Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 420 (2007) 540–552
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition

Alwin Stegeman [a,*,1], Nicholas D. Sidiropoulos [b]

[a] *Heijmans Institute of Psychological Research, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands*

[b] *Department of Electronic and Computer Engineering, Technical University of Crete, Kounoupidiana Campus, Chania – Crete 731 00, Greece*

Received 21 March 2005; accepted 4 August 2006
Available online 13 October 2006
Submitted by V. Mehrmann

Abstract

Let $\underline{\mathbf{X}}$ be a real-valued three-way array. The Candecomp/Parafac (CP) decomposition is written as $\underline{\mathbf{X}} = \underline{\mathbf{Y}}^{(1)} + \cdots + \underline{\mathbf{Y}}^{(R)} + \underline{\mathbf{E}}$, where $\underline{\mathbf{Y}}^{(r)}$ are rank-1 arrays and $\underline{\mathbf{E}}$ is a rest term. Each rank-1 array is defined by the outer product of three vectors $\mathbf{a}^{(r)}$, $\mathbf{b}^{(r)}$ and $\mathbf{c}^{(r)}$, i.e. $y_{ijk}^{(r)} = a_i^{(r)}\, b_j^{(r)}\, c_k^{(r)}$. These vectors make up the R columns of the component matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. If $2R + 2$ is less than or equal to the sum of the k-ranks of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, then the fitted part of the decomposition is unique up to a change in the order of the rank-1 arrays and rescaling/counterscaling of each triplet of vectors $(\mathbf{a}^{(r)}, \mathbf{b}^{(r)}, \mathbf{c}^{(r)})$ forming a rank-1 array. This classical result was shown by Kruskal. His proof is, however, rather inaccessible and does not seem intuitive. In order to contribute to a better understanding of CP uniqueness, this paper provides an accessible and intuitive proof of Kruskal's condition. The proof is both self-contained and compact and can easily be adapted for the complex-valued CP decomposition.

Keywords: Candecomp; Parafac; Three-way arrays; Uniqueness; Kruskal-rank condition

* Corresponding author. Tel.: +31 50 363 6193; fax: +31 50 363 6304.
E-mail addresses: a.w.stegeman@rug.nl (A. Stegeman), nikos@telecom.tuc.gr (N.D. Sidiropoulos).
[1] The author is supported by the Dutch Organisation for Scientific Research (NWO), VENI grant 451-04-102.

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.08.010

1. Introduction

The decomposition of three-way (or three-mode) arrays into rank-one three-way outer products was proposed independently by Carroll and Chang [1] (who called it Candecomp) and Harshman [2] (who called it Parafac). Candecomp/Parafac (CP) decomposes a three-way array $\underline{\mathbf{X}}$ of order $I \times J \times K$ into a fixed number of R components $\underline{\mathbf{Y}}^{(r)}$, $r = 1, \ldots, R$, and a residual term $\underline{\mathbf{E}}$, i.e.

$$\underline{\mathbf{X}} = \sum_{r=1}^{R} \underline{\mathbf{Y}}^{(r)} + \underline{\mathbf{E}}. \tag{1.1}$$

In the sequel, we will denote three-way arrays as $\underline{\mathbf{Z}}$, matrices as $\mathbf{Z}$, vectors as $\mathbf{z}$ and scalars as z. We assume all arrays, matrices, vectors and scalars to be real-valued.

CP has its origins in psychology, where it was conceived primarily as an exploratory data analysis tool. Later, CP attracted considerable interest in chemistry (see Smilde et al. [10] and references therein), and, more recently, in signal processing and telecommunications research (see Sidiropoulos [9] and references therein). For example, CP is appropriate for the analysis of fluorescence excitation–emission measurements. CP is also useful in the context of certain parameter and signal estimation problems in wireless communications, including emitter localization, carrier frequency offset estimation, and the separation of spread-spectrum communication signals. This renewed interest has helped to sustain advances in both theory and applications of CP.

Each component $\underline{\mathbf{Y}}^{(r)}$ is the outer product of three vectors $\mathbf{a}^{(r)}$, $\mathbf{b}^{(r)}$ and $\mathbf{c}^{(r)}$, i.e. $y_{ijk}^{(r)} = a_i^{(r)}\, b_j^{(r)}\, c_k^{(r)}$. This implies that each of the R components has three-way rank 1. Analogous to matrix algebra, the three-way rank of $\underline{\mathbf{X}}$ is defined (see Kruskal [5]) as the smallest number of rank-1 arrays whose sum equals $\underline{\mathbf{X}}$. Since there are no restrictions on the vectors $\mathbf{a}^{(r)}$, $\mathbf{b}^{(r)}$ and $\mathbf{c}^{(r)}$, the array $\underline{\mathbf{X}}$ has three-way rank R if and only if R is the smallest number of components for which a CP decomposition (1.1) exists with perfect fit, i.e. with an all-zero residual term $\underline{\mathbf{E}}$. For a fixed value of R, the CP decomposition (1.1) is found by minimizing the sum of squares of $\underline{\mathbf{E}}$. It may be noted that the CP decomposition is a special case of the three-mode principal component model of Tucker [15], which reduces to CP when the Tucker3 core array is $R \times R \times R$ and superdiagonal.

A CP solution is usually expressed in terms of the component matrices $\mathbf{A}$ $(I \times R)$, $\mathbf{B}$ $(J \times R)$ and $\mathbf{C}$ $(K \times R)$, which have as columns the vectors $\mathbf{a}^{(r)}$, $\mathbf{b}^{(r)}$ and $\mathbf{c}^{(r)}$, respectively. Let $\mathbf{X}_k$ $(I \times J)$ and $\mathbf{E}_k$ $(I \times J)$ denote the kth slices of $\underline{\mathbf{X}}$ and $\underline{\mathbf{E}}$, respectively. Now (1.1) can be written as

$$\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\, \mathbf{B}^{\mathrm{T}} + \mathbf{E}_k, \quad k = 1, \ldots, K, \tag{1.2}$$

where $\mathbf{C}_k$ is the diagonal matrix with the kth row of $\mathbf{C}$ as its diagonal.

The uniqueness of a CP solution is usually studied for given residuals $\underline{\mathbf{E}}$. It can be seen that the fitted part of a CP decomposition, i.e. a full decomposition of $\underline{\mathbf{X}} - \underline{\mathbf{E}}$, can only be unique up to rescaling/counterscaling and jointly permuting columns of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. For instance, suppose the rth columns of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are multiplied by scalars λ_a, λ_b and λ_c, respectively, and there holds $\lambda_a \lambda_b \lambda_c = 1$. Then the rth CP component $\underline{\mathbf{Y}}^{(r)}$ remains unchanged. Furthermore, a joint permutation of the columns of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ amounts to a new order of the R components. Hence, the residuals $\underline{\mathbf{E}}$ will be the same for the solution $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ as for the solution $(\mathbf{A}\boldsymbol{\Pi}\boldsymbol{\Lambda}_a, \mathbf{B}\boldsymbol{\Pi}\boldsymbol{\Lambda}_b, \mathbf{C}\boldsymbol{\Pi}\boldsymbol{\Lambda}_c)$, where $\boldsymbol{\Pi}$ is a permutation matrix and $\boldsymbol{\Lambda}_a$, $\boldsymbol{\Lambda}_b$ and $\boldsymbol{\Lambda}_c$ are diagonal matrices such that $\boldsymbol{\Lambda}_a\boldsymbol{\Lambda}_b\boldsymbol{\Lambda}_c = \mathbf{I}_R$. Contrary to the 2-dimensional situation, these are usually the only transformational indeterminacies in CP. When, for given residuals $\underline{\mathbf{E}}$, the CP solution $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is unique up to these indeterminacies, the solution is called *essentially unique*.

To avoid unnecessarily complicated notation, we assume (without loss of generality) the residual array $\underline{\mathbf{E}}$ to be all-zero. Hence, we consider the essential uniqueness of the full CP decomposition

$$\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}^{\mathrm{T}}, \quad k = 1, \ldots, K. \tag{1.3}$$

We introduce the following notation. Let $\circ$ denote the outer product, i.e. for vectors $\mathbf{x}$ and $\mathbf{y}$ we define $\mathbf{x} \circ \mathbf{y} = \mathbf{x}\mathbf{y}^{\mathrm{T}}$. For three vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, the product $\mathbf{x} \circ \mathbf{y} \circ \mathbf{z}$ is a three-way array with elements $x_i y_j z_k$.

We use $\otimes$ to denote the Kronecker product. The Khatri–Rao product, which is the column-wise Kronecker product, is denoted by $\odot$. It is defined as follows. Suppose matrices $\mathbf{X}$ and $\mathbf{Y}$ both have n columns. Then the product $\mathbf{X} \odot \mathbf{Y}$ also has n columns and its jth column is equal to $\mathbf{x}_j \otimes \mathbf{y}_j$, where $\mathbf{x}_j$ and $\mathbf{y}_j$ denote the jth columns of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Notice that $\mathbf{X} \odot \mathbf{Y}$ contains the columns $1, n + 2, 2n + 3, \ldots, (n - 1)n + n$ of the Kronecker product $\mathbf{X} \otimes \mathbf{Y}$.

For a matrix $\mathbf{X}$, let $\mathrm{Vec}(\mathbf{X})$ denote the column vector which is obtained by placing the columns of $\mathbf{X}$ below each other, such that the first column of $\mathbf{X}$ is on top. For a vector $\mathbf{x}$, lets $\mathrm{diag}(\mathbf{x})$ denote the diagonal matrix with its diagonal equal to $\mathbf{x}$. Notice that $\mathrm{diag}(\mathbf{x}^{\mathrm{T}}) = \mathrm{diag}(\mathbf{x})$.

Although we consider the real-valued CP decomposition, all presented results also hold in the complex-valued case. To translate the proofs to the complex-valued case, the ordinary transpose $^{\mathrm{T}}$ should be changed into the Hermitian or conjugated transpose $^{\mathrm{H}}$ (except in those cases where the transpose is due to the formulation of the CP decomposition, such as in (1.3)).

2. Conditions for essential uniqueness

The first sufficient conditions for essential uniqueness are due to Jennrich (in Harshman [2]) and Harshman [3]. For a discussion, see Ten Berge and Sidiropoulos [14]. The most general sufficient condition for essential uniqueness is due to Kruskal [5]. Kruskal's condition involves a variant of the concept of matrix rank that he introduced, which has been named k-rank after him. The k-rank of a matrix is defined as follows.

Definition 2.1. The k-rank of a matrix is the largest value of x such that every subset of x columns of the matrix is linearly independent.

We denote the k-rank and rank of a matrix $\mathbf{A}$ by $k_{\mathbf{A}}$ and $r_{\mathbf{A}}$, respectively. There holds $k_{\mathbf{A}} \leqslant r_{\mathbf{A}}$. The matrix $\mathbf{A}$ has an all-zero column if and only if $k_{\mathbf{A}} = 0$ and $\mathbf{A}$ contains proportional columns if $k_{\mathbf{A}} = 1$.

Kruskal [5] showed that a CP solution $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is essentially unique if

$$2R + 2 \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}. \tag{2.1}$$

Notice that (2.1) cannot hold when $R = 1$. However, in that case uniqueness has been proven by Harshman [3]. Ten Berge and Sidiropoulos [14] have shown that Kruskal's sufficient condition (2.1) is also necessary for $R = 2$ and $R = 3$, but not for $R > 3$. The same authors conjectured that Kruskal's condition might be necessary and sufficient for $R > 3$, provided that k-ranks and ranks of the component matrices coincide, i.e. $k_{\mathbf{A}} = r_{\mathbf{A}}$, $k_{\mathbf{B}} = r_{\mathbf{B}}$ and $k_{\mathbf{C}} = r_{\mathbf{C}}$. However, Stegeman and Ten Berge [11] refuted this conjecture. Kruskal's condition was generalized to n-way arrays ($n > 3$) by Sidiropoulos and Bro [8].

Alternative sufficient conditions for essential uniqueness have been obtained for the case where one of the component matrices has full column rank, e.g. if $r_{\mathbf{C}} = R$. See Jiang and Sidiropoulos

[4] and De Lathauwer [6], who have independently proposed the same uniqueness condition (this was noticed in Stegeman, Ten Berge and De Lathauwer [12]). The latter author also provides a corresponding algorithm. Moreover, for random component matrices **A**, **B** and **C**, [6] derives a condition for "uniqueness with probability 1" in the form of a dimensionality constraint. A link between the deterministic approach of Jiang and Sidiropoulos [4] and the random setting of De Lathauwer [6] is provided in [12].

Besides sufficient conditions for essential uniqueness, also necessary conditions can be formulated. For instance, the CP solution is not essentially unique if one of the component matrices has an all-zero column. Indeed, suppose the rth column of **A** is all-zero, then the rth component $\underline{\mathbf{Y}}^{(r)}$ is all-zero and the rth columns of **B** and **C** can be arbitrary vectors. Also, the CP solution is not essentially unique if the k-rank of one of the component matrices equals 1. This can be seen as follows. Suppose $k_{\mathbf{A}} = 1$. Then **A** has two proportional columns, i.e. $\mathbf{a}^{(s)} = \lambda \mathbf{a}^{(t)}$ for some $s \neq t$. We have

$$\begin{aligned} \underline{\mathbf{Y}}^{(s)} + \underline{\mathbf{Y}}^{(t)} &= \lambda \mathbf{a}^{(t)} \circ \mathbf{b}^{(s)} \circ \mathbf{c}^{(s)} + \mathbf{a}^{(t)} \circ \mathbf{b}^{(t)} \circ \mathbf{c}^{(t)} \\ &= \mathbf{a}^{(t)} \circ [\lambda \mathbf{b}^{(s)} | \mathbf{b}^{(t)}][\mathbf{c}^{(s)} | \mathbf{c}^{(t)}]^{\mathrm{T}} \\ &= \mathbf{a}^{(t)} \circ [\lambda \mathbf{b}^{(s)} | \mathbf{b}^{(t)}]\mathbf{U}([\mathbf{c}^{(s)} | \mathbf{c}^{(t)}]\mathbf{U}^{-\mathrm{T}})^{\mathrm{T}}, \end{aligned} \tag{2.2}$$

for any nonsingular 2×2 matrix **U**. Eq. (2.2) describes mixtures of the sth and tth columns of **B** and **C** for which the fitted part of the CP model remains unchanged. If **U** is not the product of a diagonal and a permutation matrix, then (2.2) indicates that the CP solution is not essentially unique. Since the three-way array $\underline{\mathbf{X}}$ may be "viewed from different sides", the roles of **A**, **B** and **C** are exchangeable in the sequel. Hence, it follows that

$$k_{\mathbf{A}} \geqslant 2, \quad k_{\mathbf{B}} \geqslant 2, \quad k_{\mathbf{C}} \geqslant 2, \tag{2.3}$$

is a necessary condition for essential uniqueness of (**A**, **B**, **C**).

Another necessary condition for essential uniqueness is due to Liu and Sidiropoulos [7]. Let $\mathbf{X}_k$ be as in (1.3), i.e. $\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}^{\mathrm{T}}$. Let **X** be the matrix having $\mathrm{Vec}(\mathbf{X}_k^{\mathrm{T}})$ as its kth column, $k = 1, \ldots, K$. Then **X** can be written as

$$\mathbf{X} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^{\mathrm{T}}. \tag{2.4}$$

Suppose that $(\mathbf{A} \odot \mathbf{B})$ is not of full column rank. Then there exists a nonzero vector **n** such that $(\mathbf{A} \odot \mathbf{B})\mathbf{n} = \mathbf{0}$. Adding **n** to any column of $\mathbf{C}^{\mathrm{T}}$ preserves (2.4), but produces a different solution for **C**. Moreover, we have $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^{\mathrm{T}} = (\mathbf{A} \odot \mathbf{B})(\mathbf{C} + \mathbf{z}\mathbf{n}^{\mathrm{T}})^{\mathrm{T}}$ for any vector **z**, and we can choose **z** such that one column of $(\mathbf{C} + \mathbf{z}\mathbf{n}^{\mathrm{T}})$ vanishes. It follows that $\underline{\mathbf{X}}$ has a full CP decomposition with $R - 1$ components if $(\mathbf{A} \odot \mathbf{B})$ does not have full column rank.

Hence, full column rank of $(\mathbf{A} \odot \mathbf{B})$ is necessary for essential uniqueness. By exchanging the roles of **A**, **B** and **C**, we obtain that

$$(\mathbf{A} \odot \mathbf{B}) \quad \text{and} \quad (\mathbf{C} \odot \mathbf{A}) \quad \text{and} \quad (\mathbf{B} \odot \mathbf{C}) \quad \text{have full column rank}, \tag{2.5}$$

is a necessary condition for essential uniqueness of (**A**, **B**, **C**). Note that $(\mathbf{A} \odot \mathbf{B}) = \mathbf{\Pi}(\mathbf{B} \odot \mathbf{A})$, where $\mathbf{\Pi}$ is a row-permutation matrix. Hence, in each of the three Khatri–Rao products in (2.5) the two matrices may be swapped.

In the following, we denote the column space of a matrix **A** by span(**A**) and we denote its orthogonal complement null(**A**), i.e. $\mathrm{null}(\mathbf{A}) = \{\mathbf{x} : \mathbf{x}^{\mathrm{T}}\mathbf{A} = \mathbf{0}^{\mathrm{T}}\}$.

3. Proof of Kruskal's uniqueness condition

The proof of Kruskal's condition (2.1) in Kruskal [5] is long and rather inaccessible and does not seem very intuitive. Therefore, Kruskal's condition has been partially reproved by Sidiropoulos and Bro [8] and Jiang and Sidiropoulos [4]. However, apart from [5] no other complete proof of Kruskal's condition exists in the literature. In the remaining part of this paper, a complete, accessible and intuitive proof of Kruskal's condition will be provided. Kruskal's uniqueness condition is formalized in the following theorem.

Theorem 3.1. *Let* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *be a CP solution which decomposes the three-way array* $\underline{\mathbf{X}}$ *into* R *rank-*1 *arrays. Suppose Kruskal's condition* $2R + 2 \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}$ *holds and we have an alternative CP solution* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ *also decomposing* $\underline{\mathbf{X}}$ *into* R *rank-*1 *arrays. Then there holds* $\bar{\mathbf{A}} = \mathbf{A}\boldsymbol{\Pi}\boldsymbol{\Lambda}_a$, $\bar{\mathbf{B}} = \mathbf{B}\boldsymbol{\Pi}\boldsymbol{\Lambda}_b$ *and* $\bar{\mathbf{C}} = \mathbf{C}\boldsymbol{\Pi}\boldsymbol{\Lambda}_c$, *where* $\boldsymbol{\Pi}$ *is a unique permutation matrix and* $\boldsymbol{\Lambda}_a$, $\boldsymbol{\Lambda}_b$ *and* $\boldsymbol{\Lambda}_c$ *are unique diagonal matrices such that* $\boldsymbol{\Lambda}_a\boldsymbol{\Lambda}_b\boldsymbol{\Lambda}_c = \mathbf{I}_R$.

The cornerstone in both Kruskal's and our proof of Theorem 3.1 is Kruskal's Permutation Lemma, see Section 5 of Kruskal [5]. Below, we present this lemma. Its proof is postponed until Section 4 of this paper. Let $\omega(\mathbf{z})$ denote the number of nonzero elements of the vector $\mathbf{z}$.

Lemma 3.2 (Permutation Lemma). *Let* $\mathbf{C}$ *and* $\bar{\mathbf{C}}$ *be two* $K \times R$ *matrices and let* $k_{\mathbf{C}} \geqslant 2$. *Suppose the following condition holds: for any vector* $\mathbf{x}$ *such that* $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$, *we have* $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$. *Then there exists a unique permutation matrix* $\boldsymbol{\Pi}$ *and a unique nonsingular diagonal matrix* $\boldsymbol{\Lambda}$ *such that* $\bar{\mathbf{C}} = \mathbf{C}\boldsymbol{\Pi}\boldsymbol{\Lambda}$.

The condition of the Permutation Lemma states that if a vector $\mathbf{x}$ is orthogonal to $h \geqslant r_{\bar{\mathbf{C}}} - 1$ columns of $\bar{\mathbf{C}}$, then $\mathbf{x}$ is orthogonal to at least h columns of $\mathbf{C}$. Assuming the Permutation Lemma is true, this condition has to be equivalent to: if a vector $\mathbf{x}$ is orthogonal to h columns of $\bar{\mathbf{C}}$, then $\mathbf{x}$ is orthogonal to at least h columns of $\mathbf{C}$, $1 \leqslant h \leqslant R$. As an alternative to Kruskal's proof of the Permutation Lemma, Jiang and Sidiropoulos [4] prove the equivalence of these two conditions and show that the latter condition implies $\bar{\mathbf{C}} = \mathbf{C}\boldsymbol{\Pi}\boldsymbol{\Lambda}$. In Section 4, we reconsider Kruskal's original proof of the Permutation Lemma and explain the link with the approach of Jiang and Sidiropoulos [4].

In our proof of Theorem 3.1, we need the following result of Sidiropoulos and Bro [8] on the k-rank of a Khatri–Rao product. The proof below is a shorter version of the proof in [8] and is due to Ten Berge [13].

Lemma 3.3. *Consider matrices* $\mathbf{A}$ $(I \times R)$ *and* $\mathbf{B}$ $(J \times R)$:

(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* $k_{\mathbf{A}\odot\mathbf{B}} \geqslant \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, R)$.

Proof. First, we prove (i). If $k_{\mathbf{A}} = 0$, then $\mathbf{A}$ has an all-zero column. This implies that also $(\mathbf{A} \odot \mathbf{B})$ has an all-zero column and, hence, that $k_{\mathbf{A}\odot\mathbf{B}} = 0$. The same is true if $k_{\mathbf{B}} = 0$. This completes the proof of (i).

Next, we prove (ii). Suppose $k_{\mathbf{A}} \geqslant 1$ and $k_{\mathbf{B}} \geqslant 1$. Premultiplying a matrix by a nonsingular matrix affects neither the rank nor the k-rank. We have $(\mathbf{SA}) \odot (\mathbf{TB}) = (\mathbf{S} \otimes \mathbf{T})(\mathbf{A} \odot \mathbf{B})$, where

$(\mathbf{S} \otimes \mathbf{T})$ is nonsingular if both $\mathbf{S}$ and $\mathbf{T}$ are nonsingular. Hence, premultiplying $\mathbf{A}$ and $\mathbf{B}$ by nonsingular matrices also does not affect the rank and k-rank of $(\mathbf{A} \odot \mathbf{B})$. Since $\mathbf{A}$ has no all-zero columns, a linear combination of its rows exists such that all its elements are nonzero. Hence, since both $\mathbf{A}$ and $\mathbf{B}$ have no all-zero columns, we can find nonsingular matrices $\mathbf{S}$ and $\mathbf{T}$ such that both $\mathbf{SA}$ and $\mathbf{TB}$ have at least one row with all elements nonzero. Therefore, we may assume without loss of generality that both $\mathbf{A}$ and $\mathbf{B}$ have at least one row with all elements nonzero.

Since $(\mathbf{A} \odot \mathbf{B})$ has R columns, we have either $k_{\mathbf{A}\odot\mathbf{B}} = R$ or $k_{\mathbf{A}\odot\mathbf{B}} < R$. Suppose $k_{\mathbf{A}\odot\mathbf{B}} < R$. Let n be the smallest number of linearly dependent columns of $(\mathbf{A} \odot \mathbf{B})$, i.e. $k_{\mathbf{A}\odot\mathbf{B}} = n - 1$. We collect n linearly dependent columns of $(\mathbf{A} \odot \mathbf{B})$ in the matrix $(\mathbf{A}_n \odot \mathbf{B}_n)$, where $\mathbf{A}_n$ and $\mathbf{B}_n$ contain the corresponding columns of $\mathbf{A}$ and $\mathbf{B}$. Let $\mathbf{d}$ be a nonzero vector such that $(\mathbf{A}_n \odot \mathbf{B}_n)\mathbf{d} = \mathbf{0}$. Let $\mathbf{D} = \mathrm{diag}(\mathbf{d})$, which is nonsingular by the definition of n. Note that $\mathrm{Vec}(\mathbf{B}_n\mathbf{D}\mathbf{A}_n^{\mathrm{T}}) = (\mathbf{A}_n \odot \mathbf{B}_n)\mathbf{d} = \mathbf{0}$, which implies $\mathbf{B}_n\mathbf{D}\mathbf{A}_n^{\mathrm{T}} = \mathbf{O}$. Sylvester's inequality gives

$$0 = r_{\mathbf{B}_n\mathbf{D}\mathbf{A}_n^{\mathrm{T}}} \geqslant r_{\mathbf{A}_n} + r_{\mathbf{B}_n\mathbf{D}} - n = r_{\mathbf{A}_n} + r_{\mathbf{B}_n} - n, \tag{3.1}$$

where the last equality is due to the fact that $\mathbf{D}$ is nonsingular. Writing $n = k_{\mathbf{A}\odot\mathbf{B}} + 1$, Eq. (3.1) yields

$$k_{\mathbf{A}\odot\mathbf{B}} \geqslant r_{\mathbf{A}_n} + r_{\mathbf{B}_n} - 1. \tag{3.2}$$

Clearly, $r_{\mathbf{A}_n} \geqslant k_{\mathbf{A}_n}$. Because $\mathbf{B}$ has at least one row with all elements nonzero, it follows from $(\mathbf{A}_n \odot \mathbf{B}_n)\mathbf{d} = \mathbf{0}$ that the columns of $\mathbf{A}_n$ are linearly dependent. Hence, we also have $k_{\mathbf{A}_n} \geqslant k_{\mathbf{A}}$. Equivalently, $r_{\mathbf{B}_n} \geqslant k_{\mathbf{B}_n} \geqslant k_{\mathbf{B}}$. Hence, if $k_{\mathbf{A}\odot\mathbf{B}} < R$, we have $k_{\mathbf{A}\odot\mathbf{B}} \geqslant k_{\mathbf{A}} + k_{\mathbf{B}} - 1$. This completes the proof. □

The following lemma shows that Kruskal's condition implies the necessary uniqueness conditions (2.3) and (2.5). For (2.5), the proof is due to Liu and Sidiropoulos [7].

Lemma 3.4. *If Kruskal's condition* $2R + 2 \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}$ *holds, then*

(i) $k_{\mathbf{A}} \geqslant 2$, $k_{\mathbf{B}} \geqslant 2$ *and* $k_{\mathbf{C}} \geqslant 2$,
(ii) $r_{\mathbf{A}\odot\mathbf{B}} = r_{\mathbf{C}\odot\mathbf{A}} = r_{\mathbf{B}\odot\mathbf{C}} = R$.

Proof. We first prove (i). From Kruskal's condition and $k_{\mathbf{C}} \leqslant R$ it follows that $R \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} - 2$. If $k_{\mathbf{A}} \leqslant 1$, then this implies $k_{\mathbf{B}} \geqslant R + 1$, which is impossible. Hence, $k_{\mathbf{A}} \geqslant 2$ if Kruskal's condition holds. The complete proof of (i) is obtained by exchanging the roles of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$.

Next, we prove (ii). By the proof of (i) and statement (ii) of Lemma 3.3, we have $k_{\mathbf{A}\odot\mathbf{B}} \geqslant \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, R) = R$. Hence, $r_{\mathbf{A}\odot\mathbf{B}} \geqslant k_{\mathbf{A}\odot\mathbf{B}} \geqslant R$, which implies $r_{\mathbf{A}\odot\mathbf{B}} = R$. The complete proof of (ii) is obtained by exchanging the roles of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. □

We are now ready to prove Theorem 3.1. The first part of our proof, up to Eq. (3.11), is similar to Sidiropoulos and Bro [8]. Let the CP solutions $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ and $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ be as in Theorem 3.1 and assume Kruskal's condition $2R + 2 \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}$ holds. First, we use the Permutation Lemma to show that $\mathbf{C}$ and $\bar{\mathbf{C}}$ are identical up to a column permutation and rescaling. For this, we need to prove that for any $\mathbf{x}$ such that $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$ there holds $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$.

Let $\mathbf{x} = (x_1, \ldots, x_K)^{\mathrm{T}}$. We consider the linear combination $\sum_{k=1}^{K} x_k \mathbf{X}_k$ of the slices $\mathbf{X}_k$ in (1.3). For the Vec of the transpose of this linear combination, see (2.4), we have

$$(\mathbf{A} \odot \mathbf{B})\mathbf{C}^{\mathrm{T}}\mathbf{x} = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}. \tag{3.3}$$

By Lemma 3.4, the matrix $(\mathbf{A} \odot \mathbf{B})$ has full column rank. This implies that if $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) = 0$, then also $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) = 0$. Hence, $\mathrm{null}(\bar{\mathbf{C}}) \subseteq \mathrm{null}(\mathbf{C})$. The orthogonal decomposition theorem states that any $\mathbf{y} \in \mathrm{span}(\mathbf{C})$ can be written as the sum $\mathbf{y} = \mathbf{s} + \mathbf{s}^{\perp}$, where $\mathbf{s} \in \mathrm{span}(\bar{\mathbf{C}})$ and $\mathbf{s}^{\perp} \in \mathrm{null}(\bar{\mathbf{C}})$. But since $\mathrm{null}(\bar{\mathbf{C}}) \subseteq \mathrm{null}(\mathbf{C})$, there must hold $\mathbf{s}^{\perp} = \mathbf{0}$ and $\mathbf{y} \in \mathrm{span}(\bar{\mathbf{C}})$. Hence, we have $\mathrm{span}(\mathbf{C}) \subseteq \mathrm{span}(\bar{\mathbf{C}})$ and also $r_{\mathbf{C}} \leqslant r_{\bar{\mathbf{C}}}$. This implies that if $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$, then

$$\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1 \leqslant R - r_{\mathbf{C}} + 1 \leqslant R - k_{\mathbf{C}} + 1 \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} - (R+1), \tag{3.4}$$

where the last inequality follows from Kruskal's condition.

Next, consider the linear combination $\sum_{k=1}^{K} x_k \mathbf{X}_k$ again. By (1.3), it is equal to

$$\mathbf{A}\ \mathrm{diag}(\mathbf{C}^{\mathrm{T}}\mathbf{x})\mathbf{B}^{\mathrm{T}} = \bar{\mathbf{A}}\ \mathrm{diag}(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})\bar{\mathbf{B}}^{\mathrm{T}}. \tag{3.5}$$

We have

$$\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) = r_{\mathrm{diag}(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})} \geqslant r_{\bar{\mathbf{A}}\ \mathrm{diag}(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})\bar{\mathbf{B}}^{\mathrm{T}}} = r_{\mathbf{A}\ \mathrm{diag}(\mathbf{C}^{\mathrm{T}}\mathbf{x})\mathbf{B}^{\mathrm{T}}}. \tag{3.6}$$

Let $\gamma = \omega(\mathbf{C}^{\mathrm{T}}\mathbf{x})$ and let $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ consist of the columns of $\mathbf{A}$ and $\mathbf{B}$, respectively, corresponding to the nonzero elements of $\mathbf{C}^{\mathrm{T}}\mathbf{x}$. Then $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ both have γ columns. Let $\mathbf{t}$ be the $\gamma \times 1$ vector containing the nonzero elements of $\mathbf{C}^{\mathrm{T}}\mathbf{x}$ such that $\mathbf{A}\ \mathrm{diag}(\mathbf{C}^{\mathrm{T}}\mathbf{x})\mathbf{B}^{\mathrm{T}} = \widetilde{\mathbf{A}}\ \mathrm{diag}(\mathbf{t})\widetilde{\mathbf{B}}^{\mathrm{T}}$. Sylvester's inequality now yields

$$r_{\mathbf{A}\ \mathrm{diag}(\mathbf{C}^{\mathrm{T}}\mathbf{x})\mathbf{B}^{\mathrm{T}}} = r_{\widetilde{\mathbf{A}}\ \mathrm{diag}(\mathbf{t})\widetilde{\mathbf{B}}^{\mathrm{T}}} \geqslant r_{\widetilde{\mathbf{A}}} + r_{\widetilde{\mathbf{B}}\ \mathrm{diag}(\mathbf{t})} - \gamma = r_{\widetilde{\mathbf{A}}} + r_{\widetilde{\mathbf{B}}} - \gamma, \tag{3.7}$$

where the last equality is due to the fact that $\mathbf{t}$ contains no zero elements. From the definition of the k-rank, it follows that

$$r_{\widetilde{\mathbf{A}}} \geqslant \min(\gamma, k_{\mathbf{A}}), \quad r_{\widetilde{\mathbf{B}}} \geqslant \min(\gamma, k_{\mathbf{B}}). \tag{3.8}$$

From (3.6)–(3.8) we obtain

$$\omega(\mathbf{x}^{\mathrm{T}}\bar{\mathbf{C}}) \geqslant \min(\gamma, k_{\mathbf{A}}) + \min(\gamma, k_{\mathbf{B}}) - \gamma. \tag{3.9}$$

Combining (3.4) and (3.9) yields

$$\min(\gamma, k_{\mathbf{A}}) + \min(\gamma, k_{\mathbf{B}}) - \gamma \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} - (R+1). \tag{3.10}$$

Recall that we need to show that $\gamma = \omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$. From (3.10) it follows that we are done if $\gamma < \min(k_{\mathbf{A}}, k_{\mathbf{B}})$. We will prove this by contradiction. Suppose $\gamma > \max(k_{\mathbf{A}}, k_{\mathbf{B}})$. Then (3.10) gives $\gamma \geqslant R+1$, which is impossible. Suppose next that $k_{\mathbf{A}} \leqslant \gamma \leqslant k_{\mathbf{B}}$. Then (3.10) gives $k_{\mathbf{B}} \geqslant R+1$, which is also impossible. Since $\mathbf{A}$ and $\mathbf{B}$ can be exchanged in the latter case, this shows that $\gamma < \min(k_{\mathbf{A}}, k_{\mathbf{B}})$ must hold. Therefore, $\gamma = \omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ follows from (3.10) and the Permutation Lemma yields that a unique permutation matrix $\mathbf{\Pi}_c$ and a unique nonsingular diagonal matrix $\mathbf{\Lambda}_c$ exist such that $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}_c\mathbf{\Lambda}_c$. Notice that $k_{\mathbf{C}} \geqslant 2$ follows from Lemma 3.4.

The analysis above can be repeated for $\mathbf{A}$ and $\mathbf{B}$. Hence, it follows that

$$\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}_a\mathbf{\Lambda}_a, \quad \bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}_b\mathbf{\Lambda}_b, \quad \bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}_c\mathbf{\Lambda}_c, \tag{3.11}$$

where $\mathbf{\Pi}_a$, $\mathbf{\Pi}_b$ and $\mathbf{\Pi}_c$ are unique permutation matrices and $\mathbf{\Lambda}_a$, $\mathbf{\Lambda}_b$ and $\mathbf{\Lambda}_c$ are unique nonsingular diagonal matrices. It remains to show that $\mathbf{\Pi}_a = \mathbf{\Pi}_b = \mathbf{\Pi}_c$ and that $\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c = \mathbf{I}_R$. The following lemma states that we only need to show that $\mathbf{\Pi}_a = \mathbf{\Pi}_b$.

Lemma 3.5. *Let the CP solutions* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *and* $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}})$ *be as in Theorem 3.1 and assume Kruskal's condition* $2R + 2 \leqslant k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}$ *holds. If* (3.11) *holds and* $\mathbf{\Pi}_a = \mathbf{\Pi}_b$, *then* $\mathbf{\Pi}_a = \mathbf{\Pi}_b = \mathbf{\Pi}_c$ *and* $\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c = \mathbf{I}_R$.

Proof. Let $\mathbf{\Pi}_a = \mathbf{\Pi}_b = \mathbf{\Pi}$. We have $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^{\mathrm{T}} = (\mathbf{A\Pi\Lambda}_a \odot \mathbf{B\Pi\Lambda}_b)(\mathbf{C\Pi}_c\mathbf{\Lambda}_c)^{\mathrm{T}}$, which may be written as $(\mathbf{A} \odot \mathbf{B})(\mathbf{C\Pi}_c\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c\mathbf{\Pi}^{\mathrm{T}})^{\mathrm{T}}$. Since $(\mathbf{A} \odot \mathbf{B})$ has full column rank (see Lemma 3.4), it follows that $\mathbf{C} = \mathbf{C\Pi}_c\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c\mathbf{\Pi}^{\mathrm{T}}$. The matrix $\mathbf{\Pi}_c\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c\mathbf{\Pi}^{\mathrm{T}}$ is a rescaled permutation matrix. Since $\mathbf{C}$ has no all-zero columns or proportional columns ($k_{\mathbf{C}} \geqslant 2$, see Lemma 3.4), it follows that $\mathbf{\Pi}_c\mathbf{\Pi}^{\mathrm{T}} = \mathbf{I}_R$ and $\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c = \mathbf{I}_R$. This completes the proof. □

In the remaining part of this section, we will show that if Kruskal's condition holds, then (3.11) implies $\mathbf{\Pi}_a = \mathbf{\Pi}_b$. The proof of this is based on Kruskal [5, pp. 129–132].

Consider the CP decomposition in (1.3) and assume (3.11) holds. For vectors $\mathbf{v}$ and $\mathbf{w}$ we have

$$\mathbf{v}^{\mathrm{T}}\mathbf{X}_k\mathbf{w} = \mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{C}_k(\mathbf{w}^{\mathrm{T}}\mathbf{B})^{\mathrm{T}} = \mathbf{v}^{\mathrm{T}}\mathbf{A\Pi}_a\bar{\mathbf{C}}_k\mathbf{\Lambda}_a\mathbf{\Lambda}_b(\mathbf{w}^{\mathrm{T}}\mathbf{B\Pi}_b)^{\mathrm{T}}. \tag{3.12}$$

We combine (3.12) for $k = 1, \ldots, K$ into

$$\begin{pmatrix} \mathbf{v}^{\mathrm{T}}\mathbf{X}_1\mathbf{w} \\ \vdots \\ \mathbf{v}^{\mathrm{T}}\mathbf{X}_K\mathbf{w} \end{pmatrix} = \mathbf{C}\ \mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{A})\ \mathrm{diag}(\mathbf{w}^{\mathrm{T}}\mathbf{B}) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$= \bar{\mathbf{C}}\mathbf{\Lambda}_a\mathbf{\Lambda}_b\ \mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{A\Pi}_a)\ \mathrm{diag}(\mathbf{w}^{\mathrm{T}}\mathbf{B\Pi}_b) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{3.13}$$

Let $\mathbf{a}_r$ and $\mathbf{b}_r$ denote the rth columns of $\mathbf{A}$ and $\mathbf{B}$, respectively. We define

$$\mathbf{p} = \mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{A})\ \mathrm{diag}(\mathbf{w}^{\mathrm{T}}\mathbf{B}) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{v}^{\mathrm{T}}\mathbf{a}_1 \cdot \mathbf{w}^{\mathrm{T}}\mathbf{b}_1 \\ \vdots \\ \mathbf{v}^{\mathrm{T}}\mathbf{a}_R \cdot \mathbf{w}^{\mathrm{T}}\mathbf{b}_R \end{pmatrix}. \tag{3.14}$$

Let the index function $g(x)$ be given by $\mathbf{A\Pi}_a = [\mathbf{a}_{g(1)}|\mathbf{a}_{g(2)}|\cdots|\mathbf{a}_{g(R)}]$. Analogously, let $h(x)$ be given by $\mathbf{B\Pi}_b = [\mathbf{b}_{h(1)}|\mathbf{b}_{h(2)}|\cdots|\mathbf{b}_{h(R)}]$. We define

$$\mathbf{q} = \mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{A\Pi}_a)\ \mathrm{diag}(\mathbf{w}^{\mathrm{T}}\mathbf{B\Pi}_b) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{v}^{\mathrm{T}}\mathbf{a}_{g(1)} \cdot \mathbf{w}^{\mathrm{T}}\mathbf{b}_{h(1)} \\ \vdots \\ \mathbf{v}^{\mathrm{T}}\mathbf{a}_{g(R)} \cdot \mathbf{w}^{\mathrm{T}}\mathbf{b}_{h(R)} \end{pmatrix}. \tag{3.15}$$

Eq. (3.13) can now be written as $\mathbf{Cp} = \bar{\mathbf{C}}\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{q}$. Below, we will show that if Kruskal's condition holds and $\mathbf{\Pi}_a \neq \mathbf{\Pi}_b$, then we can choose $\mathbf{v}$ 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 $\mathbf{C}$ contains n linearly dependent columns, with $1 \leqslant n \leqslant k_{\mathbf{C}} - 1$, which contradicts the definition of $k_{\mathbf{C}}$. Hence, if Kruskal's condition and (3.11) hold, then $\mathbf{\Pi}_a = \mathbf{\Pi}_b$.

Suppose $\mathbf{\Pi}_a \neq \mathbf{\Pi}_b$. Then there exists an r such that $\mathbf{a}_r$ is the sth column of $\mathbf{A\Pi}_a$, column $\mathbf{b}_r$ is the tth column of $\mathbf{B\Pi}_b$ and $s \neq t$. Hence, there exists an r such that $r = g(s) = h(t)$ and $s \neq t$. Next, we create an index set $S \subset \{1, \ldots, R\}$ for which we find a vector $\mathbf{v}$ such that $\mathbf{v}^{\mathrm{T}}\mathbf{a}_j = 0$ if $j \in S$. Equivalently, we create an index set $T \subset \{1, \ldots, R\}$ for which we find a vector $\mathbf{w}$ such that $\mathbf{w}^{\mathrm{T}}\mathbf{b}_j = 0$ if $j \in T$. The sets S and T are created as follows:

- Put $g(t)$ in S and $h(s)$ in T.
- For all $x \notin \{s, t\}$, add $g(x)$ to S if $\mathrm{card}(S) < k_{\mathbf{A}} - 1$. Otherwise, add $h(x)$ to T.

Here, card(S) denotes the number of elements in the set S. We observe the following. In row x of the vector $\mathbf{q}$ in (3.15), either $g(x) \in S$ or $h(x) \in T$, $x = 1, \ldots, R$. The index $r = g(s) = h(t)$ is neither an element of S nor an element of T. Since $k_{\mathbf{A}} - 1 \leqslant R - 1$, the set S will contain exactly $k_{\mathbf{A}} - 1$ elements. The number of elements in T equals $R - \text{card}(S) = R - k_{\mathbf{A}} + 1$, which is less than or equal to $k_{\mathbf{B}} - 1$ (see the proof of Lemma 3.4).

We choose vectors $\mathbf{v}$ and $\mathbf{w}$ such that

$$\mathbf{v}^{\mathrm{T}}\mathbf{a}_j = 0 \text{ if } j \in S, \quad \text{and} \quad \mathbf{v}^{\mathrm{T}}\mathbf{a}_r \neq 0,$$
$$\mathbf{w}^{\mathrm{T}}\mathbf{b}_j = 0 \text{ if } j \in T, \quad \text{and} \quad \mathbf{w}^{\mathrm{T}}\mathbf{b}_r \neq 0.$$

It can be seen that such $\mathbf{v}$ and $\mathbf{w}$ always exist. The vector $\mathbf{v}$ has to be chosen from the orthogonal complement of span$\{\mathbf{a}_j, j \in S\}$, which is an $(I - k_{\mathbf{A}} + 1)$-dimensional space. If the column $\mathbf{a}_r$ is orthogonal to all possible vectors $\mathbf{v}$, it lies in span$\{\mathbf{a}_j, j \in S\}$. But then we would have $k_{\mathbf{A}}$ linearly dependent columns in $\mathbf{A}$, which is not possible. Hence, a vector $\mathbf{v}$ as above can always be found. Analogous reasoning shows that this is also true for $\mathbf{w}$.

For the sets S and T and the vectors $\mathbf{v}$ and $\mathbf{w}$ above, we have $\mathbf{q} = \mathbf{0}$ and the rth element of $\mathbf{p}$ nonzero. Let $S^c = \{1, \ldots, R\} \backslash S$ and $T^c = \{1, \ldots, R\} \backslash T$. The number of nonzero elements in $\mathbf{p}$ is bounded from above by

$$\text{card}(S^c \cap T^c) \leqslant \text{card}(S^c) = R - k_{\mathbf{A}} + 1 \leqslant k_{\mathbf{C}} - 1, \tag{3.16}$$

where the last inequality follows from Kruskal's condition and $k_{\mathbf{B}} \leqslant R$. Hence, $\mathbf{Cp} = \mathbf{0}$ implies that $\mathbf{C}$ contains n linearly dependent columns, with $1 \leqslant n \leqslant k_{\mathbf{C}} - 1$, which contradicts the definition of $k_{\mathbf{C}}$. This completes the proof of Theorem 3.1.

4. Proof of Kruskal's Permutation Lemma

Here, we prove Kruskal's Permutation Lemma, i.e. Lemma 3.2. The proof presented in Section 4.1 is similar to Kruskal's original proof in [5]. However, we add more clarification to it in order to show that the proof is natural and intuitive. In Section 4.2 we discuss the link between the proof in Section 4.1 and the alternative proof of the Permutation Lemma by Jiang and Sidiropoulos [4].

4.1. Kruskal's original proof revisited

Let $\mathbf{C}$ and $\bar{\mathbf{C}}$ be two $K \times R$ matrices and let $k_{\mathbf{C}} \geqslant 2$. Suppose the following condition holds: for any vector $\mathbf{x}$ such that $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$, we have $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$. We need to show that there exists a unique permutation matrix $\mathbf{\Pi}$ and a unique nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\bar{\mathbf{C}} = \mathbf{C\Pi\Lambda}$.

From the condition above, it follows that if $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) = 0$, then also $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) = 0$. Hence, null($\bar{\mathbf{C}}$) $\subseteq$ null($\mathbf{C}$). As before, this implies span($\mathbf{C}$) $\subseteq$ span($\bar{\mathbf{C}}$) and also $r_{\mathbf{C}} \leqslant r_{\bar{\mathbf{C}}}$. Since $r_{\mathbf{C}} \geqslant k_{\mathbf{C}} \geqslant 2$, the matrix $\bar{\mathbf{C}}$ must have at least two nonzero and nonproportional columns.

Recall that the condition above states that if a vector $\mathbf{x}$ is orthogonal to $h \geqslant r_{\bar{\mathbf{C}}} - 1$ columns of $\bar{\mathbf{C}}$, then $\mathbf{x}$ is orthogonal to at least h columns of $\mathbf{C}$. Any vector $\mathbf{x}$ is orthogonal to the all-zero columns of $\bar{\mathbf{C}}$. If $\mathbf{x}$ is orthogonal to a nonzero column $\bar{\mathbf{c}}_j$ of $\bar{\mathbf{C}}$, then it is also orthogonal to all nonzero columns which are scalar multiples of $\bar{\mathbf{c}}_j$. Therefore, it makes sense to partition the columns of $\bar{\mathbf{C}}$ into the following sets.

$$G_0 = \{\text{the all-zero columns of } \bar{\mathbf{C}}\},$$
$$G_j = \{\text{column } \bar{\mathbf{c}}_j \text{ of } \bar{\mathbf{C}} \text{ and all its nonzero scalar multiples in } \bar{\mathbf{C}}\}, \quad j = 1, \ldots, M.$$

Hence, if $\mathbf{x}$ is orthogonal to a nonzero column $\bar{\mathbf{c}}_j$ of $\bar{\mathbf{C}}$, then it is orthogonal to at least $\text{card}(G_0) + \text{card}(G_j)$ columns of $\bar{\mathbf{C}}$. If $\mathbf{x}$ is orthogonal to two linearly independent columns $\bar{\mathbf{c}}_i$ and $\bar{\mathbf{c}}_j$ of $\bar{\mathbf{C}}$, then it is orthogonal to all columns in G_0, G_i and G_j and to all columns of $\bar{\mathbf{C}}$ in $\text{span}\{\bar{\mathbf{c}}_i, \bar{\mathbf{c}}_j\}$. To be able to work with such sets of columns of $\bar{\mathbf{C}}$, we define the following.

Definition 4.1. A set H_k of columns of $\bar{\mathbf{C}}$ is called a *k-dimensional column set* if H_k contains exactly k linearly independent columns of $\bar{\mathbf{C}}$ and all columns of $\bar{\mathbf{C}}$ in the span of those k columns, $k = 1, \ldots, r_{\bar{\mathbf{C}}}$. The 0-*dimensional column set* H_0 is defined as $H_0 = G_0$.

Notice that the $r_{\bar{\mathbf{C}}}$-dimensional column set contains all columns of $\bar{\mathbf{C}}$. For $k \in \{1, \ldots, r_{\bar{\mathbf{C}}} - 1\}$, there are more possibilities of forming a k-dimensional column set. For example, the 1-dimensional column sets are given by the unions $G_0 \cup G_j$, $j = 1, \ldots, M$. The span of a k-dimensional column set H_k is denoted by $\text{span}(H_k)$ and its orthogonal complement by $\text{null}(H_k)$. It is important to note that if $\mathbf{x}$ is orthogonal to some nonzero columns of $\bar{\mathbf{C}}$, then those columns form a k-dimensional column set H_k and $\mathbf{x} \in \text{null}(H_k)$, $k \geqslant 1$. This fact, together with $\text{span}(\mathbf{C}) \subseteq \text{span}(\bar{\mathbf{C}})$, yields the following result. For ease of presentation, we postpone its proof until the end of this section.

Lemma 4.2. *Let* $\mathbf{C}$ *and* $\bar{\mathbf{C}}$ *be two* $K \times R$ *matrices. Suppose that* $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ *for any vector* $\mathbf{x}$ *such that* $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$. *For* $k \in \{0, \ldots, r_{\bar{\mathbf{C}}}\}$, *let* H_k *be a k-dimensional column set. Then the number of columns of* $\mathbf{C}$ *in* $\text{span}(H_k)$ *is larger than or equal to* $\text{card}(H_k)$.

The Permutation Lemma now follows from the result of Lemma 4.2 for $k = 0$ and $k = 1$. First, consider $k = 0$. Since $\mathbf{C}$ does not have any all-zero columns ($k_{\mathbf{C}} \geqslant 2$), the number of columns of $\mathbf{C}$ in $\text{span}(H_0)$ is zero. From Lemma 4.2 it then follows that also $\text{card}(H_0)$ must be zero. Hence, $\bar{\mathbf{C}}$ contains no all-zero columns and the set G_0 is empty.

Next, consider $k = 1$. There must hold that the number of columns of $\mathbf{C}$ in $\text{span}(H_1)$ is larger than or equal to $\text{card}(H_1)$, for any 1-dimensional column set H_1. The set G_0 is empty, which implies that the 1-dimensional column sets are given by G_j, $j = 1, \ldots, M$. Hence, the number of columns of $\mathbf{C}$ in $\text{span}(G_j)$ must be larger than or equal to $\text{card}(G_j)$, for $j = 1, \ldots, M$. Since $k_{\mathbf{C}} \geqslant 2$, the number of columns of $\mathbf{C}$ contained in a particular $\text{span}(G_j)$ cannot be larger than one. This implies $\text{card}(G_j) \leqslant 1$, for $j = 1, \ldots, M$. However, from the construction of the sets G_j, it follows that $\text{card}(G_j) \geqslant 1$. Hence, $\text{card}(G_j) = 1$, for $j = 1, \ldots, M$, and $M = R$. Since any column of $\mathbf{C}$ is contained in at most one $\text{span}(G_j)$, it follows that this can only be true if $\mathbf{C}$ and $\bar{\mathbf{C}}$ have the same columns up to permutation and scalar multiplication. That is, there exists a permutation matrix $\mathbf{\Pi}$ and a nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$. It can be seen that $\mathbf{\Pi}$ and $\mathbf{\Lambda}$ are indeed unique.

In order to prove the Permutation Lemma, we need the result of Lemma 4.2 only for $k = 0$ and $k = 1$. However, as will be seen in the proof of Lemma 4.2 below, the only way to obtain the result for $k = 0$ and $k = 1$ seems to be through induction, starting at $k = r_{\bar{\mathbf{C}}}$ and subsequently decreasing the value of k. This explains why Lemma 4.2 is formulated for all $k \in \{0, \ldots, r_{\bar{\mathbf{C}}}\}$.

Finally, we prove Lemma 4.2. We need the following result.

Lemma 4.3. *Let* H_k *be a k-dimensional column set and* $\mathbf{y}$ *a* $K \times 1$ *vector. If* $\mathbf{y} \notin \text{span}(H_k)$, *then there is at most one* $(k+1)$-*dimensional column set* H_{k+1} *with* $H_k \subset H_{k+1}$ *and* $\mathbf{y} \in \text{span}(H_{k+1})$.

Proof. Suppose to the contrary that there are two different $(k+1)$-dimensional column sets $H_{k+1}^{(1)}$ and $H_{k+1}^{(2)}$ which both contain H_k and such that $\mathbf{y} \in \text{span}(H_{k+1}^{(1)})$ and $\mathbf{y} \in \text{span}(H_{k+1}^{(2)})$. Since $\mathbf{y} \notin \text{span}(H_k)$, it follows that $\text{span}\{H_k, \mathbf{y}\} = \text{span}(H_{k+1}^{(1)}) = \text{span}(H_{k+1}^{(2)})$. Without loss of generality we may assume that $H_{k+1}^{(1)}$ contains some column $\bar{\mathbf{c}}$ of $\bar{\mathbf{C}}$ and $\bar{\mathbf{c}} \notin H_{k+1}^{(2)}$. Then $\text{span}\{H_{k+1}^{(2)}, \bar{\mathbf{c}}\} = \text{span}\{H_k, \mathbf{y}, \bar{\mathbf{c}}\}$ has dimension $k+2$. But this is not possible, since $\text{span}\{H_k, \mathbf{y}\} = \text{span}(H_{k+1}^{(1)})$ has dimension $k+1$ and contains $\bar{\mathbf{c}}$. This completes the proof. □

Proof of Lemma 4.2. First, we give the proof for $k = r_{\bar{\mathbf{C}}}$ and $k = r_{\bar{\mathbf{C}}} - 1$. Then we use induction on k (going from $k+1$ to k) to complete the proof.

Let $k = r_{\bar{\mathbf{C}}}$. There holds $\text{span}(H_k) = \text{span}(\bar{\mathbf{C}})$. Hence, the number of columns of $\bar{\mathbf{C}}$ in $\text{span}(H_k)$ is equal to R. Since $\text{span}(\mathbf{C}) \subseteq \text{span}(\bar{\mathbf{C}})$, the number of columns of $\mathbf{C}$ in $\text{span}(H_k)$ is also equal to R. This completes the proof for $k = r_{\bar{\mathbf{C}}}$.

Next, we consider the case $k = r_{\bar{\mathbf{C}}} - 1$. Let H_k be an $(r_{\bar{\mathbf{C}}} - 1)$-dimensional column set. Pick a vector $\mathbf{x} \in \text{null}(H_k)$, which does not lie in $\text{null}(\bar{\mathbf{C}})$. Then $\mathbf{x}$ can be chosen from a 1-dimensional subspace. Let $\bar{q} = \text{card}(H_k)$ and let q denote the number of columns of $\mathbf{C}$ in $\text{span}(H_k)$. We have $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - \bar{q}$; we claim that in fact $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) = R - \bar{q}$. Indeed, suppose $\mathbf{x}$ is orthogonal to column $\bar{\mathbf{c}}_j$ of $\bar{\mathbf{C}}$ and $\bar{\mathbf{c}}_j$ is not included in H_k. Then $\mathbf{x}$ would lie in $\text{null}\{H_k, \bar{\mathbf{c}}_j\} = \text{null}(\bar{\mathbf{C}})$, which is a contradiction. In the same way we can show that $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) = R - q$. As before, suppose $\mathbf{x}$ is orthogonal to column $\mathbf{c}_j$ of $\mathbf{C}$ and $\mathbf{c}_j$ does not lie in $\text{span}(H_k)$. Since $\mathbf{c}_j$ lies in $\text{span}(\mathbf{C}) \subseteq \text{span}(\bar{\mathbf{C}})$, the vector $\mathbf{x}$ lies in $\text{null}\{H_k, \ \mathbf{c}_j\} = \text{null}(\bar{\mathbf{C}})$. Again, we obtain a contradiction.

Since $\bar{q} \geqslant r_{\bar{\mathbf{C}}} - 1$, we have $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) = R - \bar{q} \leqslant R - r_{\bar{\mathbf{C}}} + 1$. Hence, the condition of Lemma 4.2 implies $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$, which yields $q \geqslant \bar{q}$. This completes the proof for $k = r_{\bar{\mathbf{C}}} - 1$.

The remaining part of the proof is by induction on k. Suppose the result of Lemma 4.2 holds for $k + 1 < r_{\bar{\mathbf{C}}}$. We will show that it holds for k as well. Let H_k be a k-dimensional column set. Let $H_{k+1}^{(i)}$, $i = 1, \ldots, N$, be all $(k+1)$-dimensional column sets such that $H_k \subset H_{k+1}^{(i)}$. Since there exists a group of exactly $r_{\bar{\mathbf{C}}} - k$ linearly independent columns of $\bar{\mathbf{C}}$ which are not included in H_k, it follows that there are at least $r_{\bar{\mathbf{C}}} - k \geqslant 2$ different $(k+1)$-dimensional column sets $H_{k+1}^{(i)}$ containing H_k. Thus, $N \geqslant 2$.

As above, Let $\bar{q} = \text{card}(H_k)$ and let q denote the number of columns of $\mathbf{C}$ in $\text{span}(H_k)$. Also, let $\bar{q}_i = \text{card}(H_{k+1}^{(i)})$ and let q_i denote the number of columns of $\mathbf{C}$ in $\text{span}(H_{k+1}^{(i)})$, for $i = 1, \ldots, N$. By the induction hypothesis, we know that $q_i \geqslant \bar{q}_i$ for all i. Next, we will show that $q \geqslant \bar{q}$.

According to Lemma 4.3, a column of $\mathbf{C}$ which does not lie in $\text{span}(H_k)$, is included in at most one $\text{span}(H_{k+1}^{(i)})$, $i \in \{1, \ldots, N\}$. This implies

$$q + \sum_{i=1}^{N} (q_i - q) \leqslant R, \tag{4.1}$$

where $q_i - q$ denotes the number of columns of $\mathbf{C}$ which are included in $\text{span}(H_{k+1}^{(i)})$ but not in $\text{span}(H_k)$. Analogously, Lemma 4.3 yields that a column of $\bar{\mathbf{C}}$ which is not in H_k, is an element of at most one $H_{k+1}^{(i)}$, $i \in \{1, \ldots, N\}$. Hence,

$$\bar{q} + \sum_{i=1}^{N} (\bar{q}_i - \bar{q}) \leqslant R, \tag{4.2}$$

where $\bar{q}_i - \bar{q}$ denotes the number of columns of $\bar{\mathbf{C}}$ which are included in $H_{k+1}^{(i)}$ but not in H_k. Notice that a column of $\bar{\mathbf{C}}$ which is included in the span of some k-dimensional column set is by definition an element of the k-dimensional column set itself.

Next, we show that we have equality in (4.2). Let $\bar{\mathbf{c}}_j$ be a column of $\bar{\mathbf{C}}$ which is not an element of H_k. Then we may create $H_{k+1}^{(i)}$ such that $\text{span}(H_{k+1}^{(i)}) = \text{span}\{H_k, \bar{\mathbf{c}}_j\}$. It can be seen that $H_k \subset H_{k+1}^{(i)}$ and $\bar{\mathbf{c}}_j \in H_{k+1}^{(i)}$. Therefore, any column of $\bar{\mathbf{C}}$ not included in H_k is included in some $H_{k+1}^{(i)}$, $i \in \{1, \ldots, N\}$. This implies that we may replace (4.2) by

$$\bar{q} + \sum_{i=1}^{N} (\bar{q}_i - \bar{q}) = R. \tag{4.3}$$

Now, the result follows from Eqs. (4.1) and (4.3) and the induction hypothesis. Indeed, we have

$$(N-1)q \geqslant \left(\sum_{i=1}^{N} q_i\right) - R \geqslant \left(\sum_{i=1}^{N} \bar{q}_i\right) - R = (N-1)\bar{q}, \tag{4.4}$$

where the first inequality follows from (4.1), the second inequality follows from the induction hypothesis and the equality follows from (4.3). Since $N \geqslant 2$, Eq. (4.4) yields $q \geqslant \bar{q}$. This completes the proof of Lemma 4.2. □

4.2. Connection with the proof of Jiang and Sidiropoulos

Here, we explain the link between the proof in Section 4.1 and the alternative proof of the Permutation Lemma by Jiang and Sidiropoulos [4]. As stated in Section 3, the latter authors prove the Permutation Lemma in two steps. First, they show that the condition in the Permutation Lemma is equivalent to: if a vector $\mathbf{x}$ is orthogonal to h columns of $\bar{\mathbf{C}}$, then it is orthogonal to at least h columns of $\mathbf{C}$. Second, they show that the latter condition implies $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$. These two steps are formalized in the following lemmas.

Lemma 4.4. *Let* $\mathbf{C}$ *and* $\bar{\mathbf{C}}$ *be two* $K \times R$ *matrices and let* $k_{\mathbf{C}} \geqslant 1$. *Then* $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ *for all* $\mathbf{x}$ *if and only if* $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ *for any* $\mathbf{x}$ *such that* $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$.

Lemma 4.5. *Let* $\mathbf{C}$ *and* $\bar{\mathbf{C}}$ *be two* $K \times R$ *matrices and let* $k_{\mathbf{C}} \geqslant 2$. *Suppose that* $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ *for all* $\mathbf{x}$. *Then there exists a unique permutation matrix* $\mathbf{\Pi}$ *and a unique nonsingular diagonal matrix* $\mathbf{\Lambda}$ *such that* $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$.

Using the approach in Section 4.1, we will prove Lemma 4.4 and Lemma 4.5. For Lemma 4.4, it suffices to prove the "if" part. This can be done by using Lemma 4.2 above. Indeed, suppose that $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ for any $\mathbf{x}$ such that $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) \leqslant R - r_{\bar{\mathbf{C}}} + 1$ and let $k_{\mathbf{C}} \geqslant 1$. We need to show that $\omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \leqslant \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$ holds for all $\mathbf{x}$. Since $\text{null}(\bar{\mathbf{C}}) \subseteq \text{null}(\mathbf{C})$, the case $\mathbf{x} \in \text{null}(\bar{\mathbf{C}})$ is trivial. Also the case $\omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) = R$ is trivial. Suppose $0 < \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x}) < R$. Then the columns of $\bar{\mathbf{C}}$ orthogonal to $\mathbf{x}$ form a k-dimensional column set H_k, for some $k \in \{1, \ldots, r_{\bar{\mathbf{C}}} - 1\}$, and $\mathbf{x} \in \text{null}(H_k)$. Notice that since $\bar{\mathbf{C}}$ does not contain all-zero columns (this follows from $k_{\mathbf{C}} \geqslant 1$ and the case $k = 0$ in Lemma 4.2, see above), the value of k must be at least one. As in the proof of Lemma 4.2, we have $\text{card}(H_k) = R - \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$. Similarly, the number of columns of $\mathbf{C}$ in $\text{span}(H_k)$ equals

$R - \omega(\mathbf{C}^{\mathrm{T}}\mathbf{x})$. Thus, according to Lemma 4.2, we have $R - \omega(\mathbf{C}^{\mathrm{T}}\mathbf{x}) \geqslant \mathrm{card}(H_k) = R - \omega(\bar{\mathbf{C}}^{\mathrm{T}}\mathbf{x})$, which proves Lemma 4.4.

As in Section 4.1, Lemma 4.5 can be obtained from Lemma 4.4 and Lemma 4.2. This explains the link between the proof of the Permutation Lemma in Section 4.1 and the one of Jiang and Sidiropoulos [4].

Acknowledgments

The authors would like to thank Jos ten Berge for sharing his short-cut proof of Lemma 3.3 and for commenting on an earlier version of this manuscript.

References

[1] J.D. Carroll, J.J. Chang, Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart–Young decomposition, Psychometrika 35 (1970) 283–319.
[2] R.A. Harshman, Foundations of the Parafac procedure: models and conditions for an "explanatory" multimodal factor analysis, UCLA Working Papers in Phonetics 16 (1970) 1–84.
[3] R.A. Harshman, Determination and proof of minimum uniqueness conditions for Parafac-1, UCLA Working Papers in Phonetics 22 (1972) 111–117.
[4] T. Jiang, 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) 2625–2636.
[5] J.B. Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics, Linear Algebra Appl. 18 (1977) 95–138.
[6] L. De Lathauwer, A link between the Canonical Decomposition in multilinear algebra and simultaneous matrix diagonalization, SIAM J. Matrix Anal. Appl. 28 (2006) 642–666.
[7] X. Liu, N.D. Sidiropoulos, Cramér–Rao lower bounds for low-rank decomposition of multidimensional arrays, IEEE Trans. Signal Process. 49 (2001) 2074–2086.
[8] N.D. Sidiropoulos, R. Bro, On the uniqueness of multilinear decomposition of N-way arrays, J. Chemometr. 14 (2000) 229–239.
[9] N.D. Sidiropoulos, Low-rank decomposition of multi-way arrays: a signal processing perspective, in: Proceedings of IEEE Sensor Array and Multichannel (SAM) Signal Processing Workshop, July 18–21, Sitges, Barcelona, Spain, 2004.
[10] A. Smilde, R. Bro, P. Geladi, Multi-way Analysis: Applications in the Chemical Sciences, Wiley, 2004.
[11] A. Stegeman, J.M.F. Ten Berge, Kruskal's condition for uniqueness in Candecomp/Parafac when ranks and k-ranks coincide, Computational Statistics and Data Analysis 50 (2006) 210–220.
[12] A. Stegeman, J.M.F. Ten Berge, L. De Lathauwer, Sufficient conditions for uniqueness in Candecomp/Parafac and Indscal with random component matrices, Psychometrika 71 (2006) 219–229.
[13] 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.
[14] J.M.F. Ten Berge, N.D. Sidiropoulos, On uniqueness in Candecomp/Parafac, Psychometrika 67 (2002) 399–409.
[15] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (1966) 279–311.