ON THE UNIQUENESS OF THE CANONICAL POLYADIC DECOMPOSITION OF THIRD-ORDER TENSORS — PART I: BASIC RESULTS AND UNIQUENESS OF ONE FACTOR MATRIX *

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

Abstract. Canonical Polyadic Decomposition (CPD) of a higher-order tensor is decomposition in a minimal number of rank-1 tensors. We give an overview of existing results concerning uniqueness. We present new, relaxed, conditions that guarantee uniqueness of one factor matrix. These conditions involve Khatri-Rao products of compound matrices. We make links with existing results involving ranks and k-ranks of factor matrices. We give a shorter proof, based on properties of second compound matrices, of existing results concerning overall CPD uniqueness in the case where one factor matrix has full column rank. We develop basic material involving m-th compound matrices that will be instrumental in Part II for establishing overall CPD uniqueness in cases where none of the factor matrices has full column rank.

Key words. Canonical Polyadic Decomposition, Candecomp, Parafac, three-way array, tensor, multilinear algebra, Khatri-Rao product, compound matrix

AMS subject classifications. 15A69, 15A23

1. Introduction.

1.1. Problem statement. Throughout the paper $\mathbb{F}$ denotes the field of real or complex numbers; $(\cdot)^T$ denotes transpose; $r_{\mathbf{A}}$ and range($\mathbf{A}$) denote the rank and the range of a matrix $\mathbf{A}$, respectively; Diag($\mathbf{d}$) denotes a square diagonal matrix with the elements of a vector $\mathbf{d}$ on the main diagonal; $\omega(\mathbf{d})$ denotes the number of nonzero components of $\mathbf{d}$; C_n^k denotes the binomial coefficient, $C_n^k = \frac{n!}{k!(n-k)!}$; $\mathbf{O}_{m\times n}$, $\mathbf{0}_m$, and $\mathbf{I}_n$ are the zero $m \times n$ matrix, the zero $m \times 1$ vector, and the $n \times n$ identity matrix, respectively.

We have the following basic definitions.

DEFINITION 1.1. *A third order-tensor $\mathcal{T} \in \mathbb{F}^{I\times J\times K}$ is* rank-1 *if it equals the outer product of three nonzero vectors $\mathbf{a} \in \mathbb{F}^I$, $\mathbf{b} \in \mathbb{F}^J$ and $\mathbf{c} \in \mathbb{F}^K$, which means that $t_{ijk} = a_i b_j c_k$ for all values of the indices.*

A rank-1 tensor is also called a *simple tensor* or a *decomposable tensor*. The outer product in the definition is written as $\mathcal{T} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$.

DEFINITION 1.2. *A* Polyadic Decomposition *(PD) of a third-order tensor $\mathcal{T} \in \mathbb{F}^{I\times J\times K}$ expresses $\mathcal{T}$ as a sum of rank-*1 *terms:*

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

where $\mathbf{a}_r \in \mathbb{F}^I$, $\mathbf{b}_r \in \mathbb{F}^J$, $\mathbf{c}_r \in \mathbb{F}^K$, $1 \le r \le R$.

*Research supported by: (1) Research Council KU Leuven: GOA-Ambiorics, GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT 1/08/23, (2) F.W.O.: (a) project G.0427.10N, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization", 2007–2011), (4) EU: ERNSI.

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

[‡]Department of Electrical Engineering (ESAT), SCD–SISTA, KU Leuven, Kasteelpark Arenberg 10, postbus 2440, B-3001 Heverlee (Leuven), Belgium.

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

DEFINITION 1.3. *The* rank *of a tensor $\mathcal{T} \in \mathbb{F}^{I\times J\times K}$ is defined as the minimum number of rank-*1 *tensors in a PD of $\mathcal{T}$ and is denoted by $r_{\mathcal{T}}$.*

In general, the rank of a third-order tensor depends on $\mathbb{F}$ [21]: a tensor over $\mathbb{R}$ may have a different rank than the same tensor considered over $\mathbb{C}$.

DEFINITION 1.4. *A* Canonical Polyadic Decomposition *(CPD) of a third-order tensor $\mathcal{T}$ expresses $\mathcal{T}$ as a minimal sum of rank-*1 *terms.*

Note that $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ is a CPD of $\mathcal{T}$ if and only if $R = r_{\mathcal{T}}$.

Let us reshape $\mathcal{T}$ into a vector $\mathbf{t} \in \mathbb{F}^{IJK\times 1}$ and a matrix $\mathbf{T} \in \mathbb{F}^{IJ\times K}$ as follows: the (i, j, k)-th entry of $\mathcal{T}$ corresponds to the $((i-1)JK + (j-1)K + k)$-th entry of $\mathbf{t}$ and to the $((i-1)J + j, k)$-th entry of $\mathbf{T}$. In particular, the rank-1 tensor $\mathbf{a}\circ\mathbf{b}\circ\mathbf{c}$ corresponds to the vector $\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}$ and to the rank-1 matrix $(\mathbf{a}\otimes\mathbf{b})\mathbf{c}^T$, where "$\otimes$" denotes the *Kronecker product*:

$$\mathbf{a}\otimes\mathbf{b} = \begin{bmatrix}a_1\mathbf{b}^T & \dots & a_I\mathbf{b}^T\end{bmatrix}^T = \begin{bmatrix}a_1b_1\dots a_1b_J & \dots & a_Ib_1\dots a_Ib_J\end{bmatrix}^T.$$

Thus, (1.1) can be identified either with

$$\mathbf{t} = \sum_{r=1}^{R} \mathbf{a}_r\otimes\mathbf{b}_r\otimes\mathbf{c}_r, \tag{1.2}$$

or with the matrix decomposition

$$\mathbf{T} = \sum_{r=1}^{R} (\mathbf{a}_r\otimes\mathbf{b}_r)\mathbf{c}_r^T. \tag{1.3}$$

Further, (1.3) can be rewritten as a factorization of $\mathbf{T}$,

$$\mathbf{T} = (\mathbf{A}\odot\mathbf{B})\mathbf{C}^T, \tag{1.4}$$

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

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

It is clear that in (1.1)–(1.3) the rank-1 terms can be arbitrarily permuted and that vectors within the same rank-1 term can be arbitrarily scaled provided the overall rank-1 term remains the same. *The CPD of a tensor is unique* when it is only subject to these trivial indeterminacies.

In this paper we find sufficient conditions on the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ which guarantee that the CPD of $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ is partially unique in the following sense: the third factor matrix of any other CPD of $\mathcal{T}$ coincides with $\mathbf{C}$ up to permutation and scaling of columns. In such a case we say that *the third factor matrix of $\mathcal{T}$ is unique*. We also develop basic material involving m-th compound matrices that will be instrumental in Part II for establishing overall CPD uniqueness.

1.2. Literature overview. The CPD was introduced by F.L. Hitchcock in [14]. It has been rediscovered a number of times and called *Canonical Decomposition* (Candecomp) [1], *Parallel Factor Model* (Parafac) [11, 13], and *Topographic Components Model* [24]. Key to many applications are the uniqueness properties of the CPD.

Contrary to the matrix case, where there exist (infinitely) many rank-revealing decompositions, CPD may be unique without imposing constraints like orthogonality. Such constraints cannot always be justified from an application point of view. In this sense, CPD may be a meaningful data representation, and actually reveals a unique decomposition of the data in interpretable components. CPD has found many applications in Signal Processing [2], [3], Data Analysis [19], Chemometrics [29], Psychometrics [1], etc. We refer to the overview papers [4,7,17] and the references therein for background, applications and algorithms. We also refer to [30] for a discussion of optimization-based algorithms.

1.2.1. Early results on uniqueness of the CPD. In [11, p. 61] the following result concerning the uniqueness of the CPD is attributed to R. Jennrich.

THEOREM 1.5. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let*

$$r_{\mathbf{A}} = r_{\mathbf{B}} = r_{\mathbf{C}} = R. \tag{1.5}$$

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

Condition (1.5) may be relaxed as follows.

THEOREM 1.6. *[12] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$*, let*

$$r_{\mathbf{A}} = r_{\mathbf{B}} = R \quad \textit{and let any two columns of } \mathbf{C} \textit{ be linearly independent.}$$

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

1.2.2. Kruskal's conditions. A further relaxed result is due to J. Kruskal. To present Kruskal's theorem we recall the definition of k-rank ("k" refers to "Kruskal").

DEFINITION 1.7. *The* k*-rank of a matrix* $\mathbf{A}$ *is the largest number* $k_{\mathbf{A}}$ *such that every subset of* $k_{\mathbf{A}}$ *columns of the matrix* $\mathbf{A}$ *is linearly independent.*

Obviously, $k_{\mathbf{A}} \leq r_{\mathbf{A}}$. Note that the notion of the k-rank is closely related to the notions of girth, spark, and k-stability [23, Lemma 5.2, p. 317] and references therein. The famous Kruskal theorem states the following.

THEOREM 1.8. *[20] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2. \tag{1.6}$$

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

Kruskal's original proof was made more accessible in [32] and was simplified in [22, Theorem 12.5.3.1, p. 306]. In [25] an other proof of Theorem 1.8 is given.

Before Kruskal arrived at Theorem 1.8 he obtained results about uniqueness of one factor matrix [20, Theorem 3a–3d, p. 115–116]. These results were flawed. Here we present their corrected versions.

THEOREM 1.9. *[9, Theorem 2.3] (for original formulation see [20, Theorems 3a,b]) Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and suppose*

$$\begin{cases} k_{\mathbf{C}} \geq 1, \\ r_{\mathbf{C}} + \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq R + 2, \\ r_{\mathbf{C}} + k_{\mathbf{A}} + k_{\mathbf{B}} + \max(r_{\mathbf{A}} - k_{\mathbf{A}}, r_{\mathbf{B}} - k_{\mathbf{B}}) \geq 2R + 2. \end{cases} \tag{1.7}$$

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

Let the matrices $\mathbf{A}$ and $\mathbf{B}$ have R columns. Let $\tilde{\mathbf{A}}$ be any set of columns of $\mathbf{A}$, let $\tilde{\mathbf{B}}$ be the corresponding set of columns of $\mathbf{B}$. We will say that condition (Hm) holds for the matrices $\mathbf{A}$ and $\mathbf{B}$ if

$$H(\delta) := \min_{card(\tilde{\mathbf{A}})=\delta} [r_{\tilde{\mathbf{A}}} + r_{\tilde{\mathbf{B}}} - \delta] \geq \min(\delta, m) \quad \text{for} \quad \delta = 1, 2, \ldots, R. \tag{Hm}$$

THEOREM 1.10. *(see §4, for original formulation see [20, Theorems 3d]) Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m := R - r_{\mathbf{C}} + 2$. *Assume that*

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

(ii) (Hm) *holds for* $\mathbf{A}$ *and* $\mathbf{B}$.

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

Kruskal also obtained results about overall uniqueness that are more general than Theorem 1.8. These results will be discussed in Part II [8].

1.2.3. Uniqueness of the CPD when one factor matrix has full column rank. We say that a $K \times R$ matrix *has full column rank* if its column rank is R, which implies $K \geq R$.

Let us assume that $r_{\mathbf{C}} = R$. The following result concerning uniqueness of the CPD was obtained by T. Jiang and N. Sidiropoulos in [16]. We reformulate the result in terms of the Khatri-Rao product of the second compound matrices of $\mathbf{A}$ and $\mathbf{B}$. The k-th compound matrix of an $I \times R$ matrix $\mathbf{A}$ (denoted by $\mathcal{C}_k(\mathbf{A})$) is the $C_I^k \times C_R^k$ matrix containing the determinants of all $k \times k$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order (see Definition 2.1 and Example 2.2).

THEOREM 1.11. *[16, Condition A, p. 2628, Condition B and eqs. (16) and (17), p. 2630] Let* $\mathbf{A} \in \mathbb{F}^{I \times R}$, $\mathbf{B} \in \mathbb{F}^{J \times R}$, $\mathbf{C} \in \mathbb{F}^{K \times R}$ *and* $r_{\mathbf{C}} = R$. *Then the following statements are equivalent:*

(i) *if* $\mathbf{d} \in \mathbb{F}^R$ *is such that* $r_{\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T} \leq 1$, *then* $\omega(\mathbf{d}) \leq 1$;

(ii) *if* $\mathbf{d} \in \mathbb{F}^R$ *is such that*

$$(\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})) \begin{bmatrix} d_1 d_2 & d_1 d_3 & \dots & d_1 d_R & d_2 d_3 & \dots & d_{R-1} d_R \end{bmatrix}^T = \mathbf{0},$$

then $\omega(\mathbf{d}) \leq 1$; (U2)

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

Papers [16] and [5] contain the following more restrictive sufficient condition for CPD uniqueness, formulated differently. This condition can be expressed in terms of second compound matrices as follows.

THEOREM 1.12. *[5, Remark 1, p. 652], [16] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = R$, *and suppose*

$$\mathbf{U} = \mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}) \text{ has full column rank.} \tag{C2}$$

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

It is clear that (C2) implies (U2). If $r_{\mathbf{C}} = R$, then Kruskal's condition (1.6) is more restrictive than condition (C2).

THEOREM 1.13. *[31, Proposition 3.2, p. 215 and Lemma 4.4, p. 221] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let* $r_{\mathbf{C}} = R$. *If*

$$\begin{cases} r_{\mathbf{A}} + k_{\mathbf{B}} \geq R + 2, \\ k_{\mathbf{A}} \geq 2 \end{cases} \quad \text{or} \quad \begin{cases} r_{\mathbf{B}} + k_{\mathbf{A}} \geq R + 2, \\ k_{\mathbf{B}} \geq 2, \end{cases} \tag{K2}$$

then (C2) *holds. Hence,* $r_{\mathcal{T}} = R$ *and the CPD of* $\mathcal{T}$ *is unique.*

Theorem 1.13 is due to A. Stegeman [31, Proposition 3.2, p. 215 and Lemma 4.4, p. 221]. Recently, another proof of Theorem 1.13 has been obtained in [10, Theorem 1, p. 3477].

Assuming $r_{\mathbf{C}} = R$, the conditions of Theorems 1.8 through 1.13 are related by

$$\begin{aligned} k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2 \;&\Rightarrow (\mathrm{K2}) \;\Rightarrow (\mathrm{C2}) \;\Rightarrow (\mathrm{U2}) \\ &\Leftrightarrow r_{\mathcal{T}} = R \text{ and the CPD of } \mathcal{T} \text{ is unique.} \end{aligned} \tag{1.8}$$

1.2.4. Necessary conditions for uniqueness of the CPD. Results concerning rank and k-rank of Khatri-Rao product. It was shown in [34] that condition (1.6) is not only sufficient but also necessary for the uniqueness of the CPD if $R = 2$ or $R = 3$. Moreover, it was proved in [34] that if $R = 4$ and if the k-ranks of the factor matrices coincide with their ranks, then the CPD of $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_4$ is unique if and only if condition (1.6) holds. Passing to higher values of R we have the following theorems.

THEOREM 1.14. *[33, p. 651], [36, p. 2079, Theorem 2], [18, p. 28] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathcal{T}} = R \geq 2$, and let the CPD of $\mathcal{T}$ be unique. Then*

(i) $\mathbf{A} \odot \mathbf{B}$, $\mathbf{B} \odot \mathbf{C}$, $\mathbf{C} \odot \mathbf{A}$ *have full column rank;*

(ii) $\min(k_{\mathbf{A}}, k_{\mathbf{B}}, k_{\mathbf{C}}) \geq 2$.

THEOREM 1.15. *[6, Theorem 2.3] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathcal{T}} = R \geq 2$, and let the CPD of $\mathcal{T}$ be unique. Then the condition* (U2) *holds for the pairs* $(\mathbf{A}, \mathbf{B})$, $(\mathbf{B}, \mathbf{C})$, *and* $(\mathbf{C}, \mathbf{A})$.

Theorem 1.15 gives more restrictive uniqueness conditions than Theorem 1.14 and generalizes the implication (iii)⇒(ii) of Theorem 1.11 to CPDs with $r_{\mathbf{C}} \leq R$.

The following lemma gives a condition under which

$$\mathbf{A} \odot \mathbf{B} \text{ has full column rank.} \tag{C1}$$

LEMMA 1.16. *[10, Lemma 1, p. 3477] Let $\mathbf{A} \in \mathbb{F}^{I \times R}$ and $\mathbf{B} \in \mathbb{F}^{J \times R}$. If*

$$\begin{cases} r_{\mathbf{A}} + k_{\mathbf{B}} & \geq R+1, \\ k_{\mathbf{A}} & \geq 1 \end{cases} \quad \text{or} \quad \begin{cases} r_{\mathbf{B}} + k_{\mathbf{A}} & \geq R+1, \\ k_{\mathbf{B}} & \geq 1, \end{cases} \tag{K1}$$

then (C1) *holds.*

We conclude this section by mentioning two important corollaries that we will use.

COROLLARY 1.17. *[27, Lemma 1, p. 2382] If $k_{\mathbf{A}} + k_{\mathbf{B}} \geq R+1$, then* (C1) *holds.*

COROLLARY 1.18. *[28, Lemma 1, p. 231] If $k_{\mathbf{A}} \geq 1$ and $k_{\mathbf{B}} \geq 1$, then $k_{\mathbf{A} \odot \mathbf{B}} \geq \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, R)$.*

The proof of Corollary 1.18 in [28] was based on Corollary 1.17. Other proofs are given in [26, Lemma 1, p. 231] and [32, Lemma 3.3, p. 544]. (The proof in [32] is due to J. Ten Berge, see also [35].) All mentioned proofs are based on the Sylvester rank inequality.

1.3. Results and organization. Motivated by the conditions appearing in the various theorems of the preceding section, we formulate more general versions, depending on an integer parameter m. How these conditions, in conjunction with other assumptions, imply the uniqueness of one particular factor matrix will be the core of our work.

To introduce the new conditions we need the following notation. With a vector $\mathbf{d} = \begin{bmatrix} d_1 & \dots & d_R \end{bmatrix}^T$ we associate the vector

$$\widehat{\mathbf{d}}^m := \begin{bmatrix} d_1 \cdots d_m & d_1 \cdots d_{m-1} d_{m+1} & \dots & d_{R-m+1} \cdots d_R \end{bmatrix}^T \in \mathbb{F}^{C_R^m}, \tag{1.9}$$

whose entries are all products $d_{i_1} \cdots d_{i_m}$ with $1 \leq i_1 < \cdots < i_m \leq R$. Let us define conditions (Km), (Cm), (Um) and (Wm), which depend on matrices $\mathbf{A} \in \mathbb{F}^{I \times R}$, $\mathbf{B} \in \mathbb{F}^{J \times R}$, $\mathbf{C} \in \mathbb{F}^{K \times R}$ and an integer parameter m:

$$\begin{cases} r_{\mathbf{A}} + k_{\mathbf{B}} & \geq R+m, \\ k_{\mathbf{A}} & \geq m \end{cases} \quad \text{or} \quad \begin{cases} r_{\mathbf{B}} + k_{\mathbf{A}} & \geq R+m, \\ k_{\mathbf{B}} & \geq m \end{cases}; \tag{Km}$$

$$\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}) \qquad \text{has full column rank;} \tag{Cm}$$

$$\begin{cases} (\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))\widehat{\mathbf{d}}^m = \mathbf{0}, \\ \mathbf{d} \in \mathbb{F}^R \end{cases} \quad \Rightarrow \quad \widehat{\mathbf{d}}^m = \mathbf{0}; \tag{Um}$$

$$\begin{cases} (\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))\widehat{\mathbf{d}}^m = \mathbf{0}, \\ \mathbf{d} \in \text{range}(\mathbf{C}^T) \end{cases} \quad \Rightarrow \quad \widehat{\mathbf{d}}^m = \mathbf{0}. \tag{Wm}$$

In §2 we give a formal definition of compound matrices and present some of their properties. This basic material will be heavily used in the following sections.

In §3 we establish the following implications:

$$\begin{array}{lccccccccc} & (\mathrm{W}_m) & & (\mathrm{W}_{m\text{-}1}) & & \dots & & (\mathrm{W}_2) & & (\mathrm{W}_1) \\ \text{(Lemma 3.3)} & \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Uparrow \\ \text{(Lemma 3.7)} & (\mathrm{U}_m) & \Rightarrow & (\mathrm{U}_{m\text{-}1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{U}_2) & \Rightarrow & (\mathrm{U}_1) \\ \text{(Lemma 3.1)} & \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Updownarrow \\ \text{(Lemma 3.6)} & (\mathrm{C}_m) & \Rightarrow & (\mathrm{C}_{m\text{-}1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{C}_2) & \Rightarrow & (\mathrm{C}_1) \\ \text{(Lemma 3.8)} & \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Uparrow \\ \text{(Lemma 3.4)} & (\mathrm{K}_m) & \Rightarrow & (\mathrm{K}_{m\text{-}1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{K}_2) & \Rightarrow & (\mathrm{K}_1) \end{array} \tag{1.10}$$

as well as (Lemma 3.12)

$$\text{if } \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1, \text{ then } (\mathrm{W}_m) \Rightarrow (\mathrm{W}_{m\text{-}1}) \Rightarrow \dots \Rightarrow (\mathrm{W}_2) \Rightarrow (\mathrm{W}_1). \tag{1.11}$$

We also show in Lemmas 3.5, 3.9–3.10 that (1.10) remains valid after replacing conditions (C_m),…,(C_1) and equivalence (C_1) $\Leftrightarrow$ (U_1) by conditions (H_m),…,(H_1) and implication (H_1) $\Rightarrow$ (U_1), respectively.

Equivalence of (C_1) and (U_1) is trivial, since the two conditions are the same. The implications (K_2) $\Rightarrow$ (C_2) $\Rightarrow$ (U_2) already appeared in (1.8). The implication (K_1) $\Rightarrow$ (C_1) was given in Lemma 1.16, and the implications (K_m) $\Rightarrow$ (H_m) $\Rightarrow$ (U_m) are implicitly contained in [20]. From the definition of conditions (K_m) and (H_m) it follows that $r_{\mathbf{A}} + r_{\mathbf{B}} \geq R + m$. On the other hand, condition (C_m) may hold for $r_{\mathbf{A}} + r_{\mathbf{B}} < R + m$. We do not know examples where (H_m) holds, but (C_m) does not. We suggest that (H_m) always implies (C_m).

In §4 we present a number of results establishing the uniqueness of one factor matrix under various hypotheses including at least one of the conditions (K_m), (H_m), (C_m), (U_m) and (W_m). The results of this section can be summarized as:
if $k_{\mathbf{C}} \geq 1$ and $m = m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2$, then

$$(1.7) \;\Leftrightarrow\; \begin{array}{ccc} & (\mathrm{C}_m) & \\ \nearrow & & \searrow \\ (\mathrm{K}_m) & & (\mathrm{U}_m) \\ \searrow & & \nearrow \\ & (\mathrm{H}_m) & \end{array} \;\Rightarrow\; \begin{cases} \mathbf{A} \odot \mathbf{B} \text{ has full column rank,} \\ (\mathrm{W}_m), \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1 \end{cases} \tag{1.12}$$

$$\Rightarrow \begin{cases} \mathbf{A} \odot \mathbf{B} \text{ has full column rank,} \\ (\mathrm{W}_m), (\mathrm{W}_{m\text{-}1}), \dots, (\mathrm{W}_1) \end{cases}$$

$$\Rightarrow \quad r_{\mathcal{T}} = R \;\text{ and the third factor matrix of } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \;\text{ is unique.}$$

Thus, Theorems 1.9–1.10 are implied by the more general statement (1.12), which therefore provides new, more relaxed sufficient conditions for uniqueness of one factor matrix.

Further, compare (1.12) to (1.8). For the case $r_{\mathbf{C}} = R$, i.e., $m = 2$, uniqueness of the overall CPD has been established in Theorem 1.11. Actually, in this case overall CPD uniqueness follows easily from uniqueness of $\mathbf{C}$.

In §5 we simplify the proof of Theorem 1.11 using the material we have developed so far. In Part II [8] we will use (1.12) to generalize (1.8) to cases where possibly $r_{\mathbf{C}} < R$, i.e., $m > 2$.

2. Compound matrices and their properties. In this section we define compound matrices and present several of their properties. The material will be heavily used in the following sections.

Let

$$S_n^k := \{(i_1, \ldots, i_k) : 1 \leq i_1 < \cdots < i_k \leq n\} \tag{2.1}$$

denote the set of all k combinations of the set $\{1, \ldots, n\}$. We assume that the elements of S_n^k are ordered lexicographically. Since the elements of S_n^k can be indexed from 1 up to C_n^k, there exists an order preserving bijection

$$\sigma_{n,k} : \{1, 2, \ldots, C_n^k\} \to S_n^k = \{S_n^k(1), S_n^k(2), \ldots, S_n^k(C_n^k)\}. \tag{2.2}$$

In the sequel we will both use indices taking values in $\{1, 2, \ldots, C_n^k\}$ and multi-indices taking values in S_n^k. The connection between both is given by (2.2).

To distinguish between vectors from $\mathbb{F}^R$ and $\mathbb{F}^{C_n^k}$ we will use the subscript S_n^k, which will also indicate that the vector entries are enumerated by means of S_n^k. For instance, throughout the paper the vectors $\mathbf{d} \in \mathbb{F}^R$ and $\mathbf{d}_{S_R^m} \in \mathbb{F}^{C_R^m}$ are always defined by

$$\begin{aligned} \mathbf{d} &= \begin{bmatrix} d_1 & d_2 & \ldots & d_R \end{bmatrix} \in \mathbb{F}^R, \\ \mathbf{d}_{S_R^m} &= \begin{bmatrix} d_{(1,\ldots,m)} & \ldots & d_{(j_1,\ldots,j_m)} & \ldots & d_{(R-m+1,\ldots,R)} \end{bmatrix}^T \in \mathbb{F}^{C_R^m}. \end{aligned} \tag{2.3}$$

Note that if $d_{(i_1,\ldots,i_m)} = d_{i_1} \cdots d_{i_m}$ for all indices $i_1, \ldots, i_m$, then the vector $\mathbf{d}_{S_R^m}$ is equal to the vector $\widehat{\mathbf{d}}^m$ defined in (1.9).

Thus, $\mathbf{d}_{S_R^1} = \widehat{\mathbf{d}}^1 = \mathbf{d}$.

DEFINITION 2.1. *[15] Let $\mathbf{A} \in \mathbb{F}^{m \times n}$ and $k \leq \min(m, n)$. Denote by $\mathbf{A}(S_m^k(i), S_m^k(j))$ the submatrix at the intersection of the k rows with row numbers $S_m^k(i)$ and the k columns with column numbers $S_m^k(j)$. The C_m^k-by-C_n^k matrix whose (i, j) entry is $\det \mathbf{A}(S_m^k(i), S_n^k(j))$ is called the k-th compound matrix of $\mathbf{A}$ and is denoted by $\mathcal{C}_k(\mathbf{A})$.*

EXAMPLE 2.2. *Let*

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

Then

$$\begin{aligned} \mathcal{C}_2(\mathbf{A}) &= \begin{bmatrix} \mathcal{C}_2(\mathbf{A})_1 & \mathcal{C}_2(\mathbf{A})_2 & \mathcal{C}_2(\mathbf{A})_3 & \mathcal{C}_2(\mathbf{A})_4 & \mathcal{C}_2(\mathbf{A})_5 & \mathcal{C}_2(\mathbf{A})_6 \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{C}_2(\mathbf{A})_{(1,2)} & \mathcal{C}_2(\mathbf{A})_{(1,3)} & \mathcal{C}_2(\mathbf{A})_{(1,4)} & \mathcal{C}_2(\mathbf{A})_{(2,3)} & \mathcal{C}_2(\mathbf{A})_{(2,4)} & \mathcal{C}_2(\mathbf{A})_{(3,4)} \end{bmatrix} \end{aligned}$$

$$\begin{array}{r} \\ (1,2) \\ =(1,3) \\ (2,3) \end{array}\!\!\begin{array}{c} \begin{array}{cccccc} (1,2) & & (1,3) & & (1,4) & \end{array} \\ \begin{bmatrix} \begin{vmatrix} a_1 & 1 \\ a_2 & 0 \end{vmatrix} & \begin{vmatrix} a_1 & 0 \\ a_2 & 1 \end{vmatrix} & \begin{vmatrix} a_1 & 0 \\ a_2 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} a_1 & 1 \\ a_3 & 0 \end{vmatrix} & \begin{vmatrix} a_1 & 0 \\ a_3 & 0 \end{vmatrix} & \begin{vmatrix} a_1 & 0 \\ a_3 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} \\ \begin{vmatrix} a_2 & 0 \\ a_3 & 0 \end{vmatrix} & \begin{vmatrix} a_2 & 1 \\ a_3 & 0 \end{vmatrix} & \begin{vmatrix} a_2 & 0 \\ a_3 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} \end{bmatrix} \end{array}$$

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

Definition 2.1 immediately implies the following lemma.

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

1. $\mathcal{C}_1(\mathbf{A}) = \mathbf{A}$;
2. *If* $I = R$, *then* $\mathcal{C}_R(\mathbf{A}) = \det(\mathbf{A})$;
3. $\mathcal{C}_k(\mathbf{A})$ *has one or more zero columns if and only if* $k > k_{\mathbf{A}}$;
4. $\mathcal{C}_k(\mathbf{A})$ *is equal to the zero matrix if and only if* $k > r_{\mathbf{A}}$.

The following properties of compound matrices are well-known.

LEMMA 2.4. *[15, p. 19–22] Let* k *be a positive integer and let* $\mathbf{A}$ *and* $\mathbf{B}$ *be matrices such that* $\mathbf{AB}$, $\mathcal{C}_k(\mathbf{A})$, *and* $\mathcal{C}_k(\mathbf{B})$ *are defined. Then*

1. $\mathcal{C}_k(\mathbf{AB}) = \mathcal{C}_k(\mathbf{A})\mathcal{C}_k(\mathbf{B})$ *(Binet-Cauchy formula);*
2. *If* $\mathbf{A}$ *is nonsingular square matrix, then* $\mathcal{C}_k(\mathbf{A})^{-1} = \mathcal{C}_k(\mathbf{A}^{-1})$;
3. $\mathcal{C}_k(\mathbf{A}^T) = (\mathcal{C}_k(\mathbf{A}))^T$;
4. $\mathcal{C}_k(\mathbf{I}_n) = \mathbf{I}_{C_n^k}$;
5. *If* $\mathbf{A}$ *is an* $n \times n$ *matrix, then* $\det(\mathcal{C}_k(\mathbf{A})) = \det(\mathbf{A})^{C_{n-1}^{k-1}}$ *(Sylvester-Franke theorem).*

We will extensively use compound matrices of diagonal matrices.

LEMMA 2.5. *Let* $\mathbf{d} \in \mathbb{F}^R$, $k \leq R$, *and let* $\widehat{\mathbf{d}}^k$ *be defined by* (1.9). *Then*

1. $\widehat{\mathbf{d}}^k = \mathbf{0}$ *if and only if* $\omega(\mathbf{d}) \leq k-1$;
2. $\widehat{\mathbf{d}}^k$ *has exactly one nonzero component if and only if* $\omega(\mathbf{d}) = k$;
3. $\mathcal{C}_k(\mathrm{Diag}(\mathbf{d})) = \mathrm{Diag}(\widehat{\mathbf{d}}^k)$.

EXAMPLE 2.6. *Let* $\mathbf{d} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 \end{bmatrix}^T$ *and* $\mathbf{D} = \mathrm{Diag}(\mathbf{d})$. *Then*

$$\mathcal{C}_2(\mathbf{D}) = \mathrm{Diag}(\begin{bmatrix} d_1d_2 & d_1d_3 & d_1d_4 & d_2d_3 & d_2d_4 & d_3d_4 \end{bmatrix}) = \mathrm{Diag}(\widehat{\mathbf{d}}^2),$$
$$\mathcal{C}_3(\mathbf{D}) = \mathrm{Diag}(\begin{bmatrix} d_1d_2d_3 & d_1d_2d_4 & d_1d_3d_4 & d_2d_3d_4 \end{bmatrix}) = \mathrm{Diag}(\widehat{\mathbf{d}}^3).$$

For vectorization of a matrix $\mathbf{T} = [\mathbf{t}_1 \ \cdots \ \mathbf{t}_R]$, we follow the convention that $\mathrm{vec}(\mathbf{T})$ denotes the column vector obtained by stacking the columns of $\mathbf{T}$ on top of one another, i.e.,

$$\mathrm{vec}(\mathbf{T}) = \begin{bmatrix} \mathbf{t}_1^T & \ldots & \mathbf{t}_R^T \end{bmatrix}^T.$$

It is clear that in vectorized form, rank-1 matrices correspond to Kronecker products of two vectors. Namely, for arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$, $\mathrm{vec}(\mathbf{ba}^T) = \mathbf{a} \otimes \mathbf{b}$. For matrices $\mathbf{A}$ and $\mathbf{B}$ that both have R columns and $\mathbf{d} \in \mathbb{F}^R$, we now immediately obtain expressions

that we will frequently use:

$$\mathrm{vec}(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T) = \mathrm{vec}\left(\sum_{r=1}^{R} \mathbf{b}_r \mathbf{a}_r^T d_r\right) = \sum_{r=1}^{R} (\mathbf{a}_r \otimes \mathbf{b}_r) d_r = (\mathbf{A} \odot \mathbf{B})\mathbf{d}, \tag{2.4}$$

$$\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T = \mathbf{O} \;\Leftrightarrow\; \mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T = \mathbf{O} \;\Leftrightarrow\; (\mathbf{A} \odot \mathbf{B})\mathbf{d} = \mathbf{0}. \tag{2.5}$$

The following generalization of property (2.4) will be used throughout the paper.

LEMMA 2.7. *Let* $\mathbf{A} \in \mathbb{F}^{I \times R}$, $\mathbf{B} \in \mathbb{F}^{J \times R}$, $\mathbf{d} \in \mathbb{F}^R$, *and* $k \leq \min(I, J, R)$. *Then*

$$\mathrm{vec}(\mathcal{C}_k(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T)) = [\mathcal{C}_k(\mathbf{A}) \odot \mathcal{C}_k(\mathbf{B})]\widehat{\mathbf{d}}^k,$$

where $\widehat{\mathbf{d}}^k \in \mathbb{F}^{C_R^k}$ *is defined by* (1.9).

Proof. From Lemma 2.4 (1),(3) and Lemma 2.5 (3) it follows that

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

By (2.4),

$$\mathrm{vec}(\mathcal{C}_k(\mathbf{B})\mathrm{Diag}(\widehat{\mathbf{d}}^k)\mathcal{C}_k(\mathbf{A})^T) = [\mathcal{C}_k(\mathbf{A}) \odot \mathcal{C}_k(\mathbf{B})]\widehat{\mathbf{d}}^k. \qquad \square$$

The following Lemma contains an equivalent definition of condition (Um).

LEMMA 2.8. *Let* $\mathbf{A} \in \mathbb{F}^{I \times R}$ *and* $\mathbf{B} \in \mathbb{F}^{J \times R}$. *Then the following statements are equivalent:*

(i) *if* $\mathbf{d} \in \mathbb{F}^R$ *is such that* $r_{\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T} \leq m-1$, *then* $\omega(\mathbf{d}) \leq m-1$;

(ii) (Um) *holds.*

Proof. From the definition of the m-th compound matrix and Lemma 2.7 it follows that

$$\begin{aligned} r_{\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T} = r_{\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T} \leq m-1 \quad &\Leftrightarrow \quad \mathcal{C}_m(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T) = \mathbf{O} \\ &\Leftrightarrow \quad \mathrm{vec}(\mathcal{C}_m(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T)) = \mathbf{0} \quad \Leftrightarrow \quad [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\widehat{\mathbf{d}}^m = \mathbf{0}. \end{aligned}$$

Now the result follows from Lemma 2.5 (1). $\square$

The following three auxiliary lemmas will be used in §3.

LEMMA 2.9. *Consider* $\mathbf{A} \in \mathbb{F}^{I \times R}$ *and* $\mathbf{B} \in \mathbb{F}^{J \times R}$ *and let condition* (Um) *hold. Then* $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m$.

Proof. We prove equivalently that if $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m$ does not hold, then (Um) does not hold. Hence, we start by assuming that $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) = k < m$, which implies that there exist indices $i_1, \ldots, i_m$ such that the vectors $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_m}$ or the vectors $\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_m}$ are linearly dependent. Let

$$\mathbf{d} := \begin{bmatrix} d_1 & \ldots & d_R \end{bmatrix}^T, \qquad d_i := \begin{cases} 1, & i \in \{i_1, \ldots, i_m\}; \\ 0, & i \notin \{i_1, \ldots, i_m\}, \end{cases}$$

and let $\widehat{\mathbf{d}}^m \in \mathbb{F}^{C_R^m}$ be given by (1.9). Because of the way $\mathbf{d}$ is defined, $\widehat{\mathbf{d}}^m$ has exactly one nonzero entry, namely $d_{i_1} \cdots d_{i_m}$. We now have

$$(\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))\widehat{\mathbf{d}}^m = \mathcal{C}_m(\begin{bmatrix} \mathbf{a}_{i_1} & \ldots & \mathbf{a}_{i_m} \end{bmatrix}) \odot \mathcal{C}_m(\begin{bmatrix} \mathbf{b}_{i_1} & \ldots & \mathbf{b}_{i_m} \end{bmatrix}) d_{i_1} \cdots d_{i_m} = \mathbf{0},$$

in which the latter equality holds because of the assumed linear dependence of $\mathbf{a}_{i_1}$, $\ldots$, $\mathbf{a}_{i_m}$ or $\mathbf{b}_{i_1}$, $\ldots$, $\mathbf{b}_{i_m}$. We conclude that condition (Um) does not hold. $\square$

LEMMA 2.10. *Let $m \leq I$. Then there exists a linear mapping $\Phi^{I,m} : \mathbb{F}^I \to \mathbb{F}^{C_I^m \times C_I^{m-1}}$ such that*

$$\mathcal{C}_m([\mathbf{A}\ \mathbf{x}]) = \Phi^{I,m}(\mathbf{x})\mathcal{C}_{m-1}(\mathbf{A}) \quad \text{for all } \mathbf{A} \in \mathbb{F}^{I\times(m-1)} \text{ and for all } \mathbf{x} \in \mathbb{F}^I. \quad (2.6)$$

Proof. Since $[\mathbf{A}\ \mathbf{x}]$ has m columns, $\mathcal{C}_m([\mathbf{A}\ \mathbf{x}])$ is a vector that contains the determinants of the matrices formed by m rows. Each of these determinants can be expanded along its last column, yielding linear combinations of $(m-1)\times(m-1)$ minors, the combination coefficients being equal to entries of $\mathbf{x}$, possibly up to the sign. Overall, the expansion can be written in the form (2.6), in which $\Phi^{I,m}(\mathbf{x})$ is a $C_I^m \times C_I^{m-1}$ matrix, the nonzero entries of which are equal to entries of $\mathbf{x}$, possibly up to the sign.

More in detail, we have the following.

(i) Let $\widehat{\mathbf{A}} \in \mathbb{F}^{m\times(m-1)}$, $\widehat{\mathbf{x}} \in \mathbb{F}^m$. By the Laplace expansion theorem [15, p. 7],

$$\mathcal{C}_m([\widehat{\mathbf{A}}\ \widehat{\mathbf{x}}]) = \det([\widehat{\mathbf{A}}\ \widehat{\mathbf{x}}]) = \begin{bmatrix}\widehat{x}_m & -\widehat{x}_{m-1} & \widehat{x}_{m-2} & \dots & (-1)^{m-1}\widehat{x}_1\end{bmatrix} \mathcal{C}_{m-1}(\widehat{\mathbf{A}}).$$

Hence, Lemma 2.10 holds for $m = I$ with

$$\Phi^{m,m}(\mathbf{x}) = \begin{bmatrix}x_m & -x_{m-1} & x_{m-2} & \dots & (-1)^{m-1}x_1\end{bmatrix}.$$

(ii) Let $m < I$. Since, $\mathcal{C}_m([\mathbf{A}\ \mathbf{x}]) = \begin{bmatrix}d_1 & \dots & d_{C_I^m}\end{bmatrix}^T$, it follows from the definition of compound matrix that $d_i = \mathcal{C}_m([\widehat{\mathbf{A}}\ \widehat{\mathbf{x}}])$, where $[\widehat{\mathbf{A}}\ \widehat{\mathbf{x}}]$ is a submatrix of $[\mathbf{A}\ \mathbf{x}]$ formed by rows with the numbers $\sigma_{I,m}(i) = S_I^m(i) := (i_1, \dots, i_m)$. Let us define $\Phi_i(\mathbf{x}) \in \mathbb{F}^{1\times C_I^{m-1}}$ by

$$\begin{array}{rccccccccc}\Phi_i(\mathbf{x}) = [& 0 & \dots & 0 & x_{i_m} & 0 & \dots & 0 & (-1)^{m-1}x_{i_1} & \dots],\\ & \uparrow & \dots & \dots & \uparrow & \dots & \dots & \dots & \uparrow & \dots\\ & 1 & \dots & \dots & j_m & \dots & \dots & \dots & j_1 & \dots\end{array}$$

where

$$j_m := \sigma_{I,m-1}^{-1}((i_1, \dots, i_{m-1})), \quad \dots \quad , j_1 := \sigma_{I,m-1}^{-1}((i_2, \dots, i_m))$$

and $\sigma_{I,m-1}^{-1}$ is defined by (2.2). Then by (i),

$$d_i = \mathcal{C}_m([\widehat{\mathbf{A}}\ \widehat{\mathbf{x}}]) = \begin{bmatrix}x_{i_m} & -x_{i_{m-1}} & x_{i_{m-2}} & \dots & (-1)^{m-1}x_{i_1}\end{bmatrix}\mathcal{C}_{m-1}(\widehat{\mathbf{A}}) = \Phi_i(\mathbf{x})\mathcal{C}_{m-1}(\mathbf{A}).$$

The proof is completed by setting

$$\Phi^{I,m}(\mathbf{x}) = \begin{bmatrix}\Phi_1(\mathbf{x})\\ \vdots \\ \Phi_{C_I^m}(\mathbf{x})\end{bmatrix}. \qquad \square$$

EXAMPLE 2.11. *Let us illustrate Lemma 2.10 for $m = 2$ and $I = 4$. If $\mathbf{A} = \begin{bmatrix}a_{11} & a_{21} & a_{31} & a_{41}\end{bmatrix}^T$, then*

$$\begin{aligned}\mathcal{C}_2(\begin{bmatrix}\mathbf{A} & \mathbf{x}\end{bmatrix}) =& \mathcal{C}_2(\begin{bmatrix}a_{11} & x_1\\ a_{21} & x_2\\ a_{31} & x_3\\ a_{41} & x_4\end{bmatrix}) = \begin{bmatrix}x_2a_{11} - x_1a_{21}\\ x_3a_{11} - x_1a_{31}\\ x_4a_{11} - x_1a_{41}\\ x_3a_{21} - x_2a_{31}\\ x_4a_{21} - x_2a_{41}\\ x_4a_{31} - x_3a_{41}\end{bmatrix} = \begin{bmatrix}x_2 & -x_1 & 0 & 0\\ x_3 & 0 & -x_1 & 0\\ x_4 & 0 & 0 & -x_1\\ 0 & x_3 & -x_2 & 0\\ 0 & x_4 & 0 & -x_2\\ 0 & 0 & x_4 & -x_3\end{bmatrix}\begin{bmatrix}a_{11}\\ a_{21}\\ a_{31}\\ a_{41}\end{bmatrix}\\ =& \Phi^{4,2}(\mathbf{x})\mathcal{C}_1(\mathbf{A}).\end{aligned}$$

3. Basic implications. In this section we derive the implications in (1.10) and (1.11). We first establish scheme (1.10) by means of Lemmas 3.1, 3.2, 3.3, 3.4, 3.6, 3.7 and 3.8.

LEMMA 3.1. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $2 \leq m \leq \min(I, J)$. *Then condition* (Cm) *implies condition* (Um).

Proof. Since, by (Cm), $C_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has only the zero vector in its kernel, it does a forteriori not have an other vector in its kernel with the structure specified in (Um). □

LEMMA 3.2. *For* $\mathbf{A} \in \mathbb{F}^{I\times R}$ *and* $\mathbf{B} \in \mathbb{F}^{J\times R}$. *Then*

$$(\text{C}_1) \Leftrightarrow (\text{U}_1) \Leftrightarrow \mathbf{A} \odot \mathbf{B} \text{ has full column rank.}$$

Proof. The proof follows trivially from Lemma 2.3.1, since $\widehat{\mathbf{d}}^1 = \mathbf{d}$. □

LEMMA 3.3. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $1 \leq m \leq \min(I, J)$. *Then condition* (Um) *implies condition* (Wm) *for any matrix* $\mathbf{C} \in \mathbb{F}^{K\times R}$.

Proof. The proof trivially follows from the definitions of conditions (Um) and (Wm). □

LEMMA 3.4. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $1 < m \leq \min(I, J)$. *Then condition* (Km) *implies conditions* (Km-1),...,(K1).

Proof. Trivial. □

LEMMA 3.5. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $1 < m \leq \min(I, J)$. *Then condition* (Hm) *implies conditions* (Hm-1),...,(H1).

Proof. Trivial. □

LEMMA 3.6. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $1 < m \leq \min(I, J)$. *Then condition* (Cm) *implies conditions* (Cm-1),...,(C1).

Proof. It is sufficient to prove that (Ck) implies (Ck-1) for $k \in \{m, m-1, \ldots, 2\}$. Let us assume that there exists a vector $\mathbf{d}_{S_R^{k-1}} \in \mathbb{F}^{C_R^{k-1}}$ such that

$$[\mathcal{C}_{k-1}(\mathbf{A}) \odot \mathcal{C}_{k-1}(\mathbf{B})]\, \mathbf{d}_{S_R^{k-1}} = \mathbf{0},$$

which, by (2.5), is equivalent with

$$\mathcal{C}_{k-1}(\mathbf{A})\text{Diag}(\mathbf{d}_{S_R^{k-1}})\mathcal{C}_{k-1}(\mathbf{B})^T = \mathbf{O}.$$

Multiplying by matrices $\Phi^{I,k}(\mathbf{a}_r) \in \mathbb{F}^{C_I^k \times C_I^{k-1}}$ and $\Phi^{J,k}(\mathbf{b}_r) \in \mathbb{F}^{C_J^k \times C_J^{k-1}}$, constructed as in Lemma 2.10, we obtain

$$\Phi^{I,k}(\mathbf{a}_r)\mathcal{C}_{k-1}(\mathbf{A})\text{Diag}(\mathbf{d}_{S_R^{k-1}})\mathcal{C}_{k-1}(\mathbf{B})^T\Phi^{J,k}(\mathbf{b}_r)^T = \mathbf{O}, \qquad r = 1, \ldots, R,$$

which, by (2.5), is equivalent with

$$\left[\left(\Phi^{I,k}(\mathbf{a}_r)\mathcal{C}_{k-1}(\mathbf{A})\right) \odot \left(\Phi^{J,k}(\mathbf{b}_r)\mathcal{C}_{k-1}(\mathbf{B})\right)\right] \mathbf{d}_{S_R^{k-1}} = \mathbf{0}, \qquad r = 1, \ldots, R. \tag{3.1}$$

By (2.6),

$$\Phi^{I,k}(\mathbf{a}_r)\mathcal{C}_{k-1}([\mathbf{a}_{i_1} \ \ldots \ \mathbf{a}_{i_{k-1}}]) = \mathcal{C}_k([\mathbf{a}_{i_1} \ \ldots \ \mathbf{a}_{i_{k-1}} \ \mathbf{a}_r]) = \begin{cases} \mathbf{0}, & \text{if } r \in \{i_1, \ldots, i_{k-1}\}; \\ \pm\mathcal{C}_l(\mathbf{A})_{[i_1,i_2,\ldots,i_{k-1},r]}, & \text{if } r \notin \{i_1, \ldots, i_{k-1}\}, \end{cases} \tag{3.2}$$

where $\mathcal{C}_k(\mathbf{A})_{[i_1,i_2,\ldots,i_{k-1},r]}$ denotes the $[i_1,i_2,\ldots,i_{k-1},r]$-th column of the matrix $\mathcal{C}_k(\mathbf{A})$, in which $[i_1,i_2,\ldots,i_{k-1},r]$ denotes an ordered k-tuple. (Recall that by (2.2), the columns of $\mathcal{C}_k(\mathbf{A})$ can be enumerated with S_R^k.) Similarly,

$$\Phi^{J,k}(\mathbf{b}_r)\mathcal{C}_{k-1}([\mathbf{b}_{i_1}\ \ldots\ \mathbf{b}_{i_{k-1}}]) = \begin{cases} \mathbf{0}, & \text{if } r \in \{i_1,\ldots,i_{k-1}\}; \\ \pm\mathcal{C}_k(\mathbf{B})_{[i_1,i_2,\ldots,i_{k-1},r]}, & \text{if } r \notin \{i_1,\ldots,i_{k-1}\}. \end{cases} \tag{3.3}$$

Now, equations (3.1)–(3.3) yield

$$\sum_{\substack{1\le i_1<\cdots<i_{k-1}\le R \\ i_1,\ldots,i_{k-1}\neq r}} d_{(i_1,\ldots,i_{k-1})}\left(\mathcal{C}_k(\mathbf{A})_{[i_1,i_2,\ldots,i_{k-1},r]} \otimes \mathcal{C}_k(\mathbf{B})_{[i_1,i_2,\ldots,i_{k-1},r]}\right) = \mathbf{0}, \qquad r = 1,\ldots,R. \tag{3.4}$$

Since $\mathcal{C}_k(\mathbf{A}) \odot \mathcal{C}_k(\mathbf{B})$ has full column rank, it follows that for all $r = 1,\ldots,R$,

$$d_{(i_1,\ldots,i_{k-1})} = 0, \text{ whenever } 1 \le i_1 < \cdots < i_{k-1} \le R \text{ and } i_1,\ldots,i_{k-1} \neq r.$$

It immediately follows that $\mathbf{d}_{S_R^{k-1}} = \mathbf{0}$. Hence, $\mathcal{C}_{k-1}(\mathbf{A}) \odot \mathcal{C}_{k-1}(\mathbf{B})$ has full column rank. □

LEMMA 3.7. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $1 < m \le \min(I,J)$. *Then condition* (U$_m$) *implies conditions* (U$_{m-1}$),…,(U$_1$).

Proof. It is sufficient to prove that (U$_k$) implies (U$_{k-1}$) for $k \in \{m, m-1,\ldots,2\}$. Assume to the contrary that (U$_{k-1}$) does not hold. Then there exists a nonzero vector $\widehat{\mathbf{d}}^{k-1}$ such that $[\mathcal{C}_{k-1}(\mathbf{A}) \odot \mathcal{C}_{k-1}(\mathbf{B})]\,\widehat{\mathbf{d}}^{k-1} = \mathbf{0}$. Analogous to the proof of Lemma 3.6 we obtain that (3.4) holds with

$$d_{(i_1,\ldots,i_{k-1})} = d_{i_1}\cdots d_{i_{k-1}}, \qquad (i_1,\ldots,i_{k-1}) \in S_R^{k-1}. \tag{3.5}$$

Thus, multiplying the r-th equation from (3.4) by d_r, for $1 \le r \le R$, we obtain

$$\sum_{\substack{1\le i_1<\cdots<i_{k-1}\le R \\ i_1,\ldots,i_{k-1}\neq r}} d_{i_1}\cdots d_{i_{k-1}} d_r\left(\mathcal{C}_k(\mathbf{A})_{[i_1,i_2,\ldots,i_{k-1},r]} \otimes \mathcal{C}_k(\mathbf{B})_{[i_1,i_2,\ldots,i_{k-1},r]}\right) = \mathbf{0}. \tag{3.6}$$

Summation of (3.6) over r yields

$$k\,[\mathcal{C}_k(\mathbf{A}) \odot \mathcal{C}_k(\mathbf{B})]\,\widehat{\mathbf{d}}^{k} = \mathbf{0}. \tag{3.7}$$

Since (U$_k$) holds, (3.7) implies that

$$d_{i_1}\cdots d_{i_k} = 0, \qquad (i_1,\ldots,i_k) \in S_R^k.$$

Since $\widehat{\mathbf{d}}^{k-1}$ is nonzero, it follows that exactly $k-1$ of the R values $d_1,\ldots,d_R$ are different from zero. Therefore, $\widehat{\mathbf{d}}^{k-1}$ has exactly one nonzero component. It follows that the matrix $\mathcal{C}_{k-1}(\mathbf{A}) \odot \mathcal{C}_{k-1}(\mathbf{B})$ has a zero column. Hence, $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \le k-2$. On the other hand, Lemma 2.9 implies that $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge k$, which is a contradiction. □

The following lemma completes scheme (1.10).

LEMMA 3.8. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $m \le \min(I,J)$. *Then condition* (K$_m$) *implies condition* (C$_m$).

Proof. We give the proof for the case $r_{\mathbf{A}} + k_{\mathbf{B}} \geq R + m$ and $k_{\mathbf{A}} \geq m$; the case $r_{\mathbf{B}} + k_{\mathbf{A}} \geq R + m$ and $k_{\mathbf{B}} \geq m$ follows by symmetry. We obviously have $k_{\mathbf{B}} \geq m$.

In the case $k_{\mathbf{B}} = m$, we have $r_{\mathbf{A}} = R$. Lemma 2.4 (5) implies that the $C_I^m \times C_R^m$ matrix $C_m(\mathbf{A})$ has full column rank. The fact that $k_{\mathbf{B}} = m$ implies that every column of $C_m(\mathbf{B})$ contains at least one nonzero entry. It immediately follows that $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank.

We now consider the case $k_{\mathbf{B}} > m$.

(i) Suppose that $[\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\mathbf{d}_{S_R^m} = \mathbf{0}_{C_I^m C_J^m}$ for some vector $\mathbf{d}_{S_R^m} \in \mathbb{F}^{C_R^m}$. Then, by (2.5),

$$\mathcal{C}_m(\mathbf{A})\mathrm{Diag}(\mathbf{d}_{S_R^m})\mathcal{C}_m(\mathbf{B})^T = \mathbf{O}_{C_I^m \times C_J^m}. \tag{3.8}$$

(ii) Let us for now assume that the last $r_{\mathbf{A}}$ columns of $\mathbf{A}$ are linearly independent. We show that $d_{(k_{\mathbf{B}}-m+1,\ldots,k_{\mathbf{B}})} = 0$.

By definition of $k_{\mathbf{B}}$, the matrix $\mathbf{X} := \begin{bmatrix}\mathbf{b}_1 & \ldots & \mathbf{b}_{k_{\mathbf{B}}}\end{bmatrix}^T$ has full row rank. Hence, $\mathbf{X}\mathbf{X}^\dagger = \mathbf{I}_{k_{\mathbf{B}}}$, where $\mathbf{X}^\dagger$ denotes a right inverse of $\mathbf{X}$. Denoting

$$\mathbf{Y} := \mathbf{X}^\dagger \begin{bmatrix} \mathbf{O}_{(k_{\mathbf{B}}-m)\times m} \\ \mathbf{I}_m \end{bmatrix},$$

we have

$$\begin{aligned}\mathbf{B}^T\mathbf{Y} &= \begin{bmatrix} \mathbf{X} \\ \begin{bmatrix}\mathbf{b}_{k_{\mathbf{B}}+1} & \ldots & \mathbf{b}_R\end{bmatrix}^T \end{bmatrix} \mathbf{X}^\dagger \begin{bmatrix} \mathbf{O}_{(k_{\mathbf{B}}-m)\times m} \\ \mathbf{I}_m \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I}_{k_{\mathbf{B}}} \\ \boxplus_{(R-k_{\mathbf{B}})\times k_{\mathbf{B}}} \end{bmatrix} \begin{bmatrix} \mathbf{O}_{(k_{\mathbf{B}}-m)\times m} \\ \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{(k_{\mathbf{B}}-m)\times m} \\ \mathbf{I}_m \\ \boxplus_{(R-k_{\mathbf{B}})\times m} \end{bmatrix},\end{aligned}$$

where $\boxplus_{p\times q}$ denotes a $p \times q$ matrix that is not further specified. From the definition of the m-th compound matrix it follows that

$$\mathcal{C}_m(\mathbf{B}^T\mathbf{Y}) = \begin{bmatrix} \mathbf{0}_{C_R^m - C_{R-k_{\mathbf{B}}+m}^m} \\ 1 \\ \boxplus_{(C_{R-k_{\mathbf{B}}+m}^m - 1)\times 1} \end{bmatrix}. \tag{3.9}$$

We now have

$$\begin{aligned}\mathbf{0}_{C_I^m} &= \mathbf{O}_{C_I^m \times C_J^m} \cdot \mathcal{C}_m(\mathbf{Y}) \\ &\overset{(3.8)}{=} \mathcal{C}_m(\mathbf{A})\mathrm{Diag}(\mathbf{d}_{S_R^m})\mathcal{C}_m(\mathbf{B}^T) \cdot \mathcal{C}_m(\mathbf{Y}) \\ &= \mathcal{C}_m(\mathbf{A})\mathrm{Diag}(\mathbf{d}_{S_R^m})\mathcal{C}_m(\mathbf{B}^T\mathbf{Y}) \\ &\overset{(3.9)}{=} \mathcal{C}_m(\mathbf{A}) \begin{bmatrix} \mathbf{0}_{C_R^m - C_{R-k_{\mathbf{B}}+m}^m} \\ d_{(k_{\mathbf{B}}-m+1,\ldots,k_{\mathbf{B}})} \\ \boxplus_{(C_{R-k_{\mathbf{B}}+m}^m - 1)\times 1} \end{bmatrix}.\end{aligned} \tag{3.10}$$

Since the last $r_{\mathbf{A}}$ columns of $\mathbf{A}$ are linearly independent, Lemma 2.4 (5) implies that the $C_I^m \times C_{r_{\mathbf{A}}}^m$ matrix $\mathbf{M} = \mathcal{C}_m(\begin{bmatrix}\mathbf{a}_{R-r_{\mathbf{A}}+1} & \ldots & \mathbf{a}_R\end{bmatrix})$ has full column rank. By definition, $\mathbf{M}$ consists of the last $C_{r_{\mathbf{A}}}^m$ columns of $C_m(\mathbf{A})$. Obviously, $r_{\mathbf{A}} + k_{\mathbf{B}} \geq R + m$ implies $C_{r_{\mathbf{A}}}^m \geq C_{R-k_{\mathbf{B}}+m}^m$. Hence, the last $C_{R-k_{\mathbf{B}}+m}^m$ columns of $C_m(\mathbf{A})$ are linearly independent and the coefficient vector in (3.10) is zero. In particular, $d_{(k_{\mathbf{B}-m+1},\ldots,k_{\mathbf{B}})} = 0$.

(iii) We show that $d_{(j_1,\dots,j_m)} = 0$ for any choice of j_1, j_2, …, j_m, $1 \le j_1 < \cdots < j_m \le R$.
Since $k_{\mathbf{A}} \ge m$, the set of vectors $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_m}$ is linearly independent. Let us extend the set $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_m}$ to a basis of range($\mathbf{A}$) by adding $r_{\mathbf{A}} - m$ linearly independent columns of $\mathbf{A}$. Denote these basis vectors by $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_m}, \mathbf{a}_{j_{m+1}}, \dots, \mathbf{a}_{j_{r_{\mathbf{A}}}}$. It is clear that there exists an $R \times R$ permutation matrix $\mathbf{\Pi}$ such that the $(\mathbf{A\Pi})_{R-r_{\mathbf{A}}+1} = \mathbf{a}_{j_1}, \dots, (\mathbf{A\Pi})_R = \mathbf{a}_{j_{r_{\mathbf{A}}}}$, where here and in the sequel $(\mathbf{A\Pi})_r$ denotes the r-th column of the matrix $\mathbf{A\Pi}$. Moreover, since $k_{\mathbf{B}} - m + 1 \ge R - r_{\mathbf{A}} + 1$ we can choose $\mathbf{\Pi}$ such that it additionally satisfies $(\mathbf{A\Pi})_{k_{\mathbf{B}}-m+1} = \mathbf{a}_{j_1}$, $(\mathbf{A\Pi})_{k_{\mathbf{B}}-m+2} = \mathbf{a}_{j_2}$, …, $(\mathbf{A\Pi})_{k_{\mathbf{B}}} = \mathbf{a}_{j_m}$. We can now reason as under (ii) for $\mathbf{A\Pi}$ and $\mathbf{B\Pi}$ to obtain that $d_{(j_1,\dots,j_m)} = 0$.

(iv) From (iii) we immediately obtain that $\mathbf{d}_{S_R^m} = \mathbf{0}$. Hence, $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank.
□
We now give results that concern (Hm).

LEMMA 3.9. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $m \le \min(I, J)$. *Then condition* (Km) *implies condition* (Hm).

Proof. We give the proof for the case $r_{\mathbf{A}} + k_{\mathbf{B}} \ge R + m$ and $k_{\mathbf{A}} \ge m$; the case $r_{\mathbf{B}} + k_{\mathbf{A}} \ge R + m$ and $k_{\mathbf{B}} \ge m$ follows by symmetry. We obviously have $k_{\mathbf{B}} \ge m$.

(i) Suppose that $\delta \le m$. Then $r_{\tilde{\mathbf{A}}} = r_{\tilde{\mathbf{B}}} = \delta$. Hence, $H(\delta) = \delta$.

(ii) Suppose that $\delta \ge m$ and $\delta \ge k_{\mathbf{B}}$. Then $r_{\tilde{\mathbf{B}}} \ge k_{\mathbf{B}}$. Let $\tilde{\mathbf{A}}^c$ denote the $I \times (R - \delta)$ matrix obtained from $\mathbf{A}$ by removing the columns that are also in $\tilde{\mathbf{A}}$. Then $r_{\tilde{\mathbf{A}}} \ge r_{\mathbf{A}} - r_{\tilde{\mathbf{A}}^c}$. Hence, $r_{\tilde{\mathbf{A}}} + r_{\tilde{\mathbf{B}}} - \delta \ge r_{\mathbf{A}} - r_{\tilde{\mathbf{A}}^c} + r_{\tilde{\mathbf{B}}} - \delta \ge r_{\mathbf{A}} - (R-\delta) + k_{\mathbf{B}} - \delta \ge m$. Thus, $H(\delta) \ge m$.

(iii) Suppose that $k_{\mathbf{B}} \ge \delta \ge m$. Then $r_{\tilde{\mathbf{B}}} = \delta$. Since $r_{\tilde{\mathbf{A}}} \ge \min(\delta, k_{\mathbf{A}}) \ge m$, it follows that $r_{\tilde{\mathbf{A}}} + r_{\tilde{\mathbf{B}}} - \delta \ge m + \delta - \delta = m$. Thus, $H(\delta) \ge m$.
□

LEMMA 3.10. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $m \le \min(I, J)$. *Then condition* (Hm) *implies condition* (Um).

Proof. The following proof is based on the proof of Rank Lemma from [20, p. 121]. Let $(\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))\widehat{\mathbf{d}}^m = \mathbf{0}$ for $\widehat{\mathbf{d}}^m$ associated with $\mathbf{d} \in \mathbb{F}^R$. By Lemma 2.7, $\mathcal{C}_m(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T) = \mathcal{C}_m(\mathbf{B})\mathrm{Diag}(\widehat{\mathbf{d}}^m)\mathcal{C}_m(\mathbf{A})^T = \mathbf{O}$. Hence, $r_{\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T} \le m - 1$. Let $\omega(\mathbf{d}) = \delta$ and $d_{i_1} = \cdots = d_{i_{R-\delta}} = 0$. Form the matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$ and the vector $\tilde{\mathbf{d}}$ by dropping the columns of $\mathbf{A}$, $\mathbf{B}$ and the entries of $\mathbf{d}$ indexed by $i_1, \dots, i_{R-\delta}$. From the Sylvester rank inequality we obtain

$$\begin{aligned}\min(\delta, m) \le H(\delta) \le r_{\tilde{\mathbf{B}}} + r_{\tilde{\mathbf{A}}} - \delta &= r_{\tilde{\mathbf{B}}\mathrm{Diag}(\tilde{\mathbf{d}})} + r_{\mathrm{Diag}(\tilde{\mathbf{d}})\tilde{\mathbf{A}}} - r_{\mathrm{Diag}(\tilde{\mathbf{d}})} \\ &\le r_{\tilde{\mathbf{B}}\mathrm{Diag}(\tilde{\mathbf{d}})\tilde{\mathbf{A}}^T} = r_{\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T} \le m - 1.\end{aligned}$$

Hence, $\delta \le m - 1$. From Lemma 2.5 1) it follows that $\widehat{\mathbf{d}}^m = \mathbf{0}$. □

The remaining part of this section concerns (1.11).

Lemma 3.7 and Lemma 2.9 can be summarized as follows:

$$(\mathrm{Um}) \Rightarrow \begin{cases} (\mathrm{Uk}),\ k \le m; \\ \mathbf{A} \odot \mathbf{B} \ \text{ has full column rank } (\Leftrightarrow (\mathrm{U1})); \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m. \end{cases}$$

The following example demonstrates that similar implications do not necessarily hold for (Wm). Namely, in general, (Wm) does not imply any of the following conditions:

$$(\mathrm{Wk}) \text{ for } \ k \le m - 1,$$

$$\begin{aligned}&\mathbf{A}\odot\mathbf{B}\ \text{ has full column rank,}\\&\min(k_{\mathbf{A}},k_{\mathbf{B}})\geq m-1.\end{aligned}$$

EXAMPLE 3.11. *Let $m=2$ and let*

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

Let us show that condition (W2) *holds but condition* (W1) *does not hold. It is easy to check that*

$$\mathbf{A}\odot\mathbf{B}=\begin{bmatrix}1&0&0&1\\0&0&0&2\\0&0&0&1\\0&1&0&2\end{bmatrix},\quad \mathcal{C}_2(\mathbf{A})\odot\mathcal{C}_2(\mathbf{B})=\begin{bmatrix}1&0&2&0&1&0\end{bmatrix}.$$

Let $\mathbf{d}\in\operatorname{range}(\mathbf{C}^T)$. *Then there exist* $x_1,x_2\in\mathbb{F}$ *such that* $\mathbf{d}=\begin{bmatrix}x_2&x_2&x_1&x_2\end{bmatrix}^T$. *Hence,* $\widehat{\mathbf{d}}^2=\begin{bmatrix}x_2^2&x_1x_2&x_2^2&x_1x_2&x_2^2&x_1x_2\end{bmatrix}^T$. *Therefore*

$$\begin{aligned}&(\mathcal{C}_2(\mathbf{A})\odot\mathcal{C}_2(\mathbf{B}))\widehat{\mathbf{d}}^2=\mathbf{0}\Leftrightarrow x_2=0\Rightarrow\omega(\mathbf{d})\leq 1=2-1\Rightarrow\text{(W2) } \textit{holds},\\&(\mathbf{A}\odot\mathbf{B})\mathbf{e}_3^4=\mathbf{0},\quad \mathbf{e}_3^4\in\operatorname{range}(\mathbf{C}^T),\quad \omega(\mathbf{e}_3^4)=1>1-1\Rightarrow\text{(W1) } \textit{does not hold},\end{aligned}$$

where $\mathbf{e}_3^4=\begin{bmatrix}0&0&1&0\end{bmatrix}^T$. *In particular,* $\mathbf{A}\odot\mathbf{B}$ *does not have full column rank. Besides, since the matrix* $\mathbf{A}$ *has a zero column, it follows that* $\min(k_{\mathbf{A}},k_{\mathbf{B}})=0<m-1$.

The following lemma now establishes (1.11).

LEMMA 3.12. *Let* $\mathbf{A}\in\mathbb{F}^{I\times R}$, $\mathbf{B}\in\mathbb{F}^{J\times R}$, $\mathbf{C}\in\mathbb{F}^{K\times R}$, $1<m\leq\min(I,J)$, *and* $\min(k_{\mathbf{A}},k_{\mathbf{B}})\geq m-1$. *Then condition* (Wm) *implies conditions* (Wm-1),...,(W1).

Proof. The proof is the same as the proof of Lemma 3.7, with the difference that instead of Lemma 2.9 we use the condition $\min(k_{\mathbf{A}},k_{\mathbf{B}})\geq m-1$. ∎

4. Sufficient conditions for the uniqueness of one factor matrix. In this section we establish conditions under which a PD is canonical, with one of the factor matrices unique. We have the following formal definition.

DEFINITION 4.1. *Let $\mathcal{T}$ be a tensor of rank R. The first (resp. second or third) factor matrix of $\mathcal{T}$ is* unique *if $\mathcal{T}=[\mathbf{A},\mathbf{B},\mathbf{C}]_R=[\bar{\mathbf{A}},\bar{\mathbf{B}},\bar{\mathbf{C}}]_R$ implies that there exist an $R\times R$ permutation matrix $\mathbf{\Pi}$ and an $R\times R$ nonsingular diagonal matrix $\mathbf{\Lambda_A}$ (resp. $\mathbf{\Lambda_B}$ or $\mathbf{\Lambda_C}$) such that $\bar{\mathbf{A}}=\mathbf{A\Pi\Lambda_A}$ (resp. $\bar{\mathbf{B}}=\mathbf{B\Pi\Lambda_B}$ or $\bar{\mathbf{C}}=\mathbf{C\Pi\Lambda_C}$).*

4.1. Conditions based on (Um), (Cm), (Hm), and (Km). First, we recall Kruskal's permutation lemma, which we will use in the proof of Proposition 4.3.

LEMMA 4.2. *[16, 20, 32] Consider two matrices $\bar{\mathbf{C}}\in\mathbb{F}^{K\times\bar{R}}$ and $\mathbf{C}\in\mathbb{F}^{K\times R}$ such that $\bar{R}\leq R$ and $k_{\mathbf{C}}\geq 1$. If for every vector $\mathbf{x}$ such that $\omega(\bar{\mathbf{C}}^T\mathbf{x})\leq\bar{R}-r_{\bar{\mathbf{C}}}+1$, we have $\omega(\mathbf{C}^T\mathbf{x})\leq\omega(\bar{\mathbf{C}}^T\mathbf{x})$, then $\bar{R}=R$ and there exist a unique permutation matrix $\mathbf{\Pi}$ and a unique nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\bar{\mathbf{C}}=\mathbf{C\Pi\Lambda}$.*

We start the derivation of (1.12) with the proof of the following proposition.

PROPOSITION 4.3. *Let $\mathbf{A}\in\mathbb{F}^{I\times R}$, $\mathbf{B}\in\mathbb{F}^{J\times R}$, $\mathbf{C}\in\mathbb{F}^{K\times R}$, and let $\mathcal{T}=[\mathbf{A},\mathbf{B},\mathbf{C}]_R$. Assume that*

(i) $k_{\mathbf{C}}\geq 1$;
(ii) $m=R-r_{\mathbf{C}}+2\leq\min(I,J)$;
(iii) *condition* (Um) *holds.*

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

Proof. Let $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_{\bar{R}}$ be a CPD of $\mathcal{T}$, which implies $\bar{R} \leq R$. We have $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T$. We check that the conditions of Lemma 4.2 are satisfied. From Lemma 3.7 it follows that conditions $(\mathrm{U}_{m\text{-}1}), \ldots, (\mathrm{U}_2), (\mathrm{U}_1)$ hold. The fact that (U_1) holds, means that $\mathbf{A} \odot \mathbf{B}$ has full column rank. Hence,

$$r_{\mathbf{C}} = r_{\mathbf{C}^T} = r_{(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T} = r_{(\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T} \leq r_{\bar{\mathbf{C}}^T} = r_{\bar{\mathbf{C}}}. \tag{4.1}$$

Consider any vector $\mathbf{x} \in \mathbb{F}^K$ such that

$$\omega(\bar{\mathbf{C}}^T\mathbf{x}) := k - 1 \leq \bar{R} - r_{\bar{\mathbf{C}}} + 1,$$

as in Lemma 4.2. Then $r_{\bar{\mathbf{A}}\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{x})\bar{\mathbf{B}}^T} \leq k - 1$ and, by (4.1),

$$\bar{R} - r_{\bar{\mathbf{C}}} + 1 \leq R - r_{\mathbf{C}} + 1 = m - 1,$$

which implies $k \leq m$. We have $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{x} = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T\mathbf{x}$ so $\mathbf{A}\mathrm{Diag}(\mathbf{C}^T\mathbf{x})\mathbf{B}^T = \bar{\mathbf{A}}\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{x})\bar{\mathbf{B}}^T$. From Lemma 2.4(1) it follows

$$\begin{aligned}\mathcal{C}_k(\mathbf{A}\mathrm{Diag}(\mathbf{C}^T\mathbf{x})\mathbf{B}^T) &= \mathcal{C}_k(\bar{\mathbf{A}}\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{x})\bar{\mathbf{B}}^T)\\ &= \mathcal{C}_k(\bar{\mathbf{A}})\mathcal{C}_k(\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{x}))\mathcal{C}_k(\bar{\mathbf{B}}^T)\\ &= \mathbf{O},\end{aligned}$$

in which the latter equality follows from Lemma 2.5(1). Hence, by Lemma 2.7,

$$(\mathcal{C}_k(\mathbf{A}) \odot \mathcal{C}_k(\mathbf{B}))\widehat{\mathbf{d}}^k = \mathbf{0}$$

for $\mathbf{d} := \mathbf{C}^T\mathbf{x} \in \mathbb{F}^R$. Since condition (U_k) holds for $\mathbf{A}$ and $\mathbf{B}$, it follows that $\omega(\mathbf{C}^T\mathbf{x}) \leq k - 1 = \omega(\bar{\mathbf{C}}^T\mathbf{x})$. Hence, by Lemma 4.2, $\bar{R} = R$ and the matrices $\mathbf{C}$ and $\bar{\mathbf{C}}$ are the same up to permutation and column scaling. ∎

The implications $(\mathrm{C}_m) \Rightarrow (\mathrm{U}_m)$ and $(\mathrm{H}_m) \Rightarrow (\mathrm{U}_m)$ in scheme (1.10) lead to Corollary 4.4 and to Theorem 1.10, respectively. The implication $(\mathrm{K}_m) \Rightarrow (\mathrm{C}_m)$ in turn leads to Corollary 4.5. Clearly, conditions (C_m), (H_m), and (K_m) are more restrictive than (U_m). On the other hand, they may be easier to verify.

COROLLARY 4.4. *Let $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, $\mathbf{C} \in \mathbb{F}^{K\times R}$, and let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. Assume that*

(i) $k_{\mathbf{C}} \geq 1$;
(ii) $m = R - r_{\mathbf{C}} + 2 \leq \min(I, J)$;
(iii) $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ *has full column rank.* (C_m)

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

COROLLARY 4.5. *Let $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, $\mathbf{C} \in \mathbb{F}^{K\times R}$, and $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. Let also $m := R - r_{\mathbf{C}} + 2$. If*

$$\left\{\begin{array}{rl} r_{\mathbf{A}} + k_{\mathbf{B}} & \geq R + m,\\ k_{\mathbf{A}} & \geq m,\\ k_{\mathbf{C}} & \geq 1\end{array}\right. \quad \textit{or} \quad \left\{\begin{array}{rl} r_{\mathbf{B}} + k_{\mathbf{A}} & \geq R + m,\\ k_{\mathbf{B}} & \geq m,\\ k_{\mathbf{C}} & \geq 1,\end{array}\right. \tag{4.2}$$

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

REMARK 4.6. *Condition* (ii) *in Proposition 4.3 and Corollary 4.4 guarantees that the matrices $\mathcal{C}_m(\mathbf{A})$ and $\mathcal{C}_m(\mathbf{B})$ are defined. In Corollary 4.5,* (K_m) *cannot hold if $m = R - r_{\mathbf{C}} + 2 > \min(I, J)$.*

COROLLARY 4.7. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, $\mathbf{C} \in \mathbb{F}^{K\times R}$, *and* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Let also* $m := R - r_{\mathbf{C}} + 2$. *If*

$$\begin{cases} k_{\mathbf{C}} \geq 1, \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m, \\ \max(r_{\mathbf{A}} + k_{\mathbf{B}}, r_{\mathbf{B}} + k_{\mathbf{A}}) \geq R + m, \end{cases} \tag{4.3}$$

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

Proof. It can easily be checked that (4.3) and (4.2) are equivalent. ∎

REMARK 4.8. *It is easy to see that Corollaries 4.5 and 4.7 are equivalent to Theorem 1.9. Indeed, if* $m = R - r_{\mathbf{C}} + 2$, *then*

$$(1.7) \Leftrightarrow \begin{cases} k_{\mathbf{C}} \geq 1, \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m, \\ k_{\mathbf{A}} + k_{\mathbf{B}} + \max(r_{\mathbf{A}} - k_{\mathbf{A}}, r_{\mathbf{B}} - k_{\mathbf{B}}) \geq R + m \end{cases} \Leftrightarrow (4.3) \Leftrightarrow (4.2).$$

4.2. Conditions based on (W$_m$). In this subsection we deal with condition (W$_m$). Similar to condition (U$_m$) in Proposition 4.3, condition (W$_m$) will in Proposition 4.9 imply the uniqueness of one factor matrix. However, condition (W$_m$) is more relaxed than condition (U$_m$). Like condition (U$_m$), condition (W$_m$) may be hard to check. We give an example in which the uniqueness of one factor matrix can nevertheless be demonstrated using condition (W$_m$).

PROPOSITION 4.9. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, $\mathbf{C} \in \mathbb{F}^{K\times R}$, *and let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Assume that*

(i) $k_{\mathbf{C}} \geq 1$;
(ii) $m = R - r_{\mathbf{C}} + 2 \leq \min(I, J)$;
(iii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank;*
(iv) *conditions* (W$_m$), ..., (W$_1$) *hold.*

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

Proof. The proof is analogous to the proof of Proposition 4.3, with two points of difference. Namely, the fact that $\mathbf{A} \odot \mathbf{B}$ has full column rank does not follow from (W$_1$) but is assumed in condition (iii). Second, $\omega(\mathbf{C}^T\mathbf{x}) \leq \omega(\bar{\mathbf{C}}^T\mathbf{x})$ follows from (W$_k$) instead of (U$_k$). ∎

From Lemmas 3.7 and 3.3 it follows that Proposition 4.9 is more relaxed than Proposition 4.3.

Combining Proposition 4.9 and Lemma 3.12 we obtain the following result, which completes the derivation of scheme (1.12).

COROLLARY 4.10. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, $\mathbf{C} \in \mathbb{F}^{K\times R}$, *and let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Assume that*

(i) $k_{\mathbf{C}} \geq 1$;
(ii) $m = R - r_{\mathbf{C}} + 2 \leq \min(I, J)$;
(iii) $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m - 1$;
(iv) $\mathbf{A} \odot \mathbf{B}$ *has full column rank;*
(v) *condition* (W$_m$) *holds.*

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

EXAMPLE 4.11. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_7$, *with*

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We have

$$k_{\mathbf{A}} = k_{\mathbf{B}} = 4, \quad r_{\mathbf{A}} = r_{\mathbf{B}} = 6, \quad k_{\mathbf{C}} = 1, \quad r_{\mathbf{C}} = 4, \quad m = 5.$$

Since $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) < m$, *it follows from Lemma 2.9 that condition* (U$_{\rm m}$) *does not hold. We show that, on the other hand, condition* (W$_{\rm m}$) *does hold. One can easily check that the rank of the* 36×21 *matrix* $\mathbf{U} = \mathcal{C}_5(\mathbf{A}) \odot \mathcal{C}_5(\mathbf{B})$ *is equal to* 19. *Obviously, both the* $(1,2,3,4,5)$*-th and the* $(1,4,5,6,7)$*-th column of* $\mathbf{U}$ *are equal to zero. Hence, if* $\mathbf{U}\widehat{\mathbf{d}}^5 = \mathbf{0}$ *for* $\mathbf{d} = \begin{bmatrix} d_1 & \dots & d_7 \end{bmatrix}$, *then at most the two entries* $d_1d_2d_3d_4d_5$ *and* $d_1d_4d_5d_6d_7$ *of the vector* $\widehat{\mathbf{d}}^5$ *are nonzero. Consequently, for a nonzero vector* $\widehat{\mathbf{d}}^5$ *we have*

$$\begin{cases} d_2 = d_3 = 0, \\ d_1d_4d_5d_6d_7 \neq 0 \end{cases} \quad or \quad \begin{cases} d_6 = d_7 = 0, \\ d_1d_2d_3d_4d_5 \neq 0. \end{cases} \tag{4.4}$$

On the other hand, since $\mathbf{d} \in \operatorname{range}(\mathbf{C}^T)$, *there exists* $\mathbf{x} \in \mathbb{F}^4$ *such that*

$$\mathbf{d} = \mathbf{C}^T\mathbf{x} = \begin{bmatrix} x_1 + x_4 & x_2 & x_3 & x_1 & x_2 & x_3 & x_4 \end{bmatrix}. \tag{4.5}$$

One can easily check that set (4.4) *does not have solutions of the form* (4.5). *Thus, condition* (W$_5$) *holds.*

Corollary 1.18 implies that $\mathbf{A} \odot \mathbf{B}$ *has full column rank. Thus, by Corollary 4.10,* $r_{\mathcal{T}} = 7$ *and the third factor matrix of* $\mathcal{T}$ *is unique.*

Note that, since $k_{\mathbf{C}} = 1$, *it follows from Theorem 1.14 that the CPD* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_7$ *is not unique.*

5. Overall CPD uniqueness.

5.1. At least one factor matrix has full column rank. The results for the case $r_{\mathbf{C}} = R$ are well-studied. They are summarized in (1.8). In particular, Theorems 1.11, 1.12 and 1.13 present (U$_2$), (C$_2$) and (K$_2$), respectively, as sufficient conditions for CPD uniqueness.

The implications (i)⇔(ii) ⇔(iii) in Theorem 1.11 were proved in [16]. The core implication is (ii)⇒(iii). This implication follows almost immediately from Proposition 4.3, which establishes uniqueness of $\mathbf{C}$, as we show below. Implication (iii)⇒(i) follows from Theorem 1.15. Together with an explanation of the equivalence (i)⇔(ii) we obtain a short proof of Theorem 1.11.

Next, Theorem 1.12 follows immediately from Theorem 1.11. Theorem 1.13 in turn follows immediately from Theorem 1.12, cf. scheme (1.10).

Proof of Theorem 1.11. (i)⇔(ii): Follows from Lemma 2.8 for $m = 2$.

(ii)⇒(iii): By Proposition 4.3, $r_{\mathcal{T}} = R$ and the third factor matrix of $\mathcal{T}$ is unique. That is, for any CPD $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_R$ there exists a permutation matrix $\mathbf{\Pi}$ and a nonsingular diagonal matrix $\mathbf{\Lambda_C}$ such that $\bar{\mathbf{C}} = \mathbf{C\Pi_C\Lambda_C}$.

Then, by (1.4), $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\mathbf{\Lambda_C\Pi_C^T C}^T$. Since the matrix $\mathbf{C}$ has full column rank, it follows that $\mathbf{A} \odot \mathbf{B} = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\mathbf{\Lambda_C\Pi_C^T}$. Equating columns, we obtain that there exist nonsingular diagonal matrices $\mathbf{\Lambda_A}$ and $\mathbf{\Lambda_B}$ such that $\bar{\mathbf{A}} = \mathbf{A\Pi_C\Lambda_A}$ and $\bar{\mathbf{B}} = \mathbf{B\Pi_C\Lambda_B}$, with $\mathbf{\Lambda_A\Lambda_B\Lambda_C} = \mathbf{I}_R$. Hence, the CPD of $\mathcal{T}$ is unique.

(iii)⇒(i): follows from Theorem 1.15. □

Proof of Theorem 1.12: Condition (C2) in Theorem 1.12 trivially implies condition (U2) in Theorem 1.11. Hence, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Proof of Theorem 1.13: By Lemma 3.8, condition (K2) in Theorem 1.13 implies condition (C2). Hence, by Theorem 1.12, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

REMARK 5.1. *The results obtained in [34] (see the beginning of subsection 1.2.4) can be completed as follows. Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $2 \leq r_{\mathcal{T}} = R \leq 4$. *Assume without loss of generality that* $r_{\mathbf{C}} \geq \max(r_{\mathbf{A}}, r_{\mathbf{B}})$. *For such low tensor rank, we have now a uniqueness condition that is both necessary and sufficient. That condition is that* $r_{\mathbf{C}} = R$ *and* (U2) *holds. Also, for* $R \leq 3$, (K2), (H2), (C2), *and* (U2) *are equivalent. For these values of R, condition* (K2) *is the easiest one to check. For* $R = 4$, (H2), (C2), *and* (U2) *are equivalent. The proofs are based on a check of all possibilities and are therefore omitted.*

5.2. No factor matrix is required to have full column rank. Kruskal's original proof (also the simplified version in [32]) of Theorem 1.8 consists of three main steps. The first step is the proof of the permutation lemma (Lemma 4.2). The second and the third step concern the following two implications:

$$\begin{aligned} & k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2 \\ & \Rightarrow \begin{cases} k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \\ r_{\mathcal{T}} = R, \\ \text{every factor matrix in } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is by itself unique} \end{cases} \\ & \Rightarrow \text{the overall CPD } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is unique.} \end{aligned} \tag{5.1}$$

That is, the proof goes via demonstrating that individual factor matrices are unique. Similarly, the proof of uniqueness result (ii)⇔(iii) in Theorem 1.11 for the case $r_{\mathbf{C}} = R$, goes, via Proposition 4.3, in two steps, which correspond to the proofs of the following two equivalences:

$$\begin{aligned} \text{(U2)} \Leftrightarrow & \begin{cases} r_{\mathcal{T}} = R, \\ \text{the third factor matrix of } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is unique} \end{cases} \\ \Leftrightarrow & \text{ the CPD of } \mathcal{T} \text{ is unique.} \end{aligned} \tag{5.2}$$

Again the proof goes via demonstrating that one factor matrix is unique. Note that the second equivalence in (5.2) is almost immediate since $\mathbf{C}$ has full column rank. In contrast, the proof of the second implication in (5.1) is not trivial.

Scheme (1.12) generalizes the first implications in (5.1) and (5.2). What remains for the demonstration of overall CPD uniqueness, is the generalization of the second implications. This problem is addressed in Part II [8]. Namely, part of the discussion

in [8] is about investigating how, in cases where possibly none of the factor matrices has full column rank, uniqueness of one or more factor matrices implies overall uniqueness.

6. Conclusion. We have given an overview of conditions guaranteeing the uniqueness of one factor matrix in a PD or uniqueness of an overall CPD. We have discussed properties of compound matrices and used them to build the schemes of implications (1.10) and (1.11). For the case $r_{\mathbf{C}} = R$ we have demonstrated the overall CPD uniqueness results in (1.8) using second compound matrices. Using (1.10) and (1.11) we have obtained relaxed conditions guaranteeing the uniqueness of one factor matrix, for instance $\mathbf{C}$. The general idea is to the relax the condition on $\mathbf{C}$, no longer requiring that it has full column rank, while making the conditions on $\mathbf{A}$ and $\mathbf{B}$ more restrictive. The latter are conditions on the Khatri-Rao product of m-th compound matrices of $\mathbf{A}$ and $\mathbf{B}$, where $m > 2$. In Part II [8] we will use the results to derive relaxed conditions guaranteeing the uniqueness of the overall CPD.

7. Acknowledgments. The authors would like to thank the anonymous reviewers for their valuable comments and their suggestions to improve the presentation of the paper. The authors are also grateful for useful suggestions from Professor A. Stegeman (University of Groningen, The Netherlands).

REFERENCES

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

[2] A. Cichocki, R. Zdunek, A. H. Phan, and S. Amari, *Nonnegative Matrix and Tensor Factorizations - Applications to Exploratory Multi-way Data Analysis and Blind Source Separation.*, Wiley, 2009.

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

[4] P. Comon, X. Luciani, and A. L. F. de Almeida, *Tensor decompositions, alternating least squares and other tales*, J. Chemometrics, 23 (2009), pp. 393–405.

[5] 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.

[6] ———, *Blind separation of exponential polynomials and the decomposition of a tensor in rank–$(L_r, L_r, 1)$ terms*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1451–1474.

[7] ———, *A short introduction to tensor-based methods for Factor Analysis and Blind Source Separation*, in ISPA 2011: Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis, (2011), pp. 558–563.

[8] I. Domanov and L. De Lathauwer, *On the Uniqueness of the Canonical Polyadic Decomposition of third-order tensors — Part II: Uniqueness of the overall decomposition*, ESAT-SISTA Internal Report, 12-72, Leuven, Belgium: Department of Electrical Engineering (ESAT), KU Leuven, (2012).

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

[10] X. Guo, S. Miron, D. Brie, S. Zhu, and X. Liao, *A CANDECOMP/PARAFAC perspective on uniqueness of DOA estimation using a vector sensor array*, IEEE Trans. Signal Process., 59 (2011), pp. 3475–3481.

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

[12] ———, *Determination and Proof of Minimum Uniqueness Conditions for PARAFAC1*, UCLA Working Papers in Phonetics, 22 (1972), pp. 111–117.

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

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

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.
[16] T. Jiang and N. D. Sidiropoulos, *Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints*, IEEE Trans. Signal Process., 52 (2004), pp. 2625–2636.
[17] T. G. Kolda and B. W. Bader, *Tensor Decompositions and Applications*, SIAM Review, 51 (2009), pp. 455–500.
[18] W. P. Krijnen, *The analysis of three-way arrays by constrained Parafac methods*, DSWO Press, Leiden, 1991.
[19] P. M. Kroonenberg, *Applied Multiway Data Analysis*, Hoboken, NJ: Wiley, 2008.
[20] 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.
[21] J. B. Kruskal, *Rank, decomposition, and uniqueness for 3-way and n-way arrays*, in Multiway Data Analysis, R. Coppi and S. Bolasco., eds., Elsevier, North-Holland, (1989), pp. 7–18.
[22] J. M. Landsberg, *Tensors: Geometry and Applications*, AMS, Providence, Rhode Island, 2012.
[23] L.-H. Lim and P. Comon, *Multiarray signal processing: Tensor decomposition meets compressed sensing*, C.R. Mec., 338 (2010), pp. 311–320.
[24] J. Möcks, *Topographic components model for event-related potentials and some biophysical considerations*, IEEE Trans. Biomed. Eng., 35 (1988), pp. 482–484.
[25] J. A. Rhodes, *A concise proof of Kruskal's theorem on tensor decomposition*, Linear Algebra Appl., 432 (2010), pp. 1818–1824.
[26] N. D. Sidiropoulos and R. Bro, *On the uniqueness of multilinear decomposition of N-way arrays*, J. Chemometrics, 14 (2000), pp. 229–239.
[27] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, *Parallel Factor Analysis in Sensor Array Processing*, IEEE Trans. Signal Process., 48 (2000), pp. 2377–2388.
[28] N. D. Sidiropoulos and L. Xiangqian, *Identifiability results for blind beamforming in incoherent multipath with small delay spread*, IEEE Trans. Signal Process., 49 (2001), pp. 228–236.
[29] A.K. Smilde, R. Bro, and P. Geladi, *Multi-way analysis with applications in the chemical sciences*, J. Wiley, 2004.
[30] L. Sorber, M. Van Barel, and L. De Lathauwer, *Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-(L_r,L_r,1) terms and a new generalization*, ESAT-SISTA Internal Report, 12-37, Leuven, Belgium: Department of Electrical Engineering (ESAT), KU Leuven, (2012).
[31] A. Stegeman, *On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode*, Linear Algebra Appl., 431 (2009), pp. 211–227.
[32] A. Stegeman and N. D. Sidiropoulos, *On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition*, Linear Algebra Appl., 420 (2007), pp. 540–552.
[33] V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl., 52–53 (1983), pp. 645–685.
[34] J. Ten Berge and N. D. Sidiropoulos, *On uniqueness in Candecomp/Parafac*, Psychometrika, 67 (2002), pp. 399–409.
[35] J. M. F. Ten Berge, *The k-rank of a Khatri–Rao product*, Unpublished Note, Heijmans Institute of Psychological Research, University of Groningen, The Netherlands, (2000).
[36] L. Xiangqian and N. D. Sidiropoulos, *Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays*, IEEE Trans. Signal Process., 49 (2001), pp. 2074–2086.