Open access • Journal Article • DOI:10.1137/140964813

New Uniqueness Conditions for the Canonical Polyadic Decomposition of Third-Order Tensors — Source link

Mikael Sorensen, Lieven De Lathauwer

Published on: 01 Oct 2015 - SIAM Journal on Matrix Analysis and Applications (Society for Industrial and Applied Mathematics)

Topics: Uniqueness

SIAM J. MATRIX ANAL. APPL.
Vol. 36, No. 4, pp. 1381–1403

NEW UNIQUENESS CONDITIONS FOR THE CANONICAL POLYADIC DECOMPOSITION OF THIRD-ORDER TENSORS*

MIKAEL SØRENSEN† AND LIEVEN DE LATHAUWER†

Abstract. The uniqueness properties of the canonical polyadic decomposition (CPD) of higher-order tensors make it an attractive tool for signal separation. However, CPD uniqueness is not yet fully understood. In this paper, we first present a new uniqueness condition for a polyadic decomposition (PD) where one of the factor matrices is assumed to be known. We also show that this result can be used to obtain a new overall uniqueness condition for the CPD. In signal processing the CPD factor matrices are often constrained. Building on the preceding results, we provide a new uniqueness condition for a CPD with a columnwise orthonormal factor matrix, representing uncorrelated signals. We also obtain a new uniqueness condition for a CPD with a partial Hermitian symmetry, useful for tensors in which covariance matrices are stacked, which are common in statistical signal processing. We explain that such constraints can lead to more relaxed uniqueness conditions. Finally, we provide an inexpensive algorithm for computing a PD with a known factor matrix that is also useful for the computation of the full CPD.

Key words. tensor, polyadic decomposition, parallel factor, canonical decomposition, canonical polyadic decomposition

AMS subject classifications. 15A22, 15A23, 15A69

DOI. 10.1137/140964813

1. Introduction.

Tensor decompositions are finding more and more applications in signal processing, data analysis, and machine learning. For instance, canonical polyadic decomposition (CPD) is becoming a basic tool for signal separation. Essentially, this is due to the fact that CPD is unique under mild conditions, compared to decomposition of a matrix in rank-1 terms. Thanks to their uniqueness, the rank-1 terms are easily associated with interpretable data components. Numerous applications have been reported in independent component analysis, exploratory data analysis, wireless communication, radar, chemometrics, psychometrics, sensor array processing, and so on [1, 2, 7, 20, 21, 17, 19, 14, 15]. However, the understanding of uniqueness is lagging behind the use of CPD in practice.

Many uniqueness conditions for the CPD such as those developed in [16, 12, 27, 8, 9] are based on Kruskal's permutation lemma [16]. To prove uniqueness of a CPD, one may first prove the uniqueness of one factor matrix. In the next step, overall CPD uniqueness is obtained from the uniqueness of a polyadic decomposition (PD) with a known factor matrix. The former problem has been thoroughly studied in [8]. In this paper we focus on the latter problem.

*Received by the editors April 14, 2014; accepted for publication (in revised form) by D. P. O'Leary June 17, 2015; published electronically October 1, 2015. The research of the authors was supported by (1) Research Council KU Leuven: GOA/10/09 MaNet, CoE PFV/10/002 (OPTEC); (2) F.W.O.: project G.0830.14N, G.0881.14N; (3) the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012-2017); (4) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (339804). This paper reflects only the authors' views, and the Union is not liable for any use that may be made of the contained information.

http://www.siam.org/journals/simax/36-4/96481.html

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

Besides uniqueness, the problem of studying PD with a known factor matrix has direct relevance in applications. As an example, we mention that several wireless communication systems based on tensor-coding have been proposed (e.g., [3, 18, 4, 5]) since their conception in [22]. It suffices to say that many tensor-based wireless communication systems essentially rely on computing a PD with a known factor. However, mainly optimization-based methods such as alternating least squares (ALS) have been considered. Such iterative methods are not guaranteed to find the decomposition, even in the exact case. Thus, an optimization-based method can suffer from slow convergence (many iterations may be needed) and local minima (many initializations may be needed). On the other hand, an algebraic method is guaranteed to find the decomposition in the exact case. For sufficiently high signal-to-noise ratio (SNR), algebraic methods can provide an inexpensive but accurate estimate of the solution that can be used to initialize an optimization-based algorithm. Hence, the development of an algebraic method for computing the PD with a known factor matrix is another relevant problem that will be addressed in this paper. For this reason we will develop a constructive uniqueness proof for the PD with a known factor matrix.

The contributions of this paper are the following. We first present a new constructive uniqueness condition for a PD with a known factor matrix that leads to more relaxed conditions than those obtained in [9] and is eligible in cases where none of the involved factors have full column rank. Based on this result we propose a new, relatively easy-to-check deterministic overall uniqueness condition for the CPD that is comparable with recent relaxed deterministic overall CPD uniqueness conditions [9]. In the supplementary material we present a partial uniqueness variant. In tensor-based signal processing CPD is often constrained. Uncorrelated signals may translate into a columnwise orthonormal factor matrix [24], while stacking covariance matrices may lead to a CPD with a partial Hermitian symmetry (e.g., [2, 7]). We provide a new uniqueness condition for a CPD with a columnwise orthonormal factor matrix. We also provide a new uniqueness condition for a CPD with a partial Hermitian symmetry. We explain that such constrained CPDs can be unique under milder conditions than their unconstrained counterparts. Finally, we present an algorithm for computing a PD with a known factor matrix that does not cost much more than a single ALS iteration. Numerical experiments demonstrate that this inexpensive algorithm provides a good initialization for a subsequent optimization-based method such as ALS in difficult cases with modest to high SNR.

This paper is organized as follows. The rest of the introduction presents our notation. Sections 2 and 3 briefly review the CPD and the PD with a known factor matrix, respectively. Section 4 presents a new uniqueness condition and an algorithm for a CPD with a known factor matrix. Next, in section 5 we present new uniqueness conditions for the overall CPD and for some variants. Numerical experiments are reported in section 6. Section 7 concludes the paper. We also mention that in the supplementary material a new partial uniqueness condition for CPD and an efficient implementation of the ALS method for CPD with a known factor matrix are reported.

1.1. Notation. Vectors, matrices, and tensors are denoted by lower case boldface, upper case boldface, and upper case calligraphic letters, respectively. The rth column vector of $\mathbf{A}$ is denoted by $\mathbf{a}_r$. The symbols $\otimes$ and $\odot$ denote the Kronecker and Khatri–Rao products, defined as

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

in which $(\mathbf{A})_{mn} = a_{mn}$. The Hadamard product is denoted by $*$ and satisfies $(\mathbf{A} * \mathbf{B})_{mn} = a_{mn}b_{mn}$. The outer product of N vectors $\mathbf{a}^{(n)} \in \mathbb{C}^{I_n}$ is denoted by $\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ such that

$$\left(\mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)}\right)_{i_1,i_2,\ldots,i_N} = a^{(1)}_{i_1} a^{(2)}_{i_2} \cdots a^{(N)}_{i_N}.$$

The identity matrix, all-zero matrix, and all-zero vector are denoted by $\mathbf{I}_M \in \mathbb{C}^{M \times M}$, $\mathbf{0}_{M,N} \in \mathbb{C}^{M \times N}$, and $\mathbf{0}_M \in \mathbb{C}^M$, respectively. The symbol δ_{ij} denotes the Kronecker delta function, equal to 1 if $i = j$ and 0 if $i \neq j$.

The real part, imaginary part, transpose, conjugate, conjugate-transpose, inverse, Moore–Penrose pseudoinverse, Frobenius norm, determinant, range, and kernel of a matrix are denoted by $\mathrm{Re}\{\cdot\}$, $\mathrm{Im}\{\cdot\}$, $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, $\|\cdot\|_F$, $|\cdot|$, $\mathrm{range}\,(\cdot)$, and $\ker\,(\cdot)$, respectively.

The rank of a matrix $\mathbf{A}$ is denoted by $r\,(\mathbf{A})$ or $r_{\mathbf{A}}$. The k-rank of a matrix $\mathbf{A}$ is denoted by $k\,(\mathbf{A})$ or $k_{\mathbf{A}}$. It is equal to the largest integer $k_{\mathbf{A}}$ such that every subset of $k_{\mathbf{A}}$ columns of $\mathbf{A}$ is linearly independent. Since the rank and k-rank quantities of matrices will play an important role in this paper, it is important to notice the differences between the two. In particular, we have that $k_{\mathbf{A}} \leq r_{\mathbf{A}}$.

The cardinality of a set S is denoted by $\mathrm{card}\,(S)$.

Given $\mathbf{A} \in \mathbb{C}^{I \times J}$, then $\mathrm{Vec}\,(\mathbf{A}) \in \mathbb{C}^{IJ}$ will denote the column vector defined by $(\mathrm{Vec}\,(\mathbf{A}))_{i+(j-1)I} = (\mathbf{A})_{ij}$. Given $\mathbf{a} \in \mathbb{C}^{IJ}$, then the reverse operation is $\mathrm{Unvec}\,(\mathbf{a}) = \mathbf{A} \in \mathbb{C}^{I \times J}$ such that $(\mathbf{a})_{i+(j-1)I} = (\mathbf{A})_{ij}$. $D_k\,(\mathbf{A}) \in \mathbb{C}^{J \times J}$ denotes the diagonal matrix holding row k of $\mathbf{A} \in \mathbb{C}^{I \times J}$ on its diagonal.

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 $\mathbf{A}$. Likewise, $\mathbf{A}([1,2],[1,3])$ denotes the 2×2 submatrix $\mathbf{A}([1,2],[1,3]) = \left[\begin{smallmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{smallmatrix}\right]$.

Let $\mathbf{T} \in \mathbb{C}^{I \times R}$; then $\underline{\mathbf{T}} = \mathbf{T}\,(1:I-1,:) \in \mathbb{C}^{(I-1) \times R}$, i.e., $\underline{\mathbf{T}}$ is obtained by deleting the bottom row of $\mathbf{T}$.

The matrix that orthogonally projects onto the orthogonal complement of the column space of $\mathbf{A} \in \mathbb{C}^{I \times J}$ is denoted by

$$\mathbf{P}_{\mathbf{A}} = \mathbf{I}_I - \mathbf{F}\mathbf{F}^H \in \mathbb{C}^{I \times I},$$

where the column vectors of $\mathbf{F}$ constitute an orthonormal basis for $\mathrm{range}\,(\mathbf{A})$.

Let $C_n^k = \frac{n!}{k!(n-k)!}$ denote the binomial coefficient. The kth compound matrix of $\mathbf{A} \in \mathbb{C}^{m \times n}$ is denoted by $\mathcal{C}_k\,(\mathbf{A}) \in \mathbb{C}^{C_m^k \times C_n^k}$ and its entries correspond to the $k \times k$ minors of $\mathbf{A}$ ordered lexicographically. As an example, let $\mathbf{A} \in \mathbb{C}^{4 \times 3}$; then

$$\mathcal{C}_2\,(\mathbf{A}) = \begin{bmatrix} |\mathbf{A}\,([1,2],[1,2])| & |\mathbf{A}\,([1,2],[1,3])| & |\mathbf{A}\,([1,2],[2,3])| \\ |\mathbf{A}\,([1,3],[1,2])| & |\mathbf{A}\,([1,3],[1,3])| & |\mathbf{A}\,([1,3],[2,3])| \\ |\mathbf{A}\,([1,4],[1,2])| & |\mathbf{A}\,([1,4],[1,3])| & |\mathbf{A}\,([1,4],[2,3])| \\ |\mathbf{A}\,([2,3],[1,2])| & |\mathbf{A}\,([2,3],[1,3])| & |\mathbf{A}\,([2,3],[2,3])| \\ |\mathbf{A}\,([2,4],[1,2])| & |\mathbf{A}\,([2,4],[1,3])| & |\mathbf{A}\,([2,4],[2,3])| \\ |\mathbf{A}\,([3,4],[1,2])| & |\mathbf{A}\,([3,4],[1,3])| & |\mathbf{A}\,([3,4],[2,3])| \end{bmatrix}.$$

See [10, 8] for a discussion of compound matrices.

2. 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 some 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:

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

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 $\mathcal{X}$, i.e., a PD of $\mathcal{X}$ with a minimal number of terms is a CPD. Let us stack the vectors $\{\mathbf{a}_r\}$, $\{\mathbf{b}_r\}$, and $\{\mathbf{c}_r\}$ into the matrices

$$\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R] \in \mathbb{C}^{I\times R}, \quad \mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R] \in \mathbb{C}^{J\times R}, \quad \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{C}^{K\times R}.$$

The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are called CPD factor matrices. Let $\mathbf{X}^{(i\cdot\cdot)} \in \mathbb{C}^{J\times K}$ denote the matrix such that $\left(\mathbf{X}^{(i\cdot\cdot)}\right)_{jk} = x_{ijk}$; then $\mathbf{X}^{(i\cdot\cdot)} = \mathbf{B} D_i(\mathbf{A})\mathbf{C}^T$ and

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

Similarly, let the matrices $\mathbf{X}^{(\cdot j\cdot)} \in \mathbb{C}^{K\times I}$ be constructed such that $\left(\mathbf{X}^{(\cdot j\cdot)}\right)_{ki} = x_{ijk}$; then $\mathbf{X}^{(\cdot j\cdot)} = \mathbf{C} D_j(\mathbf{B})\mathbf{A}^T$ and

$$\mathbb{C}^{JK\times I} \ni \mathbf{X}_{(2)} := \left[\mathbf{X}^{(\cdot 1\cdot)T}, \ldots, \mathbf{X}^{(\cdot J\cdot)T}\right]^T = (\mathbf{B}\odot\mathbf{C})\,\mathbf{A}^T. \tag{2.3}$$

Finally, let $\mathbf{X}^{(\cdot\cdot k)} \in \mathbb{C}^{I\times J}$ satisfy $\left(\mathbf{X}^{(\cdot\cdot k)}\right)_{ij} = x_{ijk}$; then $\mathbf{X}^{(\cdot\cdot k)} = \mathbf{A} D_k(\mathbf{C})\mathbf{B}^T$ and

$$\mathbb{C}^{IK\times J} \ni \mathbf{X}_{(3)} := \left[\mathbf{X}^{(\cdot\cdot 1)T}, \ldots, \mathbf{X}^{(\cdot\cdot K)T}\right]^T = (\mathbf{C}\odot\mathbf{A})\,\mathbf{B}^T. \tag{2.4}$$

2.1. Uniqueness conditions for one factor matrix of a CPD. A factor matrix, say $\mathbf{C}$, of the CPD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ is said to be unique if it can be determined up to the inherent column scaling and permutation ambiguities from $\mathcal{X}$. More formally, the factor matrix $\mathbf{C}$ is unique if alternative factor matrices $\widehat{\mathbf{C}}$ satisfy the condition

$$\widehat{\mathbf{C}} = \mathbf{C}\mathbf{P}\Delta_{\mathbf{C}}\,,$$

where $\mathbf{P}$ is a permutation matrix and $\Delta_{\mathbf{C}}$ is a nonsingular diagonal matrix. Based on Kruskal's permutation lemma [16] and the properties of compound matrices [10, 8], the following theorem, Theorem 2.1, guarantees the uniqueness of one factor matrix of a CPD. Theorem 2.1 is one of the most relaxed and yet quite easy-to-check sufficient (but not necessary) uniqueness conditions that have been reported in the literature. Consequently, we will use Theorem 2.1 to obtain new overall uniqueness results.

THEOREM 2.1 (see [8]). *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in (2.1). If*

$$\begin{cases} k_{\boldsymbol{C}} \geq 1, \\ \min(I,J) \geq R - r_{\boldsymbol{C}} + 2, \\ \mathcal{C}_{R-r_{\boldsymbol{C}}+2}(\boldsymbol{A}) \odot \mathcal{C}_{R-r_{\boldsymbol{C}}+2}(\boldsymbol{B}) \quad \textit{has full column rank,} \end{cases} \tag{2.5}$$

then the rank of $\mathcal{X}$ is R and the factor matrix $\boldsymbol{C}$ is unique.

Note that the first two conditions in (2.5) are trivial. More precisely, the k-rank condition $\mathrm{k}_{\mathbf{C}} \geq 1$ ensures that none of the columns of $\mathbf{C}$ is zero while the condition $\min(I,J) \geq R - \mathrm{r}_{\mathbf{C}} + 2$ ensures that $\mathcal{C}_{R-r_{\mathbf{C}}+2}(\mathbf{A}) \odot \mathcal{C}_{R-r_{\mathbf{C}}+2}(\mathbf{B})$ is well defined.

We refer to [8] and the references therein for other conditions that guarantee the uniqueness of one factor matrix of a CPD.

2.2. Uniqueness conditions for CPD. The CPD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ is said to be unique if all the triplets $(\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}})$ satisfying (2.1) are related via

$$\widehat{\mathbf{A}} = \mathbf{A}\mathbf{P}\Delta_{\mathbf{A}}, \quad \widehat{\mathbf{B}} = \mathbf{B}\mathbf{P}\Delta_{\mathbf{B}}, \quad \widehat{\mathbf{C}} = \mathbf{C}\mathbf{P}\Delta_{\mathbf{C}},$$

where $\Delta_{\mathbf{A}}$, $\Delta_{\mathbf{B}}$, and $\Delta_{\mathbf{C}}$ are diagonal matrices satisfying $\Delta_{\mathbf{A}}\Delta_{\mathbf{B}}\Delta_{\mathbf{C}} = \mathbf{I}_R$ and $\mathbf{P}$ is a permutation matrix.

Theorem 2.2 below is one of the most relaxed and yet quite easy-to-check deterministic CPD uniqueness conditions that have been reported in the literature.

THEOREM 2.2 (see [9]). *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in* (2.1)*. If*

$$\begin{cases} \min(k_{\boldsymbol{A}}, k_{\boldsymbol{B}}) + k_{\boldsymbol{C}} \geq R+1, \\ \max(k_{\boldsymbol{A}}, k_{\boldsymbol{B}}) + k_{\boldsymbol{C}} \geq R+2, \\ \mathcal{C}_{R-r_C+2}(\boldsymbol{A}) \odot \mathcal{C}_{R-r_C+2}(\boldsymbol{B}) \ \text{ has full column rank,} \end{cases} \tag{2.6}$$

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

See [9] and the references therein for more related deterministic overall CPD uniqueness conditions.

3. PD with known factor matrix. Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ with matrix representation $\mathbf{X}_{(1)} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T$. Assume that the factor matrix $\mathbf{C}$ is known and let the pair $(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})$ yield an alternative decomposition of $\mathcal{X}$ with the same $\mathbf{C}$. The PD of $\mathcal{X}$ with the known factor matrix $\mathbf{C}$ is said to be unique if all the pairs $(\widehat{\mathbf{A}}, \widehat{\mathbf{B}})$ satisfying (2.1) are related via

$$\widehat{\mathbf{A}} = \mathbf{A}\Delta_{\mathbf{A}}, \quad \widehat{\mathbf{B}} = \mathbf{B}\Delta_{\mathbf{B}},$$

where $\Delta_{\mathbf{A}}$ and $\Delta_{\mathbf{B}}$ are diagonal matrices with property $\Delta_{\mathbf{A}}\Delta_{\mathbf{B}} = \mathbf{I}_R$.

In this paper we also consider generic uniqueness conditions for PD with a known factor matrix. Assume that the factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are randomly drawn from absolutely continuous probability measures. Then we say that the PD of $\mathcal{X}$ in (2.1) with $\mathbf{C}$ known is generically unique if the set of nonunique PDs of $\mathcal{X}$ with $\mathbf{C}$ fixed (and $\mathbf{A}$, $\mathbf{B}$ varying) is of Lebesgue measure zero. Intuitively, a generic property holds with high probability for sufficiently random data.

If the known factor matrix has full column rank, then the PD with a known factor is unique in a trivial manner and can be computed via a number of rank-1 approximations (e.g., [29, 6]). Indeed, by recognizing that the columns of $\mathbf{Y} = \mathbf{X}_{(1)}(\mathbf{C}^T)^{\dagger} = \mathbf{A} \odot \mathbf{B}$ are vector representations of the rank-1 matrices $\{\mathbf{a}_r\mathbf{b}_r^T\}$, $r \in \{1,\ldots,R\}$, it is clear that the factor matrices $\mathbf{A}$ and $\mathbf{B}$ are unique. We state this result here for completeness.

PROPOSITION 3.1. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in* (2.1)*. If $\boldsymbol{C}$ is known and has full column rank, then the PD of $\mathcal{X}$ with $\boldsymbol{C}$ known is unique.*

The following proposition, Proposition 3.2, presents a uniqueness condition for the case where the known factor matrix is not required to have full column rank.

PROPOSITION 3.2 (see [9]). *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in* (2.1)*. If $\boldsymbol{C}$ is known and*

$$k_{\boldsymbol{C}} + \min(\min(k_{\boldsymbol{A}}, k_{\boldsymbol{B}}) - 1, \max(k_{\boldsymbol{A}}, k_{\boldsymbol{B}}) - 2) \geq R, \tag{3.1}$$

then the PD of $\mathcal{X}$ with $\boldsymbol{C}$ known is unique.

Assume that $\mathbf{A}$ is randomly drawn from an absolutely continuous probability measure; then it is well known that $k_{\mathbf{A}} = \min(I, R)$ (similarly for $k_{\mathbf{B}}$ and $k_{\mathbf{C}}$). Hence, condition (3.1) is generically satisfied if $\min(K, R) + \min(\min(V, R) - 1, \min(W, R) - 2) \geq R$, where $V = \min(I, J)$ and $W = \max(I, J)$.

Except for the case where $\mathbf{C}$ has full column rank and $\max(k_{\mathbf{A}}, k_{\mathbf{B}}) = 1$, Proposition 3.2 yields a more relaxed condition than Proposition 3.1.

4. New uniqueness result for PD with a known factor matrix. In subsection 4.1 we present a new result for the case where the known factor matrix does not necessarily have full column rank, but one of the unknown factors does. Subsection 4.2 generalizes the result to the case where none of the involved factor matrices is required to have full column rank.

4.1. At least one factor matrix has full column rank. The main result of this subsection is a new uniqueness condition for PD with a known factor, stated in Theorem 4.6. Condition (4.12) in Theorem 4.6 boils down to checking the rank of some matrices. In order to assess whether the rank condition (4.12) in Theorem 4.6 is expected to be satisfied, we resort to the following tool for checking generic conditions (e.g., [13]).

LEMMA 4.1. *We are given an analytic function $f : \mathbb{C}^n \to \mathbb{C}$. If there exists a vector $\boldsymbol{x} \in \mathbb{C}^n$ such that $f(\boldsymbol{x}) \neq 0$, then the set $\{ \boldsymbol{x} | f(\boldsymbol{x}) = 0 \}$ is of Lebesgue measure zero.*

Lemma 4.1 tells us that in order to obtain a generic rank property for matrices it suffices to numerically check the rank condition for just one example. The link between matrix rank properties and Lemma 4.1 is that an $I \times R$ matrix has full column rank R if it has a nonvanishing $R \times R$ minor, where a minor is an analytic function (namely, it is a polynomial in several variables). Thus, in order to check if an $I \times R$ matrix generically has full column rank, we just need to find one $I \times R$ matrix with a nonvanishing $R \times R$ minor.

A matrix $\mathbf{A} \in \mathbb{C}^{I \times R}$ is said to be Vandermonde if it takes the form

$$\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R], \quad \mathbf{a}_r = \left[1, z_r, z_r^2, \ldots, z_r^{I-1}\right]^T, \tag{4.1}$$

where $\{z_r\}$ are called the generators of $\mathbf{A}$. Because of their simplicity, we use Vandermonde matrices in the development of the following auxiliary results, which will lead to the example for assessing whether the rank condition (4.12) in Theorem 4.6 is generically satisfied.

LEMMA 4.2. *Let $\boldsymbol{A} \in \mathbb{C}^{I \times R}$ be a Vandermonde matrix and let $\boldsymbol{B} \in \mathbb{C}^{J \times R}$; then the matrix $\boldsymbol{A} \odot \boldsymbol{B}$ generically has rank* $\min(IJ, R)$.

Proof. The result follows from a combination of Theorem 3 and Corollary 1 in [11]. ◻

LEMMA 4.3. *Let $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_R] \in \mathbb{C}^{I \times R}$ be a Vandermonde matrix with distinct generators. Let the column vectors of $\boldsymbol{U} \in \mathbb{C}^{IJ \times R}$ constitute a basis for range$(\boldsymbol{A} \odot \boldsymbol{B})$, where $\boldsymbol{B} \in \mathbb{C}^{J \times R}$. If the matrix $\underline{\boldsymbol{A}} \odot \boldsymbol{B}$ has full column rank, then the matrix*

$$\boldsymbol{G}^{(r)} = [\boldsymbol{U}, \boldsymbol{a}_r \otimes \boldsymbol{I}_J] \in \mathbb{C}^{IJ \times (R+J)} \tag{4.2}$$

has rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$. Generically, $\boldsymbol{G}^{(r)}$ has rank $R+J-1$ if $(I-1)J+1 \geq R$.

Proof. Consider a Vandermonde matrix $\mathbf{A} \in \mathbb{C}^{I \times R}$ of the form (4.1). Let us first prove that if $\underline{\mathbf{A}} \odot \mathbf{B}$ has full column rank, then $\mathbf{G}^{(r)}$ has a one-dimensional kernel. Note that all bases for range$(\mathbf{A} \odot \mathbf{B})$ are related via a right multiplication by a nonsingular matrix which does not affect the rank of $\mathbf{G}^{(r)}$. Thus, without loss of generality (w.l.o.g.) we set $\mathbf{U} = \mathbf{A} \odot \mathbf{B}$ such that $\mathbf{G}^{(r)} \in \mathbb{C}^{IJ\times(R+J)}$ becomes

$$\mathbf{G}^{(r)} = [\mathbf{A} \odot \mathbf{B}, \mathbf{a}_r \otimes \mathbf{I}_J] = \begin{bmatrix} \mathbf{B} & \mathbf{I}_J \\ \mathbf{BZ} & z_r\mathbf{I}_J \\ \vdots & \vdots \\ \mathbf{BZ}^{I-1} & z_r^{I-1}\mathbf{I}_J \end{bmatrix},$$

where $\mathbf{Z} = D_1([z_1, z_2, \dots, z_R])$. Consider the following nonsingular block-bidiagonal matrix:

$$\mathbf{F}^{(r)} = \begin{bmatrix} \mathbf{I}_J & \mathbf{0}_{J,J} & \cdots & \cdots & \mathbf{0}_{J,J} & \mathbf{0}_{J,J} \\ -z_r\mathbf{I}_J & \mathbf{I}_J & \ddots & & \vdots & \vdots \\ \mathbf{0}_{J,J} & -z_r\mathbf{I}_J & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{0}_{J,J} & \mathbf{0}_{J,J} \\ \mathbf{0}_{J,J} & & \ddots & \ddots & \mathbf{I}_J & \mathbf{0}_{J,J} \\ \mathbf{0}_{J,J} & \cdots & \cdots & \mathbf{0}_{J,J} & -z_r\mathbf{I}_J & \mathbf{I}_J \end{bmatrix} \in \mathbb{C}^{IJ\times IJ};$$

then

$$\mathbf{F}^{(r)}\mathbf{G}^{(r)} = \begin{bmatrix} \mathbf{B} & \mathbf{I}_J \\ \mathbf{B}(\mathbf{Z} - z_r\mathbf{I}_R) & \mathbf{0}_{J,J} \\ \vdots & \vdots \\ \mathbf{B}(\mathbf{Z}^{I-1} - z_r\mathbf{Z}^{I-2}) & \mathbf{0}_{J,J} \end{bmatrix}. \tag{4.3}$$

Hence, the problem of determining the rank of $\mathbf{G}^{(r)}$ reduces to finding the rank of

$$\begin{aligned} \mathbf{H}^{(r)} &= \begin{bmatrix} \mathbf{B}(\mathbf{Z} - z_r\mathbf{I}_R) \\ \vdots \\ \mathbf{B}(\mathbf{Z}^{I-1} - z_r\mathbf{Z}^{I-2}) \end{bmatrix} = \begin{bmatrix} \mathbf{BI}_R(\mathbf{Z} - z_r\mathbf{I}_R) \\ \vdots \\ \mathbf{BZ}^{I-2}(\mathbf{Z} - z_r\mathbf{I}_R) \end{bmatrix} \\ &= (\underline{\mathbf{A}} \odot \mathbf{B})(\mathbf{Z} - z_r\mathbf{I}_R). \end{aligned} \tag{4.4}$$

The rth column vector $\mathbf{h}_r^{(r)}$ of $\mathbf{H}^{(r)}$ is an all-zero vector. Thus the problem of determining the rank of $\mathbf{G}^{(r)}$ further reduces to finding the rank of

$$\widetilde{\mathbf{H}}^{(r)} = \big(\underline{\widetilde{\mathbf{A}}}^{(r)} \odot \widetilde{\mathbf{B}}^{(r)}\big)\widetilde{\mathbf{Z}}^{(r)}, \tag{4.5}$$

where

$$\begin{aligned} \widetilde{\mathbf{A}}^{(r)} &= [\mathbf{a}_1, \dots, \mathbf{a}_{r-1}, \mathbf{a}_{r+1}, \dots, \mathbf{a}_R] \in \mathbb{C}^{I\times(R-1)}, \\ \widetilde{\mathbf{B}}^{(r)} &= [\mathbf{b}_1, \dots, \mathbf{b}_{r-1}, \mathbf{b}_{r+1}, \dots, \mathbf{b}_R] \in \mathbb{C}^{J\times(R-1)}, \\ \widetilde{\mathbf{Z}}^{(r)} &= D_1([z_1 - z_r, \dots, z_{r-1} - z_r, z_{r+1} - z_r, \dots, z_R - z_r]) \in \mathbb{C}^{(R-1)\times(R-1)}. \end{aligned}$$

The deterministic part of Lemma 4.3 follows from (4.5). Indeed, the full column rank assumption on $\underline{\mathbf{A}} \odot \mathbf{B}$ implies that $\underline{\widetilde{\mathbf{A}}}^{(r)} \odot \widetilde{\mathbf{B}}^{(r)}$ has full column rank for every

$r \in \{1, \ldots, R\}$. This in turn implies that the matrix in (4.4) has rank $R-1$ and, consequently, $\mathbf{G}^{(r)}$ in (4.2) has rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$. We can now conclude that any $\mathbf{U}$ of which the columns form a basis for range $(\mathbf{A} \odot \mathbf{B})$ yields a matrix $\mathbf{G}^{(r)}$ of rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$.

Let us now prove that if $(I-1)J+1 \geq R$, then generically $\mathbf{G}^{(r)}$ has a one-dimensional kernel. Due to Lemma 4.1 we just need to find one example where the statement made in this lemma holds. Note that any minor of $\mathbf{G}^{(r)}$ is an analytic function of the elements $\mathbf{G}^{(r)}$. Due to Lemma 4.1 we know that when $(I-1)J+1 \geq R$, the matrix $\mathbf{H}^{(r)}$ given by (4.4) generically has rank $R-1$. This allows us to conclude that when $(I-1)J+1 \geq R$, the matrix $\mathbf{G}^{(r)}$ in (4.2) generically has rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$. ∎

A consequence of Lemma 4.3 is the following result, which generalizes the generic rank condition to the case where $\mathbf{A}$ is not Vandermonde.

PROPOSITION 4.4. *Given $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_R] \in \mathbb{C}^{I \times R}$, let the column vectors of $\boldsymbol{U} \in \mathbb{C}^{IJ \times R}$ constitute a basis for range $(\boldsymbol{A} \odot \boldsymbol{B})$ in which $\boldsymbol{B} \in \mathbb{C}^{J \times R}$. If $(I-1)J+1 \geq R$, then the matrix*

$$\boldsymbol{G}^{(r)} = [\boldsymbol{U}, \boldsymbol{a}_r \otimes \boldsymbol{I}_J] \in \mathbb{C}^{IJ \times (R+J)} \tag{4.6}$$

generically has rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$.

Proof. Due to Lemma 4.1 we just need to find one example where the statement made in this proposition holds. By way of example, let $\mathbf{A}$ and $\mathbf{B}$ be Vandermonde matrices and set $\mathbf{U} = \mathbf{A} \odot \mathbf{B}$. Since $(I-1)J+1 \geq R$, the matrix $\mathbf{A} \odot \mathbf{B} \in \mathbb{C}^{IJ \times R}$ generically has full column rank. Due to Lemma 4.3 we know that $\mathbf{G}^{(r)}$ in (4.6) generically has rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$. More generally, any $\mathbf{U}$ of which the columns form a basis for range $(\mathbf{A} \odot \mathbf{B})$ yields a matrix $\mathbf{G}^{(r)}$ of rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$. By invoking Lemma 4.1 we can now conclude that when $(I-1)J+1 \geq R$, the matrix $\mathbf{G}^{(r)}$ in (4.6) generically has rank $R+J-1$ $\forall r \in \{1, \ldots, R\}$. ∎

In the proof of Theorem 4.6 we also make use of the following result.

LEMMA 4.5. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in* (2.1). *Assume that the rth column vector $\boldsymbol{c}_r$ of the factor matrix $\boldsymbol{C}$ is known. If*

$$\begin{cases} \boldsymbol{B} \text{ has full column rank,} \\ r\left([\boldsymbol{C} \odot \boldsymbol{A}, \boldsymbol{c}_r \otimes \boldsymbol{I}_I]\right) = I + R - 1\,, \end{cases} \tag{4.7}$$

then the rth column vector $\boldsymbol{a}_r$ of is $\boldsymbol{A}$ unique.[1] *Generically, condition* (4.7) *is satisfied if*

$$\min\left((K-1)\,I + 1, J\right) \geq R\,. \tag{4.8}$$

Proof. Let us first prove that the deterministic condition (4.7) guarantees the uniqueness of $\mathbf{a}_r$. Note that, trivially, $\mathbf{c}_r \otimes \mathbf{a}_r \in$ range $(\mathbf{c}_r \otimes \mathbf{I}_I)$. Condition (4.7) implies that this is the only dependency between the column vectors of $[\mathbf{C} \odot \mathbf{A}, \mathbf{c}_r \otimes \mathbf{I}_I]$. Hence, we know that $r\,(\mathbf{C} \odot \mathbf{A}) = R$. Given are $\mathbf{X}_{(3)} = (\mathbf{C} \odot \mathbf{A})\,\mathbf{B}^T$ and $\mathbf{c}_r$. We know that, under condition (4.7), both the matrices $\mathbf{C} \odot \mathbf{A}$ and $\mathbf{B}$ have full column rank. Let $\mathbf{X}_{(3)} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{X}_{(3)}$, where $\mathbf{U} \in \mathbb{C}^{KI \times R}$; then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{R \times R}$ such that

$$\mathbf{U}\mathbf{M} = \mathbf{C} \odot \mathbf{A} \Leftrightarrow \mathbf{U}\mathbf{m}_r = \mathbf{c}_r \otimes \mathbf{a}_r, \quad r \in \{1, \ldots, R\}\,. \tag{4.9}$$

[1]The last condition states that $\mathbf{M}_r = [\mathbf{C} \odot \mathbf{A}, \mathbf{c}_r \otimes \mathbf{I}_I]$ has the maximal rank possible. Note that $\mathbf{M}_r$ has at least a one-dimensional kernel, since $[\mathbf{n}_r^T, \mathbf{a}_r^T]^T \in \ker(\mathbf{M}_r)$ for some $\mathbf{n}_r \in \mathbb{C}^R$.

Relation (4.9) can also be written as

$$\mathbf{G}^{(r)} \begin{bmatrix} \mathbf{m}_r \\ \mathbf{a}_r \end{bmatrix} = \mathbf{0}_{KI}\,, \quad r \in \{1, \ldots, R\}\,, \tag{4.10}$$

where

$$\mathbf{G}^{(r)} = [\mathbf{U}, -\mathbf{c}_r \otimes \mathbf{I}_I] \in \mathbb{C}^{KI \times (I+R)}. \tag{4.11}$$

Since we assume that $\mathbf{G}^{(r)}$ has a one-dimensional kernel, the rth column vector of $\mathbf{A}$ follows from (4.10). Indeed, let $\mathbf{y}_r \in \mathbb{C}^{I+R}$ denote the right singular vector associated with the zero singular value of the matrix in (4.11); then $\mathbf{a}_r = \mathbf{y}_r\,(R+1 : R+I)$.

Let us now prove that condition (4.8) generically guarantees the uniqueness of the column vector $\mathbf{a}_r$. Given are $\mathbf{X}_{(3)} = (\mathbf{C} \odot \mathbf{A})\,\mathbf{B}^T$ and $\mathbf{c}_r$. If $J \geq R$, then $\mathbf{B}$ generically has full column rank. Due to Lemma 4.2 we also know that since $KI \geq R$, $\mathbf{C} \odot \mathbf{A}$ generically has full column rank. Finally, Proposition 4.4 tells us that if $(K-1)\,J + 1 \geq R$, then the rank of $\mathbf{G}^{(r)}$ in (4.11) is generically equal to $I + R - 1$. Hence, generically we can work as in the proof of the deterministic case. ∎

THEOREM 4.6. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ in* (2.1). *Assume that the factor matrix $\boldsymbol{C}$ is known. If*

$$\begin{cases} \boldsymbol{B} \textit{ has full column rank}, \\ r\left([\boldsymbol{C} \odot \boldsymbol{A}, \boldsymbol{c}_r \otimes \boldsymbol{I}_I]\right) = I + R - 1 \quad \forall r \in \{1, \ldots, R\}\,, \end{cases} \tag{4.12}$$

then the PD of $\mathcal{X}$ with $\boldsymbol{C}$ known is unique. Generically, condition (4.12) *is satisfied if*

$$\min\left((K-1)\,I + 1, J\right) \geq R\,. \tag{4.13}$$

Proof. Let us first prove that condition (4.12) guarantees the uniqueness of the CPD of $\mathcal{X}$ with $\mathbf{C}$ known. Lemma 4.5, together with the condition (4.12), guarantees the uniqueness of the factor matrix $\mathbf{A}$. Recall also from the proof of Lemma 4.5 that the matrix $\mathbf{C} \odot \mathbf{A}$ has full column rank when condition (4.12) holds. Hence, $\mathbf{B}$ follows from

$$\mathbf{B}^T = (\mathbf{C} \odot \mathbf{A})^{\dagger}\, \mathbf{X}_{(3)} = \left(\left(\mathbf{C}^H \mathbf{C}\right) * \left(\mathbf{A}^H \mathbf{A}\right)\right)^{-1} (\mathbf{C} \odot \mathbf{A})^H\, \mathbf{X}_{(3)}\,.$$

The proof that condition (4.13) generically guarantees the uniqueness of the CPD of $\mathcal{X}$ with $\mathbf{C}$ is analogous to the proof of the generic part of Lemma 4.5. ∎

An important remark concerning the proof of Theorem 4.6 is that it is constructive, i.e., it provides us with an algorithm to compute a PD with a known factor. In other words, if one of the factor matrices of the PD is known, say $\mathbf{C}$, and the conditions stated in Theorem 4.6 are satisfied, then even if the known factor matrix does not have full column rank (as opposed to Proposition 3.1), we can find the remaining factor matrices from $\mathbf{X}_{(3)}$. An outline of the proposed method for computing a PD with a known factor matrix is given in Algorithm 1.

Note that Theorem 4.6 does not prevent $k_{\mathbf{A}} = 1$. This is a remarkable difference compared to unconstrained CPD, where $k_{\mathbf{A}} \geq 2$ is necessary for uniqueness (e.g., [27]). Note also that cases where $\mathbf{C}$ does not have full column rank and $k_{\mathbf{A}} = 1$ are not covered by Propositions 3.1 and 3.2. More precisely, if $r_{\mathbf{C}} < R$, Proposition 3.1 does not apply. Likewise, if $k_{\mathbf{C}} < R$ and $k_{\mathbf{A}} = 1$, Proposition 3.1 does not apply either. Theorem 4.6 also leads to improved generic bounds. As an example, let $I = 4$, $J = R$, $K = 3$. Proposition 3.1 generically requires that $R \leq K = 3$, while Proposition 3.1 generically requires that $R \leq 6$. Theorem 4.6 relaxes the generic bound to $R \leq 9$.

Algorithm 1. Computation of PD with known factor matrix based on Theorem 4.6.

Input: $\mathbf{X}_{(3)} = (\mathbf{C} \odot \mathbf{A})\,\mathbf{B}^T$ and $\mathbf{C}$.
  Find $\mathbf{U}$ whose column vectors constitute an orthonormal basis for range$\,(\mathbf{C} \odot \mathbf{A})$.
  Solve set of linear equations $[\mathbf{U}, -\mathbf{c}_r \otimes \mathbf{I}_I]\,\mathbf{y}_r = \mathbf{0}_{KI}, \quad r \in \{1, \ldots, R\}$.
  Set $\mathbf{a}_r = \mathbf{y}_r\,(R+1 : R+I), \quad r \in \{1, \ldots, R\}$.
  Compute $\mathbf{B}^T = \left((\mathbf{C}^H\mathbf{C}) * (\mathbf{A}^H\mathbf{A})\right)^{-1} (\mathbf{C} \odot \mathbf{A})^H\,\mathbf{X}_{(3)}$.
Output: $\mathbf{A}$ and $\mathbf{B}$.

4.2. None of the factor matrices is required to have full column rank. Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ with matrix representation

$$\mathbf{X}_{(1)} = (\mathbf{A} \odot \mathbf{B})\,\mathbf{C}^T \tag{4.14}$$

in which $\mathbf{C}$ is known. In this section we extend Theorem 4.6 and the associated Algorithm 1 to cases where none of the involved factor matrices in (4.14) has full column rank. In order to make the extension clear we outline the idea before going through the technical steps. The main idea is to reduce the PD problem in (4.14) into a PD problem with a known factor matrix and a factor matrix that has full column rank, as in Theorem 4.6. This will be accomplished by partitioning the PD in (4.14) into two parts as follows:

$$\mathbf{X}_{(1)} = (\mathbf{A}^{(S)} \odot \mathbf{B}^{(S)})\mathbf{C}^{(S)T} + (\mathbf{A}^{(S^c)} \odot \mathbf{B}^{(S^c)})\mathbf{C}^{(S^c)T}, \tag{4.15}$$

where $[\mathbf{A}^{(S)}, \mathbf{A}^{(S^c)}]$ is a column permuted version of $\mathbf{A}$ (similarly for $\mathbf{B}$ and $\mathbf{C}$). If the partitioning can be chosen such that $\mathbf{C}^{(S)}$ and $\mathbf{B}^{(S^c)}$ have full column rank, then by first projecting the rows of $\mathbf{X}_{(1)}$ onto the orthogonal complement of the row space of $\mathbf{C}^{(S)}$ we cancel the first term $(\mathbf{A}^{(S)} \odot \mathbf{B}^{(S)})\mathbf{C}^{(S)T}$ in (4.15). The uniqueness of the second term $(\mathbf{A}^{(S^c)} \odot \mathbf{B}^{(S^c)})\mathbf{C}^{(S^c)T}$ in (4.15) can subsequently be established via Theorem 4.6 and computed via Algorithm 1. Finally, by subtracting the latter term from (4.15) we can establish uniqueness of the former term via Proposition 3.1, while the computation can be carried out via rank-1 approximations.

The technical derivation is organized as follows. In Theorem 4.8 we explain that by an appropriate preprocessing step it is possible to derive a relaxed version of Theorem 4.6 in which neither $\mathbf{A}$ nor $\mathbf{B}$ is required to have full column rank. Lemma 4.7 below generalizes Lemma 4.5 to the case where none of the factor matrices is required to have full column rank. Lemma 4.7 makes use of the subsets S, S^c, T, and U of $\{1, \ldots, R\}$. They satisfy the relations $S \subseteq T \subseteq \{1, \ldots, R\}$, $S^c = \{1, \ldots, R\} \setminus S$, and $U = T \setminus S \subseteq S^c$. The relations among the sets are visualized in Figure 1.

LEMMA 4.7. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in* (2.1). *Assume that the columns of $\boldsymbol{C}$ indexed by the elements of the set T with card$(T) = Q$ are known. Stack the columns of $\boldsymbol{C}$ with index in T in $\boldsymbol{C}^{(T)} \in \mathbb{C}^{K\times card(T)}$. Let S denote a subset of T and let $S^c = \{1, \ldots, R\} \setminus S$. Stack the columns of $\boldsymbol{C}$ with index in S in $\boldsymbol{C}^{(S)} \in \mathbb{C}^{K\times card(S)}$ and stack the columns of $\boldsymbol{C}$ with index in S^c in $\boldsymbol{C}^{(S^c)} \in \mathbb{C}^{K\times(R-card(S))}$. Stack the columns of $\boldsymbol{A}$ (resp., $\boldsymbol{B}$) in the same order such that $\boldsymbol{A}^{(S)} \in \mathbb{C}^{I\times card(S)}$ (resp., $\boldsymbol{B}^{(S)} \in \mathbb{C}^{J\times card(S)}$) and $\boldsymbol{A}^{(S^c)} \in \mathbb{C}^{I\times(R-card(S))}$ (resp., $\boldsymbol{B}^{(S^c)} \in \mathbb{C}^{J\times(R-card(S))}$) are*

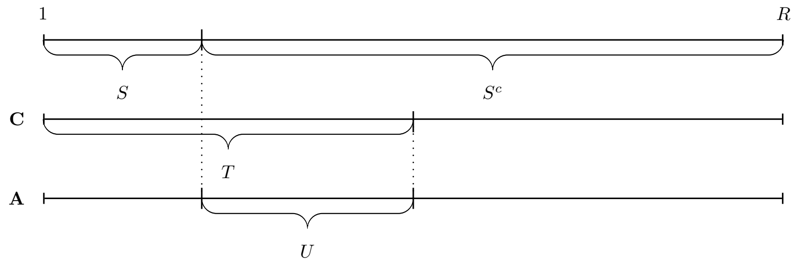

FIG. 1. *The upper line depicts the disjoint partitioning of the set* $\{1,\ldots,R\}$ *into* S *and* S^c. *The middle line depicts the known columns of* $\boldsymbol{C}$ *indexed by the elements in the set* T. *The bottom line depicts the columns of* $\boldsymbol{A}$ *indexed by the elements in the set* $U = T \setminus S$ *for which uniqueness will be established in Proposition* 4.7.

obtained. If there exists a subset $S \subseteq T$ *with* $0 \leq \mathit{card}\,(S) \leq r_{\boldsymbol{C}^{(T)}}$ *such that*[2]

$$
\begin{cases}
\boldsymbol{C}^{(S)} \ \textit{has full column rank}, \\
\boldsymbol{B}^{(S^c)} \ \textit{has full column rank}, \\
r\left(\left[\left(\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{C}^{(S^c)}\right) \odot \boldsymbol{A}^{(S^c)}, (\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{c}_r) \otimes \boldsymbol{I}_I\right]\right) = I + R - \mathit{card}\,(S) - 1 \ \ \forall r \in U,
\end{cases}
\tag{4.16}
$$

where $U = T \setminus S \subseteq S^c$, *then the column vectors* $\{\boldsymbol{a}_r\}_{r\in U}$ *of* $\boldsymbol{A}$ *are unique. Generically, condition* (4.16) *is satisfied if*

$$
\begin{cases}
R \leq \min\left(J + \min\left(K,Q\right), \frac{J(I-1)+I(K-1)+1}{I}\right) & \textit{when} \quad J < R\,, \\
R \leq (K-1)I + 1 & \textit{when} \quad J \geq R\,.
\end{cases}
\tag{4.17}
$$

Proof. Let us first prove the deterministic part of the proposition. Without loss of generality, we assume that $\mathbf{C}(1 : \mathrm{card}\,(S), 1 : \mathrm{card}\,(S))$ is nonsingular, i.e., we set $\mathbf{C}^{(S)} = \mathbf{C}(:, 1 : \mathrm{card}\,(S))$. We first project on the orthogonal complement of range$(\mathbf{C}^{(S)})$ to cancel card (S) rank-1 terms in the PD of $\mathcal{X}$. That is, we compute

$$
\mathbf{Y}_{(1)} = \mathbf{X}_{(1)}\mathbf{P}^T_{\mathbf{C}^{(S)}} = \left(\mathbf{A}^{(S^c)} \odot \mathbf{B}^{(S^c)}\right)\mathbf{C}^{(S^c)T}\mathbf{P}^T_{\mathbf{C}^{(S)}},
$$

where the relation $\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C} = \mathbf{P}_{\mathbf{C}^{(S)}}[\mathbf{C}^{(S)}, \mathbf{C}^{(S^c)}] = [\mathbf{0}_{K,\mathrm{card}(S)}, \mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}]$ was used. The tensor $\mathcal{Y}$ represented by $\mathbf{Y}_{(1)}$ also has matrix representation

$$
\mathbf{Y}_{(3)} = \left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)} \odot \mathbf{A}^{(S^c)}\right)\mathbf{B}^{(S^c)T}.
$$

We have assumed that $r([(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}, (\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r) \otimes \mathbf{I}_I]) = I + R - \mathrm{card}\,(S) - 1$. This implies that $r((\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}) = R - \mathrm{card}\,(S)$, i.e., $(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}$ has full column rank. We have also assumed that $\mathbf{B}^{(S^c)}$ has full column rank. Let $\mathbf{Y}_{(3)} = \mathbf{U}\Sigma\mathbf{V}^H$ denote the compact SVD of $\mathbf{Y}_{(3)}$ in which $\mathbf{U} \in \mathbb{C}^{KI \times (R-\mathrm{card}(S))}$; then there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{(R-\mathrm{card}(S)) \times (R-\mathrm{card}(S))}$ such that

$$
\mathbf{U}\mathbf{M} = \left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}\right) \odot \mathbf{A}^{(S^c)}
$$

[2]The last condition states that $\mathbf{M}_r = [(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}, (\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r^{(S^c)}) \otimes \mathbf{I}_I]$ has the maximal rank possible. Note that $\mathbf{M}_r$ has at least a one-dimensional kernel for every $r \in U$, since $[\mathbf{n}_r^T, \mathbf{a}_r^{(S^c)T}]^T \in \ker(\mathbf{M}_r)$ for some $\mathbf{n}_r \in \mathbb{C}^{\mathrm{card}(S^c)}$.

and

$$\mathbf{G}^{(r,S^c)} \begin{bmatrix} \mathbf{m}_{\mu(r)} \\ \mathbf{a}_r \end{bmatrix} = \mathbf{0}_{KI}\,, \quad r \in S^c, \tag{4.18}$$

in which $\mathbf{M} = \left[\mathbf{m}_{\mu(1)}, \ldots, \mathbf{m}_{\mu(\mathrm{card}(S^c))}\right]$, and $\mathbf{G}^{(r,S^c)} \in \mathbb{C}^{KI \times (I+R-\mathrm{card}(S))}$ is given by

$$\mathbf{G}^{(r,S^c)} = \left[\mathbf{U}, -\left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r\right) \otimes \mathbf{I}_I\right]\,, \quad r \in S^c. \tag{4.19}$$

We have assumed that $\mathbf{G}^{(r,S^c)}$ has rank $I + R - \mathrm{card}\,(S) - 1$ $\forall r \in U \subseteq S^c$, which implies that $\mathbf{G}^{(r,S^c)}$ has a one-dimensional kernel for every $r \in U$. This in turn means that we can obtain the column vectors $\{\mathbf{a}_r\}_{r\in U}$ from (4.18) as follows. Let $\mathbf{z}_r \in \mathbb{C}^{(I+R-\mathrm{card}(S))}$ denote the right singular vector associated with the zero singular value of the matrix given by (4.19); then

$$\mathbf{a}_r = \mathbf{z}_r\left(R - \mathrm{card}\,(S) + 1 : R - \mathrm{card}\,(S) + I\right)\,, \quad r \in U.$$

Let us now prove that condition (4.17) generically guarantees the uniqueness of the column vectors $\{\mathbf{a}_r\}_{r\in U}$.

Consider first the case where $J \geq R$. In that case we choose $S = \emptyset$ and consequently also $\mathbf{P}_{\mathbf{C}^{(S)}} = \mathbf{I}_K$. We know that $\mathbf{B}$ generically has full column rank when $J \geq R$. Due to Proposition 4.4 we also know that $[\mathbf{C} \odot \mathbf{A}, \mathbf{c}_r \otimes \mathbf{I}_I]$ generically has rank $R + I - 1$ when $(K-1)I + 1 \geq R$. This proves the second leg of (4.17).

Consider now the case where $J < R$. Due to Lemma 4.1 we need only find one example for which the first leg of (4.17) generically guarantees the uniqueness of the column vectors $\{\mathbf{a}_r\}_{r\in U}$. By way of example, we let

$$\mathbf{C} = \left[\begin{array}{c|c} \mathbf{C}^{(S)} & \mathbf{C}^{(S^c)} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{I}_{\mathrm{card}(S)} & \\ \mathbf{0}_{K-\mathrm{card}(S),\mathrm{card}(S)} & \mathbf{C}^{(S^c)} \end{array}\right]. \tag{4.20}$$

Then

$$\mathbf{P}_{\mathbf{C}^{(S)}} = \mathbf{I}_K - \left[\begin{array}{c|c} \mathbf{I}_{\mathrm{card}(S)} & \mathbf{0}_{\mathrm{card}(S),K-\mathrm{card}(S)} \\ \mathbf{0}_{K-\mathrm{card}(S),\mathrm{card}(S)} & \mathbf{0}_{K-\mathrm{card}(S),K-\mathrm{card}(S)} \end{array}\right]$$

and

$$\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)} = \begin{bmatrix} \mathbf{0}_{\mathrm{card}(S),R-\mathrm{card}(S)} \\ \mathbf{D} \end{bmatrix}, \quad \mathbf{D} \in \mathbb{C}^{(K-\mathrm{card}(S))\times(R-\mathrm{card}(S))}.$$

Note that $\mathrm{card}\,(S) = r_{\mathbf{C}^{(T)}} - m$ for some integer m with property $0 \leq m \leq r_{\mathbf{C}^{(T)}}$. We know that $\mathbf{B}^{(S^c)} \in \mathbb{C}^{J\times(R-\mathrm{card}(S))}$ generically has full column rank if

$$J \geq R - \mathrm{card}\,(S) = R - r_{\mathbf{C}^{(T)}} + m \Leftrightarrow J - R + r_{\mathbf{C}^{(T)}} \geq m. \tag{4.21}$$

Note that, since $J < R$, the condition $m \leq r_{\mathbf{C}^{(T)}}$ is automatically satisfied when (4.21) holds. In order to ensure that $m \geq 0$, the following inequality must be satisfied:

$$J + r_{\mathbf{C}^{(T)}} \geq R\,, \tag{4.22}$$

where $r_{\mathbf{C}^{(T)}} = \min(K, Q)$. Due to the structure of $\mathbf{C}$ in (4.20), the problem of determining the generic rank of

$$\left[\left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}\right) \odot \mathbf{A}^{(S^c)}, \left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r\right) \otimes \mathbf{I}_I\right]$$

reduces to finding the generic rank of

$$\left[\mathbf{D} \odot \mathbf{A}^{(S^c)}, \mathbf{d}_r \otimes \mathbf{I}_I\right] \in \mathbb{C}^{(K-\mathrm{card}(S))I \times (R-\mathrm{card}(S)+I)}. \tag{4.23}$$

Proposition 4.4 tells us that the matrix given by (4.23) generically has rank $R - \mathrm{card}\,(S) + I - 1$ if

$$(K - \mathrm{card}\,(S) - 1)I + 1 \geq R - \mathrm{card}\,(S). \tag{4.24}$$

By inserting $\mathrm{card}\,(S) = r_{\mathbf{C}^{(T)}} - m$ into (4.24) we obtain

$$(K - r_{\mathbf{C}^{(T)}} + m - 1)I + 1 \geq R - r_{\mathbf{C}^{(T)}} + m. \tag{4.25}$$

Relation (4.25) can also be expressed as

$$m \geq \frac{R + r_{\mathbf{C}^{(T)}}(I-1) - I(K-1) - 1}{I-1}. \tag{4.26}$$

Combining conditions (4.21) and (4.26) we conclude that if inequalities (4.22) and

$$\frac{R + r_{\mathbf{C}^{(T)}}(I-1) - I(K-1) - 1}{I-1} \leq J - R + r_{\mathbf{C}^{(T)}} \Leftrightarrow R \leq \frac{J(I-1) + I(K-1) + 1}{I}$$

are satisfied, then generically the column vectors $\{\mathbf{a}_r\}_{r\in U}$ are unique. This proves the first leg of (4.17). ∎

Note that there is some flexibility in the choice of $\mathrm{card}\,(S)$ in (4.16). By choosing a large $\mathrm{card}\,(S)$ we relax the constraint on $\mathbf{B}^{(S^c)}$ but also impose a stronger constraint on $\mathbf{C}$. The larger the $\mathrm{card}\,(S)$, the smaller the number of columns of $\mathbf{A}$ for which uniqueness is demonstrated.

THEOREM 4.8. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in* (2.1). *Assume that $\boldsymbol{C}$ is known. Let S denote a subset of $\{1,\ldots,R\}$ and let $S^c = \{1,\ldots,R\}\setminus S$ denote the complementary set. Stack the columns of $\boldsymbol{C}$ with index in S in $\boldsymbol{C}^{(S)} \in \mathbb{C}^{K\times card(S)}$ and stack the columns of $\boldsymbol{C}$ with index in S^c in $\boldsymbol{C}^{(S^c)} \in \mathbb{C}^{K\times(R-card(S))}$. Stack the columns of $\boldsymbol{A}$ (resp., $\boldsymbol{B}$) in the same order such that $\boldsymbol{A}^{(S)} \in \mathbb{C}^{I\times card(S)}$ (resp., $\boldsymbol{B}^{(S)} \in \mathbb{C}^{J\times card(S)}$) and $\boldsymbol{A}^{(S^c)} \in \mathbb{C}^{I\times(R-card(S))}$ (resp., $\boldsymbol{B}^{(S^c)} \in \mathbb{C}^{J\times(R-card(S))}$) are obtained. If there exists a subset S of $\{1,\ldots,R\}$ with $0 \leq card\,(S) \leq r_{\boldsymbol{C}}$ such that $\boldsymbol{C}^{(S)}$ has full column rank and*

$$\begin{cases} \boldsymbol{B}^{(S^c)} \text{ has full column rank}, \\ r\left(\left[\left(\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{C}^{(S^c)}\right)\odot \boldsymbol{A}^{(S^c)}, (\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{c}_r)\otimes \boldsymbol{I}_I\right]\right) = \alpha \quad \forall r\in S^c, \end{cases} \tag{4.27a}$$

where $\alpha = I + R - card\,(S) - 1$, or

$$\begin{cases} \boldsymbol{A}^{(S^c)} \text{ has full column rank}, \\ r\left(\left[\left(\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{C}^{(S^c)}\right)\odot \boldsymbol{B}^{(S^c)}, (\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{c}_r)\otimes \boldsymbol{I}_J\right]\right) = \beta \quad \forall r\in S^c, \end{cases} \tag{4.27b}$$

where $\beta = J + R - card\,(S) - 1$, then the PD of $\mathcal{X}$ with $\boldsymbol{C}$ known is unique. Generically, condition (4.27a) *or* (4.27b) *is satisfied if*

$$\begin{cases} R \leq \min\left(W + \min\,(K,R), \frac{W(V-1)+V(K-1)+1}{V}\right) & \text{when} \quad W < R\,, \\ R \leq (K-1)V + 1 & \text{when} \quad W \geq R\,, \end{cases} \tag{4.28}$$

where $V = \min(I, J)$ *and* $W = \max(I, J)$.

Proof. Let us first prove that condition (4.27a) guarantees the uniqueness of the CPD of $\mathcal{X}$ with $\mathbf{C}$ known. We work as follows. The fact that $\mathbf{C}$ is known allows us to project away the part of the CPD of $\mathcal{X}$ that corresponds to S. After finding $\mathbf{A}^{(S^c)}$ and $\mathbf{B}^{(S^c)}$, we subtract the part of the CPD of $\mathcal{X}$ that corresponds to S^c, which leads to the remaining $\mathbf{A}^{(S)}$ and $\mathbf{B}^{(S)}$.

In more detail, we first determine the matrix $\mathbf{A}^{(S^c)}$ in a way similar to the vectors $\{\mathbf{a}\}_{r\in U}$ in Lemma 4.7. More precisely, we set $Q = R$ and consequently $T = \{1, \ldots, R\}$, $U = \{1, \ldots, R\} \setminus S = S^c$.

The next step is to find $\mathbf{B}^{(S^c)}$. Recall that we assumed that

$$r\left(\left[\left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}\right) \odot \mathbf{A}^{(S^c)}, (\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r) \otimes \mathbf{I}_I\right]\right) = I + R - \operatorname{card}(S) - 1.$$

This implies that $r((\mathbf{P}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}) = R - \operatorname{card}(S)$, i.e., $(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}$ has full column rank. Hence, $\mathbf{B}^{(S^c)}$ follows from the relation

$$\begin{aligned}
\mathbf{B}^{(S^c)T} &= \left(\left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}\right) \odot \mathbf{A}^{(S^c)}\right)^{\dagger}\mathbf{Y}_{(3)} \\
&= \left(\left(\mathbf{C}^{(S^c)H}\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}\right) * \left(\mathbf{A}^{(S^c)H}\mathbf{A}^{(S^c)}\right)\right)^{-1}\left(\mathbf{P}\mathbf{C}^{(S^c)} \odot \mathbf{A}^{(S^c)}\right)^{H}\mathbf{Y}_{(3)}.
\end{aligned}$$

Now that $\{\mathbf{A}^{(S^c)}, \mathbf{B}^{(S^c)}, \mathbf{C}^{(S^c)}\}$ are known, we can compute

$$\mathbf{Q}_{(1)} = \mathbf{X}_{(1)} - \left(\mathbf{A}^{(S^c)} \odot \mathbf{B}^{(S^c)}\right)\mathbf{C}^{(S^c)T} = \left(\mathbf{A}^{(S)} \odot \mathbf{B}^{(S)}\right)\mathbf{C}^{(S)T}.$$

Recall also that the matrix $\mathbf{C}^{(S)}$ is assumed to have full column rank. Compute

$$\mathbf{H} = \mathbf{Q}_{(1)}\left(\mathbf{C}^{(S)T}\right)^{\dagger}.$$

The remaining columns of unknowns $\mathbf{A}^{(S)}$ and $\mathbf{B}^{(S)}$ are obtained by recognizing that the columns of $\mathbf{H}$ are vectorized rank-1 matrices, i.e., $\mathbf{h}_{\sigma(r)} = \mathbf{a}_r \otimes \mathbf{b}_r$, $r \in S$, where $\mathbf{H} = \left[\mathbf{h}_{\sigma(1)}, \ldots, \mathbf{h}_{\sigma(S)}\right]$. The proof for (4.27b) is analogous to the above proof for condition (4.27a).

The proof that condition (4.28) generically guarantees the uniqueness of the CPD of $\mathcal{X}$ with $\mathbf{C}$ known is analogous to the generic part of the proof of Lemma 4.7. Essentially, by replacing Q with R in the proof of the generic condition (4.17) in Lemma 4.7, the result follows. ∎

Note that Theorem 4.8 only requires the existence of one partitioning of $\mathbf{A}$, denoted by $[\mathbf{A}^{(S)}, \mathbf{A}^{(S^c)}]$ (similarly for $\mathbf{B}$ and $\mathbf{C}$), for which the rank conditions (4.27a) or (4.27b) are satisfied. This is different from Kruskal-type conditions (e.g., Proposition 3.2), in which the k-rank depends on all possible submatrices of a certain size. Since Theorem 4.8 requires that $\mathbf{C}^{(S)}$ and $\mathbf{B}^{(S^c)}$ or $\mathbf{A}^{(S^c)}$ have full column rank, an educated guess of what S should be can often be deduced from these dimensionality constraints, i.e., the choice of S should take into account the dimensionality constraints $\operatorname{card}(S) \leq K$ and $\operatorname{card}(S^c) \leq J$ or $\operatorname{card}(S^c) \leq I$. In the generic case, Theorem 4.8 provides us with a very simple way of assessing if a PD with a known factor matrix is expected to be unique. It suffices to check if the dimensionality condition (4.28) is satisfied.

Theorem 4.8 does not prevent that $k_{\mathbf{A}} = 1$ and/or $k_{\mathbf{B}} = 1$ and/or $r_{\mathbf{A}\odot\mathbf{B}} < R$. This is in contrast to unconstrained CPD where $k_{\mathbf{A}} \geq 2$, $k_{\mathbf{B}} \geq 2$, and $r_{\mathbf{A}\odot\mathbf{B}} = R$ are all necessary uniqueness conditions (e.g., [27]). As an example, consider the PD of

$\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in (2.1) in which $\mathbf{C}$ is known, $\mathbf{a}_2 = \mathbf{a}_1$, $\mathbf{b}_2 = \mathbf{b}_1$, $R = 7$, $I = 4$, $J = 4$, and $K = 5$. If apart from these constraints the parameters are randomly drawn, Theorem 4.8 with $S = \{1, 2\}$ and $\operatorname{card}(S) = 2$ guarantees the generic uniqueness of $\mathbf{A}$ and $\mathbf{B}$.

Theorem 4.8 also leads to improved bounds on R in generic cases without collinearities. As an example, we consider the situation where $V = \min(I, J) = 3$, $W = \max(I, J) = 7$, $K = 4$, and $\mathbf{C}$ is known. Proposition 3.1 requires that $R \leq K = 4$, while Proposition 3.2 relaxes the bound to $R \leq 6$. On the other hand, Theorem 4.8 only requires that $R \leq 8$.

We observe that the proof of Theorem 4.8 is constructive. We present the construction as Algorithm 2.

The cardinality of set S in Algorithm 2 must be chosen such that the conditions in Theorem 4.8 are satisfied. As explained in the proof, the choice of $\operatorname{card}(S)$ affects the computation in the sense that we first compute the rank-$\operatorname{card}(S^c)$ PD $\mathcal{Y} = \sum_{r\in S^c} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$, and thereafter the rank-$\operatorname{card}(S)$ PD $\mathcal{Q} = \sum_{r\in S} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$. S must be chosen such that $\mathbf{C}^{(S)}$ has full column rank, which is easy to check. S must also be chosen such that $\mathbf{B}^{(S^c)}$ has full column rank, implying that $\operatorname{card}(S) \geq R - J$ must be satisfied. This condition can numerically be verified by checking if the effective rank of $\mathbf{Y}_{(3)}$ is equal to $\operatorname{card}(S^c)$. S must also be chosen such that

$$\left[\left(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}\right) \odot \mathbf{A}^{(S^c)}, (\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r) \otimes \mathbf{I}_I\right]$$

has a one-dimensional kernel for every $r \in S^c$. This condition can numerically be verified by checking if the effective dimension of the kernel of $[\mathbf{U}, -(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r) \otimes \mathbf{I}_I]$ is one for every $r \in S^c$.

In the supplementary material we briefly explain how to extend Theorem 4.8 and Algorithm 2 to tensors of arbitrary order.

Algorithm 2. Computation of PD with known factor matrix based on Theorem 4.8.

Input: $\mathbf{X}_{(1)} = (\mathbf{A} \odot \mathbf{B})\,\mathbf{C}^T$ and $\mathbf{C}$

1. Determine $r_{\mathbf{C}}$.
2. Choose sets $S \subseteq \{1, \ldots, R\}$ and $S^c = \{1, \ldots, R\}\backslash S$ subject to $0 \leq \operatorname{card}(S) \leq r_{\mathbf{C}}$.
3. Build $\mathbf{C}^{(S)}$ and $\mathbf{C}^{(S^c)}$.
4. Find $\mathbf{F}$ whose column vectors constitute an orthonormal basis for $\left(\mathbf{C}^{(S)}\right)$.
5. Compute $\mathbf{P}_{\mathbf{C}^{(S)}} = \mathbf{I}_K - \mathbf{F}\mathbf{F}^H$, $\mathbf{D}^{(S^c)} = \mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}$, and $\mathbf{E}^{(S^c)} = \mathbf{C}^{(S^c)H}\mathbf{D}^{(S^c)}$.
6. Compute $\mathbf{Y}_{(1)} = \mathbf{X}_{(1)}\mathbf{P}^T_{\mathbf{C}^{(S)}}$.
7. Build $\mathbf{Y}_{(3)}$.
8. Find $\mathbf{U}$ whose column vectors constitute an orthonormal basis for range $\left(\mathbf{Y}_{(3)}\right)$.
9. Solve set of linear equations $[\mathbf{U}, -(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r) \otimes \mathbf{I}_I]\,\mathbf{z}_r = \mathbf{0}_{KI}, \quad r \in S^c$.
10. Set $\mathbf{a}_r = \mathbf{z}_r\,(R - \operatorname{card}(S) + 1 : R - \operatorname{card}(S) + I)\,, \quad r \in S^c\,$.
11. Compute $\mathbf{B}^{(S^c)T} = \left(\mathbf{E}^{(S^c)} * \left(\mathbf{A}^{(S^c)H}\mathbf{A}^{(S^c)}\right)\right)^{-1}\left(\mathbf{D}^{(S^c)} \odot \mathbf{A}^{(S^c)}\right)^H\mathbf{Y}_{(3)}$.
12. Compute $\mathbf{Q}_{(1)} = \mathbf{X}_{(1)} - \left(\mathbf{A}^{(S^c)} \odot \mathbf{B}^{(S^c)}\right)\mathbf{C}^{(S^c)T}$.
13. Compute $\mathbf{H} = \left[\mathbf{h}_{\sigma(1)}, \ldots, \mathbf{h}_{\sigma(S)}\right] = \mathbf{Q}_{(1)}\left(\mathbf{C}^{(S)T}\right)^{\dagger}$.
14. Compute best rank-1 approximations $\min_{\mathbf{a}_r,\mathbf{b}_r} \left\|\mathbf{h}_{\sigma(r)} - \mathbf{a}_r \otimes \mathbf{b}_r\right\|_F^2\,, \quad r \in S\,$.

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

5. New uniqueness results for overall CPD. By combining the existing result presented in subsection 2.1 with the new results presented in section 4, we will in this section derive new uniqueness conditions for the overall CPD and also some variants that are important in signal processing.

5.1. New uniqueness condition for CPD. We first present a deterministic uniqueness condition for CPD based on Theorem 4.8.

THEOREM 5.1. *Consider the PD of $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ in (2.1). Let S denote a subset of $\{1,\ldots,R\}$ and let $S^c = \{1,\ldots,R\} \setminus S$ denote the complementary set. Stack the columns of $\boldsymbol{C}$ with index in S in $\boldsymbol{C}^{(S)} \in \mathbb{C}^{K\times card(S)}$ and stack the columns of $\boldsymbol{C}$ with index in S^c in $\boldsymbol{C}^{(S^c)} \in \mathbb{C}^{K\times(R-card(S))}$. Stack the columns of $\boldsymbol{A}$ (resp., $\boldsymbol{B}$) in the same order such that $\boldsymbol{A}^{(S)} \in \mathbb{C}^{I\times card(S)}$ (resp., $\boldsymbol{B}^{(S)} \in \mathbb{C}^{J\times card(S)}$) and $\boldsymbol{A}^{(S^c)} \in \mathbb{C}^{I\times(R-card(S))}$ (resp., $\boldsymbol{B}^{(S^c)} \in \mathbb{C}^{J\times(R-card(S))}$) are obtained. If*

$$\mathcal{C}_{R-r_C+2}(\boldsymbol{A}) \odot \mathcal{C}_{R-r_C+2}(\boldsymbol{B}) \text{ has full column rank,} \tag{5.1a}$$

and if there exists a subset $S \subseteq \{1,\ldots,R\}$ with $0 \le card(S) \le r_{\boldsymbol{C}}$ such that $\boldsymbol{C}^{(S)}$ has full column rank and

$$\begin{cases} \boldsymbol{B}^{(S^c)} \text{ has full column rank,} \\ r\left(\left[\left(\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{C}^{(S^c)}\right) \odot \boldsymbol{A}^{(S^c)}, (\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{c}_r^{(S^c)}) \otimes \boldsymbol{I}_I\right]\right) = \alpha \quad \forall r \in S^c, \end{cases} \tag{5.1b}$$

where $\alpha = I + R - card(S) - 1$, or

$$\begin{cases} \boldsymbol{A}^{(S^c)} \text{ has full column rank,} \\ r\left(\left[(\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{C}^{(S^c)}) \odot \boldsymbol{B}^{(S^c)}, (\boldsymbol{P}_{\boldsymbol{C}^{(S)}}\boldsymbol{c}_r^{(S^c)}) \otimes \boldsymbol{I}_J\right]\right) = \beta \quad \forall r \in S^c, \end{cases} \tag{5.1c}$$

where $\beta = J + R - card(S) - 1$, then the rank of $\mathcal{X}$ is R, and the CPD of $\mathcal{X}$ is unique.

Proof. Note that both conditions (5.1b) and (5.1c) imply that $k_{\mathbf{C}} \ge 2$. Indeed, if $k_{\mathbf{C}} = 1$, then (i) $k_{\mathbf{C}^{(S)}} = 1$, (ii) $k_{\mathbf{C}^{(S^c)}} = 1$, or (iii) $k_{[\mathbf{c}_m,\mathbf{c}_n]} = 1$ for some $m \in S$ and $n \in S^c$. Since it is assumed that $\mathbf{C}^{(S)}$ has full column rank (i.e., $k_{\mathbf{C}^{(S)}} = \text{card}(S)$), we can rule out the first case (i). The second and third cases imply that the kernel of $[(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{A}^{(S^c)}, (\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r^{(S^c)}) \otimes \mathbf{I}_I]$ or $[(\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{C}^{(S^c)}) \odot \mathbf{B}^{(S^c)}, (\mathbf{P}_{\mathbf{C}^{(S)}}\mathbf{c}_r^{(S^c)}) \otimes \mathbf{I}_J]$ has dimension greater than one for some $r \in S^c$. Hence, cases (ii) and (iii) can also be ruled out. Theorem 2.1, together with the inequality $k_{\mathbf{C}} \ge 2 > 1$, now tells us that the rank of $\mathcal{X}$ is R and that the factor $\mathbf{C}$ is unique. Let us now consider $\mathcal{X}$ as a third-order tensor with known $\mathbf{C}$. Due to Theorem 4.8, the conditions (5.1b) or (5.1c) now guarantee the overall uniqueness of the CPD. ∎

For a discussion of Theorem 5.1, let us first consider the generic case, where factor matrices have maximal rank and k-rank. We notice that, generically, if the one factor matrix uniqueness condition in Theorem 2.1 is satisfied, then the overall uniqueness condition in Theorem 2.2 is automatically satisfied. This means that in the generic case, the one factor matrix uniqueness condition in Theorem 2.1, which corresponds to condition (5.1a), makes that Theorem 5.1 is not more relaxed than Theorem 2.2.

In order to demonstrate that Theorem 5.1 can nevertheless improve Theorem 2.2 in nongeneric cases, we consider the following example taken from [9]. Consider the third-order tensor $\mathcal{X} \in \mathbb{C}^{4\times4\times4}$ with matrix representation $\mathbf{X}_{(1)} = (\mathbf{A} \odot \mathbf{B})\,\mathbf{C}^T$ in which $\mathbf{A} \in \mathbb{C}^{4\times5}$, $\mathbf{B} \in \mathbb{C}^{4\times5}$, and $\mathbf{C} \in \mathbb{C}^{4\times5}$, where $[\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4] = \mathbf{I}_4$ and $a_{25} = 0$, $[\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3,\mathbf{b}_4] = \mathbf{I}_4$ and $b_{15} = 0$, and $[\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3,\mathbf{c}_4] = \mathbf{I}_4$ and $c_{25} = 0$. For a random choice of parameter entries we have $k_{\mathbf{A}} = 3$, $k_{\mathbf{B}} = 3$, and $k_{\mathbf{C}} = 3$. On the other hand, we also have $r_{\mathbf{A}} = 4$, $r_{\mathbf{B}} = 4$, and $r_{\mathbf{C}} = 4$. For this problem the conditions in Theorem 2.2 are not satisfied. However, the conditions (5.1a)–(5.1b) in Theorem 5.1 are generically satisfied if $\mathrm{card}(S) = 2$ and $S = \{1, 2\}$. Thus, overall uniqueness of the CPD of $\mathcal{X}$ can be established by Theorem 5.1 but not by the related Theorem 2.2. We mention that CPD uniqueness of $\mathcal{X}$ has already been established in [9] by other means. A nice property of Theorem 5.1 is that it can be extended to tensor decompositions other than the third-order CPD. For instance, in [25, 26] we demonstrate how Theorem 5.1 can be adapted to coupled CPD and CPD of tensors of arbitrary order. As a final remark, we recall that Theorem 2.1 is not necessary for one factor uniqueness. For this reason, it may be possible to further generalize Theorem 5.1 by combining Theorem 4.8 with a more relaxed single factor matrix uniqueness condition.

5.2. New uniqueness condition for CPD with partial Hermitian symmetry. We say that the CPD of a tensor $\mathcal{X} \in \mathbb{C}^{I\times I\times K}$ has partial Hermitian symmetry if $\mathbf{X}_{(1)} = (\mathbf{A} \odot \mathbf{A}^*)\,\mathbf{C}^T$. The Hermitian symmetry-constrained rank of $\mathcal{X}$ is equal to the minimal number of rank-1 terms $\mathbf{a}_r \circ \mathbf{a}_r^* \circ \mathbf{c}_r$ that yield $\mathcal{X}$ in a linear combination. This structure is very common in signal separation applications. For instance, tensors with partial Hermitian symmetry are obtained by stacking sets of covariance matrices of complex data; see [2, 7] for references to concrete applications. A uniqueness condition for the Hermitian symmetry-constrained CPD of a tensor was provided in [28] for the special case where $\mathbf{C}$ has full column rank. We will now provide a new uniqueness condition. In contrast to [28] we do not require that $\mathbf{C}$ has full column rank. In fact, by better exploiting the Hermitian symmetry, $\mathbf{C}$ may even contain collinear columns, and even $K = 1$ can be admitted in some cases. Let us define the tensor $\mathcal{Y} \in \mathbb{C}^{I\times I\times K}$ by $y_{ijk} = x_{jik}^*$ for all indices. We have $\mathcal{Z} \in \mathbb{C}^{I\times I\times 2K}$ with matrix representation

$$\mathbf{Z}_{(1)} = \left[\mathbf{X}_{(1)}, \mathbf{Y}_{(1)}\right] = (\mathbf{A} \odot \mathbf{A}^*)\,\mathbf{D}^T \in \mathbb{C}^{I^2\times 2K}, \tag{5.2}$$

where

$$\mathbf{D} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}^* \end{bmatrix} \in \mathbb{C}^{2K\times R}. \tag{5.3}$$

We may now study the uniqueness of the CPD of $\mathcal{Z}$ in order to obtain uniqueness results for $\mathcal{X}$. From relation (5.2) it is clear that if $\mathbf{A}$ has full column rank, then the Hermitian symmetry-constrained CPD of $\mathcal{X}$ is unique if $k_{\mathbf{D}} \geq 2$. Note that the latter condition does not prevent $k_{\mathbf{C}} = 1$ or even $K = 1$. Theorem 5.2 is an adaption of Theorem 5.1 to the Hermitian symmetry-constrained CPD case.

THEOREM 5.2. *Consider the PD of* $\mathcal{X} \in \mathbb{C}^{I\times I\times K}$ *with matrix representation* $\boldsymbol{X}_{(1)} = (\boldsymbol{A} \odot \boldsymbol{A}^*)\,\boldsymbol{C}^T$. *Consider also* $\boldsymbol{D}$ *in* (5.3). *Let* S *denote a subset of* $\{1,\ldots,R\}$ *and let* $S^c = \{1,\ldots,R\} \setminus S$ *denote the complementary set. Stack the columns of* $\boldsymbol{D}$ *with index in* S *in* $\boldsymbol{D}^{(S)} \in \mathbb{C}^{2K\times card(S)}$, *and stack the columns of* $\boldsymbol{D}$ *with index in* S^c *in* $\boldsymbol{D}^{(S^c)} \in \mathbb{C}^{2K\times(R-card(S))}$. *Stack the columns of* $\boldsymbol{A}$ *in the same order such that* $\boldsymbol{A}^{(S)} \in \mathbb{C}^{I\times card(S)}$ *and* $\boldsymbol{A}^{(S^c)} \in \mathbb{C}^{I\times(R-card(S))}$ *are obtained. If*

$$\mathcal{C}_{R-r_{\boldsymbol{D}}+2}(\boldsymbol{A}^*) \odot \mathcal{C}_{R-r_{\boldsymbol{D}}+2}(\boldsymbol{A}) \ \textit{has full column rank}, \tag{5.4a}$$

and if there exists a subset $S \subseteq \{1,\ldots,R\}$ *with* $0 \leq card(S) \leq r_{\boldsymbol{D}}$ *such that*

$$\text{(5.4b)} \qquad \begin{cases} \boldsymbol{D}^{(S)} \text{ has full column rank,} \\ \boldsymbol{A}^{(S^c)} \text{ has full column rank,} \\ r\left(\left[\left(\boldsymbol{P}_{\boldsymbol{D}^{(S)}}\boldsymbol{D}^{(S^c)}\right)\odot \boldsymbol{A}^{(S^c)}, \left(\boldsymbol{P}_{\boldsymbol{D}^{(S)}}\boldsymbol{d}_r^{(S^c)}\right)\otimes \boldsymbol{I}_I\right]\right) = \alpha \quad \forall r \in S^c, \end{cases}$$

where $\alpha = I + R - card(S) - 1$*, then the Hermitian symmetry-constrained rank of* $\mathcal{X}$ *is* R *and the Hermitian symmetry-constrained CPD of* $\mathcal{X}$ *is unique.*

Proof. Analogous to the proof of Theorem 5.1, condition (5.4b) implies that $k_{\mathbf{D}} \geq 2$. Consequently, conditions (5.4a) and (5.4b) together with Theorem 2.1 imply that the rank of $\mathcal{Z}$ with matrix representation (5.2) is R and that the factor $\mathbf{D}$ is unique. Since $\mathbf{D}$ is unique, $\mathbf{C}$ is unique. Due to Theorem 4.8, we know that if condition (5.4b) is satisfied, $\mathbf{A}$ is also unique. Overall, we obtain that the Hermitian symmetry-constrained CPD of $\mathcal{X}$ is unique and the Hermitian symmetry-constrained rank of $\mathcal{X}$ is R. ▯

Equations (5.2)–(5.3) are the key to understanding why partial Hermitian symmetry allows us to relax uniqueness conditions. Indeed, a condition on $\mathbf{D}$ is typically less restrictive than a condition on $\mathbf{C}$. (Note that generically $r_{\mathbf{C}} = \min(K, R)$, while, on the other hand, generically $r_{\mathbf{D}} = \min(2K, R)$; see the supplementary material.) For that reason the conditions in Theorem 5.2 are more relaxed than the conditions for unconstrained CPD in Theorems 2.2 and 5.1. For instance, Theorem 5.2 does not prevent $k_{\mathbf{C}} = 1$, i.e., it only requires that $k_{\mathbf{D}} \geq 2$. This is in contrast to the unconstrained CPD for which $k_{\mathbf{C}} \geq 2$ is a necessary uniqueness condition (e.g., [27]). Theorem 5.2 is also more relaxed than the existing result presented in [28]. In contrast to [28], Theorem 5.2 allows us to establish uniqueness in cases where $r_{\mathbf{C}} < R$, $k_{\mathbf{C}} = 1$, and even $K = 1$. Again, since Theorem 2.1 is not necessary for one factor uniqueness it may be possible to further generalize Theorem 5.2 by combining Theorem 4.8 with a more relaxed single factor matrix uniqueness condition.

A necessary condition for Hermitian symmetry-constrained CPD uniqueness is that the matrix $\mathbf{A} \odot \mathbf{A}^*$ must have full column rank. Indeed, if $\mathbf{A} \odot \mathbf{A}^*$ does not have full column rank, then for any $\mathbf{x} \in \ker(\mathbf{A} \odot \mathbf{A}^*)$ we obtain from relation (5.2) that $\mathbf{Z}_{(1)} = (\mathbf{A} \odot \mathbf{A}^*)\mathbf{D}^T = (\mathbf{A} \odot \mathbf{A}^*)(\mathbf{D}^T + \mathbf{x}[\mathbf{y}^T, \mathbf{y}^T])$, where $\mathbf{y} \in \mathbb{R}^K$. Since $\mathbf{A} \odot \mathbf{A}^*$ generically has rank $\min(I^2, R)$ [24], this requirement is not very restrictive.

5.3. New uniqueness condition for CPD with columnwise orthonormal factor matrix. In this section we consider the PD of $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times K}$ with matrix representation $\mathbf{X}_{(1)} = (\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)})\mathbf{C}^T$, in which $\mathbf{C}^H\mathbf{C} = \mathbf{I}_R$. The orthogonality constrained rank of a tensor $\mathcal{X}$ is equal to the minimal number of orthogonality constrained ($\mathbf{c}_i^H\mathbf{c}_j = \delta_{ij}$) rank-1 terms that yield $\mathcal{X}$ in a linear combination. In [24] uniqueness of CPD with a columnwise orthonormal factor matrix has been demonstrated under conditions milder than an unconstrained CPD. Orthogonality-constrained CPD is very common in statistical signal processing. Typically, the orthonormal columns of $\mathbf{C}$ model signals are uncorrelated. See [24] for an overview, applications, and references. We will now provide a new uniqueness condition for a CPD with a columnwise orthonormal factor matrix.

THEOREM 5.3. *Consider the PD of* $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times K}$ *with matrix representation*

$\boldsymbol{X}_{(1)} = (\boldsymbol{A}^{(1)} \odot \boldsymbol{A}^{(2)})\boldsymbol{C}^T$, in which $\boldsymbol{C}$ is columnwise orthonormal. Let

$$N_{\max} = \begin{cases} 2 & \text{if} \quad I_2 \geq I_1, \\ 1 & \text{if} \quad I_1 > I_2, \end{cases} \qquad N_{\min} = \begin{cases} 2 & \text{if} \quad I_2 < I_1, \\ 1 & \text{if} \quad I_1 \leq I_2, \end{cases} \tag{5.5}$$

$$I_{\max} = \max(I_1, I_2), \qquad I_{\min} = \min(I_1, I_2)\,. \tag{5.6}$$

Construct $\boldsymbol{D} = \boldsymbol{A}^{(N_{\min})*} \odot \boldsymbol{A}^{(N_{\min})} \in \mathbb{C}^{I_{\min}^2 \times R}$. Let S denote a subset of $\{1,\ldots,R\}$ and let $S^c = \{1,\ldots,R\} \setminus S$ denote the complementary set. Stack the columns of $\boldsymbol{D}$ with index in S in $\boldsymbol{D}^{(S)} \in \mathbb{C}^{I_{\min}^2 \times card(S)}$ and stack the columns of $\boldsymbol{D}$ with index in S^c in $\boldsymbol{D}^{(S^c)} \in \mathbb{C}^{I_{\min}^2 \times (R-card(S))}$. Stack the columns of $\boldsymbol{A}^{(N_{\max})}$ in the same order such that $\boldsymbol{A}^{(N_{\max},S)} \in \mathbb{C}^{I_{\max} \times card(S)}$ and $\boldsymbol{A}^{(N_{\max},S^c)} \in \mathbb{C}^{I_{\max} \times (R-card(S))}$ are obtained. If

$$\mathcal{C}_{R-r_{\boldsymbol{D}}+2}\big(\boldsymbol{A}^{(N_{\max})*}\big) \odot \mathcal{C}_{R-r_{\boldsymbol{D}}+2}\big(\boldsymbol{A}^{(N_{\max})}\big) \text{ has full column rank,} \tag{5.7a}$$

and if there exists a subset $S \subseteq \{1,\ldots,R\}$ with $0 \leq card(S) \leq r_{\boldsymbol{D}}$ such that

$$\begin{cases} \boldsymbol{D}^{(S)} \text{ has full column rank,} \\ \boldsymbol{A}^{(N_{\max},S^c)} \text{ has full column rank,} \\ r\left(\left[\left(\boldsymbol{P}_{\boldsymbol{D}^{(S)}}\boldsymbol{D}^{(S^c)}\right) \odot \boldsymbol{A}^{(N_{\max},S^c)}, \left(\boldsymbol{P}_{\boldsymbol{D}^{(S)}}\boldsymbol{d}_r^{\ (S^c)}\right) \otimes \boldsymbol{I}_{I_{\max}}\right]\right) = \alpha \quad \forall r \in S^c, \end{cases} \tag{5.7b}$$

where $\alpha = I_{\max} + R - card(S) - 1$, or

$$r\left(\left[\boldsymbol{A}^{(N_{\min})} \odot \boldsymbol{A}^{(N_{\max})}, \boldsymbol{a}_r^{(N_{\min})} \otimes \boldsymbol{I}_{I_{\max}}\right]\right) = I_{\max} + R - 1 \quad \forall r \in \{1,\ldots,R\}, \tag{5.7c}$$

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

Proof. Denote $\mathbf{X}^{(i_1\cdot\cdot)} = \mathcal{X}(i_1,:,:)$ and construct the fourth-order tensor $\mathcal{Y} \in \mathbb{C}^{I_2\times I_2\times I_1\times I_1}$ with matrix slices $\mathbb{C}^{I_2\times I_2} \ni \mathbf{Y}^{(\cdot\cdot i_3,i_4)} \triangleq \mathcal{Y}(:,:,i_3,i_4) = \mathbf{X}^{(i_3\cdot\cdot)}\mathbf{X}^{(i_4\cdot\cdot)H}$. Since $\mathbf{X}^{(i_1\cdot\cdot)} = \mathbf{A}^{(2)}D_{i_1}(\mathbf{A}^{(1)})\mathbf{C}^T$ and since $\mathbf{C}$ is columnwise orthonormal, the matrix slice $\mathbf{Y}^{(\cdot\cdot i_3,i_4)}$ admits the factorization

$$\mathbf{Y}^{(\cdot\cdot i_3,i_4)} = \mathbf{A}^{(2)}D_{i_3}\big(\mathbf{A}^{(1)}\big)D_{i_4}\big(\mathbf{A}^{(1)*}\big)\mathbf{A}^{(2)H}, \quad i_3, i_4 \in \{1,\ldots,I_1\}\,.$$

Thus, the fourth-order tensor $\mathcal{Y}$ has the following matrix decomposition:

$$\begin{aligned} \mathbf{Y} &= \left[\mathrm{Vec}\big(\mathbf{Y}^{(\cdot\cdot1,1)}\big), \mathrm{Vec}\big(\mathbf{Y}^{(\cdot\cdot1,2)}\big), \ldots, \mathrm{Vec}\big(\mathbf{Y}^{(\cdot\cdot I_1,I_1)}\big)\right] \\ &= \big(\mathbf{A}^{(2)*} \odot \mathbf{A}^{(2)}\big)\big(\mathbf{A}^{(1)*} \odot \mathbf{A}^{(1)}\big)^T. \end{aligned} \tag{5.8}$$

If the fourth-order CPD with matrix representation $\mathbf{Y}$ in (5.8) is unique, then the CPD of $\mathcal{X}$ with columnwise orthonormal factor matrix is also unique. Let us interpret (5.8) as a matrix representation of a CPD with factor matrices $\mathbf{A}^{(N_{\max})*}$, $\mathbf{A}^{(N_{\max})}$, and $\mathbf{A}^{(N_{\min})*} \odot \mathbf{A}^{(N_{\min})}$, in which the Khatri–Rao structure of the latter is in the first instance ignored.

We first explain that conditions (5.7a) and (5.7b) imply orthogonality constrained CPD uniqueness. Analogous to the proof of Theorem 5.1, condition (5.7b) implies that $k_{\mathbf{D}} \geq 2$. Condition (5.7a) together with Theorem 2.1 now tells us that the rank of $\mathcal{Y}$ is R and the factor matrix $\mathbf{A}^{(N_{\min})*} \odot \mathbf{A}^{(N_{\min})}$ is unique. Since $\mathbf{A}^{(N_{\min})*} \odot \mathbf{A}^{(N_{\min})}$ is unique, $\mathbf{A}^{(N_{\min})}$ is unique. Since $\mathbf{A}^{(N_{\min})}$ is unique and conditions (5.7b) hold, Theorem 4.8 guarantees the uniqueness of the factor matrix $\mathbf{A}^{(N_{\max})}$. The columnwise orthonormal matrix $\mathbf{C}$ follows from $\mathbf{X}_{(1)} = (\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)})\, \mathbf{C}^T$, with $\mathbf{X}_{(1)}$, $\mathbf{A}^{(1)}$, and $\mathbf{A}^{(2)}$ known, via an orthogonal Procrustes problem (e.g., [10]). We can now conclude that the orthogonality constrained CPD of $\mathcal{X}$ is unique and the orthogonality constrained rank of $\mathcal{X}$ is R.

We now explain that conditions (5.7a) and (5.7c) imply orthogonality constrained CPD uniqueness. Analogous to the proof of Theorem 5.1, condition (5.7c) implies that $k_{\mathbf{D}} \geq 2$. Condition (5.7a) together with Theorem 2.1 now tells us that the rank of $\mathcal{Y}$ is R and the factor matrix $\mathbf{A}^{(N_{\min})*} \odot \mathbf{A}^{(N_{\min})}$ is unique. This in turn implies that $\mathbf{A}^{(N_{\min})}$ is unique. Since $\mathbf{A}^{(N_{\min})}$ is unique and conditions (5.7c) hold, Theorem 4.6 guarantees the uniqueness of $\mathbf{A}^{(N_{\max})}$ and $\mathbf{C}$. Thus, the orthogonality constrained CPD of $\mathcal{X}$ is unique, and the orthogonality constrained rank of $\mathcal{X}$ is R. ∎

Theorem 5.3 leads to more relaxed uniqueness conditions than those presented in [24]. For instance, consider the PD of $\mathcal{X} \in \mathbb{C}^{7\times 4\times 17}$ with matrix representation $\mathbf{X}_{(1)} = (\mathbf{A} \odot \mathbf{B})\, \mathbf{C}^T$, in which $\mathbf{A} \in \mathbb{C}^{7\times 17}$, $\mathbf{B} \in \mathbb{C}^{4\times 17}$, and columnwise orthonormal $\mathbf{C} \in \mathbb{C}^{17\times 17}$. We randomly generate $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. For this setting the uniqueness results presented in [24] do not apply. On the other hand, using Lemma 4.1, it can be verified that conditions (5.7a) and (5.7c) in Theorem 5.3 with $\operatorname{card}(S) = 9$ are generically satisfied. Once again, since Theorem 2.1 is not necessary for one factor uniqueness it may be possible to generalize Theorem 5.3 by combining Theorem 4.8 with a more relaxed single factor matrix uniqueness condition.

6. Numerical experiments.

Let $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$ denote the rank-R tensor for which the CPD is given by (2.1) with $\mathbf{C}$ known. The goal is to estimate $\mathbf{A}$ and $\mathbf{B}$ from the observed tensor $\mathcal{T} = \mathcal{X} + \beta\mathcal{N}$, where $\mathcal{N}$ is an unstructured perturbation tensor and $\beta \in \mathbb{R}$ controls the noise level. The entries of all the involved factor matrices and perturbation tensors are randomly drawn from a Gaussian distribution with zero mean and unit variance.

The following SNR measure will be used: SNR [dB] $= 10 \log\left(\|\mathbf{X}_{(1)}\|_F^2 / \|\beta\mathbf{N}_{(1)}\|_F^2\right)$. The distance between a factor matrix, say $\mathbf{A}$, and its estimate $\widehat{\mathbf{A}}$, is measured according to the following criterion: $\mathrm{P}(\mathbf{A}) = \min_{\Lambda} \|\mathbf{A} - \widehat{\mathbf{A}}\Lambda\|_F / \|\mathbf{A}\|_F$, Λ is a diagonal matrix.

We compare Algorithms 1 and 2 with the ALS method, which alternates between the update of the two unknown factor matrices $\mathbf{A}$ and $\mathbf{B}$. See the supplementary material for an efficient implementation of ALS. We use ALS as a representative of the class of optimization algorithms; see [23] and references therein for other optimization-based algorithms. Let $f_k = \|\mathbf{T}_{(1)} - \widehat{\mathbf{T}}_{(1)}^{(k)}\|_F$, where $\widehat{\mathbf{T}}_{(1)}^{(k)}$ denotes the estimated tensor at iteration k; then we decide that the ALS method has converged when $f_k - f_{k+1} < \epsilon_{\mathrm{ALS}} = 1e-8$ or when the number of iterations exceeds 5000. The ALS method is randomly initialized. Since the best out of ten random initializations is used, the ALS method will be referred to as ALS-10. Algorithms 1 and 2 will be referred to as Alg1 and Alg2, respectively. When Alg1 and Alg2 are followed by at most 500 ALS refinement iterations they will be referred to as Alg1-ALS and Alg2-ALS, respectively. Note that the computational cost of the Alg1 and Alg2 methods is very low, not much more than one ALS iteration.

For the choice of the integer $\operatorname{card}(S)$ in Algorithm 2, the value that yields the best fit of $\mathcal{T}$ is retained. More precisely, the integer $\operatorname{card}(S)$ which minimizes

$g_{\mathrm{card}(S)} = \|\mathbf{T}_{(1)} - (\widehat{\mathbf{A}}_{\mathrm{card}(S)} \odot \widehat{\mathbf{B}}_{\mathrm{card}(S)})\mathbf{C}^T\|_F$ is retained, where $\widehat{\mathbf{A}}_{\mathrm{card}(S)}$ and $\widehat{\mathbf{B}}_{\mathrm{card}(S)}$ denote the estimates of the factor matrices $\mathbf{A}$ and $\mathbf{B}$ obtained by Algorithm 2 with the given $\mathrm{card}\,(S)$. Since the data are uniformly distributed, we just choose $S = \{1, 2, \ldots, \mathrm{card}\,(S)\}$ if $\mathrm{card}\,(S) \geq 1$ (i.e., we pick the first $\mathrm{card}\,(S)$ columns without trying other combinations) and $S = \emptyset$ if $\mathrm{card}\,(S) = 0$.

Case 1: $I, K < R$ *and* $J > R$. The model parameters are $I = 3$, $J = 10$, $K = 4$, and $R = 9$. The existing uniqueness conditions stated in Propositions 3.1 and 3.2 do not apply, indicating a difficult problem. On the other hand, Theorem 4.6 generically guarantees the uniqueness of the decomposition. The mean and standard deviation $\mathrm{P}\,(\mathbf{A})$ and $\mathrm{P}\,(\mathbf{B})$ values over 100 trials as a function of SNR can be seen in Figure 2. We notice that the Alg1 and Alg1-ALS methods perform better than the ALS-10 method. Hence, the proposed Alg1 method, possibly followed by few ALS refinement steps, seems to be a good method for computing a CPD with a known factor matrix in cases where only one unknown factor matrix has full column rank.

Case 2: $I, J, K < R$. The model parameters are $I = 4$, $J = 6$, $K = 5$, and $R = 8$. Despite the fact that both Proposition 3.2 and Theorem 4.8 generically guarantee the uniqueness of the decomposition, the problem is difficult since none of the involved factors have full column rank. The mean and standard deviation $\mathrm{P}\,(\mathbf{A})$ and $\mathrm{P}\,(\mathbf{B})$ values over 100 trials as a function of SNR can be seen in Figure 3. On average, $\mathrm{card}\,(S) = 2$, which is the smallest value for which the conditions in Theorem 4.8 are satisfied, yielded the best fit. We notice that in the presence of noise Alg2 performs worse than ALS-10. However, we also observe that Alg2-ALS and ALS-10 perform about the same. Since the Alg2-ALS method is computationally cheaper than the ALS-10 method, Alg2 is an attractive procedure for the initialization of the iterative ALS method. Overall, the Alg2-ALS method seems to be the method of choice for the case where none of the unknown factor matrices have full column rank.

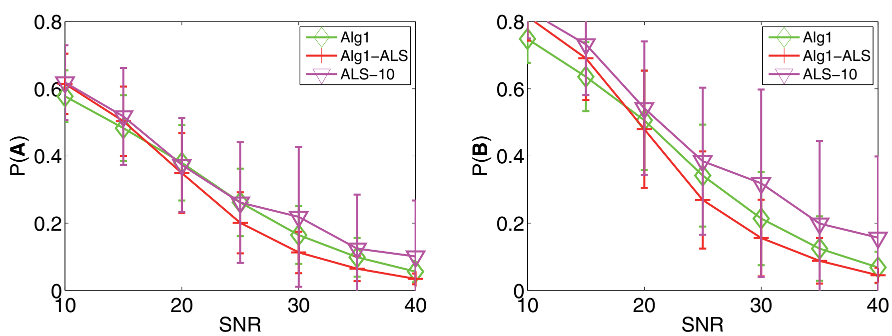

FIG. 2. *Mean and standard deviation* $P(\boldsymbol{A})$ *and* $P(\boldsymbol{B})$ *for varying SNR, case* 1.

7. Conclusion. Many problems in signal processing can be formulated as tensor decomposition problems with a known factor matrix. We mentioned applications in wireless communication. In the first part of this paper we provided a new uniqueness condition for CPD with a known factor matrix. We showed that by taking the known factor into account more relaxed uniqueness conditions can be obtained compared to unconstrained CPD. We also proposed an inexpensive algebraic method for computing a CPD with a known factor matrix. In the supplementary material an efficient implementation of the ALS method for CPD with a known factor matrix was also reported.

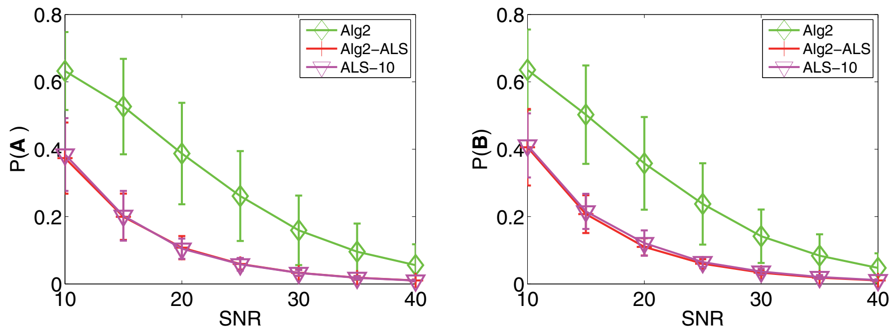

FIG. 3. *Mean and standard deviation $P(\boldsymbol{A})$ and $P(\boldsymbol{B})$ for varying SNR, case* 2.

Based on the results obtained in the first part of this paper, we provided in the second part a new versatile deterministic overall uniqueness condition for the CPD. Since the condition is flexible it can be adapted to other tensor decompositions, as demonstrated in [25, 26]. The supplementary material also contains a partial uniqueness variant of the overall uniqueness result presented in this paper. In tensor-based statistical signal processing CPDs are typically constrained, e.g., orthogonality or Hermitian-symmetry constrained. For that reason we presented new uniqueness conditions for CPD with a partial Hermitian symmetry or columnwise orthonormal factor matrix. Again, such constraints lead in some cases to more relaxed uniqueness conditions than their unconstrained CPD counterparts.

Finally, numerical experiments confirmed the practical use of the proposed Algorithms 1 and 2 for computing a CPD with a known factor matrix. More precisely, in the case where one of the unknown factor matrices had full column rank, Algorithm 1 performed better than the popular ALS method. In the case where none of the unknown factor matrices had full column rank, Algorithm 2 provided at a low computational cost a good initial value for an optimization-based algorithm.

Acknowledgments. The authors thank the anonymous reviewers for several helpful comments and, in particular, for bringing [13] to their attention.

REFERENCES

[1] A. Cichocki, D. Mandic, A.-H. Phan, C. Caifa, G. Zhou, Q. Zhao, 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.

[2] P. Comon and C. Jutten, Eds., *Handbook of Blind Source Separation, Independent Component Analysis and Applications*, Academic Press, Oxford, UK, 2010.

[3] A. L. F. de Almeida, G. Favier, and J. C. M. Mota, *Multiuser MIMO systems using space-time-frequency multiple-access PARAFAC tensor modeling*, Optimizing Wireless Communication Systems, F. Cavalcanti and S. Andersson, eds., Springer, Dordrecht, Heidelberg, London, New York, 2009, pp. 421–461.

[4] A. L. F. de Almeida and G. Favier, *Double Khatri-Rao space-time frequency coding using semi-blind PARAFAC based receiver*, IEEE Signal Process. Lett., 10 (2013), pp. 471–474.

[5] A. L. F. de Almeida, A. Y. Kibangou, S. Miron, and D. C. Araújo, *Joint data and connection topology recovery in collaborative wireless sensor networks*, in 2013 International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Piscataway, NJ, 2013, pp. 5303–5307.

[6] 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.

[7] L. De Lathauwer, *A short introduction to tensor-based methods for factor analysis and blind source separation*, in 7th International Symposium on Image and Signal Processing and Analysis (ISPA), Dubrovnik, Croatia, 2011, IEEE, Piscataway, NJ, pp. 558–563.

[8] 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.

[9] 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.

[10] R. A. Horn and C. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2013.

[11] T. Jiang, N. D. Sidiropoulos, and J. M. F. ten Berge, *Almost-sure identifiability of multidimensional harmonic retrieval*, IEEE Trans. Signal Process., 49 (2001), pp. 1849–1859.

[12] T. Jiang and N. D. Sidiropoulos, *Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints*, IEEE Trans. Signal Process., 52 (2004), pp. 2625–2636.

[13] R. C. Gunning and H. Rossi, *Analytic Functions in Several Complex Variables*, Prentice–Hall, Englewood Cliffs, NJ, 1965.

[14] P. M. Kroonenberg, *Applied Multiway Data Analysis*, John Wiley & Sons, Hoboken, NJ, 2008.

[15] T. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.

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

[17] D. Nion and N. D. Sidiropoulos, *Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar*, IEEE Trans. Signal Process., 58 (2010), pp. 5693–5705.

[18] F. Roemer and M. Haardt, *Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse*, IEEE Trans. Signal Process., 58 (2010), pp. 5720–5735.

[19] A. Smilde, R. Bro, and P. Geladi, *Multi-way Analysis: Applications in the Chemical Sciences*, John Wiley & Sons, Chichester, UK, 2004.

[20] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, *Parallel factor analysis in sensor array processing*, IEEE Trans. Signal Process., 48 (2000), pp. 2377–2388.

[21] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, *Blind PARAFAC receivers for DS-CDMA systems*, IEEE Trans. Signal Process., 48 (2000), pp. 810–823.

[22] N. D. Sidiropoulos and R. S. Budampati, *Khatri-Rao space-time codes*, IEEE Trans. Signal Process., 50 (2002), pp. 2396–2407.

[23] L. Sorber, M. Van Barel, and L. De Lathauwer, *Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-(Lr,Lr,* 1*) terms, and a new generalization*, SIAM J. Optim., 23 (2013), pp. 695–720.

[24] M. Sørensen, L. De Lathauwer, P. Comon, S. Icart, and L. Deneire, *Canonical polyadic decomposition with a columnwise orthonormal factor matrix*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1190–1213.

[25] 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.

[26] 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.

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

[28] A. Stegeman, *On uniqueness of the canonical tensor decomposition with some form of symmetry*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 561–583.

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