SIAM J. MATRIX ANAL. APPL.
Vol. 30, No. 3, pp. 1033–1066

DECOMPOSITIONS OF A HIGHER-ORDER TENSOR IN BLOCK TERMS—PART II: DEFINITIONS AND UNIQUENESS*

LIEVEN DE LATHAUWER†

Abstract. In this paper we introduce a new class of tensor decompositions. Intuitively, we decompose a given tensor block into blocks of smaller size, where the size is characterized by a set of mode-n ranks. We study different types of such decompositions. For each type we derive conditions under which essential uniqueness is guaranteed. The parallel factor decomposition and Tucker's decomposition can be considered as special cases in the new framework. The paper sheds new light on fundamental aspects of tensor algebra.

Key words. multilinear algebra, higher-order tensor, Tucker decomposition, canonical decomposition, parallel factors model

AMS subject classifications. 15A18, 15A69

DOI. 10.1137/070690729

1. Introduction.

The two main tensor generalizations of the matrix singular value decomposition (SVD) are, on one hand, the Tucker decomposition/higher-order singular value decomposition (HOSVD) [59, 60, 12, 13, 15] and, on the other hand, the canonical/parallel factor (CANDECOMP/PARAFAC) decomposition [7, 26]. These are connected with two different tensor generalizations of the concept of matrix rank. The Tucker decomposition/HOSVD is linked with the set of mode-n ranks, which generalize column rank, row rank, etc. CANDECOMP/PARAFAC has to do with rank in the meaning of the minimal number of rank-1 terms that are needed in an expansion of the matrix/tensor. In this paper we introduce a new class of tensor SVDs, which we call block term decompositions. These lead to a framework that unifies the Tucker decomposition/HOSVD and CANDECOMP/PARAFAC. Block term decompositions also provide a unifying view on tensor rank.

We study different types of block term decompositions. For each type, we derive sufficient conditions for essential uniqueness, i.e., uniqueness up to trivial indeterminacies. We derive two types of uniqueness conditions. The first type follows from a reasoning that involves invariant subspaces associated with the tensor. This type of conditions generalizes the result on CANDECOMP/PARAFAC uniqueness that is presented in [6, 40, 47, 48]. The second type generalizes Kruskal's condition for CANDECOMP/PARAFAC uniqueness, discussed in [38, 49, 54].

In the following subsection we explain our notation and introduce some basic definitions. In subsection 1.2 we recall the Tucker decomposition/HOSVD and also the CANDECOMP/PARAFAC decomposition and summarize some of their properties.

*Received by the editors May 7, 2007; accepted for publication (in revised form) by J. G. Nagy April 14, 2008; published electronically September 25, 2008. This research was supported by Research Council K.U.Leuven: GOA-Ambiorics, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1; F.W.O.: project G.0321.06 and Research Communities ICCoS, ANMMM, and MLDM; the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization," 2007–2011); and EU: ERNSI.

http://www.siam.org/journals/simax/30-3/69072.html

†Subfaculty Science and Technology, Katholieke Universiteit Leuven Campus Kortrijk, E. Sabbelaan 53, 8500 Kortrijk, Belgium (Lieven.DeLathauwer@kuleuven-kortrijk.be), and Department of Electrical Engineering (ESAT), Research Division SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium (Lieven.DeLathauwer@esat.kuleuven.be, http://homes.esat.kuleuven.be/~delathau/home.html).

In section 2 we define block term decompositions. We subsequently introduce decomposition in rank-$(L, L, 1)$ terms (subsection 2.1), decomposition in rank-(L, M, N) terms (subsection 2.2), and type-2 decomposition in rank-$(L, M, \cdot)$ terms (subsection 2.3). The uniqueness of these decompositions is studied in sections 4, 5, and 6, respectively. In the analysis we use some tools that have been introduced in [19]. These will briefly be recalled in section 3.

Several proofs of lemmas and theorems establishing Kruskal-type conditions for essential uniqueness of the new decompositions generalize results for PARAFAC presented in [54]. We stay quite close to the text of [54]. We recommend studying the proofs in [54] before reading this paper.

1.1. Notation and basic definitions.

1.1.1. Notation. We use $\mathbb{K}$ to denote $\mathbb{R}$ or $\mathbb{C}$ when the difference is not important. In this paper scalars are denoted by lowercase letters (a, b, $\ldots$), vectors are written in boldface lowercase ($\mathbf{a}$, $\mathbf{b}$, $\ldots$), matrices correspond to boldface capitals ($\mathbf{A}$, $\mathbf{B}$, $\ldots$), and tensors are written as calligraphic letters ($\mathcal{A}$, $\mathcal{B}$, $\ldots$). This notation is consistently used for lower-order parts of a given structure. For instance, the entry with row index i and column index j in a matrix $\mathbf{A}$, i.e., $(\mathbf{A})_{ij}$, is symbolized by a_{ij} (also $(\mathbf{a})_i = a_i$ and $(\mathcal{A})_{ijk} = a_{ijk}$). If no confusion is possible, the ith column vector of a matrix $\mathbf{A}$ is denoted as $\mathbf{a}_i$, i.e., $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots]$. Sometimes we will use the MATLAB colon notation to indicate submatrices of a given matrix or subtensors of a given tensor. Italic capitals are also used to denote index upper bounds (e.g., $i = 1, 2, \ldots, I$). The symbol $\otimes$ denotes the Kronecker product,

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \ldots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \ldots \\ \vdots & \vdots & \end{pmatrix}.$$

Let $\mathbf{A} = [\mathbf{A}_1\ \ldots\ \mathbf{A}_R]$ and $\mathbf{B} = [\mathbf{B}_1\ \ldots\ \mathbf{B}_R]$ be two partitioned matrices. Then the Khatri–Rao product is defined as the partitionwise Kronecker product and represented by $\odot$ [46]:

$$\mathbf{A} \odot \mathbf{B} = (\mathbf{A}_1 \otimes \mathbf{B}_1 \ldots \mathbf{A}_R \otimes \mathbf{B}_R)\,. \tag{1.1}$$

In recent years, the term "Khatri–Rao product" and the symbol $\odot$ have been used mainly in the case where $\mathbf{A}$ and $\mathbf{B}$ are partitioned into vectors. For clarity, we denote this particular, columnwise, Khatri–Rao product by $\odot_c$:

$$\mathbf{A} \odot_c \mathbf{B} = (\mathbf{a}_1 \otimes \mathbf{b}_1 \ldots \mathbf{a}_R \otimes \mathbf{b}_R)\,.$$

The column space of a matrix and its orthogonal complement will be denoted by $\mathrm{span}(\mathbf{A})$ and $\mathrm{null}(\mathbf{A})$. The rank of a matrix $\mathbf{A}$ will be denoted by $\mathrm{rank}(\mathbf{A})$ or $r_{\mathbf{A}}$. The superscripts $\cdot^T$, $\cdot^H$, and $\cdot^\dagger$ denote the transpose, complex conjugated transpose, and Moore–Penrose pseudoinverse, respectively. The operator $\mathrm{diag}(\cdot)$ stacks its scalar arguments in a square diagonal matrix. Analogously, $\mathrm{blockdiag}(\cdot)$ stacks its vector or matrix arguments in a block-diagonal matrix. For vectorization of a matrix $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots]$ we stick to the following convention: $\mathrm{vec}(\mathbf{A}) = [\mathbf{a}_1^T\ \mathbf{a}_2^T\ \ldots]^T$. The symbol δ_{ij} stands for the Kronecker delta, i.e., $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. The $(N \times N)$ identity matrix is represented by $\mathbf{I}_{N\times N}$. The $(I \times J)$ zero matrix is denoted by $\mathbf{0}_{I\times J}$. $\mathbf{1}_N$ is a column vector of all ones of length N. The zero tensor is denoted by $\mathcal{O}$.

1.1.2. Basic definitions.

Definition 1.1. *Consider* $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ *and* $\mathbf{A} \in \mathbb{K}^{J_1 \times I_1}$, $\mathbf{B} \in \mathbb{K}^{J_2 \times I_2}$, $\mathbf{C} \in \mathbb{K}^{J_3 \times I_3}$. *Then the* Tucker mode-1 product $\mathcal{T} \bullet_1 \mathbf{A}$, mode-2 product $\mathcal{T} \bullet_2 \mathbf{B}$, and mode-3 product $\mathcal{T} \bullet_3 \mathbf{C}$ *are defined by*

$$(\mathcal{T} \bullet_1 \mathbf{A})_{j_1 i_2 i_3} = \sum_{i_1=1}^{I_1} t_{i_1 i_2 i_3} a_{j_1 i_1} \qquad \forall j_1, i_2, i_3,$$

$$(\mathcal{T} \bullet_2 \mathbf{B})_{i_1 j_2 i_3} = \sum_{i_2=1}^{I_2} t_{i_1 i_2 i_3} b_{j_2 i_2} \qquad \forall i_1, j_2, i_3,$$

$$(\mathcal{T} \bullet_3 \mathbf{C})_{i_1 i_2 j_3} = \sum_{i_3=1}^{I_3} t_{i_1 i_2 i_3} c_{j_3 i_3} \qquad \forall i_1, i_2, j_3,$$

respectively [11].

In this paper we denote the Tucker mode-n product in the same way as in [10]; in the literature the symbol $\times_n$ is sometimes used [12, 13, 15].

Definition 1.2. *The* Frobenius norm *of a tensor* $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ *is defined as*

$$\|\mathcal{T}\| = \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} |t_{ijk}|^2 \right)^{\frac{1}{2}}.$$

Definition 1.3. *The outer product* $\mathcal{A} \circ \mathcal{B}$ *of a tensor* $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times \cdots \times I_P}$ *and a tensor* $\mathcal{B} \in \mathbb{K}^{J_1 \times J_2 \times \cdots \times J_Q}$ *is the tensor defined by*

$$(\mathcal{A} \circ \mathcal{B})_{i_1 i_2 \ldots i_P j_1 j_2 \ldots j_Q} = a_{i_1 i_2 \ldots i_P} b_{j_1 j_2 \ldots j_Q}$$

for all values of the indices.

For instance, the outer product $\mathcal{T}$ of three vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ is defined by $t_{ijk} = a_i b_j c_k$ for all values of the indices.

Definition 1.4. *A* mode-n vector *of a tensor* $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ *is an* I_n*-dimensional vector obtained from* $\mathcal{T}$ *by varying the index* i_n *and keeping the other indices fixed* [34].

Mode-n vectors generalize column and row vectors.

Definition 1.5. *The* mode-n rank *of a tensor* $\mathcal{T}$ *is the dimension of the subspace spanned by its mode-n vectors.*

The mode-n rank of a higher-order tensor is the obvious generalization of the column (row) rank of a matrix.

Definition 1.6. *A third-order tensor is* rank-(L, M, N) *if its mode-*1 *rank, mode-*2 *rank, and mode-*3 *rank are equal to* L, M, *and* N, *respectively.*

A rank-$(1, 1, 1)$ tensor is briefly called rank-1. This definition is equivalent to the following.

Definition 1.7. *A third-order tensor* $\mathcal{T}$ *has* rank 1 *if it equals the outer product of* 3 *vectors.*

The rank (as opposed to mode-n rank) is now defined as follows.

Definition 1.8. *The* rank *of a tensor* $\mathcal{T}$ *is the minimal number of rank-*1 *tensors that yield* $\mathcal{T}$ *in a linear combination* [38].

The following definition has proved useful in the analysis of PARAFAC uniqueness [38, 49, 51, 54].

Definition 1.9. *The* Kruskal rank *or* k-rank *of a matrix* $\mathbf{A}$*, denoted by* $\text{rank}_k(\mathbf{A})$ *or* $k_{\mathbf{A}}$*, is the maximal number* r *such that any set of* r *columns of* $\mathbf{A}$ *is linearly independent* [38].

We call a property generic when it holds with probability one when the parameters of the problem are drawn from continuous probability density functions. Let $\mathbf{A} \in \mathbb{K}^{I\times R}$. Generically, we have $k_{\mathbf{A}} = \min(I, R)$.

It will sometimes be useful to express tensor properties in terms of matrices and vectors. We therefore define standard matrix representations of a third-order tensor.

Definition 1.10. *The standard* $(JK \times I)$ *matrix representation* $(\mathcal{T})_{JK\times I} = \mathbf{T}_{JK\times I}$*,* $(KI \times J)$ *representation* $(\mathcal{T})_{KI\times J} = \mathbf{T}_{KI\times J}$*, and* $(IJ \times K)$ *representation* $(\mathcal{T})_{IJ\times K} = \mathbf{T}_{IJ\times K}$ *of a tensor* $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ *are defined by*

$$
\begin{aligned}
(\mathbf{T}_{JK\times I})_{(j-1)K+k,i} &= (\mathcal{T})_{ijk},\\
(\mathbf{T}_{KI\times J})_{(k-1)I+i,j} &= (\mathcal{T})_{ijk},\\
(\mathbf{T}_{IJ\times K})_{(i-1)J+j,k} &= (\mathcal{T})_{ijk}
\end{aligned}
$$

for all values of the indices [34].

Note that in these definitions indices to the right vary more rapidly than indices to the left. Further, the ith $(J \times K)$ matrix slice of $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ will be denoted as $\mathbf{T}_{J\times K,i}$. Similarly, the jth $(K \times I)$ slice and the kth $(I \times J)$ slice will be denoted by $\mathbf{T}_{K\times I,j}$ and $\mathbf{T}_{I\times J,k}$, respectively.

1.2. HOSVD and PARAFAC. We have now enough material to introduce the Tucker/HOSVD [12, 13, 15, 59, 60] and CANDECOMP/PARAFAC [7, 26] decompositions.

Definition 1.11. *A* Tucker decomposition *of a tensor* $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ *is a decomposition of* $\mathcal{T}$ *of the form*

$$\mathcal{T} = \mathcal{D} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C}. \tag{1.2}$$

An HOSVD is a Tucker decomposition, normalized in a particular way. The normalization was suggested in the computational strategy in [59, 60].

Definition 1.12. *An* HOSVD *of a tensor* $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ *is a decomposition of* $\mathcal{T}$ *of the form*

$$\mathcal{T} = \mathcal{D} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C}, \tag{1.3}$$

in which

- *the matrices* $\mathbf{A} \in \mathbb{K}^{I\times L}$*,* $\mathbf{B} \in \mathbb{K}^{J\times M}$*, and* $\mathbf{C} \in \mathbb{K}^{K\times N}$ *are columnwise orthonormal,*
- *the core tensor* $\mathcal{D} \in \mathbb{K}^{L\times M\times N}$ *is*
  - *all-orthogonal,*

$$
\begin{aligned}
\langle \mathbf{D}_{M\times N,l_1}, \mathbf{D}_{M\times N,l_2}\rangle &= \text{trace}(\mathbf{D}_{M\times N,l_1} \cdot \mathbf{D}^H_{M\times N,l_2}) = \sigma^{(1)^2}_{l_1} \delta_{l_1,l_2},\\
&\qquad 1 \leqslant l_1, l_2 \leqslant L,\\
\langle \mathbf{D}_{N\times L,m_1}, \mathbf{D}_{N\times L,m_2}\rangle &= \text{trace}(\mathbf{D}_{N\times L,m_1} \cdot \mathbf{D}^H_{N\times L,m_2}) = \sigma^{(2)^2}_{m_1} \delta_{m_1,m_2},\\
&\qquad 1 \leqslant m_1, m_2 \leqslant M,\\
\langle \mathbf{D}_{I\times J,n_1}, \mathbf{D}_{I\times J,n_2}\rangle &= \text{trace}(\mathbf{D}_{L\times M,n_1} \cdot \mathbf{D}^H_{L\times M,n_2}) = \sigma^{(3)^2}_{n_1} \delta_{n_1,n_2},\\
&\qquad 1 \leqslant n_1, n_2 \leqslant N,
\end{aligned}
$$

– *ordered*,

$$\sigma_1^{(1)^2} \geqslant \sigma_2^{(1)^2} \geqslant \ldots \geqslant \sigma_L^{(1)^2} \geqslant 0,$$
$$\sigma_1^{(2)^2} \geqslant \sigma_2^{(2)^2} \geqslant \ldots \geqslant \sigma_M^{(2)^2} \geqslant 0,$$
$$\sigma_1^{(3)^2} \geqslant \sigma_2^{(3)^2} \geqslant \ldots \geqslant \sigma_N^{(3)^2} \geqslant 0.$$

The decomposition is visualized in Figure 1.1.

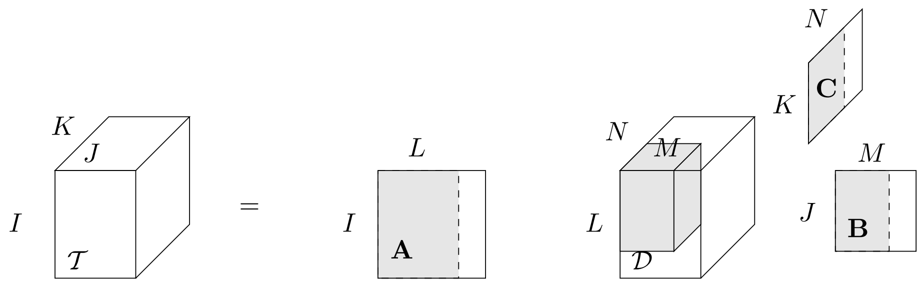

Fig. 1.1. *Visualization of the HOSVD/Tucker decomposition.*

Equation (1.3) can be written in terms of the standard $(JK \times I)$, $(KI \times J)$, and $(IJ \times K)$ matrix representations of $\mathcal{T}$ as follows:

$$\mathbf{T}_{JK\times I} = (\mathbf{B} \otimes \mathbf{C}) \cdot \mathbf{D}_{MN\times L} \cdot \mathbf{A}^T, \tag{1.4}$$
$$\mathbf{T}_{KI\times J} = (\mathbf{C} \otimes \mathbf{A}) \cdot \mathbf{D}_{NL\times M} \cdot \mathbf{B}^T, \tag{1.5}$$
$$\mathbf{T}_{IJ\times K} = (\mathbf{A} \otimes \mathbf{B}) \cdot \mathbf{D}_{LM\times N} \cdot \mathbf{C}^T. \tag{1.6}$$

The HOSVD exists for any $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$. The values L, M, and N correspond to the rank of $\mathbf{T}_{JK\times I}$, $\mathbf{T}_{KI\times J}$, and $\mathbf{T}_{IJ\times K}$, i.e., they are equal to the mode-1, mode-2 and mode-3 rank of $\mathcal{T}$, respectively. In [12] it has been demonstrated that the SVD of matrices and the HOSVD of higher-order tensors have some analogous properties.

Define $\tilde{\mathcal{D}} = \mathcal{D} \bullet_3 \mathbf{C}$. Then

$$\mathcal{T} = \tilde{\mathcal{D}} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \tag{1.7}$$

is a (normalized) *Tucker*-2 *decomposition* of $\mathcal{T}$. This decomposition is visualized in Figure 1.2.

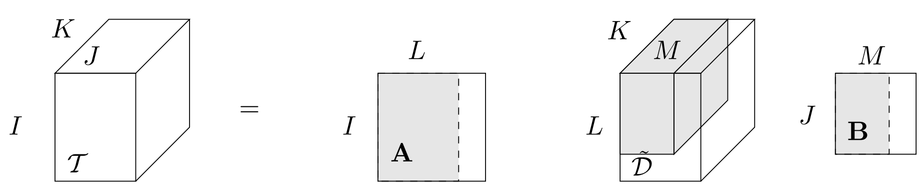

Fig. 1.2. *Visualization of the (normalized) Tucker-*2 *decomposition.*

Besides the HOSVD, there exist other ways to generalize the SVD of matrices. The most well known is CANDECOMP/PARAFAC [7, 26].

Definition 1.13. *A* canonical or parallel factor decomposition (CANDECOMP/ PARAFAC) *of a tensor $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ is a decomposition of $\mathcal{T}$ as a linear combination* *of rank-*1 *terms*:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r. \tag{1.8}$$

The decomposition is visualized in Figure 1.3.

In terms of the standard matrix representations of $\mathcal{T}$, decomposition (1.8) can be written as

$$\mathbf{T}_{JK\times I} = (\mathbf{B} \odot_c \mathbf{C}) \cdot \mathbf{A}^T, \tag{1.9}$$

$$\mathbf{T}_{KI\times J} = (\mathbf{C} \odot_c \mathbf{A}) \cdot \mathbf{B}^T, \tag{1.10}$$

$$\mathbf{T}_{IJ\times K} = (\mathbf{A} \odot_c \mathbf{B}) \cdot \mathbf{C}^T. \tag{1.11}$$

In terms of the $(J \times K)$, $(K \times I)$, and $(I \times J)$ matrix slices of $\mathcal{T}$, we have

$$\mathbf{T}_{J\times K,i} = \mathbf{B} \cdot \mathrm{diag}(a_{i1}, \dots, a_{iR}) \cdot \mathbf{C}^T, \quad i = 1, \dots, I. \tag{1.12}$$

$$\mathbf{T}_{K\times I,j} = \mathbf{C} \cdot \mathrm{diag}(b_{j1}, \dots, b_{jR}) \cdot \mathbf{A}^T, \quad j = 1, \dots, J. \tag{1.13}$$

$$\mathbf{T}_{I\times J,k} = \mathbf{A} \cdot \mathrm{diag}(c_{k1}, \dots, c_{kR}) \cdot \mathbf{B}^T, \quad k = 1, \dots, K. \tag{1.14}$$

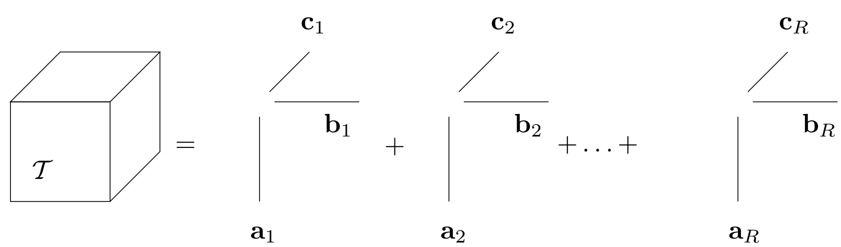

FIG. 1.3. *Visualization of the CANDECOMP/PARAFAC decomposition.*

The fully symmetric variant of PARAFAC, in which $\mathbf{a}_r = \mathbf{b}_r = \mathbf{c}_r$, $r = 1, \dots, R$, was studied in the nineteenth century in the context of invariant theory [9]. The unsymmetric decomposition was introduced by F. L. Hitchcock in 1927 [27, 28]. Around 1970, the unsymmetric decomposition was independently reintroduced in psychometrics [7] and phonetics [26]. Later, the decomposition was applied in chemometrics and the food industry [1, 5, 53]. In these various disciplines PARAFAC is used for the purpose of multiway factor analysis. The term "canonical decomposition" is standard in psychometrics, while in chemometrics the decomposition is called a parallel factors model. PARAFAC has found important applications in signal processing and data analysis [37]. In wireless telecommunications, it provides powerful means for the exploitation of different types of diversity [49, 50, 18]. It also describes the basic structure of higher-order cumulants of multivariate data on which all algebraic methods for independent component analysis (ICA) are based [8, 14, 29]. Moreover, the decomposition is finding its way to scientific computing, where it leads to a way around the curse of dimensionality [2, 3, 24, 25, 33].

To a large extent, the practical importance of PARAFAC stems from its uniqueness properties. It is clear that one can arbitrarily permute the different rank-1 terms. Also, the factors of a same rank-1 term may be arbitrarily scaled, as long as their product remains the same. We call a PARAFAC decomposition essentially unique when it is subject only to these trivial indeterminacies. The following theorem establishes a condition under which essential uniqueness is guaranteed.

THEOREM 1.14. *The PARAFAC decomposition* (1.8) *is essentially unique if*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geqslant 2R + 2. \tag{1.15}$$

This theorem was first proved for real tensors in [38]. A concise proof that also applies to complex tensors was given in [49]; in this proof, the permutation lemma of [38] was used. The result was generalized to tensors of arbitrary order in [51]. An alternative proof of the permutation lemma was given in [31]. The overall proof was reformulated in terms of accessible basic linear algebra in [54]. In [17] we derived a more relaxed uniqueness condition that applies when $\mathcal{T}$ is tall in one mode (meaning that, for instance, $K \geqslant R$).

2. Block term decompositions.

2.1. Decomposition in rank-$(L, L, 1)$ terms.

DEFINITION 2.1. *A* decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ in a sum of rank-$(L, L, 1)$ terms *is a decomposition of* $\mathcal{T}$ *of the form*

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_r, \tag{2.1}$$

in which the $(I \times J)$ *matrices* $\mathbf{E}_r$ *are rank-*L.

We also consider the decomposition of a tensor in a sum of matrix-vector outer products, in which the different matrices do not necessarily all have the same rank.

DEFINITION 2.2. *A* decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ in a sum of rank-$(L_r, L_r, 1)$ terms, $1 \leqslant r \leqslant R$, *is a decomposition of* $\mathcal{T}$ *of the form*

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_r, \tag{2.2}$$

in which the $(I \times J)$ *matrix* $\mathbf{E}_r$ *is rank-*L_r, $1 \leqslant r \leqslant R$.

If we factorize $\mathbf{E}_r$ as $\mathbf{A}_r \cdot \mathbf{B}_r^T$, in which the matrix $\mathbf{A}_r \in \mathbb{K}^{I \times L_r}$ and the matrix $\mathbf{B}_r \in \mathbb{K}^{J \times L_r}$ are rank-L_r, $r = 1, \ldots, R$, then we can write (2.2) as

$$\mathcal{T} = \sum_{r=1}^{R} (\mathbf{A}_r \cdot \mathbf{B}_r^T) \circ \mathbf{c}_r. \tag{2.3}$$

Define $\mathbf{A} = [\mathbf{A}_1 \ldots \mathbf{A}_R]$, $\mathbf{B} = [\mathbf{B}_1 \ldots \mathbf{B}_R]$, $\mathbf{C} = [\mathbf{c}_1 \ldots \mathbf{c}_R]$. In terms of the standard matrix representations of $\mathcal{T}$, (2.3) can be written as

$$\mathbf{T}_{IJ \times K} = [(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_{L_1} \ \ldots \ (\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_{L_R}] \cdot \mathbf{C}^T, \tag{2.4}$$

$$\mathbf{T}_{JK \times I} = (\mathbf{B} \odot \mathbf{C}) \cdot \mathbf{A}^T, \tag{2.5}$$

$$\mathbf{T}_{KI \times J} = (\mathbf{C} \odot \mathbf{A}) \cdot \mathbf{B}^T. \tag{2.6}$$

In terms of the matrix slices of $\mathcal{T}$, (2.3) can be written as

$$\mathbf{T}_{J \times K, i} = \mathbf{B} \cdot \operatorname{blockdiag}([(\mathbf{A}_1)_{i1} \ \ldots \ (\mathbf{A}_1)_{iL_1}]^T, \ldots, [(\mathbf{A}_R)_{i1} \ \ldots \ (\mathbf{A}_R)_{iL_R}]^T) \cdot \mathbf{C}^T, \quad i = 1, \ldots, I, \tag{2.7}$$

$$\mathbf{T}_{K \times I, j} = \mathbf{C} \cdot \operatorname{blockdiag}([(\mathbf{B}_1)_{j1} \ \ldots \ (\mathbf{B}_1)_{jL_1}], \ldots, [(\mathbf{B}_R)_{j1} \ \ldots \ (\mathbf{B}_R)_{jL_R}]) \cdot \mathbf{A}^T, \quad j = 1, \ldots, J, \tag{2.8}$$

$$\mathbf{T}_{I \times J, k} = \mathbf{A} \cdot \operatorname{blockdiag}(c_{k1}\mathbf{I}_{L_1 \times L_1}, \ldots, c_{kR}\mathbf{I}_{L_R \times L_R}) \cdot \mathbf{B}^T, \quad k = 1, \ldots, K. \tag{2.9}$$

It is clear that in (2.3) one can arbitrarily permute the different rank-$(L_r, L_r, 1)$ terms. Also, one can postmultiply $\mathbf{A}_r$ by any nonsingular $(L_r \times L_r)$ matrix $\mathbf{F}_r \in \mathbb{K}^{L_r \times L_r}$, provided $\mathbf{B}_r$ is premultiplied by the inverse of $\mathbf{F}_r$. Moreover, the factors of a same rank-$(L_r, L_r, 1)$ term may be arbitrarily scaled, as long as their product remains the same. We call the decomposition essentially unique when it is subject only to these trivial indeterminacies. Two representations $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ that are the same up to trivial indeterminacies are called essentially equal. We (partially) normalize the representation of (2.2) as follows. Scale/counterscale the vectors $\mathbf{c}_r$ and the matrices $\mathbf{E}_r$ such that $\mathbf{c}_r$ are unit-norm. Further, let $\mathbf{E}_r = \mathbf{A}_r \cdot \mathbf{D}_r \cdot \mathbf{B}_r^T$ denote the SVD of $\mathbf{E}_r$. The diagonal matrix $\mathbf{D}_r$ can be interpreted as an $(L_r \times L_r \times 1)$ tensor. Then (2.2) is equivalent to

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{c}_r. \tag{2.10}$$

Note that in this equation each term is represented in HOSVD form. The decomposition is visualized in Figure 2.1.

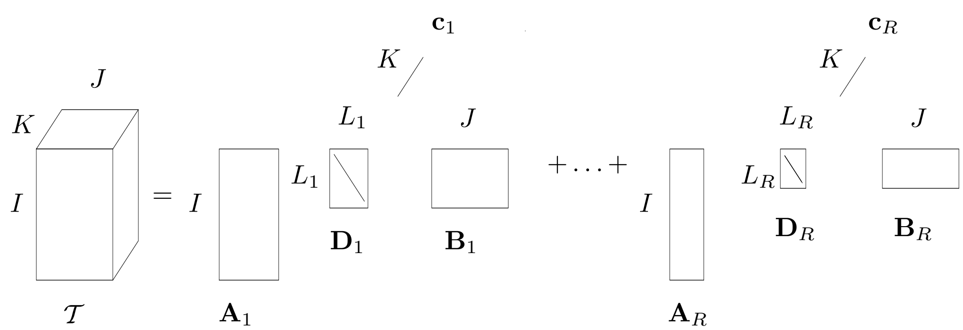

FIG. 2.1. *Visualization of the decomposition of a tensor in a sum of rank-$(L_r, L_r, 1)$ terms, $1 \leqslant r \leqslant R$.*

2.2. Decomposition in rank-(L, M, N) terms.

DEFINITION 2.3. *A decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ in a sum of rank-(L, M, N) terms is a decomposition of $\mathcal{T}$ of the form*

$$\mathcal{T} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r, \tag{2.11}$$

in which $\mathcal{D}_r \in \mathbb{K}^{L \times M \times N}$ are full rank-(L, M, N) and in which $\mathbf{A}_r \in \mathbb{K}^{I \times L}$ (with $I \geqslant L$), $\mathbf{B}_r \in \mathbb{K}^{J \times M}$ (with $J \geqslant M$), and $\mathbf{C}_r \in \mathbb{K}^{K \times N}$ (with $K \geqslant N$) are full column rank, $1 \leqslant r \leqslant R$.

Remark 1. One could also consider a decomposition in rank-(L_r, M_r, N_r) terms, where the different terms possibly have different mode-n ranks. In this paper we focus on the decomposition in rank-(L, M, N) terms.

Define partitioned matrices $\mathbf{A} = [\mathbf{A}_1 \dots \mathbf{A}_R]$, $\mathbf{B} = [\mathbf{B}_1 \dots \mathbf{B}_R]$, and $\mathbf{C} = [\mathbf{C}_1 \dots \mathbf{C}_R]$. In terms of the standard matrix representations of $\mathcal{T}$, (2.11) can be written as

$$\mathbf{T}_{JK \times I} = (\mathbf{B} \odot \mathbf{C}) \cdot \text{blockdiag}((\mathcal{D}_1)_{MN \times L}, \dots, (\mathcal{D}_R)_{MN \times L}) \cdot \mathbf{A}^T, \tag{2.12}$$

$$\mathbf{T}_{KI \times J} = (\mathbf{C} \odot \mathbf{A}) \cdot \text{blockdiag}((\mathcal{D}_1)_{NL \times M}, \dots, (\mathcal{D}_R)_{NL \times M}) \cdot \mathbf{B}^T, \tag{2.13}$$

$$\mathbf{T}_{IJ \times K} = (\mathbf{A} \odot \mathbf{B}) \cdot \text{blockdiag}((\mathcal{D}_1)_{LM \times N}, \dots, (\mathcal{D}_R)_{LM \times N}) \cdot \mathbf{C}^T. \tag{2.14}$$

It is clear that in (2.11) one can arbitrarily permute the different terms. Also, one can postmultiply $\mathbf{A}_r$ by a nonsingular matrix $\mathbf{F}_r \in \mathbb{K}^{L\times L}$, $\mathbf{B}_r$ by a nonsingular matrix $\mathbf{G}_r \in \mathbb{K}^{M\times M}$, and $\mathbf{C}_r$ by a nonsingular matrix $\mathbf{H}_r \in \mathbb{K}^{N\times N}$, provided $\mathcal{D}_r$ is replaced by $\mathcal{D}_r \bullet_1 \mathbf{F}_r^{-1} \bullet_2 \mathbf{G}_r^{-1} \bullet_3 \mathbf{H}_r^{-1}$. We call the decomposition essentially unique when it is subject only to these trivial indeterminacies. We can (partially) normalize (2.11) by representing each term by its HOSVD. The decomposition is visualized in Figure 2.2.

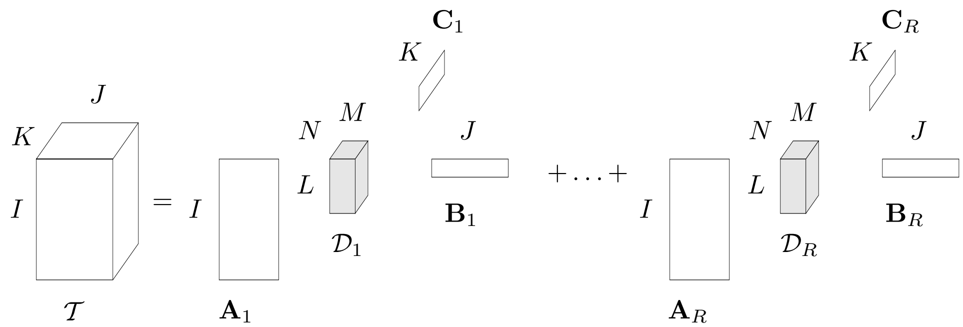

FIG. 2.2. *Visualization of the decomposition of a tensor in a sum of rank-(L,M,N) terms.*

Define $\mathcal{D} = \text{blockdiag}(\mathcal{D}_1,\dots,\mathcal{D}_R)$. Equation (2.11) can now also be seen as the multiplication of a block-diagonal core tensor $\mathcal{D}$ by means of factor matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$:

$$\mathcal{T} = \mathcal{D} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C}. \tag{2.15}$$

This alternative interpretation of the decomposition is visualized in Figure 2.3. Two representations $(\mathbf{A},\mathbf{B},\mathbf{C},\mathcal{D})$ and $(\bar{\mathbf{A}},\bar{\mathbf{B}},\bar{\mathbf{C}},\bar{\mathcal{D}})$ that are the same up to trivial indeterminacies are called essentially equal.

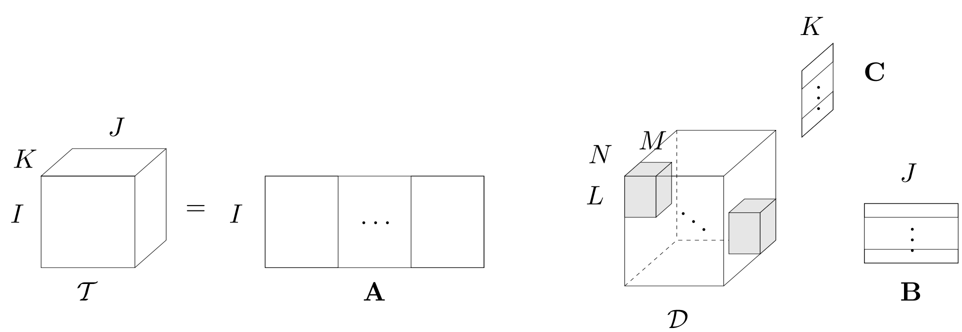

FIG. 2.3. *Interpretation of decomposition* (2.11) *in terms of the multiplication of a block-diagonal core tensor $\mathcal{D}$ by transformation matrices* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{C}$.

2.3. Type-2 decomposition in rank-$(L,M,\cdot)$ terms.

DEFINITION 2.4. *A* type-2 decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ in a sum of rank-$(L,M,\cdot)$ terms *is a decomposition of $\mathcal{T}$ of the form*

$$\mathcal{T} = \sum_{r=1}^{R} \mathcal{C}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r, \tag{2.16}$$

in which $\mathcal{C}_r \in \mathbb{K}^{L\times M\times K}$ *(with mode-*1 *rank equal to* L *and mode-*2 *rank equal to* M*) and in which* $\mathbf{A}_r \in \mathbb{K}^{I\times L}$ *(with* $I \geqslant L$*) and* $\mathbf{B}_r \in \mathbb{K}^{J\times M}$ *(with* $J \geqslant M$*) are full column rank,* $1 \leqslant r \leqslant R$*.*

Remark 2. The label "type 2" is reminiscent of the term "Tucker-2 decomposition."

Remark 3. One could also consider a type-2 decomposition in rank-$(L_r, M_r, \cdot)$ terms, where the different terms possibly have different mode-1 and/or mode-2 rank. In this paper we focus on the decomposition in rank-$(L, M, \cdot)$ terms.

Define partitioned matrices $\mathbf{A} = [\mathbf{A}_1 \dots \mathbf{A}_R]$ and $\mathbf{B} = [\mathbf{B}_1 \dots \mathbf{B}_R]$. In terms of the standard matrix representations of $\mathcal{T}$, (2.16) can be written as

$$\mathbf{T}_{IJ\times K} = (\mathbf{A} \odot \mathbf{B}) \cdot \begin{pmatrix} (\mathcal{C}_1)_{(LM\times K)} \\ \vdots \\ (\mathcal{C}_R)_{(LM\times K)} \end{pmatrix}, \tag{2.17}$$

$$\mathbf{T}_{JK\times I} = [(\mathcal{C}_1 \bullet_2 \mathbf{B}_1)_{JK\times L} \;\dots\; (\mathcal{C}_R \bullet_2 \mathbf{B}_R)_{JK\times L}] \cdot \mathbf{A}^T, \tag{2.18}$$

$$\mathbf{T}_{KI\times J} = [(\mathcal{C}_1 \bullet_1 \mathbf{A}_1)_{KI\times M} \;\dots\; (\mathcal{C}_R \bullet_1 \mathbf{A}_R)_{KI\times M}] \cdot \mathbf{B}^T. \tag{2.19}$$

Define $\mathcal{C} \in \mathbb{K}^{LR\times MR\times K}$ as an all-zero tensor, except for the entries given by

$$(\mathcal{C})_{(r-1)L+l,(r-1)M+m,k} = (\mathcal{C}_r)_{lmk} \quad \forall l, m, k, r.$$

Then (2.16) can also be written as

$$\mathcal{T} = \mathcal{C} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B}.$$

It is clear that in (2.16) one can arbitrarily permute the different terms. Also, one can postmultiply $\mathbf{A}_r$ by a nonsingular matrix $\mathbf{F}_r \in \mathbb{K}^{L\times L}$ and postmultiply $\mathbf{B}_r$ by a nonsingular matrix $\mathbf{G}_r \in \mathbb{K}^{M\times M}$, provided $\mathcal{C}_r$ is replaced by $\mathcal{C}_r \bullet_1 (\mathbf{F}_r)^{-1} \bullet_2 (\mathbf{G}_r)^{-1}$. We call the decomposition essentially unique when it is subject only to these trivial indeterminacies. Two representations $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathcal{C}})$ that are the same up to trivial indeterminacies are called essentially equal. We can (partially) normalize (2.16) by representing each term by its normalized Tucker-2 decomposition. The decomposition is visualized in Figure 2.4.

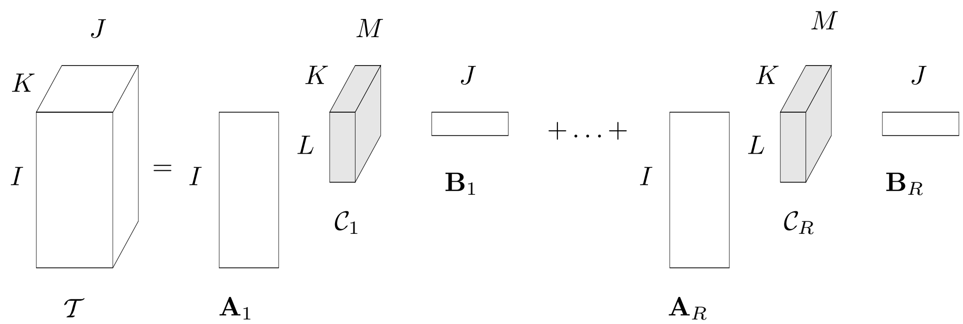

FIG. 2.4. *Visualization of the type-*2 *decomposition of a tensor in a sum of rank-*$(L, M, \cdot)$ *terms.*

3. Basic lemmas. In this section we list a number of lemmas that we will use in the analysis of the uniqueness of the block term decompositions.

Let $\omega(\mathbf{x})$ denote the number of nonzero entries of a vector $\mathbf{x}$. The following lemma was originally proposed by Kruskal in [38]. It is known as the *permutation lemma.*

It plays a crucial role in the proof of (1.15). The proof was reformulated in terms of accessible basic linear algebra in [54]. An alternative proof was given in [31]. The link between the two proofs is also discussed in [54].

LEMMA 3.1 (permutation lemma). *Consider two matrices* $\bar{\mathbf{A}}$, $\mathbf{A} \in \mathbb{K}^{I\times R}$, *that have no zero columns. If for every vector* $\mathbf{x}$ *such that* $\omega(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant R - r_{\bar{\mathbf{A}}} + 1$, *we have* $\omega(\mathbf{x}^T\mathbf{A}) \leqslant \omega(\mathbf{x}^T\bar{\mathbf{A}})$, *then there exists a unique permutation matrix* $\mathbf{\Pi}$ *and a unique nonsingular diagonal matrix* $\mathbf{\Lambda}$ *such that* $\bar{\mathbf{A}} = \mathbf{A}\cdot\mathbf{\Pi}\cdot\mathbf{\Lambda}$.

In [19] we have introduced a generalization of the permutation lemma to partitioned matrices. Let us first introduce some additional prerequisites. Let $\omega'(\mathbf{x})$ denote the number of parts of a partitioned vector $\mathbf{x}$ that are not all-zero. We call the partitioning of a partitioned matrix $\mathbf{A}$ uniform when all submatrices are of the same size. Finally, we generalize the k-rank concept to partitioned matrices [19].

DEFINITION 3.2. *The* k'-rank *of a (not necessarily uniformly) partitioned matrix* $\mathbf{A}$, *denoted by* $\mathrm{rank}_{k'}(\mathbf{A})$ *or* $k'_{\mathbf{A}}$, *is the maximal number* r *such that any set of* r *submatrices of* $\mathbf{A}$ *yields a set of linearly independent columns.*

Let $\mathbf{A} \in \mathbb{K}^{I\times LR}$ be uniformly partitioned in R matrices $\mathbf{A}_r \in \mathbb{K}^{I\times L}$. Generically, we have $k'_{\mathbf{A}} = \min(\lfloor \frac{I}{L} \rfloor, R)$.

We are now in a position to formulate the lemma that generalizes the permutation lemma.

LEMMA 3.3 (equivalence lemma for partitioned matrices). *Consider* $\bar{\mathbf{A}}$, $\mathbf{A} \in \mathbb{K}^{I\times \sum_{r=1}^{R} L_r}$, *partitioned in the same but not necessarily uniform way into* R *submatrices that are full column rank. Suppose that for every* $\mu \leqslant R - k'_{\bar{\mathbf{A}}} + 1$ *there holds that for a generic*[1] *vector* $\mathbf{x}$ *such that* $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant \mu$, *we have* $\omega'(\mathbf{x}^T\mathbf{A}) \leqslant \omega'(\mathbf{x}^T\bar{\mathbf{A}})$. *Then there exists a unique block-permutation matrix* $\mathbf{\Pi}$ *and a unique nonsingular block-diagonal matrix* $\mathbf{\Lambda}$, *such that* $\bar{\mathbf{A}} = \mathbf{A}\cdot\mathbf{\Pi}\cdot\mathbf{\Lambda}$, *where the block-transformation is compatible with the block-structure of* $\mathbf{A}$ *and* $\bar{\mathbf{A}}$.

(Compared to the presentation in [19] we have dropped the irrelevant complex conjugation of $\mathbf{x}$.)

We note that the rank $r_{\bar{\mathbf{A}}}$ in the permutation lemma has been replaced by the k'-rank $k'_{\bar{\mathbf{A}}}$ in Lemma 3.3. The reason is that the permutation lemma admits a simpler proof when we can assume that $r_{\bar{\mathbf{A}}} = k_{\bar{\mathbf{A}}}$. It is this simpler proof, given in [31], that is generalized in [19].

The following lemma gives a lower-bound on the k'-rank of a Khatri–Rao product of partitioned matrices [19].

LEMMA 3.4. *Consider partitioned matrices* $\mathbf{A} = [\mathbf{A}_1 \ \dots \ \mathbf{A}_R]$ *with* $\mathbf{A}_r \in \mathbb{K}^{I\times L_r}$, $1 \leqslant r \leqslant R$, *and* $\mathbf{B} = [\mathbf{B}_1 \ \dots \ \mathbf{B}_R]$ *with* $\mathbf{B}_r \in \mathbb{K}^{J\times M_r}$, $1 \leqslant r \leqslant R$.

(i) *If* $k'_{\mathbf{A}} = 0$ *or* $k'_{\mathbf{B}} = 0$, *then* $k'_{\mathbf{A}\odot\mathbf{B}} = 0$.

(ii) *If* $k'_{\mathbf{A}} \geqslant 1$ *and* $k'_{\mathbf{B}} \geqslant 1$, *then* $k'_{\mathbf{A}\odot\mathbf{B}} \geqslant \min(k'_{\mathbf{A}} + k'_{\mathbf{B}} - 1, R)$.

Finally, we have a lemma that says that a Khatri–Rao product of partitioned matrices is generically full column rank [19].

---

[1] We mean the following. Consider, for instance, a partitioned matrix $\bar{\mathbf{A}} = [\mathbf{a}_1 \ \mathbf{a}_2 | \mathbf{a}_3 \ \mathbf{a}_4] \in \mathbb{K}^{4\times 4}$ that is full column rank. The set $\mathsf{S} = \{\mathbf{x} | \omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant 1\}$ is the union of two subspaces, S_1 and S_2, consisting of the set of vectors orthogonal to $\{\mathbf{a}_1, \mathbf{a}_2\}$ and $\{\mathbf{a}_3, \mathbf{a}_4\}$, respectively. When we say that for a generic vector $\mathbf{x}$ such that $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant 1$, we have $\omega'(\mathbf{x}^T\mathbf{A}) \leqslant \omega'(\mathbf{x}^T\bar{\mathbf{A}})$, we mean that $\omega'(\mathbf{x}^T\mathbf{A}) \leqslant \omega'(\mathbf{x}^T\bar{\mathbf{A}})$ holds with probability one for a vector $\mathbf{x}$ drawn from a continuous probability density function over S_1 and that $\omega'(\mathbf{x}^T\mathbf{A}) \leqslant \omega'(\mathbf{x}^T\bar{\mathbf{A}})$ also holds with probability one for a vector $\mathbf{x}$ drawn from a continuous probability density function over S_2. In general, the set $\mathsf{S} = \{\mathbf{x} | \omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant \mu\}$ consists of a finite union of subspaces, where we count only the subspaces that are not contained in an other subspace. For each of these subspaces, the property should hold with probability one for a vector $\mathbf{x}$ drawn from a continuous probability density function over that subspace.

LEMMA 3.5. *Consider partitioned matrices* $\mathbf{A} = [\mathbf{A}_1 \ \dots \ \mathbf{A}_R]$ *with* $\mathbf{A}_r \in \mathbb{K}^{I \times L_r}$, $1 \leqslant r \leqslant R$, *and* $\mathbf{B} = [\mathbf{B}_1 \ \dots \ \mathbf{B}_R]$ *with* $\mathbf{B}_r \in \mathbb{K}^{J \times M_r}$, $1 \leqslant r \leqslant R$. *Generically we have that* $\operatorname{rank}(\mathbf{A} \odot \mathbf{B}) = \min(IJ, \sum_{r=1}^{R} L_r M_r)$.

4. The decomposition in rank-$(L_r, L_r, 1)$ terms. In this section we derive several conditions under which essential uniqueness of the decomposition in rank-$(L, L, 1)$ or rank-$(L_r, L_r, 1)$ terms is guaranteed. We use the notation introduced in section 2.1.

For decompositions in generic rank-$(L, L, 1)$ terms, the results of this section can be summarized as follows. We have essential uniqueness if

(i) Theorem 4.1:

$$\min(I, J) \geqslant LR \quad \text{and} \quad \mathbf{C} \text{ does not have proportional columns;} \tag{4.1}$$

(ii) Theorem 4.4:

$$K \geqslant R \quad \text{and} \quad \min\left(\left\lfloor \frac{I}{L} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{J}{L} \right\rfloor, R\right) \geqslant R + 2; \tag{4.2}$$

(iii) Theorem 4.5:

$$I \geqslant LR \quad \text{and} \quad \min\left(\left\lfloor \frac{J}{L} \right\rfloor, R\right) + \min(K, R) \geqslant R + 2 \tag{4.3}$$

or

$$J \geqslant LR \quad \text{and} \quad \min\left(\left\lfloor \frac{I}{L} \right\rfloor, R\right) + \min(K, R) \geqslant R + 2; \tag{4.4}$$

(iv) Theorem 4.7:

$$\left\lfloor \frac{IJ}{L^2} \right\rfloor \geqslant R \quad \text{and} \quad \min\left(\left\lfloor \frac{I}{L} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{J}{L} \right\rfloor, R\right) + \min(K, R) \geqslant 2R + 2. \tag{4.5}$$

First we mention a result of which the first version appeared, in a slightly different form, in [52]. The proof describes a procedure by which, under the given conditions, the components of the decomposition may be computed. This procedure is a generalization of the computation of PARAFAC from the generalized eigenvectors of the pencil $(\mathbf{T}^T_{I \times J,1}, \mathbf{T}^T_{I \times J,2})$, as explained in [20, section 1.4].

THEOREM 4.1. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *represent a decomposition of* $\mathcal{T}$ *in rank-*$(L_r, L_r, 1)$ *terms,* $1 \leqslant r \leqslant R$. *Suppose that* $\mathbf{A}$ *and* $\mathbf{B}$ *are full column rank and that* $\mathbf{C}$ *does not have proportional columns. Then* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *is essentially unique.*

Proof. Assume that $c_{21}, \dots, c_{2R}$ are different from zero and that $c_{11}/c_{21}, \dots, c_{1R}/c_{2R}$ are mutually different. (If this is not the case, consider linear combinations of matrix slices in the reasoning below.) From (2.9) we have

$$\mathbf{T}_{I \times J,1} = \mathbf{A} \cdot \operatorname{blockdiag}(c_{11}\mathbf{I}_{L_1 \times L_1}, \dots, c_{1R}\mathbf{I}_{L_R \times L_R}) \cdot \mathbf{B}^T, \tag{4.6}$$

$$\mathbf{T}_{I \times J,2} = \mathbf{A} \cdot \operatorname{blockdiag}(c_{21}\mathbf{I}_{L_1 \times L_1}, \dots, c_{2R}\mathbf{I}_{L_R \times L_R}) \cdot \mathbf{B}^T. \tag{4.7}$$

This means that the columns of $(\mathbf{A}^T)^\dagger$ are generalized eigenvectors of the pencil $(\mathbf{T}^T_{I \times J,1}, \mathbf{T}^T_{I \times J,2})$ [4, 22]. The columns of the rth submatrix of $\mathbf{A}$ are associated with the same generalized eigenvalue c_{1r}/c_{2r} and can therefore not be separated, $1 \leqslant r \leqslant R$. This is consistent with the indeterminacies of the decomposition. On the other hand, the different submatrices of $\mathbf{A}$ can be separated, as they correspond to different generalized eigenvalues. After computation of a possible matrix $\mathbf{A}$, the corresponding matrix $\mathbf{B}$ can be computed, up to scaling of its submatrices, from (4.7):

$$(\mathbf{A}^\dagger \cdot \mathbf{T}_{I\times J,2})^T = \mathbf{B} \cdot \mathrm{blockdiag}(c_{21}\mathbf{I}_{L_1\times L_1}, \dots, c_{2R}\mathbf{I}_{L_R\times L_R}).$$

Matrix $\mathbf{C}$ finally follows from (2.4):

$$\mathbf{C} = \left\{ [(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_{L_1} \ \dots \ (\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_{L_R}]^\dagger \cdot \mathbf{T}_{IJ\times K} \right\}^T . \qquad \square$$

Next, we derive generalizations of Kruskal's condition (1.15) under which essential uniqueness of $\mathbf{A}$, or $\mathbf{B}$, or $\mathbf{C}$ is guaranteed. Lemma 4.2 concerns essential uniqueness of $\mathbf{C}$. In its proof, we assume that the partitioning of $\mathbf{A}$ and $\mathbf{B}$ is uniform. Hence, the lemma applies only to the decomposition in rank-$(L, L, 1)$ terms. Lemma 4.3 concerns essential uniqueness of $\mathbf{A}$ and/or $\mathbf{B}$. This lemma applies more generally to the decomposition in rank-$(L_r, L_r, 1)$ terms. Later in this section, essential uniqueness of the decomposition of $\mathcal{T}$ will be inferred from essential uniqueness of one or more of the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$.

LEMMA 4.2. *Let $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ represent a decomposition of $\mathcal{T}$ in R rank-$(L, L, 1)$ terms. Suppose the condition*

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k_{\mathbf{C}} \geqslant 2R + 2 \tag{4.8}$$

holds and that we have an alternative decomposition of $\mathcal{T}$, represented by $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$. Then there holds $\bar{\mathbf{C}} = \mathbf{C} \cdot \mathbf{\Pi}_c \cdot \mathbf{\Lambda}_c$, in which $\mathbf{\Pi}_c$ is a permutation matrix and $\mathbf{\Lambda}_c$ a nonsingular diagonal matrix.

Proof. We work in analogy with [54]. Equality of $\mathbf{C}$ and $\bar{\mathbf{C}}$, up to column permutation and scaling, follows from the permutation lemma if we can prove that for any $\mathbf{x}$ such that $\omega(\mathbf{x}^T\bar{\mathbf{C}}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$, there holds $\omega(\mathbf{x}^T\mathbf{C}) \leqslant \omega(\mathbf{x}^T\bar{\mathbf{C}})$. This proof is structured as follows. First, we derive an upper-bound on $\omega(\mathbf{x}^T\bar{\mathbf{C}})$. Then we derive a lower-bound on $\omega(\mathbf{x}^T\bar{\mathbf{C}})$. Combination of the two bounds yields the desired result.

(i) *Derivation of an upper-bound on $\omega(\mathbf{x}^T\bar{\mathbf{C}})$.* From (2.9) we have that $\mathrm{vec}(\mathbf{T}^T_{I\times J,k}) = [(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_L \ \dots \ (\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_L] \cdot [c_{k1} \dots c_{kR}]^T$. Consider the linear combination of $(I \times J)$ slices $\sum_{k=1}^K x_k \mathbf{T}_{I\times J,k}$. Since $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ both represent a decomposition of $\mathcal{T}$, we have

$$\begin{aligned}&[(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_L \ \dots \ (\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_L] \cdot \mathbf{C}^T\mathbf{x} \\ &\quad = [(\bar{\mathbf{A}}_1 \odot_c \bar{\mathbf{B}}_1)\mathbf{1}_L \ \dots \ (\bar{\mathbf{A}}_R \odot_c \bar{\mathbf{B}}_R)\mathbf{1}_L] \cdot \bar{\mathbf{C}}^T\mathbf{x}.\end{aligned}$$

By Lemma 3.4, the matrix $\mathbf{A} \odot \mathbf{B}$ has full column rank. The matrix $[(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_L \ \dots$ $(\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_L]$ is equal to $(\mathbf{A} \odot \mathbf{B}) \cdot [\mathrm{vec}(\mathbf{I}_{L\times L})^T \ \dots \ \mathrm{vec}(\mathbf{I}_{L\times L})^T]^T$ and thus also has full column rank. This implies that if $\omega(\mathbf{x}^T\bar{\mathbf{C}}) = 0$, then also $\omega(\mathbf{x}^T\mathbf{C}) = 0$. Hence, $\mathrm{null}(\bar{\mathbf{C}}) \subseteq \mathrm{null}(\mathbf{C})$. Basic matrix algebra yields $\mathrm{span}(\mathbf{C}) \subseteq \mathrm{span}(\bar{\mathbf{C}})$ and $r_{\mathbf{C}} \leqslant r_{\bar{\mathbf{C}}}$. This implies that if $\omega(\mathbf{x}^T\bar{\mathbf{C}}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$, then

$$\omega(\mathbf{x}^T\bar{\mathbf{C}}) \leqslant R - r_{\bar{\mathbf{C}}} + 1 \leqslant R - r_{\mathbf{C}} + 1 \leqslant R - k_{\mathbf{C}} + 1 \leqslant k'_{\mathbf{A}} + k'_{\mathbf{B}} - (R+1), \tag{4.9}$$

where the last inequality corresponds to condition (4.8).

(ii) *Derivation of a lower-bound on $\omega(\mathbf{x}^T\bar{\mathbf{C}})$.* By (2.9), the linear combination of $(I \times J)$ slices $\sum_{k=1}^K x_k \mathbf{T}_{I\times J,k}$ is given by

$$\begin{aligned}&\mathbf{A} \cdot \mathrm{blockdiag}(\mathbf{x}^T\mathbf{c}_1\mathbf{I}_{L\times L}, \dots, \mathbf{x}^T\mathbf{c}_R\mathbf{I}_{L\times L}) \cdot \mathbf{B}^T \\ &\quad = \bar{\mathbf{A}} \cdot \mathrm{blockdiag}(\mathbf{x}^T\bar{\mathbf{c}}_1\mathbf{I}_{L\times L}, \dots, \mathbf{x}^T\bar{\mathbf{c}}_R\mathbf{I}_{L\times L}) \cdot \bar{\mathbf{B}}^T.\end{aligned}$$

We have

$$
\begin{aligned}
L\omega(\mathbf{x}^T\bar{\mathbf{C}}) &= r_{\text{blockdiag}(\mathbf{x}^T\bar{\mathbf{c}}_1\mathbf{I}_{L\times L},\ldots,\mathbf{x}^T\bar{\mathbf{c}}_R\mathbf{I}_{L\times L})} \\
&\geqslant r_{\bar{\mathbf{A}}\cdot\text{blockdiag}(\mathbf{x}^T\bar{\mathbf{c}}_1\mathbf{I}_{L\times L},\ldots,\mathbf{x}^T\bar{\mathbf{c}}_R\mathbf{I}_{L\times L})\cdot\bar{\mathbf{B}}^T} \\
&= r_{\mathbf{A}\cdot\text{blockdiag}(\mathbf{x}^T\mathbf{c}_1\mathbf{I}_{L\times L},\ldots,\mathbf{x}^T\mathbf{c}_R\mathbf{I}_{L\times L})\cdot\mathbf{B}^T}. \qquad (4.10)
\end{aligned}
$$

Let $\gamma = \omega(\mathbf{x}^T\mathbf{C})$ and let $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ consist of the submatrices of $\mathbf{A}$ and $\mathbf{B}$, respectively, corresponding to the nonzero elements of $\mathbf{x}^T\mathbf{C}$. Then $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ both have γL columns. Let $\mathbf{u}$ be the $(\gamma \times 1)$ vector containing the nonzero elements of $\mathbf{x}^T\mathbf{C}$ such that

$$
\mathbf{A}\cdot\text{blockdiag}(\mathbf{x}^T\mathbf{c}_1\mathbf{I}_{L\times L},\ldots,\mathbf{x}^T\mathbf{c}_R\mathbf{I}_{L\times L})\cdot\mathbf{B}^T = \tilde{\mathbf{A}}\cdot\text{blockdiag}(u_1\mathbf{I}_{L\times L},\ldots,u_\gamma\mathbf{I}_{L\times L})\cdot\tilde{\mathbf{B}}^T.
$$

Sylvester's inequality now yields

$$
\begin{aligned}
r_{\mathbf{A}\cdot\text{blockdiag}(\mathbf{x}^T\mathbf{c}_1\mathbf{I}_{L\times L},\ldots,\mathbf{x}^T\mathbf{c}_R\mathbf{I}_{L\times L})\cdot\mathbf{B}^T} &= r_{\tilde{\mathbf{A}}\cdot\text{blockdiag}(u_1\mathbf{I}_{L\times L},\ldots,u_\gamma\mathbf{I}_{L\times L})\cdot\tilde{\mathbf{B}}^T} \\
&\geqslant r_{\tilde{\mathbf{A}}} + r_{\text{blockdiag}(u_1\mathbf{I}_{L\times L},\ldots,u_\gamma\mathbf{I}_{L\times L})\cdot\tilde{\mathbf{B}}^T} - \gamma L \\
&= r_{\tilde{\mathbf{A}}} + r_{\tilde{\mathbf{B}}} - \gamma L, \qquad (4.11)
\end{aligned}
$$

where the last equality is due to the fact that $\mathbf{u}$ has no zero elements. From the definition of k′-rank, we have

$$
r_{\tilde{\mathbf{A}}} \geqslant L\min(\gamma, k'_{\mathbf{A}}), \quad r_{\tilde{\mathbf{B}}} \geqslant L\min(\gamma, k'_{\mathbf{B}}). \qquad (4.12)
$$

Combination of (4.10)–(4.12) yields the following lower-bound on $\omega(\mathbf{x}^T\bar{\mathbf{C}})$:

$$
\omega(\mathbf{x}^T\bar{\mathbf{C}}) \geqslant \min(\gamma, k'_{\mathbf{A}}) + \min(\gamma, k'_{\mathbf{B}}) - \gamma. \qquad (4.13)
$$

(iii) *Combination of the two bounds.* Combination of (4.9) and (4.13) yields

$$
\min(\gamma, k'_{\mathbf{A}}) + \min(\gamma, k'_{\mathbf{B}}) - \gamma \leqslant \omega(\mathbf{x}^T\bar{\mathbf{C}}) \leqslant k'_{\mathbf{A}} + k'_{\mathbf{B}} - (R+1). \qquad (4.14)
$$

To be able to apply the permutation lemma, we need to show that $\gamma = \omega(\mathbf{x}^T\mathbf{C}) \leqslant \omega(\mathbf{x}^T\bar{\mathbf{C}})$. By (4.14), it suffices to show that $\gamma < \min(k'_{\mathbf{A}}, k'_{\mathbf{B}})$. We prove this by contradiction. Suppose $\gamma > \max(k'_{\mathbf{A}}, k'_{\mathbf{B}})$. Then (4.14) yields $\gamma \geqslant R+1$, which is impossible. Suppose next that $k'_{\mathbf{A}} \leqslant \gamma \leqslant k'_{\mathbf{B}}$. Then (4.14) yields $k'_{\mathbf{B}} \geqslant R+1$, which is also impossible. Since $\mathbf{A}$ and $\mathbf{B}$ can be exchanged in the latter case, we have that $\gamma < \min(k'_{\mathbf{A}}, k'_{\mathbf{B}})$. Equation (4.14) now implies that $\omega(\mathbf{x}^T\mathbf{C}) \leqslant \omega(\mathbf{x}^T\bar{\mathbf{C}})$. By the permutation lemma, there exist a unique permutation matrix $\mathbf{\Pi}_c$ and a nonsingular diagonal matrix $\mathbf{\Lambda}_c$ such that $\bar{\mathbf{C}} = \mathbf{C}\cdot\mathbf{\Pi}_c\cdot\mathbf{\Lambda}_c$. ◻

In the following lemma, we prove essential uniqueness of $\mathbf{A}$ and $\mathbf{B}$ when we restrict our attention to alternative $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ that are, in some sense, "nonsingular." What we mean is that there are no linear dependencies between columns that are not imposed by the dimensionality constraints.

LEMMA 4.3. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *represent a decomposition of* $\mathcal{T}$ *in rank-*$(L_r, L_r, 1)$ *terms,* $1 \leqslant r \leqslant R$. *Suppose the condition*

$$
k'_{\mathbf{A}} + k'_{\mathbf{B}} + k_{\mathbf{C}} \geqslant 2R+2 \qquad (4.15)
$$

holds and that we have an alternative decomposition of $\mathcal{T}$, *represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$, *with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality constraints. Then there holds* $\bar{\mathbf{A}} = \mathbf{A}\cdot\mathbf{\Pi}_a\cdot\mathbf{\Lambda}_a$, *in which* $\mathbf{\Pi}_a$ *is a block permutation matrix and* $\mathbf{\Lambda}_a$ *a square* *nonsingular block-diagonal matrix, compatible with the block structure of* $\mathbf{A}$*. There also holds* $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi}_b \cdot \mathbf{\Lambda}_b$*, in which* $\mathbf{\Pi}_b$ *is a block permutation matrix and* $\mathbf{\Lambda}_b$ *a square nonsingular block-diagonal matrix, compatible with the block structure of* $\mathbf{B}$*.*

Proof. It suffices to prove the lemma for $\mathbf{A}$. The result for $\mathbf{B}$ can be obtained by switching modes. We work in analogy with the proof of Lemma 4.2. Essential uniqueness of $\mathbf{A}$ now follows from the equivalence lemma for partitioned matrices.

(i) *Derivation of an upper-bound on* $\omega'(\mathbf{x}^T\bar{\mathbf{A}})$*.* The constraint on $k'_{\bar{\mathbf{A}}}$ implies that $k'_{\bar{\mathbf{A}}} \geqslant k'_{\mathbf{A}}$. Hence, if $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant R - k'_{\bar{\mathbf{A}}} + 1$, then

$$\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant R - k'_{\bar{\mathbf{A}}} + 1 \leqslant R - k'_{\mathbf{A}} + 1 \leqslant k'_{\mathbf{B}} + k_{\mathbf{C}} - (R+1), \tag{4.16}$$

where the last inequality corresponds to condition (4.15).

(ii) *Derivation of a lower-bound on* $\omega'(\mathbf{x}^T\bar{\mathbf{A}})$. By (2.7), the linear combination of $(J \times K)$ slices $\sum_{i=1}^{I} x_i \mathbf{T}_{J\times K,i}$ is given by

$$\mathbf{B} \cdot \text{blockdiag}(\mathbf{A}_1^T\mathbf{x}, \dots, \mathbf{A}_R^T\mathbf{x}) \cdot \mathbf{C}^T = \bar{\mathbf{B}} \cdot \text{blockdiag}(\bar{\mathbf{A}}_1^T\mathbf{x}, \dots, \bar{\mathbf{A}}_R^T\mathbf{x}) \cdot \bar{\mathbf{C}}^T.$$

We have

$$\begin{aligned} \omega'(\mathbf{x}^T\bar{\mathbf{A}}) &= r_{\text{blockdiag}(\bar{\mathbf{A}}_1^T\mathbf{x},\dots,\bar{\mathbf{A}}_R^T\mathbf{x})} \\ &\geqslant r_{\bar{\mathbf{B}}\cdot\text{blockdiag}(\bar{\mathbf{A}}_1^T\mathbf{x},\dots,\bar{\mathbf{A}}_R^T\mathbf{x})\cdot\bar{\mathbf{C}}^T} \\ &= r_{\mathbf{B}\cdot\text{blockdiag}(\mathbf{A}_1^T\mathbf{x},\dots,\mathbf{A}_R^T\mathbf{x})\cdot\mathbf{C}^T}. \end{aligned} \tag{4.17}$$

Let $\gamma = \omega'(\mathbf{x}^T\mathbf{A})$ and let $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ consist of the submatrices of $\mathbf{B}\cdot\text{blockdiag}(\mathbf{A}_1^T\mathbf{x}, \dots, \mathbf{A}_R^T\mathbf{x})$ and $\mathbf{C}$, respectively, corresponding to the parts of $\mathbf{x}^T\mathbf{A}$ that are not all-zero. Then $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ both have γ columns. Sylvester's inequality now yields

$$r_{\mathbf{B}\cdot\text{blockdiag}(\mathbf{A}_1^T\mathbf{x},\dots,\mathbf{A}_R^T\mathbf{x})\cdot\mathbf{C}^T} \geqslant r_{\tilde{\mathbf{B}}} + r_{\tilde{\mathbf{C}}} - \gamma. \tag{4.18}$$

The matrix $\tilde{\mathbf{B}}$ consists of γ nonzero vectors, sampled in the column spaces of the submatrices of $\mathbf{B}$ that correspond to the parts of $\mathbf{x}^T\mathbf{A}$ that are not all-zero. From the definition of k′-rank, we have

$$r_{\tilde{\mathbf{B}}} \geqslant \min(\gamma, k'_{\mathbf{B}}). \tag{4.19}$$

On the other hand, from the definition of k-rank, we have

$$r_{\tilde{\mathbf{C}}} \geqslant \min(\gamma, k_{\mathbf{C}}). \tag{4.20}$$

Combination of (4.17)–(4.20) yields the following lower-bound on $\omega'(\mathbf{x}^T\bar{\mathbf{A}})$:

$$\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \geqslant \min(\gamma, k'_{\mathbf{B}}) + \min(\gamma, k_{\mathbf{C}}) - \gamma. \tag{4.21}$$

(iii) *Combination of the two bounds.* This is analogous to Lemma 4.2. ∎

We now use Lemmas 4.2 and 4.3, which concern the essential uniqueness of the individual matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, to establish essential uniqueness of the overall decomposition of $\mathcal{T}$. Theorem 4.4 states that if $\mathbf{C}$ is full column rank and tall (meaning that $R \leqslant K$), then its essential uniqueness implies essential uniqueness of the overall tensor decomposition. Theorem 4.5 is the equivalent for $\mathbf{A}$ (or $\mathbf{B}$). However, none of the factor matrices needs to be tall for the decomposition to be unique. A more general case is dealt with in Theorem 4.7. Its proof makes use of Lemma 4.6, guaranteeing that under a generalized Kruskal condition, $\mathbf{A}$ and $\mathbf{B}$ not only are individually essentially unique but, moreover, are subject to the same permutation of their submatrices.

We first consider essential uniqueness of a tall full column rank matrix $\mathbf{C}$.

THEOREM 4.4. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *represent a decomposition of* $\mathcal{T}$ *in* R *rank-*$(L, L, 1)$ *terms. Suppose that we have an alternative decomposition of* $\mathcal{T}$, *represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$. *If*

$$k_{\mathbf{C}} = R \quad \text{and} \quad k'_{\mathbf{A}} + k'_{\mathbf{B}} \geqslant R + 2, \tag{4.22}$$

then $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *and* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ *are essentially equal.*

Proof. From (2.4) we have

$$\begin{aligned} \mathbf{T}_{IJ\times K} &= [(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_L \ \ldots \ (\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_L] \cdot \mathbf{C}^T \\ &= [(\bar{\mathbf{A}}_1 \odot_c \bar{\mathbf{B}}_1)\mathbf{1}_L \ \ldots \ (\bar{\mathbf{A}}_R \odot_c \bar{\mathbf{B}}_R)\mathbf{1}_L] \cdot \bar{\mathbf{C}}^T. \end{aligned} \tag{4.23}$$

From Lemma 4.2 we have

$$\bar{\mathbf{C}} = \mathbf{C} \cdot \mathbf{\Pi}_c \cdot \mathbf{\Lambda}_c. \tag{4.24}$$

Since $k_{\mathbf{C}} = R$, $\mathbf{C}$ is full column rank. Substitution of (4.24) in (4.23) now yields

$$\begin{aligned} &[(\mathbf{A}_1 \odot_c \mathbf{B}_1)\mathbf{1}_L \ \ldots \ (\mathbf{A}_R \odot_c \mathbf{B}_R)\mathbf{1}_L] \\ &\quad = [(\bar{\mathbf{A}}_1 \odot_c \bar{\mathbf{B}}_1)\mathbf{1}_L \ \ldots \ (\bar{\mathbf{A}}_R \odot_c \bar{\mathbf{B}}_R)\mathbf{1}_L] \cdot \mathbf{\Lambda}_c^T \cdot \mathbf{\Pi}_c^T. \end{aligned} \tag{4.25}$$

Taking into account that $(\bar{\mathbf{A}}_r \odot_c \bar{\mathbf{B}}_r)\mathbf{1}_L$ is a vector representation of the matrix $\bar{\mathbf{A}}_r \cdot \bar{\mathbf{B}}_r^T$, $1 \leqslant r \leqslant R$, this implies that the matrices $\bar{\mathbf{A}}_r \cdot \bar{\mathbf{B}}_r^T$ are ordered in the same way as the vectors $\bar{\mathbf{c}}_r$. Furthermore, if $\bar{\mathbf{c}}_i = \lambda \mathbf{c}_j$, then $(\bar{\mathbf{A}}_i \odot_c \bar{\mathbf{B}}_i)\mathbf{1}_L = \lambda^{-1}(\mathbf{A}_j \odot_c \mathbf{B}_j)\mathbf{1}_L$, or, equivalently, $\bar{\mathbf{A}}_i \cdot \bar{\mathbf{B}}_i^T = \lambda^{-1}\mathbf{A}_j \cdot \mathbf{B}_j^T$. ∎

We now consider essential uniqueness of a tall full column rank matrix $\mathbf{A}$ or $\mathbf{B}$.

THEOREM 4.5. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *represent a decomposition of* $\mathcal{T}$ *in rank-*$(L_r, L_r, 1)$ *terms,* $1 \leqslant r \leqslant R$. *Suppose that we have an alternative decomposition of* $\mathcal{T}$, *represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$, *with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality constraints. If*

$$k'_{\mathbf{A}} = R \quad \text{and} \quad k'_{\mathbf{B}} + k_{\mathbf{C}} \geqslant R + 2 \tag{4.26}$$

or

$$k'_{\mathbf{B}} = R \quad \text{and} \quad k'_{\mathbf{A}} + k_{\mathbf{C}} \geqslant R + 2, \tag{4.27}$$

then $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *and* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ *are essentially equal.*

Proof. It suffices to prove the theorem for condition (4.26). The result for (4.27) is obtained by switching modes.

From (2.5) we have

$$\mathbf{T}_{JK\times I} = (\mathbf{B} \odot \mathbf{C}) \cdot \mathbf{A}^T = (\bar{\mathbf{B}} \odot \bar{\mathbf{C}}) \cdot \bar{\mathbf{A}}^T. \tag{4.28}$$

From Lemma 4.3 we have

$$\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a. \tag{4.29}$$

Since $k'_{\mathbf{A}} = R$, $\mathbf{A}$ is full column rank. Substitution of (4.29) in (4.28) now yields

$$\mathbf{B} \odot \mathbf{C} = (\bar{\mathbf{B}} \odot \bar{\mathbf{C}}) \cdot \mathbf{\Lambda}_a^T \cdot \mathbf{\Pi}_a^T. \tag{4.30}$$

This implies that the matrices $\bar{\mathbf{B}}_r \otimes \bar{\mathbf{c}}_r$ are ordered in the same way as the matrices $\bar{\mathbf{A}}_r$. Furthermore, if $\bar{\mathbf{A}}_i = \mathbf{A}_j \cdot \mathbf{L}$, with $\mathbf{L}$ nonsingular, then $\bar{\mathbf{B}}_i \otimes \bar{\mathbf{c}}_i = (\mathbf{B}_j \otimes \mathbf{c}_j) \cdot \mathbf{L}^{-T}$, or, equivalently, $\bar{\mathbf{B}}_i \circ \bar{\mathbf{c}}_i = (\mathbf{B}_j \cdot \mathbf{L}^{-T}) \circ \mathbf{c}_j$. □

We now prove that under a generalized Kruskal condition, the submatrices of $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in an alternative decomposition of $\mathcal{T}$ are ordered in the same way.

LEMMA 4.6. *Let $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ represent a decomposition of $\mathcal{T}$ into rank-$(L_r, L_r, 1)$ terms, $1 \leqslant r \leqslant R$. Suppose that we have an alternative decomposition of $\mathcal{T}$, represented by $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$, with $k'_{\bar{\mathbf{A}}}$ and $k'_{\bar{\mathbf{B}}}$ maximal under the given dimensionality constraints. If the condition*

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k_{\mathbf{C}} \geqslant 2R + 2 \tag{4.31}$$

holds, then $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_b$, in which $\mathbf{\Pi}$ is a block permutation matrix and $\mathbf{\Lambda}_a$ and $\mathbf{\Lambda}_b$ nonsingular block-diagonal matrices, compatible with the block structure of $\mathbf{A}$ and $\mathbf{B}$.

Proof. From Lemma 4.3 we know that $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi}_b \cdot \mathbf{\Lambda}_b$. We show that $\mathbf{\Pi}_a = \mathbf{\Pi}_b$ if (4.31) holds. We work in analogy with [38, pp. 129–132] and [54].

From (2.9) we have

$$\begin{aligned} \mathbf{T}_{I \times J, k} &= \mathbf{A} \cdot \mathrm{blockdiag}(c_{k1} \mathbf{I}_{L_1 \times L_1}, \ldots, c_{kR} \mathbf{I}_{L_R \times L_R}) \cdot \mathbf{B}^T \\ &= \bar{\mathbf{A}} \cdot \mathrm{blockdiag}(\bar{c}_{k1} \mathbf{I}_{L_1 \times L_1}, \ldots, \bar{c}_{kR} \mathbf{I}_{L_R \times L_R}) \cdot \bar{\mathbf{B}}^T. \end{aligned}$$

For vectors $\mathbf{v}$ and $\mathbf{w}$ we have

$$\begin{aligned} &(\mathbf{v}^T \mathbf{A}) \cdot \mathrm{blockdiag}(c_{k1} \mathbf{I}_{L_1 \times L_1}, \ldots, c_{kR} \mathbf{I}_{L_R \times L_R}) \cdot (\mathbf{w}^T \mathbf{B})^T \\ &\quad = (\mathbf{v}^T \bar{\mathbf{A}}) \cdot \mathrm{blockdiag}(\bar{c}_{k1} \mathbf{I}_{L_1 \times L_1}, \ldots, \bar{c}_{kR} \mathbf{I}_{L_R \times L_R}) \cdot (\mathbf{w}^T \bar{\mathbf{B}})^T \\ &\quad = (\mathbf{v}^T \mathbf{A} \mathbf{\Pi}_a) \cdot \mathbf{\Lambda}_a \cdot \mathrm{blockdiag}(\bar{c}_{k1} \mathbf{I}_{L_1 \times L_1}, \ldots, \bar{c}_{kR} \mathbf{I}_{L_R \times L_R}) \cdot \mathbf{\Lambda}_b^T \cdot (\mathbf{w}^T \mathbf{B} \mathbf{\Pi}_b)^T. \end{aligned} \tag{4.32}$$

We stack (4.32), for $k = 1, \ldots, K$, in

$$\begin{aligned} &\mathbf{C} \cdot \mathrm{blockdiag}(\mathbf{v}^T \mathbf{A}) \cdot \mathrm{blockdiag}(\mathbf{B}^T \mathbf{w}) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ &\quad = \bar{\mathbf{C}} \cdot \mathrm{blockdiag}(\mathbf{v}^T \mathbf{A} \mathbf{\Pi}_a) \cdot \mathbf{\Lambda}_a \cdot \mathbf{\Lambda}_b^T \cdot \mathrm{blockdiag}(\mathbf{\Pi}_b^T \mathbf{B}^T \mathbf{w}) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \end{aligned} \tag{4.33}$$

We define

$$\mathbf{p} = \mathrm{blockdiag}(\mathbf{v}^T \mathbf{A}) \cdot \mathrm{blockdiag}(\mathbf{B}^T \mathbf{w}) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{v}^T \mathbf{A}_1 \cdot \mathbf{B}_1^T \mathbf{w} \\ \vdots \\ \mathbf{v}^T \mathbf{A}_R \cdot \mathbf{B}_R^T \mathbf{w} \end{pmatrix}.$$

Let the index function $g(x)$ be given by $\mathbf{A}\mathbf{\Pi}_a = \left(\mathbf{A}_{g(1)}\ \mathbf{A}_{g(2)}\ \ldots\ \mathbf{A}_{g(R)}\right)$. Let a second index function $h(x)$ be given by $\mathbf{B}\mathbf{\Pi}_b = \left(\mathbf{B}_{h(1)}\ \mathbf{B}_{h(2)}\ \ldots\ \mathbf{B}_{h(R)}\right)$. We define

$$\begin{aligned}\mathbf{q} &= \operatorname{blockdiag}(\mathbf{v}^T\mathbf{A}\mathbf{\Pi}_a) \cdot \mathbf{\Lambda}_a \cdot \mathbf{\Lambda}_b^T \cdot \operatorname{blockdiag}(\mathbf{\Pi}_b^T\mathbf{B}^T\mathbf{w}) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{v}^T\mathbf{A}_{g(1)} \cdot \mathbf{\Lambda}_{a,1} \cdot \mathbf{\Lambda}_{b,1}^T \cdot \mathbf{B}_{h(1)}^T\mathbf{w} \\ \vdots \\ \mathbf{v}^T\mathbf{A}_{g(R)} \cdot \mathbf{\Lambda}_{a,R} \cdot \mathbf{\Lambda}_{b,R}^T \cdot \mathbf{B}_{h(R)}^T\mathbf{w} \end{pmatrix},\end{aligned}$$

where $\mathbf{\Lambda}_{a,r}$ and $\mathbf{\Lambda}_{b,r}$ denote the rth block of $\mathbf{\Lambda}_a$ and $\mathbf{\Lambda}_b$, respectively.

Equation (4.33) can now be written as $\mathbf{C} \cdot \mathbf{p} = \bar{\mathbf{C}} \cdot \mathbf{q}$. Below we show by contradiction that $\mathbf{\Pi}_a = \mathbf{\Pi}_b$ if (4.31) holds. If $\mathbf{\Pi}_a \neq \mathbf{\Pi}_b$, then we will be able to find vectors $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{q} = \mathbf{0}$ and $\mathbf{p} \neq \mathbf{0}$ has less than $k_{\mathbf{C}}$ nonzero elements. This implies that a set of less than $k_{\mathbf{C}}$ columns of $\mathbf{C}$ is linearly dependent, which contradicts the definition of $k_{\mathbf{C}}$.

Suppose that $\mathbf{\Pi}_a \neq \mathbf{\Pi}_b$. Then there exists an r such that $\mathbf{A}_r$ is the sth submatrix of $\mathbf{A}\mathbf{\Pi}_a$, $\mathbf{B}_r$ is the tth submatrix of $\mathbf{B}\mathbf{\Pi}_b$, and $s \neq t$. Formally, there exists an r such that $r = g(s) = h(t)$ and $s \neq t$. We now create two index sets $\mathsf{S}, \mathsf{T} \subset \{1, \ldots, R\}$ as follows:

- Put $g(t)$ in S and $h(s)$ in T.
- For $x \in \{1, \ldots, R\} \setminus \{s, t\}$, add $g(x)$ to S if $\operatorname{card}(\mathsf{S}) < k'_{\mathbf{A}} - 1$. Otherwise, add $h(x)$ to T.

The sets S and T have the following properties. Since $k'_{\mathbf{A}} - 1 \leqslant R - 1$, S contains exactly $k'_{\mathbf{A}} - 1$ elements. The set T contains $R - \operatorname{card}(\mathsf{S}) = R - k'_{\mathbf{A}} + 1$ elements. Because of (4.31) and $k_{\mathbf{C}} \leqslant R$, this is less than or equal to $k'_{\mathbf{B}} - 1$ elements. In the xth element of $\mathbf{q}$ we have either $g(x) \in \mathsf{S}$ or $h(x) \in \mathsf{T}$, $x = 1, \ldots, R$. The index $r = g(s) = h(t)$ is neither an element of S nor an element of T. Denote $\{i_1, i_2, \ldots, i_{k'_{\mathbf{A}}-1}\} = \mathsf{S}$ and $\{j_1, j_2, \ldots, j_{R-k'_{\mathbf{A}}+1}\} = \mathsf{T}$.

We choose vectors $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{v}^T\mathbf{A}_i = \mathbf{0}$ if $i \in \mathsf{S}$, $\mathbf{w}^T\mathbf{B}_j = \mathbf{0}$ if $j \in \mathsf{T}$ and $\mathbf{v}^T\mathbf{A}_r\mathbf{B}_r^T\mathbf{w} \neq 0$. This is possible for the following reasons. By the definition of $k'_{\mathbf{A}}$, $[\mathbf{A}_{i_1}\ \ldots\ \mathbf{A}_{i_{k'_{\mathbf{A}}-1}}\ \mathbf{A}_r]$ is full column rank. We have to choose $\mathbf{v}$ in $\operatorname{null}([\mathbf{A}_{i_1}\ \ldots\ \mathbf{A}_{i_{k'_{\mathbf{A}}-1}}])$. The projection of this subspace on $\operatorname{span}(\mathbf{A}_r)$ is of dimension L_r. By varying $\mathbf{v}$ in $\operatorname{null}([\mathbf{A}_{i_1}\ \ldots\ \mathbf{A}_{i_{k'_{\mathbf{A}}-1}}])$, $\mathbf{v}^T\mathbf{A}_r$ can be made equal to any vector in $\mathbb{K}^{1\times L_r}$. For instance, we can choose $\mathbf{v}$ such that $\mathbf{v}^T\mathbf{A}_r = (1\ 0\ \ldots\ 0)$. Similarly, we can choose a vector $\mathbf{w}$ in $\operatorname{null}([\mathbf{B}_{j_1}\ \ldots\ \mathbf{B}_{j_{R-k'_{\mathbf{A}}+1}}])$ satisfying $\mathbf{w}^T\mathbf{B}_r = (1\ 0\ \ldots\ 0)$.

For the vectors $\mathbf{v}$ and $\mathbf{w}$ above, we have $\mathbf{q} = \mathbf{0}$. On the other hand, the rth element of $\mathbf{p}$ is nonzero. Define $\mathsf{S}^c = \{1, \ldots, R\} \setminus \mathsf{S}$ and $\mathsf{T}^c = \{1, \ldots, R\} \setminus \mathsf{T}$. The number of nonzero entries of $\mathbf{p}$ is bounded from above by

$$\operatorname{card}(\mathsf{S}^c \cap \mathsf{T}^c) \leqslant \operatorname{card}(\mathsf{S}^c) \leqslant R - k'_{\mathbf{A}} + 1 \leqslant k_{\mathbf{C}} - 1,$$

where the last inequality is due to (4.31) and $k'_{\mathbf{B}} \leqslant R$. Hence, $\mathbf{C} \cdot \mathbf{p} = \mathbf{0}$ implies that a set of less than $k_{\mathbf{C}}$ columns of $\mathbf{C}$ is linearly dependent, which contradicts the definition of $k_{\mathbf{C}}$. This completes the proof. $\square$

THEOREM 4.7. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *represent a decomposition of* $\mathcal{T}$ *in generic rank-*$(L_r, L_r, 1)$ *terms,* $1 \leqslant r \leqslant R$. *Suppose that we have an alternative decomposition of* $\mathcal{T}$, *represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$, *with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality* *constraints. If the conditions*

$$IJ \geqslant \sum_{r=1}^{R} L_r^2, \tag{4.34}$$

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k_{\mathbf{C}} \geqslant 2R + 2 \tag{4.35}$$

hold, then $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *and* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ *are essentially equal.*

Proof. From Lemma 4.6 we have that $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_b$. Put the submatrices of $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in the same order as the submatrices of $\mathbf{A}$ and $\mathbf{B}$. After reordering, we have $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Lambda}_a$, with $\mathbf{\Lambda}_a = \text{blockdiag}(\mathbf{\Lambda}_{a,1}, \ldots, \mathbf{\Lambda}_{a,R})$, and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Lambda}_b$, with $\mathbf{\Lambda}_b = \text{blockdiag}(\mathbf{\Lambda}_{b,1}, \ldots, \mathbf{\Lambda}_{b,R})$. From (2.4) we have that

$$\begin{aligned}
\mathbf{T}_{IJ \times K} &= (\mathbf{A} \odot_c \mathbf{B}) \cdot \text{blockdiag}(\mathbf{1}_{L_1}, \ldots, \mathbf{1}_{L_R}) \cdot \mathbf{C}^T \\
&= (\mathbf{A} \odot \mathbf{B}) \cdot \text{blockdiag}(\text{vec}(\mathbf{I}_{L_1 \times L_1}), \ldots, \text{vec}(\mathbf{I}_{L_R \times L_R})) \cdot \mathbf{C}^T \\
&= (\bar{\mathbf{A}} \odot \bar{\mathbf{B}}) \cdot \text{blockdiag}(\text{vec}(\mathbf{I}_{L_1 \times L_1}), \ldots, \text{vec}(\mathbf{I}_{L_R \times L_R})) \cdot \bar{\mathbf{C}}^T \\
&= (\mathbf{A} \odot \mathbf{B}) \cdot \text{blockdiag}(\text{vec}(\mathbf{\Lambda}_{a,1} \cdot \mathbf{\Lambda}_{b,1}^T), \ldots, \text{vec}(\mathbf{\Lambda}_{a,R} \cdot \mathbf{\Lambda}_{b,R}^T)) \cdot \bar{\mathbf{C}}^T. \qquad (4.36)
\end{aligned}$$

From [19, Lemma 3.3] we have that, under condition (4.34), $\mathbf{A} \odot \mathbf{B}$ is generically full column rank. Equation (4.36) then implies that there exist nonzero scalars α_r such that $\mathbf{\Lambda}_{a,r} \cdot \mathbf{\Lambda}_{b,r}^T = \alpha_r \, \mathbf{I}_{L_r \times L_r}$ (i.e., $\mathbf{\Lambda}_{a,r} = \alpha_r \, \mathbf{\Lambda}_{b,r}^{-T}$) and $\mathbf{c}_r = \alpha_r \bar{\mathbf{c}}_r$, $1 \leqslant r \leqslant R$. In other words, $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ are equal up to trivial indeterminacies. ▯

5. The decomposition in rank-(L, M, N) terms.

In this section we study the uniqueness of the decomposition in rank-(L, M, N) terms. We use the notation introduced in section 2.2. Section 5.1 concerns uniqueness of the general decomposition. In section 5.2 we have a closer look at the special case of rank-$(2, 2, 2)$ terms.

5.1. General results.

In this section we follow the same structure as in section 4:

Theorem 5.1	corresponds to	Theorem 4.1
Lemma 5.2		Lemma 4.3
Theorem 5.3		Theorem 4.5
Lemma 5.4		Lemma 4.6
Theorem 5.5		Theorem 4.7.

For decompositions in generic rank-(L, M, N) terms, the results of this section can be summarized as follows. We have essential uniqueness if

(i) Theorem 5.1:

$$\begin{aligned}
&L = M \quad \text{and} \quad I \geqslant LR \quad \text{and} \quad J \geqslant MR \quad \text{and} \quad N \geqslant 3 \\
&\qquad \text{and} \quad \mathbf{C}_r \text{ is full column rank, } 1 \leqslant r \leqslant R;
\end{aligned} \tag{5.1}$$

(ii) Theorem 5.3:

$$I \geqslant LR \quad \text{and} \quad N > L + M - 2 \quad \text{and} \quad \min\left(\left\lfloor \frac{J}{M} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{K}{N} \right\rfloor, R\right) \geqslant R + 2; \tag{5.2}$$

or

$$J \geqslant MR \quad \text{and} \quad N > L + M - 2 \quad \text{and} \quad \min\left(\left\lfloor \frac{I}{L} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{K}{N} \right\rfloor, R\right) \geqslant R + 2. \tag{5.3}$$

(iii) Theorem 5.5:

$$N > L + M - 2 \quad \text{and} \quad \min\left(\left\lfloor \frac{I}{L} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{J}{M} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{K}{N} \right\rfloor, R\right) \geqslant 2R + 2. \tag{5.4}$$

First we have a uniqueness result that stems from the fact that the column spaces of $\mathbf{A}_r$, $1 \leqslant r \leqslant R$, are invariant subspaces of quotients of tensor slices.

THEOREM 5.1. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *represent a decomposition of* $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ *in* R *rank-*(L, L, N) *terms. Suppose that* $\operatorname{rank}(\mathbf{A}) = LR$*,* $\operatorname{rank}(\mathbf{B}) = LR$*,* $\operatorname{rank}_{k'}(\mathbf{C}) \geqslant 1$*,* $N \geqslant 3$*, and that* $\mathcal{D}$ *is generic. Then* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *is essentially unique.*

Proof. From Theorem 6.1 below we have that under the conditions specified in Theorem 5.1, a decomposition in terms of the form $\mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r$ is essentially unique. Consequently, a decomposition in terms of the form $\mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}$ is essentially unique if $\mathbf{C}$ is full column rank. A fortiori, reasoning as in the proof of Theorem 6.1, a decomposition in terms of the form $\mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r$, in which the matrices $\mathbf{C}_r$ are possibly different, is essentially unique if these matrices $\mathbf{C}_r$ are full column rank. $\square$

Remark 4. The generalization to the decomposition in rank-(L_r, L_r, N_r) terms, $1 \leqslant r \leqslant R$, is trivial.

Remark 5. In the nongeneric case, lack of uniqueness can be due to the fact that tensors $\mathcal{D}_r$ can be further block-diagonalized by means of basis transformations in their mode-1, mode-2, and mode-3 vector space. We give an example.

Example 1. Assume a tensor $\mathcal{T} \in \mathbb{K}^{12 \times 12 \times 12}$ that can be decomposed in three rank-$(4, 4, 4)$ terms as follows:

$$\mathcal{T} = \sum_{r=1}^{3} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r$$

with $\mathcal{D}_r \in \mathbb{K}^{4 \times 4 \times 4}$, $\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r \in \mathbb{K}^{12 \times 4}$, $1 \leqslant r \leqslant 3$. Now assume that $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$ can be further decomposed as follows:

$$\begin{aligned}
\mathcal{D}_1 &= \mathbf{u}_1 \circ \mathbf{v}_1 \circ \mathbf{w}_1 + \mathbf{u}_2 \circ \mathbf{v}_2 \circ \mathbf{w}_2 + \mathcal{H}_1 \bullet_1 \mathbf{E}_1 \bullet_2 \mathbf{F}_1 \bullet_3 \mathbf{G}_1, \\
\mathcal{D}_2 &= \mathbf{u}_3 \circ \mathbf{v}_3 \circ \mathbf{w}_3 + \mathcal{H}_2 \bullet_1 \mathbf{E}_2 \bullet_2 \mathbf{F}_2 \bullet_3 \mathbf{G}_2, \\
\mathcal{D}_3 &= \mathbf{u}_4 \circ \mathbf{v}_4 \circ \mathbf{w}_4 + \mathcal{H}_3 \bullet_1 \mathbf{E}_3 \bullet_2 \mathbf{F}_3 \bullet_3 \mathbf{G}_3,
\end{aligned}$$

where $\mathbf{u}_s, \mathbf{v}_s, \mathbf{w}_s \in \mathbb{K}^4$, $1 \leqslant s \leqslant 4$, $\mathbf{E}_1, \mathbf{F}_1, \mathbf{G}_1 \in \mathbb{K}^{4 \times 2}$, $\mathbf{E}_2, \mathbf{E}_3, \mathbf{F}_2, \mathbf{F}_3, \mathbf{G}_2, \mathbf{G}_3 \in \mathbb{K}^{4 \times 3}$, $\mathcal{H}_1 \in \mathbb{K}^{2 \times 2 \times 2}$, $\mathcal{H}_2, \mathcal{H}_3 \in \mathbb{K}^{3 \times 3 \times 3}$. Then we have the following alternative decomposition in three rank-$(4, 4, 4)$ terms:

$$\begin{aligned}
\mathcal{T} = {} & [(\mathbf{A}_2\mathbf{u}_3) \circ (\mathbf{B}_2\mathbf{v}_3) \circ (\mathbf{C}_2\mathbf{w}_3) + (\mathbf{A}_3\mathbf{u}_4) \circ (\mathbf{B}_3\mathbf{v}_4) \circ (\mathbf{C}_3\mathbf{w}_4) \\
& + \mathcal{H}_1 \bullet_1 (\mathbf{A}_1\mathbf{E}_1) \bullet_2 (\mathbf{B}_1\mathbf{F}_1) \bullet_3 (\mathbf{C}_1\mathbf{G}_1)] \\
& + [(\mathbf{A}_1\mathbf{u}_1) \circ (\mathbf{B}_1\mathbf{v}_1) \circ (\mathbf{C}_1\mathbf{w}_1) + \mathcal{H}_2 \bullet_1 (\mathbf{A}_2\mathbf{E}_2) \bullet_2 (\mathbf{B}_2\mathbf{F}_2) \bullet_3 (\mathbf{C}_2\mathbf{G}_2)] \\
& + [(\mathbf{A}_1\mathbf{u}_2) \circ (\mathbf{B}_1\mathbf{v}_2) \circ (\mathbf{C}_1\mathbf{w}_2) + \mathcal{H}_3 \bullet_1 (\mathbf{A}_3\mathbf{E}_3) \bullet_2 (\mathbf{B}_3\mathbf{F}_3) \bullet_3 (\mathbf{C}_3\mathbf{G}_3)].
\end{aligned}$$

We now prove essential uniqueness of $\mathbf{A}$ and $\mathbf{B}$ under a constraint on the block dimensions and a Kruskal-type condition.

LEMMA 5.2. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *represent a decomposition of* $\mathcal{T}$ *in* R *rank-*(L, M, N) *terms. Suppose that the conditions*

$$N > L + M - 2, \tag{5.5}$$

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k'_{\mathbf{C}} \geqslant 2R + 2 \tag{5.6}$$

hold and that we have an alternative decomposition of $\mathcal{T}$*, represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathcal{D}})$*, with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality constraints. For generic* $\mathcal{D}$ *there holds that* $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$*, in which* $\mathbf{\Pi}_a$ *is a block permutation matrix and* $\mathbf{\Lambda}_a$ *a square nonsingular block-diagonal matrix, compatible with the structure of* $\mathbf{A}$*. There also holds* $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi}_b \cdot \mathbf{\Lambda}_b$*, in which* $\mathbf{\Pi}_b$ *is a block permutation matrix and* $\mathbf{\Lambda}_b$ *a square nonsingular block-diagonal matrix, compatible with the structure of* $\mathbf{B}$*.*

Proof. It suffices to prove the lemma for $\mathbf{A}$. The result for $\mathbf{B}$ can be obtained by switching modes. We work in analogy with [54] and the proof of Lemma 4.2 and 4.3. We use the equivalence lemma for partitioned matrices to prove essential uniqueness of $\mathbf{A}$.

(i) *Derivation of an upper-bound on* $\omega'(\mathbf{x}^T \bar{\mathbf{A}})$. The constraint on $k'_{\bar{\mathbf{A}}}$ implies that $k'_{\bar{\mathbf{A}}} \geqslant k'_{\mathbf{A}}$. Hence, if $\omega'(\mathbf{x}^T \bar{\mathbf{A}}) \leqslant R - k'_{\bar{\mathbf{A}}} + 1$, then

$$\omega'(\mathbf{x}^T \bar{\mathbf{A}}) \leqslant R - k'_{\bar{\mathbf{A}}} + 1 \leqslant R - k'_{\mathbf{A}} + 1 \leqslant k'_{\mathbf{B}} + k_{\mathbf{C}} - (R + 1), \tag{5.7}$$

where the last inequality corresponds to condition (5.6).

(ii) *Derivation of a lower-bound on* $\omega'(\mathbf{x}^T \bar{\mathbf{A}})$. Consider $\mathcal{D}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r)$ and $\bar{\mathcal{D}}_r \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_r)$, $1 \leqslant r \leqslant R$, as $(M \times N)$ matrices. Then the linear combination of slices $\sum_{i=1}^{I} x_i \mathbf{T}_{J \times K, i}$ is given by

$$\begin{aligned} &\mathbf{B} \cdot \text{blockdiag}[\mathcal{D}_1 \bullet_1 (\mathbf{x}^T \mathbf{A}_1), \ldots, \mathcal{D}_R \bullet_1 (\mathbf{x}^T \mathbf{A}_R)] \cdot \mathbf{C}^T \\ &\quad = \bar{\mathbf{B}} \cdot \text{blockdiag}[\bar{\mathcal{D}}_1 \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_1), \ldots, \bar{\mathcal{D}}_R \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_R)] \cdot \bar{\mathbf{C}}^T. \end{aligned}$$

Taking into account that $N > M$, we have

$$\begin{aligned} M\omega'(\mathbf{x}^T \bar{\mathbf{A}}) &\geqslant r_{\text{blockdiag}[\bar{\mathcal{D}}_1 \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_1), \ldots, \bar{\mathcal{D}}_R \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_R)]} \\ &\geqslant r_{\bar{\mathbf{B}} \cdot \text{blockdiag}[\bar{\mathcal{D}}_1 \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_1), \ldots, \bar{\mathcal{D}}_R \bullet_1 (\mathbf{x}^T \bar{\mathbf{A}}_R)] \cdot \bar{\mathbf{C}}^T} \\ &= r_{\mathbf{B} \cdot \text{blockdiag}[\mathcal{D}_1 \bullet_1 (\mathbf{x}^T \mathbf{A}_1), \ldots, \mathcal{D}_R \bullet_1 (\mathbf{x}^T \mathbf{A}_R)] \cdot \mathbf{C}^T}. \end{aligned} \tag{5.8}$$

Since the tensors $\mathcal{D}_r$ are generic, and because of condition (5.5), all the $(M \times N)$ matrices $\mathcal{D}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r)$ are rank-M. (Rank deficiency would imply that $N - M + 1$ determinants are zero, while $\mathbf{x}$ provides only $L - 1$ independent parameters and an irrelevant scaling factor.) Define $(K \times M)$ matrices $\underline{\mathbf{C}}_r = \mathbf{C}_r \cdot [\mathcal{D}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r)]^T$, $1 \leqslant r \leqslant R$. Let $\gamma = \omega'(\mathbf{x}^T \mathbf{A})$ and $\underline{\mathbf{C}} = (\underline{\mathbf{C}}_1 \ \ldots \ \underline{\mathbf{C}}_R)$. Let $\tilde{\mathbf{B}}$ and $\tilde{\underline{\mathbf{C}}}$ consist of the submatrices of $\mathbf{B}$ and $\underline{\mathbf{C}}$, respectively, corresponding to the parts of $\mathbf{x}^T \mathbf{A}$ that are not all-zero. From (5.8) we have

$$M\omega'(\mathbf{x}^T \bar{\mathbf{A}}) \geqslant r_{\tilde{\mathbf{B}} \cdot \tilde{\underline{\mathbf{C}}}^T}. \tag{5.9}$$

Both $\tilde{\mathbf{B}}$ and $\tilde{\underline{\mathbf{C}}}$ have γM columns. Sylvester's inequality now yields

$$r_{\tilde{\mathbf{B}} \cdot \tilde{\underline{\mathbf{C}}}^T} \geqslant r_{\tilde{\mathbf{B}}} + r_{\tilde{\underline{\mathbf{C}}}} - \gamma M. \tag{5.10}$$

From the definition of k′-rank, we have

$$r_{\tilde{\mathbf{B}}} \geqslant M \min(\gamma, k'_{\mathbf{B}}). \tag{5.11}$$

On the other hand, $\underline{\tilde{\mathbf{C}}}$ consists of γ $(K \times M)$ submatrices, of which the columns are sampled in the column space of the corresponding submatrix of $\mathbf{C}$. From the definition of k′-rank, we must have

$$r_{\underline{\tilde{\mathbf{C}}}} \geqslant M \min(\gamma, k'_{\mathbf{C}}). \tag{5.12}$$

Combination of (5.9)–(5.12) yields the following lower-bound on $\omega'(\mathbf{x}^T\bar{\mathbf{A}})$:

$$\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \geqslant \min(\gamma, k'_{\mathbf{B}}) + \min(\gamma, k'_{\mathbf{C}}) - \gamma. \tag{5.13}$$

(iii) *Combination of the two bounds.* This is analogous to Lemma 4.2. □

If matrix $\mathbf{A}$ or $\mathbf{B}$ is tall and full column rank, then its essential uniqueness implies essential uniqueness of the overall tensor decomposition.

THEOREM 5.3. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *represent a decomposition of* $\mathcal{T}$ *in* R *rank-*(L, M, N) *terms, with* $N > L + M - 2$. *Suppose that we have an alternative decomposition of* $\mathcal{T}$, *represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathcal{D}})$, *with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality constraints. For generic* $\mathcal{D}$ *there holds that if*

$$k'_{\mathbf{A}} = R \quad \textit{and} \quad k'_{\mathbf{B}} + k'_{\mathbf{C}} \geqslant R + 2 \tag{5.14}$$

or

$$k'_{\mathbf{B}} = R \quad \textit{and} \quad k'_{\mathbf{A}} + k'_{\mathbf{C}} \geqslant R + 2, \tag{5.15}$$

then $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *and* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathcal{D}})$ *are essentially equal.*

Proof. It suffices to prove the theorem for $\mathbf{A}$. The result for $\mathbf{B}$ is obtained by switching modes. From (2.12) we have

$$\begin{aligned} \mathbf{T}_{JK \times I} &= (\mathbf{B} \odot \mathbf{C}) \cdot \operatorname{blockdiag}((\mathcal{D}_1)_{MN \times L}, \ldots, (\mathcal{D}_R)_{MN \times L}) \cdot \mathbf{A}^T \\ &= (\bar{\mathbf{B}} \odot \bar{\mathbf{C}}) \cdot \operatorname{blockdiag}((\bar{\mathcal{D}}_1)_{MN \times L}, \ldots, (\bar{\mathcal{D}}_R)_{MN \times L}) \cdot \bar{\mathbf{A}}^T. \end{aligned} \tag{5.16}$$

From Lemma 5.2 we have

$$\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a. \tag{5.17}$$

Since $k'_{\mathbf{A}} = R$, $\mathbf{A}$ is full column rank. Substitution of (5.17) in (5.16) now yields

$$\begin{aligned} &(\mathbf{B} \odot \mathbf{C}) \cdot \operatorname{blockdiag}((\mathcal{D}_1)_{MN \times L}, \ldots, (\mathcal{D}_R)_{MN \times L}) \\ &\quad = (\bar{\mathbf{B}} \odot \bar{\mathbf{C}}) \cdot \operatorname{blockdiag}((\bar{\mathcal{D}}_1)_{MN \times L}, \ldots, (\bar{\mathcal{D}}_R)_{MN \times L}) \cdot \mathbf{\Lambda}_a^T \cdot \mathbf{\Pi}_a^T. \end{aligned} \tag{5.18}$$

This implies that the matrices $(\bar{\mathbf{B}}_r \otimes \bar{\mathbf{C}}_r) \cdot (\bar{\mathcal{D}}_r)_{MN \times L}$ are permuted in the same way with respect to $(\mathbf{B}_r \otimes \mathbf{C}_r) \cdot (\mathcal{D}_r)_{MN \times L}$ as the matrices $\bar{\mathbf{A}}_r$ with respect to $\mathbf{A}_r$. Furthermore, if $\bar{\mathbf{A}}_i = \mathbf{A}_j \cdot \mathbf{F}$, then $(\bar{\mathbf{B}}_i \otimes \bar{\mathbf{C}}_i) \cdot (\bar{\mathcal{D}}_i)_{MN \times L} \cdot \mathbf{F}^T = (\mathbf{B}_j \otimes \mathbf{C}_j) \cdot (\mathcal{D}_j)_{MN \times L}$. Equivalently, we have $\bar{\mathcal{D}}_i \bullet_2 \bar{\mathbf{B}}_i \bullet_3 \bar{\mathbf{C}}_i = \mathcal{D}_j \bullet_1 \mathbf{F}^{-1} \bullet_2 \mathbf{B}_j \bullet_3 \mathbf{C}_j$. □

We now prove that under conditions (5.5) and (5.6), the submatrices of $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in an alternative decomposition of $\mathcal{T}$ are ordered in the same way.

LEMMA 5.4. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *represent a decomposition of* $\mathcal{T}$ *in* R *rank-*(L, M, N) *terms. Suppose that we have an alternative decomposition of* $\mathcal{T}$, *represented by* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathcal{D}})$, *with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality constraints. For generic* $\mathcal{D}$ *there holds that if conditions* (5.5) *and* (5.6) *hold, then* $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_a$ *and* $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_b$, *in which* $\mathbf{\Pi}$ *is a block permutation matrix and* $\mathbf{\Lambda}_a$ *and* $\mathbf{\Lambda}_b$ *square nonsingular block-diagonal matrices, compatible with the block structure of* $\mathbf{A}$ *and* $\mathbf{B}$.

Proof. From Lemma 5.2 we know that $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi}_b \cdot \mathbf{\Lambda}_b$. We show that $\mathbf{\Pi}_a = \mathbf{\Pi}_b$ if (5.5) and (5.6) hold. We work in analogy with [38, pp. 129–132], [54], and the proof of Lemma 4.6.

Since both $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathcal{D}})$ represent a decomposition of $\mathcal{T}$, we have for vectors $\mathbf{v}$ and $\mathbf{w}$,

$$\begin{aligned} \mathcal{T} \bullet_1 \mathbf{v}^T \bullet_2 \mathbf{w}^T &= \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 (\mathbf{v}^T \mathbf{A}_r) \bullet_2 (\mathbf{w}^T \mathbf{B}_r) \bullet_3 \mathbf{C}_r \\ &= \sum_{r=1}^{R} \bar{\mathcal{D}}_r \bullet_1 (\mathbf{v}^T \bar{\mathbf{A}}_r) \bullet_2 (\mathbf{w}^T \bar{\mathbf{B}}_r) \bullet_3 \bar{\mathbf{C}}_r. \end{aligned} \tag{5.19}$$

Let the index functions $g(x)$ and $h(x)$ be given by $\mathbf{A}\mathbf{\Pi}_a = \begin{pmatrix} \mathbf{A}_{g(1)} & \mathbf{A}_{g(2)} & \dots & \mathbf{A}_{g(R)} \end{pmatrix}$ and $\mathbf{B}\mathbf{\Pi}_b = \begin{pmatrix} \mathbf{B}_{h(1)} & \mathbf{B}_{h(2)} & \dots & \mathbf{B}_{h(R)} \end{pmatrix}$, respectively. Then (5.19) can be written as

$$\mathbf{C} \cdot \mathbf{p} = \bar{\mathbf{C}} \cdot \mathbf{q}, \tag{5.20}$$

in which $\mathbf{p}$ and $\mathbf{q}$ are defined by

$$\mathbf{p} = \begin{pmatrix} (\mathcal{D}_1)_{N \times LM} \cdot [(\mathbf{A}_1^T \mathbf{v}) \otimes (\mathbf{B}_1^T \mathbf{w})] \\ \vdots \\ (\mathcal{D}_R)_{N \times LM} \cdot [(\mathbf{A}_R^T \mathbf{v}) \otimes (\mathbf{B}_R^T \mathbf{w})] \end{pmatrix},$$

$$\mathbf{q} = \begin{pmatrix} (\bar{\mathcal{D}}_1)_{N \times LM} \cdot [(\mathbf{\Lambda}_{a,1}^T \bar{\mathbf{A}}_{g(1)}^T \mathbf{v}) \otimes (\mathbf{\Lambda}_{b,1}^T \bar{\mathbf{B}}_{h(1)}^T \mathbf{w})] \\ \vdots \\ (\bar{\mathcal{D}}_R)_{N \times LM} \cdot [(\mathbf{\Lambda}_{a,R}^T \bar{\mathbf{A}}_{g(R)}^T \mathbf{v}) \otimes (\mathbf{\Lambda}_{b,R}^T \bar{\mathbf{B}}_{h(R)}^T \mathbf{w})] \end{pmatrix},$$

where $\mathbf{\Lambda}_{a,r}$ and $\mathbf{\Lambda}_{b,r}$ denote the rth block of $\mathbf{\Lambda}_a$ and $\mathbf{\Lambda}_b$, respectively.

We will now show by contradiction that $\mathbf{\Pi}_a = \mathbf{\Pi}_b$. If $\mathbf{\Pi}_a \neq \mathbf{\Pi}_b$, then we will be able to find vectors $\mathbf{v}$ and $\mathbf{w}$ such that $\mathbf{q} = \mathbf{0}$ and $\mathbf{p} \neq \mathbf{0}$ has less than $k'_{\mathbf{C}}$ nonzero $(N \times 1)$ subvectors. This implies that a set of less than $k'_{\mathbf{C}}$ vectors, each sampled in the column space of a different submatrix of $\mathbf{C}$, is linearly dependent, which contradicts the definition of $k'_{\mathbf{C}}$.

Suppose that $\mathbf{\Pi}_a \neq \mathbf{\Pi}_b$. Then there exists an r such that $\mathbf{A}_r$ is the sth submatrix of $\mathbf{A}\mathbf{\Pi}_a$, $\mathbf{B}_r$ is the tth submatrix of $\mathbf{B}\mathbf{\Pi}_b$, and $s \neq t$. Formally, there exists an r such that $r = g(s) = h(t)$ and $s \neq t$. We now create two index sets $\mathsf{S}, \mathsf{T} \subset \{1, \dots, R\}$ in the same way as in the proof of Lemma 4.6.

Since $k'_{\mathbf{A}} - 1 \leqslant R - 1$, S contains exactly $k'_{\mathbf{A}} - 1$ elements. The set T contains $R - \operatorname{card}(\mathsf{S}) = R - k'_{\mathbf{A}} + 1$ elements. Because of (5.6) and $k'_{\mathbf{C}} \leqslant R$, this is less than or equal to $k'_{\mathbf{B}} - 1$ elements. In the xth element of $\mathbf{q}$ we have either $g(x) \in \mathsf{S}$ or $h(x) \in \mathsf{T}$, $x = 1, \dots, R$. The index $r = g(s) = h(t)$ is neither an element of S nor an element of T. Denote $\{i_1, i_2, \dots, i_{k'_{\mathbf{A}}-1}\} = \mathsf{S}$ and $\{j_1, j_2, \dots, j_{R-k'_{\mathbf{A}}+1}\} = \mathsf{T}$.

We choose a vector $\mathbf{v}$ such that $\mathbf{v}^T \mathbf{A}_i = \mathbf{0}$ if $i \in \mathsf{S}$, and $\mathbf{v}^T \mathbf{A}_r \neq 0$. This is always possible. The vector $\mathbf{v}$ has to be chosen in $\operatorname{null}([\mathbf{A}_{i_1} \ \dots \ \mathbf{A}_{i_{k'_{\mathbf{A}}-1}}])$, which is an $(I - (k'_{\mathbf{A}} - 1)L)$-dimensional space. If a column of $\mathbf{A}_r$ is orthogonal to all possible vectors $\mathbf{v}$, then it lies in $\operatorname{span}([\mathbf{A}_{i_1} \ \dots \ \mathbf{A}_{i_{k'_{\mathbf{A}}-1}}])$. Then we would have a contradiction with the definition of $k'_{\mathbf{A}}$. Similarly, we can choose a vector $\mathbf{w}$ such that $\mathbf{w}^T \mathbf{B}_j = \mathbf{0}$ if $j \in \mathsf{T}$, and $\mathbf{w}^T \mathbf{B}_r \neq 0$.

Because of condition (5.5), the genericity of $\mathcal{D}_r$, and the fact that $\mathbf{v}^T \mathbf{A}_r \neq 0$, the $(N \times M)$ matrix $\mathcal{D}_r \bullet_1 (\mathbf{v}^T \mathbf{A}_r)$ is rank-M. Rank deficiency would imply that $N - M + 1$

determinants are zero, while $\mathbf{v}^T\mathbf{A}_r$ provides only $L-1$ parameters and an irrelevant scaling factor. Since $\mathcal{D}_r \bullet_1 (\mathbf{v}^T\mathbf{A}_r)$ is full column rank, and since $\mathbf{w}^T\mathbf{B}_r \neq 0$, we have $\mathcal{D}_r \bullet_1 (\mathbf{v}^T\mathbf{A}_r) \bullet_2 (\mathbf{w}^T\mathbf{B}_r) \neq \mathbf{0}$. Equivalently, $(\mathcal{D}_r)_{N\times LM} \cdot [(\mathbf{A}_r^T\mathbf{v}) \otimes (\mathbf{B}_r^T\mathbf{w})] \neq \mathbf{0}$.

Define $\mathsf{S}^c = \{1,\dots,R\} \setminus \mathsf{S}$ and $\mathsf{T}^c = \{1,\dots,R\} \setminus \mathsf{T}$. The number of nonzero subvectors of $\mathbf{p}$ is bounded from above by

$$\mathrm{card}(\mathsf{S}^c \cap \mathsf{T}^c) \leqslant \mathrm{card}(\mathsf{S}^c) \leqslant R - k'_{\mathbf{A}} + 1 \leqslant k'_{\mathbf{C}} - 1,$$

where the last inequality is due to (5.6) and $k'_{\mathbf{B}} \leqslant R$. Hence, $\mathbf{C}\cdot\mathbf{p} = \mathbf{0}$ implies that a set of less than $k'_{\mathbf{C}}$ columns, each sampled in the column space of a different submatrix of $\mathbf{C}$, is linearly dependent, which contradicts the definition of $k'_{\mathbf{C}}$. This completes the proof. ∎

THEOREM 5.5. *Let* $(\mathbf{A},\mathbf{B},\mathbf{C},\mathcal{D})$ *represent a decomposition of* $\mathcal{T}$ *in* R *rank-*(L,M,N) *terms. Suppose that we have an alternative decomposition of* $\mathcal{T}$*, represented by* $(\bar{\mathbf{A}},\bar{\mathbf{B}},\bar{\mathbf{C}},\bar{\mathcal{D}})$*, with* $k'_{\bar{\mathbf{A}}}$ *and* $k'_{\bar{\mathbf{B}}}$ *maximal under the given dimensionality constraints. For generic* $\mathcal{D}$ *there holds that, if conditions* (5.5) *and* (5.6) *hold, then* $(\mathbf{A},\mathbf{B},\mathbf{C},\mathcal{D})$ *and* $(\bar{\mathbf{A}},\bar{\mathbf{B}},\bar{\mathbf{C}},\bar{\mathcal{D}})$ *are essentially equal.*

Proof. From Lemma 5.4 we have that $\bar{\mathbf{A}} = \mathbf{A}\cdot\mathbf{\Pi}\cdot\mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B}\cdot\mathbf{\Pi}\cdot\mathbf{\Lambda}_b$. Put the submatrices of $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in the same order as the submatrices of $\mathbf{A}$ and $\mathbf{B}$. After reordering, we have $\bar{\mathbf{A}} = \mathbf{A}\cdot\mathbf{\Lambda}_a$, with $\mathbf{\Lambda}_a = \mathrm{blockdiag}(\mathbf{\Lambda}_{a,1},\dots,\mathbf{\Lambda}_{a,R})$, and $\bar{\mathbf{B}} = \mathbf{B}\cdot\mathbf{\Lambda}_b$, with $\mathbf{\Lambda}_b = \mathrm{blockdiag}(\mathbf{\Lambda}_{b,1},\dots,\mathbf{\Lambda}_{b,R})$. From (2.14) we have that

$$\begin{aligned}\mathbf{T}_{IJ\times K} &= (\mathbf{A}\odot\mathbf{B})\cdot\mathrm{blockdiag}((\mathcal{D}_1)_{LM\times N},\dots,(\mathcal{D}_R)_{LM\times N})\cdot\mathbf{C}^T\\ &= (\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\cdot\mathrm{blockdiag}((\bar{\mathcal{D}}_1)_{LM\times N},\dots,(\bar{\mathcal{D}}_R)_{LM\times N})\cdot\bar{\mathbf{C}}^T\\ &= (\mathbf{A}\odot\mathbf{B})\cdot\mathrm{blockdiag}((\mathbf{\Lambda}_{a,1}\otimes\mathbf{\Lambda}_{b,1})\cdot(\bar{\mathcal{D}}_1)_{LM\times N},\dots,\\ &\qquad (\mathbf{\Lambda}_{a,R}\otimes\mathbf{\Lambda}_{b,R})\cdot(\bar{\mathcal{D}}_R)_{LM\times N})\cdot\bar{\mathbf{C}}^T.\end{aligned}\tag{5.21}$$

From Lemma 3.4 we have that $k'_{\mathbf{A}\odot\mathbf{B}} \geqslant \min(k'_{\mathbf{A}} + k'_{\mathbf{B}} - 1, R)$. From (5.6) we have that $k'_{\mathbf{A}} + k'_{\mathbf{B}} - 1 \geqslant 2R + 1 - k'_{\mathbf{C}} \geqslant R+1$. Hence, $k'_{\mathbf{A}\odot\mathbf{B}} = R$, which implies that $\mathbf{A}\odot\mathbf{B}$ is full column rank. Multiplying (5.21) from the left by $(\mathbf{A}\odot\mathbf{B})^\dagger$, we obtain that

$$(\mathbf{\Lambda}_{a,r}\otimes\mathbf{\Lambda}_{b,r})\cdot(\bar{\mathcal{D}}_r)_{LM\times N}\cdot\bar{\mathbf{C}}_r^T = (\mathcal{D}_r)_{LM\times N}\cdot\mathbf{C}_r^T, \quad 1\leqslant r\leqslant R.$$

This can be rewritten as

$$\bar{\mathcal{D}}_r \bullet_1 \mathbf{\Lambda}_{a,r} \bullet_2 \mathbf{\Lambda}_{b,r} \bullet_3 \bar{\mathbf{C}}_r = \mathcal{D}_r \bullet_3 \mathbf{C}_r, \quad 1\leqslant r\leqslant R.$$

This means that $(\mathbf{A},\mathbf{B},\mathbf{C},\mathcal{D})$ and $(\bar{\mathbf{A}},\bar{\mathbf{B}},\bar{\mathbf{C}},\bar{\mathcal{D}})$ are equal up to trivial indeterminacies. ∎

5.2. Rank-$(2,2,2)$ blocks. In the Kruskal-type results of the previous section, we have only considered rank-(L,M,N) terms for which $N > L+M-2$. Rank-$(2,2,3)$ terms, for instance, satisfy this condition. However, it would also be interesting to know whether the decomposition of a tensor in rank-$(2,2,2)$ terms is essentially unique. This special case is studied in this section.

A first result is that in $\mathbb{C}$ the decomposition of a tensor $\mathcal{T}$ in $R \geqslant 2$ rank-$(2,2,2)$ terms is not essentially unique. This is easy to understand. Assume, for instance, that $\mathcal{T}$ is the sum of two rank-$(2,2,2)$ terms $\mathcal{T}_1$ and $\mathcal{T}_2$. It is well known that in $\mathbb{C}$ the rank of rank-$(2,2,2)$ tensor is always equal to 2 [55]. Hence we have for some vectors $\mathbf{a}_r$,

$\mathbf{b}_r$, $\mathbf{c}_r$, $1 \leqslant r \leqslant 4$,

$$\begin{aligned}\mathcal{T} &= \mathcal{T}_1 + \mathcal{T}_2 \\ &= (\mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2) + (\mathbf{a}_3 \circ \mathbf{b}_3 \circ \mathbf{c}_3 + \mathbf{a}_4 \circ \mathbf{b}_4 \circ \mathbf{c}_4) \\ &= (\mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_3 \circ \mathbf{b}_3 \circ \mathbf{c}_3) + (\mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2 + \mathbf{a}_4 \circ \mathbf{b}_4 \circ \mathbf{c}_4) \\ &= \tilde{\mathcal{T}}_1 + \tilde{\mathcal{T}}_2.\end{aligned}$$

Since $\tilde{\mathcal{T}}_1$ and $\tilde{\mathcal{T}}_2$ yield an other decomposition, the decomposition of $\mathcal{T}$ in 2 rank-$(2,2,2)$ terms is not essentially unique.

Theorem 5.5 does not hold in the case of rank-$(2,2,2)$ terms because Lemma 5.2 does not hold. The problem is that in (5.8) the (2×2) matrices $\mathcal{D}_r \times_1 (\mathbf{x}^T \mathbf{A}_r)$ are not necessarily rank-2. Indeed, let λ be a generalized eigenvalue of the pencil formed by the (2×2) matrices $(\mathcal{D}_r)_{1,:,:}$ and $(\mathcal{D}_r)_{2,:,:}$. Then $\mathcal{D}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r)$ is rank-1 if $\mathbf{x}^T \mathbf{A}_r$ is proportional to $(1, -\lambda)$. As a result, (5.12) does not hold.

On the other hand, if we work in $\mathbb{R}$, the situation is somewhat different. In $\mathbb{R}$, rank-$(2,2,2)$ terms can be either rank-2 or rank-3 [30, 39, 55]. If $\mathcal{D}_r$ is rank-2 in $\mathbb{R}$, then the pencil $((\mathcal{D}_r)_{1,:,:}, (\mathcal{D}_r)_{2,:,:})$ has two real generalized eigenvalues. Conversely, if the generalized eigenvalues of $((\mathcal{D}_r)_{1,:,:}, (\mathcal{D}_r)_{2,:,:})$ are complex, then $\mathcal{D}_r$ is rank-3. (The tensor $\mathcal{D}_r$ can also be rank-3 when an eigenvalue has algebraic multiplicity two but geometric multiplicity one. This case occurs with probability zero when the entries of $\mathcal{D}_r$ are drawn from continuous probability density functions and will not further be considered in this section.) We now have the following variant of Theorem 5.5.

THEOREM 5.6. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *represent a real decomposition of* $\mathcal{T} \in \mathbb{R}^{I \times J \times K}$ *in* R *rank-*$(2,2,2)$ *terms. Suppose that the condition*

$$k'_{\mathbf{A}} + k'_{\mathbf{B}} + k'_{\mathbf{C}} \geqslant 2R + 2$$

holds and that the generalized eigenvalues of $((\mathcal{D}_r)_{1,:,:}, (\mathcal{D}_r)_{2,:,:})$ *are complex,* $1 \leqslant r \leqslant R$*. Then* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *is essentially unique.*

Proof. Under the condition on the generalized eigenvalues of $((\mathcal{D}_r)_{1,:,:}, (\mathcal{D}_r)_{2,:,:})$, the matrices $\mathcal{D}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r)$ in (5.8) are necessarily rank-2, and the reasoning in the proof of Lemma 5.2 remains valid.

On the other hand, assuming that $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi}_b \cdot \mathbf{\Lambda}_b$, only a technical modification of the proof of Lemma 5.4 is required to make sure that $\mathbf{\Pi}_a = \mathbf{\Pi}_b$ does hold. We only have to verify whether vectors $\mathbf{v}$ and $\mathbf{w}$ can be found such that $\mathbf{v}^T \mathbf{A}_i = \mathbf{0}$ if $i \in \mathsf{S}$, $\mathbf{w}^T \mathbf{B}_j = \mathbf{0}$ if $j \in \mathsf{T}$, and $(\mathcal{D}_r)_{N \times LM} \cdot [(\mathbf{A}_r^T \mathbf{v}) \otimes (\mathbf{B}_r^T \mathbf{w})] \neq \mathbf{0}$. Reasoning as in the proof of Lemma 5.4, we see that the constraint $\mathbf{v}^T \mathbf{A}_i = \mathbf{0}$, $i \in \mathsf{S}$, still leaves enough freedom for $\mathbf{v}^T \mathbf{A}_r$ to be any vector in $\mathbb{R}^2$. Equivalently, the constraint $\mathbf{w}^T \mathbf{B}_j = \mathbf{0}$, $j \in \mathsf{T}$, leaves enough freedom for $\mathbf{w}^T \mathbf{B}_r$ to be any vector in $\mathbb{R}^2$. We conclude that it is always possible to find the required vectors $\mathbf{v}$ and $\mathbf{w}$ if $\mathcal{D}_r \neq \mathcal{O}$.

Essential uniqueness of the overall tensor decomposition now follows from $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_a$ and $\bar{\mathbf{B}} = \mathbf{B} \cdot \mathbf{\Pi} \cdot \mathbf{\Lambda}_b$ in the same way as in the proof of Theorem 5.5. $\square$

From Theorem 5.6 follows that a generic decomposition in real rank-3 rank-$(2,2,2)$ terms is essentially unique provided,

$$\min\left(\left\lfloor \frac{I}{2} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{J}{2} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{K}{2} \right\rfloor, R\right) \geqslant 2R + 2.$$

Finally, we have the following variant of Theorem 5.1.

THEOREM 5.7. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *represent a real decomposition of* $\mathcal{T} \in \mathbb{R}^{I\times J\times K}$ *in* R *rank-*(L, M, N) *terms, with* $L = M = N = 2$. *Suppose that* $\text{rank}(\mathbf{A}) = 2R$, $\text{rank}(\mathbf{B}) = 2R$, $\text{rank}_{k'}(\mathbf{C}) \geqslant 1$ *and that all generalized eigenvalues of the pencil* $((\mathcal{D}_r)_{L\times M,1}, (\mathcal{D}_r)_{L\times M,2})$ *are complex,* $1 \leqslant r \leqslant R$. *Then* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{D})$ *is essentially unique.*

Proof. Consider two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^K$ for which $\mathbf{x}^T\mathbf{C}_r$ is not proportional to $\mathbf{y}^T\mathbf{C}_r$, $1 \leqslant r \leqslant R$. Since all matrices $\mathbf{C}_r$ are full column rank, this is the case for generic vectors $\mathbf{x}, \mathbf{y}$. Define $\mathbf{T}_1 = \sum_{k=1}^{K} x_k \mathbf{T}_{I\times J,k}$ and $\mathbf{T}_2 = \sum_{k=1}^{K} y_k \mathbf{T}_{I\times J,k}$. We have

$$\begin{aligned}\mathbf{T}_2 \cdot \mathbf{T}_1^\dagger = \mathbf{A} \cdot \text{blockdiag}\{&([\mathcal{D}_1 \bullet_3 (\mathbf{y}^T\mathbf{C}_1)] \cdot [\mathcal{D}_1 \bullet_3 (\mathbf{x}^T\mathbf{C}_1)]^\dagger, \ldots, \\ &[\mathcal{D}_R \bullet_3 (\mathbf{y}^T\mathbf{C}_R)] \cdot [\mathcal{D}_R \bullet_3 (\mathbf{x}^T\mathbf{C}_R)]^\dagger)\} \cdot \mathbf{A}^\dagger.\end{aligned}$$

From this equation it is clear that the column space of any $\mathbf{A}_r$ is an invariant subspace of $\mathbf{T}_2 \cdot \mathbf{T}_1^\dagger$.

Define $\mathbf{C}_r^T\mathbf{x} = \tilde{\mathbf{x}}_r$ and $\mathbf{C}_r^T\mathbf{y} = \tilde{\mathbf{y}}_r$. We have

$$\begin{aligned}\mathcal{D}_r \bullet_3 (\mathbf{x}^T\mathbf{C}_r) &= (\tilde{\mathbf{x}}_r)_1(\mathcal{D}_r)_{L\times M,1} + (\tilde{\mathbf{x}}_r)_2(\mathcal{D}_r)_{L\times M,2}, \\ \mathcal{D}_r \bullet_3 (\mathbf{y}^T\mathbf{C}_r) &= (\tilde{\mathbf{y}}_r)_1(\mathcal{D}_r)_{L\times M,1} + (\tilde{\mathbf{y}}_r)_2(\mathcal{D}_r)_{L\times M,2}.\end{aligned}$$

If there exist real values α and β, with $\alpha^2+\beta^2 = 1$, such that $\alpha\, \mathcal{D}_r \bullet_3 (\mathbf{x}^T\mathbf{C}_r) + \beta\, \mathcal{D}_r \bullet_3 (\mathbf{y}^T\mathbf{C}_r)$ is rank-1, then there also exist real values γ en μ, with $\gamma^2 + \mu^2 = 1$, such that $\gamma\, (\mathcal{D}_r)_{L\times M,1} + \mu\, (\mathcal{D}_r)_{L\times M,2}$ is rank-1. The condition on the generalized eigenvalues of the pencils $((\mathcal{D}_r)_{L\times M,1}, (\mathcal{D}_r)_{L\times M,2})$ implies thus that the blocks $[\mathcal{D}_r \bullet_3 (\mathbf{y}^T\mathbf{C}_r)] \cdot [\mathcal{D}_r \bullet_3 (\mathbf{x}^T\mathbf{C}_r)]^\dagger$ cannot be diagonalized by means of a real similarity transformation. We conclude that the only two-dimensional invariant subspaces of $\mathbf{T}_2 \cdot \mathbf{T}_1^\dagger$ are the column spaces of the matrices $\mathbf{A}_r$. In other words, $\mathbf{A}$ is essentially unique.

Essential uniqueness of the overall decomposition now follows from (2.12). Assume that we have an alternative decomposition of $\mathcal{T}$, represented by $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathcal{D}})$. We have $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$, in which $\mathbf{\Pi}_a$ is a block-permutation matrix and $\mathbf{\Lambda}_a = \text{blockdiag}(\mathbf{\Lambda}_{a,1}, \ldots, \mathbf{\Lambda}_{a,R})$ a square nonsingular block-diagonal matrix, compatible with the block structure of $\mathbf{A}$. From (2.12) we have

$$\begin{aligned}\mathbf{T}_{JK\times I} &= (\mathbf{B} \odot \mathbf{C}) \cdot \text{blockdiag}((\mathcal{D}_1)_{MN\times L}, \ldots, (\mathcal{D}_R)_{MN\times L}) \cdot \mathbf{A}^T \\ &= (\bar{\mathbf{B}} \odot \bar{\mathbf{C}}) \cdot \text{blockdiag}((\bar{\mathcal{D}}_1)_{MN\times L}, \ldots, (\bar{\mathcal{D}}_R)_{MN\times L}) \cdot \mathbf{\Pi}_a^T \cdot \mathbf{\Lambda}_a^T \cdot \mathbf{A}^T.\end{aligned}$$

Right multiplication by $(\mathbf{A}^T)^\dagger$ yields

$$\begin{aligned}&(\mathbf{B} \odot \mathbf{C}) \cdot \text{blockdiag}((\mathcal{D}_1)_{MN\times L}, \ldots, (\mathcal{D}_R)_{MN\times L}) \\ &= (\bar{\mathbf{B}} \odot \bar{\mathbf{C}}) \cdot \text{blockdiag}((\bar{\mathcal{D}}_1)_{MN\times L}, \ldots, (\bar{\mathcal{D}}_R)_{MN\times L}) \cdot \mathbf{\Pi}_a^T \cdot \mathbf{\Lambda}_a^T.\end{aligned} \tag{5.22}$$

Assume that the rth submatrix of $\mathbf{A}$ corresponds to the s-th submatrix of $\bar{\mathbf{A}}$. Then we have from (5.22) that

$$(\mathbf{B}_r \odot \mathbf{C}_r) \cdot (\mathcal{D}_r)_{MN\times L} = (\bar{\mathbf{B}}_s \odot \bar{\mathbf{C}}_s) \cdot (\bar{\mathcal{D}}_s)_{MN\times L} \cdot \mathbf{\Lambda}_{a,s}^T$$

in which $\mathbf{\Lambda}_{a,s}$ is the sth block of $\mathbf{\Lambda}_a$. Equivalently,

$$\mathcal{D}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r = \bar{\mathcal{D}}_s \bullet_1 \mathbf{\Lambda}_{a,s} \bullet_2 \bar{\mathbf{B}}_s \bullet_3 \bar{\mathbf{C}}_s.$$

This completes the proof. $\square$

6. Type-2 decomposition in rank-$(L, M, \cdot)$ terms. In this section we derive several conditions under which the type-2 decomposition in rank-$(L, M, \cdot)$ terms is unique. We use the notation introduced in section 2.3.

First we have a uniqueness result that stems from the fact that the column spaces of $\mathbf{A}_r$, $1 \leqslant r \leqslant R$, are invariant subspaces of quotients of tensor slices. This result is the counterpart of Theorem 4.1 in section 4 and Theorem 5.1 in section 5.1.

THEOREM 6.1. *Let* $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ *represent a type-*2 *decomposition of* $\mathcal{T} \in \mathbb{K}^{I\times J\times K}$ *in* R *rank-*$(L, L, \cdot)$ *terms. Suppose that* $\operatorname{rank}(\mathbf{A}) = LR$, $\operatorname{rank}(\mathbf{B}) = LR$, $K \geqslant 3$, *and that* $\mathcal{C}$ *is generic. Then* $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ *is essentially unique.*

Proof. We have

$$\mathbf{T}_{I\times J,2} \cdot \mathbf{T}^{\dagger}_{I\times J,1} = \mathbf{A} \cdot \operatorname{blockdiag}((\mathcal{C}_1)_{L\times M,2} \cdot (\mathcal{C}_1)^{\dagger}_{L\times M,1}, \dots, (\mathcal{C}_R)_{L\times M,2} \cdot (\mathcal{C}_R)^{\dagger}_{L\times M,1}) \cdot \mathbf{A}^{\dagger},$$

where $M = L$. From this equation it is clear that the column space of any $\mathbf{A}_r$ is an invariant subspace of $\mathbf{T}_{I\times J,2} \cdot \mathbf{T}^{\dagger}_{I\times J,1}$. However, any set of eigenvectors forms an invariant subspace. To determine which eigenvectors belong together, we use the third slice $\mathbf{T}_{I\times J,3}$. We have

$$\mathbf{T}_{I\times J,3} \cdot \mathbf{T}^{\dagger}_{I\times J,1} = \mathbf{A} \cdot \operatorname{blockdiag}((\mathcal{C}_1)_{L\times M,3} \cdot (\mathcal{C}_1)^{\dagger}_{L\times M,1}, \dots, (\mathcal{C}_R)_{L\times M,3} \cdot (\mathcal{C}_R)^{\dagger}_{L\times M,1}) \cdot \mathbf{A}^{\dagger}. \tag{6.1}$$

It is clear that the column space of any $\mathbf{A}_r$ is also an invariant subspace of $\mathbf{T}_{I\times J,3} \cdot \mathbf{T}^{\dagger}_{I\times J,1}$. On the other hand, because of the genericity of $\mathcal{C}$, we can interpret $(\mathcal{C}_r)_{L\times M,3} \cdot (\mathcal{C}_r)^{\dagger}_{L\times M,1}$ as $(\mathcal{C}_r)_{L\times M,2} \cdot (\mathcal{C}_r)^{\dagger}_{L\times M,1} + \mathbf{E}_r$, in which $\mathbf{E}_r \in \mathbb{K}^{L\times L}$ is a generic perturbation, $1 \leqslant r \leqslant R$. Perturbation analysis now states that the individual eigenvectors of $\mathbf{T}_{I\times J,3} \cdot \mathbf{T}^{\dagger}_{I\times J,1}$ do not correspond to those of $\mathbf{T}_{I\times J,2} \cdot \mathbf{T}^{\dagger}_{I\times J,1}$ [23, 32]. We conclude that $\mathbf{A}$ is essentially unique.

Essential uniqueness of the overall decomposition follows directly from the essential uniqueness of $\mathbf{A}$. Assume that we have an alternative decomposition of $\mathcal{T}$, represented by $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathcal{C}})$. We have $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$, in which $\mathbf{\Pi}_a$ is a block-permutation matrix and $\mathbf{\Lambda}_a$ a square nonsingular block-diagonal matrix, compatible with the block structure of $\mathbf{A}$. From (2.18) we have

$$\begin{aligned}\mathbf{T}_{JK\times I} &= [(\mathcal{C}_1 \bullet_2 \mathbf{B}_1)_{JK\times L} \ \dots \ (\mathcal{C}_R \bullet_2 \mathbf{B}_R)_{JK\times L}] \cdot \mathbf{A}^T \\ &= \left[(\bar{\mathcal{C}}_1 \bullet_2 \bar{\mathbf{B}}_1)_{JK\times L} \ \dots \ (\bar{\mathcal{C}}_R \bullet_2 \bar{\mathbf{B}}_R)_{JK\times L}\right] \cdot \bar{\mathbf{A}}^T.\end{aligned}$$

Hence,

$$\begin{aligned}&[(\mathcal{C}_1 \bullet_2 \mathbf{B}_1)_{JK\times L} \ \dots \ (\mathcal{C}_R \bullet_2 \mathbf{B}_R)_{JK\times L}] \\ &\quad = \left[(\bar{\mathcal{C}}_1 \bullet_2 \bar{\mathbf{B}}_1)_{JK\times L} \ \dots \ (\bar{\mathcal{C}}_R \bullet_2 \bar{\mathbf{B}}_R)_{JK\times L}\right] \cdot \mathbf{\Lambda}_a^T \cdot \mathbf{\Pi}_a^T.\end{aligned}$$

This implies that the matrices $(\mathcal{C}_r \bullet_2 \mathbf{B}_r)_{JK\times L}$ are ordered in the same way as the matrices $\mathbf{A}_r$. Furthermore, if $\bar{\mathbf{A}}_i = \mathbf{A}_j \cdot \mathbf{F}$, then $(\bar{\mathcal{C}}_i \bullet_2 \bar{\mathbf{B}}_i)_{JK\times L} \cdot \mathbf{F}^T = (\mathcal{C}_j \bullet_2 \mathbf{B}_j)_{JK\times L}$. Equivalently, we have $\bar{\mathcal{C}}_i \bullet_2 \bar{\mathbf{B}}_i = \mathcal{C}_j \bullet_1 \mathbf{F}^{-1} \bullet_2 \mathbf{B}_j$. This means that $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathcal{C}})$ are essentially equal. ∎

Remark 6. The generalization to the decomposition in rank-$(L_r, L_r, \cdot)$ terms, $1 \leqslant r \leqslant R$, is trivial.

Remark 7. In the nongeneric case, lack of uniqueness can be due to the fact that tensors $\mathcal{C}_r$ can be subdivided in smaller blocks by means of basis transformations in their mode-1 and mode-2 vector space. We give an example.

Example 2. Consider a tensor $\mathcal{T} \in \mathbb{K}^{10\times 10\times 5}$ that can be decomposed in two rank-$(5,5,\cdot)$ terms as follows:

$$\mathcal{T} = \sum_{r=1}^{2} \mathcal{C}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r$$

with $\mathcal{C}_r \in \mathbb{K}^{5\times 5\times 5}$, $\mathbf{A}_r \in \mathbb{K}^{10\times 5}$, and $\mathbf{B}_r \in \mathbb{K}^{10\times 5}$, $1 \leqslant r \leqslant 2$. Now assume that $\mathcal{C}_1$ and $\mathcal{C}_2$ can be further decomposed as follows:

$$\mathcal{C}_1 = \mathcal{G}_{11} \bullet_1 \mathbf{E}_{11} \bullet_2 \mathbf{F}_{11} + \mathcal{G}_{12} \bullet_1 \mathbf{E}_{12} \bullet_2 \mathbf{F}_{12},$$
$$\mathcal{C}_2 = \mathcal{G}_{21} \bullet_1 \mathbf{E}_{21} \bullet_2 \mathbf{F}_{21} + \mathcal{G}_{22} \bullet_1 \mathbf{E}_{22} \bullet_2 \mathbf{F}_{22},$$

where $\mathcal{G}_{11}, \mathcal{G}_{21} \in \mathbb{K}^{2\times 2\times 5}$, $\mathcal{G}_{12}, \mathcal{G}_{22} \in \mathbb{K}^{3\times 3\times 5}$, $\mathbf{E}_{11}, \mathbf{E}_{21}, \mathbf{F}_{11}, \mathbf{F}_{21} \in \mathbb{K}^{5\times 2}$, $\mathbf{E}_{12}, \mathbf{E}_{22}$, $\mathbf{F}_{12}, \mathbf{F}_{22} \in \mathbb{K}^{5\times 3}$. Define

$$\tilde{\mathbf{A}}_1 = [\mathbf{A}_1 \cdot \mathbf{E}_{11}\ \mathbf{A}_2 \cdot \mathbf{E}_{22}], \qquad \tilde{\mathbf{A}}_2 = [\mathbf{A}_2 \cdot \mathbf{E}_{21}\ \mathbf{A}_1 \cdot \mathbf{E}_{12}],$$
$$\tilde{\mathbf{B}}_1 = [\mathbf{B}_1 \cdot \mathbf{F}_{11}\ \mathbf{B}_2 \cdot \mathbf{F}_{22}], \qquad \tilde{\mathbf{B}}_2 = [\mathbf{B}_2 \cdot \mathbf{F}_{21}\ \mathbf{B}_1 \cdot \mathbf{F}_{12}],$$

$$(\tilde{\mathcal{C}}_1)_{1:2,1:2,:} = \mathcal{G}_{11}, \quad (\tilde{\mathcal{C}}_1)_{3:5,3:5,:} = \mathcal{G}_{22}, \quad (\tilde{\mathcal{C}}_1)_{1:2,3:5,:} = \mathcal{O}, \quad (\tilde{\mathcal{C}}_1)_{3:5,1:2,:} = \mathcal{O},$$
$$(\tilde{\mathcal{C}}_2)_{1:2,1:2,:} = \mathcal{G}_{21}, \quad (\tilde{\mathcal{C}}_2)_{3:5,3:5,:} = \mathcal{G}_{12}, \quad (\tilde{\mathcal{C}}_2)_{1:2,3:5,:} = \mathcal{O}, \quad (\tilde{\mathcal{C}}_2)_{3:5,1:2,:} = \mathcal{O}.$$

Then an alternative decomposition of $\mathcal{T}$ in rank-$(5,5,\cdot)$ terms is given by

$$\mathcal{T} = \sum_{r=1}^{2} \tilde{\mathcal{C}}_r \bullet_1 \tilde{\mathbf{A}}_r \bullet_2 \tilde{\mathbf{B}}_r. \tag{6.2}$$

For the case in which $\mathcal{C}_r \in \mathbb{R}^{2\times 2\times 2}$, $1 \leqslant r \leqslant R$, we have the following theorem.

THEOREM 6.2. *Let $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ represent a real type-2 decomposition of $\mathcal{T} \in \mathbb{R}^{I\times J\times 2}$ in R rank-$(L,M,2)$ terms with $L = M = 2$. Suppose that* $\operatorname{rank}(\mathbf{A}) = 2R$, $\operatorname{rank}(\mathbf{B}) = 2R$ *and that all generalized eigenvalues of the pencil $((\mathcal{C}_r)_{L\times M,1}, (\mathcal{C}_r)_{L\times M,2})$ are complex, $1 \leqslant r \leqslant R$. Then $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ is essentially unique.*

Proof. This theorem is a special case of Theorem 5.7. The tensors $\mathcal{D}_r$ in Theorem 5.7 correspond to $\mathcal{C}_r$, and the matrices $\mathbf{C}_r$ in Theorem 5.7 are equal to $\mathbf{I}_{2\times 2}$. ∎

In some cases, uniqueness of the decomposition can be demonstrated by direct application of the equivalence lemma for partitioned matrices. This is illustrated in the following example.

Example 3. We show that the decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{5\times 6\times 6}$ in $R = 3$ generic rank-$(2,2,\cdot)$ terms is essentially unique. Denote $I = 5$, $J = K = 6$, and $L = M = 2$. Let the decomposition be represented by $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ and let us assume the existence of an alternative decomposition, represented by $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathcal{C}})$, which is "nonsingular" in the sense that the columns of $\bar{\mathbf{A}}$ are as linearly independent as possible.

To show that $(\mathbf{A}, \mathbf{B}, \mathcal{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathcal{C}})$ are essentially equal, we first use the equivalence lemma for partitioned matrices to show that $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$, in which $\mathbf{\Pi}_a$ is a block permutation matrix and $\mathbf{\Lambda}_a$ a square nonsingular block-diagonal matrix, both consisting of (2×2) blocks. We show that for every $\mu \leqslant R - k'_{\bar{\mathbf{A}}} + 1 = 2$ there holds that for a generic vector $\mathbf{x} \in \mathbb{K}^5$ such that $\omega'(\mathbf{x}^T \bar{\mathbf{A}}) \leqslant \mu$, we have $\omega'(\mathbf{x}^T \mathbf{A}) \leqslant \omega'(\mathbf{x}^T \bar{\mathbf{A}})$. We will subsequently examine the different cases corresponding to $\mu = 0, 1, 2$.

We first derive an inequality that will prove useful. Denote by $(\mathbf{C}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r))_{M\times K}$ the $(M \times K)$ matrix formed by the single slice of $\mathbf{C}_r \bullet_1 (\mathbf{x}^T \mathbf{A}_r)$, and denote by

$(\bar{\mathbf{C}}_r \bullet_1 (\mathbf{x}^T\bar{\mathbf{A}}_r))_{M\times K}$ the $(M\times K)$ matrix formed by the single slice of $\bar{\mathbf{C}}_r \bullet_1 (\mathbf{x}^T\bar{\mathbf{A}}_r)$, $1 \leqslant r \leqslant R$. Then the $(J\times K)$ matrix formed by the single slice of $\mathcal{T} \bullet_1 \mathbf{x}^T$ is given by

$$\bar{\mathbf{B}} \cdot \begin{pmatrix} (\bar{\mathbf{C}}_1 \bullet_1 (\mathbf{x}^T\bar{\mathbf{A}}_1))_{M\times K} \\ \vdots \\ (\bar{\mathbf{C}}_R \bullet_1 (\mathbf{x}^T\bar{\mathbf{A}}_R))_{M\times K} \end{pmatrix} = \mathbf{B} \cdot \begin{pmatrix} (\mathbf{C}_1 \bullet_1 (\mathbf{x}^T\mathbf{A}_1))_{M\times K} \\ \vdots \\ (\mathbf{C}_R \bullet_1 (\mathbf{x}^T\mathbf{A}_R))_{M\times K} \end{pmatrix}.$$

For the rank of this matrix, we have

$$\begin{aligned} M\omega'(\mathbf{x}^T\bar{\mathbf{A}}) &\geqslant \operatorname{rank}\left[\bar{\mathbf{B}} \cdot \begin{pmatrix} (\bar{\mathbf{C}}_1 \bullet_1 (\mathbf{x}^T\bar{\mathbf{A}}_1))_{M\times K} \\ \vdots \\ (\bar{\mathbf{C}}_R \bullet_1 (\mathbf{x}^T\bar{\mathbf{A}}_R))_{M\times K} \end{pmatrix}\right] \\ &= \operatorname{rank}\left[\mathbf{B} \cdot \begin{pmatrix} (\mathbf{C}_1 \bullet_1 (\mathbf{x}^T\mathbf{A}_1))_{M\times K} \\ \vdots \\ (\mathbf{C}_R \bullet_1 (\mathbf{x}^T\mathbf{A}_R))_{M\times K} \end{pmatrix}\right]. \end{aligned}$$

Let $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{D}}(\mathbf{x})^T$ consist of the submatrices of $\mathbf{B}$ and

$$\begin{pmatrix} (\mathbf{C}_1 \bullet_1 (\mathbf{x}^T\mathbf{A}_1))_{M\times K} \\ \vdots \\ (\mathbf{C}_R \bullet_1 (\mathbf{x}^T\mathbf{A}_R))_{M\times K} \end{pmatrix},$$

respectively, corresponding to the nonzero subvectors of $\mathbf{x}^T\mathbf{A}$. Then we have

$$M\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \geqslant r_{\tilde{\mathbf{B}}\cdot\tilde{\mathbf{D}}(\mathbf{x})^T}.$$

Since $\mathbf{B}$ is generic, we have

$$M\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \geqslant r_{\tilde{\mathbf{D}}(\mathbf{x})^T}. \tag{6.3}$$

First, note that due to the "nonsingularity" of $\bar{\mathbf{A}}$, there does not exist a vector $\mathbf{x}$ such that $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) = 0$. This means that the case $\mu = 0$ does not present a difficulty.

Next, we consider the case $\mu = 1$. Since $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant \mu$, we have that $M\omega'(\mathbf{x}^T\bar{\mathbf{A}})$ in (6.3) is less than or equal to 2. Since $\mathbf{x}$ is orthogonal to two submatrices of $\bar{\mathbf{A}}$, the set V of vectors $\mathbf{x}$ satisfying $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant \mu$ is the union of three one-dimensional subspaces in $\mathbb{K}^5$. We prove by contradiction that for a generic $\mathbf{x} \in \mathsf{V}$, we have $\omega'(\mathbf{x}^T\mathbf{A}) \leqslant 1$. Assume first that $\omega'(\mathbf{x}^T\mathbf{A}) = 2$. Then $\tilde{\mathbf{D}}(\mathbf{x})$ in (6.3) is a (6×4) matrix. For this (6×4) matrix to be rank-2, eight independent conditions on $\mathbf{x}$ have to be satisfied. (This value is the total number of entries (i.e., 24) minus the number of independent parameters in a (6×4) rank-2 matrix (i.e., 16). The latter value can easily be determined as the number of independent parameters in, for instance, an SVD.) These conditions can impossibly be satisfied in a subset of V that is not of measure zero. We conclude that for a generic $\mathbf{x} \in \mathsf{V}$, $\omega'(\mathbf{x}^T\mathbf{A}) \neq 2$. Next assume that $\omega'(\mathbf{x}^T\mathbf{A}) = 3$. Then $\tilde{\mathbf{D}}(\mathbf{x})$ in (6.3) is a (6×6) matrix. For this matrix to be rank-2, $36 - 20 = 10$ independent conditions on $\mathbf{x}$ have to be satisfied. We conclude that for a generic $\mathbf{x}$, $\omega'(\mathbf{x}^T\mathbf{A}) \neq 3$. This completes the case $\mu = 1$.

Finally, we consider the case $\mu = 2$. We now have that $M\omega'(\mathbf{x}^T\bar{\mathbf{A}})$ in (6.3) is less than or equal to 4. Since $\mathbf{x}$ is orthogonal to one submatrix of $\bar{\mathbf{A}}$, the set V of vectors $\mathbf{x}$

satisfying $\omega'(\mathbf{x}^T\bar{\mathbf{A}}) \leqslant \mu$ is the union of three three-dimensional subspaces in $\mathbb{K}^5$. We prove by contradiction that for a generic $\mathbf{x} \in \mathsf{V}$, we have $\omega'(\mathbf{x}^T\mathbf{A}) \leqslant 2$. Assume that $\omega'(\mathbf{x}^T\mathbf{A}) = 3$. Then $\tilde{\mathbf{D}}(\mathbf{x})$ in (6.3) is a (6×6) matrix. For this matrix to be rank-4, $36 - 32 = 4$ independent conditions on $\mathbf{x}$ have to be satisfied. These conditions can impossibly be satisfied in a subset of V that is not of measure zero. This completes the case $\mu = 2$.

We conclude that the condition of the equivalence lemma for partitioned matrices is satisfied. Hence, $\bar{\mathbf{A}} = \mathbf{A} \cdot \mathbf{\Pi}_a \cdot \mathbf{\Lambda}_a$. Essential uniqueness of the decomposition follows directly from the essential uniqueness of $\mathbf{A}$; cf. the proof of Theorem 6.1.

7. Discussion and future research. In this paper we introduced the concept of block term decompositions. A block term decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ decomposes the given $(I \times J \times K)$-dimensional block in a number of blocks of smaller size. The size of a block is characterized by its mode-n rank triplet. (We mean the following. Consider a rank-(L, M, N) tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$. The observed dimensions of $\mathcal{T}$ are I, J, K. However, its inner dimensions, its inherent size, are given by L, M, N.) The number of blocks that are needed in a decomposition depends on the size of the blocks. On the other hand, the number of blocks that is allowed determines which size they should minimally be.

The concept of block term decompositions unifies HOSVD/Tucker's decomposition and CANDECOMP/PARAFAC. HOSVD is a meaningful representation of a rank-(L, M, N) tensor as a single block of size (L, M, N). PARAFAC decomposes a rank-R tensor in R scalar blocks.

In the case of matrices, column rank and row rank are equal; moreover, they are equal to the minimal number of rank-1 terms in which the matrix can be decomposed. This is a consequence of the fact that matrices can be diagonalized by means of basis transformations in their column and row space. On the other hand, tensors cannot in general be diagonalized by means of basis transformations in their mode-1, mode-2, and mode-3 vector space. This has led to the distinction between mode-n rank triplet and rank. Like HOSVD and PARAFAC, these are the two extrema in a spectrum. It is interesting to note that "the" rank of a higher-order tensor is actually a combination of the two aspects: one should specify the number of blocks *and* their size. This is not clear at the matrix level because of the lack of uniqueness of decompositions in nonscalar blocks.

Matrices can actually be diagonalized by means of orthogonal (unitary) basis transformations in their column and row space. On the other hand, by imposing orthogonality constraints on PARAFAC one obtains different (approximate) decompositions, with different properties [8, 35, 36, 42]. Generalizations to block decompositions can easily be formulated. For instance, the generalization of [8, 42] to decompositions in rank-(L, M, N) terms is simply obtained by claiming that $\mathbf{A}_r^H \cdot \mathbf{A}_s = \mathbf{0}_{L \times L}$, $\mathbf{B}_r^H \cdot \mathbf{B}_s = \mathbf{0}_{M \times M}$, and $\mathbf{C}_r^H \cdot \mathbf{C}_s = \mathbf{0}_{N \times N}$, $1 \leqslant r \neq s \leqslant R$.

Interestingly enough, the generalization of different aspects of the matrix SVD most often leads to different tensor decompositions. Although the definition of block term decompositions is very general, tensor SVDs that do not belong to this class do exist. For instance, a variational definition of singular values and singular vectors was generalized in [41]. Although Tucker's decomposition and the best rank-(L, M, N) approximation can be obtained by means of a variational approach [13, 15, 61], the general theory does not fit in the framework of block decompositions.

Block term decompositions have an interesting interpretation in terms of the decomposition of homogeneous polynomials or multilinear forms. The PARAFAC decomposition of a fully symmetric tensor (i.e., a tensor that is invariant under arbitrary index permutations) can be interpreted in terms of the decomposition of the associated homogeneous polynomial (quantic) in a sum of powers of linear forms [9]. For block term decompositions we now have the following. Given the quantic, linear forms are defined and clustered in subsets. Only within the same subset, products are admissible. The block term decomposition then decomposes the quantic in a sum of admissible products.

For instance, let $\mathcal{P} \in \mathbb{K}^{I\times I\times I}$ be fully symmetric. Let $\mathbf{x} \in \mathbb{K}^I$ be a vector of unknowns. Associate the quantic $p(\mathbf{x}) = \mathcal{P} \bullet_1 \mathbf{x}^T \bullet_2 \mathbf{x}^T \bullet_3 \mathbf{x}^T$ to $\mathcal{P}$. Let a PARAFAC decomposition of $\mathcal{P}$ be given by

$$\mathcal{P} = \sum_{r=1}^{R} d_r\, \mathbf{a}_r \circ \mathbf{a}_r \circ \mathbf{a}_r.$$

Define $y_r = \mathbf{x}^T \mathbf{a}_r$, $1 \leqslant r \leqslant R$. Then the quantic can be written as

$$p(\mathbf{y}) = \sum_{r=1}^{R} d_r\, y_r^3.$$

On the other hand, let a decomposition of $\mathcal{P}$ in rank-(L_r, L_r, L_r) terms be given by

$$\mathcal{P} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{A}_r \bullet_3 \mathbf{A}_r,$$

in which $\mathcal{D}_r \in \mathbb{K}^{L_r\times L_r\times L_r}$ and $\mathbf{A}_r \in \mathbb{K}^{I\times L_r}$, $1 \leqslant r \leqslant R$. Define $y_{lr} = \mathbf{x}^T(\mathbf{A}_r)_{:,l}$, $1 \leqslant l \leqslant L_r$, $1 \leqslant r \leqslant R$. Then the quantic can be written as

$$p(\mathbf{y}) = \sum_{r=1}^{R} \sum_{l_1,l_2,l_3=1}^{L_r} (\mathcal{D}_r)_{l_1 l_2 l_3}\, y_{l_1 r} y_{l_2 r} y_{l_3 r}.$$

In this paper we have presented EVD-based and Kruskal-type conditions guaranteeing essential uniqueness of the decompositions. Important work that remains to be done is the relaxation of the dimensionality constraints on the blocks in the Kruskal-type conditions. Some results based on simultaneous matrix diagonalization are presented in [44]. Also, we have restricted our attention to alternative decompositions that are "nonsingular." We should now check whether, for generic block terms, alternative decompositions in singular terms can exist.

It would be interesting to investigate, given the tensor dimensions I, J, and K, for which block sizes and number of blocks one obtains a generic (in the sense of existing with probability one) or a typical (in the sense of existing with probability different from zero) decomposition. In the context of PARAFAC, generic and typical rank have been studied in [55, 56, 57, 58].

In this paper we limited ourselves to the study of some algebraic aspects of block term decompositions. The computation of the decompositions, by means of alternating least squares algorithms, is addressed in [20]. Some applications are studied in [21, 43, 45].

Acknowledgments. The author wishes to thank A. Stegeman (Heijmans Institute, The Netherlands) and N. Sidiropoulos (Technical University of Crete, Greece) for sharing the manuscript of [54] before its publication. The author also wishes to thank A. Stegeman for proofreading an early version of the manuscript. A large part of this research was carried out when L. De Lathauwer was with the French Centre National de la Recherche Scientifique (C.N.R.S.).

REFERENCES

[1] C.J. Appellof and E.R. Davidson, *Strategies for analyzing data from video fluoromatric monitoring of liquid chromatographic effluents*, Anal. Chemistry, 53 (1981), pp. 2053–2056.
[2] G. Beylkin and M.J. Mohlenkamp, *Numerical operator calculus in higher dimensions*, Proc. Natl. Acad. Sci. USA, 99 (2002), pp. 10246–10251.
[3] G. Beylkin and M.J. Mohlenkamp, *Algorithms for numerical analysis in high dimensions*, SIAM J. Sci. Comput., 26 (2005), pp. 2133–2159.
[4] G. Boutry, M. Elad, G.H. Golub, and P. Milanfar, *The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 582–601.
[5] R. Bro, *PARAFAC. Tutorial and Applications*, Chemom. Intell. Lab. Syst., 38 (1997), pp. 149–171.
[6] D. Burdick, X. Tu, L. McGown, and D. Millican, *Resolution of multicomponent fluorescent mixtures by analysis of the excitation-emission-frequency array*, J. Chemometrics, 4 (1990), pp. 15–28.
[7] J. Carroll and J. Chang, *Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition*, Psychometrika, 9 (1970), pp. 267–283.
[8] P. Comon, *Independent component analysis, a new concept?* Signal Process., 36 (1994), pp. 287–314.
[9] P. Comon and B. Mourrain, *Decomposition of quantics in sums of powers of linear forms*, Signal Process., 53 (1996), pp. 93–108.
[10] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.
[11] L. De Lathauwer, *Signal Processing Based on Multilinear Algebra*, Ph.D. thesis, K.U.Leuven, Belgium, 1997.
[12] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.
[13] L. De Lathauwer, B. De Moor, and J. Vandewalle, *On the best rank-*1 *and rank-*$(R_1, R_2, \ldots, R_N)$ *approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1324–1342.
[14] L. De Lathauwer, B. De Moor, and J. Vandewalle, *An introduction to independent component analysis*, J. Chemometrics, 14 (2000), pp. 123–149.
[15] L. De Lathauwer and J. Vandewalle, *Dimensionality reduction in higher-order signal processing and rank-*$(R_1, R_2, \ldots, R_N)$ *reduction in multilinear algebra*, Linear Algebra Appl., 391 (2004), pp. 31–55.
[16] L. De Lathauwer, B. De Moor, and J. Vandewalle, *Computation of the Canonical Decomposition by means of a simultaneous generalized Schur decompositition*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 295–327.
[17] 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.
[18] L. De Lathauwer and J. Castaing, *Tensor-based techniques for the blind separation of DS-CDMA signals*, Signal Process., 87 (2007), pp. 322–336.
[19] L. De Lathauwer, *Decompositions of a higher-order tensor in block terms—Part* I: *Lemmas for partitioned matrices*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1022–1032.
[20] L. De Lathauwer and D. Nion, *Decompositions of a higher-order tensor in block terms—Part* III: *Alternating Least Squares algorithms*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1067–1083.
[21] L. De Lathauwer and A. de Baynast, *Blind deconvolution of DS-CDMA Signals by means of decomposition in rank-*$(1, L, L)$ *terms*, IEEE Trans. Signal Process., 56 (2008), pp. 1562–1571.
[22] M. Elad, P. Milanfar, and G.H. Golub, *Shape from moments—an estimation theory perspective*, IEEE Trans. Signal Process., 52 (2004), pp. 1814–1829.
[23] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University

Press, Baltimore, MD, 1996.

[24] W. Hackbush, B.N. Khoromskij, and E. Tyrtyshnikov, *Hierarchical Kronecker tensor-product approximations*, J. Numer. Math., 13 (2005), pp. 119–156.

[25] W. Hackbush and B.N. Khoromskij, *Tensor-product approximation to operators and functions in high dimension*, J. Complexity, 23 (2007), pp. 697–714.

[26] R.A. Harshman, *Foundations of the PARAFAC procedure: Model and conditions for an "explanatory" multi-mode factor analysis*, UCLA Working Papers in Phonetics, 16 (1970), pp. 1–84.

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

[28] F.L. Hitchcock, *Multiple invariants and generalized rank of a p-way matrix or tensor*, J. Math. Phys., 7 (1927), pp. 39–79.

[29] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley, New York, 2001.

[30] J. Ja'Ja', *An addendum to Kronecker's theory of pencils*, SIAM J. Appl. Math., 37 (1979), pp. 700–712.

[31] 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.

[32] T. Kato, *A Short Introduction to Perturbation theory for Linear Operators*, Springer-Verlag, New York, 1982.

[33] B.N. Khoromskij and V. Khoromskaia, *Low rank Tucker-type tensor approximation to classical potentials*, Central European J. Math., 5 (2007), pp. 523–550.

[34] H. Kiers, *Towards a standardized notation and terminology in multiway analysis*, J. Chemometrics, 14 (2000), pp. 105–122.

[35] T.G. Kolda, *Orthogonal tensor decompositions*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 243–255.

[36] T.G. Kolda, *A counterexample to the possibility of an extension of the Eckart-Young low-rank approximation theorem for the orthogonal rank tensor decomposition*, SIAM J. Matrix Analysis, 24 (2003), pp. 762–767.

[37] P.M. Kroonenberg, *Applied Multiway Data Analysis*, John Wiley, New York, 2008.

[38] 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.

[39] J.B. Kruskal, *Rank, decomposition, and uniqueness for* 3*-way and* N*-way arrays*, in Multiway Data Analysis, R. Coppi and S. Bolasco, eds., North–Holland, Amsterdam, 1989, pp. 7–18.

[40] 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.

[41] L.-H. Lim, *Singular values and eigenvalues of tensors: a variational approach*, Proceedings of the First IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2005), Puerto Vallarta, Jalisco State, Mexico, 2005, pp. 129–132.

[42] C.D. Moravitz Martin and C.F. Van Loan, *A Jacobi-type method for computing orthogonal tensor decompositions*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1219–1232.

[43] D. Nion and L. De Lathauwer, *A block factor analysis based receiver for blind multi-user access in wireless communications*, in Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006), Toulouse, France, 2006, pp. 825–828.

[44] D. Nion and L. De Lathauwer, *A tensor-based blind DS-CDMA receiver using simultaneous matrix diagonalization*, Proceedings of the Eighth IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2007), Helsinki, Finland, 2007.

[45] D. Nion and L. De Lathauwer, *Block component model based blind DS-CDMA receivers*, IEEE Trans. Signal Process., to appear.

[46] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and Its Applications*, John Wiley, New York, 1971.

[47] E. Sanchez and B.R. Kowalski, *Tensorial resolution: A direct trilinear decomposition*, J. Chemometrics, 4 (1990), pp. 29–45.

[48] 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.

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

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

[51] N. Sidiropoulos and R. Bro, *On the uniqueness of multilinear decomposition of* N*-way*

arrays, J. Chemometrics, 14 (2000), pp. 229–239.
[52] N. Sidiropoulos and G.Z. Dimić, *Blind multiuser detection in W-CDMA systems with large delay spread*, IEEE Signal Process. Letters, 8 (2001), pp. 87–89.
[53] A. Smilde, R. Bro, and P. Geladi, *Multi-way Analysis. Applications in the Chemical Sciences*, John Wiley, Chichester, UK, 2004.
[54] A. Stegeman and N.D. Sidiropoulos, *On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition*, Linear Algebra Appl., 420 (2007), pp. 540–552.
[55] J.M.F. Ten Berge and H.A.L. Kiers, *Simplicity of core arrays in three-way principal component analysis and the typical rank of $p \times q \times 2$ arrays*, Linear Algebra Appl., 294 (1999), pp. 169–179.
[56] J.M.F. Ten Berge, *The typical rank of tall three-way arrays*, Psychometrika, 65 (2000), pp. 525–532.
[57] J.M.F. Ten Berge, N.D. Sidiropoulos, and R. Rocci, *Typical rank and indscal dimensionality for symmetric three-way arrays of order $I \times 2 \times 2$ or $I \times 3 \times 3$*, Linear Algebra Appl., 388 (2004), pp. 363–377.
[58] J.M.F. Ten Berge, *Simplicity and typical rank of three-way arrays, with applications to Tucker-*3 *analysis with simple cores*, J. Chem., 18 (2004), pp. 17–21.
[59] L.R. Tucker, *The extension of factor analysis to three-dimensional matrices*, in Contributions to Mathematical Psychology, H. Gulliksen and N. Frederiksen, eds., Holt, Rinehart & Winston, New York, 1964, pp. 109–127.
[60] L.R. Tucker, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311.
[61] T. Zhang and G.H. Golub, *Rank-one approximation to high order tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550.