KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Domanov I., De Lathauwer L., ``Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition'', *SIAM Journal on Matrix Analysis and Applications (SIMAX)*, vol. 35, no. 2, Apr.-May 2014, pp. 636--660
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	insert link to the published version of your paper http://dx.doi.org/10.1137/130916084
Journal homepage	insert link to the journal homepage of your paper . https://www.siam.org/journals/simax.php
Author contact	ignat.domanov@kuleuven.be Klik hier als u tekst wilt invoeren.
IR	url in Lirias https://lirias.kuleuven.be/handle/123456789/393094

(article begins on next page)

CANONICAL POLYADIC DECOMPOSITION OF THIRD-ORDER TENSORS: REDUCTION TO GENERALIZED EIGENVALUE DECOMPOSITION *

IGNAT DOMANOV[†‡§] AND LIEVEN DE LATHAUWER[†‡§]

Abstract. Canonical Polyadic Decomposition (CPD) of a third-order tensor is decomposition in a minimal number of rank-1 tensors. We call an algorithm algebraic if it is guaranteed to find the decomposition when it is exact and if it only relies on standard linear algebra (essentially sets of linear equations and matrix factorizations). The known algebraic algorithms for the computation of the CPD are limited to cases where at least one of the factor matrices has full column rank. In the paper we present an algebraic algorithm for the computation of the CPD in cases where none of the factor matrices has full column rank. In particular, we show that if the famous Kruskal condition holds, then the CPD can be found algebraically.

Key words. Canonical Polyadic Decomposition, Candecomp/Parafac Decomposition, tensor, Khatri-Rao product, compound matrix, permanent, mixed discriminant

AMS subject classifications. 15A69, 15A23

1. Introduction.

1.1. Basic notations and terminology. Throughout the paper $\mathbb{R}$ denotes the field of real numbers and $\mathcal{T} = (t_{ijk}) \in \mathbb{R}^{I\times J\times K}$ denotes a third-order tensor with frontal slices $\mathbf{T}_1, \dots, \mathbf{T}_K \in \mathbb{R}^{I\times J}$; $r_{\mathbf{A}}$, 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 $\mathbf{A}$) is the largest number such that every subset of $k_{\mathbf{A}}$ columns of the matrix $\mathbf{A}$ is linearly independent; $\omega(\mathbf{d})$ denotes the number of nonzero entries of a vector $\mathbf{d}$; span$\{\mathbf{f}_1, \dots, \mathbf{f}_k\}$ denotes the linear span of the vectors $\mathbf{f}_1, \dots, \mathbf{f}_k$; $\mathbf{O}_{m\times n}$, $\mathbf{0}_m$, and $\mathbf{I}_n$ are the zero $m\times n$ matrix, the zero $m\times 1$ vector, and the $n\times n$ identity matrix, respectively; C_n^k denotes the binomial coefficient, $C_n^k = \frac{n!}{k!(n-k)!}$; $\mathcal{C}_m(\mathbf{A})$ (the m-th compound matrix of $\mathbf{A}$) is the matrix containing the determinants of all $m\times m$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order (see §2 for details).

The outer product $\mathbf{a}\circ\mathbf{b}\circ\mathbf{c} \in \mathbb{R}^{I\times J\times K}$ of three nonzero vectors $\mathbf{a} \in \mathbb{R}^I$, $\mathbf{b} \in \mathbb{R}^J$ and $\mathbf{c} \in \mathbb{R}^K$ is called *rank*-1 tensor ($(\mathbf{a}\circ\mathbf{b}\circ\mathbf{c})_{ijk} := a_ib_jc_k$ for all values of the indices).

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

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{1.1}$$

where $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$, $\mathbf{c}_r \in \mathbb{R}^K$, $1 \le r \le R$. If the number R of rank-1 terms in (1.1) is minimal, then (1.1) is called the *Canonical Polyadic Decomposition* (CPD) of $\mathcal{T}$ and R is called the rank of the tensor $\mathcal{T}$ (denoted by $r_{\mathcal{T}}$).

*Research supported by: (1) Research Council KU Leuven: GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT 1/08/23, (2) F.W.O.: projects G.0427.10N, G.0830.14N, G.0881.14N, (3) the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, "Dynamical systems, control and optimization", 2012–2017), (4) DBOF/10/015.

[†]Group Science, Engineering and Technology, KU Leuven - Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium (ignat.domanov, lieven.delathauwer@kuleuven-kulak.be).

[‡]Dept. of Electrical Engineering ESAT/STADIUS KU Leuven, Kasteelpark Arenberg 10, bus 2446, B-3001 Leuven-Heverlee, Belgium (lieven.delathauwer@esat.kuleuven.be).

[§]iMinds Future Health Department.

We write (1.1) as $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, where the matrices $\mathbf{A} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_R \end{bmatrix} \in \mathbb{R}^{I\times R}$, $\mathbf{B} := \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_R \end{bmatrix} \in \mathbb{R}^{J\times R}$ and $\mathbf{C} := \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_R \end{bmatrix} \in \mathbb{R}^{K\times R}$ are called the *first*, *second* and *third factor matrix* of $\mathcal{T}$, respectively.

Obviously, $\mathbf{a}\circ\mathbf{b}\circ\mathbf{c}$ has frontal slices $\mathbf{a}\mathbf{b}^T c_1, \dots, \mathbf{a}\mathbf{b}^T c_K \in \mathbb{R}^{I\times J}$. Hence, (1.1) is equivalent to the system of matrix identities

$$\mathbf{T}_k = \sum_{r=1}^{R} \mathbf{a}_r \mathbf{b}_r^T c_{kr} = \mathbf{A}\mathrm{Diag}(\mathbf{c}^k)\mathbf{B}^T, \qquad 1 \le k \le K, \tag{1.2}$$

where $\mathbf{c}^k$ denotes the k-th column of the matrix $\mathbf{C}^T$ and $\mathrm{Diag}(\mathbf{c}^k)$ denotes a square diagonal matrix with the elements of the vector $\mathbf{c}^k$ on the main diagonal.

For a matrix $\mathbf{T} = [\mathbf{t}_1 \ \cdots \ \mathbf{t}_J]$, we follow the convention that $\mathrm{vec}(\mathbf{T})$ denotes the column vector obtained by stacking the columns of $\mathbf{T}$ on top of one another, i.e., $\mathrm{vec}(\mathbf{T}) = \begin{bmatrix} \mathbf{t}_1^T & \dots & \mathbf{t}_J^T \end{bmatrix}^T$. The matrix $\mathrm{Matr}(\mathcal{T}) := \begin{bmatrix} \mathrm{vec}(\mathbf{T}_1^T) & \dots & \mathrm{vec}(\mathbf{T}_K^T) \end{bmatrix} \in \mathbb{R}^{IJ\times K}$ is called *the matricization or matrix unfolding* of $\mathcal{T}$. The inverse operation is called *tensorization*: if $\mathbf{X}$ is an $IJ \times K$ matrix, then $\mathrm{Tens}(\mathbf{X}, I, J)$ is the $I \times J \times K$ tensor such that $\mathrm{Matr}(\mathcal{T}) = \mathbf{X}$. From the well-known formula

$$\mathrm{vec}(\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T) = (\mathbf{B}\odot\mathbf{A})\mathbf{d}, \qquad \mathbf{d}\in\mathbb{R}^R \tag{1.3}$$

it follows that

$$\mathrm{Matr}(\mathcal{T}) := \begin{bmatrix} (\mathbf{A}\odot\mathbf{B})\mathbf{c}^1 & \dots & (\mathbf{A}\odot\mathbf{B})\mathbf{c}^K \end{bmatrix} = (\mathbf{A}\odot\mathbf{B})\mathbf{C}^T, \tag{1.4}$$

where "$\odot$" denotes the *Khatri-Rao product* of matrices:

$$\mathbf{A}\odot\mathbf{B} := [\mathbf{a}_1\otimes\mathbf{b}_1 \ \cdots \ \mathbf{a}_R\otimes\mathbf{b}_R] \in \mathbb{R}^{IJ\times R}$$

and "$\otimes$" denotes the *Kronecker product*: $\mathbf{a}\otimes\mathbf{b} = [a_1b_1 \dots a_1b_J \ \dots \ a_Ib_1 \dots a_Ib_J]^T$.

It is clear that in (1.1) 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.

1.2. Problem statement. The CPD was introduced by F. Hitchcock in [14] and was later referred to as *Canonical Decomposition* (Candecomp) [3], *Parallel Factor Model* (Parafac) [11, 13], and *Topographic Components Model* [26]. We refer to the overview papers [4,5,7,17], the books [18,34] and the references therein for background and applications in Signal Processing, Data Analysis, Chemometrics, and Psychometrics.

Note that in applications one most often deals with a perturbed version of (1.1):

$$\widehat{\mathcal{T}} = \mathcal{T} + \mathcal{N} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R + \mathcal{N},$$

where $\mathcal{N}$ is an unknown noise tensor and $\widehat{\mathcal{T}}$ is the given tensor. The factor matrices of $\mathcal{T}$ are approximated by a solution of the optimization problem

$$\min \|\widehat{\mathcal{T}} - [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R\|, \qquad \text{s.t.} \quad \mathbf{A}\in\mathbb{R}^{I\times R},\ \mathbf{B}\in\mathbb{R}^{J\times R},\ \mathbf{C}\in\mathbb{R}^{K\times R}, \tag{1.5}$$

where $\|\cdot\|$ denotes a suitable (usually Frobenius) norm [36].

In this paper we limit ourselves to the noiseless case. We show that under mild conditions on factor matrices the CPD is unique and can be found algebraically in the

following sense: the CPD can be computed by using basic operations on matrices, by computing compound matrices, by taking the orthogonal complement of a subspace, and by computing generalized eigenvalue decomposition. We make connections with concepts like permanents, mixed discriminants, and compound matrices, which have so far received little attention in applied linear algebra but are of interest. Our presentation is in terms of real-valued tensors for notational convenience. Complex variants are easily obtained by taking into account complex conjugations.

The heart of the algebraic approach is the following straightforward connection between CPD of a two-slice tensor and Generalized Eigenvalue Decomposition (GEVD) of a matrix pencil. Consider an $R \times R \times 2$ tensor $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, where $\mathbf{A}$ and $\mathbf{B}$ are nonsingular matrices and the matrix $\mathrm{Diag}(\mathbf{d}) := \mathrm{Diag}(\mathbf{c}^1)\mathrm{Diag}(\mathbf{c}^2)^{-1}$ is defined and has distinct diagonal entries. From the equations $\mathbf{T}_k = \mathbf{A}\mathrm{Diag}(\mathbf{c}^k)\mathbf{B}^T$, $k = 1, 2$ it follows easily that $\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{A}^{-1} = \mathbf{T}_1\mathbf{T}_2^{-1}$ and $\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{B}^{-1} = (\mathbf{T}_2^{-1}\mathbf{T}_1)^T$. Hence, the matrix $\mathrm{Diag}(\mathbf{d})$ can be found (up to permutation of its diagonal entries) from the eigenvalue decomposition of $\mathbf{T}_1\mathbf{T}_2^{-1}$ or $(\mathbf{T}_2^{-1}\mathbf{T}_1)^T$ and the columns of $\mathbf{A}$ (resp. $\mathbf{B}$) are the eigenvectors of $\mathbf{T}_1\mathbf{T}_2^{-1}$ (resp. $(\mathbf{T}_2^{-1}\mathbf{T}_1)^T$) corresponding to the R distinct eigenvalues $d_1, \ldots, d_R$. Since the matrices $\mathbf{A}$ and $\mathbf{B}$ are nonsingular, the matrix $\mathbf{C}$ can be easily found from (1.4). More generally, when $\mathbf{A}$ and $\mathbf{B}$ have full column rank and $\mathbf{C}$ does not have collinear columns, $\mathbf{A}$ and $\mathbf{B}$ follow from the GEVD of the matrix pencil $(\mathbf{T}_1, \mathbf{T}_2)$.

1.3. Previous results on uniqueness and algebraic algorithms. We say that an $I \times R$ matrix *has full column rank* if its column rank is R, which implies $I \geq R$. The following theorem generalizes the result discussed at the end of the previous subsection. Several variants of this theorem have appeared in the literature [7, 12, 21, 31, 32, 40]. The proof is essentially obtained by picking two slices (or two mixtures of slices) from $\mathcal{T}$ and computing their GEVD.

THEOREM 1.1. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and suppose that $\mathbf{A}$ and $\mathbf{B}$ have full column rank and that $k_{\mathbf{C}} \geq 2$. Then*

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

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

In Theorem 1.1 the third factor matrix plays a different role than the first and the second factor matrices. Obviously, the theorem still holds when $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are permuted. In the sequel we will present only one version of results. Taking this into account, we may say that the following result is stronger than Theorem 1.1.

THEOREM 1.2. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = R$, and suppose that $\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ has full column rank. Then*

(i) *$r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique [6, 16];*

(ii) *the CPD of $\mathcal{T}$ can be found algebraically [6].*

Computationally, we may obtain from $\mathcal{T}$ a partially symmetric 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 and work as in Theorem 1.1 to obtain $\mathbf{C}^{-T}$. The matrices $\mathbf{A}$ and $\mathbf{B}$ are subsequently easily obtained from (1.4).

Also, some algorithms for symmetric CPD have been obtained in the context of algebraic geometry. We refer to [20, 30] and references therein. Further, algebraic algorithms have been obtained for CPDs in which factor matrices are subject to constraints (such as orthogonality and Vandermonde) [37, 39].

Our discussion concerns unsymmetric CPD without constraints. Results for the partially and fully symmetric case may be obtained by setting two or all three factor matrices equal to each other, respectively.

In the remaining part of this subsection we present some results on the uniqueness of the CPD. These results will guarantee CPD uniqueness under the conditions for which we will derive algebraic algorithms. For more general results on uniqueness we refer to [8, 9]. The following result was obtained by J. Kruskal, which is little known. We present the compact version from [9]. Corollary 1.4 presents what is widely known as "Kruskal's condition" for CPD uniqueness. The known proofs of Corollary 1.4 are nonconstructive. It is one of the contributions of this paper that Kruskal type conditions not only imply the uniqueness of the CPD but guarantee that the CPD can be found algebraically (see Corollaries 1.8–1.9).

THEOREM 1.3. *[19, Theorem 4b, p. 123], [9, Corollary 1.29] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Suppose that*

$$k_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R + 2 \quad \textit{and} \quad \min(r_{\mathbf{C}} + k_{\mathbf{B}}, k_{\mathbf{C}} + r_{\mathbf{B}}) \geq R + 2. \tag{1.6}$$

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

COROLLARY 1.4. *[19, Theorem 4a, p. 123] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2. \tag{1.7}$$

Then $r_{\mathcal{T}} = R$ *and the CPD of* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *is unique.*

In [8,9] the authors obtained new sufficient conditions expressed in terms of compound matrices. We will use the following result.

THEOREM 1.5. *[9, Corollary 1.25] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m := R - r_{\mathbf{C}} + 2$. *Suppose that*

$$\max(\min(k_{\mathbf{A}}, k_{\mathbf{B}} - 1),\ \min(k_{\mathbf{A}} - 1, k_{\mathbf{B}})) + k_{\mathbf{C}} \geq R + 1, \tag{1.8}$$

$$\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}) \textit{ has full column rank.} \tag{1.9}$$

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

Since the k-rank of a matrix cannot exceed its rank (and a fortiori not its number of columns), condition (1.7) immediately implies conditions (1.6) and (1.8). It was shown in [9] that (1.6) implies (1.9) for $m = R - r_{\mathbf{C}} + 2$. Thus, Theorem 1.5 guarantees the uniqueness of the CPD under milder conditions than Theorem 1.3. Note also that statement (i) of Theorem 1.2 is the special case of Theorem 1.5 obtained for $r_{\mathbf{C}} = R$, i.e., when one of the factor matrices has full column rank.

1.4. New results. To simplify the presentation and without loss of generality we will assume throughout the paper that the third dimension of the tensor $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ coincides with $r_{\mathbf{C}}$, that is, $K = r_{\mathbf{C}}$. (This can always be achieved in a "dimensionality reduction" step: if the columns of a matrix $\mathbf{V}$ form an orthonormal basis of the row space of $\mathrm{Matr}(\mathcal{T})$ and the matrix $\mathbf{A} \odot \mathbf{B}$ has full column rank (as is always the case in the paper), then $r_{\mathbf{C}} = r_{\mathrm{Matr}(\mathcal{T})} = r_{\mathbf{V}^T\mathrm{Matr}(\mathcal{T})} = r_{\mathbf{V}^T\mathbf{C}}$, and by (1.4), the matrix $\mathrm{Matr}(\mathcal{T})\mathbf{V} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{V}$ has $r_{\mathbf{C}}$ columns, which means that the third dimension of the tensor $\mathcal{T}_{\mathbf{V}} := \mathrm{Tens}(\mathrm{Matr}(\mathcal{T})\mathbf{V}, I, J)$ is equal to $r_{\mathbf{C}}$; if the CPD $\mathcal{T}_{\mathbf{V}} = [\mathbf{A}, \mathbf{B}, \mathbf{V}^T\mathbf{C}]_R$ has been computed, then the matrix $\mathbf{C}$ can be recovered as $\mathbf{C} = \mathbf{V}(\mathbf{V}^T\mathbf{C})$).

The following theorems are the main results of the paper. In all cases we will reduce the computation to the situation as in Theorem 1.1.

THEOREM 1.6. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $m := R - r_{\mathbf{C}} + 2$. *Suppose that* $k_{\mathbf{C}} = r_{\mathbf{C}}$ *and that* (1.9) *holds. Then*

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

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

THEOREM 1.7. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $n := R - k_{\mathbf{C}} + 2$. *Suppose that*

$$\mathcal{C}_n(\mathbf{A}) \odot \mathcal{C}_n(\mathbf{B}) \text{ has full column rank.} \tag{1.10}$$

Then

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

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

Theorem 1.7 generalizes Theorem 1.6 to the case where possibly $k_{\mathbf{C}} < r_{\mathbf{C}}$. The more general situation for $\mathbf{C}$ is accommodated by tightening the condition on $\mathbf{A}$ and $\mathbf{B}$. (Indeed, (1.10) is more restrictive than (1.9) when $n > m$.) The proof of Theorem 1.7 is simple; we essentially consider a $k_{\mathbf{C}}$-slice subtensor $\bar{\mathcal{T}} = [\mathbf{A}, \mathbf{B}, \bar{\mathbf{C}}]_R$ for which $k_{\bar{\mathbf{C}}} = r_{\bar{\mathbf{C}}}$, so that Theorem 1.6 applies. (Actually, to guarantee that $k_{\bar{\mathbf{C}}} = r_{\bar{\mathbf{C}}}$, we consider a random slice-mixture.)

We also obtain the following corollaries.

COROLLARY 1.8. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Suppose that*

$$k_{\mathbf{A}} + r_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \quad \text{and} \quad k_{\mathbf{B}} + k_{\mathbf{C}} \geq R + 2. \tag{1.11}$$

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

COROLLARY 1.9. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let* $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2$. *Then the CPD of* $\mathcal{T}$ *is unique and can be found algebraically.*

Let us further explain how the theorems that we have formulated so far, relate to one another. First, we obviously have that $n = R - k_{\mathbf{C}} + 2 \geq R - r_{\mathbf{C}} + 2 = m$. Next, the following implications were proved in [8]:

$$\begin{array}{ccccc} (1.11) & \Longrightarrow & (1.10) & \Longrightarrow & \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq n \overset{\text{trivial}}{\Longrightarrow} (1.8) \\ \Downarrow \text{trivial} & & \Downarrow & & \Downarrow \quad \Uparrow \text{if } k_{\mathbf{C}} = r_{\mathbf{C}} \text{ (trivial)} \\ (1.6) & \Longrightarrow & (1.9) & \Longrightarrow & \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m \end{array} \tag{1.12}$$

The first thing that follows from scheme (1.12) is that Theorem 1.7 is indeed more general than Corollary 1.8. Corollary 1.9 follows trivially from Corollary 1.8. Next, it appears that the conditions of Theorems 1.6–1.7 are more restrictive than the conditions of Theorem 1.5. Also, the conditions of Corollary 1.8 are more restrictive than the conditions of Theorem 1.3. Hence, we immediately obtain the uniqueness of the CPD in Theorems 1.6–1.7 and Corollary 1.8. Consequently, we can limit ourselves to the derivation of the algebraic algorithms. Theorems 1.6 and 1.7 are proved in subsections 4.1 and 4.4, respectively.

1.5. Organization. We now explain how the paper is organized. Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \in \mathbb{R}^{I \times J \times K}$ with $k_{\mathbf{C}} = K$, implying $K \leq R$. In the first phase of our algorithms, we find up to column permutation and scaling the $K \times C_R^{K-1}$ matrix $\mathcal{B}(\mathbf{C})$ defined by

$$\mathcal{B}(\mathbf{C}) := \mathbf{L}\mathcal{C}_{K-1}(\mathbf{C}), \tag{1.13}$$

where

$$\mathbf{L} := \begin{bmatrix} 0 & 0 & \dots & (-1)^{K-1} \\ \vdots & \vdots & \iddots & \vdots \\ 0 & -1 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{bmatrix}. \tag{1.14}$$

The matrix $\mathcal{B}(\mathbf{C})$ can be considered as an unconventional variant of the inverse of $\mathbf{C}$:

$$\text{every column of } \mathcal{B}(\mathbf{C}) \text{ is orthogonal to exactly } K-1 \text{ columns of } \mathbf{C}, \tag{P1}$$

$$\text{any vector that is orthogonal to exactly } K-1 \text{ columns of } \mathbf{C} \text{ is proportional to a column of } \mathcal{B}(\mathbf{C}), \tag{P2}$$

$$\text{every column of } \mathbf{C} \text{ is orthogonal to exactly } C_{R-1}^{K-2} \text{ columns of } \mathcal{B}(\mathbf{C}), \tag{P3}$$

$$\text{any vector that is orthogonal to exactly } C_{R-1}^{K-2} \text{ columns of } \mathcal{B}(\mathbf{C}) \text{ is proportional to a column of } \mathbf{C}. \tag{P4}$$

Recall that every column of the classical Moore-Penrose pseudo-inverse $\mathbf{C}^\dagger \in \mathbb{R}^{R\times K}$ is orthogonal to exactly $K-1$ *rows* of $\mathbf{C}$ and vice-versa. The equality $\mathbf{C}\mathbf{C}^\dagger = \mathbf{I}_K$ works along the "long" dimension of $\mathbf{C}$. If $\mathbf{C}^\dagger$ is known, then $\mathbf{C}$ may easily be found by pseudo-inverting again, $\mathbf{C} = (\mathbf{C}^\dagger)^\dagger$. The interaction with $\mathcal{B}(\mathbf{C})$ takes place along the "short" dimension of $\mathbf{C}$, and this complicates things. Nevertheless, it is also possible to reconstruct $\mathbf{C}$ from $\mathcal{B}(\mathbf{C})$. In the second and third phase of our algorithms we use $\mathcal{B}(\mathbf{C})$ to compute CPD. The following two properties of $\mathcal{B}(\mathbf{C})$ will be crucial for our derivation.

PROPOSITION 1.10. *Let $\mathbf{C} \in \mathbb{R}^{K\times R}$ and $k_{\mathbf{C}} = K$. Then*

(i) *$\mathcal{B}(\mathbf{C})$ has no proportional columns, that is $k_{\mathcal{B}(\mathbf{C})} \geq 2$.*

(ii) *the matrices*

$$\mathcal{B}(\mathbf{C})^{(m-1)} = \underbrace{\mathcal{B}(\mathbf{C}) \odot \cdots \odot \mathcal{B}(\mathbf{C})}_{m-1}, \qquad \mathcal{B}(\mathbf{C})^{(m)} = \underbrace{\mathcal{B}(\mathbf{C}) \odot \cdots \odot \mathcal{B}(\mathbf{C})}_{m}$$

have full column rank for $m := R - K + 2$.

Sections 2–3 contain auxiliary results of which several are interesting in their own right. In Subsection 2.1 we recall the properties of compound matrices, provide an intuitive understanding of properties (P1)–(P4) and Propositions 1.10, and discuss the reconstruction of $\mathbf{C}$ from $\mathcal{B}(\mathbf{C})$. (Since the proofs of properties (P1)-(P4) and Proposition 1.10 are rather long and technical, they are included in the supplementary materials.) In Subsections 2.2–2.3 we study variants of permanental compound matrices. Let the columns of the K^m-by-C_R^m matrix $\mathcal{R}_m(\mathbf{C})$ be equal to the vectorized symmetric parts of the tensors $\mathbf{c}_{i_1} \circ \cdots \circ \mathbf{c}_{i_m}$, $1 \leq i_1 < \cdots < i_m \leq R$ and let $\operatorname{range}(\pi_S)$ denote a subspace of $\mathbb{R}^{K^m}$ that consists of vectorized versions of m-th order $K \times \cdots \times K$ symmetric tensors, yielding $\dim \operatorname{range}(\pi_S) = C_{K+m-1}^m$. We prove

$$\text{Proposition 2.13 (iii) : } \ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}\right) = \operatorname{range}(\mathcal{B}(\mathbf{C})^{(m)}), \tag{1.15}$$

where the notation $\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}$ means that we let the matrix $\mathcal{R}_m(\mathbf{C})^T$ act only on vectors from $\operatorname{range}(\pi_S)$, i.e., on $K^m \times 1$ vectorized versions of $K\times\cdots\times K$ symmetric tensors. Computationally, the subspace $\ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}\right)$ is the intersection of the subspaces $\ker(\mathcal{R}_m(\mathbf{C})^T)$ and $\operatorname{range}(\pi_S)$.

In §3 we introduce polarized compound matrices — a notion closely related to the rank detection mappings in [6, 29]. The entries of polarized compound matrices are mixed discriminants [1, 2, 22]. Using polarized compound matrices we construct a $C_I^m C_J^m \times K^m$ matrix $\mathbf{R}_m(\mathcal{T})$ from the given tensor $\mathcal{T}$ such that

$$\mathbf{R}_m(\mathcal{T}) = [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\, \mathcal{R}_m(\mathbf{C})^T. \tag{1.16}$$

Assuming that $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank and combining (1.15) with (1.16) we find the space generated by the columns of the matrix $\mathcal{B}(\mathbf{C})^{(m)}$:

$$\ker\left(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}\right) = \ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}\right) = \mathrm{range}(\mathcal{B}(\mathbf{C})^{(m)}). \tag{1.17}$$

In §4 we combine all results to obtain Theorems 1.6–1.7 and we present two algebraic CPD algorithms. Both new algorithms contain the same first phase in which we find a matrix $\mathbf{F}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. This first phase of the algorithms relies on key formula (1.17), which makes a link between the known matrix $\mathbf{R}_m(\mathcal{T})$, constructed from $\mathcal{T}$, and the unknown matrix $\mathcal{B}(\mathbf{C})$. We work as follows. We construct the matrix $\mathbf{R}_m(\mathcal{T})$ and compute the vectorized symmetric tensors in its kernel. We stack a basis of $\ker\left(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}\right)$ as columns of a matrix $\mathrm{Matr}(\mathcal{W}) \in \mathbb{R}^{K^m \times C_R^{K-1}}$, with which we associate a $K \times K^{m-1} \times C_R^{K-1}$ tensor $\mathcal{W}$. From Proposition 1.10 and Theorem 1.1 it follows that the CPD $\mathcal{W} = [\mathcal{B}(\mathbf{C}), \mathcal{B}(\mathbf{C})^{(m-1)}, \mathbf{M}]_{C_R^{K-1}}$ can be found algebraically. This allows us to find a matrix $\mathbf{F}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. In the second and third phase of the first algorithm we find the matrix $\mathbf{C}$ and the matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. For finding $\mathbf{C}$, we resort to properties (P3)–(P4). Full exploitation of the structure has combinatorial complexity and is infeasible unless the dimensions of the tensor are relatively small. As an alternative, in the second algorithm we first find the matrices $\mathbf{A}$ and $\mathbf{B}$ and then we find the matrix $\mathbf{C}$. This is done as follows. We construct the new $I \times J \times C_R^{K-1}$ tensor $\mathcal{V}$ with the matrix unfolding $\mathrm{Matr}(\mathcal{V}) := \mathrm{Matr}(\mathcal{T})\mathbf{F} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{F}$. We find subtensors of $\mathcal{V}$ such that each subtensor has dimensions $I \times J \times 2$ and its CPD can be found algebraically. Full exploitation of the structure yields $C_R^m C_m^2$ subtensors. From the CPD of the subtensors we simultaneously obtain the columns of $\mathbf{A}$ and $\mathbf{B}$, and finally we set $\mathbf{C} = \left((\mathbf{A} \odot \mathbf{B})^\dagger \mathrm{Matr}(\mathcal{T})\right)^T$.

We conclude the paper with two examples. In the first example we demonstrate how the algorithms work for a $4 \times 4 \times 4$ tensor of rank 5 for which $k_{\mathbf{A}} = k_{\mathbf{B}} = 3$. In the second example we consider a generic $6 \times 6 \times 7$ tensor of rank 9 and compare the complexity of algorithms. Note that in neither case the uniqueness of the CPDs follows from Kruskal's Theorem 1.3.

1.6. Link with [6]. Our overall derivation generalizes ideas from [6] ($K = R$). To conclude the introduction, we recall the CPD algorithm from [6] using our notations. We have $K = R$, which implies $m = 2$. First, we construct the $C_I^2 C_J^2 \times R^2$ matrix $\mathbf{R}_2(\mathcal{T})$ whose $((i-1)R + j)$-th column is computed as

$$\mathrm{Vec}\left(\, \mathcal{C}_2(\mathbf{T}_i + \mathbf{T}_j) - \mathcal{C}_2(\mathbf{T}_i) - \mathcal{C}_2(\mathbf{T}_j) \,\right), \quad 1 \le i \le j \le R,$$

where $\mathbf{T}_1, \ldots, \mathbf{T}_R \in \mathbb{R}^{I \times J}$ denote the frontal slices of $\mathcal{T}$. The entries of the $((i-1)R + j)$-th column of $\mathbf{R}_2(\mathcal{T})$ can be identified with the $C_I^2 C_J^2$ nonzero entries of the $I \times I \times J \times J$ tensor $\mathcal{P}_{ij}$ [6, p. 648]. Then we find a basis $\mathbf{w}_1, \ldots, \mathbf{w}_R \in \mathbb{R}^{R^2}$ of $E := \ker\left(\mathbf{R}_2(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}\right)$ and set $\mathbf{W} = [\mathbf{w}_1 \ \ldots \ \mathbf{w}_R]$. We note that E can be computed as the intersection of the subspaces $\ker(\mathbf{R}_2(\mathcal{T}))$ and $\mathrm{range}(\pi_S)$, where $\mathrm{range}(\pi_S)$ consists of vectorized versions of symmetric $R \times R$ matrices. In [6], the subspace E is generated by the vectors in $\mathrm{range}(\pi_S)$ that yield a zero linear combination of the R^2 tensors $\mathcal{P}_{ij}$. In the next step we recover (up to column permutation and scaling) $\mathbf{C}$ from E. This is done as follows. By (P3)–(P4), the columns of $\mathcal{B}(\mathbf{C})$ are proportional to the columns of $\mathbf{C}^{-T}$, i.e., $\mathcal{B}(\mathbf{C})^T$ is equal to the inverse of $\mathbf{C}$ up to column permutation

and scaling. Hence, by (1.17), range($\mathbf{W}$) = range($\mathbf{C}^{-T} \odot \mathbf{C}^{-T}$). Hence, there exists a nonsingular matrix $\mathbf{M}$ such that $\mathbf{W} = \left(\mathbf{C}^{-T} \odot \mathbf{C}^{-T}\right)\mathbf{M}^T$. Therefore, by (1.4), $\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T}, \mathbf{M}]_R$, where $\mathcal{W}$ denotes the $R \times R \times R$ tensor such that $\mathbf{W} =$ Matr($\mathcal{W}$). Since all factor matrices of $\mathcal{W}$ have full column rank, the CPD of $\mathcal{W}$ can be computed algebraically. Thus, we can find $\mathbf{C}^{-T}$ (and hence, $\mathbf{C}$) up to column permutation and scaling. Finally, the matrices $\mathbf{A}$ and $\mathbf{B}$ can now be easily found from Matr($\mathcal{T}$)$\mathbf{C}^{-T} = \mathbf{A} \odot \mathbf{B}$ using the fact that the columns of $\mathbf{A} \odot \mathbf{B}$ are vectorized rank-1 matrices.

2. Matrices formed by determinants and permanents of submatrices of a given matrix. Throughout the paper we will use the following multi-index notations. Let $i_1, \dots, i_k$ be integers. Then $\{i_1, \dots, i_k\}$ denotes the set with elements $i_1, \dots, i_k$ (the order does not matter) and $(i_1, \dots, i_k)$ denotes a k-tuple (the order is important). Let

$$\begin{aligned} S_n^k &= \{(i_1, \dots, i_k) : \ 1 \le i_1 < i_2 < \cdots < i_k \le n\}, \\ Q_n^k &= \{(i_1, \dots, i_k) : \ 1 \le i_1 \le i_2 \le \cdots \le i_k \le n\}, \\ R_n^k &= \{(i_1, \dots, i_k) : \ i_1, \dots, i_k \in \{1, \dots, n\}\}. \end{aligned}$$

It is well known that card $S_n^k = C_n^k$, card $Q_n^k = C_{n+k-1}^k$, and card $R_n^k = n^k$. We assume that the elements of S_n^k, Q_n^k, and R_n^k are ordered lexicographically. In the sequel we will both use indices taking values in $\{1, 2, \dots, C_n^k\}$ (resp. $\{1, 2, \dots, C_{n+k-1}^k\}$ or $\{1, 2, \dots, n^k\}$) and multi-indices taking values in S_n^k (resp. Q_n^k or R_n^k). For example,

$$\begin{gathered} S_2^2 = \{(1,2)\}, \quad Q_2^2 = \{(1,1),(1,2),(2,2)\}, \quad R_2^2 = \{(1,1),(1,2),(2,1),(2,2)\}, \\ S_2^2(1) = Q_2^2(2) = R_2^2(2), \quad Q_2^2(3) = R_2^2(4). \end{gathered}$$

Let also $P_{\{j_1,\dots,j_n\}}$ denote the set of all permutations of the set $\{j_1, \dots, j_n\}$. We follow the convention that if some of $j_1, \dots, j_n$ coincide, then the set $P_{\{j_1,\dots,j_n\}}$ contains identical elements, yielding card $P_{\{j_1,\dots,j_n\}} = n!$. For example, $P_{\{1,2,2\}} = \{\{1,2,2\},\{1,2,2\},\{2,1,2\},\{2,2,1\},\{2,1,2\},\{2,2,1\}\}$. We set $P_n := P_{\{1,\dots,n\}}$.

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Throughout the paper $\mathbf{A}((i_1, \dots, i_k), (j_1, \dots, j_k))$ denotes the submatrix of $\mathbf{A}$ at the intersection of the k rows with row numbers $i_1, \dots, i_k$ and the k columns with column numbers $j_1, \dots, j_k$.

2.1. Matrices whose entries are determinants. In this subsection we briefly discuss compound matrices. The k-th compound matrix of a given matrix is formed by $k \times k$ minors of that matrix. We have the following formal definition.

Definition 2.1. *[15] Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $k \le \min(m, n)$. The C_m^k-by-C_n^k matrix whose (i, j)-th entry is* $\det \mathbf{A}(S_m^k(i), S_n^k(j))$ *is called the k-th compound matrix of $\mathbf{A}$ and is denoted by $\mathcal{C}_k(\mathbf{A})$.*

EXAMPLE 2.2. *Let* $\mathbf{A} = [\mathbf{I}_3\ \mathbf{a}]$, *where* $\mathbf{a} = [a_1\ a_2\ a_3]^T$. *Then*

$$\mathcal{C}_2(\mathbf{A}) = \begin{array}{c} \\ (1,2) \\ (1,3) \\ (2,3) \end{array} \begin{array}{c} \begin{array}{cccccc} (1,2) & & (1,3) & & (1,4) & \end{array} \\ \begin{bmatrix} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & a_1 \\ 0 & a_2 \end{vmatrix} \\ \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & a_1 \\ 0 & a_3 \end{vmatrix} \\ \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 0 & a_2 \\ 0 & a_3 \end{vmatrix} \end{bmatrix} \end{array} \begin{array}{c} \begin{array}{cccccc} (2,3) & & (2,4) & & (3,4) & \end{array} \\ \begin{bmatrix} \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 0 & a_1 \\ 1 & a_2 \end{vmatrix} & \begin{vmatrix} 0 & a_1 \\ 0 & a_2 \end{vmatrix} \\ \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 0 & a_1 \\ 0 & a_3 \end{vmatrix} & \begin{vmatrix} 0 & a_1 \\ 1 & a_3 \end{vmatrix} \\ \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & a_2 \\ 0 & a_3 \end{vmatrix} & \begin{vmatrix} 0 & a_2 \\ 1 & a_3 \end{vmatrix} \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 1 & 0 & a_2 & 0 & -a_1 & 0 \\ 0 & 1 & a_3 & 0 & 0 & -a_1 \\ 0 & 0 & 0 & 1 & a_3 & -a_2 \end{bmatrix}.$$

Definition 2.1 immediately implies the following lemma.

LEMMA 2.3. *Let* $\mathbf{A} \in \mathbb{R}^{I \times R}$ *and* $k \leq \min(I, R)$. *Then*

(1) $\mathcal{C}_k(\mathbf{A})$ *has one or more zero columns if and only if* $k > k_{\mathbf{A}}$;

(2) $\mathcal{C}_k(\mathbf{A})$ *is equal to the zero matrix if and only if* $k > r_{\mathbf{A}}$;

(3) $\mathcal{C}_k(\mathbf{A}^T) = (\mathcal{C}_k(\mathbf{A}))^T$.

PD representation (1.2) will make us need compound matrices of diagonal matrices.

LEMMA 2.4. *Let* $\mathbf{d} \in \mathbb{R}^R$, *let* $\omega(\mathbf{d})$ *denote the number of nonzero entries of* $\mathbf{d}$, $k \leq R$, *and let* $\widehat{\mathbf{d}}^k := [d_1 \cdots d_k \ \ d_1 \cdots d_{k-1} d_{k+1} \ \ \ldots \ \ d_{R-k+1} \cdots d_R]^T \in \mathbb{R}^{C_R^k}$. *Then*

(1) $\widehat{\mathbf{d}}^k = \mathbf{0}$ *if and only if* $\omega(\mathbf{d}) \leq k-1$;

(2) $\widehat{\mathbf{d}}^k$ *has exactly one nonzero entry if and only if* $\omega(\mathbf{d}) = k$;

(3) $\mathcal{C}_k(\mathrm{Diag}(\mathbf{d})) = \mathrm{Diag}(\widehat{\mathbf{d}}^k)$.

The following result is known as Binet-Cauchy formula.

LEMMA 2.5. *[15, p. 19–22] Let* k *be a positive integer and let* $\mathbf{A}$ *and* $\mathbf{B}$ *be matrices such that* $\mathcal{C}_k(\mathbf{A})$ *and* $\mathcal{C}_k(\mathbf{B})$, *are defined. Then* $\mathcal{C}_k(\mathbf{A}\mathbf{B}^T) = \mathcal{C}_k(\mathbf{A})\mathcal{C}_k(\mathbf{B}^T)$. *If additionally* $\mathbf{d}$ *is a vector such that* $\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T$ *is defined, then* $\mathcal{C}_k(\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T) = \mathcal{C}_k(\mathbf{A})\mathrm{Diag}(\widehat{\mathbf{d}}^k)\mathcal{C}_k(\mathbf{B})^T$.

The goal of the remaining part of this subsection is to provide an intuitive understanding of properties (P1)–(P4) and Proposition 1.10.

Let $K \geq 2$, and let $\mathbf{C}$ be a $K \times K$ nonsingular matrix. By Cramer's rule and (1.13), the matrices $\det(\mathbf{C})\mathbf{C}^{-1}$ and $\mathcal{B}(\mathbf{C})$ are formed by $(K-1) \times (K-1)$ minors (also known as cofactors) of $\mathbf{C}$. It is easy to show that $\mathcal{B}(\mathbf{C}) = (\det(\mathbf{C})\mathbf{C}^{-1})^T\mathbf{L}$, where $\mathbf{L}$ is given by (1.14). It now trivially follows that every column of $\mathcal{B}(\mathbf{C})$ is a nonzero vector orthogonal to exactly $K-1$ columns of $\mathbf{C}$. Indeed,

$$\mathbf{C}^T\mathcal{B}(\mathbf{C}) = \mathbf{C}^T \det(\mathbf{C})\mathbf{C}^{-T}\mathbf{L} = \det(\mathbf{C})\mathbf{L},$$

which has precisely one non-zero entry in every column. The inverse statement holds also. Namely, if $\mathbf{x}$ is a nonzero vector that is orthogonal to exactly K $(= C_{K-1}^{K-2})$ columns of $\mathcal{B}(\mathbf{C})$ (i.e. $\omega(\mathbf{x}^T\mathcal{B}(\mathbf{C})) \leq 1$), then $\mathbf{x}$ is proportional to a column of $\mathbf{C}$. Indeed,

$$\omega(\mathbf{x}^T\mathcal{B}(\mathbf{C})) = \omega(\mathbf{x}^T \det(\mathbf{C})\mathbf{C}^{-T}\mathbf{L}) = \omega(\mathbf{x}^T\mathbf{C}^{-T}) = \omega(\mathbf{C}^{-1}\mathbf{x}) \leq 1 \Leftrightarrow \quad \mathbf{x} \text{ is proportional to a column of } \mathbf{C}. \tag{2.1}$$

Properties (P3)–(P4) generalize (2.1) for rectangular matrices and imply that, if we know $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling, then we know $\mathbf{C}$ up to column

permutation and scaling. This result will be directly used in Algorithm 1 further: we will first estimate $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling and then obtain $\mathbf{C}$ up to column permutation and scaling. Statements (P1)–(P3) are easy to show. Statement (P4) is more difficult. Since the proofs are technical, they are given in the supplementary materials.

Let us illustrate properties (P1)–(P4) and Proposition 1.10 for a rectangular matrix $\mathbf{C}$ ($K < R$).

EXAMPLE 2.6. *Let*

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

implying $k_{\mathbf{C}} = K = 3$ *and* $R = 4$. *From* (1.13) *and Example 2.2 it follows that*

$$\mathcal{B}(\mathbf{C}) = \mathbf{L}\mathcal{C}_2(\mathbf{C}) = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & -1 & -1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}.$$

One can easily check the statements of properties (P1)–(P4) *and Proposition 1.10. Note in particular that exactly* 4 *sets of* 3 *columns of* $\mathcal{B}(\mathbf{C})$ *are linearly dependent. The vectors that are orthogonal to these sets are proportional to the columns of* $\mathbf{C}$.

In our overall CPD algorithms we will find a matrix $\mathbf{F} \in \mathbb{R}^{K \times C_R^{K-1}}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. Properties (P3)–(P4) imply the following combinatorial procedure to find the third factor matrix of $\mathcal{T}$. Since the permutation indeterminacy makes that we do not know beforehand which columns of $\mathbf{F}$ are orthogonal to which columns of $\mathbf{C}$, we need to look for subsets of C_{R-1}^{K-2} columns of $\mathbf{F}$ that are linearly dependent. By properties (P3)–(P4), there exist exactly R such subsets. For each subset, the orthogonal complement yields, up to scaling, a column of $\mathbf{C}$.

2.2. Matrices whose entries are permanents.

DEFINITION 2.7. *Let* $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{n \times n}$. *Then the permanent of* $\mathbf{A}$ *is defined as*

$$\operatorname{perm} \mathbf{A} = \overset{+}{|}\mathbf{A}\overset{+}{|} = \sum_{(l_1,\dots,l_n)\in P_n} a_{1l_1}a_{2l_2}\cdots a_{nl_n} = \sum_{(l_1,\dots,l_n)\in P_n} a_{l_1 1}a_{l_2 2}\cdots a_{l_n n}.$$

The definition of the permanent of $\mathbf{A}$ differs from that of the determinant of $\mathbf{A}$ in that the signatures of the permutations are not taken into account. This makes the permanent invariant for column permutations of $\mathbf{A}$. The notations perm $\mathbf{A}$ and $\overset{+}{|}\mathbf{A}\overset{+}{|}$ are due to Minc [27] and Muir [28], respectively.

We have the following permanental variant of compound matrix.

DEFINITION 2.8. *[25] Let* $\mathbf{C} \in \mathbb{R}^{K \times R}$. *The* C_K^m*-by-*C_R^m *matrix whose* (i,j)*-th entry is* perm $\mathbf{C}(S_K^m(i), S_R^m(j))$ *is called the* m*-th permanental compound matrix of* $\mathbf{C}$ *and is denoted by* $\mathcal{PC}_m(\mathbf{C})$.

In our derivation we will also use the following two types of matrices. As far as we know, these do not have a special name.

DEFINITION 2.9. *Let* $\mathbf{C} \in \mathbb{R}^{K \times R}$. *The* C_{K+m-1}^m*-by-*C_R^m *matrix whose* (i,j)*-th entry is* perm $\mathbf{C}(Q_K^m(i), S_R^m(j))$ *is denoted by* $\mathcal{Q}_m(\mathbf{C})$.

DEFINITION 2.10. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$. *The* K^m*-by-*C_R^m *matrix whose* (i,j)*-th entry is* perm $\mathbf{C}(R_K^m(i), S_R^m(j))$ *is denoted by* $\mathcal{R}_m(\mathbf{C})$.

Note that $\mathcal{Q}_m(\mathbf{C})$ is a submatrix of $\mathcal{R}_m(\mathbf{C})$, in which the doubles of rows that are due to the permanental invariance for column permutations, have been removed.

The following lemma makes the connection between $\mathcal{Q}_m(\mathbf{C})^T$ and $\mathcal{R}_m(\mathbf{C})^T$ and permanental compound matrices.

LEMMA 2.11. *Let* $\mathbf{C} = [\mathbf{c}^1 \quad \dots \quad \mathbf{c}^K]^T \in \mathbb{R}^{K\times R}$. *Then* $\mathcal{Q}_m(\mathbf{C})^T$ *(resp.* $\mathcal{R}_m(\mathbf{C})^T$*) has columns* $\mathcal{PC}_m([\mathbf{c}^{j_1} \quad \dots \quad \mathbf{c}^{j_m}])$, *where* $(j_1,\dots,j_m) \in Q_K^m$ *(resp.* R_K^m*).*

EXAMPLE 2.12. *Let* $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. *Then*

$$\mathcal{R}_2(\mathbf{C}) = \begin{array}{c} \\ (1,1) \\ (1,2) \\ (2,1) \\ (2,2) \end{array} \begin{array}{c} \begin{array}{ccc} (1,2) & (1,3) & (2,3) \end{array} \\ \begin{bmatrix} \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix}^+ & \begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix}^+ & \begin{vmatrix} 2 & 3 \\ 2 & 3 \end{vmatrix}^+ \\ \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}^+ & \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}^+ & \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}^+ \\ \begin{vmatrix} 4 & 5 \\ 1 & 2 \end{vmatrix}^+ & \begin{vmatrix} 4 & 6 \\ 1 & 3 \end{vmatrix}^+ & \begin{vmatrix} 5 & 6 \\ 2 & 3 \end{vmatrix}^+ \\ \begin{vmatrix} 4 & 5 \\ 4 & 5 \end{vmatrix}^+ & \begin{vmatrix} 4 & 6 \\ 4 & 6 \end{vmatrix}^+ & \begin{vmatrix} 5 & 6 \\ 5 & 6 \end{vmatrix}^+ \end{bmatrix} \end{array} = \begin{bmatrix} 4 & 6 & 12 \\ 13 & 18 & 27 \\ 13 & 18 & 27 \\ 40 & 48 & 60 \end{bmatrix}.$$

The matrix $\mathcal{Q}_2(\mathbf{C})$ *is obtained from* $\mathcal{R}_2(\mathbf{C})$ *by deleting the row indexed with* $(2,1)$.

2.3. Links between matrix $\mathcal{R}_m(\mathbf{C})$, matrix $\mathcal{B}(\mathbf{C})$ and symmetrizer. Recall that the matrices $\pi_S(\mathbf{T}) := (\mathbf{T}+\mathbf{T}^T)/2$ and $(\mathbf{T}-\mathbf{T}^T)/2$ are called the symmetric part and skew-symmetric part of a square matrix $\mathbf{T}$, respectively. The equality $\mathbf{T} = (\mathbf{T}+\mathbf{T}^T)/2+(\mathbf{T}-\mathbf{T}^T)/2$ expresses the well-known fact that an arbitrary square matrix can be represented uniquely as a sum of a symmetric matrix and a skew-symmetric matrix. Similarly, with a general mth-order $K\times\cdots\times K$ tensor $\mathcal{T}$ one can uniquely associate its symmetric part $\pi_S(\mathcal{T})$ — a tensor whose entry with indices $j_1,\dots,j_m$ is equal to

$$\frac{1}{m!} \sum_{(l_1,\dots,l_m)\in P_{\{j_1,\dots,j_m\}}} (\mathcal{T})_{(l_1,\dots,l_m)} \tag{2.2}$$

(that is, to get $\pi_S(\mathcal{T})$ we should take the average of $m!$ tensors obtained from $\mathcal{T}$ by all possible permutations of the indices). The mapping π_S is called symmetrizer (also known as symmetrization map [24] or completely symmetric operator [23]; in [33] a matrix representation of π_S was called Kronecker product permutation matrix).

It is well known that mth-order $K\times\cdots\times K$ tensors can be vectorized into vectors of $\mathbb{R}^{K^m}$ in such a way that for any vectors $\mathbf{t}_1,\dots,\mathbf{t}_m \in \mathbb{R}^K$ the rank-1 tensor $\mathbf{t}_1\circ\cdots\circ\mathbf{t}_m$ corresponds to the vector $\mathbf{t}_1\otimes\cdots\otimes\mathbf{t}_m$. This allows us to consider the symmetrizer π_S on the space $\mathbb{R}^{K^m}$. In particular, by (2.2),

$$\pi_S(\mathbf{t}_1\otimes\cdots\otimes\mathbf{t}_m) = \frac{1}{m!} \sum_{(l_1,\dots,l_m)\in P_m} \mathbf{t}_{l_1}\otimes\cdots\otimes\mathbf{t}_{l_m}. \tag{2.3}$$

The following proposition makes the link between $\mathcal{B}(\mathbf{C})$ and $\mathcal{R}_m(\mathbf{C})$ and is the main result of this section.

PROPOSITION 2.13. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$, $K \leq R$, $m = R - K + 2$, *and* $k_{\mathbf{C}} \geq K - 1$. *Let also* $\mathcal{B}(\mathbf{C})$ *be defined by* (1.13) *and let* $\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}$ *denote the restriction of the mapping* $\mathcal{R}_m(\mathbf{C})^T : \mathbb{R}^{K^m} \to \mathbb{R}^{C_R^m}$ *onto* $\mathrm{range}(\pi_S)$. *Then*

(i) *The matrix* $\mathcal{R}_m(\mathbf{C})$ *has full column rank. Hence,* $\dim \mathrm{range}(\mathcal{R}_m(\mathbf{C})^T) = C_R^m$;

(ii) $\dim\left(\ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}\right)\right) = C_R^{K-1}$;

(iii) *If* $k_{\mathbf{C}} = K$, *then* $\ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}\right) = \mathrm{range}(\mathcal{B}(\mathbf{C})^{(m)})$.

In the remaining part of this subsection we prove Proposition 2.13. Readers who are mainly interested in the overall development and algorithms, can safely skip the rest of this section. We need auxiliary results and notations that we will also use in Subsection 3.3.

Let $\{\mathbf{e}_j^K\}_{j=1}^K$ denote the canonical basis of $\mathbb{R}^K$. Then $\{\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K\}_{(j_1,\ldots,j_m)\in R_K^m}$ is the canonical basis of $\mathbb{R}^{K^m}$ and by (2.3),

$$\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K) = \frac{1}{m!} \sum_{(l_1,\ldots,l_m)\in P_{\{j_1,\ldots,j_m\}}} \mathbf{e}_{l_1}^K \otimes \cdots \otimes \mathbf{e}_{l_m}^K. \tag{2.4}$$

Let the matrix $\mathbf{G} \in \mathbb{R}^{K^m \times C_{K+m-1}^m}$ be defined as follows:

$$\mathbf{G} \text{ has columns } \{\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K) : \ (j_1,\ldots,j_m) \in Q_K^m\}. \tag{2.5}$$

The following lemma follows directly from the definitions of π_S and $\mathbf{G}$ and is well known.

LEMMA 2.14. *[33] Let* π_S *and* $\mathbf{G}$ *be defined by* (2.4)–(2.5). *Then the columns of the matrix* $\mathbf{G}$ *form an orthogonal basis of* $\mathrm{range}(\pi_S)$; *in particular,* $\dim \mathrm{range}(\pi_S) = C_{K+m-1}^m$.

The following lemma explains that the matrix $\mathcal{R}_m(\mathbf{C})$ is obtained from $\mathbf{C}$ by picking all combinations of m columns, and symmetrizing the corresponding rank-1 tensor. Note that it is the symmetrization that introduces permanents.

LEMMA 2.15. *Let* $\mathbf{C} = \begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_R\end{bmatrix} \in \mathbb{R}^{K\times R}$. *Then*

$$\mathcal{R}_m(\mathbf{C}) = m!\begin{bmatrix}\pi_S(\mathbf{c}_1 \otimes \cdots \otimes \mathbf{c}_m) & \ldots & \pi_S(\mathbf{c}_{R-m+1} \otimes \cdots \otimes \mathbf{c}_R)\end{bmatrix}. \tag{2.6}$$

Proof. By (2.3), the $(i_1,\ldots,i_m)$-th entry of the vector $m!\pi_S(\mathbf{c}_{j_1} \otimes \cdots \otimes \mathbf{c}_{j_m})$ is equal to

$$\begin{aligned}\sum_{(l_1,\ldots,l_m)\in P_m} c_{i_1 j_{l_1}} \cdots c_{i_m j_{l_m}} &= \mathrm{perm}\begin{bmatrix} c_{i_1 j_1} & \ldots & c_{i_1 j_m} \\ \vdots & \vdots & \vdots \\ c_{i_m j_1} & \ldots & c_{i_m j_m}\end{bmatrix} \\ &= \mathrm{perm}\ \mathbf{C}((i_1,\ldots,i_m),(j_1,\ldots,j_m)).\end{aligned}$$

Hence, (2.6) follows from Definition 2.10. ∎

EXAMPLE 2.16. *Let the matrix* $\mathbf{C}$ *be as in Example 2.12. Then*

$$\mathcal{R}_2(\mathbf{C})^T = 2!\begin{bmatrix}\frac{1}{2!}([1\ 4]\otimes[2\ 5]+[2\ 5]\otimes[1\ 4]) \\ \frac{1}{2!}([1\ 4]\otimes[3\ 6]+[3\ 6]\otimes[1\ 4]) \\ \frac{1}{2!}([2\ 5]\otimes[3\ 6]+[3\ 6]\otimes[2\ 5])\end{bmatrix} = \begin{bmatrix}4 & 13 & 13 & 40 \\ 6 & 18 & 18 & 48 \\ 12 & 27 & 27 & 60\end{bmatrix}.$$

Let $\left\{\mathbf{e}_{(j_1,\ldots,j_m)}^{C_{K+m-1}^m}\right\}_{(j_1,\ldots,j_m)\in Q_K^m}$ denote the canonical basis of $\mathbb{R}^{C_{K+m-1}^m}$. Define the C_{K+m-1}^m-by-K^m matrix $\mathbf{H}$ as follows

$$\mathbf{H} \text{ has columns } \{\mathbf{e}_{[j_1,\ldots,j_m]}^{C_{K+m-1}^m} : \ (j_1,\ldots,j_m) \in R_K^m\}, \tag{2.7}$$

in which $[j_1, \ldots, j_m]$ denotes the ordered version of $(j_1, \ldots, j_m)$. For all K^m entries of a symmetric m-th order $K \times \cdots \times K$ tensor, the corresponding column of $\mathbf{H}$ contains a "1" at the first index combination (in lexicographic ordering) where that entry can be found. The matrix $\mathbf{H}$ can be used to "compress" symmetric $K \times \cdots \times K$ tensors by removing redundancies. The matrix $\mathbf{G}$ above does the opposite thing, so $\mathbf{G}$ and $\mathbf{H}$ act as each other's inverse. It is easy to prove that indeed $\mathbf{HG} = \mathbf{I}_{C^m_{K+m-1}}$. The relations in the following lemma reflect the same relationship and will be used in Subsection 3.3.

LEMMA 2.17. *Let* $\mathbf{C} \in \mathbb{R}^{K \times R}$ *and let the matrices* $\mathbf{G}$ *and* $\mathbf{H}$ *be defined by* (2.5) *and* (2.7)*, respectively. Then*

(i) $\mathcal{R}_m(\mathbf{C})^T = \mathcal{Q}_m(\mathbf{C})^T\mathbf{H}$*;*

(ii) $\mathcal{R}_m(\mathbf{C})^T\mathbf{G} = \mathcal{Q}_m(\mathbf{C})^T$*.*

Proof. As the proof is technical, it is given in the supplementary materials. □

Proof of Proposition 2.13. (i) Assume that there exists $\widehat{\mathbf{t}} = [t_{(1,\ldots,m)} \ \ldots \ t_{(R-m+1,\ldots,R)}]^T \in \mathbb{R}^{C^m_R}$ such that $\mathcal{R}_m(\mathbf{C})\widehat{\mathbf{t}} = \mathbf{0}$. Then, by Lemma 2.15,

$$\sum_{(p_1,\ldots,p_m)\in S^m_R} t_{(p_1,\ldots,p_m)}\pi_S(\mathbf{c}_{p_1} \otimes \cdots \otimes \mathbf{c}_{p_m}) = \mathbf{0}. \tag{2.8}$$

Let us fix $(i_1, \ldots, i_m) \in S^m_R$ and set $\{j_1, \ldots, j_{K-1}\} := \{1, \ldots, R\} \setminus \{i_1, \ldots, i_{m-1}\}$. Then $i_m \in \{j_1, \ldots, j_{K-1}\}$. Without loss of generality we can assume that $j_{K-1} = i_m$.

Since $k_{\mathbf{C}} \geq K-1$, it follows that there exists a vector $\mathbf{y}$ such that $\mathbf{y}$ is orthogonal to the vectors $\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_{K-2}}$, and $\mathbf{y}$ is not orthogonal to any of $\mathbf{c}_{i_1}, \ldots, \mathbf{c}_{i_m}$. Let $\alpha_{(p_1,\ldots,p_m)}$ denote the $(p_1, \ldots, p_m)$-th entry of the vector $\mathcal{R}_m(\mathbf{C})^T(\mathbf{y} \otimes \cdots \otimes \mathbf{y})$. Then, by Lemma 2.15,

$$\begin{aligned}\alpha_{(p_1,\ldots,p_m)} =& \pi_S(\mathbf{c}_{p_1} \otimes \cdots \otimes \mathbf{c}_{p_m})^T(\mathbf{y} \otimes \cdots \otimes \mathbf{y}) = \\ & \frac{1}{m!} \sum_{(l_1,\ldots,l_m)\in P_{\{p_1,\ldots,p_m\}}} (\mathbf{c}_{l_1}^T\mathbf{y}) \cdots (\mathbf{c}_{l_m}^T\mathbf{y}) = (\mathbf{c}_{p_1}^T\mathbf{y}) \cdots (\mathbf{c}_{p_m}^T\mathbf{y}).\end{aligned} \tag{2.9}$$

By the construction of $\mathbf{y}$, $\alpha_{(p_1,\ldots,p_m)} \neq 0$ if and only if $\{p_1, \ldots, p_m\} = \{i_1, \ldots, i_m\}$. Then, by (2.8)–(2.9),

$$\begin{aligned}0 = & \sum_{(p_1,\ldots,p_m)\in S^m_R} t_{(p_1,\ldots,p_m)}\pi_S(\mathbf{c}_{p_1} \otimes \cdots \otimes \mathbf{c}_{p_m})^T(\mathbf{y} \otimes \cdots \otimes \mathbf{y}) = \\ & \sum_{(p_1,\ldots,p_m)\in S^m_R} t_{(p_1,\ldots,p_m)}\alpha_{(p_1,\ldots,p_m)} = t_{(i_1,\ldots,i_m)}\alpha_{(i_1,\ldots,i_m)}.\end{aligned}$$

Hence, $t_{(i_1,\ldots,i_m)} = 0$. Since $(i_1, \ldots, i_m)$ was arbitrary we obtain $\widehat{\mathbf{t}} = \mathbf{0}$.

(ii) From step (i), Lemma 2.14, and Lemma 2.17 (i),(ii) it follows that

$$\begin{aligned}C^m_R =& \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T) \geq \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) = \\ & \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T\mathbf{G}) = \dim \operatorname{range}(\mathcal{Q}_m(\mathbf{C})^T) \geq \\ & \dim \operatorname{range}(\mathcal{Q}_m(\mathbf{C})^T\mathbf{H}) = \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T) = C^m_R.\end{aligned}$$

Hence, $\dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) = C^m_R$. By the ranknullity theorem,

$$\begin{aligned}\dim \ker\ (\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) =& \dim \operatorname{range}(\pi_S) - \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) = \\ & C^m_{K+m-1} - C^m_R = C^{R-K+2}_{R+1} - C^{R-K+2}_R = C^{K-1}_R.\end{aligned}$$

(iii) Let $\mathbf{y}$ denote the $(j_1,\dots,j_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$. It is clear that the vector $\underbrace{\mathbf{y}\otimes\cdots\otimes\mathbf{y}}_{m}$ is contained in $\operatorname{range}(\pi_S)$. Hence, $\operatorname{range}\left(\mathcal{B}(\mathbf{C})^{(m)}\right)\subseteq\operatorname{range}(\pi_S)$. By step (ii) and Proposition 1.10 (ii), $\dim\ker\left(\mathcal{R}_m(\mathbf{C})^T\restriction_{\operatorname{range}(\pi_S)}\right)=C_R^{K-1}=\dim\operatorname{range}\left(\mathcal{B}(\mathbf{C})^{(m)}\right)$. To complete the proof we must check that $\mathcal{R}_m(\mathbf{C})^T(\mathbf{y}\otimes\cdots\otimes\mathbf{y})=\mathbf{0}$ for all $(j_1,\dots,j_{K-1})\in S_R^{K-1}$. From the construction of the matrix $\mathcal{B}(\mathbf{C})$ it follows that $\mathbf{y}$ is orthogonal to the vectors $\mathbf{c}_{j_1},\dots,\mathbf{c}_{j_{K-1}}$. Since $(K-1)+m=R+1>R$, it follows that $(\mathbf{c}_{p_1}^T\mathbf{y})\cdots(\mathbf{c}_{p_m}^T\mathbf{y})=0$ for all $(p_1,\dots,p_m)\in S_R^m$. Hence, by (2.9), $\mathcal{R}_m(\mathbf{C})^T(\mathbf{y}\otimes\cdots\otimes\mathbf{y})=\mathbf{0}$. ∎

The following corollary of Proposition 2.13 will be used in Subsection 4.3.

COROLLARY 2.18. *Let the conditions of Proposition 2.13 hold and let $k_{\mathbf{C}}=K-1$. Then the subspace $\ker\left(\mathcal{R}_m(\mathbf{C})^T\restriction_{\operatorname{range}(\pi_S)}\right)$ cannot be spanned by vectors of the form $\{\mathbf{y}_p\otimes\mathbf{z}_p\}_{p=1}^{C_R^{K-1}}$, where $\mathbf{y}_p\in\mathbb{R}^K$ and $\mathbf{z}_p\in\mathbb{R}^{K^{m-1}}$.*

Proof. The proof is given in the supplementary materials. ∎

3. Transformation of the CPD using polarized compound matrices. In this section we derive the crucial expression (1.16). The matrix $\mathbf{R}_m(\mathcal{T})$ is constructed from polarized compound matrices of the slices of the given tensor $\mathcal{T}$. The entries of polarized compound matrices are mixed discriminants. The notions of mixed discriminants and polarized compound matrices are introduced in the first two subsections.

3.1. Mixed discriminants. The mixed discriminant is variant of the determinant that has more than one matrix argument.

DEFINITION 3.1. *[1] Let $\mathbf{T}_1,\dots,\mathbf{T}_m\in\mathbb{R}^{m\times m}$. The mixed discriminant, denoted by $\mathcal{D}(\mathbf{T}_1,\dots,\mathbf{T}_m)$, is defined as the coefficient of $x_1\cdots x_m$ in $\det(x_1\mathbf{T}_1+\cdots+x_m\mathbf{T}_m)$, that is*

$$\mathcal{D}(\mathbf{T}_1,\dots,\mathbf{T}_m)=\left.\frac{\partial^m\left(\det(x_1\mathbf{T}_1+\cdots+x_m\mathbf{T}_m)\right)}{\partial x_1\dots\partial x_m}\right|_{x_1=\cdots=x_m=0}. \tag{3.1}$$

For convenience, we have dropped the factor $1/m!$ before the fraction in (3.1). Definition 3.1 implies the following lemmas.

LEMMA 3.2. *[1] The mapping $(\mathbf{T}_1,\dots,\mathbf{T}_m)\to\mathcal{D}(\mathbf{T}_1,\dots,\mathbf{T}_m)$ is multilinear and symmetric in its arguments.*

LEMMA 3.3. *[10] Let $\mathbf{d}_1,\dots,\mathbf{d}_m\in\mathbb{R}^m$. Then $\mathcal{D}\left(\operatorname{Diag}(\mathbf{d}_1),\dots,\operatorname{Diag}(\mathbf{d}_m)\right)=\operatorname{perm}\begin{bmatrix}\mathbf{d}_1 & \dots & \mathbf{d}_m\end{bmatrix}$.*

Proof.

$$\mathcal{D}(\operatorname{Diag}(\begin{bmatrix}d_{11} & \dots & d_{m1}\end{bmatrix}),\dots,\operatorname{Diag}(\begin{bmatrix}d_{1m} & \dots & d_{mm}\end{bmatrix}))=$$
$$\left.\frac{\partial^m\left((x_1d_{11}+\cdots+x_md_{1m})\cdots(x_1d_{m1}+\cdots+x_md_{mm})\right)}{\partial x_1\dots\partial x_m}\right|_{x_1=\cdots=x_m=0}=$$
$$\sum_{(l_1,\dots,l_m)\in P_m}d_{1l_1}\cdots d_{ml_m}=\operatorname{perm}\begin{bmatrix}\mathbf{d}_1 & \dots & \mathbf{d}_m\end{bmatrix}. \qquad ∎$$

Mixed discriminants may be computed numerically from (3.1). A direct expression in terms of determinants is given in the following lemma.

LEMMA 3.4. *[2, 22] Let $\mathbf{T}_1,\dots,\mathbf{T}_m\in\mathbb{R}^{m\times m}$. Then*

$$\mathcal{D}(\mathbf{T}_1,\dots,\mathbf{T}_m)=\sum_{k=1}^{m}(-1)^{m-k}\sum_{1\le i_1<i_2<\cdots<i_k\le m}\det(\mathbf{T}_{i_1}+\cdots+\mathbf{T}_{i_k}). \tag{3.2}$$

The way in which (3.2) obtains the mixed discriminant from the determinant is an instance of a technique called polarization [20].

3.2. Polarized compound matrices. Let $m \geq 2$. In this subsection we discuss a polarized version of compound matrices, in which the mixed discriminant replaces the determinant.

DEFINITION 3.5. *Let* $\min(I,J) \geq m \geq 2$ *and let* $\mathbf{T}_1,\ldots,\mathbf{T}_m \in \mathbb{R}^{I\times J}$. *The* C_I^m*-by-*C_J^m *matrix* $\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m)$ *is defined by*

$$\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m) = \left.\frac{\partial^m\left(\mathcal{C}_m(x_1\mathbf{T}_1+\cdots+x_m\mathbf{T}_m)\right)}{\partial x_1\ldots\partial x_m}\right|_{x_1=\cdots=x_m=0}. \tag{3.3}$$

In the following lemmas we establish properties of $\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m)$.

LEMMA 3.6. *Let* $\mathbf{T}\in\mathbb{R}^{I\times J}$ *and* $\mathbf{d}_1,\ldots,\mathbf{d}_m\in\mathbb{R}^R$. *Then*

(i) *the mapping* $(\mathbf{T}_1,\ldots,\mathbf{T}_m)\to\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m)$ *is multilinear and symmetric in its arguments;*

(ii) *an equivalent expression for* $\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m)$ *is*

$$\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m) = \sum_{k=1}^{m}(-1)^{m-k}\sum_{1\leq i_1<i_2<\cdots<i_k\leq m}\mathcal{C}_m(\mathbf{T}_{i_1}+\cdots+\mathbf{T}_{i_k});$$

(iii) $\mathcal{F}_{m-1}(\mathbf{T},\ldots,\mathbf{T}) = m!\mathcal{C}_m(\mathbf{T})$;

(iv) $r_{\mathbf{T}}\leq m-1$ *if and only if* $\mathcal{F}_{m-1}(\mathbf{T},\ldots,\mathbf{T})=\mathbf{O}$;

(v) $\mathcal{F}_{m-1}\left(\mathrm{Diag}(\mathbf{d}_1),\ldots,\mathrm{Diag}(\mathbf{d}_m)\right) = \mathrm{Diag}\left(\mathcal{PC}_m(\begin{bmatrix}\mathbf{d}_1 & \ldots & \mathbf{d}_m\end{bmatrix})\right)$.

Proof. From Definitions 2.1 and 3.5 it follows that the (i,j)-th entry of the matrix $\mathcal{F}_{m-1}(\mathbf{T}_1,\ldots,\mathbf{T}_m)$ is equal to $\mathcal{D}(\mathbf{T}_1(S_I^m(i),S_J^m(j)),\ldots,\mathbf{T}_m(S_I^m(i),S_J^m(j)))$. Hence, statements (i) and (ii) follow from Lemma 3.2 and Lemma 3.4, respectively. Statement (iii) follows from (3.3). Statement (iv) follows from (iii) and Lemma 2.3 (2). Finally, (v) follows from Lemma 2.4, statement (ii), and Lemma 3.3. □

EXAMPLE 3.7.

$$\begin{aligned}\mathcal{F}_2(\mathbf{T}_1,\mathbf{T}_2,\mathbf{T}_3) &= \mathcal{C}_3(\mathbf{T}_1+\mathbf{T}_2+\mathbf{T}_3)-\\ &\mathcal{C}_3(\mathbf{T}_1+\mathbf{T}_2)-\mathcal{C}_3(\mathbf{T}_1+\mathbf{T}_3)-\mathcal{C}_3(\mathbf{T}_2+\mathbf{T}_3)+\mathcal{C}_3(\mathbf{T}_1)+\mathcal{C}_3(\mathbf{T}_2)+\mathcal{C}_3(\mathbf{T}_3).\end{aligned} \tag{3.4}$$

REMARK 3.8. *The polarized compound matrix is a matrix representation of the higher-order tensor obtained by the low-rank detection mapping in [6,29]. More specifically, in [6] a rank-*1 *detection mapping* ($m=2$) *was used to compute the CPD and in [29] a rank-*$(L,L,1)$ *detection mapping* (m *arbitrary*) *was used to compute the decomposition in rank-*$(L,L,1)$ *terms. Statement* (iv) *of Lemma 3.6 explains the terminology.*

The following counterpart of Lemma 2.5 holds for polarized compound matrices.

LEMMA 3.9. *Let* $\mathbf{A}\in\mathbb{R}^{I\times R}$, $\mathbf{B}\in\mathbb{R}^{J\times R}$, $\mathbf{d}_1,\ldots,\mathbf{d}_m\in\mathbb{R}^R$, *and* $m\leq\min(I,J,R)$. *Then*

$$\begin{aligned}&\mathcal{F}_{m-1}\left(\mathbf{A}\mathrm{Diag}(\mathbf{d}_1)\mathbf{B}^T,\ldots,\mathbf{A}\mathrm{Diag}(\mathbf{d}_m)\mathbf{B}^T\right) =\\ &\qquad\mathcal{C}_m(\mathbf{A})\mathrm{Diag}\left(\mathcal{PC}_m(\begin{bmatrix}\mathbf{d}_1 & \ldots & \mathbf{d}_m\end{bmatrix})\right)\mathcal{C}_m(\mathbf{B})^T.\end{aligned} \tag{3.5}$$

Proof. From Lemma 3.6 (ii) and Lemma 2.5 we have

$$\begin{aligned}&\mathcal{F}_{m-1}\left(\mathbf{A}\mathrm{Diag}(\mathbf{d}_1)\mathbf{B}^T,\ldots,\mathbf{A}\mathrm{Diag}(\mathbf{d}_m)\mathbf{B}^T\right) =\\ &\qquad\mathcal{C}_m(\mathbf{A})\mathcal{F}_{m-1}\left(\mathrm{Diag}(\mathbf{d}_1),\ldots,\mathrm{Diag}(\mathbf{d}_m)\right)\mathcal{C}_m(\mathbf{B})^T.\end{aligned} \tag{3.6}$$

Now (3.5) follows from (3.6) and Lemma 3.6 (v). □

3.3. Transformation of the tensor. We stack polarized compound matrices obtained from the slices of a given tensor in matrices $\mathbf{R}_m(\mathcal{T})$ and $\mathbf{Q}_m(\mathcal{T})$. In $\mathbf{R}_m(\mathcal{T})$ we consider all slice combinations, while in $\mathbf{Q}_m(\mathcal{T})$ we avoid doubles by taking into account the invariance of polarized compound matrices under permutation of their arguments. In our algorithms we will work with the smaller matrix $\mathbf{Q}_m(\mathcal{T})$ while in the theoretical development we will use $\mathbf{R}_m(\mathcal{T})$.

DEFINITION 3.10. *Let $\mathcal{T}$ be an $I \times J \times K$ tensor with frontal slices $\mathbf{T}_1, \dots, \mathbf{T}_K \in \mathbb{R}^{I\times J}$. The $(j_1, \dots, j_m)$-th column of the $C_I^m C_J^m$-by-K^m (resp. $C_I^m C_J^m$-by-C_{K+m-1}^m) matrix $\mathbf{R}_m(\mathcal{T})$ (resp. $\mathbf{Q}_m(\mathcal{T})$) equals* $\operatorname{vec}\left(\mathcal{F}_{m-1}(\mathbf{T}_{j_1}, \dots, \mathbf{T}_{j_m})\right)$*, where $(j_1, \dots, j_m) \in R_K^m$ (resp. Q_K^m).*

Let $\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}$ denote the restriction of the mapping $\mathbf{R}_m(\mathcal{T}) : \mathbb{R}^{K^m} \to \mathbb{R}^{C_I^m C_J^m}$ onto $\operatorname{range}(\pi_S)$. In the following lemma we express the matrices $\mathbf{R}_m(\mathcal{T})$ and $\mathbf{Q}_m(\mathcal{T})$ via the factor matrices of $\mathcal{T}$ and make a link between the kernel of $\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}$ and $\mathbf{Q}_m(\mathcal{T})$. These results are key to our overall derivation.

LEMMA 3.11. *Let $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, $\mathbf{C} \in \mathbb{R}^{K\times R}$, and $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. Then for $m \le \min(I, J, K, R)$,*

(i) $\mathbf{R}_m(\mathcal{T}) = [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\, \mathcal{R}_m(\mathbf{C})^T$;

(ii) $\mathbf{Q}_m(\mathcal{T}) = [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\, \mathcal{Q}_m(\mathbf{C})^T$;

(iii) $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}) = \mathbf{G} \ker(\mathbf{Q}_m(\mathcal{T}))$, *where $\mathbf{G}$ is defined in* (2.5).

Proof. (i) Let $\mathbf{c}^1, \dots, \mathbf{c}^K$ be the columns of the matrix $\mathbf{C}^T$. Recall that the frontal slices of $\mathcal{T}$ can be expressed as in (1.2). Then, by Lemma 3.9 and identity (1.3),

$$\begin{aligned}
&\operatorname{vec}\left(\mathcal{F}_{m-1}\left(\mathbf{A}\operatorname{Diag}(\mathbf{c}^{j_1})\mathbf{B}^T, \dots, \mathbf{A}\operatorname{Diag}(\mathbf{c}^{j_m})\mathbf{B}^T\right)\right) = \\
&\operatorname{vec}\left(\mathcal{C}_m(\mathbf{A})\operatorname{Diag}\left(\mathcal{PC}_m(\begin{bmatrix}\mathbf{c}^{j_1} & \dots & \mathbf{c}^{j_m}\end{bmatrix})\right)\mathcal{C}_m(\mathbf{B})^T\right) = \\
&\qquad [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\, \mathcal{PC}_m(\begin{bmatrix}\mathbf{c}^{j_1} & \dots & \mathbf{c}^{j_m}\end{bmatrix}).
\end{aligned}$$

Now (i) and (ii) follow from Definition 3.10 and Lemma 2.11.

(iii) From (i), (ii), and Lemma 2.17 (ii) it follows that $\mathbf{R}_m(\mathcal{T})\mathbf{G} = \mathbf{Q}_m(\mathcal{T})$. Since, by Lemma 2.14, $\operatorname{range}(\pi_S) = \operatorname{range}(\mathbf{G})$ we obtain (iii). ∎

4. Overall results and algorithms.

4.1. Algorithm 1 and Theorem 1.6. Overall, Algorithm 1 goes now as follows. We first compute $\mathbf{Q}_m(\mathcal{T})$ from $\mathcal{T}$, determine its null space, which, after symmetrization, yields $\ker\left(\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}\right)$, as explained in Lemma 3.11 (iii). The following lemma makes now, for a particular choice of m, a connection with $\mathcal{B}(\mathbf{C})$.

LEMMA 4.1. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $m := R - K + 2$, and let $\mathbf{R}_m(\mathcal{T})$ be defined in Definition 3.10. Assume that $k_{\mathbf{C}} = K$ and that $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank. Then*

(i) $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}) = \operatorname{range}(\mathcal{B}(\mathbf{C})^{(m)})$;

(ii) $\dim\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}) = C_R^{K-1}$;

Proof. Since $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank, it follows from Lemma 3.11 (i) that $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\operatorname{range}(\pi_S)}) = \ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}\right)$. Statements (i) and (ii) now follow from Proposition 2.13 (iii) and (ii), respectively. ∎

So far, we have obtained from $\mathcal{T}$ a basis for the column space of $\mathcal{B}(\mathbf{C})^{(m)}$. The following lemma explains that the basis vectors may be stacked in a tensor that has $\mathcal{B}(\mathbf{C})$ as factor matrix. Moreover, the CPD may be computed by a GEVD as in Theorem 1.1.

LEMMA 4.2. *Suppose that the conditions of Lemma 4.1 hold. Let* $\mathbf{W}$ *be a* $K^m \times C_R^{K-1}$ *matrix such that*

$$\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}) = \mathrm{range}\ (\mathbf{W}) \tag{4.1}$$

and let $\mathcal{W}$ *be the* $K \times K^{m-1} \times C_R^{K-1}$ *tensor such that* $\mathbf{W} = \mathrm{Matr}(\mathcal{W})$. *Then*

(i) *there exists a nonsingular* $C_R^{K-1} \times C_R^{K-1}$ *matrix* $\mathbf{M}$ *such that*

$$\mathcal{W} = \left[\mathcal{B}(\mathbf{C}), \mathcal{B}(\mathbf{C})^{(m-1)}, \mathbf{M}\right]_{C_R^{K-1}}; \tag{4.2}$$

(ii) $r_{\mathcal{W}} = C_R^{K-1}$ *and the CPD of* $\mathcal{W}$ *is unique and can be found algebraically.*

Proof. (i) From Lemma 4.1 (ii) and (4.1) it follows that there exists a nonsingular $C_R^{K-1} \times C_R^{K-1}$ matrix $\mathbf{M}$, such that $\mathbf{W} = \mathcal{B}(\mathbf{C})^{(m)}\mathbf{M}^T = \left(\mathcal{B}(\mathbf{C}) \odot \mathcal{B}(\mathbf{C})^{(m-1)}\right)\mathbf{M}^T$. Hence, by (1.4), (4.2) holds.

(ii) From Proposition 1.10 it follows that $k_{\mathcal{B}(\mathbf{C})} \geq 2$ and that the matrix $\mathcal{B}(\mathbf{C})^{(m-1)}$ has rank C_R^{K-1}. The statement now follows from Theorem 1.1. □

After finding $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling, we may find $\mathbf{C}$ as explained in Subsection 2.1. The following Lemma completes the proof of Theorem 1.6 (ii). Its proof shows how the other factor matrices may be determined once $\mathbf{C}$ has been obtained. The computation involves another CPD of the form in Theorem 1.1. The result is a variant of [38, Theorem 3.8]; in this step of the derivation we do not assume that the decomposition is canonical.

LEMMA 4.3. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and the* $K \times R$ *matrix* $\mathbf{C}$ *be known. Assume that* $k_{\mathbf{C}} = K \geq 2$, *and that* $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) + k_{\mathbf{C}} \geq R + 2$. *Then the matrices* $\mathbf{A}$, $\mathbf{B}$ *can be found algebraically up to column scaling.*

Proof. We obviously have $k_{\mathbf{C}} = r_{\mathbf{C}} = K$. Let $\mathbf{X} = \begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_K\end{bmatrix}$. By multiplying with $\mathbf{X}^{-1}$ we will create a CPD with $K-2$ terms less than R. It is clear that the matrix formed by the first two rows of $\mathbf{X}^{-1}\mathbf{C}$ has the form $\begin{bmatrix}\mathbf{I}_2 & \mathbf{O}_{2\times(K-2)} & \mathbf{Y}\end{bmatrix}$, where $\mathbf{Y}$ is a $2 \times (R-K)$ matrix. Define $\widetilde{\mathbf{A}} := \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_{K+1} & \dots & \mathbf{a}_R\end{bmatrix}$, $\widetilde{\mathbf{B}} := \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_{K+1} & \dots & \mathbf{b}_R\end{bmatrix}$, and $\widetilde{\mathbf{C}} = \begin{bmatrix}\mathbf{I}_2 & \mathbf{Y}\end{bmatrix}$. Let also $\widetilde{\mathcal{T}}$ denote the $I \times J \times 2$ tensor such that $\mathrm{Matr}(\widetilde{\mathcal{T}})$ coincides with the first two columns of the matrix $\mathrm{Matr}(\mathcal{T})\mathbf{X}^{-T}$. From (1.4) it follows that $\mathrm{Matr}(\mathcal{T})\mathbf{X}^{-T} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{X}^{-T}$. Hence, $\mathrm{Matr}(\widetilde{\mathcal{T}}) = (\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}})\widetilde{\mathbf{C}}^T$ or $\widetilde{\mathcal{T}} = [\widetilde{\mathbf{A}}, \widetilde{\mathbf{B}}, \widetilde{\mathbf{C}}]_{R-K+2}$, which is of the desired form.

It is easy to show that $\widetilde{\mathcal{T}}$ satisfies the conditions of Theorem 1.1, which means that its rank is $R-K+2$, that its CPD is unique and that the factor matrices may be found algebraically. The indeterminacies in $\widetilde{\mathcal{T}} = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_{R-K+2}$ are limited to the existence of a permutation matrix $\mathbf{P}$ and a nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\widetilde{\mathbf{C}} = \widehat{\mathbf{C}}\mathbf{P}\mathbf{\Lambda}$ and $\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}} = (\widehat{\mathbf{A}} \odot \widehat{\mathbf{B}})\mathbf{P}\mathbf{\Lambda}^{-1}$.

So far we have algebraically found the columns of the matrices $\mathbf{A}$, $\mathbf{B}$, and hence $\mathbf{A} \odot \mathbf{B}$, with indices in $I := \{1, 2, K+1, \dots, R\}$. Let $\bar{\mathbf{A}}$, $\bar{\mathbf{B}}$, and $\bar{\mathbf{C}}$ be the submatrices of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, respectively, formed by the columns with indices in $\{3, \dots, K\}$. We now subtract the rank-1 terms that we already know to obtain $\mathcal{T} - \sum_{r \in I} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r = \left[\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}\right]_{K-2} =: \bar{\mathcal{T}}$ or $(\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T = \mathrm{Matr}(\bar{\mathcal{T}})$. Since the matrix $\bar{\mathbf{C}}$ has full column rank, the columns of the matrix $\mathbf{A} \odot \mathbf{B}$ with indices in $\{3, \dots, K\}$ coincide with the columns of $\mathrm{Matr}(\bar{\mathcal{T}})\bar{\mathbf{C}}^\dagger$. Now that also the columns of $\mathbf{A} \odot \mathbf{B}$ with indices in $\{3, \dots, K\}$ have been found, $\mathbf{a}_r$ and $\mathbf{b}_r$ are easily obtained by understanding that $\mathbf{a}_r \otimes \mathbf{b}_r = \mathrm{vec}(\mathbf{b}_r\mathbf{a}_r^T)$, $r = 1, \dots, R$. □

The overall procedure that constitutes the proof of Theorem 1.6 (ii) is summarized in Algorithm 1. Phase 2 is formulated in a way that has combinatorial complexity

Algorithm 1 (Computation of $\mathbf{C}$, then $\mathbf{A}$ and $\mathbf{B}$)

Input: $\mathcal{T} \in \mathbb{R}^{I\times J\times K}$ and $R \geq 2$ 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 $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $k_{\mathbf{C}} = K \geq 2$, and $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank for $m = R - K + 2$.

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

Phase 1 (based on Lemma 4.2): 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

Apply Lemma 3.11 (iii) to find the $K^m \times C_R^{K-1}$ matrix $\mathbf{W}$ such that (4.1) *holds*

1: Construct the $C_I^m C_J^m$-by-C_{K+m-1}^m matrix $\mathbf{Q}_m(\mathcal{T})$ by Definition 3.10

2: Find $\bar{\mathbf{w}}_1, \ldots, \bar{\mathbf{w}}_{C_R^{K-1}}$ which form a basis of $\ker(\mathbf{Q}_m(\mathcal{T}))$

3: $\mathbf{W} \leftarrow \mathbf{G}\begin{bmatrix}\bar{\mathbf{w}}_1 & \ldots & \bar{\mathbf{w}}_{C_R^{K-1}}\end{bmatrix}$, where $\mathbf{G} \in \mathbb{R}^{K^m\times C_{K+m-1}^m}$ is defined by (2.5)

Apply Theorem 1.1 (ii) *to find* $\mathbf{F}$

4: $\mathcal{W} \leftarrow \mathrm{Tens}(\mathbf{W}, K, K^{m-1})$

5: Compute the CPD $\mathcal{W} = [\mathbf{F}, \mathbf{F}_2, \mathbf{F}_3]_{C_R^{K-1}}$ ($\mathbf{F}_2$ *and* $\mathbf{F}_3$ *are a by-product*) (GEVD)

Phase 2 (based on properties (P3)–(P4)**):** Find the matrix $\mathbf{C}$

6: Compute R subsets of C_{R-1}^{K-2} columns of $\mathbf{F}$ that are linearly dependent

7: Compute $\mathbf{c}_1, \ldots, \mathbf{c}_R$ as orthogonal complements to sets found in step 6

Phase 3 (based on Lemma 4.3): Find the matrices $\mathbf{A}$ and $\mathbf{B}$

Apply Theorem 1.1 (ii) *to find the columns of $\mathbf{S} := \mathbf{A}\odot\mathbf{B}$ with indices in* $\{1, 2, K+1, \ldots, R\}$

8: $\mathbf{Z} = \begin{bmatrix}\mathbf{z}_1 & \ldots & \mathbf{z}_K\end{bmatrix} \leftarrow \mathrm{Matr}(\mathcal{T})\begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_K\end{bmatrix}^{-T}$

9: $\widetilde{\mathcal{T}} \leftarrow \mathrm{Tens}(\begin{bmatrix}\mathbf{z}_1 & \mathbf{z}_2\end{bmatrix}, I, J)$

10: Compute the CPD $\widetilde{\mathcal{T}} = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_{R-K+2}$ (GEVD)

11: $\widetilde{\mathbf{C}} \leftarrow \begin{bmatrix}\mathbf{I}_2 & \mathbf{Y}\end{bmatrix}$, where $\mathbf{Y}$ is the $2 \times R$ submatrix in upper right-hand corner of $\begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_K\end{bmatrix}^{-1}\mathbf{C}$

12: Compute permutation matrix $\mathbf{P}$ and diagonal matrix $\mathbf{\Lambda}$ such that $\widetilde{\mathbf{C}} = \widehat{\mathbf{C}}\mathbf{P}\mathbf{\Lambda}$

13: $\begin{bmatrix}\mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_{K+1} \ldots & \mathbf{s}_R\end{bmatrix} \leftarrow (\widehat{\mathbf{A}} \odot \widehat{\mathbf{B}})\mathbf{P}\mathbf{\Lambda}^{-1}$

Find the columns of $\mathbf{S}$ with indices in $\{3, \ldots, K\}$

14: $\mathbf{Z} \leftarrow \mathrm{Matr}(\mathcal{T}) - (\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}})\begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_{K+1} & \ldots & \mathbf{c}_R\end{bmatrix}^T$

15: $\begin{bmatrix}\mathbf{s}_3 & \ldots & \mathbf{s}_K\end{bmatrix} \leftarrow \mathbf{Z}\begin{bmatrix}\mathbf{c}_3 & \ldots & \mathbf{c}_K\end{bmatrix}^\dagger$

16: Find the columns of $\mathbf{A}$ and $\mathbf{B}$ from the equations $\mathbf{a}_r \otimes \mathbf{b}_r = \mathbf{s}_r$, $r = 1, \ldots, R$

and quickly becomes computationally infeasible. The amount of work may be reduced by exploiting the dependencies in $\mathbf{F}$ only partially.

4.2. Algorithm 2. We derive an algorithmic variant that further reduces the computational cost. This algorithm is given in Algorithm 2 below. While Algorithm 1 first determines $\mathbf{C}$ and then finds $\mathbf{A}$ and $\mathbf{B}$, Algorithm 2 works the other way around. The basic idea is as follows. Like in Algorithm 1, we first find a matrix $\mathbf{F}$ that is equal to $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. If $\mathbf{C}$ is square, we have from Subsection 2.1 that $\mathcal{B}(\mathbf{C}) = \det(\mathbf{C})\mathbf{C}^{-T}\mathbf{L}$ and multiplication of $\mathcal{T}$ with $\mathbf{F}^T$ in the third mode yields a tensor of which every frontal slice is a rank-1 matrix, proportional to $\mathbf{a}_r\mathbf{b}_r^T$ for some $r \in \{1, \ldots, R\}$. On the other hand, if $\mathbf{C}$ is rectangular ($K < R$), then multiplication with $\mathbf{F}^T$ yields a tensor of which all slices are rank-$(R-K+1)$ matrices,

Algorithm 2 (Computation of $\mathbf{A}$ and $\mathbf{B}$, then $\mathbf{C}$)

Input: $\mathcal{T} \in \mathbb{R}^{I\times J\times K}$ and $R \geq 2$ 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 $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $k_{\mathbf{C}} = K \geq 2$, and $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank for $m = R - K + 2$.

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

Phase 1 (based on Lemma 4.2): 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

1–5: Identical to Algorithm 1

Phase 2 (based on Lemma 4.4): Find the matrices $\mathbf{A}$ and $\mathbf{B}$

6: $\mathcal{V} \leftarrow \text{Tens}(\mathbf{V}, I, J)$, where $\mathbf{V} = \text{Matr}(\mathcal{T})\mathbf{F}$
Let $\mathbf{V}_1, \ldots, \mathbf{V}_{C_R^{K-1}}$ *denote the frontal slices of* $\mathcal{V}$
Let $\mathcal{V}_{ij}$ *denote the tensor with frontal slices* $\mathbf{V}_i$ *and* $\mathbf{V}_j$
Find $C_R^m C_m^2$ *pairs* (i, j) *such that* $r_{\mathcal{V}_{ij}} = m$

7: $\mathcal{J} \leftarrow \{(i, j) : \ r_{[\mathbf{V}_i\ \mathbf{V}_j]} = r_{[\mathbf{V}_i^T\ \mathbf{V}_j^T]} = m,\ 1 \leq i < j \leq C_R^{K-1}\}$
Apply Theorem 1.1 (ii) *to find* $C_R^m C_m^2$ *sets of* m *columns of* $\mathbf{A}$ *and* $\mathbf{B}$ *each*

8: Find the CPD $\mathcal{V}_{ij} = [\mathbf{A}_{ij}, \mathbf{B}_{ij}, \mathbf{C}_{ij}]_m$ for each $(i, j) \in \mathcal{J}$ (GEVD)
($\mathbf{C}_{ij}$ *are a by-product*)

9: $\widetilde{\mathbf{A}} \leftarrow$ the $(I \times mC_R^m C_m^2)$ matrix formed by the columns of the matrices $\mathbf{A}_{ij}$

10: $\widetilde{\mathbf{B}} \leftarrow$ the $(I \times mC_R^m C_m^2)$ matrix formed by the columns of the matrices $\mathbf{B}_{ij}$

11: Choose $r_1, \ldots, r_R$ such that the sets $\{\widetilde{\mathbf{a}}_{r_1}, \ldots, \widetilde{\mathbf{a}}_{r_R}\}$ and $\{\widetilde{\mathbf{b}}_{r_1}, \ldots, \widetilde{\mathbf{b}}_{r_R}\}$ do not contain collinear vectors

12: $\mathbf{A} \leftarrow [\widetilde{\mathbf{a}}_{r_1} \quad \ldots \quad \widetilde{\mathbf{a}}_{r_R}]$, $\mathbf{B} \leftarrow [\widetilde{\mathbf{b}}_{r_1} \quad \ldots \quad \widetilde{\mathbf{b}}_{r_R}]$

Phase 3: Find the matrix $\mathbf{C}$

13: $\mathbf{C} \leftarrow \left((\mathbf{A} \odot \mathbf{B})^{\dagger}\text{Matr}(\mathcal{T})\right)^T$

generated by $R - K + 1$ rank-1 matrices $\mathbf{a}_r\mathbf{b}_r^T$. If we choose slices that have all but one rank-1 matrix in common, these form a tensor that is as in Theorem 1.1 and of which the CPD yields $R-K+2$ columns of $\mathbf{A}$ and $\mathbf{B}$. The result is formalized in the following lemma. The second statement implies that we do not have to compute the CPD to verify whether a slice combination is suitable.

LEMMA 4.4. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, *the matrix* $\mathbf{F} \in \mathbb{R}^{K\times C_R^{K-1}}$ *coincide with* $\mathcal{B}(\mathbf{C})$ *up to column permutation and scaling,* $\mathcal{V} = [\mathbf{A}, \mathbf{B}, \mathbf{F}^T\mathbf{C}]_R$, $k_{\mathbf{C}} = K$, $m := R - K + 2$, *and* $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ *have full column rank. Let also* $\mathbf{y}_1, \ldots, \mathbf{y}_{C_R^{K-1}}$ *denote the columns of* $\mathbf{C}^T\mathbf{F}$ *and* $\mathbf{V}_1, \ldots, \mathbf{V}_{C_R^{K-1}}$ *denote the frontal slices of* $\mathcal{V}$. *Then the following statements are equivalent.*

(i) *The matrix* $[\mathbf{y}_i\ \mathbf{y}_j]$ *has exactly* $K - 2$ *zero rows.*
(ii) *The matrices* $[\mathbf{V}_i\ \mathbf{V}_j]$ *and* $[\mathbf{V}_i^T\ \mathbf{V}_j^T]$ *have rank* m.
(iii) *The tensor* $\mathcal{V}_{ij}$ *formed by the frontal slices* $\mathbf{V}_i$ *and* $\mathbf{V}_j$ *has rank* m.

The CPD $\mathcal{V}_{ij} = \left[\begin{bmatrix}\mathbf{a}_{p_1} & \ldots & \mathbf{a}_{p_m}\end{bmatrix}, \begin{bmatrix}\mathbf{b}_{p_1} & \ldots & \mathbf{b}_{p_m}\end{bmatrix}, \widehat{\mathbf{C}}\right]_m$ *can be found algebraically by Theorem 1.1* (ii) *and the indices* $p_1, \ldots, p_m$ *are uniquely defined by the pair* (i, j).

Proof. Since the proof is technical, it is given in the supplementary materials. □

To summarize, we first find a matrix $\mathbf{F} \in \mathbb{R}^{K\times C_R^{K-1}}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. We construct a tensor $\mathcal{V}$ from the tensor $\mathcal{T}$ and the matrix $\mathbf{F}$ as in Lemma 4.4 and choose slice combinations for which $[\mathbf{V}_i\ \mathbf{V}_j]$ and

$[\mathbf{V}_i^T\ \mathbf{V}_j^T]$ have rank m. For each such slice combination the CPD of the corresponding tensor yields m columns of $\mathbf{A}$ and $\mathbf{B}$. In this way we obtain all columns of $\mathbf{A}$ and $\mathbf{B}$. Overall, there exist exactly $C_R^m C_m^2$ pairs (i,j) such that (i)–(iii) hold. The amount of work can be reduced by finding enough, instead of all, tensors $\mathcal{V}_{ij}$ that yield columns of $\mathbf{A}$ and $\mathbf{B}$. The matrix $\mathbf{C}$ is finally obtained by $\mathbf{C} = \big((\mathbf{A} \odot \mathbf{B})^{\dagger}\mathrm{Matr}(\mathcal{T})\big)^T$.

4.3. Discussion of working conditions. One may wonder what happens if the conditions in Theorem 1.6 are not satisfied, or, the other way around, under which circumstances the algorithms will fail. It turns out that, at least if the tensor rank is known, the crucial steps are steps 2 and 5 of Phase 1. If these do not pose problems, then the overall algorithms will work. Step 2 poses a problem when $\dim\ker(\mathbf{Q}_m(\mathcal{T})) \geq C_R^{K-1}$. This indicates that $k_{\mathbf{C}} < K-1$ and/or that $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ does not have full column rank. If step 2 does not pose a problem, but step 5 does, then $k_{\mathbf{C}} = K-1$. This is formalized in the following lemma.

LEMMA 4.5. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ be a CPD of $\mathcal{T}$ and $m = R - K + 2$. Let the matrix $\mathbf{Q}_m(\mathcal{T})$ be defined by Definition 3.10, and let the tensor $\mathcal{W}$ be constructed in steps 3–4 of Phase 1.*

(i) *If $\dim\ker(\mathbf{Q}_m(\mathcal{T})) = C_R^{K-1}$, then $k_{\mathbf{C}} \geq K-1$ and the matrix $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank;*

(ii) *If additionally, $r_{\mathcal{W}} = C_R^{K-1}$ and $\mathcal{W} = [\mathbf{F}, \mathbf{F}_2, \mathbf{F}_3]_{C_R^{K-1}}$, where $k_{\mathbf{F}} \geq 2$ and the matrices $\mathbf{F}_2$ and $\mathbf{F}_3$ have full column rank, then $k_{\mathbf{C}} = K$.*

Proof. (i) The proof is by contradiction. Assume that $k_{\mathbf{C}} \leq K-2$. Then $\mathbf{C}$ has $K-1$ columns that are linearly dependent. Without loss of generality we can assume that these columns are $\mathbf{c}_m, \ldots, \mathbf{c}_R$. Then the columns $\{\pi_s(\mathbf{c}_1 \otimes \cdots \otimes \mathbf{c}_{m-1} \otimes \mathbf{c}_k)\}_{k=m}^R$ of the matrix $\mathcal{R}_m(\mathbf{C})$ are also linearly dependent. Hence, by Lemma 2.17 (ii), the $C_R^m \times C_{K+m-1}^m$ matrix $\mathcal{Q}_m(\mathbf{C})^T$ has linearly dependent rows, which implies that $\dim\big(\ker\big(\mathcal{Q}_m(\mathbf{C})^T\big)\big) \geq C_{K+m-1}^m - C_R^m + 1 = C_R^{K-1} + 1$. On the other hand, by Lemma 3.11 (ii), $\dim(\ker(\mathbf{Q}_m(\mathcal{T}))) \geq \dim\big(\ker\big(\mathcal{Q}_m(\mathbf{C})^T\big)\big)$, which is a contradiction with $\dim(\ker(\mathbf{Q}_m(\mathcal{T}))) = C_R^{K-1}$.

We have proved that $k_{\mathbf{C}} \geq K-1$. By Proposition 2.13 (i), the matrix $\mathcal{R}_m(\mathbf{C})$ has full column rank. Since by Lemma 2.17 (i), $\mathcal{R}_m(\mathbf{C}) = \mathbf{H}^T\mathcal{Q}_m(\mathbf{C})$, it follows that the matrix $\mathcal{Q}_m(\mathbf{C})$ also has full column rank. Hence, $\dim\big(\ker\big(\mathcal{Q}_m(\mathbf{C})^T\big)\big) = C_{K+m-1}^m - C_R^m = C_R^{K-1}$. Hence, by Lemma 3.11 (ii),

$$\dim(\ker(\mathbf{Q}_m(\mathcal{T}))) = \dim(\ker(\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))) + \dim\big(\ker\big(\mathcal{Q}_m(\mathbf{C})^T\big)\big).$$

Since, by assumption, $\dim\ker(\mathbf{Q}_m(\mathcal{T})) = C_R^{K-1}$, the matrix $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank.

(ii) It is clear that the columns of the matrix $\mathbf{F} \odot \mathbf{F}_2$ form a basis of $\mathrm{range}(\mathbf{W})$. By construction of $\mathbf{W}$, $\mathrm{range}(\mathbf{W}) = \ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)})$, and by Lemma 3.11 (i), $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}) = \ker\big(\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}\big)$. Hence, by Corollary 2.18, $k_{\mathbf{C}} = K$. ∎

4.4. Theorem 1.7. It remains to prove Theorem 1.7 (ii). In the proof we construct a new tensor $\bar{\mathcal{T}}$ that has the same first two factor matrices as $\mathcal{T}$ and the CPD of which can be found by Algorithm 1 or Algorithm 2. Although, by construction of $\bar{\mathcal{T}}$, its frontal slices are random linear combinations of the frontal slices of $\mathcal{T}$, we still call the overall procedure "algebraic" because the proof of Theorem 1.1 is also based on the same random slice mixture idea (see [21] and references therein).

Proof. Let the matrix $\mathbf{C}$ have K rows, $\mathbf{X}$ be a $k_{\mathbf{C}} \times K$ matrix, and $\bar{\mathcal{T}} := [\mathbf{A}, \mathbf{B}, \mathbf{XC}]_R$. Then $\mathbf{XC} \in \mathbb{R}^{k_{\mathbf{C}} \times R}$ and by (1.4), $\mathrm{Matr}(\mathcal{T})\mathbf{X}^T := (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{X}^T = (\mathbf{A} \odot \mathbf{B})(\mathbf{XC})^T = \mathrm{Matr}(\bar{\mathcal{T}})$. Thus, the multiplication of the third factor matrix of $\mathcal{T}$ by $\mathbf{X}$ from the left is equivalent to the multiplication of the matrix unfolding $\mathrm{Matr}(\mathcal{T})$ by $\mathbf{X}^T$ from the right.

(i) Assume that $\mathbf{X}$ is such that $r_{\mathbf{XC}} = k_{\mathbf{XC}} = k_{\mathbf{C}}$. Then by Theorem 1.6, the CPD of $\bar{\mathcal{T}}$ is unique and can be found algebraically. In particular, the matrix $\mathbf{A} \odot \mathbf{B}$ has full column rank and can be found up to column permutation and scaling. Hence, $\mathbf{C} = \big((\mathbf{A} \odot \mathbf{B})^{\dagger}\mathrm{Matr}(\mathcal{T})\big)^T$, and the proof is completed.

(ii) It remains to present a construction of the matrix $\mathbf{X}$ such that $k_{\mathbf{XC}} = k_{\mathbf{C}}$. It is clear that $k_{\mathbf{XC}} \le k_{\mathbf{C}}$. We claim that $k_{\mathbf{XC}} = k_{\mathbf{C}}$ for generic $\mathbf{X}$. Namely,

$$\mu\{\mathrm{vec}(\mathbf{X}) : \ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K}, \ \ k_{\mathbf{XC}} < k_{\mathbf{C}}\} = 0, \tag{4.3}$$

where μ denotes the Lebesgue measure on $\mathbb{R}^{k_{\mathbf{C}}K}$. It is well known that the zero set of a nonzero polynomial has Lebesgue measure zero. Hence, for a nonzero vector $\mathbf{f} \in \mathbb{R}^{C_K^{k_{\mathbf{C}}}}$ we obtain

$$\mu\{\mathrm{vec}(\mathbf{X}) : \ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K}, \ \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathbf{f} = 0\} = 0. \tag{4.4}$$

From Lemma 2.3 (1) it follows that the matrix $\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C})$ has all columns nonzero. By Lemma 2.5, $k_{\mathbf{XC}} < k_{\mathbf{C}}$ if and only if the vector $\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C}) = \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{XC})$ has a zero entry. Hence, by (4.4),

$$\begin{aligned} &\{\mathrm{vec}(\mathbf{X}) : \ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K}, \ \ k_{\mathbf{XC}} < k_{\mathbf{C}}\} = \\ &\{\mathrm{vec}(\mathbf{X}) : \ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K}, \ \ \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C}) \text{ has a zero entry}\} = \\ &\qquad \bigcup_{\mathbf{f} \text{ is a column of } \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C})} \{\mathrm{vec}(\mathbf{X}) : \ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K}, \ \ \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathbf{f} = 0\}. \end{aligned} \tag{4.5}$$

Now (4.3) follows from (4.5) and (4.4). ∎

Let the conditions of Theorem 1.7, Corollary 1.8 or Corollary 1.9 hold. The following procedure for computing the CPD follows from the proof of Theorem 1.7 (ii). First, we generate a random $k_{\mathbf{C}} \times K$ matrix $\mathbf{X}$ and set $\bar{\mathbf{T}} = \mathrm{Matr}(\mathcal{T})\mathbf{X}^T$. Then the $I \times J \times k_{\mathbf{C}}$ tensor $\bar{\mathcal{T}} := \mathrm{Tens}(\bar{\mathbf{T}}, I, J)$ satisfies the conditions of Theorem 1.6 (ii) with K replaced by $k_{\mathbf{C}}$. Hence, the CPD $\bar{\mathcal{T}} = [\mathbf{A}, \mathbf{B}, \mathbf{XC}]_R$ can be found by Algorithm 1 or Algorithm 2. Finally, the matrix $\mathbf{C}$ is obtained by $\mathbf{C} = \big((\mathbf{A} \odot \mathbf{B})^{\dagger}\mathrm{Matr}(\mathcal{T})\big)^T$.

4.5. Examples. EXAMPLE 4.6. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_5$ *with*

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Since condition (1.6) *does not hold, the rank and uniqueness of the CPD do not follow from Kruskal's Theorem 1.3. One can easily check that the conditions of Theorem 1.6 hold for* $m = 5-4+2 = 3$. *Hence, the factor matrices of* $\mathcal{T}$ *can be found by Algorithms 1 and 2.*

Phase 1 of Algorithms 1 and 2. *The frontal slices of* $\mathcal{T}$ *are*

$$\mathbf{T}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{T}_2 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{T}_3 = \mathbf{T}_1^T, \quad \mathbf{T}_4 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We construct the $C_4^3C_4^3$-by-C_6^3 (or 16-by-20) matrix $\mathbf{Q}_3(\mathcal{T})$ by Definition 3.10. For instance, the $(1,2,3)$-rd (or the 6-th) column of $\mathbf{Q}_3(\mathcal{T})$ is equal to $\mathrm{vec}(\mathcal{F}_2(\mathbf{T}_1,\mathbf{T}_2,\mathbf{T}_3))$, where $\mathcal{F}_2(\mathbf{T}_1,\mathbf{T}_2,\mathbf{T}_3)$ is computed by (3.4) and equals

$$\mathcal{F}_2(\mathbf{T}_1,\mathbf{T}_2,\mathbf{T}_3) = -\begin{bmatrix} 0 & -3 & 0 & 3 \\ 0 & 1 & 1 & 0 \\ 3 & 4 & 1 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}.$$

The full matrix $\mathbf{Q}_3(\mathcal{T})$ is given in the supplementary materials. It can be checked that $\ker(\mathbf{Q}_3(\mathcal{T})) = \mathrm{range}(\overline{\mathbf{W}})$, where

$$\overline{\mathbf{W}} = [\mathbf{e}_1^{20} \quad \mathbf{e}_{11}^{20} \quad \mathbf{e}_{17}^{20} \quad \mathbf{e}_{20}^{20} \quad \mathbf{e}_{2,-5}^{20} \quad \mathbf{e}_{4,-10}^{20} \quad \mathbf{e}_{3,-8}^{20} \quad \mathbf{e}_{13,-16}^{20} \quad \mathbf{e}_{12,-14}^{20} \quad \mathbf{e}_{18,-19}^{20}]$$

and $\mathbf{e}_{i,-j}^{20} := \mathbf{e}_i^{20} - \mathbf{e}_j^{20}$. Let $\mathbf{G}$ be the 64×20 matrix defined by (2.5). We denote by $\mathcal{W}$ the $4 \times 16 \times 10$ tensor such that $\mathrm{Matr}(\mathcal{W}) = \mathbf{G}\overline{\mathbf{W}}$. We find algebraically the CPD $\mathcal{W} = [\mathbf{F}, \mathbf{F}_2, \mathbf{F}_3]_{10}$ with

$$\mathbf{F} = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

$\mathbf{F}_2 = \mathbf{F} \odot \mathbf{F}$ and some nonsingular matrix $\mathbf{F}_3$. In the sequel we will use only the fact that $\mathbf{F}$ coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling.

Phase 2 and 3 of Algorithm 1. *There are 210 4×6 submatrices of $\mathbf{F}$. In Phase 2 of Algorithm 1 we pick the five submatrices that have rank 3. One can easily see that these submatrices are $[\mathbf{f}_1\ \mathbf{f}_2\ \mathbf{f}_3\ \mathbf{f}_5\ \mathbf{f}_8\ \mathbf{f}_{10}]$, $[\mathbf{f}_1\ \mathbf{f}_2\ \mathbf{f}_4\ \mathbf{f}_6\ \mathbf{f}_9\ \mathbf{f}_{10}]$, $[\mathbf{f}_1\ \mathbf{f}_3\ \mathbf{f}_4\ \mathbf{f}_7\ \mathbf{f}_8\ \mathbf{f}_9]$, $[\ \mathbf{f}_2\ \mathbf{f}_4\ \mathbf{f}_5\ \mathbf{f}_6\ \mathbf{f}_7\ \mathbf{f}_8]$, $[\mathbf{f}_3\ \mathbf{f}_5\ \mathbf{f}_6\ \mathbf{f}_7\ \mathbf{f}_9\ \mathbf{f}_{10}]$. Their left kernels have dimension 1 and are spanned by the norm one vectors $\widehat{\mathbf{c}}_1 = [0\ 0\ 1\ 0]^T$, $\widehat{\mathbf{c}}_2 = [0.5\ 0.5\ 0.5\ 0.5]^T$, $\widehat{\mathbf{c}}_3 = [0\ 1\ 0\ 0]^T$, $\widehat{\mathbf{c}}_4 = [0\ 0\ 0\ 1]^T$, $\widehat{\mathbf{c}}_5 = [1\ 0\ 0\ 0]^T$, respectively. The matrix formed by these vectors coincides with the matrix $\mathbf{C}$ up to column permutation and scaling.*

Let us demonstrate how Phase 3 of Algorithm 1 works. One can easily see that the vectors $\mathbf{z}_1 = [-1\ 0\ 1\ 0]^T$ and $\mathbf{z}_2 = [2\ 0\ 0\ 0]^T$ coincide with the first two columns of the matrix $[\widehat{\mathbf{c}}_1\ \widehat{\mathbf{c}}_2\ \widehat{\mathbf{c}}_3\ \widehat{\mathbf{c}}_4]^{-T}$. We have

$$\mathrm{Matr}(\widetilde{\mathcal{T}}) = \mathrm{Matr}(\mathcal{T})[\mathbf{z}_1\ \mathbf{z}_2] = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T[\mathbf{z}_1\ \mathbf{z}_2] = (\mathbf{A} \odot \mathbf{B}) \begin{bmatrix} 0 & -1 & 0 & 3 & 0 \\ 2 & 2 & 0 & 0 & 0 \end{bmatrix}^T.$$

or $\widetilde{\mathcal{T}} = \left[[\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_4], [\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_4], \begin{bmatrix} 0 & -1 & 3 \\ 2 & 2 & 0 \end{bmatrix}\right]_3$. Thus, computing algebraically the CPD of $\widetilde{\mathcal{T}}$ we find the vectors $\mathbf{a}_1 \otimes \mathbf{b}_1$, $\mathbf{a}_2 \otimes \mathbf{b}_2$, and $\mathbf{a}_4 \otimes \mathbf{b}_4$. The vectors $\mathbf{a}_3 \otimes \mathbf{b}_3$ and $\mathbf{a}_5 \otimes \mathbf{b}_5$ are found by

$$[\mathbf{a}_3 \otimes \mathbf{b}_3\ \ \mathbf{a}_5 \otimes \mathbf{b}_5] = \big(\mathrm{Matr}(\mathcal{T}) - [\mathbf{a}_1 \otimes \mathbf{b}_1\ \ \mathbf{a}_2 \otimes \mathbf{b}_2\ \ \mathbf{a}_4 \otimes \mathbf{b}_4][\mathbf{c}_1\ \ \mathbf{c}_2\ \ \mathbf{c}_4]^T\big)\,[\mathbf{c}_3\ \ \mathbf{c}_5]^{\dagger,T}.$$

Phase 2 and 3 of Algorithm 2. *We construct the $4 \times 4 \times 10$ tensor $\mathcal{V}$ with matrix unfolding $\mathrm{Matr}(\mathcal{V}) = \mathrm{Matr}(\mathcal{T})\mathbf{F}$. Let $\mathbf{V}_1, \ldots, \mathbf{V}_{10}$ denote the frontal slices of $\mathcal{V}$ and let $\mathcal{V}_{ij}$ denote the $4 \times 4 \times 2$ tensor with frontal slices $\mathbf{V}_i$ and $\mathbf{V}_j$. We construct the set*

$$\begin{aligned}\mathcal{J} :=& \{(i,j) :\ \text{the matrices } [\mathbf{V}_i\ \mathbf{V}_j] \text{ and } [\mathbf{V}_i^T\ \mathbf{V}_j^T] \text{ have rank } 3,\ 1 \le i < j \le 10\} = \\ & \{(1,2),\ (1,3),\ (1,4),\ (1,8),\ (1,9),\ (1,10),\ (2,4),\ (2,5),\ (2,6),\ (2,8),\ (2,10),\\ & \ (3,5),\ (3,7),\ (3,8),\ (3,9),\ (3,10),\ (4,6),\ (4,7),\ (4,8),\ (4,9),\\ & \ (5,6),\ (5,7),\ (5,8),\ (5,10),\ (6,7),\ (6,9),\ (6,10),\ (7,8),\ (7,9),\ (9,10)\}.\end{aligned}$$

For $(i,j) \in \mathcal{J}$, $\mathcal{V}_{ij}$ *has rank* 3 *and the CPD can be computed algebraically. For instance,* $\mathrm{Matr}(\mathcal{V}_{12}) = \mathrm{Matr}(\mathcal{T})[\mathbf{f}_1\ \mathbf{f}_2] = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T[\mathbf{f}_1\ \mathbf{f}_2]$. *Since* $\mathbf{C}^T[\mathbf{f}_1\ \mathbf{f}_2] = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 \end{bmatrix}^T$, *we have* $\mathcal{V}_{12} = \left[[\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3], [\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3], \begin{bmatrix} 2 & 1 & 0 \\ 0 & -1 & 2 \end{bmatrix}\right]_3$. *In this way for each pair* $(i,j) \in \mathcal{J}$ *we estimate up to column scaling three columns of* $\mathbf{A}$ *and the corresponding columns of* $\mathbf{B}$. *If we store all the estimates of columns of* $\mathbf{A}$ *and* $\mathbf{B}$ *in* 4×90 *matrices* $\widetilde{\mathbf{A}}$ *and* $\widetilde{\mathbf{B}}$ *then* $\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}}$ *will contain* 5 *clusters of* 18 *collinear columns. Taking the cluster centers we get a matrix* $\mathbf{Z}$ *which coincides with* $\mathbf{A} \odot \mathbf{B}$ *up to column scaling and permutation. Finally, the matrix* $\left(\mathbf{Z}^{\dagger}\mathrm{Matr}(\mathcal{T})\right)^T$ *coincides with* $\mathbf{C}$ *up to column scaling and the same permutation.*

EXAMPLE 4.7. *It was shown in [9] that the conditions of Theorem 1.6 hold for a generic* $6 \times 6 \times 7$ *tensor of rank* 9. *This case is beyond Kruskal's bound. Let* $\mathbf{F}$ *be the* 7×84 *matrix produced by Phase 1 of Algorithms 1 and 2. Each column of the third factor matrix of the tensor is orthogonal to exactly* 42 *columns of the matrix* $\mathbf{F}$. *Since* C_{84}^{42} *is of order* 10^{24}, *Phase 2 as presented in Algorithm 1 is computationally infeasible. On the other hand, in Phase 2 of Algorithm 2 we check the rank of* $2C_{84}^2 = 3486$ *matrices of the size* 6×12 *each. Then we have to find algebraically the CPD of* $C_9^4 C_4^2 = 756$ *rank-*4 *tensors with dimensions* $6 \times 6 \times 2$, *which is equivalent with the computation of the GEVD of the associated matrix pencils. Moreover, one may further limit the amount of work by only determining subsets of* $\mathcal{J}$. *We implemented Algorithm 2 in MATLAB 2008a and we did experiments on a computer with Intel(R) Core(TM) T9600 Duo 2.80GHz CPU and 4GB memory running Windows Vista. The simulations demonstrate that with a suboptimal implementation, it takes less than* 9 *seconds to compute the CPD of a generic* $6 \times 6 \times 7$ *tensor of rank* 9.

5. Conclusion. We have proposed two algorithms to compute CPD. Both algorithms are algebraic in the sense that they rely only on standard linear algebra and reduce the problem to the computation of GEVD. The reduction exploits properties of (polarized) compound matrices and permanents. The derivation spans the possibilities from [19] to [6, 16] and covers cases beyond Kruskal's bound.

In this paper we have limited ourselves to exact CPD. In applications, CPD most often only approximates the given (noisy) tensor. A first observation is that the "exact result" could be used to initialize iterative algorithms for problem (1.5). We also note that (4.2) may be interpreted as the CPD of a partially symmetric tensor of order $m+1$ of which the first m factor matrices are equal and parameterized by $\mathbf{C}$. This is a structure that can be handled by current algorithms in Tensorlab [35]. These algorithms are optimization-based and are not formally guaranteed to find the solution. However, they show excellent performance in practice. So far, we have computed $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)})$ and then we have fitted the CPD structure to the result. Numerically, we could go a step further and take the Khatri-Rao structure into account in the computation of the kernel itself, with the kernel vectors parameterized by $\mathbf{C}$ and $\mathbf{M}$. One may also investigate whether the Khatri-Rao structure and the structure of $\mathbf{R}_m(\mathcal{T})$ may be exploited to avoid the computation of the mixed discriminants, so that one obtains an algorithm that works directly on $\mathcal{T}$. Since numerical aspects lead to a different type of study, we choose to defer them to an other paper.

REFERENCES

[1] A. D. ALEKSANDROV, *Zur Theorie der gemischten Volumina von konvexen Körpern.* IV. *Die gemischten Diskriminanten und die gemischten Volumina (Russian) (German summary)*,

Mat. Sb., 3 (45), no. 2 (1938), pp. 227–251.
[2] R.B. Bapat, *Mixed discriminants of positive semidefinite matrices*, Linear Algebra Appl., 126 (1989), pp. 107–124.
[3] J. Carroll and J.-J. Chang, *Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition*, Psychometrika, 35 (1970), pp. 283–319.
[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*, ESAT-STADIUS Internal Report, 13-235, Leuven, Belgium: Department of Electrical Engineering (ESAT), KU Leuven, (2013).
[5] P. Comon, X. Luciani, and A. L. F. de Almeida, *Tensor decompositions, alternating least squares and other tales*, J. Chemometrics, 23 (2009), pp. 393–405.
[6] L. De Lathauwer, *A Link Between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 642–666.
[7] ———, *A short introduction to tensor-based methods for factor analysis and blind source separation*, in ISPA 2011: Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis, (2011), pp. 558–563.
[8] I. Domanov and L. De Lathauwer, *On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors — Part I: Basic Results and Uniqueness of One Factor Matrix*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 855–875.
[9] ———, *On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors— Part II: Overall Uniqueness*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 876–903.
[10] G. P. Egorychev, *Proof of the van der Waerden conjecture for permanents*, Siberian Math. J., 22 (1981), pp. 854–859.
[11] R. A. Harshman, *Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis*, UCLA Working Papers in Phonetics, 16 (1970), pp. 1–84.
[12] ———, *Determination and Proof of Minimum Uniqueness Conditions for PARAFAC1*, UCLA Working Papers in Phonetics, 22 (1972), pp. 111–117.
[13] R. A. Harshman and M. E. Lundy, *Parafac: Parallel factor analysis*, Comput. Stat. Data Anal., (1994), pp. 39–72.
[14] F. L. Hitchcock, *The expression of a tensor or a polyadic as a sum of products*, J. Math. Phys., 6 (1927), pp. 164–189.
[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.
[16] T. Jiang and N. D. Sidiropoulos, *Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints*, IEEE Trans. Signal Process., 52 (2004), pp. 2625–2636.
[17] T. G. Kolda and B. W. Bader, *Tensor Decompositions and Applications*, SIAM Review, 51 (2009), pp. 455–500.
[18] P. M Kroonenberg, *Applied Multiway Data Analysis*, Hoboken, NJ: Wiley, 2008.
[19] J. B. Kruskal, *Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra Appl., 18 (1977), pp. 95–138.
[20] J. M. Landsberg, *Tensors: Geometry and Applications*, AMS, Providence, Rhode Island, 2012.
[21] S. E. Leurgans, R. T. Ross, and R. B. Abel, *A decomposition for three-way arrays*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1064–1083.
[22] J. S. Lomont and M. S. Cheena, *A multilinearity property of determinant functions*, Linear and Multilinear Algebra, 14 (1983), pp. 199–223.
[23] M. Marcus, *The Hadamard Theorem for Permanents*, Proc. Amer. Math. Soc., 15 (1964), pp. pp. 967–973.
[24] ———, *Finite dimensional multilinear algebra*, no. pt. 1 in Pure and applied mathematics, M. Dekker, 1973.
[25] M. Marcus and H. Minc, *On the relation between the determinant and the permanent*, Illinois J. Math., 5 (1961), pp. 376–381.
[26] J. Möcks, *Topographic components model for event-related potentials and some biophysical considerations*, IEEE Trans. Biomed. Eng., 35 (1988), pp. 482–484.
[27] H. Minc, *Permanents*, Encyclopedia of Mathematics and its Applications, Addison-Wesley Publishing Company, Advanced Book Program, 1978.
[28] T. Muir, *A treatise on the theory of determinants: with graduated sets of exercises for use in colleges and schools*, Macmillan and Co., 1882.
[29] D. Nion and L. De Lathauwer, *A Study of the Decomposition of a Third-Order Tensor in Rank-$(L, L, 1)$ Terms*, ESAT-STADIUS Internal Report, 11-239, Leuven, Belgium: De-

partment of Electrical Engineering (ESAT), KU Leuven, (2011).
[30] Luke Oeding and Giorgio Ottaviani, *Eigenvectors of tensors and algorithms for Waring decomposition*, Journal of Symbolic Computation, 54 (2013), pp. 9–35.
[31] E. Sanchez and B.. Kowalski, *Tensorial resolution: A direct trilinear decomposition*, J. Chemometrics, 4 (1990), pp. 29–45.
[32] R: Sands and F. Young, *Component models for three-way data: An alternating least squares algorithm with optimal scaling features*, Psychometrika, 45 (1980), pp. 39–67.
[33] James R. Schott, *Kronecker product permutation matrices and their application to moment matrices of the normal distribution*, J. Multivar. Anal., 87 (2003), pp. 177 – 190.
[34] A.K. Smilde, R. Bro, and P. Geladi, *Multi-way analysis with applications in the chemical sciences*, J. Wiley, 2004.
[35] L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab v2.0*, Available online, January 2014. URL: http://www.tensorlab.net/.
[36] ———, *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 (2013), pp. 695–720.
[37] M. Sørensen and L. De Lathauwer, *Blind Signal Separation via Tensor Decomposition with Vandermonde Factor Part I: Canonical Polyadic Decomposition*, IEEE Trans. Signal Process., 61 (2013), pp. 5507–5519.
[38] ———, *New Uniqueness Conditions for the Canonical Polyadic Decomposition of Higher-Order Tensors*, ESAT-STADIUS Internal Report, 11-239, Leuven, Belgium: Department of Electrical Engineering (ESAT), KU Leuven, (2013).
[39] M. Sørensen, L. De Lathauwer, P. Comon, S. Icart, and L. Deneire, *Canonical Polyadic Decomposition with a Columnwise Orthonormal Factor Matrix*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1190–1213.
[40] J. Ten Berge and J. Tendeiro, *The link between sufficient conditions by Harshman and by Kruskal for uniqueness in Candecomp/Parafac*, J. Chemometrics, 23 (2009), pp. 321–323.

CANONICAL POLYADIC DECOMPOSITION OF THIRD-ORDER TENSORS: REDUCTION TO GENERALIZED EIGENVALUE DECOMPOSITION

IGNAT DOMANOV AND LIEVEN DE LATHAUWER

supplementary materials

S.1. Supplementary material related to Proposition 1.10.

Recall that the $K \times C_R^{K-1}$ matrix $\mathcal{B}(\mathbf{C})$ is defined by

$$\mathcal{B}(\mathbf{C}) := \mathbf{L}\mathcal{C}_{K-1}(\mathbf{C}), \tag{S.1.1}$$

where

$$\mathbf{L} := \begin{bmatrix} 0 & 0 & \dots & (-1)^{K-1} \\ \vdots & \vdots & \iddots & \vdots \\ 0 & -1 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{bmatrix}. \tag{S.1.2}$$

We start with a trivial Lemma.

LEMMA S.1.1. *Let $\mathbf{x} \in \mathbb{R}^K$, $\mathbf{C} \in \mathbb{R}^{K\times R}$, $k_{\mathbf{C}} \geq K-1$, and let $\mathcal{B}(\mathbf{C})$ be defined by* (S.1.1). *Then*

(i)

$$\mathbf{x}^T\mathcal{B}(\mathbf{C}) = \left[\det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{x}\end{bmatrix} \quad \dots \quad \det\begin{bmatrix}\mathbf{c}_{R-K+2} & \dots & \mathbf{c}_R & \mathbf{x}\end{bmatrix}\right];$$

(ii) *$\mathcal{B}(\mathbf{C})$ has no zero columns;*

(iii) *the $(i_1,\dots,i_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$ is orthogonal to $\operatorname{span}\{\mathbf{c}_{i_1},\dots,\mathbf{c}_{i_{K-1}}\}$ for $(i_1,\dots,i_{K-1}) \in S_R^{K-1}$;*

(iv) *if $k_{\mathbf{C}} = K$, then $\mathcal{B}(\mathbf{C})$ has no proportional columns, that is $k_{\mathcal{B}(\mathbf{C})} \geq 2$.*

Proof. (i) From (S.1.1) it follows that the $(i_1,\dots,i_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$ is equal to $\mathbf{L}\mathcal{C}_{K-1}\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}}\end{bmatrix}$. By the Laplace expansion theorem

$$\mathbf{y}^T\mathbf{L}\mathcal{C}_{K-1}\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}}\end{bmatrix} = \det\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}} & \mathbf{y}\end{bmatrix}, \qquad \mathbf{y} \in \mathbb{R}^K. \tag{S.1.3}$$

Now, the statement (i) follows from (S.1.3) by setting $\mathbf{y} = \mathbf{x}$.

(ii) Since the vectors $\mathbf{c}_{i_1},\dots,\mathbf{c}_{i_{K-1}}$ are linearly independent in $\mathbb{R}^K$, it follows that there exists a vector $\mathbf{y}$ such that $\det\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}} & \mathbf{y}\end{bmatrix} \neq 0$. Hence, by (S.1.3), the $(i_1,\dots,i_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$ is nonzero.

(iii) follows from (S.1.3) and the fact that $\det\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}} & \mathbf{y}\end{bmatrix} = 0$ if and only if $\mathbf{y} \in \operatorname{span}\{\mathbf{c}_{i_1},\dots,\mathbf{c}_{i_{K-1}}\}$.

(iv) Assume that the $(i_1,\dots,i_{K-1})$-th and the $(j_1,\dots,j_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$ are proportional to a nonzero vector $\mathbf{t}$. Set $E := \operatorname{span}\{\mathbf{c}_{i_1},\dots,\mathbf{c}_{i_{K-1}},\mathbf{c}_{j_1},\dots,\mathbf{c}_{j_{K-1}}\}$. Since $k_{\mathbf{C}} = K$, it follows that $E = \mathbb{R}^K$. On the other hand, by (iii), the nonzero vector $\mathbf{t}$ is orthogonal to E, which is a contradiction. □

LEMMA S.1.2. *Let $\mathbf{C} \in \mathbb{R}^{K\times R}$, $K \leq R$, and $k_{\mathbf{C}} = K$. Let also $\mathcal{B}(\mathbf{C})$ be defined by* (S.1.1). *Then*

(i) *every column of the $R \times C_R^{K-1}$ matrix $\mathbf{C}^T\mathcal{B}(\mathbf{C})$ has exactly $K-1$ zero entries. Namely, if $\mathbf{d} = \begin{bmatrix}d_1 & \dots & d_R\end{bmatrix}$ is the $(j_1,\dots,j_{K-1})$-th column of $\mathbf{C}^T\mathcal{B}(\mathbf{C})$, then $d_r = 0$ if and only if $r \in \{j_1,\dots,j_{K-1}\}$;*

(ii) *every row of the* $R \times C_R^{K-1}$ *matrix* $\mathbf{C}^T\mathcal{B}(\mathbf{C})$ *has exactly* C_{R-1}^{K-2} *zero entries. Namely, if* $\widehat{\mathbf{d}} = \begin{bmatrix} d_{(1,\dots,K-1)} & \dots & d_{(R-K+2,\dots,R)} \end{bmatrix}$ *is the* r*-th row of* $\mathbf{C}^T\mathcal{B}(\mathbf{C})$*, then* $d_{(j_1,\dots,j_{K-1})} = 0$ *if and only if* $r \in \{j_1, \dots, j_{K-1}\}$*;*

(iii) *the matrix*

$$\mathcal{B}(\mathbf{C})^{(l)} := \underbrace{\mathcal{B}(\mathbf{C}) \odot \cdots \odot \mathcal{B}(\mathbf{C})}_{l}$$

has full column rank for $l \geq R - K + 1$*.*

Proof. (i)-(ii) By Lemma S.1.1 (i),

$$\mathbf{C}^T\mathcal{B}(\mathbf{C}) = \mathbf{C}^T\mathbf{L}\mathcal{C}_{K-1}(\mathbf{C}) = \begin{bmatrix} \det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{c}_1\end{bmatrix} & \dots & \det\begin{bmatrix}\mathbf{c}_{R-K+2} & \dots & \mathbf{c}_R & \mathbf{c}_1\end{bmatrix} \\ \vdots & & \vdots & & \vdots \\ \det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{c}_R\end{bmatrix} & \dots & \det\begin{bmatrix}\mathbf{c}_{R-K+2} & \dots & \mathbf{c}_R & \mathbf{c}_R\end{bmatrix} \end{bmatrix}. \tag{S.1.4}$$

Since $k_{\mathbf{C}} = K$, it follows that

$$\det\begin{bmatrix}\mathbf{c}_{j_1} & \dots & \mathbf{c}_{j_{K-1}} & \mathbf{c}_r\end{bmatrix} = 0 \Leftrightarrow r \in \{j_1, \dots, j_{K-1}\}. \tag{S.1.5}$$

The results now easily follow from (S.1.4)-(S.1.5).

(iii) It is sufficient to consider the case $l = R-K+1$. The result for $l \geq R-K+1$ then follows directly from the definition of the Khatri-Rao product.

Suppose that $\mathcal{B}(\mathbf{C})^{(R-K+1)}\widehat{\mathbf{t}} = \mathbf{0}$ for $\widehat{\mathbf{t}} = \begin{bmatrix} t_{(1,\dots,K-1)} & \dots & t_{(R-K+2,\dots,R)} \end{bmatrix} \in \mathbb{R}^{C_R^{K-1}}$. We show that $t_{(j_1,\dots,j_{K-1})} = 0$ for all $(j_1, \dots, j_{K-1}) \in S_R^{K-1}$. We fix $(j_1, \dots, j_{K-1}) \in S_R^{K-1}$ and set $\{i_1, \dots, i_{R-K+1}\} = \{1, \dots, R\} \setminus \{j_1, \dots, j_{K-1}\}$. Since $k_{\mathbf{C}} = K$, (S.1.5) holds. In particular,

$$\alpha_{(j_1,\dots,j_{K-1})} := \prod_{l=1}^{R-K+1} \det\begin{bmatrix}\mathbf{c}_{j_1} & \dots & \mathbf{c}_{j_{K-1}} & \mathbf{c}_{i_l}\end{bmatrix} \neq 0. \tag{S.1.6}$$

Let $\mathbf{f} := \mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_{R-K+1}}$. Then by (S.1.5), we have

$$0 = \mathbf{f}^T\mathbf{0} = \mathbf{f}^T\left(\mathcal{B}(\mathbf{C})^{(R-K+1)}\widehat{\mathbf{t}}\right) = \left(\mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_{R-K+1}}\right)^T\Big(\underbrace{\mathcal{B}(\mathbf{C}) \odot \cdots \odot \mathcal{B}(\mathbf{C})}_{R-K+1}\Big)\widehat{\mathbf{t}} =$$

$$\left[\prod_{l=1}^{R-K+1} \det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{c}_{i_l}\end{bmatrix} \quad \dots \quad \prod_{l=1}^{R-K+1} \det\begin{bmatrix}\mathbf{c}_{R-K+2} & \dots & \mathbf{c}_R & \mathbf{c}_{i_l}\end{bmatrix}\right]\widehat{\mathbf{t}} =$$

$$\left[0 \quad \dots \quad 0 \quad \prod_{l=1}^{R-K+1} \det\begin{bmatrix}\mathbf{c}_{j_1} & \dots & \mathbf{c}_{j_{K-1}} & \mathbf{c}_{i_l}\end{bmatrix} \quad 0 \quad \dots \quad 0\right]\widehat{\mathbf{t}} =$$

$$\alpha_{(j_1,\dots,j_{K-1})} \cdot t_{(j_1,\dots,j_{K-1})}.$$

Hence, by (S.1.6), $t_{(j_1,\dots,j_{K-1})} = 0$. Thus, $\widehat{\mathbf{t}} = \mathbf{0}$. Therefore, the matrix $\mathcal{B}(\mathbf{C})^{(R-K+1)}$ has full column rank. □

Proof of Proposition 1.10.

(i) follows from Lemma S.1.1 (iv).

(ii) follows from Lemma S.1.2 (iii). □

S.2. Supplementary material related to properties (P1)–(P4).
We will say that condition (K,R) holds if for any $K \times R$ matrix $\mathbf{C}$ with $k_{\mathbf{C}} = K$ and for any nonzero vector $\mathbf{x} \in \mathbb{R}^K$, the implication in the following scheme holds

$$\begin{aligned} \omega(\mathbf{x}^T \mathcal{B}(\mathbf{C})) \le C_R^{K-1} - C_{R-1}^{K-2} = C_{R-1}^{K-1} \Leftrightarrow \\ \mathbf{x} \text{ is orthogonal to at least } C_{R-1}^{K-2} \text{ columns of } B(\mathbf{C}) \Rightarrow \\ \mathbf{x} \text{ is proportional to a column of } \mathbf{C}. \end{aligned} \tag{S.2.1}$$

Note that the equivalence "$\Leftrightarrow$" and the implication opposite to "$\Rightarrow$" in (S.2.1) follow from the definition of $\omega(\cdot)$ and Lemma S.1.2 (ii), respectively. It can be easily checked that (2,R) holds for $R \ge 2$ and that (K,K) holds. Our goal is to show that (K,R) holds for $R \ge K$. We need the following lemma.

LEMMA S.2.1. *Suppose that both conditions* (K-1,R-1) *and* (K,R-1) *hold. Then condition* (K,R) *holds.*

Proof. Let $\mathbf{x} \in \mathbb{R}^K$ be a nonzero vector such that $\omega(\mathbf{x}^T \mathcal{B}(\mathbf{C})) \le C_{R-1}^{K-1}$ and let

$$\begin{aligned} \mathcal{J} &:= \{(j_1, \dots, j_{K-1}) : \ (j_1, \dots, j_{K-1}) \in S_R^{K-1}, \ \det\begin{bmatrix} \mathbf{c}_{j_1} & \dots & \mathbf{c}_{j_{K-1}} & \mathbf{x} \end{bmatrix} = 0\}, \\ \mathcal{J}_1 &:= \{(1, j_2, \dots, j_{K-1}) : \ (1, j_2, \dots, j_{K-1}) \in \mathcal{J}\}, \qquad \overline{\mathcal{J}}_1 := \mathcal{J} \setminus \mathcal{J}_1. \end{aligned}$$

Then by Lemma S.1.1 (i) and (S.2.1), card $\mathcal{J} \ge C_{R-1}^{K-2}$. We consider two cases: card $\overline{\mathcal{J}}_1 \ge C_{R-2}^{K-2}$ and card $\overline{\mathcal{J}}_1 < C_{R-2}^{K-2}$. If card $\overline{\mathcal{J}}_1 \ge C_{R-2}^{K-2}$, then

$$\omega(\mathbf{x}^T \mathcal{B}(\begin{bmatrix} \mathbf{c}_2 & \dots & \mathbf{c}_R \end{bmatrix})) \le C_{R-1}^{K-1} - \text{card } \overline{\mathcal{J}}_1 \le C_{R-1}^{K-1} - C_{R-2}^{K-2} = C_{R-2}^{K-1}.$$

Since condition (K,R-1) holds, it follows that either $\mathbf{x}$ is zero vector, or $\mathbf{x}$ is proportional to one of the vectors $\mathbf{c}_2, \dots, \mathbf{c}_R$. Let us consider the case card $\overline{\mathcal{J}}_1 < C_{R-2}^{K-2}$. We have

$$\text{card } \mathcal{J}_1 = \text{card } \mathcal{J} - \text{card } \overline{\mathcal{J}}_1 > C_{R-1}^{K-2} - C_{R-2}^{K-2} = C_{R-2}^{K-3}. \tag{S.2.2}$$

By Lemma S.1.1 (i) and (S.2.1), there exist numbers $\alpha_{p_1}, \alpha_{p_{j_2}}, \dots, \alpha_{p_{j_{K-1}}}$ such that

$$\mathbf{x} = \alpha_{p_1} \mathbf{c}_1 + \sum_{q=2}^{K-1} \alpha_{p_{j_q}} \mathbf{c}_{j_q}, \qquad p \in \{1, \dots, \text{card } \mathcal{J}_1\}, \qquad (1, j_2, \dots, j_{K-1}) \in \mathcal{J}_1. \tag{S.2.3}$$

Let $\mathbf{T} : \ \mathbb{R}^K \to \mathbb{R}^{K-1}$ be a linear mapping with $\ker(\mathbf{T}) = \text{span } \{\mathbf{c}_1\}$. We set

$$\widetilde{\mathbf{x}} := \mathbf{T}\mathbf{x} \in \mathbb{R}^{K-1}, \qquad \widetilde{\mathbf{C}} := \mathbf{T}\begin{bmatrix} \mathbf{c}_2 & \dots & \mathbf{c}_R \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{c}}_1 & \dots & \widetilde{\mathbf{c}}_{R-1} \end{bmatrix} \in \mathbb{R}^{(K-1)\times(R-1)}.$$

If $\mathbf{T}\mathbf{x} = \mathbf{0}$, then either $\mathbf{x} = \mathbf{0}$, or $\mathbf{x}$ is proportional to $\mathbf{c}_1$. Hence, we can assume that $\widetilde{\mathbf{x}}$ is a nonzero vector. Since $k_{\mathbf{C}} = K$ and $\ker(\mathbf{T}) = \text{span } \{\mathbf{c}_1\}$, it follows that $k_{\widetilde{\mathbf{C}}} = K - 1$. Let us apply $\mathbf{T}$ to (S.2.3)

$$\widetilde{\mathbf{x}} = \sum_{q=2}^{K-1} \alpha_{p_{j_q}} \mathbf{T}\mathbf{c}_{j_q} = \sum_{q=2}^{K-1} \alpha_{p_{j_q}} \widetilde{\mathbf{c}}_{j_q - 1}, \quad p \in \{1, \dots, \text{card } \mathcal{J}_1\}, \quad (1, j_2, \dots, j_{K-1}) \in \mathcal{J}_1.$$

Hence,

$$\det\begin{bmatrix} \widetilde{\mathbf{c}}_{j_2 - 1} & \dots & \widetilde{\mathbf{c}}_{j_{K-1} - 1} & \mathbf{x} \end{bmatrix} = 0, \quad (1, j_2, \dots, j_{K-1}) \in \mathcal{J}_1.$$

By Lemma S.1.1 (i) and (S.2.2),

$$\omega(\widetilde{\mathbf{x}}^T\mathcal{B}(\widetilde{\mathbf{C}})) \leq C_{R-1}^{K-2} - \text{card } \mathcal{J}_1 \leq C_{R-1}^{K-2} - C_{R-2}^{K-3} = C_{R-2}^{K-2}.$$

Since condition (K-1,R-1) holds and $\widetilde{\mathbf{x}} \neq \mathbf{0}$, it follows that $\widetilde{\mathbf{x}}$ is proportional to a column of $\widetilde{\mathbf{C}}$. Hence, $\mathbf{x}$ is proportional to one of the vectors $\mathbf{c}_2, \dots, \mathbf{c}_R$. ∎

LEMMA S.2.2. *Let $R \geq K$. Then condition* (K,R) *holds.*

Proof. The proof is by induction on $k = 2, \dots, K$. For $k = 2$ the result is trivial. Suppose that condition (k,R) holds for $R \geq k$. We prove that condition (k+1,R) holds for $R \geq k+1$. Since (k,k) holds for $k \geq 2$, by Lemma S.2.1, we have

$$\begin{aligned}
&\text{conditions (k,k+1) and (k+1,k+1) imply condition (k+1,k+2),}\\
&\text{conditions (k,k+2) and (k+1,k+2) imply condition (k+1,k+3),}\\
&\qquad\qquad\vdots\\
&\text{conditions (k,R-1) and (k+1,R-1) imply condition (k+1,R).}
\end{aligned}$$

∎

Proof of properties (P1)–(P4).

(P1) follows from Lemma S.1.2 (i).

(P2) follows from Lemma S.1.1 (iii).

(P3) follows from Lemma S.1.2 (ii).

(P4) We prove that

$$\begin{aligned}
&\mathbf{x} \text{ is orthogonal to } C_{R-1}^{K-2} \text{ columns of } B(\mathbf{C}) \Leftrightarrow\\
&\mathbf{x} \text{ is proportional to a column of } \mathbf{C}.
\end{aligned}$$

The result follows from Lemma S.1.2 (ii) and Lemma S.2.2, respectively.

∎

S.3. Supplementary material related to Lemma 2.17.

Proof of Lemma 2.17.

(i) From the definitions of the matrices $\mathcal{Q}_m(\mathbf{C})^T$ and $\mathcal{R}_m(\mathbf{C})^T$ (see Definitions 2.9 and 2.10) it follows that the matrix $\mathcal{Q}_m(\mathbf{C})^T$ is obtained from $\mathcal{R}_m(\mathbf{C})^T$ be removing columns that are repeated. Thus, (i) just expresses the following fact: $(j_1, \dots, j_m)$-th column of the matrix $\mathcal{Q}_m(\mathbf{C})^T$ coincides with the $(l_1, \dots, l_m)$-th column of the matrix $\mathcal{R}_m(\mathbf{C})^T$ whenever $(l_1, \dots, l_m) \in P_{\{j_1,\dots,j_m\}}$.

(ii) From Lemma 2.15 and definition of the matrix $\mathbf{G}$ (2.5) it follows that the $((i_1, \dots, i_m), (j_1, \dots, j_m))$-th entry of the matrix $\mathcal{R}_m(\mathbf{C})^T\mathbf{G}$ is equal to

$$\begin{aligned}
&m!\,(\pi_S(\mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_m}))^T \left(\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K)\right) =\\
&\frac{1}{m!} \sum_{(q_1,\dots,q_m)\in P_{\{i_1,\dots,i_m\}}} \sum_{(l_1,\dots,l_m)\in P_{\{j_1,\dots,jm\}}} (\mathbf{c}_{q_1}^T\mathbf{e}_{l_1}^K) \cdot \cdots \cdot (\mathbf{c}_{q_m}^T\mathbf{e}_{l_m}^K) =\\
&\frac{1}{m!} \sum_{(q_1,\dots,q_m)\in P_{\{i_1,\dots,i_m\}}} \text{perm } \mathbf{C}((j_1, \dots, j_m), (q_1, \dots, q_m)).
\end{aligned}$$

Since perm $\mathbf{C}((j_1, \dots, j_m), (q_1, \dots, q_m)) =$ perm $\mathbf{C}((j_1, \dots, j_m), (i_1, \dots, i_m))$ for all $(q_1, \dots, q_m) \in P_{\{i_1,\dots,i_m\}}$, it follows that

$$(\pi_S(\mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_m}))^T \left(\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K)\right) = \text{perm } \mathbf{C}((j_1, \dots, j_m), (i_1, \dots, i_m)).$$

The equality in (ii) follows now from Definition 2.9. ∎

S.4. Supplementary material related to Corollary 2.18.

Proof of Corollary 2.18.

By Proposition 2.13 (ii), $\dim\left(\ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}\right)\right) = C_R^{K-1}$. We prove that the equation $\mathcal{R}_m(\mathbf{C})^T\mathbf{x} = \mathbf{0}$ has at most $C_R^{K-1}-(K-1)$ solutions of the form $\mathbf{x} := \mathbf{y}\otimes\mathbf{z} \in \mathrm{range}(\pi_S)$. Denote by $\mathcal{X}$ the m-th order symmetric $K\times\cdots\times K$ tensor whose vectorized version coincides with $\mathbf{x}$. Since $\mathcal{X}$ is symmetric, all its $K \times K^{m-1}$ matrix unfoldings coincide with the rank-1 matrix $\mathbf{y}\mathbf{z}^T$ and hence are rank-1. It is well known that this is possible if and only if $\mathcal{X}$ is itself a rank-1 tensor. Hence, the vector $\mathbf{x}$ is proportional to the vector $\mathbf{y}\otimes\cdots\otimes\mathbf{y}$ and $\mathcal{R}_m(\mathbf{C})^T\mathbf{x} = \mathcal{R}_m(\mathbf{C})^T(\mathbf{y}\otimes\cdots\otimes\mathbf{y}) = \mathbf{0}$. Hence, by (2.9), $(\mathbf{c}_{p_1}^T\mathbf{y})\cdots(\mathbf{c}_{p_m}^T\mathbf{y}) = 0$ for $1 \le p_1 < \cdots < p_m \le R$. Consequently, $\mathbf{y}$ is orthogonal to at least $R-(m-1) = K-1$ columns of $\mathbf{C}$. Since $k_{\mathbf{C}} = K-1$, the orthogonal complement of these $K-1$ columns is one-dimensional and, hence, by Lemma S.1.1 (iii), $\mathbf{y}$ is proportional to a column of $\mathcal{B}(\mathbf{C})$. On the other hand, since $k_{\mathbf{C}} = K-1$, $\mathbf{C}$ has K columns that are linearly dependent, and, hence, it follows from Lemma S.1.1 (iii) that at least K columns of $\mathcal{B}(\mathbf{C})$ are pairwise proportional. Hence, there exist at most $C_R^{K-1}-(K-1)$ nonzero solutions of the equation $\mathcal{R}_m(\mathbf{C})^T(\mathbf{y}\otimes\cdots\otimes\mathbf{y}) = \mathbf{0}$, which completes the proof. ∎

S.5. Supplementary material related to Lemma 4.4.

We need the following Lemma.

LEMMA S.5.1. *Let $\widehat{\mathbf{A}}$ be any set of $\widehat{m}$ columns of $\mathbf{A}$, let $\widehat{\mathbf{B}}$ be the corresponding set of columns of $\mathbf{B}$. Assume that the matrix $\mathcal{C}_m(\mathbf{A})\odot\mathcal{C}_m(\mathbf{B})$ has full column rank. Then*

(i) $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m$.

(ii) $\max(r_{\widehat{\mathbf{A}}}, r_{\widehat{\mathbf{B}}}) \ge \min(\widehat{m}, m+1)$.

Proof. Since $\mathcal{C}_m(\mathbf{A})\odot\mathcal{C}_m(\mathbf{B})$ has full column rank it follows that all columns of $\mathcal{C}_m(\mathbf{A})$ and $\mathcal{C}_m(\mathbf{B})$ are nonzero. Hence, (i) follows from Lemma 2.3 (1). If $\widehat{m} \le m$, then (ii) follows from (i). If $\widehat{m} > m$, then, by Lemma 2.5, $r_{\mathcal{C}_m(\widehat{\mathbf{A}})} = C_{r_{\widehat{\mathbf{A}}}}^m$ and $r_{\mathcal{C}_m(\widehat{\mathbf{B}})} = C_{r_{\widehat{\mathbf{B}}}}^m$. Hence,

$$C_{r_{\widehat{\mathbf{A}}}}^m C_{r_{\widehat{\mathbf{B}}}}^m = r_{\mathcal{C}_m(\widehat{\mathbf{A}})} r_{\mathcal{C}_m(\widehat{\mathbf{B}})} = r_{\mathcal{C}_m(\widehat{\mathbf{A}})\otimes\mathcal{C}_m(\widehat{\mathbf{B}})} \ge r_{\mathcal{C}_m(\widehat{\mathbf{A}})\odot\mathcal{C}_m(\widehat{\mathbf{B}})} = C_{\widehat{m}}^m > 1, \tag{S.5.1}$$

where the last equality holds since the matrix $\mathcal{C}_m(\widehat{\mathbf{A}})\odot\mathcal{C}_m(\widehat{\mathbf{B}})$ has full column rank. The statement (ii) for $\widehat{m} > m$ now follows from (S.5.1). ∎

Proof of Lemma 4.4.

Without loss of generality we may assume that $\mathbf{F}$ coincides with $\mathcal{B}(\mathbf{C})$. In the proof we will associate indices $i, j \in \{1,\dots,C_R^{K-1}\}$ with their multi-index analogues $(i_1,\dots,i_{K-1}), (j_1,\dots,j_{K-1}) \in S_R^{K-1}$. By definition set

$$\{p_1,\dots,p_{\widehat{m}}\} = \{1,\dots,R\} \setminus \{\{i_1,\dots,i_{K-1}\} \cap \{j_1,\dots,j_{K-1}\}\}.$$

Since

$$\begin{gathered}\mathrm{card}\,\{\{i_1,\dots,i_{K-1}\} \cap \{j_1,\dots,j_{K-1}\}\} \le K-2,\\ \{p_1,\dots,p_{\widehat{m}}\} \subset \{\{1,\dots,R\} \setminus \{i_1,\dots,i_{K-1}\}\} \cup \{\{1,\dots,R\} \setminus \{j_1,\dots,j_{K-1}\}\}\end{gathered}$$

it follows that

$$m = R-(K-2) \le \widehat{m}, \tag{S.5.2}$$

$$\widehat{m} \le \min(R, 2(R-(K-1))) = \min(R, 2m-2). \tag{S.5.3}$$

We will show that the statements (i)–(iii) are all equivalent to the condition $\widehat{m} = m$.

(i) $\Leftrightarrow \widehat{m} = m$: follows from Lemma S.1.2 (i).

(ii) $\Leftrightarrow \widehat{m} = m$: let the vectors $\widehat{\mathbf{y}}_i$ and $\widehat{\mathbf{y}}_j$ (resp. the matrices $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{B}}$) be formed by the entries of the vectors $\mathbf{y}_i$ and $\mathbf{y}_j$ (resp. by the columns of the matrices $\mathbf{A}$ and $\mathbf{B}$) with indices $p_1, \dots, p_{\widehat{m}}$. Then

$$\mathbf{V}_i = \mathbf{A}\mathrm{Diag}(\mathbf{y}_i)\mathbf{B}^T = \widehat{\mathbf{A}}\mathrm{Diag}(\widehat{\mathbf{y}}_i)\widehat{\mathbf{B}}^T, \quad \mathbf{V}_j = \mathbf{A}\mathrm{Diag}(\mathbf{y}_j)\mathbf{B}^T = \widehat{\mathbf{A}}\mathrm{Diag}(\widehat{\mathbf{y}}_j)\widehat{\mathbf{B}}^T. \tag{S.5.4}$$

Hence,

$$[\mathbf{V}_i\ \mathbf{V}_j] = \widehat{\mathbf{A}}[\mathrm{Diag}(\widehat{\mathbf{y}}_i)\widehat{\mathbf{B}}^T\ \mathrm{Diag}(\widehat{\mathbf{y}}_j)\widehat{\mathbf{B}}^T] = \widehat{\mathbf{A}}(\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}})^T, \tag{S.5.5}$$

$$[\mathbf{V}_i^T\ \mathbf{V}_j^T] = \widehat{\mathbf{B}}[\mathrm{Diag}(\widehat{\mathbf{y}}_i)\widehat{\mathbf{A}}^T\ \mathrm{Diag}(\widehat{\mathbf{y}}_j)\widehat{\mathbf{A}}^T] = \widehat{\mathbf{B}}(\widehat{\mathbf{C}} \odot \widehat{\mathbf{A}})^T, \tag{S.5.6}$$

where $\widehat{\mathbf{C}} := [\widehat{\mathbf{y}}_i\ \widehat{\mathbf{y}}_j]^T$. We claim that

$$\text{the matrices } \ \widehat{\mathbf{C}} \odot \widehat{\mathbf{B}} \ \text{ and } \ \widehat{\mathbf{C}} \odot \widehat{\mathbf{A}} \ \text{ have full column rank.} \tag{S.5.7}$$

From the construction of the matrix $\widehat{\mathbf{C}}$ it follows that there exists an $\widehat{m} \times \widehat{m}$ permutation matrix $\mathbf{P}$ such that

$$(\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}})\mathbf{P} = (\widehat{\mathbf{C}}\mathbf{P}) \odot (\widehat{\mathbf{B}}\mathbf{P}) = \begin{bmatrix} * & \dots & * & * & \dots & * & 0 & \dots & 0 \\ \underbrace{0 & \dots & 0}_{\widehat{m}-m+1} & \underbrace{* & \dots & *}_{2m-\widehat{m}-2} & \underbrace{* & \dots & *}_{\widehat{m}-m+1} \end{bmatrix} \odot (\widehat{\mathbf{B}}\mathbf{P}) =$$

$$\begin{bmatrix} * \cdot \widetilde{\mathbf{b}}_1 & \dots & * \cdot \widetilde{\mathbf{b}}_{\widehat{m}-m+1} & * \cdot \widetilde{\mathbf{b}}_{\widehat{m}-m+2} & \dots & * \cdot \widetilde{\mathbf{b}}_{m-1} & 0 & \dots & 0 \\ 0 & \dots & 0 & * \cdot \widetilde{\mathbf{b}}_{\widehat{m}-m+2} & \dots & * \cdot \widetilde{\mathbf{b}}_{m-1} & * \cdot \widetilde{\mathbf{b}}_m & \dots & * \cdot \widetilde{\mathbf{b}}_{\widehat{m}} \end{bmatrix},$$

where $*$ denotes a nonzero value, $\widetilde{\mathbf{b}}_1, \dots, \widetilde{\mathbf{b}}_{\widetilde{m}}$ denote the columns of the matrix $\widehat{\mathbf{B}}\mathbf{P}$, and where we use the dimensionality constraints in (S.5.2)–(S.5.3). By Lemma S.5.1 (i), $k_{\widehat{\mathbf{B}}} \geq k_{\mathbf{B}} \geq m$. Hence, the matrix $(\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}})\mathbf{P}$ has full column rank. Since the matrix $\mathbf{P}$ is nonsingular, it follows that the matrix $\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}}$ also has full column rank. In a similar fashion one can prove that the matrix $\widehat{\mathbf{C}} \odot \widehat{\mathbf{A}}$ has full column rank. From (S.5.5)–(S.5.7) it follows that $r_{[\mathbf{V}_i\ \mathbf{V}_j]} = r_{\widehat{\mathbf{A}}}$ and $r_{[\mathbf{V}_i^T\ \mathbf{V}_j^T]} = r_{\widehat{\mathbf{B}}}$. Using Lemma S.5.1 we have

$$\widehat{m} \geq \max(r_{[\mathbf{V}_i\ \mathbf{V}_j]}, r_{[\mathbf{V}_i^T\ \mathbf{V}_j^T]}) = \max(r_{\widehat{\mathbf{A}}}, r_{\widehat{\mathbf{B}}}) \geq \min(\widehat{m}, m+1), \tag{S.5.8}$$

$$\min(r_{[\mathbf{V}_i\ \mathbf{V}_j]}, r_{[\mathbf{V}_i^T\ \mathbf{V}_j^T]}) = \min(r_{\widehat{\mathbf{A}}}, r_{\widehat{\mathbf{B}}}) \geq \min(k_{\widehat{\mathbf{A}}}, k_{\widehat{\mathbf{B}}}) \geq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m. \tag{S.5.9}$$

The equivalence (ii) $\Leftrightarrow \widehat{m} = m$ now easily follows from (S.5.8)–(S.5.9).

(iii) $\Leftrightarrow \widehat{m} = m$: by (S.5.4), $\mathcal{V}_{ij} = [\mathbf{A}, \mathbf{B}, [\mathbf{y}_i\ \mathbf{y}_j]^T]_R = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_{\widehat{m}}$. Hence, $\widehat{m} \geq r_{\mathcal{V}_{ij}}$. On the other hand, from (S.5.5)-(S.5.6) it follows that $r_{\mathcal{V}_{ij}} \geq \max(r_{[\mathbf{V}_i\ \mathbf{V}_j]}, r_{[\mathbf{V}_i^T\ \mathbf{V}_j^T]})$. Hence, by (S.5.8), $r_{\mathcal{V}_{ij}} \geq \min(\widehat{m}, m+1)$. The equivalence (iii) $\Leftrightarrow \widehat{m} = m$ now follows from the inequalities $\widehat{m} \geq r_{\mathcal{V}_{ij}} \geq \min(\widehat{m}, m+1)$. ∎

S.6. Supplementary material related to Example 4.6.
The full matrix $\mathbf{Q}_3(\mathcal{T})$ is given by

$$\mathbf{Q}_3(\mathcal{T}) = -\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 6 & 0 & 0 & 0 & 0 & 0 & 12 & 0 & 6 & 12 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & -4 & -4 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{3} & 0 & 0 & 0 & 0 & 0 & -6 & 0 & -6 & 3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{3} & 0 & 0 & 2 & 0 & 0 & 6 & 0 & 6 & 11 & 0 & 0 & 4 & 4 & 0 \\
0 & -6 & 0 & 0 & -6 & \mathbf{-3} & -3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & \mathbf{1} & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & \mathbf{4} & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 4 & 0 & \mathbf{0} & 2 & 0 & 2 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & \mathbf{1} & 0 & 2 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & \mathbf{1} & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 6 & 0 & 4 & 6 & \mathbf{3} & 11 & 0 & 2 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.$$