) \text{ has full column rank,} \end{cases} \tag{2.5}$$

then the rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique [21, 7, 9]. Generically, condition (2.5) is satisfied if $C_R^2 \leq C_I^2 C_J^2$ and $R \leq K$ [7, 34].

Moreover, under condition (2.5) the CPD of $\mathcal{X}$ can be converted into the CPD of an $(R \times R \times R)$ tensor $\mathcal{M}$ of rank R, even in cases where $\max(I, J) < R$ [7, 11]. The latter CPD can be computed by means of a standard EVD (e.g., [23]).

2.2. Coupled CPD. We say that a collection of tensors $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$, $n \in \{1, \ldots, N\}$, admits an R-term coupled PD if each tensor $\mathcal{X}^{(n)}$ can be written as [30]:

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r\,, \quad n \in \{1, \ldots, N\}, \tag{2.6}$$

with factor matrices

$$\mathbf{A}^{(n)} = [\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}],\ \mathbf{B}^{(n)} = [\mathbf{b}_1^{(n)}, \ldots, \mathbf{b}_R^{(n)}] \ \text{ and } \ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]\,.$$

We define the coupled rank of $\{\mathcal{X}^{(n)}\}$ as the minimal number of coupled rank-1 tensors $\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r$ that yield $\{\mathcal{X}^{(n)}\}$ in a linear combination. Assume that the coupled rank of $\{\mathcal{X}^{(n)}\}$ is R, then (2.6) will be called the coupled CPD of $\{\mathcal{X}^{(n)}\}$.

2.2.1. Matrix Representation. The coupled PD or CPD of $\{\mathcal{X}^{(n)}\}$ given by (2.6) has the following matrix representation

$$\mathbf{X} = \left[\mathbf{X}^{(1)T}, \ldots, \mathbf{X}^{(N)T}\right]^T = \mathbf{F}\mathbf{C}^T \in \mathbb{C}^{(\sum_{n=1}^{N} I_n J_n) \times K}, \tag{2.7}$$

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

2.2.2. Uniqueness. The coupled rank-1 tensors in (2.6) can be arbitrarily permuted and the vectors within the same coupled rank-1 tensor can be arbitrarily scaled provided the overall coupled rank-1 term remains the same. We say that the coupled CPD is unique when it is only subject to these trivial indeterminacies. Uniqueness conditions for the coupled CPD were derived in [30]. Theorem 2.3 extends Theorem 2.2 to coupled CPD. It makes use of the matrix

$$\mathbf{G}^{(N)} = \begin{bmatrix} C_2\left(\mathbf{A}^{(1)}\right) \odot C_2\left(\mathbf{B}^{(1)}\right) \\ \vdots \\ C_2\left(\mathbf{A}^{(N)}\right) \odot C_2\left(\mathbf{B}^{(N)}\right) \end{bmatrix} \in \mathbb{C}^{\left(\sum_{n=1}^{N} C_{I_n}^2 C_{J_n}^2\right) \times C_R^2}. \tag{2.8}$$

THEOREM 2.3. *Consider the coupled PD of $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$, $n \in \{1, \ldots, N\}$ in (2.6). If*

$$\begin{cases} \mathbf{C} \text{ in (2.7) has full column rank,} \\ \boldsymbol{G}^{(N)} \text{ in (2.8) has full column rank,} \end{cases} \tag{2.9}$$

then the coupled rank of $\{\mathcal{X}^{(n)}\}$ *is* R *and the coupled CPD of* $\{\mathcal{X}^{(n)}\}$ *is unique [30]. Generically, condition* (2.9) *is satisfied if* $C_R^2 \leq \sum_{n=1}^{N} C_{I_n}^2 C_{J_n}^2$ *and* $R \leq K$.

Furthermore, if condition (2.9) is satisfied, then the coupled CPD of $\mathcal{X}$ can be converted into the CPD of an $(R \times R \times R)$ tensor $\mathcal{M}$ of rank R [30], which in turn can be computed by means of a standard EVD.

3. Fiber sampled rank decompositions. In this section a link between low-rank decompositions of tensors that have missing fibers and coupled decompositions is presented. In particular, we consider the CPD of a tensor that has missing fibers. The overall idea is to interpret a tensor with missing fibers as a collection of (possibly partially overlapping) full tensors. Fiber sampling also leads to a new multi-grid (sub)sampling scheme suitable for large-scale tensor decomposition problems.

3.1. CPD with missing fibers. Consider the CPD model with missing entries:

$$\mathcal{Y} = \mathcal{W} * \mathcal{X} = \mathcal{W} * \left(\sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \right), \tag{3.1}$$

where $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$ denotes the observed incomplete tensor and $\mathcal{W} \in \{0,1\}^{I \times J \times K}$ is a binary tensor where $'0'$ and $'1'$ indicate that the corresponding entry of $\mathcal{X}$ is unknown and known, respectively. The goal is to recover the CPD factors $\{\mathbf{a}_r\}, \{\mathbf{b}_r\}, \{\mathbf{c}_r\}$ of $\mathcal{X}$, observing $\mathcal{Y}$. In this section we focus on the special case of missing fibers. We say that the tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ is missing a fiber if for some pair $(i,j) \in \{1,\ldots,I\} \times \{1,\ldots,J\}$ the vector $\mathbf{x}_{ij\bullet} \in \mathbb{C}^K$, defined by $(\mathbf{x}_{ij\bullet})_k = x_{ijk}$, is unobserved. A key observation is that in the case of missing fibers, $\mathcal{W}$ in (3.1) has a particular structure. Namely, if the fiber $\mathbf{x}_{pq\bullet}$ is missing, then $w_{pqk} = 0$, $\forall k \in \{1,\ldots,K\}$. The structure of $\mathcal{W}$ obviously affects the conditions under which the CPD of $\mathcal{X}$ may be recovered, observing $\mathcal{Y}$.

3.1.1. Necessary conditions for CPD recovery. Recall that a necessary condition for the recovery of $\mathbf{C}$ from $\mathbf{X} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T$ is that $\mathbf{A} \odot \mathbf{B}$ has full column rank (e.g., [35]). Here we extend the condition to the case of missing entries. From (2.2), we know that (3.1) has following matrix representation

$$\mathbf{Y} = \mathbf{W} * \mathbf{X} = \mathbf{W} * \left((\mathbf{A} \odot \mathbf{B})\, \mathbf{C}^T \right). \tag{3.2}$$

Let $\mathbf{y}_k$, $\mathbf{w}_k$ and $\mathbf{x}_k$ denote the kth column of $\mathbf{Y}$, $\mathbf{W}$ and $\mathbf{X}$, respectively. Then expression (3.2) can be written as

$$\mathbf{y}_k = \mathbf{w}_k * \mathbf{x}_k = \mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})\tilde{\mathbf{c}}_k, \quad k \in \{1,\ldots,K\}, \tag{3.3}$$

where $\mathbf{D}_{\mathbf{W}_k} = D_1\left(\mathbf{w}_k^T\right)$ and $\widetilde{\mathbf{C}} = [\tilde{\mathbf{c}}_1,\ldots,\tilde{\mathbf{c}}_k] = \mathbf{C}^T$. From (3.3) it is clear that in order to uniquely recover $\mathbf{C}$ from $\mathcal{Y}$, the matrices $\{\mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})\}$ must have full column rank. Indeed, if $\mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})$ does not have full column rank, then for any $\mathbf{x} \in \ker\left(\mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})\right)$ we have $\mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})\,\tilde{\mathbf{c}}_k = \mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})\,(\tilde{\mathbf{c}}_k + \mathbf{x})$. The following proposition summarizes the result.

PROPOSITION 3.1. *Consider the PD of* $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ *in* (2.1) *observed via* $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$ *in* (3.1). *For a unique recovery of* $\boldsymbol{C}$ *from* $\mathcal{Y}$ *it is necessary that* $\boldsymbol{D}_{\boldsymbol{W}_k}(\boldsymbol{A} \odot \boldsymbol{B})$ *in* (3.3) *has full column rank for every* $k \in \{1,\ldots,K\}$.

In the special case of missing fibers $\{\mathbf{x}_{pq\bullet}\}$, the indicator pattern $\mathbf{D}_{\mathbf{W}_k} = D_1\left(\mathbf{w}_k^T\right)$ is the same for every column $\mathbf{x}_k$ of $\mathbf{X}$, i.e., $\mathbf{D}_{\mathbf{W}_1} = \mathbf{D}_{\mathbf{W}_k}$, $\forall k \in \{1,\ldots,K\}$. Assume that $IJ - F$ fibers of $\mathcal{X}$ are missing, then relation (3.2) reduces to

$$\mathbf{Y}_{\text{sub}} = \mathbf{S}_{\mathbf{W}_1}\mathbf{X} = \mathbf{\Phi}(\mathbf{A},\mathbf{B})\mathbf{C}^T \in \mathbb{C}^{F \times K}, \tag{3.4}$$

where $\mathbf{Y}_{\text{sub}}$ is the observed submatrix of $\mathbf{Y}$, $\mathbf{\Phi}(\mathbf{A},\mathbf{B}) := \mathbf{S}_{\mathbf{W}_1}(\mathbf{A}\odot\mathbf{B}) \in \mathbb{C}^{F\times R}$ and $\mathbf{S}_{\mathbf{W}_1} \in \mathbb{C}^{F\times IJ}$ is the row-selection matrix that selects the F rows of $\mathbf{X}$ that have not been zeroed by $\mathbf{D}_{\mathbf{W}_1}$. The necessary condition for recovery of $\mathbf{C}$ from $\mathbf{Y}_{\text{sub}}$ reduces now to $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ having full column rank. Note that $\mathbf{S}_{\mathbf{W}_1}^T\mathbf{\Phi}(\mathbf{A},\mathbf{B}) = \mathbf{D}_{\mathbf{W}_1}(\mathbf{A}\odot\mathbf{B})$.

Besides the recovery of $\mathbf{C}$ from $\mathbf{Y}_{\text{sub}}$, we also consider the recovery of $\mathbf{A}$ and $\mathbf{B}$ from $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ in the case of missing fibers. Let $\boldsymbol{\varphi}_r \in \mathbb{C}^{IJ}$ denote the rth column of $\mathbf{D}_{\mathbf{W}_1}(\mathbf{A}\odot\mathbf{B})$. Then we define the matrix $\mathbf{P}^{(r)} \in \mathbb{C}^{I\times J}$ with missing entries according to

$$p_{ij}^{(r)} = \begin{cases} (\boldsymbol{\varphi}_r)_{(i-1)I+j}, & \text{if } (\mathbf{w}_1)_{(i-1)I+j} = 1, \\ \text{indeterminate}, & \text{if } (\mathbf{w}_1)_{(i-1)I+j} = 0. \end{cases} \tag{3.5}$$

The support of $\mathbf{P}^{(r)}$, denoted by $\mathcal{P}^{(r)}$, consists of all pairs (i,j) for which the entry $p_{ij}^{(r)}$ of $\mathbf{P}^{(r)}$ is determined: $(i,j) \in \mathcal{P}^{(r)}$ if $(\mathbf{w}_1)_{(i-1)I+j} = 1$, i.e., if $\mathbf{x}_{ij\bullet}$ is observed.

The term $\mathbf{a}_r\mathbf{b}_r^T$ can be obtained by looking for a rank-1 matrix that completes $\mathbf{P}^{(r)}$. Clearly, the rank-1 matrix $\mathbf{a}_r\mathbf{b}_r^T$ completes $\mathbf{P}^{(r)}$, meaning that in our setting a rank-1 completion matrix always exists. However, if there exists a row or a column of $\mathbf{P}^{(r)}$ containing only indeterminate entries, then there exists an alternative rank-1 matrix $\widetilde{\mathbf{a}}_r\widetilde{\mathbf{b}}_r^T$ that completes $\mathbf{P}^{(r)}$, which is not identical to $\mathbf{a}_r\mathbf{b}_r^T$. Hence, every row and column of $\mathbf{P}^{(r)}$ must at least have one determinate entry. More formally, for a unique recovery of $\mathbf{A}$ and $\mathbf{B}$ from $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ it is necessary that $\mathbf{P}^{(r)}$ has the property

$$\begin{cases} \forall i \in \{1,\dots,I\},\, \exists j' \in \{1,\dots,J\} :\, p_{ij'}^{(r)} \text{ is determinate } \Leftrightarrow (\mathbf{w}_1)_{(i-1)I+j'} = 1, \\ \forall j \in \{1,\dots,J\},\, \exists i' \in \{1,\dots,I\} :\, p_{i'j}^{(r)} \text{ is determinate } \Leftrightarrow (\mathbf{w}_1)_{(i'-1)I+j} = 1. \end{cases} \tag{3.6}$$

On the other hand, it is also clear that since the rank-1 matrix $\mathbf{a}_r\mathbf{b}_r^T$ contains $I+J-1$ free variables, $\mathbf{P}^{(r)}$ must contain at least $I+J-1$ determinate entries for the completion to be unique.

3.1.2. Sufficient condition for the recovery of one factor matrix. In the rest of the paper we limit the study to the cases where $\mathbf{C}$ in (3.2) has full column rank, which is most often the case in practice. Proposition 3.2 below extends a well-known necessary and sufficient uniqueness condition [21] to the case of missing fibers. In short, it states that if none of the column vectors $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ in (3.4) can be written as a linear combination of its remaining vectors, then $\mathbf{C}$ can be uniquely recovered from $\mathbf{Y}_{\text{sub}}$. Proposition 3.2 below makes use of the binary diagonal matrix $\mathbf{D}_{\text{sel}} \in \{0,1\}^{C_I^2C_J^2\times C_I^2C_J^2}$ which holds the vector $\mathbf{d}_{\text{sel}} \in \{0,1\}^{C_I^2C_J^2}$ on its diagonal, i.e., $\mathbf{D}_{\text{sel}} = D_1(\mathbf{d}_{\text{sel}}^T)$. The entries of the vector

$$\mathbf{d}_{\text{sel}} = [w_{(1,2),(1,2)},\ w_{(1,2),(1,3)},\ \dots,\ w_{(I-1,I),(J-1,J)}]^T \tag{3.7}$$

are given by

$$w_{(u,v),(p,q)} = \begin{cases} 1, \text{ if fibers } \mathbf{x}_{up\bullet}, \mathbf{x}_{uq\bullet}, \mathbf{x}_{vp\bullet} \text{ and } \mathbf{x}_{vq\bullet} \text{ are observed}, \\ 0, \text{ otherwise.} \end{cases}$$

PROPOSITION 3.2. *Consider a tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$, partially observed as $\mathcal{Y} = \mathcal{W}*\mathcal{X}$, and its PD given by* (2.1). *Assume that $\boldsymbol{C}$ has full column rank. Let $\boldsymbol{d}^{(2)} \in \mathbb{C}^{C_R^2}$*

be the vector given by (2.3) *and let* $\boldsymbol{D}_{sel} = D_1(\boldsymbol{d}_{sel}^T)$ *in which* $\boldsymbol{d}_{sel}$ *is defined as in* (3.7). *The rank of* $\mathcal{X}$ *is* R *and the factor matrix* $\boldsymbol{C}$ *is unique if and only if*

$$\boldsymbol{D}_{sel}(C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B}))\boldsymbol{d}^{(2)} = \boldsymbol{0} \Rightarrow \omega(\boldsymbol{d}) \leq 1. \tag{3.8}$$

Proof.

Sufficiency of condition (3.8)**.** Let the triplet $(\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}})$ be an alternative decomposition of (2.1) composed of $\widehat{R} \leq R$ terms so that

$$\mathbf{Y} = \boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{C}^T = \boldsymbol{\Phi}(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})\widehat{\mathbf{C}}^T. \tag{3.9}$$

Using Kruskal's permutation lemma [22] (see also [35]), uniqueness of $\mathbf{C}$ can be established.[2] Briefly, in cases where $\mathbf{C}$ has full column rank, Kruskal's permutation lemma states that if $\omega(\mathbf{C}^T\mathbf{z}) \leq \omega(\widehat{\mathbf{C}}^T\mathbf{z})$ for every vector $\mathbf{z} \in \mathbb{C}^K$ such that $\omega(\widehat{\mathbf{C}}^T\mathbf{z}) \leq 1$, then $\mathbf{C} = \widehat{\mathbf{C}}\boldsymbol{\Pi}\Delta_{\mathbf{C}}$, where $\boldsymbol{\Pi}$ is an $(R \times R)$ column permutation matrix and $\Delta_{\mathbf{C}}$ is an $(R \times R)$ nonsingular diagonal matrix. In more detail, it will be that condition (3.8) implies that $\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})$ has full column rank. This fact together with the assumption that $\mathbf{C}$ has full column rank implies that $\widehat{\mathbf{C}}$ must also have full column rank (recall that $\widehat{R} \leq R \leq K$). Denote $\mathbf{d} = \mathbf{C}^T\mathbf{z}$ and $\widehat{\mathbf{d}} = \widehat{\mathbf{C}}^T\mathbf{z}$. Kruskal's permutation lemma now guarantees uniqueness of $\mathbf{C}$ if $\omega(\mathbf{d}) \leq \omega(\widehat{\mathbf{d}})$ for every $\omega(\widehat{\mathbf{d}}) \leq R - r_{\widehat{\mathbf{C}}} + 1 = 1$. Thus, we only have to verify that this condition holds for the two cases $\omega(\widehat{\mathbf{d}}) = 0$ and $\omega(\widehat{\mathbf{d}}) = 1$.

Case $\omega(\widehat{\mathbf{d}}) = 0$**.** Let us first consider the case $\omega(\widehat{\mathbf{d}}) = 0 \Leftrightarrow \widehat{\mathbf{C}}^T\mathbf{z} = \mathbf{0}_K$. At the end of this subsection it will be explained that condition (3.8) implies that any column of $\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})$ can only be written as a trivial linear combination of its columns. In detail, assume that there exists a vector $\mathbf{d} \in \mathbb{C}^R$ with properties $\omega(\mathbf{d}) \geq 2$ and $\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \mathbf{0}$, then it will be shown that the latter property implies that $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))\mathbf{d}^{(2)} = \mathbf{0}$, which contradicts condition (3.8). We conclude that if condition (3.8) is satisfied, then $\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \mathbf{0} \Rightarrow \omega(\mathbf{d}) \leq 1$, which in turn implies that $\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})$ has full column rank. Hence, we know from (3.9) that

$$\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{C}^T\mathbf{z} = \boldsymbol{\Phi}(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})\widehat{\mathbf{C}}^T\mathbf{z} = \mathbf{0} \Leftrightarrow \mathbf{C}^T\mathbf{z} = \mathbf{0}_K\,,$$

where we took into account that $\widehat{\mathbf{d}} = \widehat{\mathbf{C}}^T\mathbf{z} = \mathbf{0}_K$. In other words, we must have that $\mathbf{d} = \mathbf{C}^T\mathbf{z} = \mathbf{0}_K$ for all $\mathbf{z} \in \mathbb{C}^K$ such that $\omega(\widehat{\mathbf{d}}) = 0$. We conclude that the inequality condition $0 = \omega(\mathbf{C}^T\mathbf{z}) \leq \omega(\widehat{\mathbf{C}}^T\mathbf{z}) = 0$ in Kruskal's permutation lemma is satisfied.

Case $\omega(\widehat{\mathbf{d}}) = 1$**.** Consider again a vector $\mathbf{z} \in \mathbb{C}^K$ so that from (3.9) we obtain

$$\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{C}^T\mathbf{z} = \boldsymbol{\Phi}(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})\widehat{\mathbf{C}}^T\mathbf{z} \Leftrightarrow \boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \boldsymbol{\Phi}(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})\widehat{\mathbf{d}}, \tag{3.10}$$

where $\mathbf{d} = \mathbf{C}^T\mathbf{z}$ and $\widehat{\mathbf{d}} = \widehat{\mathbf{C}}^T\mathbf{z}$. Assume that the vector $\mathbf{z} \in \mathbb{C}^K$ is chosen so that $\omega(\widehat{\mathbf{d}}) = \omega(\widehat{\mathbf{C}}^T\mathbf{z}) = 1$. Reshaping the vector given by (3.10) into an $(I \times J)$ matrix

[2] Kruskal's permutation lemma is not limited to cases where $\mathbf{C}$ has full column rank. In fact, it allows us to extend Proposition 3.2 to cases where $\mathbf{C}$ does not have full column rank.

yields

$$\mathbf{W}^{(\cdot\cdot\bullet)} * \left(\sum_{r=1}^{R} \mathbf{a}_r \mathbf{b}_r^T d_r \right) = \mathbf{W}^{(\cdot\cdot\bullet)} * \left(\sum_{r=1}^{R} \widehat{\mathbf{a}}_r \widehat{\mathbf{b}}_r^T \widehat{d}_r \right), \tag{3.11}$$

where $\mathbf{W}^{(\cdot\cdot\bullet)} \in \{0,1\}^{I \times J}$ is the binary matrix defined according to

$$(\mathbf{W}^{(\cdot\cdot\bullet)})_{ij} = \begin{cases} 1, & \text{if fiber } \mathbf{x}_{ij\bullet} \text{ is sampled,} \\ 0, & \text{otherwise.} \end{cases}$$

Since $\omega(\widehat{\mathbf{d}}) = 1$ all observed (2×2) submatrices of (3.11) must be rank deficient. In other words, the second-order minors of all observed (2×2) submatrices of (3.11) must vanish. The second-order minors of (3.11) are of the form

$$\begin{vmatrix} \sum_{r=1}^{R} w_{i_1 j_1}^{(\cdot\cdot\bullet)} a_{i_1 r} b_{j_1 r} \cdot d_r & \sum_{r=1}^{R} w_{i_1 j_2}^{(\cdot\cdot\bullet)} a_{i_1 r} b_{j_2 r} \cdot d_r \\ \sum_{r=1}^{R} w_{i_2 j_1}^{(\cdot\cdot\bullet)} a_{i_2 r} b_{j_1 r} \cdot d_r & \sum_{r=1}^{R} w_{i_2 j_2}^{(\cdot\cdot\bullet)} a_{i_2 r} b_{j_2 r} \cdot d_r \end{vmatrix} . \tag{3.12}$$

We will now provide a condition in terms of vanishing second-order minors. In words, we will interpret the incomplete tensor decomposition of $\mathcal{Y}$ with missing fibers as a set of $C_I^2 C_J^2$ coupled $2 \times 2 \times K$ tensor decompositions and only retain the complete tensors within this set. More formally, let Γ denote the set of all quadruples (i_1, i_2, j_1, j_2) with properties $i_1 < i_2$ and $j_1 < j_2$ so that $\operatorname{card}(\Gamma) = C_I^2 C_J^2$. The elements in Γ will be used to index submatrices of $\mathbf{A}$ and $\mathbf{B}$. Namely, if $\gamma = (i_1, i_2, j_1, j_2)$, then

$$\mathbf{A}^{(i_1,i_2)} := \mathbf{A}([i_1, i_2], :) \in \mathbb{C}^{2 \times R} \quad \text{and} \quad \mathbf{B}^{(j_1,j_2)} := \mathbf{B}([j_1, j_2], :) \in \mathbb{C}^{2 \times R} .$$

In words, $\mathbf{A}^{(i_1,i_2)}$ consists of the i_1-th and the i_2-th rows of $\mathbf{A}$. (Similarly for $\mathbf{B}^{(j_1,j_2)}$.) Note that we are essentially interpreting the CPD of $\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$ as a coupled CPD involving all possible tensors of the form $\mathcal{X}^{(i_1,i_2,j_1,j_2)} := \mathcal{X}([i_1, i_2], [j_1, j_2], :) = \sum_{r=1}^{R} \mathbf{a}_r^{(i_1,i_2)} \circ \mathbf{b}_r^{(j_1,j_2)} \circ \mathbf{c}_r$. (More details about this interpretation will be provided in the proof of Proposition 3.3.) The vanishing of the second-order minors implies that for every $\gamma \in \Gamma$ with property $w_{i_1 j_1}^{(\cdot\cdot\bullet)} = w_{i_1 j_2}^{(\cdot\cdot\bullet)} = w_{i_2 j_1}^{(\cdot\cdot\bullet)} = w_{i_2 j_2}^{(\cdot\cdot\bullet)} = 1$, the observed (2×2) minor of (3.11) must satisfy the relation

$$\left| \sum_{r=1}^{R} \mathbf{a}_r^{(i_1,i_2)} \mathbf{b}_r^{(j_1,j_2)T} d_r \right| = 0 . \tag{3.13}$$

Stacking the equations (3.13) as rows of the matrix $\mathbf{G} \in \mathbb{C}^{C_I^2 C_J^2 \times C_R^2}$ in accordance to the lexicographical rule yields:

$$\mathbf{G}\mathbf{d}^{(2)} = \mathbf{0} , \tag{3.14}$$

where $\mathbf{d}^{(2)} = [d_1 d_2, d_1 d_3, \ldots, d_{R-1} d_R]^T \in \mathbb{C}^{C_R^2}$ and

$$\mathbf{G} = \begin{bmatrix} w_{(1,2),(1,2)} \cdot C_2\left(\mathbf{A}^{(1,2)}\right) * C_2\left(\mathbf{B}^{(1,2)}\right) \\ w_{(1,2),(1,3)} \cdot C_2\left(\mathbf{A}^{(1,2)}\right) * C_2\left(\mathbf{B}^{(1,3)}\right) \\ \vdots \\ w_{(I-1,I),(J-1,J)} \cdot C_2\left(\mathbf{A}^{(I-1,I)}\right) * C_2\left(\mathbf{B}^{(J-1,J)}\right) \end{bmatrix} = \mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})), \tag{3.15}$$

in which the binary diagonal matrix $\mathbf{D}_{\text{sel}}$ is defined according to (3.7). Condition (3.8) now implies that the inequality condition $\omega(\mathbf{C}^T\mathbf{z}) = \omega(\mathbf{d}) \leq \omega(\widehat{\mathbf{C}}^T\mathbf{z}) = \omega(\widehat{\mathbf{d}}) = 1$ in Kruskal's permutation lemma is satisfied. We conclude that condition (3.8) is sufficient for the uniqueness of $\mathbf{C}$.

From relations (3.13)–(3.14) it is also clear that there cannot exist a vector $\mathbf{d} \in \mathbb{C}^R$ with properties $\omega(\mathbf{d}) \geq 2$ and $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \mathbf{0}$, since it would imply that $\mathbf{G}\mathbf{d}^{(2)} = \mathbf{0}$ and thereby contradict condition (3.8). This also explains that if condition (3.8) is satisfied, then $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ has full column rank.

Necessity of condition (3.8)**.** Assume now on the contrary that condition (3.8) is not satisfied, i.e., there exists a vector $\mathbf{d} \in \mathbb{C}^R$ with properties $\omega(\mathbf{d}) \geq 2$ and $\mathbf{G}\mathbf{d}^{(2)} = 0$. The identity $\mathbf{G}\mathbf{d}^{(2)} = 0$ means that for every quadruple (i_1, i_2, j_1, j_2), the 2-by-2 determinant $\left|\sum_{r=1}^R \mathbf{a}_r^{(i_1,i_2)} \mathbf{b}_r^{(j_1,j_2)T} d_r\right|$ vanish (cf. equations (3.13)–(3.14)) and consequently there exist vectors $\widehat{\mathbf{a}}_r \in \mathbb{C}^I$ and $\widehat{\mathbf{b}}_r \in \mathbb{C}^J$ so that

$$\mathbf{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \sum_{r=1}^R D_{\mathbf{W}_1}(\mathbf{a}_r \otimes \mathbf{b}_r) d_r = D_{\mathbf{W}_1}(\widehat{\mathbf{a}}_r \otimes \widehat{\mathbf{b}}_r).$$

In other words, columns of $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ can be written as non-trivial linear combinations of the remaining columns, that is $D_{\mathbf{W}_1}(\mathbf{a}_r \otimes \mathbf{b}_r) = \sum_{s \neq r} D_{\mathbf{W}_1}(\mathbf{a}_s \otimes \mathbf{b}_s)\alpha_s$ for some scalars $\alpha_s \in \mathbb{C}$ with property $\prod_{s \neq r} \alpha_s \neq 0$. This implies that $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ does not have full column rank and that there exists a vector $\mathbf{y} \in \ker(\mathbf{\Phi}(\mathbf{A}, \mathbf{B}))$ and a vector $\mathbf{x} \in \mathbb{C}^K$ such that $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})(\mathbf{C} + \mathbf{x}\mathbf{y}^T)^T$ has rank $R - 1$, which contradicts condition (3.8). We conclude that condition (3.8) is also necessary for the uniqueness of $\mathbf{C}$. ∎

A problem with Proposition 3.2 is that it may be difficult to check condition (3.8) in practice. However, if $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ has full column rank, then obviously $\omega(\mathbf{d}) = 0$ and consequently $\omega(\mathbf{d}) \leq 1$, implying that condition (3.8) is automatically satisfied. Based on this fact, an easy-to-check sufficient condition is stated in Proposition 3.3. The proof of Proposition 3.3 will be explained in terms of a coupled CPD of the submatrices $\mathbf{X}^{(n)}$ extracted from $\mathbf{X}$, each admitting the factorization $\mathbf{X}^{(n)} = (\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)})\mathbf{C}^T$. This interpretation will naturally lead to the notion of multi-grid sampling for low-rank tensor decompositions.

PROPOSITION 3.3. *Consider a tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, partially observed as $\mathcal{Y} = \mathcal{W} * \mathcal{X}$, and its PD given by* (2.1)*. Let $\boldsymbol{D}_{sel} = D_1(\boldsymbol{d}_{sel}^T)$ in which $\boldsymbol{d}_{sel}$ is defined as in* (3.7)*. If*

$$\begin{cases} \boldsymbol{C} \text{ has full column rank,} \\ \boldsymbol{D}_{sel}(C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B})) \text{ has full column rank,} \end{cases} \tag{3.16}$$

then the rank of $\mathcal{X}$ is R and the factor matrix $\boldsymbol{C}$ is unique. Generically, condition (3.16) *is satisfied if $C_R^2 \leq \omega(\boldsymbol{d}_{sel})$ and $R \leq K$.*

Proof. Consider all $C_I^2 C_J^2$ possible $2 \times 2 \times K$ subtensors of $\mathcal{Y}$ of the form $\mathcal{Y}_{(u,v),(p,q)} = \mathcal{Y}([u,v],[p,q],:)$ with matrix representations

$$\mathbf{Y}_{(u,v),(p,q)} = \mathbf{D}_{(u,v),(p,q)}(\mathbf{A}^{(u,v)} \odot \mathbf{B}^{(p,q)})\mathbf{C}^T,$$

where $\mathbf{A}^{(u,v)} = \mathbf{A}([u,v],:) \in \mathbb{C}^{2\times R}$, $\mathbf{B}^{(p,q)} = \mathbf{B}([p,q],:) \in \mathbb{C}^{2\times R}$ and $\mathbf{D}_{(u,v),(p,q)}$ is the associated submatrix of $\mathbf{D}_{\mathbf{W}_1}$. Construct tensors $\mathcal{Z}_{(u,v),(p,q)} \in \mathbb{C}^{2\times 2\times K}$ with matrix representation

$$\mathbf{Z}_{(u,v),(p,q)} = \mathbf{S}_{(u,v),(p,q)}\mathbf{Y}_{(u,v),(p,q)} = \mathbf{S}_{(u,v),(p,q)}(\mathbf{A}^{(u,v)} \odot \mathbf{B}^{(p,q)})\mathbf{C}^T, \tag{3.17}$$

where $\mathbf{S}_{(u,v),(p,q)} \in \mathbb{R}^{4\times 4}$ is given by

$$\mathbf{S}_{(u,v),(p,q)} = \begin{cases} \mathbf{I}_4, \text{ if fibers } \mathbf{x}_{up\bullet}, \mathbf{x}_{uq\bullet}, \mathbf{x}_{vp\bullet} \text{ and } \mathbf{x}_{vq\bullet} \text{ are sampled,} \\ \mathbf{0}, \text{ otherwise.} \end{cases}$$

The coupled CPD of $\{\mathcal{Z}_{(u,v),(p,q)}\}$ has the following matrix representation

$$\begin{bmatrix} \mathbf{Z}_{(1,2),(1,2)} \\ \mathbf{Z}_{(1,2),(1,3)} \\ \vdots \\ \mathbf{Z}_{(I-1,I),(J-1,J)} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{(1,2),(1,2)}(\mathbf{A}^{(1,2)} \odot \mathbf{B}^{(1,2)}) \\ \mathbf{S}_{(1,2),(1,3)}(\mathbf{A}^{(1,2)} \odot \mathbf{B}^{(1,3)}) \\ \vdots \\ \mathbf{S}_{(I-1,I),(J-1,J)}(\mathbf{A}^{(I-1,I)} \odot \mathbf{B}^{(J-1,J)}) \end{bmatrix} \mathbf{C}^T. \tag{3.18}$$

The matrix (2.8) associated with $\{\mathcal{Z}_{(u,v),(p,q)}\}$ takes the form (3.15). Due to Theorem 2.3, we can conclude from (3.15) that if condition (3.16) is satisfied, then the rank of $\mathcal{X}$ is R and the factor matrix $\mathbf{C}$ is unique.

We now explain that condition (3.16) is generically satisfied if $C_R^2 \leq \omega(\mathbf{d}_{\text{sel}})$ and $R \leq K$. The latter inequality is obvious. The former inequality follows from the fact that $C_2(\mathbf{A}) \odot C_2(\mathbf{B})$ generically has full column rank if $C_I^2 C_J^2 \geq C_R^2$ [34] (i.e., the row-dimension of $C_2(\mathbf{A}) \odot C_2(\mathbf{B})$ exceeds its column-dimension). The result now immediately follows if $C_R^2 \leq \omega(\mathbf{d}_{\text{sel}})$, implying that the number of non-zeroed rows of $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ still exceeds the column-dimension. □

3.1.3. Sufficient condition for the recovery of overall CPD. Note that if the full column rank matrix $\mathbf{C}$ is unique, then $\boldsymbol{\Phi}(\mathbf{A},\mathbf{B}) = \mathbf{Y}_{\text{sub}}(\mathbf{C}^T)^\dagger$ is also unique. Thus, the remaining problem of determining whether $\mathbf{A}$ and $\mathbf{B}$ are unique reduces to checking if the partial matrix $\mathbf{P}^{(r)}$ given by (3.5) admits a unique rank-1 completion for every $r \in \{1,\ldots,R\}$.

Intuitively, an incomplete matrix $\mathbf{Y} = \mathbf{W} * \mathbf{X}$ has a unique rank-1 completion if $\mathbf{X}$ has rank 1 and the indicator pattern of $\mathbf{W}$ is "connected" in the sense that $\mathbf{Y}$ cannot be partitioned into "disconnected" rank-1 submatrices. As an example, consider the incomplete rank-1 matrix

$$\mathbf{Y} = \mathbf{W} * \mathbf{X} = \begin{bmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & * & * \\ x_{21} & x_{22} & \alpha & * \\ * & * & x_{33} & x_{34} \\ * & * & x_{43} & x_{44} \end{bmatrix},$$

where $\mathbf{Y}_{mn} \in \mathbb{C}^{2\times 2}$ and $*$ denotes an indeterminate entry. If $\alpha = *$, then the rank-1 completion of $\mathbf{Y}$ is not unique, as explained next. The matrix $\left[\begin{smallmatrix}\mathbf{a}\\\mathbf{c}\end{smallmatrix}\right]\left[\begin{smallmatrix}\mathbf{b}\\\mathbf{d}\end{smallmatrix}\right]^T$ yields a

rank-1 completion of $\mathbf{Y}$ for every tuple $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$ with properties $\mathbf{Y}_{11} = \mathbf{a}\mathbf{b}^T$ and $\mathbf{Y}_{22} = \mathbf{c}\mathbf{d}^T$. This means that the tuple $\{\beta\mathbf{a}, \beta^{-1}\mathbf{b}, \zeta\mathbf{c}, \zeta^{-1}\mathbf{d}\}$ with $\beta \neq \zeta$ yields the rank-1 completion $\left[\begin{smallmatrix}\beta\mathbf{a}\\ \zeta\mathbf{c}\end{smallmatrix}\right]\left[\begin{smallmatrix}\beta^{-1}\mathbf{b}\\ \zeta^{-1}\mathbf{d}\end{smallmatrix}\right]^T$ of $\mathbf{Y}$. Since $\left[\begin{smallmatrix}\mathbf{a}\\ \mathbf{c}\end{smallmatrix}\right]\left[\begin{smallmatrix}\mathbf{b}\\ \mathbf{d}\end{smallmatrix}\right]^T$ is not related to $\left[\begin{smallmatrix}\beta\mathbf{a}\\ \zeta\mathbf{c}\end{smallmatrix}\right]\left[\begin{smallmatrix}\beta^{-1}\mathbf{b}\\ \zeta^{-1}\mathbf{d}\end{smallmatrix}\right]^T$ up to scaling/counterscaling the rank-1 completion of $\mathbf{Y}$ is not unique. On the other hand, if $\alpha = x_{23}$, then there is an additional constraint on this tuple. More precisely, if $\mathbf{Y}_{22} = \mathbf{c}\mathbf{d}^T$, then we must have that $a_2 d_1 = y_{23} = x_{23}$. This connection makes that the rank-1 submatrices $\mathbf{Y}_{11}$ and $\mathbf{Y}_{22}$ are dependent on each other. It is now expected that the rank-1 completion of $\mathbf{Y}$ is unique. This assessment is confirmed by a result proposed in [16] and stated in the next paragraph.

A necessary and sufficient condition for the existence and the uniqueness of a rank-1 completion of an incomplete matrix was obtained in [16]. The basic idea in [16] is to regard an incomplete matrix, say $\mathbf{P}^{(r)}$ defined by (3.5), as the weights associated with the edges of a weighted bipartite graph denoted by $G^{(r)}$. The two groups of vertices associated with $G^{(r)}$ are the row indices $1, \ldots, I$ and the columns indices $1, \ldots, J$. The edges of $G^{(r)}$ correspond to the elements in $\mathcal{P}^{(r)}$, where the edge associated with $(i, j) \in \mathcal{P}^{(r)}$ is weighted by $p_{ij}^{(r)}$. In our setting we are only interested in the uniqueness result that was obtained in [16]. We summarize an adapted version of this result to our setting in Lemma 3.4 below. Lemma 3.4 considers the restriction of $\mathcal{P}^{(r)}$ of which the support is $\widetilde{\mathcal{P}}^{(r)} = \{(i, j) \in \mathcal{P}^{(r)} \mid p_{i,j}^{(r)} \neq 0\}$. The associated restricted bipartite graph with non-zero weights will be denoted by $\widetilde{G}^{(r)}$.

LEMMA 3.4. *An incomplete rank-1 matrix $\boldsymbol{P}^{(r)}$ defined by* (3.5)*, with support $\mathcal{P}^{(r)}$ and satisfying property* (3.6)*, admits a unique rank-1 completion if and only if*

$$\widetilde{G}^{(r)} \text{ is a connected graph.} \tag{3.19}$$

Since $\mathbf{C}$ has full column rank, Propositions 3.2–3.7 ensure the uniqueness of $\boldsymbol{\Phi}(\mathbf{A}, \mathbf{B})$, which in turn ensures the existence of a rank-1 completion of $\mathbf{P}^{(r)}$. Hence, if condition (3.19) is satisfied we also know that the factor matrices $\mathbf{A}$ and $\mathbf{B}$ are unique. Since a bipartite graph is a simple graph, condition (3.19) can easily be checked via the properties of the incidence or adjacency matrix of the restricted bipartite graph $\widetilde{G}^{(r)}$ (e.g., [2]).

The combination of Proposition 3.2 and Lemma 3.4 yield the following overall uniqueness condition, which is an extension of Theorem 2.1 to the missing fibers case.

THEOREM 3.5. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in* (2.1) *observed via $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$ in* (3.1)*. Let $\boldsymbol{D}_{sel} = D_1(\boldsymbol{d}_{sel}^T)$ in which $\boldsymbol{d}_{sel}$ is defined as in* (3.7) *and let $\widetilde{G}^{(r)}$ be the restricted bipartite graph of $\boldsymbol{P}^{(r)}$ defined by* (3.5) *and with property* (3.6)*. Let $\boldsymbol{d}^{(2)} \in \mathbb{C}^{C_R^2}$ be the vector given by* (2.3)*. Assume that $\boldsymbol{C}$ has full column rank. The rank of $\mathcal{X}$ is R, the CPD of $\mathcal{X}$ is unique, and the rank-1 terms $\{\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r\}_{r=1}^R$ can be recovered from $\mathcal{Y}$ if and only if*

$$\begin{cases} D_{sel}(C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B}))\boldsymbol{d}^{(2)} = \boldsymbol{0} \Rightarrow \omega(\boldsymbol{d}) \leq 1, & \text{(3.20a)} \\ \widetilde{G}^{(r)} \text{ is a connected graph for every } r \in \{1, \ldots, R\}. & \text{(3.20b)} \end{cases}$$

Even though Theorem 3.5 provides a necessary and sufficient uniqueness condition for the cases where $\mathbf{C}$ has full column rank, condition (3.20a) may be hard to check in practice. In contrast, the combination of Proposition 3.3 and Lemma 3.4 lead to the following easy-to-check sufficient overall uniqueness condition.

THEOREM 3.6. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in* (2.1) *observed via $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$ in* (3.1)*. Let $\boldsymbol{D}_{sel} = D_1(\boldsymbol{d}_{sel}^T)$ in which $\boldsymbol{d}_{sel}$ is defined as in* (3.7) *and let $\widetilde{G}^{(r)}$*

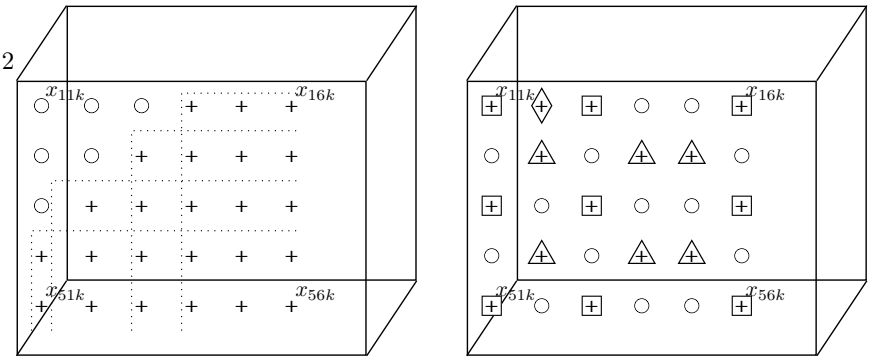


FIG. 3.1. *Multi-grid fiber sampling of a tensor with systematically missing fibers (left) and subsampling of a tensor using multi-grid fiber sampling (right), where '+' represents an observed fiber and '◯' represents an unobserved fiber. The two multi-grids in the subsampled tensor are denoted by '□' and '△'. The '◇' fiber connects the two grids and ensures that the overall CPD is unique.*

be the restricted bipartite graph of $\boldsymbol{P}^{(r)}$ *defined by* (3.5) *and with property* (3.6)*. If*

$$\begin{cases} \boldsymbol{C} \text{ in } (2.1) \text{ has full column rank,} \\ \boldsymbol{D}_{sel}(C_2(\boldsymbol{A}) \odot C_2(\boldsymbol{B})) \text{ has full column rank,} \\ \widetilde{G}^{(r)} \text{ is a connected graph for every } r \in \{1, \ldots, R\}, \end{cases} \tag{3.21}$$

then the rank of $\mathcal{X}$ *is* R*, the CPD of* $\mathcal{X}$ *is unique, and the rank-1 terms* $\{\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r\}_{r=1}^{R}$ *can be recovered from* $\mathcal{Y}$*.*

Furthermore, if condition (3.21) is satisfied, then the rank-1 terms can be recovered via an EVD, as will be explained in Section 4. Theorem 3.6 can be interpreted as a generalization of Theorem 2.2 to the case of missing fibers.

3.2. Multi-grid sampling. In large scale tensor decomposition problems it can be expensive to process all observed fibers. A simple solution is the multi-grid sampling scheme described next that instead of considering all possible $2 \times 2 \times K$ subtensors of $\mathcal{X}$, only considers N grid-structured subtensors of $\mathcal{X}$, denoted by $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(N)}$. More precisely, each subtensor is defined by

$$\mathcal{X}^{(n)} = \mathcal{X}([\mu(1), \ldots, \mu(I_n)], [\gamma(1), \ldots, \gamma(J_n)], :) \in \mathbb{C}^{I_n \times J_n \times K}, \tag{3.22}$$

in which I_n and J_n denote the row and column dimensions of $\mathcal{X}^{(n)}$, respectively. The chosen row entries of $\mathcal{X}$ are indexed by $\mu(1), \ldots, \mu(I_n)$ and the chosen column entries of $\mathcal{X}$ are indexed by $\gamma(1), \ldots, \gamma(J_n)$. Due to the grid structure, the matrix unfolding of $\mathcal{X}^{(n)}$ can be written as

$$\mathbf{X}^{(n)} = \mathbf{Y}_{\text{sub}}([\sigma(1), \ldots, \sigma(I_n J_n)], :) = (\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)})\mathbf{C}^T, \tag{3.23}$$

where $\sigma(1), \ldots, \sigma(I_n J_n)$ denote the indices of the associated observed rows of $\mathbf{X}$, and

$$\mathbf{A}^{(n)} = \mathbf{A}([\mu(1), \ldots, \mu(I_n)], :) \in \mathbb{C}^{I_n \times R}, \tag{3.24}$$

$$\mathbf{B}^{(n)} = \mathbf{B}([\gamma(1), \ldots, \gamma(J_n)], :) \in \mathbb{C}^{J_n \times R}. \tag{3.25}$$

Overall, we obtain a set of tensors $\{\mathcal{X}^{(n)}\}_{n=1}^{N}$ of which the matrix representation $\mathbf{X}_{\text{tot}} \in \mathbb{C}^{(\sum_{n=1}^{N} I_n J_n) \times K}$ admits the coupled decomposition

$$\mathbf{X}_{\text{tot}} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(1)} \odot \mathbf{B}^{(1)} \\ \vdots \\ \mathbf{A}^{(N)} \odot \mathbf{B}^{(N)} \end{bmatrix} \mathbf{C}^T. \tag{3.26}$$

In multi-grid sampling the relations between the factors in $\{\mathbf{A}^{(n)}\}$ and in $\{\mathbf{B}^{(n)}\}$ are ignored, such that from (3.1) we obtain the coupled PD

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r \,, \quad n \in \{1, \ldots, N\}, \tag{3.27}$$

where $\mathbf{a}_r^{(n)} \in \mathbb{C}^{I_n}$ and $\mathbf{b}_r^{(n)} \in \mathbb{C}^{J_n}$. We have now turned a large-scale CPD into a smaller-scaled coupled CPD. Proposition 3.7 below provides a uniqueness condition based on the proposed multi-grid sampling approach.

PROPOSITION 3.7. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in* (2.1) *observed via $\mathcal{Y} \in \mathbb{C}^{I \times J \times K}$ in* (3.1)*. Let $\boldsymbol{G}^{(N)}$ in* (2.8) *be built from the factor matrices $\{\boldsymbol{A}^{(n)}, \boldsymbol{B}^{(n)}\}$ in* (3.26) *and let $\widetilde{G}^{(r)}$ be the restricted bipartite graph of $\boldsymbol{P}^{(r)}$ defined by* (3.5) *and with property* (3.6)*. If*

$$\begin{cases} \boldsymbol{C} \text{ in (2.1) has full column rank,} \\ \boldsymbol{G}^{(N)} \text{ in (2.8) has full column rank,} \\ \widetilde{G}^{(r)} \text{ is a connected graph for every } r \in \{1, \ldots, R\}, \end{cases} \tag{3.28}$$

then the rank of $\mathcal{X}$ is R, the CPD of $\mathcal{X}$ is unique, and the rank-1 terms $\{\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r\}_{r=1}^{R}$ can be recovered from $\mathcal{Y}$.

We mention again that if condition (3.28) is satisfied, then the rank-1 terms can be recovered via an EVD, as will be explained in Section 4. Proposition 3.7 is more restrictive than Theorem 3.6, i.e., condition (3.28) implies condition (3.21). Indeed, the matrix obtained by removing any redundant rows of $\mathbf{G}^{(N)}$ corresponds to a submatrix of $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$. On the other hand, condition (3.28) supports the development of multi-grid sampling methods that allow us to consider low-rank tensors of considerable size. Briefly, assume that F observed fibers of $\mathcal{X}$ in (3.1) are taken into account. These fibers can be stored in the matrix $\mathbf{H} \in \mathbb{C}^{F \times K}$ with factorization

$$\mathbf{H} = \begin{bmatrix} \mathbf{x}_{i_1 j_1 \bullet} \\ \vdots \\ \mathbf{x}_{i_F j_F \bullet} \end{bmatrix} = \boldsymbol{\phi}_{\mathbf{H}}(\mathbf{A}, \mathbf{B}) \mathbf{C}^T, \quad \boldsymbol{\phi}_{\mathbf{H}}(\mathbf{A}, \mathbf{B}) = \begin{bmatrix} \mathbf{a}_{i_1 \bullet} * \mathbf{b}_{j_1 \bullet} \\ \vdots \\ \mathbf{a}_{i_F \bullet} * \mathbf{b}_{j_F \bullet} \end{bmatrix}. \tag{3.29}$$

Note that this fiber sampling approach only requires the storage of $(K+2)F$ values, namely the KF entries of $\mathbf{H}$ and the $2F$ coordinates associated with the sampled fibers. If $(K+2)F << IJK$, then this is a significant reduction in terms of storage. Observe also that the computation of an R-term CPD amounts to solving a system of IJK (rank-1 structured) equations in $R(I+J+K-1)$ variables. In the case of fiber sampling, only FK of the involved equations are considered. In practice, we typically have that $R(I+J+K-1) << IJK$ so that a significant reduction of the complexity is possible with a limited loss of accuracy.

3.3. Illustrations.

Systematically missing fibers. Consider the tensor $\mathcal{X} \in \mathbb{C}^{5 \times 6 \times K}$ with frontal slices of the form in Figure 3.1 (left) in which the fibers $\mathbf{x}_{11\bullet}, \mathbf{x}_{21\bullet}, \mathbf{x}_{31\bullet}, \mathbf{x}_{21\bullet}, \mathbf{x}_{22\bullet}$ and $\mathbf{x}_{13\bullet}$ are missing. Since the tensor is incomplete, CPD-specific uniqueness conditions and algorithms do not apply (e.g., Theorem 2.2). Theorem 3.6 generically guarantees the uniqueness of the CPD of $\mathcal{X}$ despite the missing fibers if $R \leq 11$ and $R \leq K$. Let

us now consider multi-grid sampling. As an example, consider the coupled CPD of $\{\mathcal{X}^{(n)}\}_{n=1}^{4}$, where

$$\mathcal{X}^{(1)} = \mathcal{X}(:, [4\ 5\ 6], :) \in \mathbb{C}^{5\times 3\times K}, \qquad \mathcal{X}^{(3)} = \mathcal{X}([3\ 4\ 5], [2\ 3\ 4\ 5\ 6], :) \in \mathbb{C}^{3\times 5\times K},$$
$$\mathcal{X}^{(2)} = \mathcal{X}([2\ 3\ 4\ 5], [3\ 4\ 5\ 6], :) \in \mathbb{C}^{4\times 4\times K}, \mathcal{X}^{(4)} = \mathcal{X}([4\ 5], :, :) \in \mathbb{C}^{2\times 6\times K},$$

in which the associated frontal slices are bordered by dotted lines in Figure 3.1 (left). Proposition 3.7 now generically guarantees the uniqueness of the CPD of $\mathcal{X}$ despite the missing fibers if $R \leq 11$ and $R \leq K$. (The reason that Theorem 3.6 and Proposition 3.7 yield the same upper bound on R is that the above coupled CPD of $\{\mathcal{X}^{(n)}\}_{n=1}^{4}$ takes all the observed $(2 \times 2 \times K)$ subtensors of $\mathcal{X}$ into account.)

Subsampling of tensors. Multi-grid fiber sampling can also be used for sparse sampling of tensors. As an example, consider the tensor $\mathcal{X} \in \mathbb{C}^{5\times 6\times K}$ with frontal slices of the form depicted in Figure 3.1 (right) in which two sampling grids are used:

$$\mathcal{X}^{(1)} := \mathcal{X}([1\ 3\ 5], [1\ 3\ 6], :) \in \mathbb{C}^{3\times 3\times K},$$
$$\mathcal{X}^{(2)} := \mathcal{X}([2\ 4], [2\ 4\ 5], :) \in \mathbb{C}^{2\times 3\times K}.$$

The fibers associated with the tensors $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$ are surrounded by squared and triangular frames, respectively. In order to ensure that the CPD of $\mathcal{X}$ can be recovered from the incomplete tensor $\mathcal{Y}$, additional fibers of $\mathcal{X}$ may have to be considered. In this example, the fiber $\mathbf{x}_{12\bullet}$ surrounded by a diamond shaped frame in Figure 3.1 (right) ensures that the sampling grids $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$ are "connected". More precisely, the exclusion of $\mathbf{x}_{12\bullet}$ from the set of sampled fibers would lead to a non-unique CPD of $\mathcal{X}$. The reason is that the bipartite graph $\widetilde{G}^{(r)}$ associated with

$$\widetilde{\mathcal{P}}^{(r)} = \{(m,1),(m,3),(m,6),(n,3),(n,4),(n,5)\}_{m\in\{1,3,5\}}^{n\in\{2,4,6\}} \text{ with property } (1,2) \notin \widetilde{\mathcal{P}}^{(r)}$$

is not connected. Thanks to the extra fiber $\mathbf{x}_{12\bullet}$, Proposition 3.7 now generically guarantees the uniqueness of the CPD of $\mathcal{X}$ via the coupled CPD of $\{\mathcal{X}^{(1)}, \mathcal{X}^{(2)}\}$ if $R \leq 5$ and $R \leq K$, despite the fact that only 16 out of 30 fibers have been sampled.

4. An algorithm for fiber sampled rank decompositions. An algorithm that transforms the CPD problem into a Simultaneous matrix Diagonalization (SD) problem was proposed in [7] and further elaborated on in [11]. In [31] it was extended to coupled decompositions. In this section we explain how to adapt it to the case of missing fibers. The presented SD algorithm will rely on Theorem 3.6.

4.1. Step 1: Find basis for range $(\mathbf{\Phi}(\mathbf{A}, \mathbf{B}))$. Assume that the conditions in Theorem 3.6 are satisfied, which implies that $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ in (3.4) has full column rank. The first step of the SD method is to find a basis for range $(\mathbf{\Phi}(\mathbf{A}, \mathbf{B}))$. Let $\mathbf{Y}_{\text{sub}} = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{V}_{\text{sub}}^{H}$ denote the compact SVD of $\mathbf{Y}_{\text{sub}}$ given by (3.4) where $\mathbf{U}_{\text{sub}} \in \mathbb{C}^{F\times R}$ and $\mathbf{V}_{\text{sub}} \in \mathbb{C}^{K\times R}$ are columnwise orthonormal and $\mathbf{\Sigma}_{\text{sub}} \in \mathbb{C}^{R\times R}$ is diagonal. We know that range $(\mathbf{Y}_{\text{sub}})$ = range $(\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}})$ and that there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R\times R}$ such that

$$\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}} = \mathbf{\Phi}(\mathbf{A}, \mathbf{B}) \cdot \mathbf{F}^{T} \text{ and } \mathbf{V}_{\text{sub}}^{*}\mathbf{F}^{-1} = \mathbf{C}. \tag{4.1}$$

The left equation is of the same form as (3.4) but it only involves $R \leq K$ columns.

4.2. Step 2: From CPD with missing fibers to SD.

4.2.1. SD method for computation of standard CPD. The conversion step to SD will make use of the following identity valid for $\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{f}_r \in \mathbb{C}^{I \times J \times R}$ [11]:

$$(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))\mathbf{\Psi}_2(\mathbf{F})^T = \mathbf{R}_2(\mathcal{X}), \tag{4.2}$$

where $\mathbf{\Psi}_2(\mathbf{F}) \in \mathbb{C}^{R^2 \times C_R^2}$ and $\mathbf{R}_2(\mathcal{X}) \in \mathbb{C}^{C_I^2 C_J^2 \times R^2}$. The columns of $\mathbf{\Psi}_2(\mathbf{F})$ are lexicographically indexed by the pair (r_1, r_2) so that the column associated with the pair (r_1, r_2) is preceding the column associated with the pair (r_1', r_2') if and only if either $r_1' > r_1$ or $r_1' = r_1$ and $r_2' > r_2$. The column of $\mathbf{\Psi}_2(\mathbf{F})$ associated with the pair (r_1, r_2) is given by

$$\frac{1}{2}(\mathbf{f}_{r_1} \otimes \mathbf{f}_{r_2} + \mathbf{f}_{r_2} \otimes \mathbf{f}_{r_1}), \quad 1 \le r_1 < r_2 \le R.$$

Likewise, the columns of $\mathbf{R}_2(\mathcal{X}) \in \mathbb{C}^{C_I^2 C_J^2 \times R^2}$ are lexicographically indexed by the pair (r_1, r_2), and given by

$$\mathrm{Vec}\left(C_2(\mathcal{X}(:,:,r_1) + \mathcal{X}(:,:,r_2)) - C_2(\mathcal{X}(:,:,r_1)) - C_2(\mathcal{X}(:,:,r_2))\right).$$

Note that $\mathbf{R}_2(\mathcal{X})$ has only C_{R+1}^2 distinct columns and that its rows correspond to vectorized symmetric matrices. We refer to [11, 31] for further explanations.

Let $\pi_S : \mathbb{C}^{R^2} \to \mathbb{C}^{R^2}$ be a symmetrization mapping:

$$\pi_S(\mathrm{Vec}\,(\mathbf{F})) = \mathrm{Vec}\left((\mathbf{F} + \mathbf{F}^T)/2\right), \quad \mathbf{F} \in \mathbb{C}^{R \times R}.$$

Denote

$$W = \ker(\mathbf{\Psi}_2(\mathbf{F})^T) \cap \mathrm{range}(\pi_S). \tag{4.3}$$

Then it was also shown in [11] that if $\mathbf{F}$ has full column rank, then $\dim \mathrm{range}(\mathbf{\Psi}_2(\mathbf{F})^*) = R(R-1)/2$, $\dim W = R$, and $[\mathbf{x}_1, \ldots, \mathbf{x}_R]$ coincides with $\mathbf{F}^{-T}$ up to a permutation and column scaling if and only if

$$\mathbf{x}_1 \otimes \mathbf{x}_1, \ldots, \mathbf{x}_R \otimes \mathbf{x}_R \quad \text{form a basis of } W. \tag{4.4}$$

4.2.2. SD method in the case of missing fibers. Observe that each row of $\mathbf{R}_2(\mathcal{X})$ depends only on two rows of $\mathbf{A}$ and $\mathbf{B}$. This is the key property that allows us to extend the SD method to the case of missing fibers. More precisely, define $\beta_{r_1,r_2} = (r_2 - 1)R + r_1$ and $\alpha_{i_1,i_2,j_1,j_2} = \frac{(j_1(2j_2 - j_1 - 1) - 2)I_n(I_n - 1)}{4} + \frac{i_1(2i_2 - i_1 - 1)}{2}$, then the $(\alpha_{i_1,i_2,j_1,j_2}, \beta_{r_1,r_2})$-th entry of $\mathbf{R}_2(\mathcal{X})$ is equal to

$$x_{i_1 j_1 r_1} x_{i_2 j_2 r_2} + x_{i_1 j_1 r_2} x_{i_2 j_2 r_1} - x_{i_1 j_2 r_1} x_{i_2 j_1 r_2} - x_{i_1 j_2 r_2} x_{i_2 j_1 r_1}$$

for all $1 \le i_1 < i_2 \le I$, $1 \le j_1 < j_2 \le J$ and $1 \le r_1, r_2 \le R$.

Let $\widetilde{\widetilde{\mathbf{U}}}_{\mathrm{sub}} = \mathbf{S}_{\mathbf{W}_1}^T \mathbf{U}_{\mathrm{sub}} \mathbf{\Sigma}_{\mathrm{sub}} \in \mathbb{C}^{IJ \times R}$ denote the matrix obtained by inserting zeros in the unobserved rows of $\mathbf{X}$. By a proper reshaping of $\widetilde{\mathbf{U}}_{\mathrm{sub}}$ we obtain the compressed variant of $\mathcal{Y}$ in (3.1):

$$\widetilde{\mathcal{U}}_{\mathrm{sub}} = \mathcal{W} * \left(\sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{f}_r\right) \in \mathbb{C}^{I \times J \times R}. \tag{4.5}$$

Denote $\mathbf{R}_2 = \mathbf{D}_{\text{sel}}\mathbf{R}_2(\widetilde{\mathcal{U}}_{\text{sub}})$, in which the diagonal matrix $\mathbf{D}_{\text{sel}}$ will zero out the rows of $\mathbf{R}_2(\widetilde{\mathcal{U}}_{\text{sub}})$ that correspond to unobserved rows of $\mathbf{R}_2(\mathcal{U}_{\text{sub}})$. The combination of (4.2) and (4.5) now yields

$$\mathbf{R}_{\text{sub}} := \mathbf{D}_{\text{sel}}\mathbf{R}_2(\widetilde{\mathcal{U}}_{\text{sub}}) = \mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})) \cdot \mathbf{\Psi}_2(\mathbf{F})^T. \tag{4.6}$$

Let $\mathbf{S}_{\text{red}} \in \mathbb{C}^{\omega(\mathbf{d}_{\text{sel}}) \times C_I^2 C_J^2}$ denote the row-selection matrix that selects the rows of $\mathbf{R}_2$ that have not been zeroed by $\mathbf{D}_{\text{sel}}$. Then we obtain the reduced matrix $\mathbf{R}_{\text{red}} = \mathbf{S}_{\text{red}}\mathbf{R}_2 \in \mathbb{C}^{\omega(\mathbf{d}_{\text{sel}}) \times R^2}$ in which the rows of $\mathbf{R}_2$ that have been zeroed by $\mathbf{D}_{\text{sel}}$ are dropped. This leads to reduced version of (4.6):

$$\mathbf{R}_{\text{red}} := \mathbf{S}_{\text{red}}\mathbf{R}_2(\widetilde{\mathcal{U}}_{\text{sub}}) = \mathbf{S}_{\text{red}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})) \cdot \mathbf{\Psi}_2(\mathbf{F})^T. \tag{4.7}$$

Since the matrix $\mathbf{G}_{\text{sel}}^{(N)}$ has full column rank by assumption, it follows that

$$W = \ker(\mathbf{\Psi}_2(\mathbf{F})^T) \cap \text{range}(\pi_S) = \ker(\mathbf{R}_{\text{red}}) \cap \text{range}(\pi_S).$$

Since the rows of the matrix $\mathbf{R}_{\text{red}}$ in (4.7) are vectorized symmetric matrices, we have that range $\left(\mathbf{R}_{\text{red}}^T\right) \subseteq \text{range}(\pi_S)$. Consequently, a set of R basis vectors $\mathbf{m}_1, \ldots, \mathbf{m}_R \in \mathbb{C}^{R^2}$ for W can be obtained from a submatrix of $\mathbf{R}_{\text{red}}$ with C_{R+1}^2 distinct columns. In more detail, define the set $S_R = \{(r_1, r_2) \ : \ 1 \leq r_1 \leq r_2 \leq R\}$ in which the C_{R+1}^2 elements are ordered lexicographically and consider the mapping $f : \mathbb{N}^2 \to \{1, 2, \ldots, C_{R+1}^2\}$ that returns the position of its argument in the set S_R. Let $\mathbf{Q}_{\text{red}} \in \mathbb{C}^{\omega(\mathbf{D}_{\text{sel}}) \times C_{R+1}^2}$ denote a matrix consisting of the C_{R+1}^2 distinct columns of $\mathbf{R}_{\text{red}}$ and constructed as follows:

$$\mathbf{Q}_{\text{red}} = [\mathbf{Q}_{(1)} \ , \ 2 \cdot \mathbf{Q}_{(2)}], \tag{4.8}$$

where

$$\mathbf{Q}_{(1)} = [\mathbf{R}_{\text{red}}(:, f(1,1)), \ldots, \mathbf{R}_{\text{red}}(:, f(R,R))],$$
$$\mathbf{Q}_{(2)} = [\mathbf{R}_{\text{red}}(:, f(1,2)), \mathbf{R}_{\text{red}}(:, f(1,3)), \ldots, \mathbf{R}_{\text{red}}(:, f(R-1,R))].$$

A basis $\{\mathbf{m}_r\}$ for W can now be built from the R right singular vectors associated with the R smallest singular values of $\mathbf{Q}_{\text{red}}$. Alternatively, since $\ker(\mathbf{Q}_{\text{red}}) = \ker(\mathbf{Q}_{\text{red}}^H \cdot \mathbf{Q}_{\text{red}})$ the basis can also be found from the R right singular vectors associated with the R smallest singular values of the matrix $\mathbf{Q}_{\text{red}}^H \cdot \mathbf{Q}_{\text{red}}$.

Let the columns of $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ constitute a basis for W, then due to relation (4.4) we can conclude that there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R \times R}$ such that

$$(\mathbf{F}^{-T} \odot \mathbf{F}^{-T})\mathbf{N}^T = \mathbf{M}. \tag{4.9}$$

Clearly, (4.9) corresponds to a matrix representation of the CPD $\mathcal{M} = \sum_{r=1}^R \mathbf{g}_r \circ \mathbf{g}_r \circ \mathbf{n}_r \in \mathbb{C}^{R \times R \times R}$ in which $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_R] = \mathbf{F}^{-T}$. The latter CPD can be interpreted as a SD problem[3], which in turn can be solved via an EVD (e.g., [18, 23]). Indeed, let $\mathbf{M}^{(\cdot\cdot k)} \in \mathbb{C}^{R \times R}$ denote a matrix slice of $\mathcal{M}$ such that $(\mathbf{M}^{(\cdot\cdot k)})_{ij} = (\mathcal{M})_{ijk}$, then $\mathbf{M}^{(\cdot\cdot k)} = \mathbf{G}D_k(\mathbf{N})\mathbf{G}^T$ and we obtain the standard (generalized) EVD relation $\mathbf{M}^{(\cdot\cdot k_1)} \cdot \mathbf{F} = \mathbf{M}^{(\cdot\cdot k_2)} \cdot \mathbf{F} \cdot (D_{k_1}(\mathbf{N})D_{k_2}(\mathbf{N})^{-1})$, where $1 \leq k_1 \neq k_2 \leq R$.

[3]The rank-R tensor $\mathcal{M} = \sum_{r=1}^R \mathbf{g}_r \circ \mathbf{g}_r \circ \mathbf{n}_r \in \mathbb{C}^{R \times R \times R}$ can be interpreted as a collection of R matrix slices $\mathbf{M}^{(\cdot\cdot 1)}, \ldots, \mathbf{M}^{(\cdot\cdot R)} \in \mathbb{C}^{R \times R}$, each with properties $(\mathbf{M}^{(\cdot\cdot k)})_{ij} = \mathcal{M}_{ijk}$ and $\mathbf{M}^{(\cdot\cdot k)} = \mathbf{G}D_k(\mathbf{N})\mathbf{G}^T$. It is now clear that the problem of finding the CPD of $\mathcal{M}$ can be understood as a SD problem.

Variable	Description
$\mathbf{p}_j^{(r)}$	jth column of the $(I \times J)$ incomplete matrix $\mathbf{P}^{(r)} = [\mathbf{p}_1^{(r)}, \ldots, \mathbf{p}_J^{(r)}]$.
$\Theta^{(r)}$	The column index set of the non-zero columns in $\mathbf{P}^{(r)}$, i.e., $\Theta^{(r)} = \{j \in \{1, \ldots, J\} \mid \mathbf{p}_j^{(r)} \neq \mathbf{0}\}$.
$I_{j,r}$	Number of determinate entries of column vector $\mathbf{p}_j^{(r)}$.
$\mu(i)$	Index of the ith determinate entry of $\mathbf{p}_j^{(r)}$, i.e., $p_{\mu(i),j}^{(r)} = a_{\mu(i),r} b_{j,r}$, $\forall i \in \{1, \ldots, I_{j,r}\}$.
$I_{j,r}^c$	Number of indeterminate entries of $\mathbf{p}_j^{(r)}$.
$\upsilon(i)$	Index of the ith indeterminate entry of $\mathbf{p}_j^{(r)}$, i.e., entry $p_{\upsilon(i),j}^{(r)}$ is unknown for every $i \in \{1, \ldots, I_{j,r}^c\}$.
$\mathbf{z}_j^{(r)}$	Vector in which the indeterminate entries of $\mathbf{p}_j^{(r)}$ have been replaced by zeros, i.e., $z_{ij}^{(r)} = p_{ij}^{(r)}$ if $i \in \{\mu(1), \ldots, \mu(I_{j,r})\}$ and $z_{ij}^{(r)} = 0$ otherwise.

TABLE 4.1
Overview of the variables used in Section 4.3.

4.3. Step 3: From SD to rank-1 subspace identification. After $\mathbf{F}^{-1}$ has been found we obtain $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ and $\mathbf{C}$ via (4.1). If all fibers of $\mathcal{X}$ have been sampled, then $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{A} \odot \mathbf{B}$ and the pair $\{\mathbf{a}_r, \mathbf{b}_r\}$ can be obtained via the rank-1 matrix approximation $\min_{\mathbf{a}_r, \mathbf{b}} \|\mathbf{P}^{(r)} - \mathbf{a}_r \mathbf{b}_r^T\|_F^2$, where the entries of the matrix $\mathbf{P}^{(r)}$ given by (3.5) are all observed.

Assume now that some of the entries of $\mathbf{P}^{(r)} = [\mathbf{p}_1^{(r)}, \ldots, \mathbf{p}_J^{(r)}]$ are missing (i.e., $\exists\ (i,j) \notin \mathcal{P}^{(r)}$). Then $\{\mathbf{a}_r, \mathbf{b}_r\}$ cannot be found from a best rank-1 approximation of $\mathbf{P}^{(r)}$. We will now develop an efficient subspace method for the case of missing entries that again reduces to a best rank-1 approximation of a matrix. The derivation will make use of several variables which are listed in Table 4.1.

Let $\mathbf{X}^{(r)}$ denote a rank-1 completed version of $\mathbf{P}^{(r)}$. Observe that every non-zero column of $\mathbf{X}^{(r)} = [\mathbf{x}_1^{(r)}, \ldots, \mathbf{x}_J^{(r)}] = \mathbf{a}_r \mathbf{b}_r^T \in \mathbb{C}^{I \times J}$ generates a one-dimensional subspace with property $\mathbf{a}_r \in \text{range}\left(\mathbf{x}_j^{(r)}\right)$, $\forall j \in \Theta^{(r)} = \{j \in \{1, \ldots, J\} \mid \mathbf{p}_j^{(r)} \neq \mathbf{0}\}$. Let $I_{j,r}$ denote the number of known entries of $\mathbf{p}_j^{(r)}$ Similarly, let $I_{j,r}^c = I - I_{j,r}$ denote the number of unknown entries. Every column $\mathbf{p}_j^{(r)}$ of $\mathbf{P}^{(r)}$ generates a subspace of dimension $I_{j,r}^c + 1$ if we consider scaled versions of $\mathbf{p}_j^{(r)}$ and if moreover we let the indeterminate entries of $\mathbf{p}_j^{(r)}$ vary. More precisely, let $\mathbf{z}_j^{(r)} \in \mathbb{C}^I$ denote the column vector in which the indeterminate entries of $\mathbf{p}_j^{(r)}$ have been replaced by zeros, i.e.,

$$z_{ij}^{(r)} = \begin{cases} p_{ij}^{(r)}, & \text{if } p_{ij}^{(r)} \text{ is determinate,} \\ 0, & \text{if } p_{ij}^{(r)} \text{ is indeterminate.} \end{cases}$$

Further, let the indeterminate entries of $\mathbf{p}_j^{(r)}$ be indexed by $\upsilon(1), \ldots, \upsilon(I_{j,r}^c)$. Then the columns of $\mathbf{P}_j^{(r)} = [\mathbf{z}_j^{(r)}, \mathbf{e}_{\upsilon(1)}^{(I)}, \ldots, \mathbf{e}_{\upsilon(I_{j,r}^c)}^{(I)}]$ constitute a basis for range $\left(\mathbf{P}_j^{(r)}\right)$. Since $\mathbf{a}_r \in \text{range}\left(\mathbf{P}_j^{(r)}\right)$, $\forall j \in \Theta^{(r)}$, we have that

$$\mathbf{a}_r \in S^{(r)} := \bigcap_{j \in \Theta^{(r)}} \text{range}\left(\mathbf{P}_j^{(r)}\right).$$

We first explain that if the conditions in Theorem 3.6 are satisfied, then $\mathbf{a}_r$ can be recovered from the subspace $S^{(r)}$.

4.3.1. Dimension and range of subspace $S^{(r)} = \bigcap_{j\in\Theta^{(r)}} \text{range}(\mathbf{P}_j^{(r)})$. We know that under the conditions in Theorem 3.6 the rank-1 completion of $\mathbf{P}^{(r)}$ is unique. This implies that $S^{(r)} = \text{range}\,(\mathbf{a}_r)$. Indeed, if $S^{(r)} \neq \text{range}\,(\mathbf{a}_r)$, then there exist scalars $\alpha_j \in \mathbb{C}$, $j \in \Theta^{(r)}$, and a vector $\widetilde{\mathbf{a}}_r \in S^{(r)}$ with property $\widetilde{\mathbf{a}}_r \notin \text{range}\,(\mathbf{a}_r)$ such that $\mathbf{p}_j^{(r)} = \widetilde{\mathbf{w}}_j^{(r)} * (\widetilde{\mathbf{a}}_r \alpha_j)$, $\forall j \in \Theta^{(r)}$, where $\widetilde{w}_{ij}^{(r)} = 1$ if $z_{ij}^{(r)} \neq 0$ and zero elsewhere. This leads to an alternative rank-1 completion of $\mathbf{P}^{(r)}$, contradicting Theorem 3.6. To summarize, if the conditions in Theorem 3.6 are satisfied, then the subspace $S^{(r)}$ is one-dimensional and $\mathbf{a}_r$ can be recovered from it.

4.3.2. Recovery of $\mathbf{a}_r$ from $S^{(r)}$ via subspace method. Note that for every $j \in \Theta^{(r)}$, the columns of $\mathbf{Z}_j^{(r)} = [\mathbf{z}_j^{(r)}/\|\mathbf{z}_j^{(r)}\|_F, \mathbf{e}_{\upsilon(1)}^{(I)}, \ldots, \mathbf{e}_{\upsilon(I_{j,r}^c)}^{(I)}]$ constitute an orthonormal basis for range $\left(\mathbf{P}_j^{(r)}\right)$. Define $\mathbf{Z}_{\text{total}}^{(r)} = [\mathbf{Z}_1^{(r)}, \ldots, \mathbf{Z}_J^{(r)}]$ in which $\mathbf{Z}_j^{(r)} = \emptyset$ if $j \notin \Theta^{(r)}$ (i.e., $\mathbf{Z}_j^{(r)}$ is dropped if $j \notin \Theta^{(r)}$). Then the left principal singular vector of $\mathbf{Z}^{(r)}$ is a basis vector for $S^{(r)}$ (e.g., [42]). Thus, if $\mathbf{u}^{(r)}$ denotes the left principal singular vector of $\mathbf{Z}_{\text{total}}^{(r)}$, then we set $\mathbf{a}_r = \mathbf{u}^{(r)}$. Furthermore, since range $\left(\mathbf{Z}_{\text{total}}^{(r)}\right) = \text{range}\left(\mathbf{Z}_{\text{total}}^{(r)} \cdot \mathbf{Z}_{\text{total}}^{(r)H}\right)$, the left principal singular vector $\mathbf{u}^{(r)}$ can be found more efficiently from the matrix $\mathbf{Z}_{\text{total}}^{(r)} \cdot \mathbf{Z}_{\text{total}}^{(r)H} = \sum_{j\in\Theta^{(r)}} \widetilde{\mathbf{Z}}_j^{(r)} \in \mathbb{C}^{I\times I}$, where $\widetilde{\mathbf{Z}}_j^{(r)} := \mathbf{Z}_j^{(r)} \cdot \mathbf{Z}_j^{(r)H} \in \mathbb{C}^{I\times I}$ with

$$\widetilde{\mathbf{Z}}_j^{(r)}([\mu(1), \ldots, \mu(I_{j,r})], [\mu(1), \ldots, \mu(I_{j,r})]) = \frac{\mathbf{z}_j^{(r)}\mathbf{z}_j^{(r)H}}{\|\mathbf{z}_j^{(r)}\|_F^2}, \tag{4.10}$$

$$\widetilde{\mathbf{Z}}_j^{(r)}(:, [\upsilon(1), \ldots, \upsilon(I_{j,r}^c)]) = \mathbf{I}_I(:, [\upsilon(1), \ldots, \upsilon(I_{j,r}^c)]) = \left[\mathbf{e}_{\upsilon(1)}^{(I)}, \ldots, \mathbf{e}_{\upsilon(I_{j,r}^c)}^{(I)}\right], \tag{4.11}$$

in which $\mu(1), \ldots, \mu(I_{j,r})$ denote the indices of the observed entries of $\mathbf{p}_j^{(r)}$. Note that (4.10) only amounts to the computation of an $(I_{j,r} \times I_{j,r})$ rank-1 matrix while (4.11) boils down to the determination of the diagonal entries of $\widetilde{\mathbf{Z}}_j^{(r)}$, i.e., $(\widetilde{\mathbf{Z}}_j^{(r)})_{kk} = 1$ if $k \in \{\upsilon(1), \ldots, \upsilon(I_{j,r}^c)\}$ and $(\widetilde{\mathbf{Z}}_j^{(r)})_{kk} = 0$ otherwise.

4.3.3. Recovery of $\mathbf{b}_r$ from $\mathbf{P}^{(r)}$ given $\mathbf{a}_r$. Define $\widehat{\mathbf{a}}_r = \mathbf{a}_r([\mu(1), \ldots, \mu(I_{j,r})])$, where $\mu(1), \ldots, \mu(I_{j,r})$ denote the indices of the observed entries of $\mathbf{p}_j^{(r)}$. Then the non-zero entries of $\mathbf{b}_r$ follows from the relation $b_{jr} = \widehat{\mathbf{a}}_r^H \widehat{\mathbf{z}}_j^{(r)} / (\widehat{\mathbf{a}}_r^H \widehat{\mathbf{a}}_r)$, $j \in \Theta^{(r)}$, where $\widehat{\mathbf{z}}_j^{(r)} = [z_{\mu(1),j}^{(r)} \ldots, z_{\mu(I_{j,r}),j}^{(r)}]^T$ is the part of $\mathbf{z}_j^{(r)}$ that is associated with the observed entries of $\mathbf{p}_j^{(r)}$.

4.4. Summary and discussion. The overall procedure is summarized as Algorithm 1. If the conditions in Theorem 3.6 are satisfied, then Algorithm 1 is guaranteed to find the decomposition in the exact case. In the inexact case, it can be used as an initialization for an optimization-based method (e.g., [44]).

Note that the complexity of solving the SD problem in step 4 via an EVD is only of the order $\mathcal{O}(R^3)$. Hence, for large dimensions, the computational cost of Algorithm 1 will be dominated by the dimensionality reduction (i.e., step 1) and

the construction of $\mathbf{Q}_{\text{red}}$ (i.e., steps 2 and 3). The complexity of step 1 can be significantly reduced by utilizing randomized compression [17] or column/row/fiber subset selection [14] factorization techniques. The complexity of steps 2 and 3 can be reduced by considering only part of the structure as in multi-grid sampling, cf. Section 3.1.2.

Algorithm 1 SD method for CPD with missing fibers.

Input: $\mathbf{Y}_{\text{sub}} = \mathbf{S}_{\mathbf{W}_1}\mathbf{X}$ of the form (3.4).

1. Compute SVD $\mathbf{Y}_{\text{sub}} = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{V}_{\text{sub}}^H$ (and find R from $\mathbf{\Sigma}_{\text{sub}}$).
2. Build $\mathbf{Q}_{\text{red}}$ in (4.8) from $\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$.
3. Find a matrix $\mathbf{M}$ whose columns form a basis for W in (4.3) via $\ker(\mathbf{Q}_{\text{red}})$.
4. Solve SD problem $(\mathbf{F}^{-T} \odot \mathbf{F}^{-T})\mathbf{N}^T = \mathbf{M}$.
5. Compute $\mathbf{C} = \mathbf{V}^*\mathbf{F}^{-1}$ and $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{F}^T$.
6. Compute $[\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_R] = \mathbf{S}_{\mathbf{W}_1}^T\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$.
7. Build $\mathbf{P}^{(r)}$ defined by (3.5) from $\boldsymbol{\varphi}_r$, $r \in \{1, \ldots, R\}$.
8. Obtain $\{\mathbf{a}_r, \mathbf{b}_r\}$ via rank-1 completion of $\mathbf{P}^{(r)}$, $r \in \{1, \ldots, R\}$.

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

4.5. Remark on more general incomplete tensors. Let us briefly explain that by combining the previous results with matrix completion, fiber sampling can easily be extended to cases where the entries are missing more randomly. We mean that the sampled fibers $\mathbf{x}_{ij\bullet}$ may themselves be incomplete (i.e., some of the entries of the fibers are missing). Let $\mathbf{S}_{\text{sel}} \in \mathbb{C}^{F \times IJ}$ denote a row-selection matrix that selects $F \leq IJ$ fibers of the incomplete observation matrix $\mathbf{Y}$ given by (3.2) so that

$$\mathbf{Y}_{\text{sel}} = \mathbf{S}_{\text{sel}}\mathbf{Y} = \mathbf{S}_{\text{sel}}(\mathbf{W} * \mathbf{X}) = \mathbf{W}_{\text{sel}} * (\mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B})\mathbf{C}^T), \tag{4.12}$$

where $\mathbf{W}_{\text{sel}} = \mathbf{S}_{\text{sel}}\mathbf{W} \in \{0, 1\}^{F \times K}$ and $\mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B}) = \mathbf{S}_{\text{sel}}(\mathbf{A} \odot \mathbf{B}) \in \mathbb{C}^{F \times R}$. From (4.12) it is clear that the differences between complete and incomplete fiber sampling are that in the latter case (i) a basis for range $(\mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B}))$ is not immediately accessible and that (ii) additional conditions for the recovery of $\mathbf{C}$ are required. However, these two issues do not pose a major problem. Briefly, assume that we dispose of a matrix $\mathbf{U}_{\text{sel}} \in \mathbb{C}^{F \times R}$ whose columns span range $(\mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B}))$. Such a matrix can be found using relatively well developed methods for matrix completion.[4] Then the problem is readily reduced to a CPD completion problem of the type we have studied so far. Indeed, since the columns of $\mathbf{U}_{\text{sel}}$ span range $(\mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B}))$, there exists a nonsingular matrix $\mathbf{F} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}_{\text{sel}} = \mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B})\mathbf{F}^T, \tag{4.13}$$

which is a compressed variant of (3.4). Assuming that the conditions in Theorem 3.6 are satisfied, $\{\mathbf{A}, \mathbf{B}, \mathbf{F}\}$ can be obtained via the CPD of a tensor that has missing fibers. If additionally $\mathbf{D}_{\mathbf{W}_k}(\mathbf{A} \odot \mathbf{B})$ has full column rank for every $k \in \{1, \ldots, K\}$, then the remaining factor $\mathbf{C}$ can again be recovered via the relation (3.3), despite the incompleteness of the sampled fibers.

[4] A simple subspace intersection method that can be used for the determination of a basis for range $(\mathbf{\Phi}_{\text{sel}}(\mathbf{A}, \mathbf{B}))$ has been discussed in Subsection 4.3. However, more sophisticated methods for finding subspaces from incomplete data have been developed; see for instance [4, 20, 27, 28, 26] and references therein.

4.6. Remark on dimensionality reduction. CPD algorithms usually involve a preprocessing step in which the tensor is orthogonally compressed. Such a dimensionality reduction step can easily be integrated into multi-grid sampling as well. Consider a multi-grid sampling version of a tensor $\mathcal{X}$ of the form (3.27) that involves the observed subtensors $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(N)}$. Since all fibers of $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(N)}$ are observed, standard truncated multilinear SVD [39, 6] can be used in a preprocessing step. In short, let $\mathbf{U}_{\mathbf{A}}^{(n)} \in \mathbb{C}^{I_n \times \widehat{I}_n}$ with $\widehat{I}_n < I_n$ and $\mathbf{U}_{\mathbf{B}}^{(n)} \in \mathbb{C}^{J_n \times \widehat{J}_n}$ with $\widehat{J}_n < J_n$ be compression matrices obtained via a multilinear SVD of $\mathcal{X}^{(n)}$. Then by replacing $\mathbf{X}^{(n)}$ with its compressed version $\widetilde{\mathbf{X}}^{(n)} = (\mathbf{U}_{\mathbf{A}}^{(n)H} \otimes \mathbf{U}_{\mathbf{B}}^{(n)H})\mathbf{X}^{(n)}$, the factor matrix $\mathbf{C}$ of $\mathcal{X}$ can be found via the coupled CPD of a set of smaller sized tensors $\widetilde{\mathcal{X}}^{(n)} \in \mathbb{C}^{\widehat{I}_n \times \widehat{J}_n \times K}$, $n \in \{1, \ldots, N\}$, while $\mathbf{A}$ and $\mathbf{B}$ follow from $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{Y}_{\text{sub}}\mathbf{C}^{-T}$ via rank-1 completion.

5. Numerical experiments. We demonstrate the behaviour of fiber sampling based on synthetic data in MATLAB. We consider the CPD of an $(I \times J \times K)$ tensor $\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$. The goal is to estimate the factor matrices from the observed tensor $\mathcal{Y} = \mathcal{W} * (\mathcal{X} + \mathcal{N})$, where $\mathcal{W}$ is a binary observation indicator tensor and $\mathcal{N}$ is an unstructured perturbation tensor. The entries of the factor matrices in the CPD of $\mathcal{X}$ and the perturbation tensor $\mathcal{N}$ are randomly drawn from a Gaussian distribution with zero mean and unit variance. The $(0, 1)$-pattern of $\mathcal{W}$ depends on the experiment.

The following Signal-to-Noise Ratio (SNR) measure will be used:

$$\text{SNR} = 10 \log \left(\|\mathbf{W} * \mathbf{X}\|_F^2 \,/\, \|\mathbf{W} * \mathbf{N}\|_F^2 \right).$$

The performance evaluation will be based on the distance between a factor matrix $\mathbf{A}$ and its estimate, $\widehat{\mathbf{A}}$. The distance is measured according to the criterion:

$$\text{P}(\mathbf{A}) = \min_{\Pi\Lambda} \left\| \mathbf{A} - \widehat{\mathbf{A}}\Pi\Lambda \right\|_F \,/\, \|\mathbf{A}\|_F ,$$

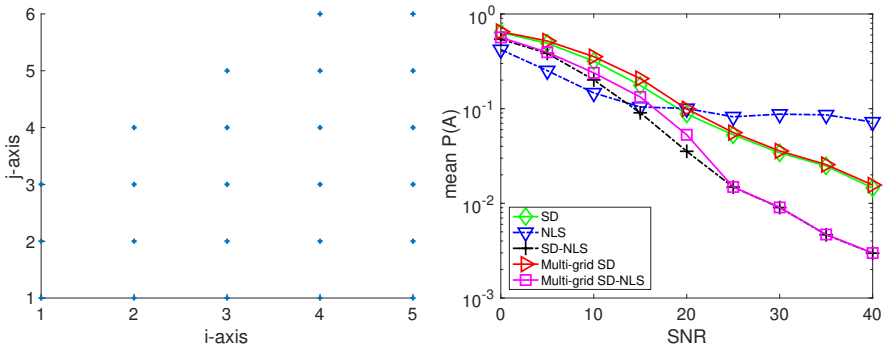
where Π and Λ denote a permutation matrix and a diagonal matrix, respectively.

5.1. Case 1: Systematically missing fibers. Consider the rank-R tensor $\mathcal{X} \in \mathbb{C}^{5 \times 6 \times K}$ depicted in figure 5.1 (left) in which the fibers $\mathbf{x}_{11\bullet}, \mathbf{x}_{21\bullet}, \mathbf{x}_{31\bullet}, \mathbf{x}_{21\bullet}, \mathbf{x}_{22\bullet}$ and $\mathbf{x}_{13\bullet}$ are missing. We set $R = 5$ and $K = 40$. We compare (i) Algorithm 1, (ii) a multi-grid sampled version of Algorithm 1 in which only the grids depicted in figure 5.1 (left) are used in the computation of $\mathbf{C}$, and (iii) the randomly initialized NLS method `sdf_nls.m` in [44]. The methods will be referred to as 'SD', 'Multi-grid SD' and 'NLS', respectively. In addition, we also consider the NLS method initialized by the estimates obtained by the SD and Multi-grid SD methods. They will be referred to as 'SD-NLS' and 'Multi-grid SD-NLS'. The fiber sampling pattern and the mean $\text{P}(\mathbf{A})$ values over 100 trials as a function of SNR can be seen in figure 5.1. We observe that the algebraic SD and Multi-grid SD methods performed about same. The reason is that the Multi-grid SD method already exploits all $(2 \times 2 \times K)$ rank-1 structured subtensors of $\mathcal{X}$. We also observe that the optimization based NLS method is sensitive w.r.t. its initialization and that initialization by a SD method leads to a considerable improvement.

5.2. Case 2: Randomly missing fibers. Consider now the rank-5 tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ with $I = J = K = 50$. We randomly sample 15 percent of the fibers of $\mathcal{X}$. In order to make sure that the CPD factor matrices can be recovered from the fiber

sampled version of $\mathcal{X}$, the border fibers $\{\mathbf{x}_{i1\bullet}\}_{i=1}^{I}$ and $\{\mathbf{x}_{1j\bullet}\}_{j=1}^{J}$ are also sampled. In total, 459 out of 2500 fibers are sampled. For the Multi-grid SD method $N = 10$ grids of size (10×10) are randomly selected, meaning that not more than 100 fibers are used in the computation of $\mathbf{C}$. In other words, the Multi-grid SD method solves the CPD recovery problem via a coupled CPD of a set of small N observation tensors $\mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(N)}$ that have fibers missing. The fiber sampling pattern and mean $\mathrm{P}(\mathbf{A})$ values over 100 trials as a function of SNR can be seen in figure 5.2. The algebraic SD method performs better than the Multi-grid SD method. The reason is the latter method does not take all the low-rank structure of $\mathcal{Y}$ into account. Again, we observe that the optimization based NLS method is sensitive w.r.t. its intialization.

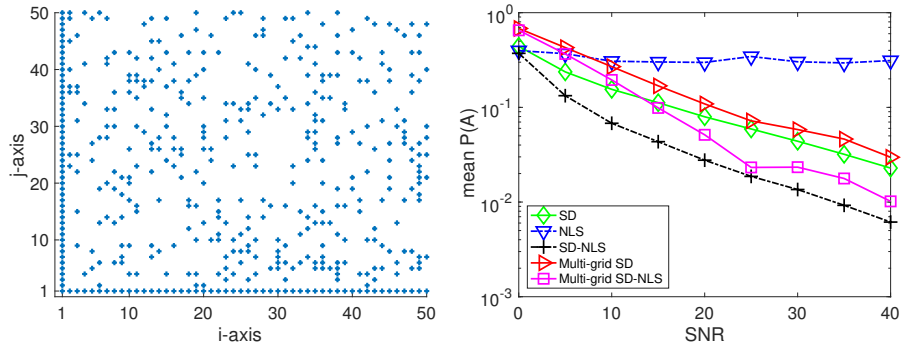
5.3. Case 3: Multi-grid fiber sampling. Consider the rank-10 tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ with large dimensions $I = J = K = 1000$. Note that storage of the full $(1000\times1000\times1000)$ tensor in double precision floating-points would require 8 gigabytes and that the algebraic SD method for the computation of its CPD would have a complexity of the order $\mathcal{O}(I^2J^2R^2) = 10^{10}$. A more practical approach is to consider multi-grid sampling in which $N = 5$ grids of size (10×10) are randomly sampled from $\mathcal{X}$. In order to guarantee the recovery of the CPD factors of $\mathcal{X}$, the border fibers $\{\mathbf{x}_{i1\bullet}\}_{i=1}^{I}$ and $\{\mathbf{x}_{1j\bullet}\}_{j=1}^{J}$ are also sampled. In total $I+J-1+\sum_{n=1}^{N}(I_nJ_n) = 2499$ out of $IJ = 1.000.000$ possible fibers are sampled. Note that in this multi-grid sampling scheme only $\sum_{n=1}^{N}(I_nJ_n) = 500$ fibers are used in the computation of $\mathbf{C}$. We also refine the estimate of $\mathbf{C}$ obtained by SD using the basic Alternating Least Squares (ALS) method for coupled CPD. The latter method will be referred to as 'Multi-grid SD-ALS'. The fiber sampling pattern and the mean $\mathrm{P}(\mathbf{A})$ values over 100 trials as a function of SNR can be seen in figure 5.3. We observe that the algebraic Multi-grid SD and refined 'Multi-grid SD-ALS' methods perform about same. The reason is that the coupled CPD of the sampled CPD is by itself highly overdetermined and consequently little is gained by the extra ALS refinement step.



(a) Fiber sampling pattern. (b) Mean $\mathrm{P}(\mathbf{A})$.

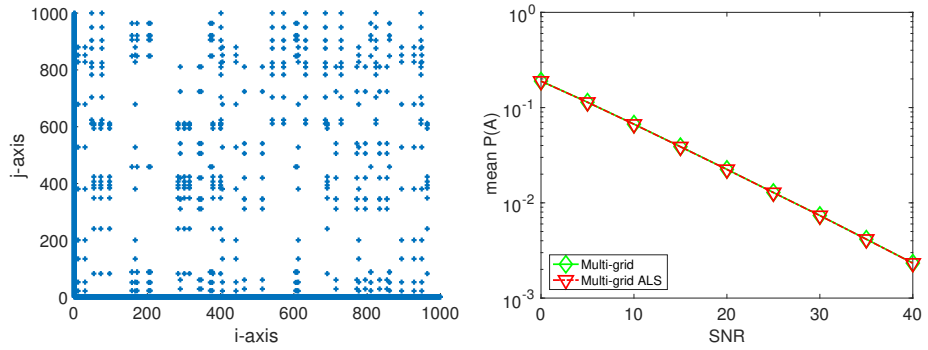
FIG. 5.1. *Fiber sampling pattern and mean $P(\boldsymbol{A})$ over 100 trials while SNR is varying from* 10 *to* 40 *dB and mask, case 1.*

6. Conclusion. In this paper we studied a fiber sampled version of the CPD. This led to a generalization of the conventional CPD modeling framework to the case of missing fibers, including uniqueness conditions and an algebraic algorithm. More precisely, fiber sampling allowed us to reduce a tensor decomposition problem involving missing fibers into simpler matrix completion problems via a matrix EVD.



(a) Fiber sampling pattern. (b) Mean P (**A**).

FIG. 5.2. *Fiber sampling pattern and mean* $P(\boldsymbol{A})$ *over 100 trials while SNR is varying from* 10 *to* 40 *dB and mask, case 2.*



(a) Fiber sampling pattern. (b) Mean P (**A**).

FIG. 5.3. *Fiber sampling pattern and mean* $P(\boldsymbol{A})$ *over 100 trials while SNR is varying from* 10 *to* 40 *dB and mask, case 3.*

Fiber sampling also led to a new multi-grid sampling scheme that can turn a large-scale CPD problem into a small-scale coupled CPD problem. Combining the presented results with matrix completion it is possible to extend fiber sampling to incomplete tensor decomposition problems in which the sampling pattern is unstructured. Even though we limited the discussion to the CPD model, the result can be extended to other low-rank tensor decomposition models, such as the block term decomposition [8]. In the supplementary material in [33] we illustrate this for fiber sampled block term and hierarchical rank decompositions.

Acknowledgements: The authors thank Professor Alwin Stegeman for early proof reading and for several helpful comments.

REFERENCES

[1] E. ACAR, D. M. DUNLAVY AND T. G. KOLDA AND M. MØRUP, *Scalable tensor factorizations for incomplete data*, Chemometr. Intell. Lab., 106(2011), pp. 41-56.

[2] R. B. BAPAT, *Graphs and matrices*, Springer, 2011.

[3] J. D. Carroll and J. J. Chang, *Analysis of individual differences in multidimensional scaling via N-way generalization of Eckart-Young decomposition*, Psychometrika, 35(1970), pp. 283-319.

[4] E. Candès and B. Recht, *Exact matrix completion via convex optimization*, Found. Comput. Math., 9(2009), pp. 717-772.

[5] L. Chiantini and G. Ottaviani, *On generic identifiability of 3-tensors of small rank*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 855–875.

[6] L. De Lathauwer, B. De Moor and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.

[7] L. De Lathauwer, *A Link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 642–666.

[8] L. De Lathauwer, *Decomposition of a higher-order tensor in block terms — Part II: Definitions and uniqueness*, SIAM J. Matrix Anal. Appl., 30(2008), pp. 1033-1066.

[9] 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., 34 (2013), pp. 855–875.

[10] I. Domanov and L. De Lathauwer, *On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Overall uniqueness*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 876–903.

[11] I. Domanov and L. De Lathauwer, *Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 636–660.

[12] I. Domanov and L. De Lathauwer, *Generic uniqueness condition for the canonical polyadic decomposition and Indscal*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 1567-1589.

[13] I. Domanov and L. De Lathauwer, *Canonical polyadic decomposition of third-order tensors: Relaxed uniqueness conditions and algebraic algorithm*, Linear Algebra Appl., 513 (2017), pp. 342–375.

[14] P. Drineas and M. W. Mahoney, *A randomized algorithm for a tensor-based generalization of the singular value decomposition*, Linear Algebra and Appl., 420 (2007), pp. 553–571.

[15] S. Gandy, B. Recht and I. Yamada, *Tensor completion and low-n-rank tensor recovery via convex optimization*, Inverse Problems 27 (2011).

[16] D. Hadwin, K. J. Harrison and J. A. Ward, *Rank-one completions of partial matrices and completely rank-nonincreasing linear functionals*, P. Am. Math. Soc., 136(2006), pp. 2169-2178.

[17] N. Halko and P. G. Martinsson and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Review, 53(2011), pp. 217-288.

[18] R. A. Harshman, *Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis*, UCLA Working Papers in Phonetics, 16(1970), pp. 1–84.

[19] I. Ibraghimov, *Application of the three-way decomposition for matrix compression*, Numer. Linear Algebra Appl., 9(2002), pp. 551-565.

[20] D. W. Jacobs, *Linear fitting with missing data for structure-from-motion*, Computer Vision and Image Understanding, 81(2001), pp. 57–81.

[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 Processing, 52(2004), pp. 2625–2636.

[22] J. B. Kruskal, *Three-way arrays: Rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics*, Linear Algebra and its Applications, 18 (1977), pp. 95–138.

[23] S. Leurgans, R. T. Ross, and R. B. Abel, *A decomposition for three-way arrays*, SIAM J. Matrix Anal. Appl., 14(1993), pp. 1064–1083.

[24] M. Mayzel, J. Rosenlow, L. Isaksson and V. Y. Orekhov, *Time-resolved multidimensional NMR with non-uniform sampling*, J. Biomol. NMR, 58(2014), pp. 129–1393.

[25] V. Y. Orekhov, I. Ibraghimov and M. Billeter, *Optimizing resolution in multidimensional NMR by three-way decomposition*, J. Biomol. NMR, 27(203), pp. 165–173.

[26] D. L. Pimentel-Alarcón, N. Boston and R. D. Nowak, *A characterization of deterministic sampling patterns for low-rank matrix completion*, IEEE J. Sel. Topics Signal Process., 10 (2016), pp. 623–636.

[27] B. Recht, *A simpler approach to matrix completion*, Journal of Machine Learning Research, 12(2011), pp. 3413-3430.

[28] B. Recht, M. Fazel and P. A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev., 52(2011), pp. 471–501.

[29] 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, 65(2017), pp. 3551–3582.

[30] 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., 36(2015), pp. 496–522.

[31] 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., 36(2015), pp. 1015-1045.

[32] M. Sørensen and L. De Lathauwer, *Shift invariance, incomplete arrays and coupled CPD: a case study*, Proc. of The Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2016), Rio de Janeiro, Brazil, July 2016.

[33] M. Sørensen and L. De Lathauwer, *Fiber sampling approach to canonical polyadic decomposition and tensor completion*, Technical Report 15-151, ESAT-STADIUS, KU Leuven, Belgium, 2017.

[34] 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, 71(2006), pp. 219–229.

[35] A. Stegeman and N. D. Sidiropoulos, *On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition*, Linear Algebra and its Applications, 420 (2007), pp. 540–552.

[36] A. Stegeman, *On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode*, Linear Algebra and its Applications, 431(2009), pp. 211-227.

[37] V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra and its Applications, 52 (1983), pp. 645–685.

[38] G. Tomasi and R. Bro, *PARAFAC and missing values*, Chemometr. Intell. Lab., 75(2005), pp. 163-180.

[39] L. R. Tucker, *Some mathematical notes on the three-mode factor analysis*, Psychometrika, 31(1966), pp. 279–311.

[40] V. Tugarinov, L. E. Kay, I. Ibraghimov and V. Y. Orekhov, *High-resolution four-dimensional 1H–13C NOE spectroscopy using Methyl-TROSY, sparse data acquisition, and multidimensional decomposition*, J. Am. Chem. Soc., 127(2005), pp. 2767-2775.

[41] M. Vandecappelle, M. Boussé, F. Van Eeghem and L. De Lathauwer, *Tensor decompositions for graph clustering*, Internal Report 16-170, ESAT-STADIUS, KU Leuven (Leuven, Belgium), 2016.

[42] A.-J. van der Veen, *Blind subspace-based uplink receiver algorithm for wideband CDMA*, Proc. IEEE VTC'99-Fall, Amsterdam, The Netherlands, September 1999.

[43] N. Vervliet, O. Debals, L. Sorber and L. De Lathauwer, *Breaking the curse of dimensionality using decompositions of incomplete tensors*, IEEE Signal Process. Mag., 31(2014), pp. 71–79.

[44] N. Vervliet, O. Debals, L. Sorber, M. Van Barel and L. De Lathauwer, *Tensorlab v3.0*, Available online, March 2016, http://www.tensorlab.net/.

[45] M. Yuan and C.-H. Zhang, *On tensor completion via nuclear norm minimization*, Foundations of Computational Mathematics, 16(2016), pp. 1031–1068.

[46] V. Yu, V. Y. Orekhov and V. A. Jaravine, *Analysis of non-uniformly sampled spectra with multi-dimensional decompositions*, Prog. Nucl. Magn. Reson. Spectrosc., 59(2011), pp. 271–292.

SUPPLEMENTARY MATERIAL

MIKAEL SØRENSEN * AND LIEVEN DE LATHAUWER †

S.1. Block Term Decomposition (BTD) with missing fibers. In this section we briefly explain that the CPD with missing fibers framework presented in [10] can also be extended to the BTD case.

S.1.1. Model. Assume that $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ admits the BTD [1]:

$$\mathcal{X} = \sum_{r=1}^{R}\sum_{l=1}^{L} \mathbf{a}_l^{(r)} \circ \mathbf{b}_l^{(r)} \circ \mathbf{c}^{(r)} = \sum_{r=1}^{R} \left(\mathbf{A}^{(r)}\mathbf{B}^{(r)T}\right) \circ \mathbf{c}^{(r)}, \tag{S.1.1}$$

where $\mathbf{A}^{(r)} = [\mathbf{a}_1^{(r)}, \ldots, \mathbf{a}_L^{(r)}] \in \mathbb{C}^{I\times L}$ and $\mathbf{B}^{(r)} = [\mathbf{a}_1^{(r)}, \ldots, \mathbf{a}_L^{(r)}] \in \mathbb{C}^{I\times L}$. Consider now the BTD model (S.1.1) with missing data (cf. [10, Eq. (3.1)]):

$$\mathcal{Y} = \mathcal{W} * \mathcal{X} = \mathcal{W} * \left(\sum_{r=1}^{R}\sum_{l=1}^{L} \mathbf{a}_l^{(r)} \circ \mathbf{b}_l^{(r)} \circ \mathbf{c}^{(r)}\right), \tag{S.1.2}$$

where $\mathcal{W} \in \{0,1\}^{I\times J\times K}$ is a binary tensor where $'0'$ and $'1'$ indicate that the corresponding entry of $\mathcal{X}$ is unknown and known, respectively. The matrix representation of (S.1.2) takes the form (cf. [10, Eq. (3.2)]):

$$\mathbf{Y} = \mathbf{W} * \mathbf{X} = \mathbf{W} * \left(\mathbf{E}\mathbf{C}^{(\mathrm{red})T}\right), \tag{S.1.3}$$

where

$$\mathbf{E} = \left[\mathrm{Vec}\left(\mathbf{B}^{(1)}\mathbf{A}^{(1)T}\right), \ \ldots, \ \mathrm{Vec}\left(\mathbf{B}^{(R)}\mathbf{A}^{(R)T}\right)\right] \in \mathbb{C}^{IJ\times R}, \tag{S.1.4}$$

$$\mathbf{C}^{(\mathrm{red})} = [\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(R)}] \in \mathbb{C}^{K\times R}. \tag{S.1.5}$$

In the case of missing fibers, (S.1.3) reduces to (cf. [10, Eq. (3.3)]):

$$\mathbf{Y} = \mathbf{D}_{\mathbf{W}_1}\mathbf{X} = \mathbf{D}_{\mathbf{W}_1}\mathbf{E}\mathbf{C}^{(\mathrm{red})T}, \tag{S.1.6}$$

where we recall that $\mathbf{D}_{\mathbf{W}_1} = D_1\left(\mathbf{w}_1^T\right)$ in which $\mathbf{w}_1$ denotes the first column of $\mathbf{W}$ in (S.1.3). Relation (S.1.6) can also be written more compactly as (cf. [10, Eq. (3.4)]):

$$\mathbf{Y}_{\mathrm{sub}} = \mathbf{S}_{\mathbf{W}_1}\mathbf{X} = \mathbf{\Phi}(\mathbf{A},\mathbf{B})(\mathbf{I}_R \otimes \mathbf{1}_L)\mathbf{C}^{(\mathrm{red})T}, \tag{S.1.7}$$

*mr.mikael.sorensen@gmail.com.

†KU Leuven - E.E. Dept. (ESAT) - STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, the Group Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium, and iMinds Medical IT, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, {Mikael.Sorensen, Lieven.DeLathauwer}@kuleuven.be.

Research supported by: (1) Research Council KU Leuven: C1 project C16/15/059-nD, (2) F.W.O.: project G.0830.14N, G.0881.14N, (3) the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012–2017), (4) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

where we recall that $\mathbf{S}_{\mathbf{W}_1} \in \mathbb{C}^{F \times IJ}$ is the row-selection matrix that selects the F rows of $\mathbf{X}$ that have not been zeroed by $\mathbf{D}_{\mathbf{W}_1}$ and that $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{S}_{\mathbf{W}_1} (\mathbf{A} \odot \mathbf{B})$ in which

$$\mathbf{A} = [\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(R)}] \in \mathbb{C}^{I \times RL},$$
$$\mathbf{B} = [\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(R)}] \in \mathbb{C}^{J \times RL}.$$

S.1.2. Uniqueness. In this subsection we explain that Propositions 3.2 and 3.3 in [10] can be generalized to the BTD case (S.1.6). Let us limit the extension to the sufficient condition stated in Proposition 3.3. It will be based on the matrix $\mathbf{G}_{\text{sel}}^{(L)}$ defined by (S.1.12) below.

Define the sets

$$S_I^{L+1} = \{(s_1, \ldots, s_{L+1}) \mid 1 \leq s_1 < \cdots < s_{L+1} \leq I\}\,, \tag{S.1.8}$$
$$U_J^{L+1} = \{(u_1, \ldots, u_{L+1}) \mid 1 \leq u_1 < \cdots < u_{L+1} \leq J\}\,, \tag{S.1.9}$$

in which the elements are ordered lexicographically. The elements of S_I^{L+1} are indexed by $S_I^{L+1}(1), \ldots, S_I^{L+1}(C_I^{L+1})$. Similarly, the elements of U_J^{L+1} are indexed by $U_J^{L+1}(1), \ldots, U_J^{L+1}(C_J^{L+1})$. Consider also the subtensors

$$\mathcal{X}^{(p,q)} = \mathcal{X}([S_I^{L+1}(p)], [U_J^{L+1}(q)], :), \tag{S.1.10}$$

in which $[(s_1, \ldots, s_{L+1})]$ and $[(u_1, \ldots, u_{L+1})]$ mean $[s_1, \ldots, s_{L+1}]$ and $[u_1, \ldots, u_{L+1}]$, respectively. The construction of $\mathbf{G}_{\text{sel}}^{(L)}$ will also be based on the column-selection matrix $\mathbf{P} \in \mathbb{C}^{C_{RL}^{L+1} \times (C_{R+L}^{L+1} - R)}$. Here we only state how $\mathbf{P}$ is defined, see [9] for more explanation of the construction. The $C_{R+L}^{L+1} - R$ columns of $\mathbf{P}$ are indexed by the lexicographically ordered tuples in the set

$$\Gamma_c = \{(j_1, \ldots, j_{L+1}) \mid 1 \leq j_1 \leq \cdots \leq j_{L+1} \leq R\} \setminus \{(j, \ldots, j)\}_{j=1}^{R}.$$

We will also use the mapping $f_c : \mathbb{N}^{L+1} \to \{1, 2, \ldots, C_{R+L}^{L+1} - R\}$ that returns the position of its argument in the set Γ_c. Similarly, the C_{RL}^{L+1} rows of $\mathbf{P}$ are indexed by the lexicographically ordered tuples in the set

$$\Gamma_r = \{(i_1, \ldots, i_{L+1}) \mid 1 < i_1 < \cdots < i_{L+1} \leq RL\}.$$

Likewise, we define the mapping $f_r : \mathbb{N}^{L+1} \to \{1, 2, \ldots, C_{RL}^{L+1}\}$ that returns the position of its argument in the set Γ_r. The entries of $\mathbf{P}$ are given by

$$(\mathbf{P})_{f_r(i_1,\ldots,i_{L+1}), f_c(j_1,\ldots,j_{L+1})} = \begin{cases} 1, & \text{if } \lceil \frac{i_1}{L} \rceil = j_1, \ldots, \lceil \frac{i_{L+1}}{L} \rceil = j_{L+1}, \\ 0, & \text{otherwise.} \end{cases} \tag{S.1.11}$$

We can now define $\mathbf{G}_{\text{sel}}^{(L)} \in \mathbb{C}^{C_I^{L+1} C_J^{L+1} \times (C_{R+L}^{L+1} - R)}$ as follows

$$\mathbf{G}_{\text{sel}}^{(L)} = \mathbf{D}_{\text{sel}}^{(L)} \left(C_{L+1}(\mathbf{A}) \odot C_{L+1}(\mathbf{B})\right) \mathbf{P}, \tag{S.1.12}$$

in which the diagonal matrix $\mathbf{D}_{\text{sel}}^{(L)} = D_1(\mathbf{d}_{\text{sel}}^{(L)T})$ in (S.1.12) is given by

$$\mathbf{d}_{\text{sel}}^{(L)} = [w_{(1,1)}, w_{(1,2)}, \ldots, w_{(C_I^{L+1}, C_J^{L+1})}]^T, \tag{S.1.13}$$

where

$$w_{(p,q)} = \begin{cases} 1, & \text{if all fibers in } \mathcal{X}^{(p,q)} \text{ are sampled,} \\ 0, & \text{otherwise.} \end{cases}$$

PROPOSITION S.1.1. *Consider a tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$, partially observed as $\mathcal{Y} = \mathcal{W} * \mathcal{X}$, and its BTD is given by* (S.1.1). *If*

$$\begin{cases} \boldsymbol{C}^{(red)} \text{ in (S.1.3) has full column rank,} \\ \boldsymbol{G}_{sel}^{(L)} \text{ in (S.1.12) has full column rank,} \end{cases} \tag{S.1.14}$$

then the BTD rank of $\mathcal{X}$ is R and the factor matrix $\boldsymbol{C}^{(red)}$ is unique.

The extension of Theorems 3.6 and 3.7 in [10] to the BTD case (S.1.6) is more complicated. The reason is that it is harder to obtain uniqueness conditions for the rank-L completion of the matrix

$$\mathbf{P}^{(r)} = \mathbf{N} * (\mathbf{A}^{(r)}\mathbf{B}^{(r)T}), \quad n_{ij} = \begin{cases} 1, & \text{if fiber } \mathcal{X}_{ij\bullet} \text{ is sampled,} \\ 0, & \text{otherwise.} \end{cases} \tag{S.1.15}$$

In short, $\mathbf{A}^{(r)}\mathbf{B}^{(r)T}$ is the unique rank-L completion of $\mathbf{P}^{(r)}$ if the solution to the systems of determinantal equations

$$C_{L+1}(\mathbf{P}^{(r)}) = \mathbf{0} \tag{S.1.16}$$

is unique. Here uniqueness obviously means that if $C_{L+1}(\widehat{\mathbf{A}}^{(r)}\widehat{\mathbf{B}}^{(r)T}) = \mathbf{0}$, then $\widehat{\mathbf{A}}^{(r)} = \mathbf{A}^{(r)}\mathbf{F}_r$ and $\widehat{\mathbf{B}}^{(r)} = \mathbf{B}^{(r)}\mathbf{F}_r^{-1}$, where $\mathbf{F}_r \in \mathbb{C}^{R\times R}$ is a nonsingular matrix. More practical sufficient conditions that will ensure that the solution to (S.1.16) is unique can be found in the references in [10].

S.1.3. Computation. In this section we briefly extend [10, Algorithm 1] to the BTD case where condition (S.1.14) in Proposition S.1.1 is satisfied.

S.1.3.1. Conversion to SD.

SD method for computation of BTD. Consider

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{A}^{(r)}\mathbf{B}^{(r)T} \circ \mathbf{f}_r = \sum_{r=1}^{R}\sum_{l=1}^{L} \mathbf{a}_l^{(r)} \circ \mathbf{b}_l^{(r)} \circ \mathbf{f}_r \in \mathbb{C}^{I\times J\times R}$$

in which $\mathbf{A}^{(r)} \in \mathbb{C}^{I\times L}$ and $\mathbf{B}^{(r)} \in \mathbb{C}^{J\times L}$. The BTD extension will make use of the matrix $\mathbf{R}_{L+1}(\mathcal{X}) \in \mathbb{C}^{C_I^{L+1}C_J^{L+1}\times R^{L+1}}$ whose

$$\big((r_{L+1}-1)R^L + (r_L-1)R^{L-1} + \cdots + (r_2-1)R + r_1\big) - \text{th column}$$

is given by

$$\text{Vec}\left(\sum_{k=1}^{L+1}(-1)^{L+1-k} \sum_{1\le p_1<p_2<\cdots<p_k\le L+1} \mathbf{Z}_{(p_1,\ldots,p_k)}\right),$$

where $\mathbf{Z}_{(p_1,\ldots,p_k)} = \mathcal{C}_{L+1}\big(\mathcal{X}(:,:,r_{p_1}) + \cdots + \mathcal{X}(:,:,r_{p_k})\big)$. Define

$$\mathbf{F}^{(\text{ext})} = [\mathbf{1}_L^T \otimes \mathbf{f}_1, \ldots, \mathbf{1}_L^T \otimes \mathbf{f}_R]. \tag{S.1.17}$$

Then we will also make use of the matrix $\mathbf{\Psi}_{L+1}(\mathbf{F}^{(\text{ext})}) \in \mathbb{C}^{R^{L+1} \times C_{RL}^{L+1}}$ with columns

$$(L+1)!\pi_S(\mathbf{f}_{i_1}^{(\text{ext})} \otimes \cdots \otimes \mathbf{f}_{i_m}^{(\text{ext})}), \quad 1 \leq i_1 < \cdots < i_{L+1} \leq LR,$$

ordered lexicographically w.r.t. to the entries of the tuple $(i_1, \ldots, i_{L+1})$, and where $\pi_S : \mathbb{C}^{R^{L+1}} \to \mathbb{C}^{R^{L+1}}$ is the symmetrization mapping

$$\pi_S(\mathbf{f}_1^{(\text{ext})} \otimes \cdots \otimes \mathbf{f}_{L+1}^{(\text{ext})}) = \frac{1}{(L+1)!} \sum_{(l_1,\ldots,l_{L+1}) \in P_{L+1}} \mathbf{f}_{l_1}^{(\text{ext})} \otimes \cdots \otimes \mathbf{f}_{l_{L+1}}^{(\text{ext})},$$

in which P_{L+1} denotes the set of all permutations of the set $\{1, \ldots, L+1\}$. See [2, 9] for further explanations. We end up with the identity [9]:

$$[(\mathcal{C}_{L+1}(\mathbf{A}) \odot \mathcal{C}_{L+1}(\mathbf{B}))\, \mathbf{P}]\, \mathbf{\Psi}_{L+1}^{(\text{dis})}(\mathbf{F}^{(\text{ext})})^T = \mathbf{R}_{L+1}(\mathcal{X}), \tag{S.1.18}$$

with factor matrices $\mathbf{A} = [\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(R)}] \in \mathbb{C}^{I \times RL}$ and $\mathbf{B} = [\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(R)}] \in \mathbb{C}^{J \times RL}$, and where $\mathbf{P} \in \mathbb{C}^{C_{RL}^{L+1} \times (C_{R+L}^{L+1} - R)}$ is the column-selection matrix defined according to (S.1.11) and $\mathbf{\Psi}_{L+1}^{(\text{dis})}(\mathbf{F}^{(\text{ext})}) = \mathbf{\Psi}_{L+1}(\mathbf{F}^{(\text{ext})})\mathbf{P}^T$ consists of the distinct columns of $\mathbf{\Psi}_{L+1}(\mathbf{F}^{(\text{ext})})$).

Denote $W = \ker(\mathbf{\Psi}_{L+1}^{(\text{dis})}(\mathbf{F}^{(\text{ext})})^T) \cap \Pi_s \subset \mathbb{C}^{R^{L+1}}$, where

$$\Pi_s = \text{range}\left(\{\pi_S(\mathbf{t}_1 \otimes \cdots \otimes \mathbf{t}_{L+1}) \mid \ \mathbf{t}_1, \ldots, \mathbf{t}_{L+1} \in \mathbb{C}^R\}\right).$$

It was explained in [9] that if $\mathbf{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_R] \in \mathbb{C}^{R \times R}$ has full column rank, then $\mathbf{\Psi}_{L+1}^{(\text{dis})}(\mathbf{F}^{(\text{ext})})^*$ has full column rank, $\dim W = R$, and $[\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_1, \ldots, \mathbf{x}_R \otimes \cdots \otimes \mathbf{x}_R]$ coincides with $\mathbf{F}^{-T} \odot \cdots \odot \mathbf{F}^{-T}$ up to a permutation and column scaling if and only if

$$\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_1, \ldots, \mathbf{x}_R \otimes \cdots \otimes \mathbf{x}_R \quad \text{form a basis of } W. \tag{S.1.19}$$

SD method in the case of missing fibers. Consider now the BTD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ of the form (S.1.1) and with matrix representation (S.1.7). Let $\mathbf{Y}_{\text{sub}} = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{V}_{\text{sub}}^H$ denote the compact SVD of $\mathbf{Y}_{\text{sub}}$, then there exists a nonsingular matrix $\mathbf{F}$ such that

$$\widetilde{\mathbf{U}}_{\text{sub}} := \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}} = \mathbf{\Phi}(\mathbf{A}, \mathbf{B})(\mathbf{I}_R \otimes \mathbf{1}_L) \cdot \mathbf{F}^T \ \text{ and } \ \mathbf{V}_{\text{sub}}^* \mathbf{F}^{-1} = \mathbf{C}^{(\text{red})}. \tag{S.1.20}$$

The combination of (S.1.12), (S.1.17), (S.1.18) and (S.1.20) now yields

$$\mathbf{R}_{\text{red}}^{(L)} = \mathbf{S}_{\text{red}}^{(L)} \mathbf{R}_{L+1}(\widetilde{\mathcal{U}}_{\text{sub}}) \mathbf{G}_{\text{red}}^{(L)} \cdot \mathbf{\Psi}_{L+1}^{(\text{dis})}(\mathbf{F}^{(\text{ext})})^T, \tag{S.1.21}$$

where we recall that $\widetilde{\mathcal{U}}_{\text{sub}} = \mathcal{W} * \left(\sum_{r=1}^R \mathbf{A}^{(r)}\mathbf{B}^{(r)T} \circ \mathbf{f}_r\right)$ and where $\mathbf{S}_{\text{red}}^{(L)} \in \mathbb{C}^{\omega(\mathbf{d}_{\text{red}}^{(L)}) \times C_I^{L+1} C_J^{L+1}}$ denotes the row-selection matrix that only retains the rows of $\mathbf{D}_{\text{sel}}^{(L)} \mathbf{R}_{L+1}(\widetilde{\mathcal{U}}_{\text{sub}})$ that have not been zeroed out by the indicator matrix $\mathbf{D}_{\text{sel}}^{(L)}$, and

$$\mathbf{G}_{\text{red}}^{(L)} = \mathbf{S}_{\text{red}}^{(L,1)} \left(C_{L+1}\left(\mathbf{A}^{(1)}\right) \odot C_{L+1}\left(\mathbf{B}^{(1)}\right)\right) \mathbf{P} \in \mathbb{C}^{\omega(\mathbf{D}_{\text{sel}}^{(L)}) \times C_R^2}$$

is the reduced version of the matrix $\mathbf{G}_{\text{sel}}^{(L)}$ in (S.1.12). Since the matrix $\mathbf{G}_{\text{sel}}^{(L)}$ has full column rank by assumption, it follows that

$$W = \ker(\mathbf{\Psi}_{L+1}^{(\text{dis})}(\mathbf{F}^{(\text{ext})})^T) \cap \Pi_s = \ker(\mathbf{R}_{\text{red}}^{(L)}) \cap \Pi_s. \tag{S.1.22}$$

Let the columns of $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ constitute a basis for W obtained via relation (S.1.22), then there exists a nonsingular matrix $\mathbf{\Gamma} \in \mathbb{C}^{R \times R}$ such that

$$(\mathbf{F}^{-T} \odot \cdots \odot \mathbf{F}^{-T})\mathbf{\Gamma} = \mathbf{M}. \tag{S.1.23}$$

As in the CPD case, the SD factorization in (S.1.23) can in the exact case be computed via a GEVD.

S.1.3.2. Rank-L subspace identification. Once the matrix $\mathbf{F}^{-T}$ has been computed, we obtain $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})(\mathbf{I}_R \otimes \mathbf{1}_L)$ and $\mathbf{C}^{(\text{red})}$ via relation (S.1.20). Compute $[\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_R] = \mathbf{Y}(\mathbf{C}^{(\text{red})T})^{\dagger}$, where we recall that $\mathbf{Y}$ is given by (S.1.6). The goal is now to determine the rank-L subspace of the matrix $\mathbf{A}^{(r)}\mathbf{B}^{(r)T}$ from the incomplete matrix $\mathbf{P}^{(r)}$ built from $\boldsymbol{\varphi}_r$ in accordance to

$$p_{ij}^{(r)} = \begin{cases} (\boldsymbol{\varphi}_r)_{(i-1)I+j}, & \text{if } (\mathbf{w}_1)_{(i-1)I+j} = 1, \\ \text{indeterminate}, & \text{if } (\mathbf{w}_1)_{(i-1)I+j} = 0. \end{cases} \tag{S.1.24}$$

Here we will explain how to extend the rank-1 matrix completion method presented in [10] to the rank-L case. See also [10] for references to more sophisticated matrix completion methods.

From $\mathbf{P}^{(r)}$ we can extract M submatrices, each composed of $\widetilde{J} \geq L$ columns:

$$\mathbf{P}_m^{(r)} = \mathbf{P}^{(r)}\left(:,\ [\gamma_m(1), \ldots, \gamma_m(\widetilde{J})]\right) \in \mathbb{C}^{I \times \widetilde{J}}, \tag{S.1.25}$$

where $\gamma_m(1), \ldots, \gamma_m(\widetilde{J})$ denote the indices of the associated columns of $\mathbf{P}^{(r)}$. (Typically we set $\widetilde{J} = L$.) Assume that $\mathbf{P}_m^{(r)}$ consists of $P_{m,r}$ rows containing only observed entries, indexed by $\mu(1), \ldots, \mu(P_{m,r})$. The remaining $P_{m,r}^c = I - P_{m,r}$ rows of $\mathbf{P}_m^{(r)}$ are indexed by $\upsilon(1), \ldots, \upsilon(P_{m,r}^c)$. Let $\mathbf{Z}_m^{(r)} \in \mathbb{C}^{P_{m,r} \times \widetilde{J}}$ denote the matrix in which the rows of $\mathbf{P}_m^{(r)}$ containing unobserved entries are replaced by all-zero row-vectors. Partition $\mathbf{Z}_m$ into a non-zero and zero part:

$$\hat{\mathbf{Z}}_m^{(r)} = \begin{bmatrix} \mathbf{Z}_m^{(r)}(\mu(1), :) \\ \vdots \\ \mathbf{Z}_m^{(r)}(\mu(P_{m,r}), :) \end{bmatrix} \neq \mathbf{0}, \quad \check{\mathbf{Z}}_m^{(r)} = \begin{bmatrix} \mathbf{Z}_m^{(r)}(\upsilon(1), :) \\ \vdots \\ \mathbf{Z}_m^{(r)}(\upsilon(P_{m,r}^c), :) \end{bmatrix} = \mathbf{0}\ .$$

As in the rank-1 completion case, we define

$$\text{range}\left(\mathbf{P}_m^{(r)}\right) := \text{range}\left([\hat{\mathbf{Z}}_m^{(r)}, \mathbf{e}_{\upsilon(1)}^{(I)}, \ldots, \mathbf{e}_{\upsilon(P_{m,r}^c)}^{(I)}]\right).$$

Dimension and range of $S_m^{(r)} := \bigcap_{m=1}^{M}$ *range* $\left(\boldsymbol{P}_m^{(r)}\right)$. Since $\text{range}\left(\mathbf{A}^{(r)}\right) \subseteq S_m^{(r)}$, $\forall m \in \{1, \ldots, M\}$ we have that $\text{range}\left(\mathbf{A}^{(r)}\right) \subseteq S_m^{(r)}$. As a consequence, if $S_m^{(r)}$ spans an L-dimensional subspace, then $\text{range}\left(\mathbf{A}^{(r)}\right) = S_m^{(r)}$. In other words, $\mathbf{A}^{(r)}\mathbf{B}^{(r)T}$ can be recovered from $\mathbf{P}^{(r)}$. Note in contrast to the rank-1 case, we here make the additional assumption that $S_m^{(r)}$ spans an L-dimensional subspace.

Recovery of $\boldsymbol{A}^{(r)}$ from $S_m^{(r)}$ via subspace method. As in the rank-1 matrix completion case, we let $\hat{\mathbf{Q}}_m^{(r)} \in \mathbb{C}^{P_{m,r} \times \widetilde{J}_m}$ denote an orthonormal basis for range $\left(\hat{\mathbf{Z}}_m^{(r)}\right)$ with $\widetilde{J}_m \leq \widetilde{J}$ (e.g., obtained via the QR factorization of $\hat{\mathbf{Z}}_m^{(r)}$). Build $\widetilde{\mathbf{Z}}_m^{(r)} \in \mathbb{C}^{I \times \widetilde{J}_m}$ as follows:

$$\widetilde{\mathbf{Z}}_m^{(r)}([\mu(1), \ldots, \mu(P_{m,r})], :) = \hat{\mathbf{Q}}_m^{(r)} ,$$
$$\widetilde{\mathbf{Z}}_m^{(r)}([\upsilon(1), \ldots, \upsilon(P_{m,r}^c)], :) = \mathbf{0} .$$

The columns of $\mathbf{N}_m^{(r)} = [\widetilde{\mathbf{Z}}_m^{(r)}, \mathbf{e}_{\upsilon(1)}^{(I)}, \ldots, \mathbf{e}_{\upsilon(P_{m,r}^c)}^{(I)}] \in \mathbb{C}^{I \times (\widetilde{J}_m + P_{m,r}^c)}$ constitute an orthonormal basis for range $(\mathbf{P}_m)$. Stacking for all $m \in \{1, \ldots, M\}$ yields the matrix $\mathbf{N}^{(r)} = [\mathbf{N}_1^{(r)}, \ldots, \mathbf{N}_M^{(r)}] \in \mathbb{C}^{I \times \sum_m^M (\widetilde{J}_m + P_{m,r}^c)}$ from which an orthonormal basis for $S_m^{(r)} = \bigcap_{m=1}^M \text{range}\left(\mathbf{P}_m^{(r)}\right)$ can be derived. Since range $\left(\mathbf{N}^{(r)}\right) = \text{range}\left(\mathbf{N}^{(r)} \cdot \mathbf{N}^{(r)H}\right)$, the basis can be found more efficiently from the matrix $\mathbf{N}^{(r)} \cdot \mathbf{N}^{(r)H} = \sum_{m=1}^M \widetilde{\mathbf{N}}_m^{(r)} \in \mathbb{C}^{I \times I}$, where $\widetilde{\mathbf{N}}_m^{(r)} := \mathbf{N}_m^{(r)} \cdot \mathbf{N}_m^{(r)H} \in \mathbb{C}^{I \times I}$ is given by

$$\widetilde{\mathbf{N}}_m^{(r)}([\mu(1), \ldots, \mu(P_{m,r})], [\mu(1), \ldots, \mu(P_{m,r})]) = \hat{\mathbf{Q}}_m^{(r)} \cdot \hat{\mathbf{Q}}_m^{(r)H},$$
$$\widetilde{\mathbf{N}}_m^{(r,n)}(:, [\upsilon(1), \ldots, \upsilon(P_{m,r}^c)]) = \mathbf{I}_J(:, [\upsilon(1), \ldots, \upsilon(P_{m,r}^c)]).$$

Thus, if range $\left(\mathbf{N} \cdot \mathbf{N}^H\right) = S_m^{(r)} = \bigcap_{m=1}^M \text{range}\,(\mathbf{P}_m)$ spans an L-dimensional subspace, then range $\left(\mathbf{N} \cdot \mathbf{N}^H\right) = \text{range}\left(\mathbf{A}^{(r)}\right)$, implying that $\mathbf{A}^{(r)}$ can be can be obtained from, say the SVD of $\mathbf{N} \cdot \mathbf{N}^H$.

Recovery of $\boldsymbol{B}^{(r)}$ from $\boldsymbol{P}^{(r)}$ given $\boldsymbol{A}^{(r)}$. Once $\mathbf{A}^{(r)}$ has been obtained, then $\mathbf{B}^{(r)}$ follows immediately from the relation

$$\mathbf{B}^{(r)}(j, :) = (\mathbf{A}^{(r)}([\mu(1), \ldots, \mu(I_{j,r})], :)^\dagger \cdot \mathbf{P}^{(r)}([\mu(1), \ldots, \mu(I_{j,r})], j),$$

where the indices $\mu(1), \ldots, \mu(I_{j,r})$ denote the observered entries of the jth column of $\mathbf{P}^{(r)}$.

Remark on fiber sampling pattern. Note that there are $C_J^{\widetilde{J}}$ submatrices of $\mathbf{P}^{(r)}$ of the form (S.1.25). Consequently, we usually have that $M << C_J^{\widetilde{J}}$. For the case of arbitrarily sampled fibers, one can resort to random sampling (e.g., pick out $Q >> M$ submatrices $\mathbf{P}_1^{(r)}, \ldots, \mathbf{P}_Q^{(r)}$ from $\mathbf{P}^{(r)}$ and retain the ones containing L rows with only observed entries). In cases where the fiber sampling pattern can be designed, one may resort to a structured sampling scheme that ensures that the M matrices $\mathbf{P}_1^{(r)}, \ldots, \mathbf{P}_M^{(r)}$ given by (S.1.25) all consist of L rows with only observed entries (e.g., using multi-grid sampling with $\min(I, J) \geq L$).

S.1.3.3. Summary. The overall procedure for BTD with missing fibers is summarized as Algorithm 1.

S.2. Hierarchical rank decompositions. Another example where fiber sampling can be useful is in the case of matrix factorizations of the form

$$\mathbf{X} = \mathbf{M}\mathbf{S}^T \in \mathbb{C}^{I \times L}, \quad \mathbf{M} \in \mathbb{C}^{I \times R}, \quad \mathbf{S} \in \mathbb{C}^{L \times R}, \tag{S.2.1}$$

where the columns of $\mathbf{M}$ are subject to a more locally defined low-rank structure. We note in passing that hierarchical structured decompositions have received a lot

Algorithm 1 SD method for BTD with missing fibers.

Input: $\mathbf{Y}_{\text{sub}} = \mathbf{S}_{\mathbf{W}_1}\mathbf{X}$ of the form (S.1.7).

1. Compute SVD $\mathbf{Y}_{\text{sub}} = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{V}_{\text{sub}}^H$ (and find R from $\mathbf{\Sigma}_{\text{sub}}$).
2. Build $\mathbf{R}_{\text{red}}^{(L)}$ in (S.1.21) from $\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$.
3. Find a basis matrix $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ for $\ker(\mathbf{R}_{\text{red}}^{(L)}) \cap \Pi_S$.
4. Solve SD problem $(\mathbf{F}^{-T} \odot \cdots \odot \mathbf{F}^{-T})\mathbf{\Gamma} = \mathbf{M}$.
5. Compute $\mathbf{C}^{(\text{red})} = \mathbf{V}^*\mathbf{F}^{-1}$ and $\mathbf{\Phi}^{(\text{red})}(\mathbf{E}) = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{F}^T$.
6. Compute $[\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_R] = \mathbf{S}_{\mathbf{W}_1}^T\mathbf{\Phi}^{(\text{red})}(\mathbf{E})$.
7. Build $\mathbf{P}^{(r)}$ defined by (S.1.24) from $\boldsymbol{\varphi}_r$, $r \in \{1, \ldots, R\}$.
8. Obtain $\{\mathbf{A}^{(r)}\mathbf{B}^{(r)T}\}$ via rank-L completion of $\mathbf{P}^{(r)}$, $r \in \{1, \ldots, R\}$.

Output: $\{\mathbf{A}^{(r)}\mathbf{B}^{(r)T}\}$ and $\mathbf{C}^{(\text{red})}$.

of attention in *data compression* via hierarchical Tucker-type decompositions (e.g., [7, 6, 4, 5]). However, hierarchical Tucker-type decompositions do not allow an easy interpretation in terms of signal components, as required for *data analysis*.

Local low-rank decomposition. Note that both the CPD and BTD models assume a *global* low-rank structure on the columns of $\mathbf{M}$ in (S.2.1). However, $\mathbf{m}_r$ may be subject to *local* low-rank structures. More precisely, it is assumed that there exist row-selection matrices $\mathbf{R}^{(n)} \in \mathbb{C}^{I_n \times I}$, $n \in \{1, \ldots, N\}$ such that from (S.2.1) we obtain a collection of low-rank structured matrices:

$$\mathbf{X}^{(n)} = \mathbf{R}^{(n)}\mathbf{X} = \mathbf{M}^{(n)}\mathbf{S}^T, \quad n \in \{1, \ldots, N\}, \tag{S.2.2}$$

in which the locally defined matrices $\mathbf{M}^{(n)} = \mathbf{R}^{(n)}\mathbf{M}$ are subject to a low-rank constraint (e.g., $\mathbf{m}_r^{(n)} = \mathbf{a}_r^{(n)} \otimes \mathbf{b}_r^{(n)}$ or $\mathbf{m}_r = \text{Vec}(\mathbf{B}_r^{(n)}\mathbf{A}_r^{(n)T})$). Since the factorizations $\{\mathbf{M}^{(n)}\mathbf{S}^T\}$ are connected via the coupling matrix $\mathbf{S}$ the different local low-rank structures can be combined via coupled tensor decompositions. We will illustrate this idea using the Hierarchical Rank Decomposition (HRD) model presented below.

Definition of HRD model. Consider the simple case where $\mathbf{m}_r$ can be partitioned into double indexed subvectors $\mathbf{m}_r^{(m,n)}$ supporting locally defined low-rank structures of the form $\mathbf{m}_r^{(m,n)} = \mathbf{a}_r^{(m,n)} \otimes \mathbf{b}_r^{(m,n)} \otimes \mathbf{c}_r^{(n)}$. Remark that BTD variants are also possible (e.g., $\mathbf{m}_r^{(m,n)} = \text{Vec}\left(\mathbf{B}_r^{(m,n)}\mathbf{A}_r^{(m,n)T}\right) \otimes \mathbf{c}_r^{(n)}$).

Consider the following HRD of the tensors $\mathcal{X}^{(m,n)} \in \mathbb{C}^{I_{m,n} \times J_{m,n} \times K_n \times L}$, $m \in \{1, \ldots, M_n\}$, $n \in \{1, \ldots, N\}$:

$$\mathcal{X}^{(m,n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(m,n)} \circ \mathbf{b}_r^{(m,n)} \circ \mathbf{c}_r^{(n)} \circ \mathbf{s}_r \,, \tag{S.2.3}$$

in which each tensor $\mathcal{X}^{(m,n)}$ has been constructed from (reshaped) submatrices of $\mathbf{X}$ in (S.2.1). The exact reshaping and choice of submatrices of $\mathbf{X}$ is problem dependent. Note that the columns of $\mathbf{S}$ affect the full data set $\mathbf{X}$, that the columns of $\mathbf{C}^{(n)}$ appear in $\mathcal{X}^{(m,n)}$, $m \in \{1, \ldots, M_n\}$, and that the columns of $\mathbf{A}^{(m,n)}$ and $\mathbf{B}^{(m,n)}$ are specific for the tensor $\mathcal{X}^{(m,n)}$.

We define the hierarchical rank of $\{\mathcal{X}^{(m,n)}\}$ as the minimal number of hierarchical rank-1 tensors $\mathbf{a}_r^{(m,n)} \circ \mathbf{b}_r^{(m,n)} \circ \mathbf{c}_r^{(n)} \circ \mathbf{s}_r$ that yield $\{\mathcal{X}^{(m,n)}\}$ in a linear combination. The hierarchical rank-1 tensors in (S.2.3) can be arbitrarily permuted and the vectors within the same hierarchical rank-1 tensor can be arbitrarily scaled provided the

overall term remains the same. We say that the HRD is unique when it is only subject to these trivial indeterminacies.

S.2.1. HRD via fiber sampling. Observe that the HRD model (S.2.3) consists of two structures, namely higher-order and nested PD structures. In short, we first consider the two structures separately and thereafter combines the results.

S.2.1.1. Exploiting nested PD structure in $\{\mathcal{X}^{(m,n)}\}$. For every $n \in \{1,\dots,N\}$, the coupled PD of the subset of tensors $\{\mathcal{X}^{(m,n)}\}_{1\leq m\leq M_n}$ in $\{\mathcal{X}^{(m,n)}\}$ has the following matrix representation

$$\mathbf{Y}_{(1)}^{(n)} := \begin{bmatrix} \mathbf{X}^{(1,n)} \\ \vdots \\ \mathbf{X}^{(M_n,n)} \end{bmatrix} = \left(\mathbf{F}_{(1)}^{(n)} \odot \mathbf{C}^{(n)}\right)\mathbf{S}^T, \tag{S.2.4}$$

where

$$\mathbf{F}_{(1)}^{(n)} = \begin{bmatrix} \mathbf{F}_{(1)}^{(1,n)} \\ \vdots \\ \mathbf{F}_{(1)}^{(M_n,n)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(1,n)} \odot \mathbf{B}^{(1,n)} \\ \vdots \\ \mathbf{A}^{(M_n,n)} \odot \mathbf{B}^{(M_n,n)} \end{bmatrix} \in \mathbb{C}^{(\sum_{m=1}^{M_n} I_{m,n}J_{m,n})\times R}, \tag{S.2.5}$$

and $\mathbf{F}_{(1)}^{(m,n)} = \mathbf{A}^{(m,n)} \odot \mathbf{B}^{(m,n)}$. For now we will ignore the Khatri-Rao structure within $\mathbf{F}_{(1)}^{(n)}$, so that (S.2.4) corresponds to a matrix representation of the PD of a third-order tensor $\mathcal{Y}_{(1)}^{(n)} \in \mathbb{C}^{(\sum_{m=1}^{M_n} I_{m,n}J_{m,n})\times K_n\times L}$ in which $\{\mathbf{X}^{(m,n)}\}_{1\leq m\leq M_n}$ are stacked. Overall, for varying n, the coupled PD of $\mathcal{Y}_{(1)}^{(1)},\dots,\mathcal{Y}_{(1)}^{(N)}$ has the following matrix representation

$$\begin{bmatrix} \mathbf{Y}_{(1)}^{(1)} \\ \vdots \\ \mathbf{Y}_{(1)}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{(1)}^{(1)} \odot \mathbf{C}^{(1)} \\ \vdots \\ \mathbf{F}_{(1)}^{(N)} \odot \mathbf{C}^{(N)} \end{bmatrix} \mathbf{S}^T. \tag{S.2.6}$$

The coupled PD of $\{\mathcal{Y}_{(1)}^{(n)}\}$, and expressed by (S.2.6), captures the nested PD structure within the HRD of $\{\mathcal{X}^{(m,n)}\}$.

S.2.1.2. Exploiting higher-order PD structure in $\{\mathcal{X}^{(m,n)}\}$. We now impose that $\mathbf{F}_{(1)}^{(m,n)} = \mathbf{A}^{(m,n)} \odot \mathbf{B}^{(m,n)}$, i.e., we impose that (S.2.3) has fourth-order structure rather than third-order structure. Let $\mathbf{P}_{(2)}^{(m,n)}$ and $\mathbf{P}_{(3)}^{(m,n)}$ denote row-permutation matrices such that

$$\begin{aligned} &\mathbf{P}_{(2)}^{(m,n)}(\mathbf{A}^{(m,n)} \odot \mathbf{B}^{(m,n)} \odot \mathbf{C}^{(n)}) = \mathbf{F}_{(2)}^{(m,n)} \odot \mathbf{A}^{(m,n)}, \qquad \mathbf{F}_{(2)}^{(m,n)} = \mathbf{C}^{(n)} \odot \mathbf{B}^{(m,n)}, \\ &\mathbf{P}_{(3)}^{(m,n)}(\mathbf{A}^{(m,n)} \odot \mathbf{B}^{(m,n)} \odot \mathbf{C}^{(n)}) = \mathbf{F}_{(3)}^{(m,n)} \odot \mathbf{B}^{(m,n)}, \qquad \mathbf{F}_{(3)}^{(m,n)} = \mathbf{C}^{(n)} \odot \mathbf{A}^{(m,n)}. \end{aligned}$$

Define the tensors $\mathcal{Y}_{(2)}^{(m,n)} \in \mathbb{C}^{J_{m,n}K_n\times I_{m,n}\times L}$ and $\mathcal{Y}_{(3)}^{(m,n)} \in \mathbb{C}^{I_{m,n}K_n\times J_{m,n}\times L}$, for which the associated PDs admit the following matrix representations

$$\mathbf{Y}_{(2)}^{(m,n)} = \mathbf{P}_{(2)}^{(m,n)}\mathbf{X}^{(m,n)} = \left(\mathbf{F}_{(2)}^{(m,n)} \odot \mathbf{A}^{(m,n)}\right)\mathbf{S}^T, \tag{S.2.7}$$

$$\mathbf{Y}_{(3)}^{(m,n)} = \mathbf{P}_{(3)}^{(m,n)}\mathbf{X}^{(m,n)} = \left(\mathbf{F}_{(3)}^{(m,n)} \odot \mathbf{B}^{(m,n)}\right)\mathbf{S}^T, \tag{S.2.8}$$

where $1 \leq m \leq M_n$, $1 \leq n \leq N$ and where the Khatri-Rao structure within $\mathbf{F}_{(2)}^{(m,n)}$ and $\mathbf{F}_{(3)}^{(m,n)}$ has been ignored. Overall, we obtain the coupled PD of $\{\mathcal{Y}_{(2)}^{(m,n)}, \mathcal{Y}_{(3)}^{(m,n)}\}_{n=1,\ldots,N}^{m=1,\ldots,M_n}$, which captures the higher-order PD structures within the HRD of $\{\mathcal{X}^{(m,n)}\}$ in (S.2.3).

S.2.1.3. Combination of nested and higher-order PD structures. From (S.2.6)–(S.2.8) it is clear that the HRD problem (S.2.3) can be translated into a coupled CPD problem of $\{\mathcal{Y}_{(1)}^{(n)}, \mathcal{Y}_{(2)}^{(m,n)}, \mathcal{Y}_{(3)}^{(m,n)}\}_{n=1,\ldots,N}^{m=1,\ldots,M_n}$, which takes both the nested and higher-order PD structures in (S.2.3) into account. In other words, fiber sampling has allowed us to transform the HRD problem into a coupled CPD problem. Consequently, uniqueness conditions and algorithms developed for coupled CPD (e.g., [8, 9]) can also be used to solve HRD problems.

As an example, define

$$\mathbf{G} = \left[\mathbf{H}_{(1)}^T,\ \mathbf{H}_{(2)}^T,\ \mathbf{H}_{(3)}^T\right]^T, \tag{S.2.9}$$

where

$$\mathbf{H}_{(1)} = \begin{bmatrix} C_2(\mathbf{F}_{(1)}^{(1)}) \odot C_2(\mathbf{C}^{(1)}) \\ \vdots \\ C_2(\mathbf{F}_{(1)}^{(N)}) \odot C_2(\mathbf{C}^{(N)}) \end{bmatrix},$$

$$\mathbf{H}_{(2)} = \begin{bmatrix} \mathbf{H}_{(2)}^{(1)} \\ \vdots \\ \mathbf{H}_{(2)}^{(N)} \end{bmatrix}, \mathbf{H}_{(2)}^{(n)} = \begin{bmatrix} \mathbf{H}_{(2)}^{(1,n)} \\ \vdots \\ \mathbf{H}_{(2)}^{(M_n,n)} \end{bmatrix}, \mathbf{H}_{(2)}^{(m,n)} = C_2(\mathbf{F}_{(2)}^{(m,n)}) \odot C_2(\mathbf{A}^{(m,n)}),$$

$$\mathbf{H}_{(3)} = \begin{bmatrix} \mathbf{H}_{(3)}^{(1)} \\ \vdots \\ \mathbf{H}_{(3)}^{(N)} \end{bmatrix}, \mathbf{H}_{(3)}^{(n)} = \begin{bmatrix} \mathbf{H}_{(3)}^{(1,n)} \\ \vdots \\ \mathbf{H}_{(3)}^{(M_n,n)} \end{bmatrix}, \mathbf{H}_{(3)}^{(m,n)} = C_2(\mathbf{F}_{(3)}^{(m,n)}) \odot C_2(\mathbf{B}^{(m,n)}),$$

We summarize the result in Proposition S.2.1 below.

PROPOSITION S.2.1. *Consider the HRD of* $\mathcal{X}^{(m,n)} \in \mathbb{C}^{I_{m,n} \times J_{m,n} \times K_n \times L}$, $m \in \{1, \ldots, M_n\}$, $n \in \{1, \ldots, N\}$ *in* (S.2.3). *If*

$$\begin{cases} \boldsymbol{S} \text{ in (S.2.3) has full column rank,} \\ \boldsymbol{G} \text{ in (S.2.9) has full column rank,} \end{cases} \tag{S.2.10}$$

then the hierarchical rank of $\{\mathcal{X}^{(m,n)}\}$ *is* R *and the HRD of* $\{\mathcal{X}^{(m,n)}\}$ *is unique.*

To summarize, from (S.2.3) we built several tensors each exploiting a nested or higher-order defined low-rank structure in $\{\mathcal{Y}^{(m,n)}\}$.

REFERENCES

[1] L. DE LATHAUWER, *Decomposition of a higher-order tensor in block terms — Part II: Definitions and uniqueness*, SIAM J. Matrix Anal. Appl., 30(2008), pp. 1033-1066.

[2] I. DOMANOV AND L. DE LATHAUWER, *Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 636–660.

[3] R. A. HARSHMAN, *Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis*, UCLA Working Papers in Phonetics, 16(1970), pp. 1–84.

[4] L. Grasedyck, *Hierarchical singular value decomposition of tensors*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2029–2054.
[5] W. Hackbusch, *Tensor spaces and numerical tensor calculus*, Springer, 2012.
[6] I. Oseledets and E. E. Tyrtyshnikov, *Breaking the curse of dimensionality, or how to use SVD in many dimensions*, SIAM J. Sci. Comput., 31(2009), pp. 3744–3750.
[7] J. Salmi, A. Richter and V. Koivunen, *Sequential unfolding SVD for tensors with applications in array signal processing*, IEEE Trans. Signal Processing, 57(2009), pp. 4719–4733.
[8] 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., 36(2015), pp. 496–522.
[9] 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., 36(2015), pp. 1015-1045.
[10] M. Sørensen and L. De Lathauwer, *Fiber sampling approach to canonical polyadic decomposition and tensor completion*, Internal Report 15-151, ESAT-STADIUS, KU Leuven (Leuven, Belgium), 2017.