Citation/Reference	Domanov I., De Lathauwer L., ``Canonical polyadic decomposition of third-order tensors: relaxed uniqueness conditions and algebraic algorithm'', *Linear Algebra and its Applications*, vol. 513, Jan. 2017, pp. 342-375.
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	http://dx.doi.org/10.1016/j.laa.2016.10.019
Journal homepage	http://www.journals.elsevier.com/linear-algebra-and-its-applications
Author contact	ignat.domanov@kuleuven-kulak.be Klik hier als u tekst wilt invoeren.
IR	https://lirias.kuleuven.be/handle/123456789/554392

Canonical polyadic decomposition of third-order tensors: relaxed uniqueness conditions and algebraic algorithm☆

Ignat Domanov[a,b,c], Lieven De Lathauwer[a,b,c]

[a] *Group Science, Engineering and Technology, KU Leuven - Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium*
[b] *Dept. of Electrical Engineering ESAT/STADIUS KU Leuven, Kasteelpark Arenberg 10, bus 2446, B-3001 Leuven-Heverlee, Belgium*
[c] *iMinds Medical IT*

Abstract

Canonical Polyadic Decomposition (CPD) of a third-order tensor is a minimal decomposition into a sum of rank-1 tensors. We find new mild deterministic conditions for the uniqueness of individual rank-1 tensors in CPD and present an algorithm to recover them. We call the algorithm "algebraic" because it relies only on standard linear algebra. It does not involve more advanced procedures than the computation of the null space of a matrix and eigen/singular value decomposition. Simulations indicate that the new conditions for uniqueness and the working assumptions for the algorithm hold for a randomly generated $I \times J \times K$ tensor of rank $R \geq K \geq J \geq I \geq 2$ if R is bounded as $R \leq (I + J + K - 2)/2 + (K - \sqrt{(I - J)^2 + 4K})/2$ at least for the dimensions that we have tested. This improves upon the famous Kruskal bound for uniqueness $R \leq (I + J + K - 2)/2$ as soon as $I \geq 3$.

In the particular case $R = K$, the new bound above is equivalent to the bound $R \leq (I - 1)(J - 1)$ which is known to be necessary and sufficient for the generic uniqueness of the CPD. An existing algebraic algorithm (based on

☆Research supported by: (1) Research Council KU Leuven: C1 project c16/15/059-nD, CoE PFV/10/002 (OPTEC), PDM postdoc grant; (2) F.W.O.: project G.0830.14N, G.0881.14N; (3) the Belgian Federal Science Policy Office: IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012-2017); (4) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

simultaneous diagonalization of a set of matrices) computes the CPD under the more restrictive constraint $R(R-1) \leq I(I-1)J(J-1)/2$ (implying that $R < (J-\frac{1}{2})(I-\frac{1}{2})/\sqrt{2}+1$). On the other hand, optimization-based algorithms fail to compute the CPD in a reasonable amount of time even in the low-dimensional case $I=3$, $J=7$, $K=R=12$. By comparison, in our approach the computation takes less than 1 sec. We demonstrate that, at least for $R \leq 24$, our algorithm can recover the rank-1 tensors in the CPD up to $R \leq (I-1)(J-1)$.

Keywords: canonical polyadic decomposition, CANDECOMP/PARAFAC decomposition, CP decomposition, tensor, uniqueness of CPD, uni-mode uniqueness, eigenvalue decomposition, singular value decomposition
2000 MSC: 15A69, 15A23

---

1. Introduction

Let $\mathbb{F}$ denote the field of real or complex numbers and $\mathcal{T} \in \mathbb{F}^{I\times J\times K}$ denote a third-order tensor with entries t_{ijk}. By definition, $\mathcal{T}$ is *rank*-1 if it equals the outer product of three nonzero vectors $\mathbf{a} \in \mathbb{F}^I$, $\mathbf{b} \in \mathbb{F}^J$, and $\mathbf{c} \in \mathbb{F}^K$: $\mathcal{T} = \mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}$, which means that $t_{ijk} = a_i b_j c_k$ for all values of indices.

A *Polyadic Decomposition* of $\mathcal{T}$ expresses $\mathcal{T}$ as a sum of rank-1 terms:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r, \qquad \left(\text{or } t_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr}\right) \tag{1}$$

where

$$\mathbf{a}_r = [a_{1r}\ \dots\ a_{Ir}]^T \in \mathbb{F}^I,\ \mathbf{b}_r = [b_{1r}\ \dots\ b_{Jr}]^T \in \mathbb{F}^J,\ \mathbf{c}_r = [c_{1r}\ \dots\ c_{Kr}]^T \in \mathbb{F}^K.$$

If the number R of rank-1 terms in (1) is minimal, then (1) is called the *Canonical Polyadic Decomposition* (CPD) of $\mathcal{T}$ and R is called the *rank* of $\mathcal{T}$ (denoted by $r_{\mathcal{T}}$). It is clear that in (1) the rank-1 terms can be arbitrarily permuted and that vectors within the same rank-1 term can be arbitrarily scaled provided the overall rank-1 term remains the same. The CPD of a tensor *is unique* when it is only subject to these trivial indeterminacies.

We write (1) as $\mathcal{T} = [\mathbf{A},\mathbf{B},\mathbf{C}]_R$, where the matrices $\mathbf{A} := \begin{bmatrix}\mathbf{a}_1 & \dots & \mathbf{a}_R\end{bmatrix} \in \mathbb{F}^{I\times R}$, $\mathbf{B} := \begin{bmatrix}\mathbf{b}_1 & \dots & \mathbf{b}_R\end{bmatrix} \in \mathbb{F}^{J\times R}$ and $\mathbf{C} := \begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_R\end{bmatrix} \in \mathbb{F}^{K\times R}$ are called the *first*, *second* and *third factor matrix* of $\mathcal{T}$, respectively. It may happen that the CPD of a tensor $\mathcal{T}$ is not unique but that nevertheless, for any two

CPDs $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_R$, the factor matrices in a certain mode, say the matrices $\mathbf{C}$ and $\bar{\mathbf{C}}$, coincide up to column permutation and scaling. We say that *the third factor matrix of* $\mathcal{T}$ *is unique*. For instance, it is well known that if two or more columns of the third factor matrix of $\mathcal{T}$ have collinear vectors, then the CPD is not unique. Nevertheless, the third factor matrix can still be unique [7, Example 4.11].

The literature shows some variation in terminology. The CPD was introduced by F.L. Hitchcock in [16] and was later referred to as Canonical Decomposition (CANDECOMP) [2], Parallel Factor Model (PARAFAC) [12, 15], and Topographic Components Model [21]. Uniqueness of one factor matrix is called *uni-mode uniqueness* in [11, 24]. Uniqueness of the CPD is often called *essential uniqueness* in engineering papers and *specific identifiability* in algebraic geometry papers. It is its uniqueness properties that make CPD a basic tool for signal separation and data analysis, with many concrete applications in telecommunication, array processing, machine learning, etc. [4, 5, 18, 22].

The contribution of this paper is twofold. First, we find very mild conditions for uniqueness of CPD and, second, we provide an *algebraic* algorithm for its computation, i.e. an algorithm that recovers the CPD from $\mathcal{T}$ by means of conventional linear algebra (basically by taking the orthogonal complement of a subspace and computing generalized eigenvalue decomposition (GEVD)).

Algebraic algorithms are important from a computational point view in the following sense. In practice, the factor matrices of $\mathcal{T}$ are most often obtained as the solution of the optimization problem

$$\min \|\widehat{\mathcal{T}} - [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R\|, \qquad \text{s.t.} \quad \mathbf{A} \in \mathbb{F}^{I\times R},\ \mathbf{B} \in \mathbb{F}^{J\times R},\ \mathbf{C} \in \mathbb{F}^{K\times R},$$

where $\|\cdot\|$ denotes a suitable norm [23]. The limitations of this approach are not very well-known. Algebraic algorithms may provide a good initial guess. In Example 10 we illustrate that even in a small-scale problem such as the CPD of a rank-12 tensor of dimensions $3 \times 7 \times 12$, the optimization approach may require many initializations and iterations, although the solution can be computed algebraically without a problem.

Basic notation and conventions. Throughout the paper C_n^k denotes the binomial coefficient,

$$C_n^k = \begin{cases} \frac{n!}{k!(n-k)!}, & \text{if } k \le n, \\ 0, & \text{if } k > n; \end{cases}$$

$r_{\mathbf{A}}$, range($\mathbf{A}$), and ker($\mathbf{A}$) denote the rank, the range, and the null space of a matrix $\mathbf{A}$, respectively; $k_{\mathbf{A}}$ (the k-rank of $\mathbf{A}$ [14, p. 162]) is the largest number such that every subset of $k_{\mathbf{A}}$ columns of the matrix $\mathbf{A}$ is linearly independent; "$\odot$" and "$\otimes$" denote the Khatri-Rao and Kronecker product, respectively:

$$\mathbf{A} \odot \mathbf{B} = [\mathbf{a}_1 \otimes \mathbf{b}_1 \ \dots \ \mathbf{a}_R \otimes \mathbf{b}_R],$$
$$\mathbf{a} \otimes \mathbf{b} = [a_1 b_1 \dots a_1 b_j \ \dots \ a_I b_1 \dots a_I b_J]^T.$$

It is well known that PD (1) can be rewritten in a matrix form as

$$\mathbf{R}_{1,0}(\mathcal{T}) := \begin{bmatrix} \mathbf{T}_1 \\ \vdots \\ \mathbf{T}_I \end{bmatrix} = \begin{bmatrix} \mathbf{B}\mathrm{Diag}(\mathbf{a}^1)\mathbf{C}^T \\ \vdots \\ \mathbf{B}\mathrm{Diag}(\mathbf{a}^I)\mathbf{C}^T \end{bmatrix} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T \in \mathbb{F}^{IJ \times K}, \qquad (2)$$

where $\mathbf{T}_i := (t_{ijk})_{j,k=1}^{J,K}$ denotes the ith horizontal slice of $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K}$, $\mathbf{a}^i := [a_{i1} \ \dots \ a_{i_R}]$ denotes the ith row of $\mathbf{A} \in \mathbb{F}^{I \times R}$, and $\mathrm{Diag}(\mathbf{a}^i)$ denotes a square diagonal matrix with the elements of the vector $\mathbf{a}^i$ on the main diagonal.

To simplify the presentation and w.l.o.g. we will assume throughout that the third dimension K coincides with $r_{\mathbf{C}}$, yielding $r_{\mathbf{C}} = K \leq R$. This can always be achieved in a "dimensionality reduction" step (see, for instance, [9, Subsection 1.4]).

2. Previous results, new contribution, and organization of the paper

To explain our contribution, we first briefly recall previous results on uniqueness conditions and algebraic algorithms. (We refer the readers to [7–9] and references therein for recent results and a detailed overview of early results.)

2.1. At least two factor matrices have full column rank

We say that a matrix has full column rank if its columns are linearly independent, implying that it cannot have more columns than rows. The following result is well-known and goes back to Kronecker and Weierstrass.

Theorem 1. *[13, 20] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and suppose that the matrices $\mathbf{B}$ and $\mathbf{C}$ have full column rank and that any two columns of $\mathbf{A}$ are linearly independent:*

$$r_{\mathbf{B}} = r_{\mathbf{C}} = R, \qquad k_{\mathbf{A}} \geq 2. \qquad (3)$$

Then $r_{\mathcal{T}} = R$, the CPD of $\mathcal{T}$ is unique and can be found algebraically.

Theorem 1 is the heart of the algebraic algorithms presented in [9] and also in this paper. To give an idea of how the CPD in Theorem 1 is computed, let us consider the particular case of $2 \times R \times R$ tensors. Then, by (3), $\mathbf{B}$ and $\mathbf{C}$ are $R \times R$ nonsingular matrices. For simplicity we also assume that the second row of $\mathbf{A}$ does not contain zero entries. By (2), PD (1) can be rewritten as

$$\mathbf{T}_1 = \mathbf{B}\mathrm{Diag}(\mathbf{a}^1)\mathbf{C}^T \quad \text{and} \quad \mathbf{T}_2 = \mathbf{B}\mathrm{Diag}(\mathbf{a}^2)\mathbf{C}^T, \tag{4}$$

which easily implies that

$$\mathbf{T}_1\mathbf{T}_2^{-1} = \mathbf{B}\mathbf{D}\mathbf{B}^{-1}, \quad \mathbf{T}_1^T\mathbf{T}_2^{-T} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1},$$

where $\mathbf{D} = \mathrm{Diag}(\mathbf{a}^1)\mathrm{Diag}(\mathbf{a}^2)^{-1}$. By the assumption $k_{\mathbf{A}} \geq 2$, the diagonal entries of $\mathbf{D}$ are distinct. Hence, the matrices $\mathbf{B}$ and $\mathbf{C}$ can be uniquely identified up to permutation and column scaling from the eigenvalue decomposition of $\mathbf{T}_1\mathbf{T}_2^{-1}$ and $\mathbf{T}_1^T\mathbf{T}_2^{-T}$, respectively. One can then easily recover $\mathbf{A}$ from (4). Note that, in general, the matrices $\mathbf{B}$ and $\mathbf{C}$ in Theorem 1 can be obtained from the GEVD of the matrix pencil $(\mathbf{T}_1, \mathbf{T}_2)$.

2.2. At least one factor matrix has full column rank

In this subsection we assume that only the third factor matrix of $\mathcal{T}$ has full column rank. It was shown in [17] that PD (1) is unique if and only if

$$r_{\mathbf{A}\mathrm{Diag}(\boldsymbol{\lambda})\mathbf{B}^T} \geq 1 \text{ for all } \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_R) \text{ with at least two nonzero entries.} \tag{5}$$

Condition (5) is not easy to check for a specific tensor. The following condition is more restrictive but easy to check [6, 9]. We denote by $\mathcal{C}_m(\mathbf{A}) \in \mathbb{R}^{C_I^m \times C_R^m}$ the mth compound matrix of $\mathbf{A} \in \mathbb{F}^{I \times R}$, i.e. the matrix containing the determinants of all $m \times m$ submatrices of $\mathbf{A}$ arranged with the submatrix index sets in lexicographic order.

Theorem 2. *[6, 9] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and suppose that*

$$\text{the matrices } \mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}) \text{ and } \mathbf{C} \text{ have full column rank.} \tag{6}$$

Then $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique.

It was shown in [6, 9] that the assumptions in Theorem 2 also imply an algebraic algorithm. The algorithm is based on the following relation between $\mathcal{T}$ and its factor matrices:

$$\widetilde{\mathbf{R}}_{2,0}(\mathcal{T}) = (\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}))\mathbf{S}_{2,0}(\mathbf{C})^T, \tag{7}$$

in which $\widetilde{\mathbf{R}}_{2,0}(\mathcal{T})$ denotes an $C_I^2 C_J^2 \times R^2$ matrix whose

$$((j_1(2j_2 - j_1 - 1) - 2)I(I-1)/4 + i_1(2i_2 - i_1 - 1)/2, (r_2 - 1)R + r_1)\text{-th}$$
$$(1 \le i_1 < i_2 \le I,\ 1 \le j_1 < j_2 \le J,\ 1 \le r_1, r_2 \le R)$$

entry is equal to $t_{i_1 j_1 r_1} t_{i_2 j_2 r_2} + t_{i_1 j_1 r_2} t_{i_2 j_2 r_1} - t_{i_1 j_2 r_1} t_{i_2 j_1 r_2} - t_{i_1 j_2 r_2} t_{i_2 j_1 r_1}$ and $\mathbf{S}_{2,0}(\mathbf{C})$ denotes an $R^2 \times C_R^2$ matrix that has columns $\frac{1}{2}(\mathbf{c}_{r_1} \otimes \mathbf{c}_{r_2} + \mathbf{c}_{r_2} \otimes \mathbf{c}_{r_1})$, $1 \le r_1 < r_2 \le R$. Computationally, the identity (7) is used as follows. First, the subspace $\ker(\widetilde{\mathbf{R}}_{2,0}(\mathcal{T}))$ is used to construct an auxiliary $R \times R \times R$ tensor $\mathcal{W}$ that has CPD $\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T}, \mathbf{M}]_R$ in which both $\mathbf{C}^{-T}$ and $\mathbf{M}$ have full column rank. The CPD of $\mathcal{W}$ is computed as in Theorem 1, which gives the matrix $\mathbf{C}^{-T}$. The third factor matrix of $\mathcal{T}$, $\mathbf{C}$, is obtained from $\mathbf{C}^{-T}$ and the first two factor matrices $\mathbf{A}$ and $\mathbf{B}$ can be easily found from $\mathbf{R}_{1,0}(\mathcal{T})\mathbf{C}^{-T} = \mathbf{A} \odot \mathbf{B}$ (see (2)) using the fact that the columns of $\mathbf{A} \odot \mathbf{B}$ are vectorized rank-1 matrices.

2.3. None of the factor matrices is required to have full column rank

The following result is known as Kruskal's theorem. It is the most well-known result on uniqueness of the CPD.

Theorem 3. *[19] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and suppose that*

$$2R + 2 \le k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}. \tag{8}$$

Then $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique.

In [7, 8] we presented several generalizations of uniqueness Theorems 2 and 3. In [9] we showed that the CPD can be computed algebraically under a much weaker assumption than (8).

Theorem 4. *[9, Theorem 1.7] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and suppose that*

$$\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}) \text{ has full column rank for } m = R - k_{\mathbf{C}} + 2. \tag{9}$$

Then $r_{\mathcal{T}} = R$, the CPD of $\mathcal{T}$ is unique and can be computed algebraically.

The algorithm in [9] is based on the following extension of (7):

$$\widetilde{\mathbf{R}}_{m,0}(\mathcal{T}) = (\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))\mathbf{S}_{m,0}(\mathbf{C})^T, \tag{10}$$

where the $C_I^m C_J^m \times K^m$ matrix $\widetilde{\mathbf{R}}_{m,0}(\mathcal{T})$ is constructed from the given tensor $\mathcal{T}$ and the $C_R^m \times K^m$ matrix $\mathbf{S}_{m,0}(\mathbf{C})$ depends in a certain way on $\mathbf{C}$. We refer the reader to [9] for details on the algorithm. Here we just mention that assumption (9) guarantees that the matrix $\mathbf{C}$ can be recovered from the subspace $\ker(\widetilde{\mathbf{R}}_{m,0}(\mathcal{T}))$.

2.4. Generic uniqueness results from algebraic geometry

So far we have discussed deterministic conditions, which are expressed in terms of particular $\mathbf{A}, \mathbf{B}, \mathbf{C}$. On the other hand, generic conditions are expressed in terms of dimensions and rank and hold "with probability one". Formally, we say that the CPD of a generic $I \times J \times K$ tensor of rank R is unique if

$$\mu\{(\mathbf{A}, \mathbf{B}, \mathbf{C}) : \text{the CPD of the tensor } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is not unique}\} = 0,$$

where μ denotes the Lebesgue measure on $\mathbb{F}^{(I+J+K)R}$.

It is known from algebraic geometry that if $2 \leq I \leq J \leq K \leq R$, then each of the following conditions implies that the CPD of a generic $I \times J \times K$ tensor of rank R is unique:

$$R \leq \frac{I + J + 2K - 2 - \sqrt{(I-J)^2 + 4K}}{2} \quad \text{(see [10, Proposition 1.6])}, \tag{11}$$

$$R \leq \frac{IJK}{I + J + K - 2} - K, \ 3 \leq I, \ \mathbb{F} = \mathbb{C} \quad \text{(see [1, Corollary 6.2])}, \tag{12}$$

$$R \leq 2^{\alpha+\beta-2} \leq \frac{IJ}{4} \quad \text{(see [3, Theorem 1.1])}, \tag{13}$$

where α and β are maximal integers such that $2^\alpha \leq I$ and $2^\beta \leq J$. Bounds (11)–(13) complement each other. If $R = K$, then bound (11) is equivalent to

$$R \leq (I-1)(J-1). \tag{14}$$

If $\mathbb{F} = \mathbb{C}$, then (14) is not only sufficient but also necessary, i.e., the decomposition is generically not unique for $R > (I-1)(J-1)$ [3, Proposition 2.2].

2.5. Generic versions of deterministic uniqueness conditions

Theorems 2–4, taken from [6, 9], give deterministic conditions under which the CPD is unique and can be computed algebraically. Generic counterparts of condition (6) and Kruskal's bound (8), for the case where $\max(I, J, K) \leq R$, are given by

$$C_R^2 \leq C_I^2 C_J^2 \text{ and } R \leq K \qquad \text{(see [6]) and} \qquad (15)$$

$$2R + 2 \leq I + J + K \qquad \text{(trivial)}, \qquad (16)$$

respectively. We are not aware of a generic counterpart of condition (9), but, obviously, (9) may hold only if the number of columns of the matrix $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ does not exceed the number of rows, i.e., if

$$C_R^m \leq C_I^m C_J^m, \text{ where } m = R - K + 2. \qquad (17)$$

It can be verified that the algebraic geometry based bound (11) significantly improves bounds (15)–(17) if $\min(I, J) \geq 3$. For instance, if $R = K$, then bound (11) is equivalent to (14), as has been mentioned earlier, while (15) and (16) reduce to $R \leq (J - \frac{1}{2})(I - \frac{1}{2})/\sqrt{2} + 1$ and $R \leq I + J - 1$, respectively.

2.6. Our contribution and organization of the paper

In this paper we further extend results from [6–9], narrowing the gap with what is known from algebraic geometry. Namely, we present new deterministic conditions that guarantee that the CPD is unique and can be computed algebraically. Although we do not formally prove that generically the condition coincides with (11), in our simulations we have been able to find the factor matrices by algebraic means up to the latter bound (Examples 9 and 16). Moreover, the algebraic scheme is shown to outperform numerical optimization (Example 10).

Key to our derivation is the following generalization of (2), (7), and (10):

$$\mathbf{R}_{m,l}(\mathcal{T}) := \mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B}) \mathbf{S}_{m+l}(\mathbf{C})^T, \quad m \geq 1, \quad l \geq 0, \qquad (18)$$

in which the matrices $\mathbf{R}_{m,l}(\mathcal{T})$, $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$, and $\mathbf{S}_{m+l}(\mathbf{C})$ are constructed from the tensor $\mathcal{T}$, the matrices $\mathbf{A}$ and $\mathbf{B}$, and the matrix $\mathbf{C}$, respectively. The precise definitions of these matrices are deferred to Section 3, as they require additional technical notations. In order to maintain the easy flow of the text presentation, the proof of (18) is given in Appendix A. The following scheme illustrates the links and shows that, to obtain our new results, we use (18) for $m \geq 2$ and $l \geq 1$:

$$\begin{array}{cccccccc}
(2) & \xLeftarrow{m=1,l=0} & (18) & \xRightarrow{m\geq 2,l=0} & (10) & \xRightarrow{m=2,l=0} & (7) & \\
\Downarrow & & \Downarrow{\scriptstyle m\geq 2,l\geq 1} & & \Downarrow & & & \searrow \\
\text{Theorem 1} & & \text{new results} & & \text{Theorem 4} & \Longrightarrow & \text{Theorem 3} & \text{Theorem 2}
\end{array}$$

(To clarify the link between (18) and (10), we need to mention that the matrices $\widetilde{\mathbf{R}}_{m,0}(\mathcal{T})$ and $\mathcal{C}_m(\mathbf{A})\odot\mathcal{C}_m(\mathbf{B})$ in (10) are obtained by removing the zero and redundant rows of the matrices $\mathbf{R}_{m,0}(\mathcal{T})$ and $\mathbf{\Phi}_{m,0}(\mathbf{A},\mathbf{B})$, respectively).

Our main results on uniqueness and algebraic algorithms for CPD are formulated, explained, and illustrated in Sections 4 (Theorem 8) and 5 (Theorems 11–15). Namely, in Sections 4 and 5 we generalize results mentioned in Subsections 2.2 and 2.3, respectively. In particular, Theorem 8 in Section 4 is the special case of Theorem 15 in Section 5, where the third factor matrix has full column rank, i.e. $r_{\mathbf{C}} = K = R$. For reasons of readability, in our presentation we proceed from the easy Section 4 ($r_{\mathbf{C}} = R$) to the more difficult Section 5 ($r_{\mathbf{C}} \leq R$). The proofs related to Sections 4 and 5 are given in Section 6 and Appendix B. In Section 6 we go from complicated to easy, i.e., in Subsections 6.1–6.5 we first prove the results related to Section 5 and then we derive Theorem 8 from Theorem 15 in Subsection 6.6. The paper is concluded in Section 7.

Our presentation is in terms of real-valued tensors and real-valued factor matrices for notational convenience. Complex variants are easily obtained by taking into account complex conjugations.

3. Construction of the matrices $\mathbf{R}_{m,l}(\mathcal{T})$, $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})$, and $\mathbf{S}_{m+l}(\mathbf{C})$

Let us first introduce some additional notation. Throughout the paper $P_{\{l_1,\dots,l_k\}}$ denotes the set of all permutations of the set $\{l_1,\dots,l_k\}$. We follow the convention that if some of the values $l_1,\dots,l_k$ coincide, then the cardinality of $P_{\{l_1,\dots,l_k\}}$ is counted taken into account multiplicities, so that always card $P_{\{l_1,\dots,l_k\}} = k!$. For instance, $P_{\{1,1,1\}}$ consists of six identical entries $(1,1,1)$. One can easily check that any integer from $\{1,\dots,I^{m+l}J^{m+l}\}$ can be uniquely represented as $(\tilde{i}-1)J^{m+l}+\tilde{j}$ and that any integer from

$\{1, \ldots, K^{m+l}\}$ can be uniquely represented as $\tilde{k}$, where

$$\tilde{i} := 1 + \sum_{p=1}^{m+l} (i_p - 1) I^{m+l-p}, \qquad i_1, \ldots, i_{m+l} \in \{1, \ldots, I\}, \tag{19}$$

$$\tilde{j} := 1 + \sum_{p=1}^{m+l} (j_p - 1) J^{m+l-p}, \qquad j_1, \ldots, j_{m+l} \in \{1, \ldots, J\}, \tag{20}$$

$$\tilde{k} := 1 + \sum_{p=1}^{m+l} (k_p - 1) K^{m+l-p}, \qquad k_1, \ldots, k_{m+l} \in \{1, \ldots, K\}. \tag{21}$$

These expressions are useful for switching between tensor, matrix and vector representations. We can now define $\mathbf{R}_{m,l}(\mathcal{T})$ as follows.

Definition 5. *Let $\mathcal{T} \in \mathbb{R}^{I \times J \times K}$. The $I^{m+l} J^{m+l}$-by-K^{m+l} matrix whose $((\tilde{i} - 1) J^{m+l} + \tilde{j}, \tilde{k})$th entry is*

$$\frac{1}{m!(m+l)!} \sum_{\substack{(s_1,\ldots,s_{m+l}) \in \\ P_{\{k_1,\ldots,k_{m+l}\}}}} \det \begin{bmatrix} t_{i_1 j_1 s_1} & \ldots & t_{i_1 j_m s_m} \\ \vdots & \vdots & \vdots \\ t_{i_m j_1 s_1} & \ldots & t_{i_m j_m s_m} \end{bmatrix} \prod_{p=1}^{l} t_{i_{m+p} j_{m+p} s_{m+p}} \tag{22}$$

is denoted by $\mathbf{R}_{m,l}(\mathcal{T})$.

The matrices $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$ and $\mathbf{S}_{m+l}(\mathbf{C})$ will have $M(m, l, R)$ columns, where

$$M(m, l, R) := C_R^m C_{m+l-1}^{m-1} + C_R^{m+1} C_{m+l-1}^{m} + \cdots + C_R^{m+l} C_{m+l-1}^{m+l-1}.$$

The columns of these matrices are indexed by $(m + l)$-tuples $(r_1, \ldots, r_{m+l})$ such that

$$\begin{gathered} 1 \leq r_1 \leq r_2 \leq \cdots \leq r_{m+l} \leq R \text{ and} \\ \text{the set } \{r_1, \ldots, r_{m+l}\} \text{ contains at least } m \text{ distinct elements.} \end{gathered} \tag{23}$$

It is easy to show that there indeed exist $M(m, l, R)$ $(m + l)$-tuples which satisfy condition (23). We follow the convention that the $(m+l)$-tuples in (23) are ordered lexicographically: the $(m + l)$-tuple $(r'_1, \ldots, r'_{m+l})$ is preceding the $(m + l)$-tuple $(r''_1, \ldots, r''_{m+l})$ if and only if either $r'_1 < r''_1$ or there exists $k \in \{1, \ldots, m + l - 1\}$ such that $r'_1 = r''_1, \ldots r'_k = r''_k$ and $r'_{k+1} < r''_{k+1}$.

We can now define $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$ and $\mathbf{S}_{m+l}(\mathbf{C})$ as follows.

Definition 6. *Let* $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$. *The* $I^{m+l}J^{m+l}$*-by-*$M(m,l,R)$ *matrix whose* $((\tilde{i}-1)J^{m+l}+\tilde{j},(r_1,\dots,r_{m+l}))$*th entry is*

$$\frac{1}{(m!)^2}\sum_{\substack{(s_1,\dots,s_{m+l})\in\\ P_{\{r_1,\dots,r_{m+l}\}}}} \det\begin{bmatrix} a_{i_1s_1} & \dots & a_{i_1s_m}\\ \vdots & \vdots & \vdots\\ a_{i_ms_1} & \dots & a_{i_ms_m}\end{bmatrix}\cdot\det\begin{bmatrix} b_{j_1s_1} & \dots & b_{j_1s_m}\\ \vdots & \vdots & \vdots\\ b_{j_ms_1} & \dots & b_{j_ms_m}\end{bmatrix}\cdot$$
$$a_{i_{m+1}s_{m+1}}\cdots a_{i_{m+l}s_{m+l}}\cdot b_{j_{m+1}s_{m+1}}\cdots b_{j_{m+l}s_{m+l}}$$

is denoted by $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})$.

Definition 7. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$. *The* K^{m+l}*-by-*$M(m,l,R)$ *matrix whose* $(r_1,\dots,r_{m+l})$*th column is*

$$\frac{1}{(m+l)!}\sum_{(s_1,\dots,s_{m+l})\in\ P_{\{r_1,\dots,r_{m+l}\}}} \mathbf{c}_{s_1}\otimes\cdots\otimes\mathbf{c}_{s_{m+l}} \qquad (24)$$

is denoted by $\mathbf{S}_{m+l}(\mathbf{C})$.

4. At least one factor matrix of $\mathcal{T}$ has full column rank

In this section we generalize results from Subsection 2.2, i.e. we assume that the matrix $\mathbf{C}$ has full column rank and without loss of generality $r_{\mathbf{C}} = K = R$. The more general case $r_{\mathbf{C}} = K \le R$ is handled in Theorem 15 in Section 5. The goal of this section is to explain why and how the algebraic algorithm works in the relatively easy but important case $r_{\mathbf{C}} = K = R$, so that in turn Section 5 will be more accessible.

It can be shown that for $l = 0$, condition (25) in Theorem 8 below reduces to condition (6). Thus, Theorem 2 is the special case of Theorem 8 corresponding to $l = 0$. The simulations in Example 9 below indicate that it is always possible to find some $l \ge 0$ so that (25) also covers (5). Although there is no general proof, this suggests that (5) can always be verified by checking (25) for some $l \ge 0$. This would imply that Algorithm 1 can compute the CPD of a generic tensor up to the necessary condition $R \le (I-1)(J-1)$. Example 9 confirms this up to $R \le 24$.

Let $S^{m+l}(\mathbb{R}^{K^{m+l}}) \subset \mathbb{R}^{K^{m+l}}$ denote the subspace spanned by all vectors of the form $\mathbf{x}\otimes\cdots\otimes\mathbf{x}$, where $\mathbf{x}\in\mathbb{R}^K$ is repeated $m+l$ times. In other words, $S^{m+l}(\mathbb{R}^{K^{m+l}})$ contains vectorized versions of all $K\times\cdots\times K$ symmetric tensors of order $m+l$, yielding $\dim S^{m+l}(\mathbb{R}^{K^{m+l}}) = C^{m+l}_{K+m+l-1}$. We have the following result.

Theorem 8. *Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = K = R$, $l \geq 0$, *and let the matrix* $\mathbf{R}_{2,l}(\mathcal{T})$ *be defined as in Definition 5. Assume that*

$$\dim\left(\ker(\mathbf{R}_{2,l}(\mathcal{T}))\bigcap S^{2+l}(\mathbb{R}^{K^{2+l}})\right) = R. \tag{25}$$

Then

(1) $r_{\mathcal{T}} = R$ *and the CPD of* $\mathcal{T}$ *is unique; and*

(2) the CPD of $\mathcal{T}$ *can be found algebraically.*

Condition (25) in Theorem 8 means that the intersection of $\ker(\mathbf{R}_{2,l}(\mathcal{T}))$ and $S^{2+l}(\mathbb{R}^{K^{2+l}})$ has the minimal possible dimension. Indeed, by (18), Definition 7, and the assumption $r_{\mathbf{C}} = K = R$, we have that the intersection contains at least R linearly independent vectors:

$$\ker(\mathbf{R}_{2,l}(\mathcal{T}))\bigcap S^{2+l}(\mathbb{R}^{K^{2+l}}) = \ker(\mathbf{\Phi}_{2,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{2+l}(\mathbf{C})^T)\bigcap S^{2+l}(\mathbb{R}^{K^{2+l}}) \supseteq$$
$$\ker(\mathbf{S}_{2+l}(\mathbf{C})^T)\bigcap S^{2+l}(\mathbb{R}^{K^{2+l}}) \ni \mathbf{x}\otimes\cdots\otimes\mathbf{x}, \quad \mathbf{x} \text{ is a column of } \mathbf{C}^{-T}.$$

The procedure that constitutes the proof of Theorem 8(2) is summarized as Algorithm 1. Let us comment on the different steps. From Definition 5 it follows that the rows of the matrix $\mathbf{R}_{2,l}(\mathcal{T})$ are vectorized versions of $K\times\cdots\times K$ symmetric tensors of order $2+l$. Consistently, in step 2, we find the vectors $\mathbf{w}_1,\dots,\mathbf{w}_R$ that form a basis of the orthogonal complement to $\text{range}(\mathbf{R}_{2,l}(\mathcal{T})^T)$ in the space $S^{2+l}(\mathbb{R}^{R^{2+l}})$. In other words, $\text{span}\{\mathbf{w}_1,\dots,\mathbf{w}_R\} = \ker(\mathbf{R}_{2,l}(\mathcal{T}))\bigcap S^{2+l}(\mathbb{R}^{K^{2+l}})$. If this subspace has minimal dimension, then its structure provides a key to the estimation of $\mathbf{C}$. Indeed, we have already explained that the minimal subspace is given by

$$\ker(\mathbf{R}_{2,l}(\mathcal{T}))\bigcap S^{2+l}(\mathbb{R}^{R^{2+l}}) = \text{range}\left(\underbrace{\mathbf{C}^{-T}\odot\cdots\odot\mathbf{C}^{-T}}_{2+l}\right). \tag{26}$$

In steps 4–5 we recover $\mathbf{C}^{-T}$ from $\mathbf{W}$ using (26) as follows. By (26), there exists a unique nonsingular $R\times R$ matrix $\mathbf{M}$ such that

$$\mathbf{W} = \left(\mathbf{C}^{-T}\odot\cdots\odot\mathbf{C}^{-T}\right)\mathbf{M}^T. \tag{27}$$

In step 4, we construct the tensor $\mathcal{W}$ whose vectorized frontal slices are the vectors $\mathbf{w}_1,\dots,\mathbf{w}_R$. Reshaping both sides of (27) we obtain the CPD

$\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T} \odot \cdots \odot \mathbf{C}^{-T}, \mathbf{M}]_R$. In step 5, we find the CPD by means of a GEVD using the fact that all factor matrices of $\mathcal{W}$ have full column rank, i.e., we have reduced the problem to a situation that is covered by the basic Theorem 1. Finally, in step 6 we recover $\mathbf{A}$ and $\mathbf{B}$ from $\mathbf{R}_{1,0}(\mathcal{T})\mathbf{C}^{-T} = \mathbf{A} \odot \mathbf{B}$ using the fact that the columns of $\mathbf{A} \odot \mathbf{B}$ are vectorized rank-1 matrices.

Algorithm 1 (Computation of CPD, $K = R$ (see Theorem 8(ii)))

Input: $\mathcal{T} \in \mathbb{R}^{I \times J \times R}$ and $l \geq 0$ with the property that there exist $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$, and $\mathbf{C} \in \mathbb{R}^{R \times R}$ such that $R \geq 2$, $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = R$, and (25) holds.

Output: Matrices $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$ and $\mathbf{C} \in \mathbb{R}^{R \times R}$ such that $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$

1: Construct the $I^{2+l}J^{2+l} \times R^{2+l}$ matrix $\mathbf{R}_{2,l}(\mathcal{T})$ by Definition 5.
2: Find $\mathbf{w}_1, \ldots, \mathbf{w}_R$ that form a basis of $\ker(\mathbf{R}_{2,l}(\mathcal{T})) \bigcap S^{2+l}(\mathbb{R}^{R^{2+l}})$
3: $\mathbf{W} \leftarrow [\mathbf{w}_1 \ \ldots \ \mathbf{w}_R]$
4: Reshape the $R^{2+l} \times R$ matrix $\mathbf{W}$ into an $R \times R^{1+l} \times R$ tensor $\mathcal{W}$
5: Compute the CPD
$\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T} \odot \cdots \odot \mathbf{C}^{-T}, \mathbf{M}]_R$ ($\mathbf{M}$ *is a by-product*) (GEVD)
6: Find the columns of $\mathbf{A}$ and $\mathbf{B}$ from the equation $\mathbf{A} \odot \mathbf{B} = \mathbf{R}_{1,0}(\mathcal{T})\mathbf{C}^{-T}$

The following example demonstrates that the CPD can effectively be computed by Algorithm 1 for $R \leq \min((I-1)(J-1), 24)$.

Example 9. *We consider $I \times J \times (I-1)(J-1)$ tensors generated as a sum of $R = (I-1)(J-1)$ random rank-1 tensors. More precisely, the tensors are generated by a PD $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ in which the entries of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are independently drawn from the standard normal distribution $N(0,1)$. We try different values $l = 0, 1, \ldots,$ until condition (25) is met (assuming that this will be the case for some $l \geq 0$). We test all cases $I \times J \times (I-1)(J-1)$ such that $I \geq 3$, $J \geq 3$, and $(I-1)(J-1) \leq 24$. The results are shown in Table 1. In all cases (25) indeed holds for some $l \leq 2$; the actual value of l does not depend on the random trial, i.e., it is constant for tensors of the same dimensions and rank. By comparison, the algebraic algorithm from [6, 9] is limited to the cases where $l = 0$, which is not always sufficient to reach the bound $R \leq (I-1)(J-1)$. In our implementation, we retrieved the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_R$ from the R-dimensional null space of a $C^{2+l}_{R+l+1} \times C^{2+l}_{R+l+1}$ positive semi-definite matrix $\mathbf{Q}$. The storage of $\mathbf{Q}$ is the main bottleneck in our implementation. To give some insight in the complexity of the algorithm we* *included the computational time (averaged over 100 random tensors) and the size of* $\mathbf{Q}$ *in the table. We implemented Algorithm 1 in MATLAB 2014a (the implementation was not optimized), and we did experiments on a computer with Intel Core 2 Quad CPUQ9650 3.00 GHz×4 and 8GB memory running Ubuntu 12.04.5 LTS.*

Table 1: Values of parameter l in Theorem 8 and computational cost of Algorithm 1 for $I \times J \times (I-1)(J-1)$ tensors of rank $R = (I-1)(J-1) \leq 24$ (see Example 9 for details). Note that the CPD is not generically unique if $R > (I-1)(J-1)$ (see Subsection 2.4). In all cases a value of l is found such that Algorithm 1 can be used. The rows with $l \geq 1$ are new results.

dimensions of $\mathcal{T}$	l	C_{R+l+1}^{2+l}	computational time (sec)
$3 \times 3 \times 4$	0	10	0.02
$3 \times 4 \times 6$	0	21	0.035
$3 \times 5 \times 8$	0	36	0.051
$3 \times 6 \times 10$	0	55	0.074
$3 \times 7 \times 12$	1	364	0.403
$3 \times 8 \times 14$	1	560	0.796
$3 \times 9 \times 16$	1	816	1.498
$3 \times 10 \times 18$	1	1140	2.617
$3 \times 11 \times 20$	1	1540	5.032
$3 \times 12 \times 22$	1	2024	7.089
$3 \times 13 \times 24$	1	2600	11.084
$4 \times 4 \times 9$	0	45	0.06
$4 \times 5 \times 12$	1	364	0.401
$4 \times 6 \times 15$	1	680	1.096
$4 \times 7 \times 18$	2	5985	30.941
$4 \times 8 \times 21$	2	10626	93.03
$4 \times 9 \times 24$	2	17550	360.279
$5 \times 5 \times 16$	1	816	1.473
$5 \times 6 \times 20$	2	8855	64.116
$5 \times 7 \times 24$	2	17550	351.968

The next example illustrates that Algorithm 1 may outperform optimization algorithms.

Example 10. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_{12} \in \mathbb{R}^{3\times 7\times 12}$*, with*

$$\mathbf{A} = hankel((1,2,3),(3,5,7,0,6,6,7,9,0,8,2,1)^T),$$
$$\mathbf{B} = [\mathbf{I}_7\ hankel((1,2,3,4,5,6,7),(7,0,1,2,3)^T)], \qquad \mathbf{C} = \mathbf{I}_{12},$$

where $hankel(\mathbf{c}, \mathbf{r}^T)$ *denotes a Hankel matrix whose first column is* $\mathbf{c}$ *and whose last row is* $\mathbf{r}^T$*. It turns out that* (25) *holds for* $l = 1$*. It takes less than* 1 *second to compute the CPD of* $\mathcal{T}$ *by Algorithm 1. On the other hand, it proves to be very difficult to find the CPD by means of numerical optimization. Among other optimization-based algorithms we tested the Gauss-Newton dogleg trust region method [23]. The algorithm was restarted* 500 *times from various random initial positions. In only 4 cases the residual* $\|\mathcal{T} - [\mathbf{A}_{est}, \mathbf{B}_{est}, \mathbf{C}_{est}]_{12}\|/\|\mathcal{T}\|$ *after* 10000 *iterations was of the order of* 0.0001 *and in all cases the estimated factor matrices were far from the true matrices. Other optimization-based algorithms did not yield better results.*

5. None of the factor matrices is required to have full column rank

In this subsection we consider the PD $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and extend results of the previous subsection to the case $r_{\mathbf{C}} = K \leq R$.

5.1. Results on uniqueness of one factor matrix and overall CPD

We have two results on uniqueness of the third factor matrix.

Theorem 11. *Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$*,* $r_{\mathbf{C}} = K \leq R$*,* $m = R - K + 2$*, and* $l_1, \dots, l_m$ *be nonnegative integers. Let also the matrices* $\mathbf{\Phi}_{1,l_1}(\mathbf{A}, \mathbf{B}), \dots, \mathbf{\Phi}_{m,l_m}(\mathbf{A}, \mathbf{B})$ *and* $\mathbf{S}_{1+l_1}(\mathbf{C}), \dots, \mathbf{S}_{m+l_m}(\mathbf{C})$ *be defined as in Definition 6 and Definition 7, respectively. Let* $\mathbf{U}_1, \dots, \mathbf{U}_m$ *be matrices such that their columns form bases for* $\operatorname{range}(\mathbf{S}_{1+l_1}(\mathbf{C})^T), \dots, \operatorname{range}(\mathbf{S}_{m+l_m}(\mathbf{C})^T)$*, respectively. Assume that*

(i) $k_{\mathbf{C}} \geq 1$*; and*

(ii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank; and*

(iii) $\mathbf{\Phi}_{1,l_1}(\mathbf{A}, \mathbf{B})\mathbf{U}_1, \dots, \mathbf{\Phi}_{m,l_m}(\mathbf{A}, \mathbf{B})\mathbf{U}_m$ *have full column rank.*

Then $r_{\mathcal{T}} = R$ *and the third factor matrix of* $\mathcal{T}$ *is unique.*

According to the following theorem the set of matrices in (iii) in Theorem 11 can be reduced to a single matrix if $R \leq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) + K - 1$.

Theorem 12. *Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$*,* $r_{\mathbf{C}} = K \leq R$*,* $m = R - K + 2$*, and* $l \geq 0$*. Let also the matrices* $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$ *and* $\mathbf{S}_{m+l}(\mathbf{C})$ *be defined as in Definition 6 and Definition 7, respectively. Let* $\mathbf{U}_m$ *be a matrix such that its columns form a basis for* $\operatorname{range}(\mathbf{S}_{m+l}(\mathbf{C})^T)$*. Assume that*

(i) $k_{\mathbf{C}} \geq 1$*; and*

(ii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank; and*

(iii) $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m - 1$*; and*

(iv) *the matrix* $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})\mathbf{U}_m$ *has full column rank.*

Then $r_{\mathcal{T}} = R$ *and the third factor matrix of* $\mathcal{T}$ *is unique.*

The assumptions in Theorems 11 and 12 complement each other as follows: in Theorem 11 we do not require that the condition $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1$ holds while in Theorem 12 we do not require that the matrices $\mathbf{\Phi}_{k,l_k}(\mathbf{A}, \mathbf{B})\mathbf{U}_k$, $1 \leq k \leq m - 1$ have full column rank.

It was shown in [8, Proposition 1.20] that if $\mathcal{T}$ has two PDs $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \mathbf{C}]_R$ that share the factor matrix $\mathbf{C}$ and if the condition

$$\max(\min(k_{\mathbf{A}}, k_{\mathbf{B}} - 1),\ \min(k_{\mathbf{A}} - 1, k_{\mathbf{B}})) + k_{\mathbf{C}} \geq R + 1 \tag{28}$$

holds, then both PDs consist of the same rank-one terms. Thus, combining Theorems 11–12 with [8, Proposition 1.20] we directly obtain the following result on uniqueness of the overall CPD.

Theorem 13. *Let the assumptions in Theorem 11 or Theorem 12 hold and let condition* (28) *be satisfied. Then* $r_{\mathcal{T}} = R$ *and the CPD of tensor* $\mathcal{T}$ *is unique.*

5.2. Algebraic algorithm for CPD

We have the following result on algebraic computation.

Theorem 14. *Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$*,* $r_{\mathbf{C}} = K \leq R$*,* $m = R - K + 2$*, and* $l \geq 0$*. Let also the matrices* $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$ *and* $\mathbf{S}_{m+l}(\mathbf{C})$ *be defined as in Definition 6 and Definition 7, respectively. Let* $\mathbf{U}_m$ *be a matrix such that its columns form a basis for* $\operatorname{range}(\mathbf{S}_{m+l}(\mathbf{C})^T)$*. Assume that*

(i) $k_{\mathbf{C}} = K$*; and*

(ii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank; and*

(iii) *the matrix* $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})\mathbf{U}_m$ *has full column rank.*

Then $r_{\mathcal{T}} = R$, *the CPD of* $\mathcal{T}$ *is unique and can be found algebraically.*

The assumptions in Theorem 14 are more restrictive than the assumptions in Theorem 13 as will be clear from Section 6. Hence, the statement on rank and uniqueness in Theorem 14 follows from Theorem 13. To prove the statement on algebraic computation we will explain in Section 6 that Theorem 14 can be reformulated as follows (see Section 4 for the definition of $S^{m+l}(\mathbb{R}^{K^{m+l}})$).

Theorem 15. *Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = K \leq R$, $m = R - K + 2$, *and* $l \geq 0$. *Let also the matrix* $\mathbf{R}_{m,l}(\mathcal{T})$ *be defined as in Definition 5. Assume that*

(i) $k_{\mathbf{C}} = K$; *and*

(ii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank; and*

(iii) $\dim \left(\ker(\mathbf{R}_{m,l}(\mathcal{T})) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) \right) = C_R^{K-1}$.

Then $r_{\mathcal{T}} = R$, *the CPD of* $\mathcal{T}$ *is unique and can be found algebraically.*

Note that if $k_{\mathbf{C}} = K$, then by (18) and Lemma 22 (i) below,

$$\begin{aligned}
&\dim \left(\ker(\mathbf{R}_{m,l}(\mathcal{T})) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) \right) = \\
&\dim \left(\ker(\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})\mathbf{S}_{m+l}(\mathbf{C})^T) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) \right) \geq \\
&\dim \left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) \right) = C_R^{K-1}.
\end{aligned} \tag{29}$$

Thus, assumption (iii) of Theorem 15 means that we require the subspace to have the minimal possible dimension. That is, we suppose that the factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are such that the multiplication by $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$ in (18) does not increase the overlap between $\ker(\mathbf{S}_{m+l}(\mathbf{C})^T)$ and $S^{m+l}(\mathbb{R}^{K^{m+l}})$. In other words, we suppose that the multiplication by $\mathbf{\Phi}_{m,l}(\mathbf{A}, \mathbf{B})$ does not cause additional vectorized $K \times \cdots \times K$ symmetric tensors of order $m + l$ to be part of the null space of $\mathbf{R}_{m,l}(\mathcal{T})$. This is key to the derivation. By the assumption, as we will explain further in this section, the only vectorized symmetric tensors in the null space of $\mathbf{R}_{m,l}(\mathcal{T})$ admit a direct connection with the factor matrix $\mathbf{C}$, from which $\mathbf{C}$ may be retrieved. On the other hand, the null space of $\mathbf{R}_{m,l}(\mathcal{T})$ can obviously be computed from the given tensor $\mathcal{T}$.

The algebraic procedure based on Theorem 15 consists of three phases and is summarized in Algorithm 2. In the first phase we find the $K \times C_R^{K-1}$ matrix $\mathbf{F}$ such that

$$\text{every column of } \mathbf{F} \text{ is orthogonal to exactly } K-1 \text{ columns of } \mathbf{C} \text{ and} \tag{30}$$

$$\text{any vector that is orthogonal to exactly } K-1 \text{ columns of } \mathbf{C} \text{ is proportional to a column of } \mathbf{F}. \tag{31}$$

Since $k_{\mathbf{C}} = K$ any $K-1$ columns of $\mathbf{C}$ define a unique column of $\mathbf{F}$ (up to

Algorithm 2 (Computation of CPD, $K \leq R$ (see Theorem 15))

Input: $\mathcal{T} \in \mathbb{R}^{I\times J\times K}$ and $l \geq 0$ with the property that there exist $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, and $\mathbf{C} \in \mathbb{R}^{K\times R}$ such that $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and assumptions (i)–(iii) in Theorem 15 hold.

Output: Matrices $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$ and $\mathbf{C} \in \mathbb{R}^{R\times R}$ such that $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$

Phase 1: Find the matrix $\mathbf{F} \in \mathbb{R}^{K\times C_R^{K-1}}$ such that $\mathbf{F}$ coincides with $\mathcal{B}(\mathbf{C})$ up to (unknown) column permutation and scaling

1: Construct the $I^{m+l}J^{m+l} \times K^{m+l}$ matrix $\mathbf{R}_{m,l}(\mathcal{T})$ by Definition 5.
2: Find $\mathbf{w}_1, \dots, \mathbf{w}_{C_R^{K-1}}$ that form a basis of $\ker(\mathbf{R}_{m,l}(\mathcal{T})) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})$
3: $\mathbf{W} \leftarrow [\mathbf{w}_1 \ \dots \ \mathbf{w}_{C_R^{K-1}}]$
4: Reshape the $K^{m+l} \times C_R^{K-1}$ matrix $\mathbf{W}$ into an $K \times K^{m+l-1} \times C_R^{K-1}$ tensor $\mathcal{W}$
5: Compute the CPD

$$\mathcal{W} = [\mathbf{F}, \mathbf{F}\odot\cdots\odot\mathbf{F}, \mathbf{M}]_{C_R^{K-1}} \quad (\mathbf{M} \textit{ is a by-product}) \tag{GEVD}$$

Phase 2 and Phase 3 (can be taken verbatim from [9, Algorithms 1,2])

scaling). Thus, (30)–(31) define the matrix $\mathbf{F}$ up to column permutation and scaling. A special representation of $\mathbf{F}$ (called $\mathcal{B}(\mathbf{C})$) was studied in [9]. It was shown in [9] that the matrix $\mathbf{F}$ can be considered as an unconventional variant of the inverse of $\mathbf{C}$:

$$\text{every column of } \mathbf{C} \text{ is orthogonal to exactly } C_{R-1}^{K-2} \text{ columns of } \mathbf{F}, \tag{32}$$

$$\begin{aligned}\text{any vector that is orthogonal to exactly } C_{R-1}^{K-2} \text{ columns of } \mathbf{F}\\ \text{is proportional to a column of } \mathbf{C}.\end{aligned} \tag{33}$$

(Note that, since $k_{\mathbf{C}} = K$, multiplication by the Moore–Penrose pseudo-inverse $\mathbf{C}^{\dagger}$ yields $\mathbf{C}\mathbf{C}^{\dagger} = \mathbf{I}_K$. In contrast, for $\mathbf{F}$ we consider the product $\mathbf{FC}$.) It can be shown (see Lemma 23) that under the assumptions in Theorems 14–15:

$$k_{\mathbf{F}} \geq 2, \qquad \text{the matrix } \mathbf{F}^{(m+l-1)} \text{ has full column rank and} \tag{34}$$

$$\ker(\mathbf{R}_{m,l}(\mathcal{T})) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) = \text{range}\left(\mathbf{F}^{(m+l)}\right), \tag{35}$$

where

$$\mathbf{F}^{(m+l-1)} := \underbrace{\mathbf{F} \odot \cdots \odot \mathbf{F}}_{m+l-1}, \qquad \mathbf{F}^{(m+l)} := \underbrace{\mathbf{F} \odot \cdots \odot \mathbf{F}}_{m+l}. \tag{36}$$

If $K = R$ (as in Subsection 4), then $m = R - K + 2 = 2$, (35) coincides with (26) ($\mathbf{F}$ coincides with $\mathbf{C}^{-T}$ up to column permutation and scaling), and the first phase of Algorithm 2 coincides with steps 1–5 of Algorithm 1. For $K < R$ (implying $m > 2$) we work as follows. From Definition 5 it follows that the rows of the matrix $\mathbf{R}_{m,l}(\mathcal{T})$ are vectorized versions of $K \times \cdots \times K$ symmetric tensors of order $m + l$. Thus, in step 2, we find the vectors $\mathbf{w}_1, \dots, \mathbf{w}_{C_R^{K-1}}$ that form a basis of the orthogonal complement to $\text{range}(\mathbf{R}_{m,l}(\mathcal{T})^T)$ in the space $S^{m+l}(\mathbb{R}^{K^{m+l}})$ (the existence of such a basis follows from assumption (iii) of Theorem 15). By (35), there exists a unique nonsingular $C_R^{K-1} \times C_R^{K-1}$ matrix $\mathbf{M}$ such that

$$\mathbf{W} = \mathbf{F}^{(m+l)}\mathbf{M}^T. \tag{37}$$

In step 4, we construct the tensor $\mathcal{W}$ whose vectorized frontal slices are the vectors $\mathbf{w}_1, \dots, \mathbf{w}_{C_R^{K-1}}$. Reshaping both sides of (37) we obtain the CPD $\mathcal{W} = [\mathbf{F}, \mathbf{F}^{(m+l-1)}, \mathbf{M}]_R$ in which the matrices $\mathbf{F}^{(m+l-1)}$ and $\mathbf{M}$ have full column rank and $k_{\mathbf{F}} \geq 2$. By Theorem 1, the CPD of $\mathcal{W}$ can be computed by means of GEVD.

In the second and third phase we use $\mathbf{F}$ to find $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$. There are two ways to do this. The first way is to find $\mathbf{C}$ from $\mathbf{F}$ by (32)–(33) and then to recover $\mathbf{A}$ and $\mathbf{B}$ from $\mathcal{T}$ and $\mathbf{C}$. The second way is to find $\mathbf{A}$ and $\mathbf{B}$ from $\mathcal{T}$ and $\mathbf{F}$ and then to recover $\mathbf{C}$. The second and third phase were thoroughly discussed in [9] and can be taken verbatim from [9, Algorithms 1 and 2].

Example 16. *Table 2 contains some examples of CPDs which can be computed by Algorithm 2 and cannot be computed by algorithms from [9]. The tensors were generated by a PD* $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *in which the entries of* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{C}$ *are independently drawn from the standard normal distribution* $N(0, 1)$. *We try different values* $l = 0, 1, \dots,$ *until condition* (iii) *in Theorem 15 is met (assuming that this will be the case for some* $l \geq 0$*). In our implementation, we retrieved the vectors* $\mathbf{w}_1, \dots, \mathbf{w}_{C_R^{K-1}}$ *from the* C_R^{K-1}*-dimensional null space of a* $C_{R+l+1}^{m+l} \times C_{R+l+1}^{m+l}$ *positive semi-definite matrix* $\mathbf{Q}$. *The storage of* $\mathbf{Q}$ *is the main bottleneck in our implementation. To give some insight in the complexity of the algorithm we included the computational time (averaged over 100 random tensors) and the size of* $\mathbf{Q}$ *in the table.*

Uniqueness of the CPDs follows from Theorem 15. By comparison, the results of [8] guarantee uniqueness only for rows 1–4 (see [8, Table 3.1]).

Table 2: Upper bounds on R under which the CPD of a generic $I \times J \times K$ tensor can be computed by Algorithm 2 (see Example 16 details).

dimensions of $\mathcal{T}$	R	m	l	C_{R+l+1}^{m+l}	computational time (sec)
$4 \times 5 \times 6$	7	3	1	126	0.182
$5 \times 7 \times 7$	9	4	1	462	1.598
$6 \times 9 \times 8$	11	5	1	1716	28.616
$7 \times 7 \times 7$	10	5	1	924	8.192
$4 \times 6 \times 8$	9	3	1	330	0.63
$4 \times 7 \times 10$	11	3	1	715	2.352
$5 \times 6 \times 6$	8	4	2	462	1.256
$5 \times 7 \times 8$	10	4	2	1716	14.552

6. Proofs related to Sections 4 and 5

In this section we 1) prove Theorems 11 and 12; 2) show that the assumptions in Theorem 14 are more restrictive than the assumptions in Theorem 13, which implies the statement on uniqueness in Theorem 14; 3) prove that assumption (iii) in Theorem 14 is equivalent to assumption (iii) in Theorem 15; 4) prove statements (34)–(35); 5) prove Theorem 8.

6.1. Proofs of Theorems 11 and 12

In the sequel, $\omega(\lambda_1, \dots, \lambda_R)$ denotes the number of nonzero entries of $[\lambda_1 \ \dots \ \lambda_R]^T$. The following condition (Wm) was introduced in [7, 8] in terms of m-th compound matrices. In this paper we will use the following (equivalent) definition of (Wm) .

Definition 17. *We say that condition* (Wm) *holds for the triplet of matrices* $(\mathbf{A}, \mathbf{B}, \mathbf{C}) \in \mathbb{R}^{I\times R} \times \mathbb{R}^{J\times R} \times \mathbb{R}^{K\times R}$ *if* $\omega(\lambda_1, \dots, \lambda_R) \leq m-1$ *whenever*

$$r_{\mathbf{A}\mathrm{Diag}(\lambda_1,\dots,\lambda_R)\mathbf{B}^T} \leq m-1 \quad \textit{and} \quad [\lambda_1 \ \dots \ \lambda_R]^T \in \mathrm{range}(\mathbf{C}^T). \tag{38}$$

Since the rank of the product $\mathbf{A}\mathrm{Diag}(\lambda_1, \dots, \lambda_R)\mathbf{B}^T$ does not exceed the rank of the factors and $r_{\mathrm{Diag}(\lambda_1,\dots,\lambda_R)} = \omega(\lambda_1, \dots, \lambda_R)$, we always have the implication

$$\omega(\lambda_1, \dots, \lambda_R) \leq m-1 \quad \Rightarrow \quad r_{\mathbf{A}\mathrm{Diag}(\lambda_1,\dots,\lambda_R)\mathbf{B}^T} \leq m-1. \tag{39}$$

By Definition 17, condition (Wm) holds for the triplet $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ if and only if the opposite implication in (39) holds for all $[\lambda_1 \ \dots \ \lambda_R] \in \mathrm{range}(\mathbf{C}^T) \subset \mathbb{R}^R$.

The following results on rank and uniqueness of one factor matrix have been obtained in [7].

Proposition 18. *(see [7, Proposition 4.9]) Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = K \leq R$. *Assume that*

(i) $k_{\mathbf{C}} \geq 1$;

(ii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank;*

(iii) conditions (Wm), …, (W1) *hold for the triplet of matrices* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$.

Then $r_{\mathcal{T}} = R$ *and the third factor matrix of* $\mathcal{T}$ *is unique.*

Proposition 19. *(see [7, Corollary 4.10]) Let* $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = K \leq R$. *Assume that*

(i) $k_{\mathbf{C}} \geq 1$;

(ii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank;*

(iii) $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1$;

(iv) condition (W$_{\text{m}}$) *holds for the triplet of matrices* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$.

Then $r_{\mathcal{T}} = R$ *and the third factor matrix of* $\mathcal{T}$ *is unique.*

One can easily notice the similarity between the assumptions in Theorems 11–12 and the assumptions in Propositions 18–19. The proofs of Theorems 11–12 follow from Propositions 18–19 and the following lemma.

Lemma 20. *Let* $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, *and* $\mathbf{C} \in \mathbb{R}^{K\times R}$, $r_{\mathbf{C}} = K \leq R$, $k \leq m = R - K + 2$, *and let* l *be a nonnegative integer. Let also the matrix* $\mathbf{\Phi}_{k,l}(\mathbf{A}, \mathbf{B})$ *be defined as in Definition 6, the matrix* $\mathbf{S}_{k+l}(\mathbf{C})$ *be defined as in Definition 7, and* $\mathbf{U}$ *be a matrix such that its columns form a basis for* $\text{range}(\mathbf{S}_{k+l}(\mathbf{C})^T)$. *Assume that*

$$\text{the matrix } \mathbf{\Phi}_{k,l}(\mathbf{A}, \mathbf{B})\mathbf{U} \text{ has full column rank.} \tag{40}$$

Then condition (W$_{\text{k}}$) *holds for the triplet of matrices* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$.

Proof. Let (38) hold for $m = k$. We need to show that $\omega(\lambda_1, \dots, \lambda_R) \leq k-1$. Since $[\lambda_1 \ \dots \ \lambda_R]^T \in \text{range}(\mathbf{C}^T)$ and $r_{\mathbf{C}} = K$, there exists a unique vector $\mathbf{x} \in \mathbb{R}^K$ such that $[\lambda_1 \ \dots \ \lambda_R] = \mathbf{x}^T\mathbf{C}$. Hence, we need to show that $\mathbf{x}$ is orthogonal to at least $R - k + 1$ columns of $\mathbf{C}$.

By (38), there exist $\tilde{\mathbf{A}} \in \mathbb{R}^{I\times R}$ and $\tilde{\mathbf{B}} \in \mathbb{R}^{J\times R}$ such that

$$\mathbf{A}\text{Diag}(\lambda_1, \dots, \lambda_R)\mathbf{B}^T = \tilde{\mathbf{A}}\tilde{\mathbf{B}}^T \tag{41}$$

and $\max(r_{\tilde{\mathbf{A}}}, r_{\tilde{\mathbf{B}}}) \leq k - 1$. Since $\mathbf{a}\mathbf{b}^T\lambda = \mathbf{a} \otimes \mathbf{b} \otimes \lambda$, we can consider (41) as an equality of two PDs of an $I \times J \times 1$ tensor

$$\sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \lambda_r = \sum_{r=1}^{R} \tilde{\mathbf{a}}_r \otimes \tilde{\mathbf{b}}_r \otimes 1.$$

Hence, by (18),

$$\begin{aligned} \mathbf{\Phi}_{k,l}(\mathbf{A}, \mathbf{B})\mathbf{S}_{k+l}(\mathbf{x}^T\mathbf{C})^T = \mathbf{\Phi}_{k,l}(\mathbf{A}, \mathbf{B})\mathbf{S}_{k+l}([\lambda_1 \ \dots \ \lambda_R])^T = \\ \mathbf{\Phi}_{k,l}(\tilde{\mathbf{A}}, \tilde{\mathbf{B}})\mathbf{S}_{k+l}([1 \ \dots \ 1])^T. \end{aligned} \tag{42}$$

Since $\max(r_{\tilde{\mathbf{A}}}, r_{\tilde{\mathbf{B}}}) \leq k-1$, it follows from Definition 6 that $\mathbf{\Phi}_{k,l}(\tilde{\mathbf{A}}, \tilde{\mathbf{B}})$ is the zero matrix (cf. explanation at the end of Section Appendix A). Besides, it easily follows from Definition 7 that

$$\mathbf{S}_{k+l}(\mathbf{C})^T(\underbrace{\mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{k+l}) = \mathbf{S}_{k+l}(\mathbf{x}^T\mathbf{C})^T.$$

Thus, (42) takes the form

$$\mathbf{\Phi}_{k,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{k+l}(\mathbf{C})^T(\mathbf{x}\otimes\cdots\otimes\mathbf{x}) = \mathbf{\Phi}_{k,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{k+l}(\mathbf{x}^T\mathbf{C})^T = \mathbf{0}.$$

Hence, by (40), the vector $\mathbf{x}\otimes\cdots\otimes\mathbf{x}$ is orthogonal to the range of $S_{k+l}(\mathbf{C})$. In particular,

$$(\mathbf{x}\otimes\cdots\otimes\mathbf{x})^T \sum_{\substack{(s_1,\dots,s_{k+l})\in\\ P_{\{r_1,\dots,r_k,\dots,r_k\}}}} \mathbf{c}_{s_1}\otimes\cdots\otimes\mathbf{c}_{s_{m+l}} =$$

$$(\mathbf{x}^T\mathbf{c}_{r_1})\cdots(\mathbf{x}^T\mathbf{c}_{r_{k-1}})(\mathbf{x}^T\mathbf{c}_{r_k})^{l+1} = 0$$

for all $(k+l)$-tuples $(r_1,\dots,r_k,\dots,r_k)$ such that $1 \le r_1 < \cdots < r_k \le R$, yielding that $\mathbf{x}$ is orthogonal to at least $R-k+1$ columns of $\mathbf{C}$. □

6.2. Proof of statement on rank and uniqueness in Theorem 14

In Lemma 21 below we prove that $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m$. It is clear that condition $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m$ and assumption (i) in Theorem 14 imply assumption (iii) in Theorem 12 and condition (28). Hence, by Theorem 13, $r_{\mathcal{T}} = R$ and the CPD of tensor $\mathcal{T}$ is unique.

Lemma 21. *Let assumptions* (i) *and* (iii) *in Theorem 14 hold. Then* $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m$.

Proof. Assume to the contrary that $k_{\mathbf{A}} < m$ or $k_{\mathbf{B}} < m$. W.l.o.g. we assume that the first m columns of $\mathbf{A}$ are linearly dependent. We will get a contradiction with assumption (iii) by constructing a nonzero vector $\mathbf{f} \in \text{range}(\mathbf{S}_{m+l}(\mathbf{C})^T)$ such that $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{f} = \mathbf{0}$. Since $k_{\mathbf{C}} = K$, there exists $\mathbf{x} \in \mathbb{R}^K$ such that

$$\mathbf{x}^T\mathbf{c}_1 \ne 0, \dots, \mathbf{x}^T\mathbf{c}_m \ne 0, \quad \mathbf{x}^T\mathbf{c}_{m+1} = \cdots = \mathbf{x}^T\mathbf{c}_R = 0. \tag{43}$$

We set $\mathbf{f} = \mathbf{S}_{m+l}(\mathbf{C})^T(\underbrace{\mathbf{x}\otimes\cdots\otimes\mathbf{x}}_{m+l})$ and we index the entries of $\mathbf{f}$ by $(m+l)$-tuples as in (23). One can easily show that $\mathbf{f}$ has entries $(\mathbf{x}^T\mathbf{c}_{r_1})\dots(\mathbf{x}^T\mathbf{c}_{r_{m+l}})$. Hence, by (43),

$$(\mathbf{x}^T\mathbf{c}_{r_1})\dots(\mathbf{x}^T\mathbf{c}_{r_{m+l}}) = 0, \text{ if } \{r_1,\dots,r_{m+l}\}\setminus\{1,\dots,m\} \ne \emptyset,$$
$$(\mathbf{x}^T\mathbf{c}_{r_1})\dots(\mathbf{x}^T\mathbf{c}_{r_{m+l}}) \ne 0, \text{ if } \{r_1,\dots,r_{m+l}\}\setminus\{1,\dots,m\} = \emptyset.$$

On the other hand, by Definition 6 and the assumption of linear dependence of the vectors $\mathbf{a}_1,\dots,\mathbf{a}_m$, the columns of $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})$ indexed by the $(m+l)$-tuples (23) such that $\{r_1,\dots,r_{m+l}\}\setminus\{1,\dots,m\} = \emptyset$ are zero. Hence, $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{f} = \mathbf{0}$. □

6.3. Properties of the matrix $\mathbf{S}_{m+l}(\mathbf{C})^T$

The following auxiliary Lemma will be used in Subsections 6.4 and 6.5. Since the proof is rather long and technical, it is included in Appendix B.

Lemma 22. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$, $k_{\mathbf{C}} = K$, $m = R - K + 2$, $l \geq 0$, *let* $\mathbf{F}$ *satisfy* (30)–(31), *and let* $\mathbf{F}^{(m+l)}$ *be defined by* (36). *Then*

(i) $\dim\left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) = C_R^{K-1}$;

(ii) $\ker(\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) = \operatorname{range}\left(\mathbf{F}^{(m+l)}\right)$;

(iii) $\operatorname{range}(\mathbf{S}_{m+l}(\mathbf{C})^T) = \mathbf{S}_{m+l}(\mathbf{C})^T(S^{m+l}(\mathbb{R}^{K^{m+l}}))$;

(iv) $\dim\left(\operatorname{range}(\mathbf{S}_{m+l}(\mathbf{C})^T)\right) = C_{K+m+l-1}^{m+l} - C_R^{K-1}$.

6.4. Proof of equivalence of Theorems 14 and 15

We prove that assumption (iii) in Theorem 14 is equivalent to assumption (iii) in Theorem 15. By (29), it is sufficient to prove that

$$\begin{gathered}\dim\left(\ker(\mathbf{R}_{m,l}(\mathcal{T}))\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) \geq C_R^{K-1} + 1 \Leftrightarrow \\ \text{the matrix } \ \mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{U}_m \text{ does not have full column rank.}\end{gathered} \tag{44}$$

To prove (44) we will use the following result for $\mathbf{X} := \mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})$, $\mathbf{Y} := \mathbf{S}_{m+l}(\mathbf{C})^T$, and $E := S^{m+l}(\mathbb{R}^{K^{m+l}})$: if E is a subspace and $\mathbf{X}$ and $\mathbf{Y}$ are matrices such that $\mathbf{XY}$ is defined, then

$$\begin{gathered}\dim\left(\ker(\mathbf{XY})\cap E\right) \geq \dim\left(\ker(\mathbf{Y})\cap E\right) + 1 \quad \Leftrightarrow \\ \text{there exists a nonzero vector } \mathbf{f} \in E \setminus \ker(\mathbf{Y}) \text{ such that } \ \mathbf{XYf} = 0.\end{gathered} \tag{45}$$

We have

$$\dim\left(\ker(\mathbf{R}_{m,l}(\mathcal{T}))\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) \geq C_R^{K-1} + 1 \qquad \overset{(18)}{\Longleftrightarrow}$$

$$\left\{\begin{aligned}\dim\left(\ker(\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) \geq C_R^{K-1} + 1 = \\ \dim\left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) + 1\end{aligned}\right\} \overset{(45)}{\Longleftrightarrow}$$

$$\left\{\begin{aligned}\text{there exists a nonzero vector } \mathbf{f} \in S^{m+l}(\mathbb{R}^{K^{m+l}}) \setminus \ker(\mathbf{S}_{m+l}(\mathbf{C})^T) \\ \text{such that } \ \mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{m+l}(\mathbf{C})^T\mathbf{f} = 0\end{aligned}\right\} \Longleftrightarrow$$

$$\text{the matrix } \ \mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{U}_m \text{ does not have full column rank,}$$

where the equality in the second statement holds by Lemma 22 (i) and the last equivalence follows from $\operatorname{range}(\mathbf{U}_m) = \operatorname{range}(\mathbf{S}_{m+l}(\mathbf{C})^T)$.

6.5. Proof of the statement on algebraic computation in Theorem 14

The overall procedure that constitutes the proof of the statement on algebraic computation is summarized in Algorithm 2 and explained in Subsection 5.2. In this subsection we prove statements (34)–(35).

Lemma 23. *Let assumptions* (i) *and* (iii) *in Theorem 15 hold and let* $\mathbf{F}$ *satisfy* (30)–(31). *Then* (34)–(35) *hold.*

Proof. The implication $k_{\mathbf{C}} = K \Rightarrow$ (34) was proved in [9, Proposition 1.10]. In Subsection 6.4 we proved that assumption (iii) in Theorem 14 holds. By (18), Theorem 14 (iii), and Lemma 22 we have

$$\begin{aligned}
&\ker(\mathbf{R}_{m,l}(\mathcal{T}))\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) = \\
&\ker(\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) = \\
&\ker(\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) = \operatorname{range}\left(\mathbf{F}^{(m+l)}\right),
\end{aligned}$$

which completes the proof of (35). □

6.6. Proof of Theorem 8

We check the assumptions in Theorem 15 for $m = 2$. Assumption (i) holds since $r_{\mathbf{C}} = K = R$ implies $k_{\mathbf{C}} = K$ and assumption (iii) coincides with (25). To prove assumption (ii) we assume to the contrary that $(\mathbf{A} \odot \mathbf{B})[\lambda_1\ \dots\ \lambda_R]^T = \mathbf{0}$. Then $r_{\mathbf{A}\mathrm{Diag}(\lambda_1,\dots,\lambda_R)\mathbf{B}^T)} = 0$. In Subsection 6.4 we explained that assumption (iii) in Theorem 14 also holds. Hence, by Lemma 20, condition (W2) holds for the triplet $(\mathbf{A}, \mathbf{B}, \mathbf{C})$. Hence, at most one of the values $\lambda_1, \dots, \lambda_R$ is not zero. If such a λ_r exists, then $\mathbf{a}_r = \mathbf{0}$ or $\mathbf{b}_r = \mathbf{0}$ yielding that $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) = 0$. On the other hand, by Lemma 21, $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq 2$, which is a contradiction. Hence, $\lambda_1 = \cdots = \lambda_R = 0$.

7. Discussion

A number of conditions (called (Km), (Cm), (Um), and (Wm)) for uniqueness of CPD of a specific tensor have been proposed in [7, 8]. It was shown that each subsequent condition in (Km), …, (Wm) is more general than the preceding one, but harder to use. Verification of conditions (Km) and (Cm) reduces to the computation of matrix rank. In contrast, conditions (Um) and (Wm) are not easy to check for a specific tensor but hold automatically for generic tensors of certain dimensions and rank [10].

In this paper we have proposed new sufficient conditions for uniqueness that can be verified by the computation of matrix rank, are more relaxed than (Km) and (Cm), but that cannot be more relaxed than (Wm). Nevertheless, examples illustrate that in many cases the new conditions may be considered as an "easy to check analogue" of (U2) ($\Leftrightarrow$ (W2)) and (Wm).

We have also proposed an algorithm to compute the factor matrices. The algorithm relies only on standard linear algebra, and has as input the tensor $\mathcal{T}$, the tensor rank R, and a nonnegative integer parameter l. The algorithm basically reduces the problem to the construction of a $C_{K+m+l-1}^{m+l} \times C_{K+m+l-1}^{m+l}$ matrix $\mathbf{Q}$, the computation of its C_R^{K-1}-dimensional null space, and the GEVD of a $C_R^{K-1} \times C_R^{K-1}$ matrix pencil, where $m = R - K + 2$. For $l = 0$, Algorithms 1 and 2 coincide with algorithms from [6] and [9], respectively. Our derivation is different from the derivations in [6] and [9] but has the same structure: from the CPD $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ we derive a set of equations that depend only on $\mathbf{C}$; we find $\mathbf{C}$ from the new system by means of GEVD, and then recover $\mathbf{A}$ and $\mathbf{B}$ from $\mathcal{T}$ and $\mathbf{C}$.

It is interesting to note that the new algorithm (with $l = 1$) computes the CPD of a generic $3 \times 7 \times 12$ tensor of rank 12 in less than 1 second while optimization-based algorithms (we checked a Gauss-Newton dogleg trust region method) fail to find the solution in a reasonable amount of time.

We have demonstrated that our algorithm (with $l \leq 2$) can find the CPD of a generic $I \times J \times K$ tensor of rank R if $R \leq K \leq (I-1)(J-1)$ and $R \leq 24$. We conjecture that the algorithm (possibly with $l \geq 3$) can also find the CPD for $R \geq 25$. (It is known that the CPD of a generic tensor is not unique if $R > (I-1)(J-1)$). In that case the $C_{K+m+l-1}^{m+l} \times C_{K+m+l-1}^{m+l}$ matrix $\mathbf{Q}$ becomes large and the computation, as it is proposed in the paper, becomes infeasible. Since the null space of $\mathbf{Q}$ is just R-dimensional the approach may possibly be scaled by using iterative methods to compute the null space.

Appendix A. Derivation of identity (18)

Let $\mathcal{T} = (t_{ijk})_{i,j,k=1}^{I,J,K} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. In this section we establish a link between the matrix $\mathbf{R}_{m,l}(\mathcal{T})$ defined in subsection 2 and the factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. We show that the matrix $\mathbf{R}_{m,l}(\mathcal{T})$ is obtained from $\mathcal{T}$ by taking the following steps: 1) taking the $(m + l)$th Kronecker power of $\mathcal{T}$; 2) making two partial skew-symmetrizations and one partial symmetrization of the result; 3) reshaping the result into an $I^{m+l}J^{m+l} \times K^{m+l}$ matrix. The main identity is obtained by applying steps 1)–3) to the both sides of (1).

Appendix A.1. Step 1: Kronecker product power of $\mathcal{T}$

The Kronecker product square of $\mathcal{T}$, $\mathcal{T}^{(2)} := \mathcal{T} \otimes \mathcal{T}$, is an $I \times J \times K$ block-tensor whose (i, j, k)th block is the $I \times J \times K$ tensor $t_{ijk}\mathcal{T}$. Equivalently, $\mathcal{T}^{(2)}$ is an $I^2 \times J^2 \times K^2$ tensor whose $(\tilde{i}, \tilde{j}, \tilde{k}) := (i_1 - 1)I + i_2, (j_1 - 1)J + j_2, (k_1 - 1)K + k_2)$th entry is

$$t^{(2)}_{\tilde{i}\tilde{j}\tilde{k}} = t_{i_1 j_1 k_1} t_{i_2 j_2 k_2}.$$

Similarly, the $(l + m)$-th Kronecker product power of $\mathcal{T}$,

$$\mathcal{T}^{(m+l)} := \underbrace{\mathcal{T} \otimes \cdots \otimes \mathcal{T}}_{m+l},$$

is an $I^{m+l} \times J^{m+l} \times K^{m+l}$ tensor whose $(\tilde{i}, \tilde{j}, \tilde{k})$th entry is

$$t^{(m+l)}_{\tilde{i}\tilde{j}\tilde{k}} = t_{i_1 j_1 k_1} t_{i_2 j_2 k_2} \cdots t_{i_{m+l} j_{m+l} k_{m+l}}, \tag{A.1}$$

where $\tilde{i}$,$\tilde{j}$, and $\tilde{k}$ are defined in (19), (20), and (21), respectively. One can easily check that if $\mathcal{T} = \sum\limits_{r=1}^{R} \mathbf{a}_r \circledast \mathbf{b}_r \circledast \mathbf{c}_r$, then

$$\mathcal{T}^{(m+l)} = \sum_{r_1, \dots, r_{m+l}=1}^{R} (\mathbf{a}_{r_1} \otimes \cdots \otimes \mathbf{a}_{r_{m+l}}) \circledast (\mathbf{b}_{r_1} \otimes \cdots \otimes \mathbf{b}_{r_{m+l}}) \circledast (\mathbf{c}_{r_1} \otimes \cdots \otimes \mathbf{c}_{r_{m+l}}).$$

Appendix A.2. Step 2: two partial skew-symmetrizations and one partial symmetrization of a reshaped version of $\mathcal{T}^{(m+l)}$

Recall that a higher-order tensor is said to be *symmetric (resp. skew-symmetric) with respect to a given group of indices* or *partially symmetric (resp. skew-symmetric)* if its coordinates do not alter by an arbitrary permutation of these indices (resp. if the sign changes with every interchange of two arbitrary indices in the group).

Let us recall the operations of (complete) symmetrization and skew-symmetrization. With a general kth-order $L \times \cdots \times L$ tensor $\mathcal{N}$ one can associate its symmetric part $S^k(\mathcal{N})$ and skew-symmetric part $\Lambda^k(\mathcal{N})$ as follows. By construction, $S^k(\mathcal{N})$ is a tensor whose entry with indices $l_1, \dots, l_k$ is equal to

$$\frac{1}{k!} \sum_{(p_1,\dots,p_k) \in P_{\{l_1,\dots,l_k\}}} n_{p_1 \dots p_k}. \tag{A.2}$$

That is, to get $S^k(\mathcal{N})$ we should take the average of $k!$ tensors obtained from $\mathcal{N}$ by all possible permutations of the indices. Similarly, $\Lambda^k(\mathcal{N})$ is a tensor whose entry with indices $l_1, \dots, l_k$ is equal to

$$\begin{cases} \frac{1}{k!} \sum\limits_{(p_1,\dots,p_k)\in P_{\{l_1,\dots,l_k\}}} \operatorname{sgn}(p_1,\dots,p_k) n_{p_1\dots p_k}, & \text{if } l_1,\dots,l_k \text{ are distinct,} \\ 0, & \text{otherwise,} \end{cases} \tag{A.3}$$

where

$\operatorname{sgn}(p_1, \dots, p_k)$ denotes the signature of the permutation $(p_1, \dots, p_k)$.

The definition of $\Lambda^k(\mathcal{N})$ differs from that of $S^k(\mathcal{N})$ in that the signatures of the permutations are taken into account and that the entries of $\Lambda^k(\mathcal{N})$ with repeated indices are necessarily zeros. One can easily check that if $\mathcal{N} = \mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k$ (that is, $\mathcal{N}$ is the kth order rank-1 tensor), then

$$S^k(\mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k) = \frac{1}{k!} \sum_{(p_1,\dots,p_k)\in P_{\{1,\dots,k\}}} \mathbf{d}_{p_1} \otimes \dots \otimes \mathbf{d}_{p_k}, \tag{A.4}$$

$$\Lambda^k(\mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k) = \frac{1}{k!} \sum_{(p_1,\dots,p_k)\in P_{\{1,\dots,k\}}} \sigma(p_1,\dots,p_k)\mathbf{d}_{p_1} \otimes \dots \otimes \mathbf{d}_{p_k}. \tag{A.5}$$

Partial (skew-)symmetrization is a (skew-)symmetrization with respect to a given group of indices. Instead of presenting the formal definitions we illustrate both notions for an $M \times L \times L$ tensor $\mathcal{N} = (n_{ml_1l_2})_{m,l_1,l_2=1}^{M,L,L}$. Partial symmetrization with respect to the group of indices $\{2, 3\}$ maps the tensor $\mathcal{N}$ to a tensor that we denote by $(\mathbf{I}_M \otimes S^2)\mathcal{N}$, whose entry with indices (m, l_1, l_2) is equal to

$$\sum_{(p_1,p_2)\in P_{\{l_1,l_2\}}} n_{ml_1l_2} = n_{ml_1l_2} + n_{ml_2l_1}.$$

Similarly, by $(\mathbf{I}_M \otimes \Lambda^2)\mathcal{N}$ we denote the tensor whose entry with indices (m, l_1, l_2) is equal to

$$\begin{cases} \sum\limits_{(p_1,p_2)\in P_{\{l_1,l_2\}}} \operatorname{sgn}(p_1,p_2) n_{ml_1l_2} = n_{ml_1l_2} - n_{ml_2l_1}, & \text{if } \ l_1 \neq l_2, \\ 0, & \text{if } \ l_1 = l_2. \end{cases}$$

If $\mathcal{N} = \mathbf{d}_1 \otimes \mathbf{d}_2 \otimes \mathbf{d}_3 \in \mathbb{R}^{M\times L\times L}$, then

$$(\mathbf{I}_M \otimes S^2)(\mathbf{d}_1 \otimes \mathbf{d}_2 \otimes \mathbf{d}_3) = \mathbf{d}_1 \otimes S^2(\mathbf{d}_2 \otimes \mathbf{d}_3) = \mathbf{d}_1 \otimes \mathbf{d}_2 \otimes \mathbf{d}_3 + \mathbf{d}_1 \otimes \mathbf{d}_3 \otimes \mathbf{d}_2, \tag{A.6}$$

$$(\mathbf{I}_M \otimes \Lambda^2)(\mathbf{d}_1 \otimes \mathbf{d}_2 \otimes \mathbf{d}_3) = \mathbf{d}_1 \otimes \Lambda^2(\mathbf{d}_2 \otimes \mathbf{d}_3) = \mathbf{d}_1 \otimes \mathbf{d}_2 \otimes \mathbf{d}_3 - \mathbf{d}_1 \otimes \mathbf{d}_3 \otimes \mathbf{d}_2. \tag{A.7}$$

Thus, operations $(\mathbf{I}_M \otimes S^2)$ and $(\mathbf{I}_M \otimes \Lambda^2)$ symmetrize and skew-symmetrize the horizontal slices of $\mathcal{N}$.

Let us reshape the tensor $\mathcal{T}^{(m+l)}$ into an $I\times\cdots\times I\times J\times\cdots\times J\times K\times\cdots\times K$ (each letter is repeated $m+l$ times) tensor $\widehat{\mathcal{T}}^{(m+l)}$ as:

$$\widehat{\mathcal{T}}^{(m+l)} = \sum_{r_1,\dots,r_{m+l}=1}^{R} (\mathbf{a}_{r_1}\otimes\dots\otimes\mathbf{a}_{r_{m+l}})\otimes(\mathbf{b}_{r_1}\otimes\dots\otimes\mathbf{b}_{r_{m+l}})\otimes(\mathbf{c}_{r_1}\otimes\dots\otimes\mathbf{c}_{r_{m+l}}). \tag{A.8}$$

Then the entries of $\widehat{\mathcal{T}}^{(m+l)}$ are given by

$$\widehat{t}^{(m+l)}_{i_1\dots i_{m+l}j_1\dots j_{m+l}k_1\dots k_{m+l}} = t_{i_1j_1k_1}t_{i_2j_2k_2}\cdots t_{i_{m+l}j_{m+l}k_{m+l}}. \tag{A.9}$$

From (A.1) and (A.9) it follows that $\widehat{\mathcal{T}}^{(m+l)}$ is just a higher-order representation of $\mathcal{T}^{(m+l)}$.

A new tensor $\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}$ is obtained from $\widehat{\mathcal{T}}^{(m+l)}$ by applying two partial skew-symmetrizations and one partial symmetrization as follows:

$$\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S} := \left[(\Lambda^m \otimes \mathbf{I}_I \otimes \dots \otimes \mathbf{I}_I) \otimes (\Lambda^m \otimes \mathbf{I}_J \otimes \dots \otimes \mathbf{I}_J) \otimes S^{m+l}\right] \widehat{\mathcal{T}}^{(m+l)}. \tag{A.10}$$

To obtain $\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}$ we first skew-symmetrize $\widehat{\mathcal{T}}^{(m+l)}$ with respect to the group of indices $\{1,\dots,m\}$ (the first m "I" dimensions), then we skew-symmetrize the result with respect to the group of indices $\{m+l+1,\dots,2m+l\}$ (the first m "J" dimensions), and, finally, we symmetrize the result with respect to the group of indices $\{2m+2l+1,\dots,3m+3l\}$ (all "K" dimensions). From (A.2), (A.3), and (A.9), it follows that the $(i_1,\dots,i_{m+l},j_1,\dots,j_{m+l},k_1,\dots,k_{m+l})$th entry of the tensor $\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}$ is equal to zero if some index is repeated in $i_1,\dots,i_m$ or $j_1,\dots,j_m$ and is equal to

$$\frac{1}{(m+l)!} \sum_{\substack{(s_1,\dots,s_{m+l})\in \\ P_{\{k_1,\dots,k_{m+l}\}}}} \left[\frac{1}{m!} \sum_{\substack{(q_1,\dots,q_m)\in \\ P_{\{j_1,\dots,j_m\}}}} \operatorname{sgn}(q_1,\dots,q_m)\times \right.$$
$$\left.\left(\frac{1}{m!} \sum_{\substack{(p_1,\dots,p_m)\in \\ P_{\{i_1,\dots,i_m\}}}} \operatorname{sgn}(p_1,\dots,p_m) \prod_{u=1}^{m} t_{p_uq_us_u} \prod_{v=1}^{l} t_{i_{m+v}j_{m+v}s_{m+v}}\right)\right] =$$

$$\frac{1}{(m+l)!}\sum_{\substack{(s_1,\dots,s_{m+l})\in\\ P_{\{k_1,\dots,k_{m+l}\}}}}\left[\frac{1}{m!}\sum_{\substack{(q_1,\dots,q_m)\in\\ P_{\{j_1,\dots,j_m\}}}}\operatorname{sgn}(q_1,\dots,q_m)\times\right.$$

$$\left.\frac{1}{m!}\det\begin{bmatrix} t_{i_1q_1s_1} & \dots & t_{i_1q_ms_m}\\ \vdots & \vdots & \vdots\\ t_{i_mq_1s_1} & \dots & t_{i_mq_ms_m}\end{bmatrix}\prod_{v=1}^{l} t_{i_{m+v}j_{m+v}s_{m+v}}\right] =$$

$$\frac{1}{m!(m+l)!}\sum_{\substack{(s_1,\dots,s_{m+l})\in\\ P_{\{k_1,\dots,k_{m+l}\}}}}\det\begin{bmatrix} t_{i_1j_1s_1} & \dots & t_{i_1j_ms_m}\\ \vdots & \vdots & \vdots\\ t_{i_mj_1s_1} & \dots & t_{i_mj_ms_m}\end{bmatrix}\prod_{v=1}^{l} t_{i_{m+v}j_{m+v}s_{m+v}},$$

otherwise (we used twice the Leibniz formula for the determinant). Thus, by (22), the tensor $\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}$ and the matrix $\mathbf{R}_{m,l}(\mathcal{T})$ have the same entries (in step 3 it will be shown that $\mathbf{R}_{m,l}(\mathcal{T})$ is a matrix unfolding of $\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}$).

Let us apply partial skew-symmetrizations and partial symmetrization to the right-hand side of (A.8): from (A.8), (A.10), (see also (A.6)–(A.7) for the properties of the outer product) it follows that

$$\begin{aligned}\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S} = &\sum_{r_1,\dots,r_{m+l}=1}^{R}\big[\Lambda^m(\mathbf{a}_{r_1}\otimes\dots\otimes\mathbf{a}_{r_m})\otimes\mathbf{a}_{r_{m+1}}\otimes\dots\otimes\mathbf{a}_{r_{m+l}}\otimes\\ &\Lambda^m(\mathbf{b}_{r_1}\otimes\dots\otimes\mathbf{b}_{r_m})\otimes\mathbf{b}_{r_{m+1}}\otimes\dots\otimes\mathbf{b}_{r_{m+l}}\big]\otimes S^{m+l}(\mathbf{c}_{r_1}\otimes\dots\otimes\mathbf{c}_{r_{m+l}}) = \\ &\sum_{r_1,\dots,r_{m+l}=1}^{R}\mathcal{F}^{\mathbf{A},\mathbf{B}}_{r_1,\dots,r_{m+l}}\otimes S^{m+l}(\mathbf{c}_{r_1}\otimes\dots\otimes\mathbf{c}_{r_{m+l}}),\end{aligned}\tag{A.11}$$

where the expressions $\Lambda^m(\mathbf{a}_{r_1}\otimes\dots\otimes\mathbf{a}_{r_m})$ and $\Lambda^m(\mathbf{b}_{r_1}\otimes\dots\otimes\mathbf{b}_{r_m})$ are defined in (A.5), the expression $S^{m+l}(\mathbf{c}_{r_1}\otimes\dots\otimes\mathbf{c}_{r_{m+l}})$ is defined in (A.4), and, by definition,

$$\begin{aligned}\mathcal{F}^{\mathbf{A},\mathbf{B}}_{r_1,\dots,r_{m+l}} :=&\Lambda^m(\mathbf{a}_{r_1}\otimes\dots\otimes\mathbf{a}_{r_m})\otimes\mathbf{a}_{r_{m+1}}\otimes\dots\otimes\mathbf{a}_{r_{m+l}}\otimes\\ &\Lambda^m(\mathbf{b}_{r_1}\otimes\dots\otimes\mathbf{b}_{r_m})\otimes\mathbf{b}_{r_{m+1}}\otimes\dots\otimes\mathbf{b}_{r_{m+l}}\end{aligned}$$

(recall that the vectors $\mathbf{a}_r$ and $\mathbf{b}_r$ are columns of the matrices $\mathbf{A}$ and $\mathbf{B}$, respectively).

Note that by construction, $S^{m+l}(\mathbf{c}_{r_1}\otimes\dots\otimes\mathbf{c}_{r_{m+l}})$ is a completely symmetric tensor (that is, the expression $S^{m+l}(\mathbf{c}_{r_1}\otimes\dots\otimes\mathbf{c}_{r_{m+l}})$ does not change after any permutation of the vectors $\mathbf{c}_{r_1}, \dots, \mathbf{c}_{r_{m+l}}$). Taking this fact into account we can group the summands in (A.11) as follows

$$\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S} = \sum_{1 \le r_1 \le \dots \le r_{m+l} \le R} \left(\sum_{\substack{(p_1,\dots,p_{m+l}) \in \\ P_{\{r_1,\dots,r_{m+l}\}}}} \mathcal{F}^{\mathbf{A},\mathbf{B}}_{p_1,\dots,p_{m+l}} \right) {\scriptstyle\otimes}\, S^{m+l}(\mathbf{c}_{r_1} {\scriptstyle\otimes} \dots {\scriptstyle\otimes}\, \mathbf{c}_{r_{m+l}}). \tag{A.12}$$

Appendix A.3. Step 3: Reshaping (unfolding) of $\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}$ *into the matrix* $\mathbf{R}_{m,l}(\mathcal{T})$

We define the matricization operation

$$\text{Matr}: \ \mathbb{R}^{I\times\dots\times I\times J\times\dots\times J\times K\times\dots\times K} \ \to \ \mathbb{R}^{I^{m+l}J^{m+l}\times K^{m+l}}$$

as follows: the $(i_1, \dots, i_{m+l}, j_1, \dots, j_{m+l}, k_1, \dots, k_{m+l})$th entry of a tensor is mapped to the $((\tilde{i}-1)J^{m+l} + \tilde{j}, \tilde{k})$th entry of a matrix, where $\tilde{i}$, $\tilde{j}$, and $\tilde{k}$ are defined in (19), (20), and (21), respectively. One can easily verify that

$$\begin{aligned} &\text{Matr}\left(\mathbf{a}_{i_1} {\scriptstyle\otimes} \dots {\scriptstyle\otimes}\, \mathbf{a}_{i_{m+l}} {\scriptstyle\otimes}\, \mathbf{b}_{j_1} {\scriptstyle\otimes} \dots {\scriptstyle\otimes}\, \mathbf{b}_{j_{m+l}} {\scriptstyle\otimes}\, \mathbf{c}_{k_1} {\scriptstyle\otimes} \dots {\scriptstyle\otimes}\, \mathbf{c}_{k_{m+l}}\right) = \\ &\qquad \left[\mathbf{a}_{i_1} \otimes \dots \otimes \mathbf{a}_{i_{m+l}} \otimes \mathbf{b}_{j_1} \otimes \dots \otimes \mathbf{b}_{j_{m+l}}\right] (\mathbf{c}_{k_1} \otimes \dots \otimes \mathbf{c}_{k_{m+l}})^T \end{aligned} \tag{A.13}$$

and that $\mathbf{R}_{m,l}(\mathcal{T}) = \text{Matr}(\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S})$.

What is left to show is that the matricization of the right-hand side of (A.12) coincides with the matrix $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})\mathbf{S}_{m+l}(\mathbf{C})^T$. In the sequel, when no confusion is possible, we will use S^k and Λ^k to denote "symmetrization" and "skew-symmetrization" of vector representations of a certain tensor: if $\mathbf{d}_1, \dots, \mathbf{d}_k \in \mathbb{R}^L$, then the vectors $S^k(\mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k)$ and $\Lambda^k(\mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k)$ are computed in the same way as in (A.4)–(A.5) but with "${\scriptstyle\otimes}$" replaced by "$\otimes$":

$$\begin{aligned} S^k(\mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k) &= \frac{1}{k!} \sum_{(p_1,\dots,p_k) \in P_{\{1,\dots,k\}}} \mathbf{d}_{p_1} \otimes \dots \otimes \mathbf{d}_{p_k}, \\ \Lambda^k(\mathbf{d}_1 \otimes \dots \otimes \mathbf{d}_k) &= \frac{1}{k!} \sum_{(p_1,\dots,p_k) \in P_{\{1,\dots,k\}}} \sigma(p_1,\dots,p_k) \mathbf{d}_{p_1} \otimes \dots \otimes \mathbf{d}_{p_k}. \end{aligned} \tag{A.14}$$

Hence, by (A.11), (A.12), and (A.13)

$$\begin{aligned} &\mathbf{R}_{m,l}(\mathcal{T}) = \text{Matr}(\widehat{\mathcal{T}}^{(m+l)}_{\Lambda\Lambda S}) = \\ &\sum_{1 \le r_1 \le \dots \le r_{m+l} \le R} \left(\sum_{\substack{(s_1,\dots,s_{m+l}) \in \\ P_{\{r_1,\dots,r_{m+l}\}}}} \mathbf{f}^{\mathbf{A},\mathbf{B}}_{s_1,\dots,s_{m+l}} \right) S^{m+l}(\mathbf{c}_{r_1} \otimes \dots \otimes \mathbf{c}_{r_{m+l}})^T = \end{aligned} \tag{A.15}$$

$$\sum_{1\le r_1\le\cdots\le r_{m+l}\le R} \phi(\mathbf{A},\mathbf{B})_{r_1,\dots,r_{m+l}} S^{m+l}(\mathbf{c}_{r_1}\otimes\cdots\otimes\mathbf{c}_{r_{m+l}})^T,$$

where

$$\begin{aligned}\mathbf{f}^{\mathbf{A},\mathbf{B}}_{s_1,\dots,s_{m+l}} :=& \Lambda^m(\mathbf{a}_{s_1}\otimes\cdots\otimes\mathbf{a}_{s_m})\otimes\mathbf{a}_{s_{m+1}}\otimes\cdots\otimes\mathbf{a}_{s_{m+l}}\otimes\\ &\Lambda^m(\mathbf{b}_{s_1}\otimes\cdots\otimes\mathbf{b}_{s_m})\otimes\mathbf{b}_{s_{m+1}}\otimes\cdots\otimes\mathbf{b}_{s_{m+l}},\\ \phi(\mathbf{A},\mathbf{B})_{r_1,\dots,r_{m+l}} :=& \sum_{\substack{(s_1,\dots,s_{m+l})\in\\ P_{\{r_1,\dots,r_{m+l}\}}}} \mathbf{f}^{\mathbf{A},\mathbf{B}}_{s_1,\dots,s_{m+l}}.\end{aligned}$$

We show that $\phi(\mathbf{A},\mathbf{B})_{r_1,\dots,r_{m+l}}$ is the zero vector if the set $\{r_1,\dots,r_{m+l}\}$ has fewer than m distinct elements and that $\phi(\mathbf{A},\mathbf{B})_{r_1,\dots,r_{m+l}}$ is a column of $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})$ otherwise. From (A.14) and the Leibniz formula for the determinant it follows that the entries of the vector $\Lambda^k(\mathbf{d}_1\otimes\cdots\otimes\mathbf{d}_k)$ are all possible $k\times k$ minors of the matrix $\mathbf{D} := [\mathbf{d}_1\ \dots\ \mathbf{d}_k]$ divided by $k!$. In particular, if some of the vectors $\mathbf{d}_i$ coincide, then $\Lambda^k(\mathbf{d}_1\otimes\cdots\otimes\mathbf{d}_k)$ is the zero vector. Hence, the vector $\mathbf{f}^{\mathbf{A},\mathbf{B}}_{s_1,\dots,s_{m+l}}$ has entries

$$\frac{1}{(m!)^2}\det\begin{bmatrix} a_{i_1 s_1} & \dots & a_{i_1 s_m}\\ \vdots & \vdots & \vdots\\ a_{i_m s_1} & \dots & a_{i_m s_m}\end{bmatrix}\cdot\det\begin{bmatrix} b_{j_1 s_1} & \dots & b_{j_1 s_m}\\ \vdots & \vdots & \vdots\\ b_{j_m s_1} & \dots & b_{j_m s_m}\end{bmatrix}\cdot$$
$$a_{i_{m+1}s_{m+1}}\cdots a_{i_{m+l}s_{m+l}}\cdot b_{j_{m+1}s_{m+1}}\cdots b_{j_{m+l}s_{m+l}},$$

where $i_1,\dots,i_{m+l}\in\{1,\dots,I\}$ and $j_1,\dots,j_{m+l}\in\{1,\dots,J\}$. In particular, if the set $\{r_1,\dots,r_{m+l}\}$ has fewer than m distinct elements, then $\mathbf{f}^{\mathbf{A},\mathbf{B}}_{s_1,\dots,s_{m+l}}$ are zero vectors for all $(s_1,\dots,s_{m+l})\in P_{\{r_1,\dots,r_{m+l}\}}$, yielding that $\phi(\mathbf{A},\mathbf{B})_{r_1,\dots,r_{m+l}}$ is the zero vector. Hence, by Definition 6, the matrix $\mathbf{\Phi}_{m,l}(\mathbf{A},\mathbf{B})$ has columns $\phi(\mathbf{A},\mathbf{B})_{r_1,\dots,r_{m+l}}$, where $(r_1,\dots,r_{m+l})$ satisfies (23). Thus, (A.15) coincides with (18).

Appendix B. Proof of Lemma 22

The proof of Lemma 22 is based on the following simple generalization of the rank-nullity theorem and relies on two bounds that will be obtained in in Appendix B.1 and Appendix B.2, respectively.

Lemma 24. *Let* $\mathbf{X}$ *be a matrix and* E *be a subspace such that* $\operatorname{range}(\mathbf{X}^T)\subseteq E$. *Then*

(i) $\dim(\ker(\mathbf{X}) \cap E) + \dim(\mathbf{X}(E)) = \dim E$*;*

(ii) $\operatorname{range}(\mathbf{X}) = \mathbf{X}(E)$*,*

where the subspace $\mathbf{X}(E)$ *denotes the image of* E *under* $\mathbf{X}$*.*

Proof. Let $\mathbf{P}$ be a matrix whose columns form a basis for the subspace E. Then $\dim(\ker(\mathbf{X}) \cap E) = \dim(\ker(\mathbf{XP}))$, $\mathbf{X}(E) = \operatorname{range}(\mathbf{XP})$, and the matrix $\mathbf{XP}$ has $\dim E$ columns. Hence, by the rank-nullity theorem,

$$\dim(\ker(\mathbf{X}) \cap E) + \dim(\mathbf{X}(E)) = \dim(\ker(\mathbf{XP})) + \dim(\mathbf{XP}) = \dim E.$$

Since, $\operatorname{range}(\mathbf{X}^T) \subseteq \operatorname{range}(\mathbf{P})$, it follows that $\operatorname{range}(\mathbf{X}) = \operatorname{range}(\mathbf{XP}) = \mathbf{X}(E)$. □

Proof of Lemma 22. We set $\mathbf{X} = \mathbf{S}_{m+l}(\mathbf{C})^T$ and $E = S^{m+l}(\mathbb{R}^{K^{m+l}})$. Then statement (iii) follows from Lemma 24 (ii). By Lemma 24 (i),

$$\dim\left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) + r_{\mathbf{S}_{m+l}(\mathbf{C})^T} = C^{m+l}_{K+m+l-1}. \qquad \text{(B.1)}$$

In Appendix B.1 and Appendix B.2 we prove that the summands in (B.1) are bounded as

$$\dim\left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right) \geq C^{K-1}_R, \text{ and}$$
$$r_{\mathbf{S}_{m+l}(\mathbf{C})^T} \geq C^{m+l}_{R+l+1} - C^{m-1}_R,$$

respectively. Since $C^{K-1}_R + (C^{m+l}_{R+l+1} - C^{m-1}_R) = C^{m-1}_R + (C^{m+l}_{K+m+l-1} - C^{m-1}_R) = C^{m+l}_{K+m+l-1}$, statements (i) and (iv) follow from (B.1). Statement (ii) follows from statement (i) and Lemma 25 below. □

Appendix B.1. A lower bound on $\dim\left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right)$

In this subsection we prove the following result.

Lemma 25. *Let* $\mathbf{C} \in \mathbb{R}^{K \times R}$*,* $k_{\mathbf{C}} = K$*,* $m = R - K + 2$*,* $l \geq 0$*, let* $\mathbf{F}$ *satisfy* (30)–(31)*, and let* $\mathbf{F}^{(m+l)}$ *be defined by* (36)*. Then*

(i) *The matrix* $\mathbf{F}^{(m+l)}$ *has full column rank, that is* $r_{\mathbf{F}^{(m+l)}} = C^{K-1}_R$*;*

(ii) $\ker(\mathbf{S}_{m+l}(\mathbf{C})^T) \bigcap S^{m+l}(\mathbb{R}^{K^{m+l}}) \supset \operatorname{range}\left(\mathbf{F}^{(m+l)}\right)$.

In particular, $\dim\left(\ker(\mathbf{S}_{m+l}(\mathbf{C})^T)\bigcap S^{m+l}(\mathbb{R}^{K^{m+l}})\right)\geq C_R^{K-1}$.

Proof. Statement (i) was proved in [9, Proposition 1.10].

Let $\mathbf{f}$ be a column of the matrix $\mathbf{F}$. Then $\mathbf{f}^{(m+l)}$ is a column of $\mathbf{F}^{(m+l)}$. It is clear that $\mathbf{f}^{(m+l)}\in S^{m+l}(\mathbb{R}^{K^{m+l}})$. To prove (ii) we need to show that $\mathbf{S}_{m+l}(\mathbf{C})^T\mathbf{f}^{(m+l)}=\mathbf{0}$. By Definition 7, the $(r_1,\dots,r_{m+l})$th entry of the vector $\mathbf{S}_{m+l}(\mathbf{C})^T\mathbf{f}^{(m+l)}$ is

$$\left(\frac{1}{(m+l)!}\sum_{\substack{(s_1,\dots,s_{m+l})\in\\ P_{\{r_1,\dots,r_{m+l}\}}}}\mathbf{c}_{s_1}\otimes\cdots\otimes\mathbf{c}_{s_{m+l}}\right)^T\mathbf{f}^{(m+l)}=$$

$$\frac{1}{(m+l)!}\sum_{\substack{(s_1,\dots,s_{m+l})\in\\ P_{\{r_1,\dots,r_{m+l}\}}}}(\mathbf{c}_{s_1}^T\mathbf{f})\cdots(\mathbf{c}_{s_{m+l}}^T\mathbf{f})=(\mathbf{c}_{r_1}^T\mathbf{f})\cdots(\mathbf{c}_{r_{m+l}}^T\mathbf{f}).$$

Since the vector $\mathbf{f}$ is orthogonal to exactly $K-1$ columns of $\mathbf{C}$, the fact that at least m indices of $r_1,\dots,r_{m+l}$ are distinct, and $m+l\geq R-K+2+l>R-(K-1)$ it follows that $(\mathbf{c}_{r_1}^T\mathbf{f})\cdots(\mathbf{c}_{r_{m+l}}^T\mathbf{f})=0$, which completes the proof of (ii). □

Appendix B.2. A lower bound on the rank of the matrix $\mathbf{S}_{m+l}(\mathbf{C})$

We need some additional notation. Let $m\geq 2$, $l\geq 0$, $p\geq 0$ and $m\leq m+l-p\leq R$. With an $(m+l-p)$-tuple

$$(i_1,\dots,i_{m+l-p})\text{ such that }1\leq i_1<\cdots<i_{m+l-p}\leq R$$

we associate the set

$$E_{i_1,\dots,i_{m+l-p}}:=\{(i_1,\dots,i_{m+l-p},i_{q_1},\dots,i_{q_p}):\ 1\leq q_1\leq\cdots\leq q_p\leq 2+l-p\}.$$

In other words, the set $E_{i_1,\dots,i_{m+l-p}}$ consists of the $(m+l)$-tuples that are obtained by merging the $(m+l-p)$-tuple $(i_1,\dots,i_{m+l-p})$ with p-combinations with repetitions of the set $\{i_1,\dots,i_{2+l-p}\}$. It is clear that for a fixed p there exist C_R^{m+l-p} sets $E_{i_1,\dots,i_{m+l-p}}$ and each set $E_{i_1,\dots,i_{m+l-p}}$ contains $C_{2+l-p+p-1}^p=C_{l+1}^p$ $(m+l)$-tuples. Let E be the union of all sets $E_{i_1,\dots,i_{m+l-p}}$. Then the set E contains exactly

$$C_{l+1}^0C_R^{m+l}+C_{l+1}^1C_R^{m+l-1}+\cdots+C_{l+1}^lC_R^m=C_{R+l+1}^{m+l}-C_R^{m-1}$$

$(m+l)$-tuples (we follow the convention that $C_R^{m+l-p} := 0$ if $m+l-p > R$). Since, by construction, each $(m+l)$-tuple of $E_{i_1,\dots,i_{m+l-p}}$ contains exactly $m+l-p \geq m$ distinct elements, it follows that each $(m+l)$-tuple of E contains at least m distinct elements. Let $\mathbf{S}_{m+l}^E(\mathbf{C})$ denote the $K^m \times (C_{R+l+1}^{m+l} - C_R^{m-1})$ matrix with columns (24), where $(r_1, \dots, r_{m+l}) \in E$. Then $\mathbf{S}_{m+l}^E(\mathbf{C})$ is a submatrix of $\mathbf{S}_{m+l}(\mathbf{C})$. We have the following lemma.

Lemma 26. *Let $\mathbf{C} \in \mathbb{R}^{K\times R}$, $k_{\mathbf{C}} = K$, and $m = R-K+2$. Then the matrix $\mathbf{S}_{m+l}^E(\mathbf{C})$ has full column rank. In particular, $r_{\mathbf{S}_{m+l}(\mathbf{C})} \geq C_{R+l+1}^{m+l} - C_R^{m-1}$.*

Proof. Suppose that there exists $\mathbf{f} \in \mathbb{R}^{C_{R+l+1}^{m+l} - C_R^{m-1}}$ such that $\mathbf{S}_{m+l}^E(\mathbf{C})\mathbf{f} = \mathbf{0}$. We show that $\mathbf{f} = \mathbf{0}$. We assume that the entries of $\mathbf{f}$ are indexed by $(m+l)$-tuples $(r_1, \dots, r_{m+l}) \in E$, that is, in $\mathbf{S}_{m+l}^E(\mathbf{C})\mathbf{f} = \mathbf{0}$ the column of $\mathbf{S}_{m+l}^E(\mathbf{C})$ associated with the $(m+l)$-tuple $(i_1, \dots, i_{m+l-p}, i_{q_1}, \dots, i_{q_p})$ is multiplied by $f_{i_1,\dots,i_{m+l-p},i_{q_1},\dots,i_{q_p}}$.

To show that all entries $f_{i_1,\dots,i_{m+l-p},i_{q_1},\dots,i_{q_p}}$ are zero we proceed by induction on $p = l, l-1, \dots, \max(0, m+l-R)$: in the pth step we assume that the identities

$$f_{i_1,\dots,i_{m+l-\tilde{p}},i_{q_1},\dots,i_{q_{\tilde{p}}}} = 0, \quad \text{where}$$
$$1 \leq i_1 < \dots < i_{m+l-\tilde{p}} \leq R, \quad 1 \leq q_1 \leq \dots \leq q_{\tilde{p}} \leq 2 + l - \tilde{p}$$

hold for $\tilde{p} = l, l-1, \dots, p-1$ and prove that the identities hold for $\tilde{p} = p$.

(i) Induction hypothesis: $p = l$. We show that

$$f_{i_1,\dots,i_m,i_{q_1},\dots,i_{q_l}} = 0 \quad \text{for} \quad 1 \leq i_1 < \dots < i_m \leq R, \quad 1 \leq q_1 \leq \dots \leq q_l \leq 2.$$

We give the proof for the case $i_1 = 1, \dots, i_m = m$, the other cases follow similarly. Thus, we show that $f_{1,\dots,m,q_1,\dots,q_l} = 0$ for $1 \leq q_1 \leq \dots \leq q_l \leq 2$.

Since $k_{\mathbf{C}} = K$, the square matrix $\widetilde{\mathbf{C}} := [\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_{m+1} \dots\ \mathbf{c}_R]$ is nonsingular. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ denote the first and the second column of $\widetilde{\mathbf{C}}^{-T}$, respectively. Then

$$\begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{bmatrix} \widetilde{\mathbf{C}} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}. \tag{B.2}$$

Let $\mathbf{x} = t_1\mathbf{u}_1 + t_2\mathbf{u}_2$. Then the vector $\mathbf{x}^{(m+l)} := \mathbf{x} \otimes \dots \otimes \mathbf{x}$ is orthogonal to the columns of the matrix $\mathbf{S}_{m+l}^E(\mathbf{C})$ indexed by the $(m+l)$-tuples $(r_1, \dots, r_{m+l}) \in E \setminus E_{1,\dots,m}$. Indeed, if $\{r_1, \dots, r_{m+l}\} \setminus \{1, \dots, m\} \neq \emptyset$, then by (24) and (B.2),

$$\frac{\mathbf{x}^{(m+l)T}}{(m+l)!} \sum_{\substack{(s_1,\dots,s_{m+l}) \in \\ P_{\{r_1,\dots,r_{m+l}\}}}} \mathbf{c}_{s_1} \otimes \dots \otimes \mathbf{c}_{s_{m+l}} = (\mathbf{x}^T\mathbf{c}_{r_1}) \cdots (\mathbf{x}^T\mathbf{c}_{r_{m+l}}) = 0. \tag{B.3}$$

Hence

$$0 = \mathbf{x}^{(m+l)T}\mathbf{S}^{E}_{m+l}(\mathbf{C})\mathbf{f} = \sum_{(r_1,\dots,r_{m+l})\in E_{1,\dots,m}} (\mathbf{x}^T\mathbf{c}_{r_1})\cdots(\mathbf{x}^T\mathbf{c}_{r_{m+l}})f_{r_1,\dots,r_{m+l}} =$$
$$\sum_{1\le q_1\le\cdots\le q_l\le 2} (\mathbf{x}^T\mathbf{c}_1)\cdots(\mathbf{x}^T\mathbf{c}_m)(\mathbf{x}^T\mathbf{c}_{q_1})\cdots(\mathbf{x}^T\mathbf{c}_{q_l})f_{1,\dots,m,q_1,\dots,q_l} =$$
$$(\mathbf{x}^T\mathbf{c}_1)\cdots(\mathbf{x}^T\mathbf{c}_m)\sum_{1\le q_1\le\cdots\le q_l\le 2}(\mathbf{x}^T\mathbf{c}_{q_1})\cdots(\mathbf{x}^T\mathbf{c}_{q_l})f_{1,\dots,m,q_1,\dots,q_l}. \tag{B.4}$$

Since $k_{\mathbf{C}} = K$, at most one of the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ can be orthogonal to any of the vectors $\mathbf{c}_3,\dots,\mathbf{c}_m$. Hence,

$$(\mathbf{x}^T\mathbf{c}_1)\cdots(\mathbf{x}^T\mathbf{c}_m) = t_1t_2(t_1\mathbf{u}_1^T\mathbf{c}_3 + t_2\mathbf{u}_2^T\mathbf{c}_3)\cdots(t_1\mathbf{u}_1^T\mathbf{c}_m + t_2\mathbf{u}_2^T\mathbf{c}_m) \neq 0$$

for generic $t_1, t_2 \in \mathbb{R}$. Hence, by (B.4),

$$\sum_{1\le q_1\le\cdots\le q_l\le 2}(\mathbf{x}^T\mathbf{c}_{q_1})\cdots(\mathbf{x}^T\mathbf{c}_{q_l})f_{1,\dots,m,q_1,\dots,q_l} = 0 \tag{B.5}$$

for generic $t_1, t_2 \in \mathbb{R}$. By construction of $\mathbf{x}$, the $l+1$ products $(\mathbf{x}^T\mathbf{c}_{q_1})\cdots(\mathbf{x}^T\mathbf{c}_{q_l})$, $1 \le q_1 \le \cdots \le q_l \le 2$, coincide with the monomials $t_1^l t_2^0, t_1^{l-1}t_2^1, \dots, t_1^0 t_2^l$. Thus, identity (B.5) expresses the fact that a polynomial in t_1 and t_2 with coefficients $f_{1,\dots,m,q_1,\dots,q_l}$ vanishes for generic $t_1, t_2 \in \mathbb{R}$. It is well known that this is possible only if the polynomial is identically zero, yielding that $f_{1,\dots,m,q_1,\dots,q_l} = 0$ for $1 \le q_1 \le \cdots \le q_l \le 2$.

(ii) Inductive step. We show that

$$f_{i_1,\dots,i_{m+l-p},i_{q_1},\dots,i_{q_p}} = 0 \ \text{ for } \quad 1 \le i_1 < \cdots < i_{m+l-p} \le R, \quad 1 \le q_1 \le \cdots \le q_p \le 2 + l - p$$

or, equivalently, that $f_{i_1,\dots,i_{m+l-p},i_{q_1},\dots,i_{q_p}} = 0$ for

$$(i_1,\dots,i_{m+l-p},i_{q_1},\dots,i_{q_p}) \in \bigcup_{1\le i_1<\cdots<i_{m+l-p}\le R} E_{i_1,\dots,i_{m+l-p}}.$$

We give the proof for the case $i_1 = 1, \dots, i_{m+l-p} = m+l-p$, the other cases follow similarly. Thus, we show that $f_{1,\dots,m+l-p,q_1,\dots,q_p} = 0$ for $1 \le q_1 \le \cdots \le q_p \le 2 + l - p$. The derivation is very similar to that of the induction hypothesis.

Since $k_{\mathbf{C}} = K$, the $K \times K$ matrix $\widetilde{\mathbf{C}} := [\mathbf{c}_1 \ \dots \mathbf{c}_{2+l-p} \ \mathbf{c}_{m+l-p+1} \dots \ \mathbf{c}_R]$ is nonsingular. Let $\mathbf{u}_1, \dots, \mathbf{u}_{2+l-p}$ denote the first $2 + l - p$ columns of $\widetilde{\mathbf{C}}^{-T}$. Then

$$\begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_{2+l-p}^T \end{bmatrix} \widetilde{\mathbf{C}} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{bmatrix}. \tag{B.6}$$

Let $\mathbf{x} = t_1\mathbf{u}_1 + \dots + t_{2+l-p}\mathbf{u}_{2+l-p}$. Let

$$E_{\tilde{p}<p} := \bigcup_{\tilde{p}<p} \ \bigcup_{1 \le i_1 < \dots < i_{m+l-\tilde{p}} \le R} E_{i_1,\dots,i_{m+l-\tilde{p}}}$$

and let the sets $E_{\tilde{p}>p}$ and $E_{\tilde{p}=p}$ be defined similarly. Then $E = E_{\tilde{p}<p} \cup E_{\tilde{p}>p} \cup E_{\tilde{p}=p}$. Then, by (24), (B.3), and (B.6), the vector $\mathbf{x}^{(m+l)} := \mathbf{x} \otimes \dots \otimes \mathbf{x}$ is orthogonal to the columns of the matrix $\mathbf{S}_{m+l}^E(\mathbf{C})$ indexed by the $(m + l)$-tuples

$$(r_1, \dots, r_{m+l}) \in E_{\tilde{p}<p} \cup (E_{\tilde{p}=p} \setminus E_{1,\dots,m+l-p}).$$

Hence, similarly to (B.4) we obtain

$$0 = \mathbf{x}^{(m+l)T}\mathbf{S}_{m+l}^E(\mathbf{C})\mathbf{f} = \sum_{(r_1,\dots,r_{m+l}) \in E} (\mathbf{x}^T\mathbf{c}_{r_1}) \cdots (\mathbf{x}^T\mathbf{c}_{r_{m+l}}) f_{r_1,\dots,r_{m+l}} = \sum_{(r_1,\dots,r_{m+l}) \in E_{\tilde{p}>p} \cup E_{1,\dots,m+l-p}} (\mathbf{x}^T\mathbf{c}_{r_1}) \cdots (\mathbf{x}^T\mathbf{c}_{r_{m+l}}) f_{r_1,\dots,r_{m+l}}.$$

Since, by the induction assumption, $f_{r_1,\dots,r_{m+l}} = 0$ for $(r_1, \dots, r_{m+l}) \in E_{\tilde{p}>p}$, we have

$$\begin{aligned} 0 = &\sum_{(r_1,\dots,r_{m+l}) \in E_{1,\dots,m+l-p}} (\mathbf{x}^T\mathbf{c}_{r_1}) \cdots (\mathbf{x}^T\mathbf{c}_{r_{m+l}}) f_{r_1,\dots,r_{m+l}} = \\ &\sum_{1 \le q_1 \le \dots \le q_p \le 2+l-p} (\mathbf{x}^T\mathbf{c}_1) \cdots (\mathbf{x}^T\mathbf{c}_{m+l-p})(\mathbf{x}^T\mathbf{c}_{q_1}) \cdots (\mathbf{x}^T\mathbf{c}_{q_p}) f_{1,\dots,m+l-p,q_1,\dots,q_p} = \\ &(\mathbf{x}^T\mathbf{c}_1) \cdots (\mathbf{x}^T\mathbf{c}_{m+l-p}) \sum_{1 \le q_1 \le \dots \le q_p \le 2+l-p} (\mathbf{x}^T\mathbf{c}_{q_1}) \cdots (\mathbf{x}^T\mathbf{c}_{q_p}) f_{1,\dots,m+l-p,q_1,\dots,q_p}. \end{aligned} \tag{B.7}$$

Since $k_{\mathbf{C}} = K$, at most $1+l-p$ of the vectors $\mathbf{u}_1, \dots, \mathbf{u}_{2+l-p}$ can be orthogonal to any of the vectors $\mathbf{c}_{3+l-p}, \ldots, \mathbf{c}_{m+l-p}$. Hence,

$$\begin{aligned}&(\mathbf{x}^T\mathbf{c}_1)\cdots(\mathbf{x}^T\mathbf{c}_{m+l-p}) =\\&t_1\cdots t_{2+l-p}(t_1\mathbf{u}_1^T\mathbf{c}_{3+l-p}+\cdots+t_{2+l-p}\mathbf{u}_{2+l-p}^T\mathbf{c}_{3+l-p})\cdots\\&(t_1\mathbf{u}_1^T\mathbf{c}_{m+l-p}+\cdots+t_{2+l-p}\mathbf{u}_{2+l-p}^T\mathbf{c}_{m+l-p})\neq 0\end{aligned} \tag{B.8}$$

for generic $t_1, \ldots, t_{2+l-p} \in \mathbb{R}$. Hence, by (B.7),

$$\sum_{1\le q_1\le\cdots\le q_p\le 2+l-p}(\mathbf{x}^T\mathbf{c}_{q_1})\cdots(\mathbf{x}^T\mathbf{c}_{q_p})f_{1,\ldots,m+l-p,q_1,\ldots,q_p}=0 \tag{B.9}$$

for generic $t_1, \ldots, t_{2+l-p} \in \mathbb{R}$. By construction of $\mathbf{x}$, the C_{l+1}^p products $(\mathbf{x}^T\mathbf{c}_{q_1})\cdots(\mathbf{x}^T\mathbf{c}_{q_p})$, $1 \le q_1 \le \cdots \le q_p \le 2+l-p$, coincide with the monomials $\{t_1^{\alpha_1}\cdots t_{2+l-p}^{\alpha_{2+l-p}}\}_{\alpha_1+\cdots+\alpha_{2+l-p}=p}$. Thus, identity (B.9) expresses the fact that a polynomial in $t_1, \ldots, t_{2+l-p}$ with coefficients $f_{1,\ldots,m+l-p,q_1,\ldots,q_p}$ vanishes for generic $t_1, \ldots, t_{2+l-p} \in \mathbb{R}$. It is well known that this is possible only if the polynomial is identically zero, yielding that $f_{1,\ldots,m+l-p,q_1,\ldots,q_p} = 0$ for $1 \le q_1 \le \cdots \le q_p \le 2+l-p$. □

References

[1] C. Bocci, L. Chiantini, and G. Ottaviani. Refined methods for the identifiability of tensors. *Ann. Mat. Pur. Appl.*, pages 1–12, 2013.

[2] J. Carroll and J.-J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35:283–319, 1970.

[3] L. Chiantini and G. Ottaviani. On generic identifiability of 3-tensors of small rank. *SIAM J. Matrix Anal. Appl.*, 33(3):1018–1037, 2012.

[4] A. Cichocki, D. Mandic, C. Caiafa, A-H. Phan, 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.*, 32:145–163, March 2015.

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

[6] L. De Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM J. Matrix Anal. Appl.*, 28:642–666, 2006.

[7] 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.*, 34:855–875, 2013.

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

[9] I. Domanov and L. De Lathauwer. Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition. *SIAM J. Matrix Anal. Appl.*, 35(2):636–660, 2014.

[10] I. Domanov and L. De Lathauwer. Generic uniqueness conditions for the canonical polyadic decomposition and INDSCAL. *SIAM J. Matrix Anal. Appl.*, 36(4):1567–1589, 2015.

[11] X. Guo, S. Miron, D. Brie, and A. Stegeman. Uni-Mode and Partial Uniqueness Conditions for CANDECOMP/PARAFAC of Three-Way Arrays with Linearly Dependent Loadings. *SIAM J. Matrix Anal. Appl.*, 33:111–129, 2012.

[12] R. A. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis. *UCLA Working Papers in Phonetics*, 16:1–84, 1970.

[13] R. A. Harshman. Determination and Proof of Minimum Uniqueness Conditions for PARAFAC1. *UCLA Working Papers in Phonetics*, 22:111–117, 1972.

[14] R. A. Harshman and M. E. Lundy. The PARAFAC model for three-way factor analysis and multidimensional scaling. *Research methods for multimode data analysis*, pages 122–215, 1984.

[15] R. A. Harshman and M. E. Lundy. Parafac: Parallel factor analysis. *Comput. Stat. Data Anal.*, pages 39–72, 1994.

[16] F. L. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.*, 6:164–189, 1927.

[17] 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(9):2625–2636, September 2004.

[18] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51:455–500, 2009.

[19] J. B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.*, 18:95–138, 1977.

[20] S. E. Leurgans, R. T. Ross, and R. B. Abel. A decomposition for three-way arrays. *SIAM J. Matrix Anal. Appl.*, 14(4):1064–1083, October 1993.

[21] J. Möcks. Topographic components model for event-related potentials and some biophysical considerations. *IEEE Trans. Biomed. Eng.*, 35:482–484, 1988.

[22] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos. Tensor decomposition for signal processing and machine learning. *Tech. Report 16-34, ESAT-STADIUS, KU Leuven (Leuven, Belgium), 2016.*

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

[24] L. Zhang, T.-Z. Huang, Q.-F. Zhu, and L. Feng. Uni-mode uniqueness conditions for CANDECOMP/PARAFAC decomposition of n-way arrays with linearly dependent loadings. *Linear Algebra Appl.*, 439(7):1918–1928, 2013.