SIAM J. MATRIX ANAL. APPL.
Vol. 40, No. 3, pp. 888–917
© 2019 Society for Industrial and Applied Mathematics

FIBER SAMPLING APPROACH TO CANONICAL POLYADIC DECOMPOSITION AND APPLICATION TO 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 difficult to analyze and process using conventional tensor techniques. In this paper we focus on the basic canonical polyadic decomposition (CPD). We propose an algebraic framework for finding the CPD of tensors that have missing fibers. This includes extensions of multilinear algebraic as well as generic uniqueness conditions originally developed for the CPD of fully observed tensors. Computationally, we reduce the CPD of a tensor with missing fibers to relatively simple matrix completion problems via a matrix eigenvalue decomposition (EVD). Under the given conditions, the EVD-based algorithm is guaranteed to return the exact CPD. The derivation establishes connections with so-called coupled CPDs, an emerging concept that has proven to be of great interest in a range of array processing and signal processing applications. It will become clear that relatively few fibers are needed in order to compute the CPD. This makes fiber sampling interesting for large scale tensor decompositions. Numerical experiments show that the algebraic framework may significantly speed up more common optimization-based computation schemes for the estimation of the CPD of incomplete noisy data tensors.

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

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

DOI. 10.1137/17M1140790

1. Introduction.

Tensor methods have found many applications, including in psychometrics, chemometrics, signal processing, machine learning, and data compression; see [9, 35] 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., [26, 47, 1, 53, 35]). 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 [47], in electroencephalogram (EEG) processing based on time-frequency-measurement analysis where some of the electrodes are malfunctioning [1], and in nuclear magnetic resonance (NMR) spectroscopy due to nonuniform sampling [32, 49, 57, 31]. A more recent application of fiber sampled CPD is tensor-based

*Received by the editors July 26, 2017; accepted for publication (in revised form) May 21, 2019; published electronically July 30, 2019.
https://doi.org/10.1137/17M1140790
Funding: The work of the authors was supported by the Research Council KU Leuven C1 project C16/15/059-nD, by the FWO grant G.0F67.18N (EOS SeLMA), and by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC advanced grant BIOTENSORS grant 339804. This paper reflects only the authors' view and the Union is not liable for any use that may be made of the contained information.

†Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904 (ms8tz@virginia.edu). At the time of the research, Mikael Sørensen was with the Department of Electrical Engineering (ESAT), KU Leuven, Kortrijk, 8500, Belgium.
‡KU Leuven - E.E. Dept. (ESAT) - STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium, and the Group Science, Engineering and Technology, KU Leuven-Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium (Lieven.DeLathauwer@kuleuven.be).

graph clustering in which higher-order network structures are exploited [50]. 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 [38] 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., [20]). The difference is that the latter problem is concerned with finding the missing entries of the incomplete tensor and not necessarily with finding the factors of an underlying CPD. Another notable difference is that many of the existing matrix/tensor completion methods (e.g., [7, 33, 34, 20, 56]) 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 [36, 37] interpretation that reduces the problem to a set of relatively simple matrix completion problems. The approach leads to a new easy-to-check and yet powerful uniqueness condition.

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 the missing fibers. As mentioned in [9, 35], the CPD has found many applications, ranging from signal processing and statistics to chemometrics and psychometrics. The presented framework for CPD of tensors that have missing fibers is relevant for many of these applications. We will in particular consider the tensor completion problem where we want to find the CPD of a large scale tensor that has low rank and where only a subset of its fibers is observed. Numerical experiments demonstrate that the proposed EVD-based method can provide an efficient initialization for more conventional optimization-based methods for incomplete tensor decompositions, which is of practical interest.

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. 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. In section 5 we explain that fiber sampling can be used to compute the CPD of a large-scale tensor via the coupled CPD of a set of smaller-scaled tensors. Numerical experiments are reported in section 6, where it is demonstrated that the subsampling 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, determinant, rank, range, and kernel of a matrix $\mathbf{A}$ are denoted by $\mathbf{a}_r$, $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^{-1}$, $\mathbf{A}^\dagger$, $\|\mathbf{A}\|_F$, $|\mathbf{A}|$, $r_{\mathbf{A}}$, $\operatorname{range}(\mathbf{A})$, and $\ker(\mathbf{A})$, respectively.

Let $\mathbf{A} \in \mathbb{C}^{m\times p}$ and $\mathbf{B} \in \mathbb{C}^{n\times q}$; then the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$ is defined as

$$\mathbf{A} \otimes \mathbf{B} := \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1p}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mp}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{mn\times pq},$$

in which $(\mathbf{A})_{mn} = a_{mn}$. Similarly, let $\mathbf{A} \in \mathbb{C}^{m\times p}$ and $\mathbf{B} \in \mathbb{C}^{n\times p}$; then the Khatri–Rao product of $\mathbf{A}$ and $\mathbf{B}$ is defined as

$$\mathbf{A} \odot \mathbf{B} := [\mathbf{a}_1 \otimes \mathbf{b}_1 \ \cdots \ \mathbf{a}_p \otimes \mathbf{b}_p\,] \in \mathbb{C}^{mn\times p}.$$

Let $C_R^2 = \frac{R(R-1)}{2}$; then the second compound matrix of $\mathbf{A} \in \mathbb{C}^{I\times R}$ is denoted by $C_2(\mathbf{A}) \in \mathbb{C}^{C_I^2 \times C_R^2}$. It is the matrix containing the determinants of all 2×2 submatrices of $\mathbf{A}$; see [25] for further details. As an example, if $\mathbf{A} \in \mathbb{C}^{3\times 3}$, then

$$C_2(\mathbf{A}) = \begin{bmatrix} \left|\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right| & \left|\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}\right| & \left|\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix}\right| \\ \left|\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix}\right| & \left|\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}\right| & \left|\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix}\right| \\ \left|\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}\right| & \left|\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}\right| & \left|\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right| \end{bmatrix} \in \mathbb{C}^{3\times 3}.$$

The symbol $*$ denotes the Hadamard product, e.g., in the case of third-order tensors, we have that

$$(\mathcal{A} * \mathcal{B})_{ijk} = a_{ijk} b_{ijk}.$$

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 nonzero entries of a vector $\mathbf{x}$ is denoted by $\omega(\mathbf{x})$ in the tensor decomposition literature, dating back to the work of Kruskal [28]. Let $\mathrm{diag}(\mathbf{a}) \in \mathbb{C}^{J\times J}$ denote a diagonal matrix that holds a column vector $\mathbf{a} \in \mathbb{C}^{J\times 1}$ or a row vector $\mathbf{a} \in \mathbb{C}^{1\times J}$ on its diagonal. In some cases a diagonal matrix is holding row k of $\mathbf{A} \in \mathbb{C}^{I\times J}$ on its diagonal. This will be denoted by $D_k(\mathbf{A}) \in \mathbb{C}^{J\times J}$. Let $\mathbf{I}_N \in \mathbb{C}^{N\times N}$ 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 $\mathbf{A}$. Submatrices constructed from two row vectors of a matrix will play an important role in this paper. The submatrix consisting of the i_1th and i_2th rows of $\mathbf{A} \in \mathbb{C}^{I\times R}$ is denoted by $\mathbf{A}^{(i_1,i_2)} := \mathbf{A}([i_1, i_2], :) \in \mathbb{C}^{2\times R}$. Vectors extracted from a tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ will also play an important role. In particular, the mode-3 fiber $\mathbf{x}_{ij\bullet} \in \mathbb{C}^K$, defined by $(\mathbf{x}_{ij\bullet})_k = x_{ijk}$, will occur frequently throughout the paper.

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 nonzero 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 [24, 6]:

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

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 $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R] \in \mathbb{C}^{I\times R}$, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R] \in \mathbb{C}^{J\times R}$, and $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{K\times R}$. The 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 [12, 13, 17, 35] and references therein.

2.1.3. Basic CPD: All factor matrices have full column rank. In many practical problems, all three factor matrices of the CPD can be assumed to have full column rank. This will be referred to as the "basic CPD" problem. It is well known that the CPD in (2.1) with full column rank $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ is unique and can be computed via an EVD involving only two matrix slices of $\mathcal{X}$ (e.g., [24, 29]). Let us briefly elaborate on the latter property since it will play an important role in this paper. In detail, from the matrix factorization $\mathbf{X}^{(i\cdot\cdot)} = \mathbf{B}D_i(\mathbf{A})\mathbf{C}^T$ we obtain the standard (generalized) EVD relation

$$\mathbf{X}^{(i_1\cdot\cdot)}\cdot\mathbf{F}\cdot D_{i_2}(\mathbf{A}) = \mathbf{X}^{(i_2\cdot\cdot)}\cdot\mathbf{F}\cdot D_{i_1}(\mathbf{A}), \quad 1\le i_1\neq i_2\le I, \tag{2.3}$$

where $\mathbf{F} = (\mathbf{C}^T)^\dagger$. (Here we assume for simplicity that the generalized eigenvalues are distinct; see [29, 46] for the extension to the case of repeated generalized eigenvalues where $a_{i_1r}a_{i_2s} = a_{i_1s}a_{i_2r}$ for some $r\neq s$.) In our setting, it will become clear that the recovery of $\mathbf{F}$ via the EVD (2.3) will be sufficient to guarantee the uniqueness and the recovery of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. It is not difficult to see that once $\mathbf{F}$ is known, then $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ follow. In section 4.3.1 we provide details on how this can be implemented.

2.1.4. More general CPD: Only one factor matrix is required to have full column rank. For the tensor with missing fibers problem that will be discussed in this paper, it will become clear that the assumption that all factor matrices have full column rank will turn out to be too restrictive. We will now consider the more relaxed case where only one factor matrix of the CPD is required to have full column rank. For this case the following necessary and sufficient uniqueness condition was obtained in [27] and later reformulated in terms of compound matrices in [12]. It makes use of the vector

$$\mathbf{f}(\mathbf{d}) = [d_1d_2, d_1d_3, \ldots, d_{R-1}d_R]^T \in \mathbb{C}^{C_R^2}. \tag{2.4}$$

Note that $\mathbf{f}(\mathbf{d})$ consists of all distinct entries $d_r\cdot d_s$ of $\mathbf{d}\otimes\mathbf{d}$ with $r<s$.

THEOREM 2.1. *Consider the PD of $\mathcal{X}\in\mathbb{C}^{I\times J\times K}$ in (2.1). Assume that $\mathbf{C}$ has full column rank. The rank of $\mathcal{X}$ is R and the CPD of $\mathcal{X}$ is unique if and only if the following implication holds:*

$$(C_2(\mathbf{A})\odot C_2(\mathbf{B}))\mathbf{f}(\mathbf{d}) = \mathbf{0} \Rightarrow \omega(\mathbf{d})\le 1 \tag{2.5}$$

for all structured vectors $\mathbf{f}(\mathbf{d})$ of the form (2.4). Generically,[1] the implication holds if and only if $R\le(I-1)(J-1)$ [45, 8, 15].

[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 (2.5) can be hard to check. Observe that if $C_2(\mathbf{A}) \odot C_2(\mathbf{B})$ in (2.5) has full column rank, then $\mathbf{f}(\mathbf{d}) = \mathbf{0}$ 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*

$$\begin{cases} \mathbf{C} \text{ has full column rank,} \\ C_2(\mathbf{A}) \odot C_2(\mathbf{B}) \text{ has full column rank,} \end{cases} \tag{2.6}$$

then the rank of $\mathcal{X}$ *is* R *and the CPD of* $\mathcal{X}$ *is unique* [27, 10, 12]. *Generically, condition* (2.6) *is satisfied if* $C_R^2 \leq C_I^2 C_J^2$ *and* $R \leq K$ [10, 43].

The matrix $C_2(\mathbf{A}) \odot C_2(\mathbf{B})$ has full column rank if and only if the $(C_R^2 \times C_R^2)$ matrix $(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))^H (C_2(\mathbf{A}) \odot C_2(\mathbf{B})) = C_2(\mathbf{A}^H\mathbf{A}) * C_2(\mathbf{B}^H\mathbf{B})$ is nonsingular. Consequently, checking condition (2.6) essentially amounts to checking if the latter $(C_R^2 \times C_R^2)$ matrix is nonsingular.

In contrast to the "basic CPD," the CPD with only a single full column rank factor cannot directly be computed via a matrix EVD. Fortunately, under condition (2.6) the "more general CPD" of $\mathcal{X}$ can be converted into a "basic CPD" of an $(R \times R \times R)$ tensor $\mathcal{M}$ of rank R, even in cases where $\max(I, J) < R$ [10, 14]. As already mentioned in section 2.1.3, the latter CPD can be computed by means of a standard EVD. In section 4.2 it will be explained how to construct the tensor $\mathcal{M}$ from $\mathcal{X}$ and how to retrieve the CPD factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ of $\mathcal{X}$ from the CPD of $\mathcal{M}$.

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 [36]

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r\,, \quad n \in \{1, \ldots, N\}, \tag{2.7}$$

with factor matrices

$$\begin{aligned} \mathbf{A}^{(n)} &= [\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}] \in \mathbb{C}^{I_n \times R}, \\ \mathbf{B}^{(n)} &= [\mathbf{b}_1^{(n)}, \ldots, \mathbf{b}_R^{(n)}] \in \mathbb{C}^{J_n \times R}, \\ \mathbf{C} &= [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{K \times R}. \end{aligned}$$

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 (2.7) will be called the coupled CPD of $\{\mathcal{X}^{(n)}\}$. We mention that the more recent coupled CPD framework has already found interesting applications in engineering and science, such as in multidimensional harmonic retrieval [40, 41] and in sensor array processing [42].

2.2.1. Matrix representation. The coupled PD or CPD of $\{\mathcal{X}^{(n)}\}$ given by (2.7) has the following matrix representation:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{(1)} \odot \mathbf{B}^{(1)} \\ \vdots \\ \mathbf{A}^{(N)} \odot \mathbf{B}^{(N)} \end{bmatrix} \mathbf{C}^T \in \mathbb{C}^{(\sum_{n=1}^{N} I_n J_n) \times K}. \tag{2.8}$$

2.2.2. Uniqueness. The coupled rank-1 tensors in (2.7) 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 [36]. Theorem 2.3 extends Theorem 2.2 to coupled CPD. It makes use of the matrix

$$\mathbf{G}^{(N)} = \begin{bmatrix} C_2(\mathbf{A}^{(1)}) \odot C_2(\mathbf{B}^{(1)}) \\ \vdots \\ C_2(\mathbf{A}^{(N)}) \odot C_2(\mathbf{B}^{(N)}) \end{bmatrix} \in \mathbb{C}^{\left(\sum_{n=1}^{N} C_{I_n}^2 C_{J_n}^2\right) \times C_R^2}. \tag{2.9}$$

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* (2.7). *If*

$$\begin{cases} \mathbf{C} \text{ in (2.8) has full column rank,} \\ \mathbf{G}^{(N)} \text{ in (2.9) has full column rank,} \end{cases} \tag{2.10}$$

then the coupled rank of $\{\mathcal{X}^{(n)}\}$ *is* R *and the coupled CPD of* $\{\mathcal{X}^{(n)}\}$ *is unique* [36]. *Generically, condition* (2.10) *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 (2.10) is satisfied, then the coupled CPD of $\{\mathcal{X}^{(n)}\}$ can be converted into the CPD of an $(R \times R \times R)$ tensor $\mathcal{M}$ of rank R [36], which in turn can be computed by means of a standard EVD. Again, more details will be provided in section 4.2.

The *objective of this paper* is to extend the results discussed in this section to the case of tensors that are not fully observed. More precisely, in section 3 we extend the sufficient uniqueness conditions stated in Theorems 2.1–2.3 to the case of tensors that have missing fibers. Likewise, in section 4 we extend the algebraic algorithms associated with Theorems 2.2 and 2.3 to the case of missing fibers. In section 5 we discuss fiber subsampling variants that are of interest for large scale tensor decompositions.

3. Uniqueness of CPD of tensor with missing fibers. 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.

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 mode-3 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 mode-3 vector $\mathbf{x}_{ij\bullet} \in \mathbb{C}^K$ of $\mathcal{X}$, defined by $(\mathbf{x}_{ij\bullet})_k = x_{ijk}$, is unobserved. (We note in passing that we could just as well have considered missing mode-1 or mode-2 fibers of $\mathcal{X}$.) 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 $\mathcal{Y}$.

3.2. Necessary conditions for the uniqueness of the CPD. Recall that a necessary condition for uniqueness of the CPD factor matrix $\mathbf{C}$ in (2.2) is that $\mathbf{A} \odot \mathbf{B}$ has full column rank (e.g., [44]). Here we extend the condition to the case of missing entries. From (2.2), we know that (3.1) has the 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

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

where $\widetilde{\mathbf{C}} = [\tilde{\mathbf{c}}_1, \ldots, \tilde{\mathbf{c}}_k] = \mathbf{C}^T$. From (3.3) it is clear that $\mathbf{C}$ can only be unique if the matrices $\{\operatorname{diag}(\mathbf{w}_k)\,(\mathbf{A} \odot \mathbf{B})\}$ have full column rank. Indeed, if $\operatorname{diag}(\mathbf{w}_k)\,(\mathbf{A} \odot \mathbf{B})$ does not have full column rank, then for any $\mathbf{x} \in \ker(\operatorname{diag}(\mathbf{w}_k)\,(\mathbf{A} \odot \mathbf{B}))$ we have $\operatorname{diag}(\mathbf{w}_k)\,(\mathbf{A} \odot \mathbf{B})\,\tilde{\mathbf{c}}_k = \operatorname{diag}(\mathbf{w}_k)\,(\mathbf{A} \odot \mathbf{B})\,(\tilde{\mathbf{c}}_k + \mathbf{x})$. 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). A necessary condition for uniqueness of the factor matrix $\mathbf{C}$ is that $\operatorname{diag}(\mathbf{w}_k)\,(\mathbf{A} \odot \mathbf{B})$ 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 $\operatorname{diag}(\mathbf{w}_k)$ is the same for every column $\mathbf{x}_k$ of $\mathbf{X}$, i.e., $\operatorname{diag}(\mathbf{w}_1) = \operatorname{diag}(\mathbf{w}_k)\ \forall k \in \{1, \ldots, K\}$. 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 $\mathbf{S}_{\mathbf{W}_1}^T \mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \operatorname{diag}(\mathbf{w}_1)\,(\mathbf{A} \odot \mathbf{B})$.

Besides the uniqueness of the CPD factor $\mathbf{C}$ of $\mathbf{Y}_{\text{sub}}$, we also consider whether there is more than one choice of matrices $\mathbf{A}$ and $\mathbf{B}$ to obtain $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$. Let $\boldsymbol{\varphi}_r \in \mathbb{C}^{IJ}$ denote the rth column of $\operatorname{diag}(\mathbf{w}_1)\,(\mathbf{A} \odot \mathbf{B})$. 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 have at least one determinate entry. More formally, for $\mathbf{A}$ and $\mathbf{B}$ in $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ to be unique, it is necessary that $\mathbf{P}^{(r)}$ has the property

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

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.

3.3. Sufficient condition for 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. Theorem 3.2 below extends a well-known sufficient uniqueness condition [27] 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 $\mathbf{Y}_{\text{sub}}$. Theorem 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., $\mathbf{D}_{\text{sel}} = \text{diag}\,(\mathbf{d}_{\text{sel}})$. The entries of the vector

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

are given by

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

THEOREM 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 $\mathbf{C}$ has full column rank. Let $\mathbf{D}_{sel} = diag\,(\mathbf{d}_{sel})$ in which $\mathbf{d}_{sel}$ is defined as in* (3.7). *The rank of $\mathcal{X}$ is R and the factor matrix $\mathbf{C}$ is unique if the following implication holds:*

$$\mathbf{D}_{sel}(C_2\,(\mathbf{A}) \odot C_2\,(\mathbf{B}))\mathbf{f}(\mathbf{d}) = \mathbf{0} \Rightarrow \omega(\mathbf{d}) \leq 1 \tag{3.8}$$

for all structured vectors $\mathbf{f}(\mathbf{d})$ of the form (2.4). *Generically, the rank of $\mathcal{X}$ is R and the factor matrix $\mathbf{C}$ is unique if*

$$R \leq F - (I+J) \quad \textit{and} \quad R \leq K, \tag{3.9}$$

where $F = \omega(\mathbf{w}_1)$ denotes the number of observed fibers.

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

$$\mathbf{Y} = \mathbf{\Phi}(\mathbf{A},\mathbf{B})\mathbf{C}^T = \mathbf{\Phi}(\widehat{\mathbf{A}},\widehat{\mathbf{B}})\widehat{\mathbf{C}}^T. \tag{3.10}$$

Using Kruskal's permutation lemma [28] (see also [44]), 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 $\mathbf{C} = \widehat{\mathbf{C}}\mathbf{\Pi}\Delta_{\mathbf{C}}$, where $\mathbf{\Pi}$ 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 $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ 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 $\omega(\widehat{\mathbf{d}}) = 1$.

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

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 $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ 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 $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \mathbf{0}$; then it will be shown that the latter property implies that $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))\mathbf{f}(\mathbf{d}) = \mathbf{0}$, which contradicts condition (3.8). We conclude that if condition (3.8) is satisfied, then $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \mathbf{0} \Rightarrow \omega(\mathbf{d}) \leq 1$, which in turn implies that $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ has full column rank. Hence, we know from (3.10) that

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

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 (3.10) we obtain

$$\mathbf{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{C}^T\mathbf{z} = \mathbf{\Phi}(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})\widehat{\mathbf{C}}^T\mathbf{z} \Leftrightarrow \mathbf{\Phi}(\mathbf{A}, \mathbf{B})\mathbf{d} = \mathbf{\Phi}(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})\widehat{\mathbf{d}}, \tag{3.11}$$

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 (3.11) into an $(I \times J)$ matrix yields

$$\mathbf{W}^{(\cdot\bullet)} * \left(\sum_{r=1}^{R} \mathbf{a}_r \mathbf{b}_r^T d_r \right) = \mathbf{W}^{(\cdot\bullet)} * \left(\sum_{r=1}^{R} \widehat{\mathbf{a}}_r \widehat{\mathbf{b}}_r^T \widehat{d}_r \right), \tag{3.12}$$

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

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

Since $\omega(\widehat{\mathbf{d}}) = 1$, all observed (2×2) submatrices of (3.12) must be rank deficient. In other words, the second-order minors of all observed (2×2) submatrices of (3.12) must vanish. The second-order minors of (3.12) are of the form

$$\begin{vmatrix} \sum_{r=1}^{R} w_{i_1 j_1}^{(\cdot\bullet)} a_{i_1 r} b_{j_1 r} \cdot d_r & \sum_{r=1}^{R} w_{i_1 j_2}^{(\cdot\bullet)} a_{i_1 r} b_{j_2 r} \cdot d_r \\ \sum_{r=1}^{R} w_{i_2 j_1}^{(\cdot\bullet)} a_{i_2 r} b_{j_1 r} \cdot d_r & \sum_{r=1}^{R} w_{i_2 j_2}^{(\cdot\bullet)} a_{i_2 r} b_{j_2 r} \cdot d_r \end{vmatrix}. \tag{3.13}$$

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 $\text{card}(\Gamma) = C_I^2 C_J^2$. 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_1th and the i_2th 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 Theorem 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 (3.12) must satisfy the relation

$$\left| \sum_{r=1}^{R} \mathbf{a}_r^{(i_1,i_2)} \mathbf{b}_r^{(j_1,j_2)T} d_r \right| = 0. \tag{3.14}$$

Stacking the equations (3.14) as rows of the matrix $\mathbf{G} \in \mathbb{C}^{C_I^2 C_J^2 \times C_R^2}$ in accordance with the lexicographical rule yields

$$\mathbf{G}\mathbf{f}(\mathbf{d}) = \mathbf{0}, \tag{3.15}$$

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

$$\begin{aligned} \mathbf{G} &= \begin{bmatrix} w_{(1,2),(1,2)} \cdot C_2\left(\mathbf{A}^{(1,2)}\right) * C_2\left(\mathbf{B}^{(1,2)}\right) \\ w_{(1,2),(1,3)} \cdot C_2\left(\mathbf{A}^{(1,2)}\right) * C_2\left(\mathbf{B}^{(1,3)}\right) \\ \vdots \\ w_{(I-1,I),(J-1,J)} \cdot C_2\left(\mathbf{A}^{(I-1,I)}\right) * C_2\left(\mathbf{B}^{(J-1,J)}\right) \end{bmatrix} \\ &= \mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})), \end{aligned} \tag{3.16}$$

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 (3.14)–(3.15) 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 this would imply that $\mathbf{G}\mathbf{f}(\mathbf{d}) = \mathbf{0}$ 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.

Generic sufficiency of condition (3.8). For generic $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ we can resort to an algebraic geometry–based tool for checking generic uniqueness of structured matrix factorizations of the form $\mathbf{Y} = \mathbf{M}\mathbf{C}^T$, in which the entries of the matrix $\mathbf{M}$ can be parameterized by rational functions [16]. More precisely, in our setting where $\mathbf{C}$ generically has full column rank, the decomposition of $\mathbf{Y} = \operatorname{diag}(\mathbf{w}_1)(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T$ with $\mathbf{M} = \operatorname{diag}(\mathbf{w}_1)(\mathbf{A} \odot \mathbf{B})$ is generically unique if the number of parameterized rank-1 terms is bounded by $R \leq \widehat{N} - \widehat{l} - 1$ [16, Theorem 1], where $\widehat{l}$ is an upper bound on the number of variables needed to parameterize the vector $\operatorname{diag}(\mathbf{w}_1)(\mathbf{a} \otimes \mathbf{b})$, and $\widehat{N}$ is a lower bound on the dimension of the vector space spanned by the vectors in the set

$$\{\operatorname{diag}(\mathbf{w}_1)(\mathbf{a} \otimes \mathbf{b}) \,|\, \mathbf{a} \in \mathbb{C}^I, \mathbf{b} \in \mathbb{C}^J\}. \tag{3.17}$$

Clearly, $\widehat{l} = I + J - 1$ is a valid upper bound; i.e., $\widehat{l}$ can be taken equal to the number of free variables in $\mathbf{a} \otimes \mathbf{b}$. It is well known that the matrix $\mathbf{A} \odot \mathbf{B}$ with columns $\mathbf{a}_r = [1 \; x_r \; \ldots \; x_r^{I-1}]^T$ and $\mathbf{b}_r = [1 \; x_r^I \; \ldots \; x_r^{(J-1)I}]^T$ is a Vandermonde matrix with full rank if $x_r \neq x_s \; \forall r \neq s$. This implies that the vectors in the set $\{\mathbf{a} \otimes \mathbf{b} \,|\, \mathbf{a} \in \mathbb{C}^I, \mathbf{b} \in \mathbb{C}^J\}$ span the entire IJ-dimensional space. Consequently, we also know that the vectors in the set (3.17) also span the entire $\omega(\mathbf{w}_1)$-dimensional space, i.e., $\widehat{N} = \omega(\mathbf{w}_1)$, which in turn leads to the inequality condition (3.9). ∎

A problem with Theorem 3.2 is that it may be difficult to check condition (3.8), which requires us to verify that all vectors in the kernel of $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ are structured vectors of the form $\mathbf{f}(\mathbf{d})$ with property $\omega(\mathbf{d}) \leq 1$. For generic factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, condition (3.8) generically holds if the condition (3.9) is satisfied. Note that the generic condition (3.9) is easy to check but that it is not necessary. For example, if $I = J = K = R = 2$ and all fibers have been sampled ($\omega(\mathbf{w}_1) = 4$), then condition (3.9) is not satisfied, despite the fact that the CPD is obviously unique (e.g., condition (2.6) in Theorem 2.2 is satisfied). Note also that 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 Theorem 3.3 that can also be used in the case of deterministic (nongeneric) factor matrices. The proof of Theorem 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 $\mathbf{X}^{(n)} = (\mathbf{A}^{(n)} \odot \mathbf{B}^{(n)})\mathbf{C}^T$.

THEOREM 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 $\mathbf{D}_{sel} = diag(\mathbf{d}_{sel})$ in which $\mathbf{d}_{sel}$ is defined as in* (3.7). *If*

$$\begin{cases} \mathbf{C} \text{ has full column rank,} \\ \mathbf{D}_{sel}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})) \text{ has full column rank,} \end{cases} \tag{3.18}$$

then the rank of $\mathcal{X}$ is R and the factor matrix $\mathbf{C}$ is unique. Generically, condition (3.18) *is satisfied if $C_R^2 \leq \omega(\mathbf{d}_{sel})$ 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

$$\mathbf{Z}_{(u,v),(p,q)} = \mathbf{S}_{(u,v),(p,q)}\mathbf{Y}_{(u,v),(p,q)} = \mathbf{S}_{(u,v),(p,q)}(\mathbf{A}^{(u,v)} \odot \mathbf{B}^{(p,q)})\mathbf{C}^T, \tag{3.19}$$

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:

$$\begin{bmatrix} \mathbf{Z}_{(1,2),(1,2)} \\ \mathbf{Z}_{(1,2),(1,3)} \\ \vdots \\ \mathbf{Z}_{(I-1,I),(J-1,J)} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{(1,2),(1,2)}(\mathbf{A}^{(1,2)} \odot \mathbf{B}^{(1,2)}) \\ \mathbf{S}_{(1,2),(1,3)}(\mathbf{A}^{(1,2)} \odot \mathbf{B}^{(1,3)}) \\ \vdots \\ \mathbf{S}_{(I-1,I),(J-1,J)}(\mathbf{A}^{(I-1,I)} \odot \mathbf{B}^{(J-1,J)}) \end{bmatrix} \mathbf{C}^T. \tag{3.20}$$

The matrix (2.9) associated with $\{\mathcal{Z}_{(u,v),(p,q)}\}$ takes the form (3.16). Due to Theorem 2.3, we can conclude from (3.16) that if condition (3.18) is satisfied, then the rank of $\mathcal{X}$ is R and the factor matrix $\mathbf{C}$ is unique.

We now explain that condition (3.18) 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 $C_I^2 C_J^2 \geq C_R^2$ [43] (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 nonzeroed rows of $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ still exceeds the column-dimension. ∎

As in Theorem 2.2, checking if the $(C_I^2 C_J^2 \times C_R^2)$ matrix $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ has full column rank is equivalent to checking if the smaller $(C_R^2 \times C_R^2)$ matrix

$$(\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})))^H (\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})))$$
$$= \sum_{1 \leq i_1 < i_2 \leq I} \sum_{1 \leq j_1 < j_2 \leq J} w_{(i_1,i_2),(j_1,j_2)} C_2\left(\mathbf{A}^{(i_1,i_2)H} \mathbf{A}^{(i_1,i_2)}\right) * C_2\left(\mathbf{B}^{(j_1,j_2)H} \mathbf{B}^{(j_1,j_2)}\right)$$

is nonsingular, where we recall that $w_{(i_1,i_2),(j_1,j_2)}$ corresponds to a diagonal entry of $\mathbf{D}_{\text{sel}}$, defined according to (3.7).

3.4. Sufficient uniqueness condition for overall CPD. Note that if the full column rank matrix $\mathbf{C}$ is unique, then $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{Y}_{\text{sub}}(\mathbf{C}^T)^\dagger$ 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 [22] where a necessary and sufficient condition for the existence and the uniqueness of a rank-1 completion of an incomplete matrix was obtained. The basic idea in [22] is to think of an incomplete matrix, say $\mathbf{P}^{(r)}$ defined by (3.5), as holding the weights associated with the edges of a weighted bipartite graph. In our setting we already assume that there exists a rank-1 completion of $\mathbf{P}^{(r)}$, i.e., the existence question has already been answered. For the uniqueness result it suffices to associate the incomplete matrix with an unweighted bipartite graph, denoted by $P^{(r)}$. The two groups of vertices associated with $P^{(r)}$ are the row indices $1, \ldots, I$ and the column indices $1, \ldots, J$. Let $\mathcal{P}^{(r)}$ denote the edge set associated with the bipartite graph $P^{(r)}$. For our purpose, the weight of the edge associated with $(i, j) \in \mathcal{P}^{(r)}$ is

given by $p_{ij}^{(r)}$, as defined in (3.5). We adapt the result from [22] to our setting in Lemma 3.4 below. Lemma 3.4 considers the restriction of $\mathcal{P}^{(r)}$, the support of which is $\widetilde{\mathcal{P}}^{(r)} = \{(i,j) \in \mathcal{P}^{(r)} \mid p_{i,j}^{(r)} \neq 0\}$.

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

$$\widetilde{P}^{(r)} \text{ is a connected graph.} \tag{3.21}$$

Property (3.6) ensures that every row and every column of $\mathbf{P}^{(r)}$ has been sampled at least once while the restriction to $\widetilde{\mathcal{P}}^{(r)}$ ensures that the sampled entries are nonzero. Since $\mathbf{C}$ has full column rank, Theorems 3.2 and 3.3 and Proposition 5.1 ensure the uniqueness of $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$, which in turn ensures the existence of a rank-1 completion of $\mathbf{P}^{(r)}$. Hence, if condition (3.21) 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 (3.21) can easily be checked via the properties of the incidence or adjacency matrix of the restricted bipartite graph $\widetilde{P}^{(r)}$ (e.g., [3]).

To illustrate the importance of condition (3.21), we consider the tensor $\mathcal{Y} = \mathcal{W} * \mathcal{X} \in \mathbb{C}^{5\times 6\times K}$ with frontal slices of the form depicted in Figure 3.1(right) in which the observed fibers of $\mathcal{X}$ correspond to the following two subtensors of $\mathcal{X}$:

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

The fibers associated with $\mathcal{X}^{(1)}$ 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 $\mathcal{X}$ need 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 tensors $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$ are "connected." Not including $\mathbf{x}_{12\bullet}$ would lead to a nonunique CPD of $\mathcal{X}$. The reason is that the bipartite graph $\widetilde{P}^{(r)}$ 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 $\mathbf{x}_{12\bullet}$, the combination of condition (3.9) and Lemma 3.4 generically guarantees the uniqueness of the CPD of $\mathcal{X}$ if $R \leq 5$ and $R \leq K$, despite the fact that only 16 out of 30 fibers have been sampled.

Formally, the combination of Theorem 3.2 and Lemma 3.4 yields the following overall uniqueness condition, which is an extension of the sufficient condition stated in 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* $\mathbf{D}_{sel} = diag(\mathbf{d}_{sel})$ *in which* $\mathbf{d}_{sel}$ *is defined as in* (3.7)*, and let* $\widetilde{P}^{(r)}$ *be the restricted bipartite graph of* $\mathbf{P}^{(r)}$ *defined by* (3.5) *and with property* (3.6)*. Assume that* $\mathbf{C}$ *has full column rank. The rank of* $\mathcal{X}$ *is* R *and the CPD of* $\mathcal{X}$ *is unique if*

$$\begin{cases} D_{sel}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))\mathbf{f}(\mathbf{d}) = \mathbf{0} \Rightarrow \omega(\mathbf{d}) \leq 1 \quad \forall \mathbf{d} \in \mathbb{C}^R, & \text{(3.22a)} \\ \widetilde{P}^{(r)} \text{ is a connected graph for every } r \in \{1,\ldots,R\}, & \text{(3.22b)} \end{cases}$$

where $\mathbf{f}(\mathbf{d})$ *is a structured vector of the form* (2.4)*.*

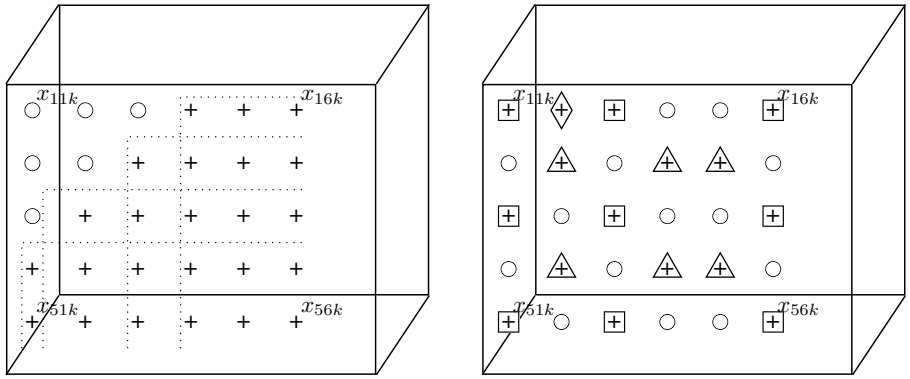


FIG. 3.1. *Tensor with intrinsically missing fibers (left) and a sparsely fiber sampled tensor (right), where + represents an observed fiber and ◯ represents an unobserved fiber. The two subtensors $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$ in the tensor $\mathcal{X}$ are denoted by □ and △. The ◇ fiber connects the two grids and ensures that the overall CPD is unique.*

Even though Theorem 3.5 provides a sufficient uniqueness condition for the cases where $\mathbf{C}$ has full column rank, it may be hard to check whether the implication in (3.22a) holds. In contrast, the combination of Theorem 3.3 and Lemma 3.4 leads to the following sufficient overall uniqueness condition, which just requires checking whether $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ in (3.22a) has full column rank.

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 $\mathbf{D}_{sel} = diag(\mathbf{d}_{sel})$ in which $\mathbf{d}_{sel}$ is defined as in* (3.7), *and let $\widetilde{P}^{(r)}$ be the restricted bipartite graph of $\mathbf{P}^{(r)}$ defined by* (3.5) *and with property* (3.6). *If*

$$\begin{cases} \mathbf{C} \text{ in (2.1) has full column rank,} \\ \mathbf{D}_{sel}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})) \text{ has full column rank,} \\ \widetilde{P}^{(r)} \text{ is a connected graph for every } r \in \{1,\dots,R\}, \end{cases} \tag{3.23}$$

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

Theorem 3.6 can be interpreted as a generalization of Theorem 2.2 to the case of missing fibers.

Let us end this section by comparing the uniqueness conditions stated in Theorems 3.5 and 3.6 on an example. Consider the tensor $\mathcal{Y} = \mathcal{W} * \mathcal{X} \in \mathbb{C}^{5\times 6\times K}$ with frontal slices of the form in Figure 3.1(left) where 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. (The dotted lines in the figure will first be used in section 5.2.) Since the tensor is incomplete, CPD-specific uniqueness conditions and algorithms do not apply (e.g., Theorem 2.2). Conditions (3.22a)–(3.22b) in Theorem 3.5 generically guarantee the uniqueness of the CPD of $\mathcal{X}$ despite the missing fibers if $R \le 13$ and $R \le K$. In more detail, the inequalities in (3.9) ensure that condition (3.22a) is generically satisfied, which in turn ensures the uniqueness of $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ and $\mathbf{C}$. Lemma 3.4 states that if the graph connectivity condition (3.22b) in Theorem 3.5 is satisfied, then $\mathbf{A}$ and $\mathbf{B}$ are also unique. By similar reasoning Theorem 3.6 generically guarantees the uniqueness of the CPD of $\mathcal{X}$ if $R \le 11$ and $R \le K$.

4. An algorithm for CPD of tensor with missing fibers. In [10] an algorithm was proposed to compute the "more general CPD" discussed in section 2.1.4,

where only one factor matrix is required to have full column rank. It was further elaborated on in [14]. In [37] it was extended to coupled decompositions. In this section we will extend the algorithm to the case of missing fibers. The presented algebraic algorithm will rely on Theorem 3.6. Before discussing it in detail, we provide a roadmap and explain the overall idea.

Step 1: By capitalizing on the low-rank structure of the $(IJ \times K)$ matrix $\mathbf{Y}_{\text{sub}} = \mathbf{S}_{\mathbf{W}_1}\mathbf{X} = \mathbf{\Phi}(\mathbf{A},\mathbf{B})\mathbf{C}^T$, we will in section 4.1 first compress it to a smaller $(IJ \times R)$ matrix $\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$ whose columns form a basis for range $(\mathbf{\Phi}(\mathbf{A},\mathbf{B}))$. The matrix $\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$ can be found via the SVD of $\mathbf{Y}_{\text{sub}}$.

Step 2: Next, by capitalizing on the Khatri–Rao structure of $\mathbf{A} \odot \mathbf{B}$, we will find $\mathbf{\Phi}(\mathbf{A},\mathbf{B}) = \mathbf{S}_{\mathbf{W}_1}(\mathbf{A}\odot\mathbf{B})$ from range $(\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}})$. More precisely, in section 4.2 we will explain that the knowledge of range $(\mathbf{\Phi}(\mathbf{A},\mathbf{B}))$ enables us to transform the original CPD problem with missing fibers, with possibly $I < R$ and/or $J < R$, into the rank-R CPD of an $(R \times R \times R)$ tensor

$$\mathcal{M} = \sum_{r=1}^{R} \mathbf{g}_r \circ \mathbf{g}_r \circ \mathbf{n}_r \in \mathbb{C}^{R\times R\times R}, \tag{4.1}$$

in which the factor matrix $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_R]$ has the property $\mathbf{\Phi}(\mathbf{A},\mathbf{B}) = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{G}$. In other words, $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ and $\mathbf{C} = (\mathbf{\Phi}(\mathbf{A},\mathbf{B})^\dagger\mathbf{Y}_{\text{sub}})^T$ can be obtained once the CPD of $\mathcal{M}$ has been found. To compute the basic CPD (4.1) in the exact case, only two matrix slices of $\mathcal{M}$ are needed. 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 (generalized) EVD relation

$$\mathbf{M}^{(\cdot\cdot k_1)} \cdot \mathbf{F} \cdot D_{k_2}(\mathbf{N}) = \mathbf{M}^{(\cdot\cdot k_2)} \cdot \mathbf{F} \cdot D_{k_1}(\mathbf{N}), \quad 1 \le k_1 \ne k_2 \le R, \tag{4.2}$$

where $\mathbf{F} = \mathbf{G}^{-T}$. In the inexact case, the result obtained from the EVD (4.2) can be used to initialize an optimization-based method that computes a refined estimate of $\mathbf{G}$ via the CPD of $\mathcal{M}$. The optimization-based method uses all R matrix slices $\mathbf{M}^{(\cdot\cdot 1)}, \ldots, \mathbf{M}^{(\cdot\cdot R)}$, and not just two matrix slices $\mathbf{M}^{(\cdot\cdot k_1)}$ and $\mathbf{M}^{(\cdot\cdot k_2)}$.

Step 3: Finally, in section 4.3 we will explain that from $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ we can determine $\mathbf{A}$ and $\mathbf{B}$ via a set of R decoupled rank-1 matrix completion problems.

As a final note, we mention that in the context of inexact tensor decompositions, it is common practice to use an exact decomposition method for the initialization of an optimization-based method. Thus, in our setting, the CPD factors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ obtained via the above three steps can be used to initialize an optimization-based method that further refines the estimates by fitting the factors to the observed tensor $\mathcal{Y}$ given by (3.1).

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. This further implies that a basis for $\mathbf{\Phi}(\mathbf{A},\mathbf{B})$ can easily be found via, for instance, an SVD. 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

$$\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}} = \mathbf{\Phi}(\mathbf{A},\mathbf{B}) \cdot \mathbf{F}^T \quad \text{and} \quad \mathbf{V}_{\text{sub}}^*\mathbf{F}^{-1} = \mathbf{C}. \tag{4.3}$$

The left equation is of the same form as (3.4), but it only involves $R \leq K$ columns. From (4.3) it is clear that once the unknown matrix $\mathbf{F}$ will have been found, $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ and $\mathbf{C}$ will immediately follow. In the next step we explain how to obtain $\mathbf{F}$ from $\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$. As a matter of fact, $\mathbf{F}^{-T}$ will assume the role of $\mathbf{G}$ in the roadmap (step 2).

4.2. Step 2: From CPD with missing fibers, possibly also with $I < R$ and/or $J < R$, to basic CPD. In section 4.2.1 we review the procedure presented in [10, 14] to transform a CPD, possibly with $I < R$ and/or $J < R$, into a basic CPD that involves only nonsingular factor matrices. Next, in section 4.2.2, we explain how to extend it to the case of missing fibers.

4.2.1. Algebraic EVD method for CPD of fully observed tensor. Consider the rank-R tensor

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{f}_r \in \mathbb{C}^{I \times J \times R}, \tag{4.4}$$

where for simplicity we have set $K = R$ (i.e., we assume that a compression has been carried out if K is larger than R). Let $\mathbf{X} = (\mathbf{A} \odot \mathbf{B})\mathbf{F}^T$ denote its matrix representation. The overall idea is to look for rank-1 structured vectors in range $(\mathbf{X}) =$ range $(\mathbf{A} \odot \mathbf{B})$. Concretely, we look for vectors $\mathbf{g}_1, \ldots, \mathbf{g}_R$ with property $\mathbf{X}\mathbf{g}_r = \mathbf{a}_r \otimes \mathbf{b}_r$. Thanks to the structure of $\mathbf{X}$, the matrix $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_R] = \mathbf{F}^{-T}$ can be found from range $(\mathbf{X})$ if condition (2.10) in Theorem 2.3 is satisfied, as explained in [10, 14] and reviewed in this section.

Consider first the rank-1 tensor $\mathcal{X}^{(R=1)} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{f} \in \mathbb{C}^{I \times J \times R}$. The basic idea in [10, 14] is to exploit the rank-1 property of $\mathcal{X}^{(R=1)}$; i.e., we have that

$$\begin{vmatrix} x^{(R=1)}_{i_1,j_1,k} & x^{(R=1)}_{i_1,j_2,k} \\ x^{(R=1)}_{i_2,j_1,k} & x^{(R=1)}_{i_2,j_2,k} \end{vmatrix} = a_{i_1}a_{i_2}b_{j_1}b_{j_2}f_kf_k - a_{i_1}a_{i_2}b_{j_1}b_{j_2}f_kf_k = 0, \tag{4.5}$$

where $1 \leq j_1 < j_2 \leq J$, $1 \leq i_1 < i_2 \leq I$, and $1 \leq k \leq K$. Relation (4.5) just states that all 2×2 minors of the rank-1 matrix $f_k\mathbf{a}\mathbf{b}^T$ vanish.

Let us now consider the rank-R tensor $\mathcal{X}$ given by (4.4). Consider also the expression

$$x_{i_1j_1k_1}x_{i_2j_2k_2} + x_{i_1j_1k_2}x_{i_2j_2k_1} - x_{i_1j_2k_1}x_{i_2j_1k_2} - x_{i_1j_2k_2}x_{i_2j_1r_1}, \tag{4.6}$$

where $1 \leq i_1 < i_2 \leq I$, $1 \leq j_1 < j_2 \leq J$, and $1 \leq k_1, k_2 \leq R$. Let us stack the set of expressions (4.6) into the matrix $\mathbf{R}_2(\mathcal{X}) \in \mathbb{C}^{C_I^2 C_J^2 \times R^2}$, defined as follows. Let $\beta_{k_1,k_2} = (k_2 - 1)R + k_1$ and $\alpha_{i_1,i_2,j_1,j_2} = \frac{(j_1(2j_2-j_1-1)-2)I(I-1)}{4} + \frac{i_1(2i_2-i_1-1)}{2}$ so that the $(\alpha_{i_1,i_2,j_1,j_2}, \beta_{k_1,k_2})$th entry of $\mathbf{R}_2(\mathcal{X})$ is equal to (4.6). Furthermore, let the columns of $\mathbf{R}_2(\mathcal{X}) \in \mathbb{C}^{C_I^2 C_J^2 \times R^2}$ be indexed lexicographically by the pair (k_1, k_2) so that the column associated with the pair (k_1, k_2) precedes the column associated with the pair (k_1', k_2') if and only if either $k_1' > k_1$ or $k_1' = k_1$ and $k_2' > k_2$. The column of $\mathbf{R}_2(\mathcal{X}) \in \mathbb{C}^{C_I^2 C_J^2 \times R^2}$ given by the pair (k_1, k_2) is equal to

$$\text{Vec}\left(C_2(\mathcal{X}(:,:,k_1) + \mathcal{X}(:,:,k_2)) - C_2(\mathcal{X}(:,:,k_1)) - C_2(\mathcal{X}(:,:,k_2))\right). \tag{4.7}$$

From (4.7) it can be verified that $\mathbf{R}_2(\mathcal{X})$ has C_{R+1}^2 distinct columns and that its rows correspond to vectorized symmetric matrices. From (4.6) one can also obtain that [14]

$$\mathbf{R}_2(\mathcal{X}) = (C_2(\mathbf{A}) \odot C_2(\mathbf{B}))\mathbf{\Psi}_2(\mathbf{F})^T, \tag{4.8}$$

where the columns of $\mathbf{\Psi}_2(\mathbf{F}) \in \mathbb{C}^{R^2 \times C_R^2}$ are also lexicographically indexed by the pair (k_1, k_2) so that the column of $\mathbf{\Psi}_2(\mathbf{F})$ associated with the pair (k_1, k_2) is given by $\frac{1}{2}(\mathbf{f}_{k_1} \otimes \mathbf{f}_{k_2} + \mathbf{f}_{k_2} \otimes \mathbf{f}_{k_1})$, $1 \leq k_1 < k_2 \leq R$. Taking into account that all (2×2) minors of a rank-1 matrix vanish (cf. (4.5)), a technical derivation allows us to verify that $\mathbf{g}_1 \otimes \mathbf{g}_1, \ldots, \mathbf{g}_R \otimes \mathbf{g}_R \in \ker(\mathbf{R}_2(\mathcal{X}))$. Furthermore, if condition (2.6) in Theorem 2.2 is satisfied, then it can be shown that the linearly independent vectors $\mathbf{g}_1 \otimes \mathbf{g}_1, \ldots, \mathbf{g}_R \otimes \mathbf{g}_R$ are the *only* rank-1 structured vectors in $\ker(\mathbf{R}_2(\mathcal{X}))$. Consequently, the subspace

$$W := \ker(\mathbf{R}_2(\mathcal{X})) \cap \text{range}(\pi_S) = \ker(\mathbf{\Psi}_2(\mathbf{F})^T) \cap \text{range}(\pi_S) = \text{range}(\mathbf{G} \odot \mathbf{G}) \tag{4.9}$$

is R-dimensional, where $\text{range}(\pi_S)$ denotes the subspace of vectorized symmetric matrices. The symmetry restriction in (4.9) on the vectors in $\ker(\mathbf{R}_2(\mathcal{X}))$ is obtained by the symmetrization mapping $\pi_S : \mathbb{C}^{R^2} \to \mathbb{C}^{R^2}$, defined as

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

Let the columns of $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ constitute a basis for W; then we can conclude from (4.9) that there exists a nonsingular matrix $\mathbf{N} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{M} = (\mathbf{G} \odot \mathbf{G})\mathbf{N}^T, \tag{4.10}$$

where $\mathbf{G} = \mathbf{F}^{-T}$. Clearly, (4.10) corresponds to a matrix representation of the basic CPD of $\mathcal{M}$ in (4.1), which in the exact case can be computed via an EVD, as mentioned in section 2.1.3. In other words, once $\mathbf{M}$ in (4.10) is obtained from, for instance, the SVD of $\mathbf{R}_2(\mathcal{X})$, $\mathbf{G}$ follows. In the next section we will extend the above procedure to the case of missing fibers and provide details on how to compute $\mathbf{M}$ from $\mathbf{R}_2(\mathcal{X})$.

4.2.2. Extension to the case of missing fibers. Observe that each row of $\mathbf{R}_2(\mathcal{X})$ in (4.8) depends only on two rows of $\mathbf{A}$ and $\mathbf{B}$. This is the key property that allows us to extend the approach in section 4.2.1 to the case of missing fibers. More precisely, each of the $C_I^2 C_J^2$ rows of $\mathbf{R}_2(\mathcal{X})$ is constructed from a distinct $(2 \times 2 \times R)$ subtensor of $\mathcal{X}$. In the missing fibers case we will just work with the $\omega(\mathbf{d}_{\text{sel}})$ $(2 \times 2 \times R)$ subtensors that are observed. Briefly, 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 $\mathbf{R}_2(\mathcal{X})$ that are observed. The missing fibers version of (4.8) is then

$$\mathbf{R}_{\text{red}} := \mathbf{S}_{\text{red}}\mathbf{R}_2(\mathcal{X}) = \mathbf{S}_{\text{red}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B})) \cdot \mathbf{\Psi}_2(\mathbf{F})^T. \tag{4.11}$$

Since we assume that condition (3.23) is satisfied, $\mathbf{S}_{\text{red}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ has full column rank. Consequently, in the case of missing fibers, we can just obtain a basis for the subspace W by working with $\mathbf{R}_{\text{red}}$ in (4.11) instead of $\mathbf{R}_2(\mathcal{X})$ in (4.9). Since the rows of the matrix $\mathbf{R}_{\text{red}}$ are vectorized symmetric matrices, we have that $\text{range}\left(\mathbf{R}_{\text{red}}^T\right) \subseteq \text{range}(\pi_S)$. 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:

$$\mathbf{Q}_{\text{red}} = [\mathbf{Q}_{(1)} \ , \ 2 \cdot \mathbf{Q}_{(2)}], \tag{4.12}$$

where

$$\mathbf{Q}_{(1)} = [\mathbf{R}_{\text{red}}(:, f(1,1)), \ldots, \mathbf{R}_{\text{red}}(:, f(R,R))], \tag{4.13}$$

$$\mathbf{Q}_{(2)} = [\mathbf{R}_{\text{red}}(:, f(1,2)), \mathbf{R}_{\text{red}}(:, f(1,3)), \ldots, \mathbf{R}_{\text{red}}(:, f(R-1,R))]. \tag{4.14}$$

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 eigenvectors associated with the R smallest eigenvalues of the matrix $\mathbf{Q}_{\text{red}}^H \cdot \mathbf{Q}_{\text{red}}$. Let the columns of $\mathbf{M}$ constitute a basis for $\ker(\mathbf{Q}_{\text{red}})$; then it admits the factorization (4.10), which clearly 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 $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_R] = \mathbf{F}^{-T}$. As mentioned earlier, the latter CPD can be computed via an EVD, implying that $\mathbf{F}$ can be recovered by means of standard linear algebra methods (e.g., [21]).

4.3. Step 3: From basic CPD to factor matrices A, B, C. After $\mathbf{F}^{-1}$ has been found we obtain $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$ and $\mathbf{C}$ via (4.3). What remains is to compute $\mathbf{A}$ and $\mathbf{B}$. We will now explain how to extract the remaining unknown factors $\mathbf{A}$ and $\mathbf{B}$ from $\mathbf{\Phi}(\mathbf{A}, \mathbf{B})$. In order to make clear how this step is carried out, let us first review how it can be done in the case where no fibers are missing.

4.3.1. Rank-1 factorization method for finding factor matrices A and B. Recall first that if all fibers of $\mathcal{X}$ have been sampled, then $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{A} \odot \mathbf{B}$ and relation (4.3) simplifies to

$$\mathbf{P} := \mathbf{U}_{\text{sub}} \mathbf{\Sigma}_{\text{sub}} \mathbf{G}^T = \mathbf{A} \odot \mathbf{B} \text{ and } \mathbf{V}_{\text{sub}}^* \mathbf{G} = \mathbf{C}, \tag{4.15}$$

where $\mathbf{G} = \mathbf{F}^{-1}$ has been obtained via relation (4.10). Let $\mathbf{P}^{(r)} \in \mathbb{C}^{I \times J}$ denote the reshaped version of the rth column of $\mathbf{P}$ so that

$$\mathbf{P}^{(r)} = \mathbf{a}_r \mathbf{b}_r^T, \quad r \in \{1, \ldots, R\}. \tag{4.16}$$

Note that $\mathbf{P}^{(r)}$ is equal to the matrix given by (3.5) with all entries of the matrix observed. From (4.16) it is clear that the pair $\{\mathbf{a}_r, \mathbf{b}_r\}$ can be obtained via the rank-1 factorization of $\mathbf{P}^{(r)}$. In the inexact case, this is achieved via the rank-1 matrix approximation

$$\min_{\|\mathbf{a}_r\|_F = 1, \mathbf{b}} \|\mathbf{P}^{(r)} - \mathbf{a}_r \mathbf{b}_r^T\|_F^2, \quad r \in \{1, \ldots, R\}. \tag{4.17}$$

In the next section it will become clear that in the case of missing entries, the best rank-1 matrix approximation (4.17) cannot directly be used. We will now discuss an alternative approach that can easily be extended to the case of missing entries. From (4.16) it is clear that the columns of $\mathbf{P}^{(r)}$ are all proportional to $\mathbf{a}_r$, i.e.,

$$\mathbf{a}_r \in R^{(r)} := \bigcap_{j=1}^{J} \text{range}\left(\mathbf{p}_j^{(r)}\right), \quad r \in \{1, \ldots, R\}. \tag{4.18}$$

Let the columns of $\mathbf{N}_j^{(r)} \in \mathbb{C}^{I \times (I-1)}$ form a basis for the orthogonal complement of $\mathbf{p}_j^{(r)}$; then

$$\text{range}\,(\mathbf{a}_r) \in \bigcap_{j=1}^{J} \text{range}\left(\mathbf{p}_j^{(r)}\right) \Leftrightarrow \mathbf{N}_j^{(r)H} \mathbf{a}_r = \mathbf{0}\, \forall j \in \{1, \ldots, J\}. \tag{4.19}$$

It is now clear that the vector $\mathbf{a}_r$ in (4.16) can also be obtained by solving the set of homogeneous linear equations in (4.19), i.e.,

$$\mathbf{a}_r^H[\mathbf{N}_1^{(r)},\ldots,\mathbf{N}_J^{(r)}]=\mathbf{0},\quad r\in\{1,\ldots,R\}.$$

In the inexact case, $\mathbf{a}_r$ can be obtained from the SVD of the $((I-1)J\times I)$ matrix $[\mathbf{N}_1^{(r)T},\ldots,\mathbf{N}_J^{(r)T}]^T$. Alternatively, $\mathbf{a}_r$ can found by solving the eigenvalue problem

$$\min_{\|\mathbf{a}_r\|_F=1}\sum_{j=1}^{J}\|\mathbf{N}_j^{(r)H}\mathbf{a}_r\|_F^2=\min_{\|\mathbf{a}_r\|_F=1}\mathbf{a}_r^H\left(\sum_{j=1}^{J}\mathbf{N}_j^{(r)}\mathbf{N}_j^{(r)H}\right)\mathbf{a}_r,\ r\in\{1,\ldots,R\},\tag{4.20}$$

which only involves the smaller $(I\times I)$ matrix $\sum_{j=1}^{J}\mathbf{N}_j^{(r)}\mathbf{N}_j^{(r)H}$. Once $\mathbf{a}_r$ has been obtained from (4.20), $\mathbf{b}_r$ follows immediately. In the next section we will explain that this approach is also appropriate for the case of missing entries.

4.3.2. Extension to the case of missing fibers. 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. More precisely, will now explain how to find the subspace $R^{(r)}$ in (4.18) using the incomplete matrix $\mathbf{P}^{(r)}$. The derivation will make use of several variables which are listed in Table 4.1.

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. 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. 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

$$\mathbf{P}_j^{(r)}=[\mathbf{z}_j^{(r)},\mathbf{e}_{\upsilon(1)}^{(I)},\ldots,\mathbf{e}_{\upsilon(I_{j,r}^c)}^{(I)}]\in\mathbb{C}^{I\times(1+I_{j,r}^c)}$$

constitute a basis for range $\left(\mathbf{P}_j^{(r)}\right)$. Since $\mathbf{a}_r\in$ range $\left(\mathbf{P}_j^{(r)}\right)\ \forall j\in\Theta^{(r)}$, we have that

$$\mathbf{a}_r\in S^{(r)}:=\bigcap_{j\in\Theta^{(r)}}\text{range}\left(\mathbf{P}_j^{(r)}\right).\tag{4.21}$$

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

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 $S^{(r)}=\text{range}\,(\mathbf{a}_r)=R^{(r)}$. 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 $z_{ij}^{(r)}\neq0$ 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 so that $\mathbf{a}_r$ can be found.

TABLE 4.1
Overview of the variables used in section 4.3.2.

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 nonzero 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 $\mathbf{p}_j^{(r)}$.
$I_{j,r}^c$	Number of indeterminate entries of $\mathbf{p}_j^{(r)}$. $(I_{j,r}^c = I - I_{j,r})$.
$\upsilon(i)$	Index of the ith indeterminate entry of $\mathbf{p}_j^{(r)}$, i.e., entry $p_{\upsilon(i),j}^{(r)}$ is unknown for every $i \in \{1, \ldots, I_{j,r}^c\}$.
$\mu(i)$	Index of the ith determinate entry of $\mathbf{p}_j^{(r)}$, i.e., $p_{\mu(i),j}^{(r)} = a_{\mu(i),r} b_{j,r}$ $\forall i \in \{1, \ldots, I_{j,r}\}$.
$\mathbf{z}_j^{(r)}$	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.

Recovery of $\mathbf{a}_r$. Similarly to (4.19), we know from (4.21) that

$$\mathbf{a}_r \in S^{(r)} \Leftrightarrow \mathbf{N}_j^{(r)H} \mathbf{a}_r = \mathbf{0} \quad \forall j \in \Theta^{(r)}, \tag{4.22}$$

where the columns of $\mathbf{N}_j^{(r)} \in \mathbb{C}^{I \times (I-1-I_{j,r}^c)}$ form a basis for the orthogonal complement of $\mathbf{P}_j^{(r)}$. The vector $\mathbf{a}_r$ can now be obtained from the SVD of the $((\sum_{j \in \Theta^{(r)}} (I - 1 - I_{j,r}^c) \times I)$ matrix $[\mathbf{N}_1^{(r)T}, \ldots, \mathbf{N}_J^{(r)T}]^T$. Note that for every $j \in \Theta^{(r)}$, the columns of

$$\mathbf{Z}_j^{(r)} = [\mathbf{z}_j^{(r)} / \|\mathbf{z}_j^{(r)}\|_F, \mathbf{e}_{\upsilon(1)}^{(I)}, \ldots, \mathbf{e}_{\upsilon(I_{j,r}^c)}^{(I)}] \in \mathbb{C}^{I \times (1 + I_{j,r}^c)}$$

constitute an orthonormal basis for range $\left(\mathbf{P}_j^{(r)}\right)$. This property together with the equivalence relation in (4.22) now implies that (e.g., [51])

$$\arg \min_{\|\mathbf{a}_r\|_F = 1} \sum_{j \in \Theta^{(r)}} \|\mathbf{N}_j^{(r)H} \mathbf{a}_r\|_F^2 = \arg \max_{\|\mathbf{a}_r\|_F = 1} \|\mathbf{a}_r^H \mathbf{Z}_{\text{total}}^{(r)}\|_F^2, \quad r \in \{1, \ldots, R\}, \tag{4.23}$$

where $\mathbf{Z}_{\text{total}}^{(r)} = [\mathbf{Z}_1^{(r)}, \ldots, \mathbf{Z}_J^{(r)}] \in \mathbb{C}^{I \times \sum_{j \in \Theta^{(r)}} (1 + I_{j,r}^c)}$. Consequently, the dominant left singular vector of $\mathbf{Z}_{\text{total}}^{(r)}$ is a basis vector for $S^{(r)}$. Thus, denoting the dominant left singular vector of $\mathbf{Z}_{\text{total}}^{(r)}$ by $\mathbf{u}^{(r)}$, we can also set $\mathbf{a}_r = \mathbf{u}^{(r)}$. Furthermore, if $I < \sum_{j \in \Theta^{(r)}} (1 + I_{j,r}^c)$, then, since range $\left(\mathbf{Z}_{\text{total}}^{(r)}\right) = \text{range}\left(\mathbf{Z}_{\text{total}}^{(r)} \cdot \mathbf{Z}_{\text{total}}^{(r)H}\right)$, $\mathbf{u}^{(r)}$ can also be found as the dominant eigenvector of the smaller matrix $\mathbf{Z}_{\text{total}}^{(r)} \cdot \mathbf{Z}_{\text{total}}^{(r)H} = \sum_{j \in \Theta^{(r)}} \mathbf{Z}_j^{(r)} \cdot \mathbf{Z}_j^{(r)H} \in \mathbb{C}^{I \times I}$.

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 nonzero entries of $\mathbf{b}_r$ follow 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. Note also that in the exact case, the rank R can be determined from the SVD of $\mathbf{Y}_{\text{sub}}$; i.e., R is equal to the rank of $\mathbf{Y}_{\text{sub}}$.

The complexity of solving the CPD 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 [23] or column/row/fiber subset selection [18] factorization techniques. The complexity of steps 2 and 3 can be reduced by considering only part of the structure as in subsampling; cf. section 5.2. As mentioned in section 2.1.3, the CPD in step 4 can be computed via an EVD; i.e., only two matrix slices $\mathbf{M}^{(\cdot\cdot k_1)}$ and $\mathbf{M}^{(\cdot\cdot k_2)}$ of the tensor $\mathcal{M}$ with matrix factorization $\mathbf{M} = (\mathbf{G} \odot \mathbf{G})\mathbf{N}^T$ are needed. Likewise, as mentioned in section 4.3.2, the rank-1 matrix completion problem in step 8 can be reduced to an eigenvalue problem (e.g., [21]).

In the inexact case, we can refine the EVD solution obtained in step 4 by an optimization-based method for CPD (e.g., [55]) that takes all the R matrix slices $\{\mathbf{M}^{(\cdot\cdot r)}\}$ of the tensor $\mathcal{M}$ into account. Finally, in the inexact case, the output $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ of Algorithm 1 can also be refined by an optimization-based method (e.g., [55]) that fits them to the tensor $\mathcal{Y}$.

Algorithm 1. Algebraic method for CPD with missing fibers.

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

1. Compute SVD $\mathbf{Y}_{\text{sub}} = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{V}_{\text{sub}}^H$.
2. Build $\mathbf{Q}_{\text{red}}$ in (4.12) from $\mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$; i.e., set $\mathbf{R}_{\text{red}} = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}$ in (4.13)–(4.14).
3. Collect in a matrix $\mathbf{M}$ the R right singular vectors associated with the R smallest singular values of $\mathbf{Q}_{\text{red}}$.
4. Solve CPD problem $\mathbf{M} = (\mathbf{G} \odot \mathbf{G})\mathbf{N}^T$ via EVD.
5. Compute $\mathbf{C} = \mathbf{V}^*\mathbf{G}^T$ and $\mathbf{\Phi}(\mathbf{A}, \mathbf{B}) = \mathbf{U}_{\text{sub}}\mathbf{\Sigma}_{\text{sub}}\mathbf{G}^{-1}$.
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 matrix completion of $\mathbf{P}^{(r)}$, $r \in \{1, \ldots, R\}$.

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

A difference between Algorithm 1 and the more conventional optimization-based nonlinear least squares (NLS) methods (e.g., [1, 30, 35, 53, 54]) is that the former is guaranteed to find the factor matrices in the exact case, while the latter can better take additive noise perturbation terms into account by fitting the data to the CPD model in a least squares sense. Another difference is the computational complexity of the two approaches. In Algorithm 1 the complexity is dominated by the construction of $\mathbf{Q}_{\text{red}}$, which is of the order $\mathcal{O}(\omega(\mathbf{d}_{\text{sel}})R^2)$, and the computation of the CPD of $\mathbf{M}$ via an EVD, which is of the order $\mathcal{O}(R^3)$. In large scale cases, NLS methods for incomplete CPD typically have a complexity of the order $\mathcal{O}(it_{CG} \cdot N_{ke} \cdot R)$ per iteration, in which N_{ke} denotes the number of nonzero entries in $\mathcal{Y}$ and it_{CG} is usually a small number [54]. For small scale problems where $I = J = K$, NLS methods for incomplete CPD typically have a complexity of the order $\mathcal{O}(R^2N_{ke} + I^3R^3)$ per iteration [54]. Discussions about the complexity of optimization-based methods for the computation of the CPD of an incomplete tensor can be found in [30, 54].

5. Fiber sampling and large scale tensor decompositions. In sections 3 and 4 we presented uniqueness conditions and an algorithm for the CPD of a tensor with missing fibers. In this section we explain that the results can be used to find the CPD of very large tensors (which can in principle be fully given). In short, we will show that fiber sampling enables us to turn a large scale tensor decomposition problem into a small scale tensor decomposition problem. An obvious application is

CPD-based tensor completion.

5.1. Motivation. A common way to deal with large scale data sets in numerical linear algebra is to employ randomized algorithms that sample a small subset of the rows/columns of the given matrix (e.g., [19]). The reason why such a randomized subset selection approach works is that a small sample of rows/columns contains a lot of information about the row/column space of the given matrix. Likewise, in large scale tensor decomposition problems it can be expensive to process the full tensor. Fortunately, if (i) $F \geq I + J + R$ fibers of length $K \geq R$ have been (randomly) sampled so that condition (3.9) is satisfied and (ii) the fibers have been sampled in such a way that the graph connectivity condition (3.21) in Lemma 3.4 is satisfied, then the CPD of $\mathcal{X}$ can generically be determined via its fiber sampled version $\mathcal{Y}$. Thus, only a fraction of the fibers are needed in order to find the CPD of $\mathcal{X}$. More concretely, assume that F observed fibers of $\mathcal{X}$ in (3.1) are taken into account. These fibers can be stored in a matrix $\mathbf{H} \in \mathbb{C}^{F \times K}$ with factorization

$$\mathbf{H} = \begin{bmatrix} \mathbf{x}_{i_1 j_1 \bullet} \\ \vdots \\ \mathbf{x}_{i_F j_F \bullet} \end{bmatrix} = \boldsymbol{\phi}_{\mathbf{H}}(\mathbf{A}, \mathbf{B}) \mathbf{C}^T, \quad \boldsymbol{\phi}_{\mathbf{H}}(\mathbf{A}, \mathbf{B}) = \begin{bmatrix} \mathbf{a}_{i_1 \bullet} * \mathbf{b}_{j_1 \bullet} \\ \vdots \\ \mathbf{a}_{i_F \bullet} * \mathbf{b}_{j_F \bullet} \end{bmatrix}. \tag{5.1}$$

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. Since only $F \geq I + J - 1$ fibers need to be sampled in order to obtain a connected bipartite graph $P^{(r)}$, we can conclude from the preceding discussion that as few as $F \geq I + J + R$ mode-3 fibers can be sufficient to find the CPD of $\mathcal{X}$. 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 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.

Besides guaranteeing uniqueness, we want to sample fibers in such a way that the factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be found with relative ease. In a way, this is the opposite problem of the one we dealt with in sections 3 and 4. In sections 3 and 4, we extracted full subtensors $\mathcal{X}^{(n)}$ from a partially observed tensor $\mathcal{X}$. The coupled CPD of $\{\mathcal{X}^{(n)}\}$ provided insight into the uniqueness and led to an EVD-based algorithm. We considered subtensors of minimal size $2 \times 2 \times K$ in order to maximally exploit the available structure. To reduce the computational complexity, one could consider only a subset of the $2 \times 2 \times K$ subtensors of $\mathcal{X}$. As a matter of fact, one could sample only larger subtensors and ignore the fine structure that can only be captured by $2 \times 2 \times K$ subtensors. In the remainder of this section we briefly formalize the approach. Determination of the optimal subtensor size and "crosses" on which to sample is outside the scope of this paper. Finally, note that fiber sampling can be used both in cases where tensor $\mathcal{X}$ in principle could be fully observed and in cases where this is not possible. (As an example of the latter, in some applications parts of $\mathcal{X}$ do not make sense physically and hence need to be considered as intrinsically missing.)

5.2. Result for fiber sampling. Let $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$ denote a subtensor $\mathcal{X}$ that consists of all fibers $\mathbf{x}_{ij\bullet}$ on an $I_n \times J_n$ grid where

$$(i, j) \in \{i_{1,n}, \ldots, i_{I_n,n}\} \times \{j_{1,n}, \ldots, j_{J_n,n}\} \subset \{1, \ldots, I\} \times \{1, \ldots, J\}.$$

By sampling a set of N such subtensors $\mathcal{X}^{(n)} \in \mathbb{C}^{I_n \times J_n \times K}$ of $\mathcal{X}$ we obtain the coupled PD

$$\mathcal{X}^{(n)} = \sum_{r=1}^{R} \mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r \,, \quad n \in \{1, \ldots, N\}, \tag{5.2}$$

where $\mathbf{a}_r^{(n)} \in \mathbb{C}^{I_n}$ is a vector that consists of the entries $i_{1,n}, \ldots, i_{I_n,n}$ of $\mathbf{a}_r$ and $\mathbf{b}_r^{(n)} \in \mathbb{C}^{J_n}$ is a vector that consists of the entries $j_{1,n}, \ldots, j_{J_n,n}$ of $\mathbf{b}_r$, i.e.,

$$\mathbf{a}_r^{(n)} = [a_{i_{1,n}}, \ldots, a_{i_{I_n,n}}]^T \in \mathbb{C}^{I_n} \quad \text{and} \quad \mathbf{b}_r^{(n)} = [b_{j_{1,n}}, \ldots, b_{j_{J_n,n}}]^T \in \mathbb{C}^{J_n}.$$

For simplicity, the relations between the factors in $\{\mathbf{A}^{(n)}\}$ and in $\{\mathbf{B}^{(n)}\}$ are ignored. We have now turned a large-scale CPD into a smaller-scaled coupled CPD. Proposition 5.1 below is the fiber sampling counterpart of Theorem 3.6. The derivation of the uniqueness condition (5.3) in Proposition 5.1 is analogous to the derivation of the uniqueness condition (3.23) in Theorem 3.6.

PROPOSITION 5.1. *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). *Let $\mathbf{G}^{(N)}$ in* (2.9) *be built from the factor matrices $\mathbf{A}^{(n)} = [\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}] \in \mathbb{C}^{I_n \times R}$ and $\mathbf{B}^{(n)} = [\mathbf{b}_1^{(n)}, \ldots, \mathbf{b}_R^{(n)}] \in \mathbb{C}^{J_n \times R}$ in* (5.2)*, and let $\widetilde{P}^{(r)}$ be the restricted bipartite graph of $\mathbf{P}^{(r)}$ defined by* (3.5) *and with property* (3.6). *If*

$$\begin{cases} \mathbf{C} \text{ in (2.1) has full column rank,} \\ \mathbf{G}^{(N)} \text{ in (2.9) has full column rank,} \\ \widetilde{P}^{(r)} \text{ is a connected graph for every } r \in \{1, \ldots, R\}, \end{cases} \tag{5.3}$$

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

Note that the conditions ensuring uniqueness in Proposition 5.1 are more restrictive than in Theorem 3.6; i.e., condition (5.3) implies condition (3.23). This comes as no surprise as in Proposition 5.1 uniqueness needs to be inferred from only a part of $\mathcal{X}$. Indeed, the matrix obtained by removing any redundant rows of $\mathbf{G}^{(N)}$ in (5.3) corresponds to a submatrix of $\mathbf{D}_{\text{sel}}(C_2(\mathbf{A}) \odot C_2(\mathbf{B}))$ in (3.23). In other words, direct sampling of full subtensors $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(N)}$ of $\mathcal{X}$ is the same as only considering a subset of the observed $2 \times 2 \times K$ subtensors of $\mathcal{X}$.

As an example, consider again the tensor $\mathcal{Y} = \mathcal{W} * \mathcal{X} \in \mathbb{C}^{5 \times 6 \times K}$ with frontal slices of the form in Figure 3.1(left), where 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 intrinsically missing. Let us now sample the following four subtensors of $\mathcal{X}$:

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

in which the associated frontal slices are bordered by dotted lines in Figure 3.1(left). If $\mathbf{G}^{(N)}$ in (5.3) is constructed from the subtensors $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$, then Proposition 5.1 generically guarantees the uniqueness of the CPD of $\mathcal{X}$ if $R \leq 8$ and $R \leq K$. On the other hand, if $\mathbf{G}^{(N)}$ is constructed from the subtensors $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(4)}$, then Proposition 5.1 now generically guarantees the uniqueness of the CPD of $\mathcal{X}$ if $R \leq 9$ and $R \leq K$. Finally, if $\mathbf{G}^{(N)}$ is constructed from all four subtensors $\mathcal{X}^{(1)}$, $\mathcal{X}^{(2)}$, $\mathcal{X}^{(3)}$, and $\mathcal{X}^{(4)}$, then Proposition 5.1 generically guarantees the uniqueness of the CPD of $\mathcal{X}$ if $R \leq 11$ and $R \leq K$. The latter bound actually coincides with the upper bound on R ensured by Theorem 3.6. (The reason that Theorem 3.6 and Proposition 5.1 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.)

6. Numerical experiments. We demonstrate the behavior 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.

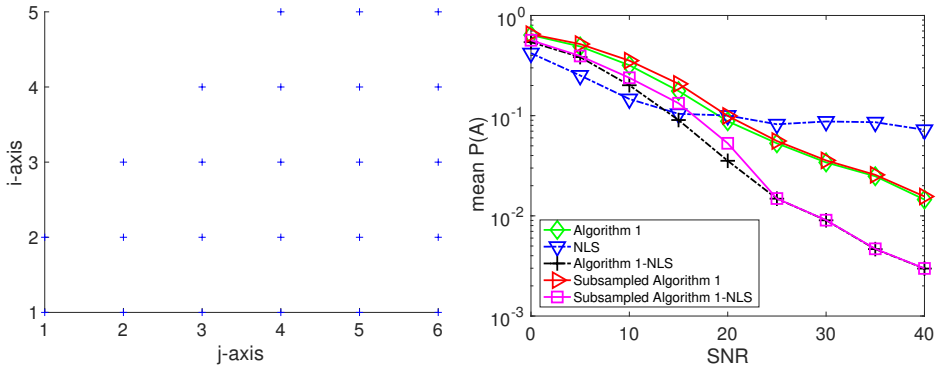
The following signal-to-noise ratio (SNR) measure will be used:

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

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

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

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

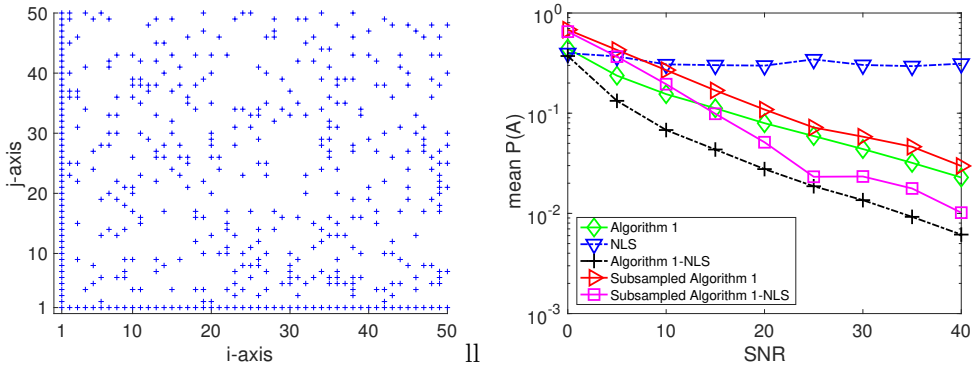


(a) Fiber sampling pattern. (b) Mean $\mathrm{P}(\mathbf{A})$.

FIG. 6.1. *Fiber sampling pattern (a blue cross + at position (i,j) indicates that fiber $\mathbf{x}_{ij\bullet}$ has been sampled) and mean $P(\mathbf{A})$ over 100 trials while SNR is varying from 10 to 40 dB, case 1. (Color available online.)*

6.1. Case 1: Systematically missing fibers. Consider the rank-R tensor $\mathcal{X} \in \mathbb{C}^{5\times 6\times K}$ 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, as depicted in Figure 6.1(left). We set $R = 5$ and $K = 40$. We will now investigate the performance of Algorithm 1 in the presence of perturbation noise. In Algorithm 1 the CPD in step 4 in will numerically be computed by the extended QZ method [52], while the rank-1 matrix completion problem in step 8 will numerically be solved via the built-in `svd.m` function in MATLAB. We compare (i) Algorithm 1, (ii) a subsampled version of Algorithm 1 in which only the four subtensors holding the fibers $\{\mathbf{x}_{ij\bullet}\}_{i\in\{1,2\}}^{j\in\{1,2,3,4,5,6\}}$, $\{\mathbf{x}_{ij\bullet}\}_{i\in\{1,2,3\}}^{j\in\{2,3,4,5,6\}}$, $\{\mathbf{x}_{ij\bullet}\}_{i\in\{1,2,3,4\}}^{j\in\{3,4,5,6\}}$, and $\{\mathbf{x}_{ij\bullet}\}_{i\in\{1,2,3,4,5\}}^{j\in\{4,5,6\}}$ of $\mathcal{X}$, as depicted in Figure 3.1(left), are used in the computation of $\mathbf{C}$, and (iii) the randomly initialized NLS method `sdf_nls.m` in [55]. The methods will be referred to as "Algorithm 1," "Subsampled Algorithm 1," and "NLS," respectively. In addition, we also consider the NLS method initialized by estimates obtained by Algorithm 1 and the Subsampled Algorithm 1. They will be referred to as "Algorithm 1-NLS"

and "Subsampled Algorithm 1-NLS." 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 6.1(right). We observe that the algebraic Algorithm 1 and Subsampled Algorithm 1 methods performed about the same. The reason is that the Subsampled Algorithm 1 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 with respect to its initialization and that initialization by an Algorithm 1 method leads to a considerable improvement. We can even observe that at high SNR the algebraic algorithms, which were derived for an exact decomposition, obtain better precision than the NLS algorithm that tries to minimize the difference between the tensor and its CPD approximation.

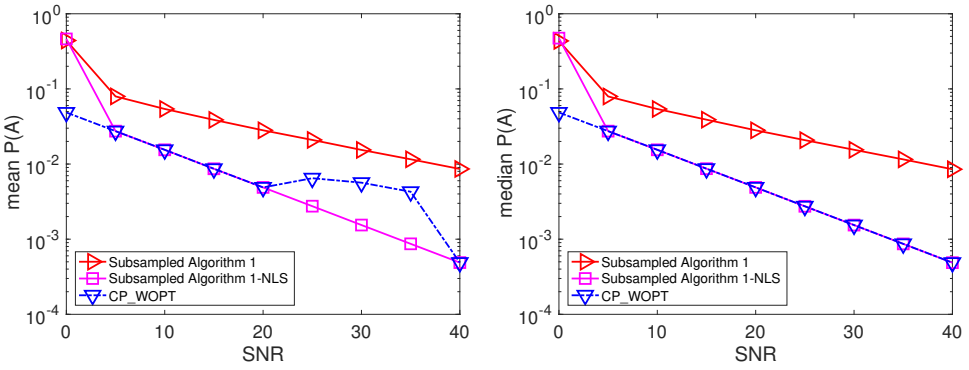


(a) Fiber sampling pattern. (b) Mean $\mathrm{P}(\mathbf{A})$.

FIG. 6.2. *Fiber sampling pattern (a blue cross + at position (i, j) indicates that fiber $\mathbf{x}_{ij\bullet}$ has been sampled) and mean $P(\mathbf{A})$ over 100 trials while SNR is varying from 10 to 40 dB, case 2. (Color available online.)*

6.2. Case 2: Small sized tensor with randomly missing fibers. Consider now a 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 $\mathcal{X}$. All the row border fibers $\mathbf{x}_{11\bullet}, \ldots, \mathbf{x}_{I1\bullet}$ and all the column border fibers $\mathbf{x}_{11\bullet}, \ldots, \mathbf{x}_{1J\bullet}$ are also sampled. This is just one way to ensure that the graph connectivity condition in Theorem 3.6 is satisfied. In total, 459 out of 2500 fibers are sampled. For the Subsampled Algorithm 1, $N = 10$ subtensors $\mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(N)}$ of size $(10 \times 10 \times K)$ are randomly selected from $\mathcal{Y}$.[3] This means that not more than 100 fibers are used in the computation of $\mathbf{C}$. In other words, the Subsampled Algorithm 1 solves the CPD recovery problem via a coupled CPD of $\mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(N)}$ that has 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 6.2. The algebraic Algorithm 1 performs better than the Subsampled Algorithm 1. The reason is the latter does not take all of the low-rank structure of $\mathcal{Y}$ into account. On the other hand, we observe that the loss of accuracy as a consequence of using fewer fibers in the Subsampled Algorithm 1 is very moderate. Furthermore, the Subsampled Algorithm 1 typically ran more than 50 times faster than Algorithm 1. Again, we observe that the optimization-based NLS method is sensitive with respect to its initialization and that both Algorithm 1 and Subsampled Algorithm 1 provided good starting values that significantly improved the performance.

[3] More strategic and data dependent sampling schemes could be considered (e.g., [18, 48, 23, 5, 4]), but that is outside the scope of the paper.

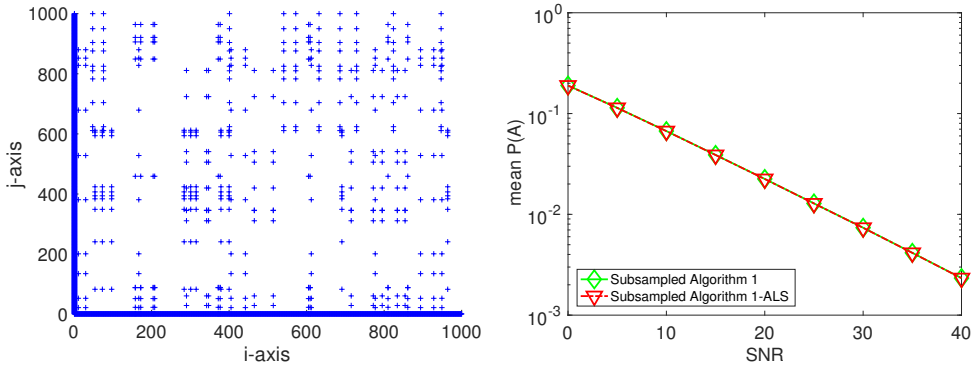


(a) Mean P (**A**). (b) Median P (**A**).

FIG. 6.3. *Mean* $P(\mathbf{A})$ *and median* $P(\mathbf{A})$ *over* 100 *trials while SNR is varying from* 10 *to* 40 *dB, case* 3.

6.3. Case 3: Medium sized tensor with randomly missing fibers. Consider a rank-10 tensor $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ with dimensions $I = J = K = 100$. We randomly sample about 46 percent of the fibers of $\mathcal{X}$. Again, all the row border fibers $\mathbf{x}_{11\bullet}, \ldots, \mathbf{x}_{I1\bullet}$ and all the column border fibers $\mathbf{x}_{11\bullet}, \ldots, \mathbf{x}_{1J\bullet}$ are sampled. In total, 4609 out of 10000 fibers are sampled. Using all 4609 fibers in the construction of $\mathbf{Q}_{\text{red}}$ in step 2 in Algorithm 1 is considered to be too costly. Consequently, we only consider the subsampled version of Algorithm 1 in which $N = 10$ subtensors $\mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(N)}$ of size $(10 \times 10 \times K)$ are randomly selected from $\mathcal{Y}$ and used in the construction of $\mathbf{Q}_{\text{red}}$ and computation of $\mathbf{C}$. (Note that all 4609 fibers are still used in the subsequent computation of $\mathbf{A}$ and $\mathbf{B}$.) We also consider the NLS method `cp_wopt.m` in [2] that uses the 4609 sampled fibers in the computation of all the CPD factor matrices. The method can be seen as a representation for the more commonly used optimization-based procedures for decomposing an incomplete tensor; see [1, 53, 54, 35] for references and details. It will be referred to as "CP_WOPT." The mean P (**A**) and median P (**A**) values over 100 trials as a function of SNR can be seen in Figure 6.3. Here the median P (**A**) value is less affected by outliers in which CP_WOPT did not perform well due to its dependence on the initialization. We also observe that the Subsampled Algorithm 1-NLS and CP_WOPT methods perform about the same and that the initialization by the algebraic Subsampled Algorithm 1 can provide a good starting value for an iterative optimization method. The Subsampled Algorithm 1 typically ran more than 7 times faster than Subsampled Algorithm 1-NLS, which in turn typically ran more than 5 times faster than CP_WOPT.

6.4. Case 4: Subsampling of large tensor. Consider a 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 $(1000 \times 1000 \times 1000)$ tensor in double precision floating points would require 8 gigabytes and that the algebraic Algorithm 1 for the computation of its CPD would have a complexity of the order $\mathcal{O}(I^2J^2R^2) = 10^{10}$. It is more practical to sample $N = 5$ subtensors $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(N)}$ of size $(10\times10\times K)$ from $\mathcal{X}$. In order to guarantee the recovery of the CPD factors of $\mathcal{X}$, the row border fibers $\mathbf{x}_{11\bullet}, \ldots, \mathbf{x}_{I1\bullet}$ and the column border fibers $\mathbf{x}_{11\bullet}, \ldots, \mathbf{x}_{1J\bullet}$ 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 subsampling scheme only the $\sum_{n=1}^{N}(I_nJ_n) = 500$ fibers of $\mathcal{X}$ that correspond to the fibers in $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(N)}$



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

FIG. 6.4. *Fiber sampling pattern (a blue cross + at position (i,j) indicates that fiber $\mathbf{x}_{ij\bullet}$ has been sampled) and mean $P(\mathbf{A})$ over 100 trials while SNR is varying from 10 to 40 dB, case 4. (Color available online.)*

are used in the computation of $\mathbf{C}$. We also refine the estimate of $\mathbf{C}$ obtained by Algorithm 1 using the basic alternating least squares (ALS) method for coupled CPD of $\mathcal{X}^{(1)},\ldots,\mathcal{X}^{(N)}$. The latter method will be referred to as "Subsampled Algorithm 1-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 6.4. We observe that the algebraic "Subsampled Algorithm 1" and refined "Subsampled Algorithm 1-ALS" methods perform about same. The reason is that the coupled CPD of $\mathcal{X}^{(1)},\ldots,\mathcal{X}^{(N)}$ is by itself highly overdetermined and consequently little is gained by the extra ALS refinement step.

7. 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. We also explained that fiber sampling is of interest for large-scale low-rank tensor decomposition problems where we deliberately sample only a few fibers. Numerical experiments showed that the presented algebraic algorithm can efficiently provide a good initialization for optimization-based methods commonly used for incomplete tensor decompositions. 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 [11]. In the supplementary material in [39] we illustrate this for fiber sampled block term and hierarchical rank decompositions.

Acknowledgment. The authors thank Professor Alwin Stegeman for early proofreading and for several helpful comments.

REFERENCES

[1] E. ACAR, D. M. DUNLAVY, T. G. KOLDA, AND M. MØRUP, *Scalable tensor factorizations for incomplete data*, Chemometr. Intell. Lab., 106 (2011), pp. 41–56.
[2] B. W. BADER, T. G. KOLDA, ET AL., *MATLAB Tensor Toolbox Version* 3.0-*dev*, available online at https://www.tensortoolbox.org, 2017.
[3] R. B. BAPAT, *Graphs and Matrices*, Springer, New York, 2011.

[4] C. Battaglino, G. Ballard, and T. G. Kolda, *A Practical Randomized CP Tensor Decomposition*, preprint, https://arxiv.org/abs/1701.06600, 2017.
[5] D. J. Biagioni, D. Beylkin, and G. Beylkin, *Randomized interpolative decomposition of separated representations*, J. Comput. Phys., 281 (2015), pp. 116–134.
[6] 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.
[7] E. Candès and B. Recht, *Exact matrix completion via convex optimization*, Found. Comput. Math., 9 (2009), pp. 717–772.
[8] L. Chiantini and G. Ottaviani, *On generic identifiability of 3-tensors of small rank*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1018–1037, https://doi.org/10.1137/110829180.
[9] A. Cichocki, D. Mandic, C. Caifa, A.-H. Phan, G. Zhou, and L. De Lathauwer, *Tensor decompositions for signal processing applications: From two-way to multiway component analysis*, IEEE Signal Process. Mag., 32 (2015), pp. 145–163.
[10] 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, https://doi.org/10.1137/040608830.
[11] 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, https://doi.org/10.1137/070690729.
[12] 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, https://doi.org/10.1137/120877234.
[13] I. Domanov and L. De Lathauwer, *On the uniqueness of the canonical polyadic decomposition of third-order tensors—part* II: *Uniqueness of the overall decomposition*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 876–903, https://doi.org/10.1137/120877258.
[14] 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, https://doi.org/10.1137/130916084.
[15] 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, https://doi.org/10.1137/140970276.
[16] I. Domanov and L. De Lathauwer, *Generic uniqueness of a structured matrix factorization and applications in blind source separation*, IEEE J. Selected Topics Signal Process., 10 (2016), pp. 701–711.
[17] 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.
[18] P. Drineas and M. W. Mahoney, *A randomized algorithm for a tensor-based generalization of the singular value decomposition*, Linear Algebra Appl., 420 (2007), pp. 553–571.
[19] A. Frieze, R. Kannan, and S. Vempala, *Fast Monte-Carlo algorithms for finding low-rank approximations*, J. ACM, 51 (2004), pp. 1025–1041.
[20] S. Gandy, B. Recht, and I. Yamada, *Tensor completion and low-n-rank tensor recovery via convex optimization*, Inverse Problems, 27 (2011), 025010.
[21] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, MD, 2013.
[22] D. Hadwin, K. J. Harrison, and J. A. Ward, *Rank-one completions of partial matrices and completely rank-nonincreasing linear functionals*, Proc. Amer. Math. Soc., 134 (2006), pp. 2169–2178.
[23] N. Halko, P. G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288, https://doi.org/10.1137/090771806.
[24] 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.
[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2013.
[26] I. Ibraghimov, *Application of the three-way decomposition for matrix compression*, Numer. Linear Algebra Appl., 9 (2002), pp. 551–565.
[27] T. Jiang and N. D. Sidiropoulos, *Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear model with constant modulus constraints*, IEEE Trans. Signal Process., 52 (2004), pp. 2625–2636.

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

[29] 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, https://doi.org/10.1137/0614071.

[30] T. Maehara, K. Hayashi, and K.-I. Kawarabayashi, *Expected tensor decomposition with stochastic gradient descent*, in Proceedings of the 13th AAAI Conference on Artificial Intelligence (AAAI'16), Phoenix, AZ, 2016, pp. 1919–1925.

[31] 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.

[32] 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.

[33] B. Recht, *A simpler approach to matrix completion*, J. Mach. Learn. Res., 12 (2011), pp. 3413–3430.

[34] B. Recht, M. Fazel, and P. A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev., 52 (2010), pp. 471–501, https://doi.org/10.1137/070697835.

[35] 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 Process., 65 (2017), pp. 3551–3582.

[36] 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, https://doi.org/10.1137/140956853.

[37] 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, https://doi.org/10.1137/140956865.

[38] M. Sørensen and L. De Lathauwer, *Shift invariance, incomplete arrays and coupled CPD: A case study*, in Proceedings of the 9th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2016), Rio de Janeiro, Brazil, 2016, pp. 1–5.

[39] M. Sørensen and L. De Lathauwer, *Fiber Sampling Approach to Canonical Polyadic Decomposition and Application to Tensor Completion*, Technical Report 15-151, ESAT-STADIUS, KU Leuven, Belgium, 2017.

[40] M. Sørensen and L. De Lathauwer, *Multidimensional harmonic retrieval via coupled canonical polyadic decomposition—part* I: *Model and identifiability*, IEEE Trans. Signal Process., 65 (2017), pp. 517–527.

[41] M. Sørensen and L. De Lathauwer, *Multidimensional harmonic retrieval via coupled canonical polyadic decomposition—part* II: *Algorithm and multirate sampling*, IEEE Trans. Signal Process., 65 (2017), pp. 528–539.

[42] M. Sørensen, I. Domanov, and L. De Lathauwer, *Coupled canonical polyadic decompositions and multiple shift invariance in array processing*, IEEE Trans. Signal Process., 66 (2018), pp. 3665–3680.

[43] 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.

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

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

[46] J. M. F. Ten Berge and J. N. Tendeiro, *The link between sufficient conditions by Harshman and by Kruskal for uniqueness in Candecomp/Parafac*, J. Chemometrics, 23 (2009), pp. 321–323.

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

[48] C. E. Tsourakakis, *MACH: Fast randomized tensor decompositions*, in Proceedings of the 2010 SIAM International Conference on Data Mining, SIAM, Philadelphia, 2010, pp. 689–700, https://doi.org/10.1137/1.9781611972801.60.

[49] 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. Amer. Chem. Soc., 127 (2005), pp. 2767–2775.

[50] 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.

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

[52] A.-J. Van Der Veen and A. Paulraj, *An analytical constant modulus algorithm*, IEEE Trans. Signal Process., 44 (1996), pp. 1136–1155.

[53] 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.

[54] N. Vervliet, O. Debals, and L. De Lathauwer, *Canonical Polyadic Decomposition of Incomplete Tensors with Linearly Constrained Factors*, Technical Report 16–172, ESAT-STADIUS, KU Leuven, Belgium, 2016.

[55] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab v*3.0, available at https://www.tensorlab.net/, 2016.

[56] M. Yuan and C.-H. Zhang, *On tensor completion via nuclear norm minimization*, Found. Comput. Math., 16 (2016), pp. 1031–1068.

[57] 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.