KU LEUVEN

ARENBERG DOCTORAL SCHOOL
KU Leuven - Kulak

Study of Canonical Polyadic Decomposition of Higher-Order Tensors

Ignat Domanov

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Engineering

September 2013

Study of Canonical Polyadic Decomposition of Higher-Order Tensors

Ignat DOMANOV

Supervisory Committee:
Prof. dr. Yves Willems, chair
Prof. dr. Lieven De Lathauwer, supervisor
Prof. dr. Marc Van Barel
Prof. dr. Joos Vandewalle
Prof. dr. Sabine Van Huffel
Prof. dr. Pierre Comon
(GIPSA-Lab, CNRS, France)
Prof. dr. Eugene E. Tyrtyshnikov
(Russian Academy of Sciences, Russia)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Engineering

September 2013

© KU Leuven - Kulak
E. Sabbelaan 53, B-8500 Kortrijk (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotocopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the publisher.

D/2013/7515/105
ISBN 978-94-6018-719-3

Preface

It is a great pleasure to express my thanks to the people and organizations who supported me during the PhD years.

First, I would like to express my gratitude to my promoter, Professor Lieven De Lathauwer, who has consistently encouraged me in this study and provided me precious suggestions and advice. Lieven, thanks for being my promoter, for guiding me towards Engineering and for a lot of time you made for me.

I would like to thank my committee members for all of their help and support in the completion of PhD degree. In addition, I am grateful to Professors Marc Van Barel, Joos Vandewalle, and Sabine Van Huffel, who introduced me to Numerical Linear Algebra, System Identification, and Biomedical Data Processing during my Pre-Doc and PhD years.

I would like to give my special thanks to my wife Olesya for her care and patience which made it possible for the thesis to be completed.

I would also like to mention the following funding organizations that contributed during my PhD years: (1) Research Council KU Leuven: GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT 1/08/23, DBOF/10/015; (2) F.W.O.: project G.0427.10N; (3) the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, "Dynamical systems, control and optimization", 2012–2017)

Abstract

In many applications signals or data vary with respect to several parameters (such as spatial coordinates, velocity, time, frequency, temperature, etc.) and are therefore naturally represented by higher-order arrays of numerical values, which are called higher-order tensors.

Matrices are tensors of order two. By definition, a matrix is rank-1 if its columns (or equivalently, rows) are proportional. A decomposition of matrix into a minimal number (known as the rank of the matrix) of rank-1 matrices is not unique. Indeed, there are many ways to write a matrix as a product of two matrices and every such factorization can be associated with a decomposition of a matrix into a sum of rank-1 matrices.

On the other hand, the factorization of a matrix into two or more structured matrices can be unique. The uniqueness of structured matrix factorizations (and hence, the uniqueness of decompositions of a matrix into a sum of structured rank-1 matrices) has found many applications in engineering. A famous example is the singular value decomposition (SVD) — a factorization into three matrices: a matrix with orthonormal columns, a diagonal matrix with positive values on the main diagonal, and a matrix with orthonormal rows. It is well known that the SVD is unique if and only if the diagonal entries of the second matrix are distinct.

In several cases the constraints on factor matrices that guarantee uniqueness of matrix decompositions are hard to justify from an application point of view.

By definition, a tensor is rank-1 if its columns (resp. rows, fibers, etc.) are proportional. A decomposition of a tensor in a minimal number of rank-1 tensors is called the Canonical Polyadic Decomposition (CPD), and the number of terms in the CPD is called the rank of tensor.

The CPD was introduced by F. Hitchcock in 1927 but it was not in use until 1970 when it was rediscovered as the Canonical Decomposition (Candecomp) in

Psychometrics and the Parallel Factor Model (Parafac) in Chemometrics. The interest in CPD from the side of data analysts and engineers was due to the remarkable property that CPD is unique under very mild conditions. Moreover, under simple constraints on the tensor dimensions the CPD is unique with probability one ("generic uniqueness").

Over the years CPD has found many applications in Signal Processing, Data Analysis, Machine Learning, Chemometrics, Psychometrics, etc.

For instance, in Chemometrics, one wishes to estimate the spectra and the concentrations of the chemicals present in a given solution. This can be done by decomposing a third-order tensor, which mathematically splits the mixture into a sum of rank-1 tensors that correspond to spectra of the chemicals and their concentrations.

In Chapter 1 we introduce CPD and recall some well-known applications in Chemometrics and Signal Processing.

In Chapter 2 we consider the case where at least one factor matrix is unique. The situation where one factor matrix is unique but the overall CPD is not, is typical for tensors with collinear loading vectors in some mode. We employ linear algebra tools to obtain new results on compound matrices. Then we obtain new conditions guaranteeing minimality of the number of terms and uniqueness of one factor matrix.

In Chapter 3 we study the overall uniqueness of the CPD. We also obtain new results on generic uniqueness of the structured CPD, i.e., the case where factor matrices analytically depend on some parameters. This includes the cases of partially symmetric tensors, tensors with Hankel, Toeplitz or Vandermonde factor matrices and cases where some entries of factor matrices are not allowed to change.

In Chapter 4 we present two algorithms for the computation of CPD. Both algorithms work under mild conditions on factor matrices (for instance, under the famous Kruskal condition) and reduce the problem to a generalized eigenvalue decomposition.

In the thesis we restrict ourselves to third-order tensors. New results for tensors of order higher than three can be easily derived from the third-order case by reshaping the higher-order tensor into a third-order tensor, with partial loss of structure.

Beknopte samenvatting

In veel toepassingen variëren signalen of data in functie van enkele parameters (bv. ruimtecoördinaten, snelheden, tijd, frequentie, temperatuur, enz.), en worden daarom op natuurlijke wijze voorgesteld door hogere-orde numerieke tabellen die tensoren genoemd worden.

Matrices zijn tensoren van tweede orde. Een matrix is per definitie van rang 1 indien de kolommen (of rijen) lineair afhankelijk zijn. Een ontbinding van een matrix in een minimaal aantal rang-1 matrices is niet uniek. Inderdaad, er bestaan veel manieren om een matrix te schrijven als een product van twee matrices en iedere dergelijke factorisatie kan geassocieerd worden met een matrixontbinding in een som van rang-1 matrices.

Anderzijds kan de factorisatie van een matrix in twee of meer gestructureerde matrices uniek zijn. De uniciteit van een gestructureerde-matrixfactorisatie (en dus de uniciteit van ontbindingen van een matrix in een som van gestructureerde rang-1 matrices) kent veel toepassingen in ingenieurswetenschappen. Een bekend voorbeeld is de Singuliere-Waardenontbinding (SWO) — een factorisatie in drie matrices: een matrix met orthonormale kolommen, een diagonale matrix met positieve waarden op de hoofddiagonaal en een matrix met orthonormale rijen. Het is algemeen bekend dat de SWO uniek is als en slechts als de waarden op de hoofddiagonaal van de tweede matrix onderling verschillend zijn.

De beperkingen op factormatrices die de uniciteit van matrixontbindingen garanderen zijn in verschillende toepassingen moeilijk te verantwoorden.

Een tensor is per definitie van rang 1 indien de kolommen (respectievelijk rijen, vezels, enz.) lineair afhankelijk zijn. Een ontbinding van een tensor in een minimaal aantal rang-1 tensoren wordt de Canonieke Polyadische Ontbinding (CPO) genoemd, en het aantal termen in de CPO wordt gedefinieerd als de rang van de tensor.

De CPO werd in 1927 geïntroduceerd door F. Hitchcock maar werd pas echt

gebruikt sinds 1970, toen ze werd herondekt als de Canonieke Ontbinding (Canonical Decomposition of Candecomp) in psychometrie en het Parallelle Factormodel (Parallel Factor Model of Parafac) in linguïstiek. De interesse in CPO vanwege gegevensanalisten en ingenieurs heeft te maken met de milde voorwaarden voor uniciteit. Onder eenvoudige beperkingen op de tensordimensies is de CPO zelfs uniek met waarschijnlijkheid één (generieke uniciteit).

Doorheen de jaren heeft CPO veel toepassingen gevonden in signaalverwerking, data-analyse, machineleren, chemometrie, psychometrie, enz.

In chemometrie bijvoorbeeld wil men de spectra en de concentraties schatten van chemicaliën die aanwezig zijn in een gegeven oplossing. Dit kan men bereiken met een ontbinding van een tensor van derde orde, die het mengsel op een wiskundige manier splitst in een som van tensoren van rang 1, dewelke de spectra en de concentraties van de chemicaliën weergeven.

In Hoofdstuk 1 introduceren we de CPO en bekijken we enkele bekende toepassingen uit de chemometrie en signaalverwerking.

In Hoofdstuk 2 bespreken we de situatie waarbij minstens één factormatrix uniek is. De situatie waarbij een enkele factormatrix uniek is maar de CPO zelf niet, is typisch voor tensoren met collineaire vectoren in een bepaalde mode. We gebruiken enkele instrumenten uit de lineaire algebra om nieuwe resultaten te bekomen voor zogenaamde compound matrices. We verkrijgen nieuwe voorwaarden die minimaliteit van het aantal termen en de uniciteit van een factormatrix garanderen.

In Hoofdstuk 3 bestuderen we de algemene uniciteit van CPO. We verkrijgen ook nieuwe resultaten voor de generieke uniciteit van de gestructureerde CPO, d.w.z. het geval waarbij factormatrices analytisch afhangen van enkele parameters. Dit omvat de gevallen van gedeeltelijk symmetrische tensoren, tensoren met Hankel, Toeplitz of Vandermonde factormatrices en gevallen waarin sommige waarden van factormatrices niet mogen veranderen.

In Hoofdstuk 4 presenteren we twee algoritmen voor de berekening van de CPO. Beide algoritmen werken onder milde voorwaarden op de factormatrices (bijvoorbeeld onder de bekende Kruskal voorwaarde) en reduceren het probleem tot een veralgemeende eigenwaardenontbinding.

In deze thesis beperken we ons tot derde-orde tensoren. Nieuwe resultaten voor hogere-orde tensoren kunnen eenvoudig afgeleid worden uit het derde-orde geval door de hogere-orde tensor te hervormen in een derde-orde tensor, met een gedeeltelijk verlies van structuur.

Contents

Abstract iii

Contents vii

List of Figures xi

List of Tables xiii

1 Overview 1

1.1 Introduction 1

1.1.1 Canonical Polyadic Decomposition 1

1.1.2 Rank of the tensor 3

1.1.3 Uniqueness of CPD 4

1.1.4 Algorithms for computation of CPD 5

1.2 Some applications of the CPD 5

1.2.1 Fluorescence spectroscopy [21] 5

1.2.2 Blind identification of CDMA systems [20] 6

1.2.3 Blind estimation of MISO FIR channels [8, 9], [6, 7] . . 7

1.2.4 SOBIUM family of problems and INDSCAL [5, 3] . . . 8

1.3 Contributions of the thesis and overview 9

1.4 Guide for the user 12
1.4.1 Unstructured CPD 12
1.4.2 Structured CPD 13
1.4.3 Tensors of order higher than 3 13
Bibliography 15

2 On the Uniqueness of the Canonical Polyadic Decomposition of third-order tensors — Part I: Basic Results and Uniqueness of One Factor Matrix 17

2.1 Introduction 18
2.1.1 Problem statement 18
2.1.2 Literature overview 19
2.1.3 Results and organization 24
2.2 Compound matrices and their properties 26
2.3 Basic implications 31
2.4 Sufficient conditions for the uniqueness of one factor matrix . . 38
2.4.1 Conditions based on (Um), (Cm), (Hm), and (Km) 38
2.4.2 Conditions based on (Wm) 41
2.5 Overall CPD uniqueness 43
2.5.1 At least one factor matrix has full column rank 43
2.5.2 No factor matrix is required to have full column rank . 44
2.6 Conclusion 45
2.7 Acknowledgments 45
Bibliography 45

3 On the Uniqueness of the Canonical Polyadic Decomposition of third-order tensors — Part II: Uniqueness of the overall decomposition 49

3.1 Introduction 49

3.1.1 Problem statement 49
3.1.2 Literature overview 52
3.1.3 Results and organization 57
3.2 Equality of PDs with common factor matrices 62
3.2.1 One factor matrix in common 62
3.2.2 Two factor matrices in common 66
3.3 Overall CPD uniqueness 68
3.4 Application to tensors with symmetric frontal slices and Indscal 72
3.5 Uniqueness beyond (W_m) 75
3.6 Generic uniqueness 79
3.6.1 Generic uniqueness of unconstrained CPD 79
3.6.2 Generic uniqueness of SFS-CPD 82
3.6.3 Examples 82
3.7 Conclusion 87
3.8 Acknowledgments 87
Bibliography 88

4 Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition 91

4.1 Introduction 91
4.1.1 Basic notations and terminology 91
4.1.2 Problem statement 93
4.1.3 Previous results on uniqueness and algebraic algorithms 94
4.1.4 New results and organization 96
4.2 Matrices formed by determinants and permanents of submatrices of a given matrix 100
4.2.1 Matrices whose entries are determinants 101
4.2.2 Matrices whose entries are permanents 103

4.2.3 Links between matrix $\mathcal{R}_m(\mathbf{C})$, matrix $\mathcal{B}(\mathbf{C})$ and symmetrizer . . . 105
4.3 Transformation of the CPD using polarized compound matrices 109
4.3.1 Mixed discriminants . . . 110
4.3.2 Polarized compound matrices . . . 111
4.3.3 Transformation of the tensor . . . 112
4.4 Overall results and algorithms . . . 113
4.5 Conclusion . . . 122
Bibliography . . . 122

5 Conclusion **127**
5.1 Conclusion . . . 127
5.2 Future work . . . 128
Bibliography . . . 129

Appendix to Chapter 4 **131**
A.1 Supplementary material related to Proposition 4.1.10 . . . 131
A.2 Supplementary material related to properties (P1)–(P4) . . . 134
A.3 Supplementary material related to Lemma 4.2.17 . . . 136
A.4 Supplementary material related to Lemma 4.4.4 . . . 137
A.5 Supplementary material related to Example 4.4.6 . . . 139

List of Figures

1.1 2-by-2-by-2 rank-1 tensor. 2

1.2 CPD of an I-by-J-by-K tensor. 2

1.3 CPD of diagonal $2 \times 2 \times 2$ tensor. 5

1.4 CPD of the third-order tensor containing fourth-order output cumulants for single-input single-output system. 8

1.5 Links between Chapters 2-4. 14

List of Tables

1.1 Decompositions of objects into a sum of simple terms 1

3.1 Upper bound $k(I)$ on R under which generic uniqueness of the CPD of an $I \times I \times I$ tensor is guaranteed by Theorem 3.1.19. . 57

3.2 Some cases where the rank and the uniqueness of the CPD of $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ may be easily obtained from Proposition 3.1.22 or its Corollary 3.1.23 (see Example 3.3.5). Matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are generated randomly. Simulations indicate that the dimensions of $\mathbf{A}$ and $\mathbf{B}$ cause the dimension of $\ker(\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))$ to be equal to 1. Thus, (Um) and (Wm) may be easily checked. . . . 72

3.3 Upper bounds on R under which generic uniqueness of the CPD of an $I \times I \times (2I - 1)$ tensor is guaranteed by Proposition 3.6.6. 84

3.4 Upper bounds on R under which generic uniqueness of the CPD (left and right value) and SFS-CPD (middle and right value) of an $I \times I \times K$ tensor is guaranteed by Proposition 3.1.31 (left), Proposition 3.6.8 (middle), and Kruskal's Theorems 3.1.8–3.1.10 (right). The values shown in bold correspond to the results that were not yet covered by Kruskal's Theorems 3.1.8–3.1.10 or Proposition 3.1.15 ($m = 2$). 86

Chapter 1

Overview

1.1 Introduction

1.1.1 Canonical Polyadic Decomposition

Decompositions of complex objects into a sum of simple terms are ubiquitous in engineering and natural sciences. Some well-known decompositions from Computer Science, Mathematics, Optics, and Chemistry are given in Table 1.1. Objects which vary with respect to time, space or frequency can be treated as

Table 1.1: Decompositions of objects into a sum of simple terms

Class of objects	**Simple terms**	**Example**
natural numbers	distinct powers of 2	$21 = 1 + 2^2 + 2^4$
vector of $\mathbb{R}^3$	basis vectors: $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$	$(2, 3, 0) = 2\mathbf{e}_1 + 3\mathbf{e}_2$
colors	red, green, blue	orange=red $+\frac{1}{2}$ green
chemical compounds	chemical elements	$H_2O = H_2 + O$

signals. *Blind Signal Separation* and *Data Analysis* consist of the splitting of signals and data into meaningful, interpretable components. The most well-known instance of this is probably the cocktail party problem: at a party we succeed in understanding the words spoken to us, while actually several people are speaking at the same time. Signal separation is of importance in application areas as audio processing, image processing, telecommunication

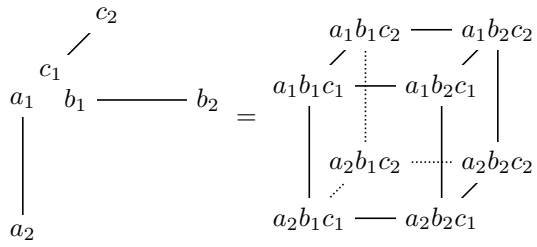


Figure 1.1: 2-by-2-by-2 rank-1 tensor.

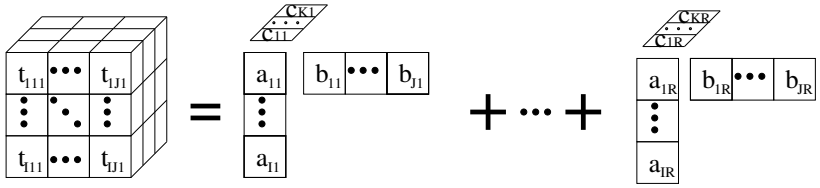


Figure 1.2: CPD of an I-by-J-by-K tensor.

(OFDM, CDMA, ...), array processing, biomedical problems (EEG, ECG, fMRI, ...), data mining, chemometrics, econometrics, bioinformatics, astrophysics, and so on.

Vectors, matrices, and higher-order tensors are numerical counterparts of signals. Higher-order tensors can be thought of as arrays of (real or complex) numbers of which the entries are indexed by more than two values.

The roles of simple terms in the decompositions of matrices and tensors are played by rank-1 matrices and rank-1 tensors, respectively.

Definition 1.1.1. *An I-by-J-by-K tensor $\mathcal{T} = (t_{ijk})$ is* rank-1 *if there exist three nonzero vectors $\mathbf{a} = [a_1 \ \ldots \ a_I]$, $\mathbf{b} = [b_1 \ \ldots \ b_J]$, and $\mathbf{c} = [c_1 \ \ldots \ c_K]$ such that $t_{ijk} = a_i b_j c_k$ for all values of the indices.*

Definition 1.1.2. Canonical Polyadic Decomposition (CPD) *of tensor $\mathcal{T}$ is the decomposition in a minimal number R of rank-1 terms. The number R is called the rank of $\mathcal{T}$ and denoted by $r_{\mathcal{T}}$.*

Figures 1.1 and 1.2 illustrate 2-by-2-by-2 rank-1 tensor and the CPD of an I-by-J-by-K tensor, respectively.

The CPD was introduced by F.L. Hitchcock in [14] and was later referred to as *Canonical Decomposition* (Candecomp) [2], *Parallel Factor Model* (Parafac) [10, 11], and *Topographic Components Model* [19].

We conclude this subsection by mentioning the principal advantages to work with tensors.

Firstly, in many applications the measured data can be naturally stored in a tensor whose entries have hidden dependencies. For instance, if the slices of a tensor correspond to the same object but measured at different time instances, then one can expect dependencies between the entries along the dimension which correspond to time. If the entries of a third-order tensor depend on coordinates on a XYZ grid, then one can expect dependencies between the entries that are close to each other. Multi-way data analysis (in particular, CPD) models data in all dimensions and hence, allows one to handle dependencies between entries better than two-way analysis.

Secondly, two-way methods for extracting information from one or two-way data typically work under strong assumptions. Consider the problem of blind channel identification: recover the transmitted message from the received message without any prior knowledge about the unknown channel. It is well-known that methods based on second-order statistics are blind to the changes induced by the channel. To recover the channel one should make additional assumptions (for instance, to consider only minimum phase channels). Such restrictions are not necessary when applying multi-way methods to tensors whose entries are higher-order statistics (HOS) (for instance, see Subsection 1.2.3). Through the computation of HOS, the matrix problem is transformed into a tensor problem.

Another example of tensorization concerns the analysis of ElectroEncephaloGraphy (EEG) data. In EEG the data sets are in the form of $channels \times time$. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) have been frequently used for the analysis of EEG data. However, PCA and ICA techniques do not take into account the frequency content of the signals in specific time periods across different channels. Wavelet based techniques allow one to transform two-way EEG data into a three-way tensor with modes $channels \times time \times frequency$ which can be further analyzed using CPD.

1.1.2 Rank of the tensor

It is a well-known fact that the rank of a matrix coincides with its column and row rank. As a consequence there exist many methods to compute the rank of a matrix. The most popular one is based on the Singular Value Decomposition

(SVD) of a matrix. SVD also solves analytically the problem of optimal low-rank approximation (Eckart-Young theorem).

The rank of a tensor is a much trickier concept. Column, row, and fiber ranks of a tensor are defined similarly to the matrix case. These numbers are not necessarily equal to each other nor to the rank of a tensor. Besides, the rank of a tensor depends on the field of scalars. For instance, if we assume that the entries of the 2-by-2-by-2 tensor $\mathcal{T}$ are independently drawn from the standard normal distribution (mean 0, variance 1), then the probability that $r_{\mathcal{T}} = 2$ is $\pi/4 \approx 0.79$ and the probability that $r_{\mathcal{T}} = 3$ is $1 - \pi/4 \approx 0.21$. In contrast, if the entries of $\mathcal{T}$ are independently drawn from the complex normal distribution, then the probability that $r_{\mathcal{T}} = 2$ is 1 [18, 1].

We also note that there is no direct analogue of the Eckart-Young theorem for tensors and that the computation of the rank of a tensor is NP-hard [12, 13].

1.1.3 Uniqueness of CPD

The decompositions given in Table 1.1 are unique: for a given object one can uniquely reconstruct the type and the amount of the simple terms that it consists of.

The matrix analogue of the CPD is known as the dyadic decomposition. In contrast to the decompositions given in Table 1.1, the dyadic decomposition of a matrix of rank greater than 1 is never unique. For instance, if $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and a, b, c, and d are numbers such that $ad - bc = 1$, then

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -2b \\ -c & 2a \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \begin{bmatrix} d & -2b \end{bmatrix} + \begin{bmatrix} b \\ d \end{bmatrix} \begin{bmatrix} -c & 2a \end{bmatrix}.$$

To guarantee the uniqueness of the dyadic decomposition one should impose additional constraints on rank-1 terms. For instance, the orthogonality constraints lead to the unique (singular value) decomposition of $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

Such constraints cannot always be justified from an application point of view. Contrary to the matrix case, CPD may be unique without imposing constraints. For instance, the CPD presented on Figure 1.3 is unique without imposing any additional constraints on its rank-1 terms.

Thanks to its uniqueness, CPD is currently becoming a standard tool for signal separation and data analysis, with concrete applications in telecommunication, array processing, machine learning, etc. [3, 16, 23].

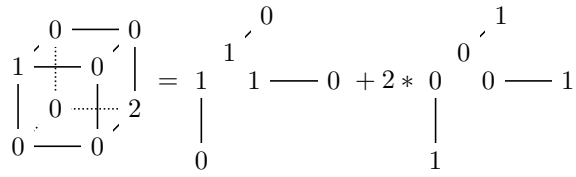


Figure 1.3: CPD of diagonal $2 \times 2 \times 2$ tensor.

1.1.4 Algorithms for computation of CPD

The known algorithms for the exact computation of CPD work under assumptions that the CPD is unique and that at least one of the tensor dimensions is greater than or equal to the rank of tensor. If these assumptions do not hold, then optimization-based algorithms may be used [22]. In this case the convergence of algorithms to the global minimum is not guaranteed. Moreover, some tensors can be approximated relatively well by tensors with strictly lower rank, which may imply that the global minimum does not exist.

1.2 Some applications of the CPD

In some applications (e.g., chemometrics and psychometrics), the tensor is obtained by just arranging data in an array and rank-1 terms in the CPD decomposition have empirical meaning. In other applications (signal processing) one should first construct the tensor based on the specific properties of data or system. We present examples which demonstrate both cases.

1.2.1 Fluorescence spectroscopy [21]

We have I samples of solutions, each contains the same chemicals at different concentrations. We need to identify the number of chemicals and to recover their concentrations.

In fluorescence spectroscopy, each sample is excited by light at J different wavelengths and then the emitted light is measured at K wavelengths. The data is stored in an $I \times J \times K$ array $\mathcal{T} = (t_{ijk})$, where t_{ijk} is the intensity of sample i at emission wavelength j and excitation wavelength k. We assume that the CPD of $\mathcal{T}$ is unique and can be computed. Then the number of chemicals coincides with the rank of $\mathcal{T}$, each rank-1 term corresponds to the unique

chemical, each horizontal slice of a particular rank-1 term is proportional to excitation-emission matrix of the chemical, and the coefficient of proportionality shows the concentration of the chemical in each sample.

1.2.2 Blind identification of CDMA systems [20]

Multiple access in telecommunication refers to the situation where several users transmit information over the same channel. In Time (resp. Frequency) Division Multiple Access (TDMA, resp. FDMA) each user uses his own time slots (resp. one or several frequency bands). In Code Division Multiple Access (CDMA) every user is allocated the entire spectrum all of the time. Simply speaking, in CDMA, each user codes the message with his own unique code. If the codes are known at the receiver, then there exists a simple procedure for extracting the coded messages from the mixture.

Let us give a linear algebra interpretation of the CDMA procedure in the case of R users and one receive antenna. With the r-th user we associate the $J \times 1$ vector $\mathbf{c}_r$ (code). If the the r-th user needs to send a message (just a vector $\mathbf{s}_r := [s_r[1] \ \ldots \ s_r[K]]^T$), then he transmits a vector $\mathbf{s}_r \otimes \mathbf{c}_r$. Suppose that the R users simultaneously send messages $\mathbf{s}_1, \ldots, \mathbf{s}_R$. Then the receiver gets the linear combination $\mathbf{x} = a_1\mathbf{s}_1 \otimes \mathbf{c}_1 + \cdots + a_R\mathbf{s}_R \otimes \mathbf{c}_R$, where the numbers a_r are responses of the receiver to the r-th user. If the vectors $\mathbf{c}_1, \ldots, \mathbf{c}_R$ are linearly independent and are known at the receiver, then the vectors $a_1\mathbf{s}_1, \ldots, a_R\mathbf{s}_R$ can be immediately obtained from the observed vector $\mathbf{x}$.

Let us consider the case with I receive antennas. Let a_{ir} be the response of antenna i to user r. Then antenna i receives the vector $\mathbf{x}_i = a_{i1}\mathbf{s}_1 \otimes \mathbf{c}_1 + \cdots + a_{iR}\mathbf{s}_R \otimes \mathbf{c}_R$. With each vector $\mathbf{x}_i$ we associate the $J \times K$ matrix $\mathbf{X}_i$ such that $\mathbf{x}_i$ is the vectorized version of $\mathbf{X}_i$. We form the $I \times J \times K$ tensor $\mathcal{X}$ with horizontal slices $\mathbf{X}_1, \ldots, \mathbf{X}_K$. It can be shown (see for instance, Subsection 4.1.1) that

$$\mathcal{X} = \mathcal{X}_1 + \cdots + \mathcal{X}_R, \tag{1.1}$$

where

$$\begin{aligned} &\mathcal{X}_r \ \text{ denotes the rank-1 tensor formed by the vectors} \\ &\qquad\qquad [a_{1r} \ \ldots \ a_{Ir}]^T, \quad \mathbf{c}_r, \quad \mathbf{s}_r. \end{aligned} \tag{1.2}$$

Suppose that the CPD of $\mathcal{X}$ is unique and is given by (1.1). Then the number of users R, their codes, their transmitted messages and the transfer function of overall system (the matrix $\mathbf{A} = (a_{ij})$) can be obtained from the rank-1 terms (1.2).

1.2.3 Blind estimation of MISO FIR channels [8, 9], [6, 7]

We consider a radio communication scenario with P transmit antennas and one receive antenna. We show that if the transmitted signals share the same carrier frequency and are not separated in time, then, in some cases, it is still possible to obtain information about the propagation channel.

We assume that the scenario can be modeled as a Multiple-Input Single-Output (MISO) system where the output signal $y[n]$ is the result of a linear combination of nonobserved input signals $s_p[n]$ filtered by unknown Finite Impulse Response FIR filters $\mathbf{h}_p = (h_{p,0}, \ldots, h_{p,L})^T$, $p \in [1, P]$ of the same length. In the presence of additive noise $v[n]$ the received signal $y[n]$ can be written as follows:

$$\begin{cases} y[n] = x_1[n] + \cdots + x_P[n] + v[n], \\ x_p[n] = (\mathbf{h}_p * s_p)[n] := \sum_{l=0}^{L} h_{p,l} s_p[n-l] \end{cases}.$$

We consider the problem of blind channel identification: identify channel parameters $\{h_{p,l}\}_{p,l=1,0}^{P,L}$ using only the system output $y[n]$.

The solution is based on the CPD of the $(2L+1)$-by-$(2L+1)$-by-$(2L+1)$ tensor $\mathcal{C} = (c_{ijk})$, where $c_{ijk} := \mathrm{cum}[y^*(n), y(n+i), y^*(n+j), y(n+k)]$ for $(i,j,k) \in [-L, L]^3$, "*" denotes the complex conjugation, and $cum(y_1, y_2, y_3, y_4)$ denotes the fourth-order cumulant of zero-mean signals y_1, y_2, y_3, y_4:

$$\begin{aligned} cum(y_1, y_2, y_3, y_4) :=& E(y_1^* y_2 y_3^* y_4) - E(y_1^* y_2) E(y_3^* y_4) - \\ & E(y_1^* y_3^*) E(y_2 y_4) - E(y_1^* y_4) E(y_2 y_3^*). \end{aligned}$$

Under certain statistical assumptions on the input signals and the noise, the Barlett-Brillinger-Rosenblatt formulae implies:

$$c_{ijk} = \gamma_{4,s} \sum_{p=1}^{P} \sum_{l=0}^{L} h_{p,l}^* h_{p,l+i} h_{p,l+j}^* h_{p,l+k}, \qquad (i,j,k) \in [-L, L]^3, \tag{1.3}$$

where $\gamma_{4,s}$ is the kurtosis of input signals. Equations (1.3) can be interpreted as a CPD of highly symmetric tensor $\mathcal{C}$. Moreover, the rank of $\mathcal{C}$ coincides with LP, the CPD of $\mathcal{C}$ is unique under mild conditions on the channel coefficients, and each rank-1 term corresponds to some channel and defines it up to a unitary scalar.

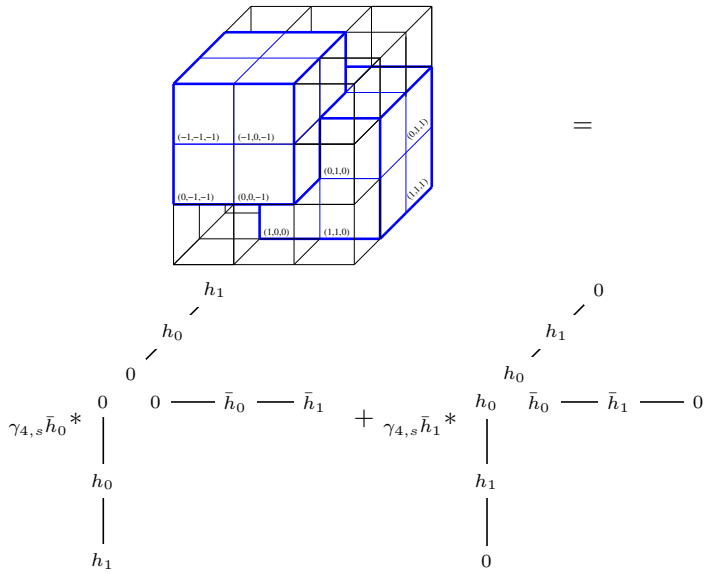


Figure 1.4: CPD of the third-order tensor containing fourth-order output cumulants for single-input single-output system.

In the case with one input and the channel $\mathbf{h} = [h_0\ h_1]$ (i.e $P = 1$ and $L = 1$), system (1.3) contains 15 equations

$$\begin{cases} c_{-1,-1,-1} = c_{0,1,0} = \overline{c}_{0,0,1} = \overline{c}_{1,0,0} & = \gamma_{4,s} h_0^2 \overline{h}_0 \overline{h}_1, \\ c_{-1,0,-1} = \overline{c}_{1,0,-1} & = \gamma_{4,s} h_0^2 \overline{h}_1^2, \\ c_{-1,-1,0} = c_{0,-1,-1} = c_{0,1,0} = c_{1,1,0} & = \gamma_{4,s} h_0 \overline{h}_0 h_1 \overline{h}_1, \\ c_{-1,0,0} = c_{0,0,-1} = \overline{c}_{0,-1,0} = \overline{c}_{1,1,1} & = \gamma_{4,s} h_0 \overline{h}_1^2 h_1, \\ c_{0,0,0} & = \gamma_{4,s} (h_0^2 \overline{h}_0^2 + h_1^2 \overline{h}_1^2). \end{cases} \tag{1.4}$$

The CPD of the 3-by-3-by-3 tensor $\mathcal{C}$ is shown on Figure 1.4, where the 15 "blue" entries correspond to entries that appear in system (1.4). It is easy to see that the channel $[h_0\ h_1]$ can be reconstructed from any rank-1 term up to a unitary scalar.

1.2.4 SOBIUM family of problems and INDSCAL [5, 3]

In this subsection we consider the CPD based approach for the Second-Order Blind Identification of Underdetermined Mixtures (SOBIUM).

We consider a system described by the following model

$$\mathbf{x} = \mathbf{A}\mathbf{s},$$

where $\mathbf{x}$ is the I-dimensional vector of observations, $\mathbf{s}$ is the R-dimensional unknown source vector and $\mathbf{A}$ is the I-by-R unknown mixing matrix. We assume that the sources are mutually uncorrelated but individually correlated in time, so the task corresponds to the Independent Component Analysis (ICA). We consider the case where the number of sources exceed the number of sensors ($R > I$). Our goal is to identify the matrix $\mathbf{A}$.

First we consider the case where all quantities involved take their values in the complex field. It is known that the spatial covariance matrices of the observations satisfy

$$\mathbf{C}_1 = E(\mathbf{x}_t \mathbf{x}_{t+\tau_1}^H) = \mathbf{A}\mathbf{D}_1\mathbf{A}^H,$$

$$\vdots$$

$$\mathbf{C}_K = E(\mathbf{x}_t \mathbf{x}_{t+\tau_K}^H) = \mathbf{A}\mathbf{D}_K\mathbf{A}^H,$$

in which $\mathbf{D}_k = E(\mathbf{s}_t \mathbf{s}_{t+\tau_k}^H)$ is the R-by-R diagonal matrix with the elements of the vector $\mathbf{d}^k$ on the main diagonal and $(\cdot)^H$ denotes the conjugate transpose. Let $\mathcal{T}$ be the I-by-J-by-K tensor with frontal slices $\mathbf{C}_1, \ldots, \mathbf{C}_K$ and let $\mathbf{a}_1, \ldots, \mathbf{a}_R$ and $\mathbf{d}_1, \ldots, \mathbf{d}_R$ denote the columns of the matrices $\mathbf{A}$ and $\begin{bmatrix}\mathbf{d}^1 & \ldots & \mathbf{d}^K\end{bmatrix}^T$, respectively. As in (1.1)–(1.2) we obtain $\mathcal{T} = \mathcal{T}_1 + \cdots + \mathcal{T}_R$, where

$$\mathcal{T}_r \quad \text{denotes the rank-1 tensor formed by the vectors} \quad \mathbf{a}_r, \quad \mathbf{d}_r, \quad \mathbf{a}_r^*. \tag{1.5}$$

If R is the rank of $\mathcal{T}$ and if the CPD of $\mathcal{T}$ is unique, then the columns of the matrix $\mathbf{A}$ can be found from (1.5).

If all quantities involved take their values in the real field, then the tensor $\mathcal{T}$ and the rank-1 tensors $\mathcal{T}_r$ have symmetric frontal slices. Such decompositions correspond to the individual differences scaling (INDSCAL) model, as introduced by Carroll and Chang [2].

1.3 Contributions of the thesis and overview

Results on rank and uniqueness of the CPD

The uniqueness of CPD is not yet completely understood, but for third-order tensors we find new sufficient conditions which guarantee the rank and the uniqueness of the CPD. Our results cover most cases of practical interest.

Suppose that the CPD of a tensor $\mathcal{X}$ is not unique but all CPDs of $\mathcal{X}$ share the same factor matrix in some mode. In this case we say that this factor matrix of $\mathcal{X}$ is unique. It is well known that if at least two rank-1 terms in the CPD of $\mathcal{X}$ have collinear vectors in some mode, then CPD is not unique. Nevertheless, the factor matrix in the same mode can still be unique, and hence, can be computed.

In Chapter 2 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 Chapter 3 for establishing overall CPD uniqueness in cases where none of the factor matrices has full column rank. A fortiori, the results of Chapter 2 also guarantee the rank of tensor.

In the context of the CDMA system considered in Subsection 1.2.2 the results obtained in Chapter 2 may be used in the following situations.

(1) If some Directions Of Arrival (DOA) are close, then some columns in the transfer function matrix $\mathbf{A}$ are close to collinear. We can still identify the number of users and the matrix $\mathbf{A}$.

(2) When the same message arrives from several DOAs, we can identify the number of paths and the transmitted messages.

New uniqueness results also imply the possibility to use shorter code sequences, and to have more users and fewer receive antennas in the system. Note also that in the two cases above the CPD is not unique (i.e. the full identification based on the CPD model is not possible).

In Chapter 3, based on results from Chapter 2, we establish overall CPD uniqueness in cases where none of the factor matrices has full column rank. We obtain uniqueness conditions involving Khatri-Rao products of compound matrices and Kruskal-type conditions. We consider both deterministic and generic uniqueness. We also discuss uniqueness of INDSCAL and other constrained polyadic decompositions.

Suppose that K is the largest dimension of the I-by-J-by-K tensor of rank R. Roughly speaking, in the literature easy-to-check deterministic sufficient condition for the uniqueness of the CPD are available for cases where at least

one of the following conditions holds

$$R \leq \frac{I+J+K-2}{2}, \tag{1.6}$$

$$R(R-1) \leq \frac{I(I-1)J(J-1)}{2}, \quad K \geq R. \tag{1.7}$$

If $K \geq R$, then (1.7) gives a much more relaxed bound on R than (1.6). On the other hand, the condition $K \geq R$ may be restrictive in some applications, and then only (1.6) is known.

The bound on R obtained in Chapter 3 is

$$C_R^m \leq C_I^m C_J^m, \quad m = R-K+2, \quad K \leq R, \tag{1.8}$$

where C_n^k denotes the binomial coefficient, $C_n^k = \frac{n!}{k!(n-k)!}$. Condition (1.8) generalizes (1.7) for $K \leq R$ and is more relaxed than (1.6).

We consider the case of (non canonical) polyadic decomposition with one or two known (up to permutation and column scaling) factor matrices. We find conditions for the identifiability of the remaining factor matrices.

We also obtain results on generic uniqueness of (structured) CPD: the uniqueness of a generic (random) $I \times J \times K$ tensor of rank R. The known results on generic uniqueness of unstructured CPDs include bounds (1.6)–(1.7) and some more relaxed bounds obtained recently in algebraic geometry.

We present new, easy-to-check sufficient conditions that guarantee generic uniqueness of the structured CPD. The conditions are formulated in terms of parameters that describe the structure of factor matrices. The structured CPDs may include: tensors with Hankel factor matrices, factor matrices with fixed entries, factor matrices satisfying constant modulus constraints etc. Note that for these kinds of structure the algebraic geometry methods most probably will fail.

Algorithms for explicit computation of the CPD

The most well-known condition guaranteeing uniqueness of the CPD was derived by J. Kruskal [17]. Kruskal's derivation is not constructive, i.e., it does not yield an algorithm. In many applications, there are reasonable bounds on the rank of the tensor at hand. For cases where the rank is known not to exceed one of the tensor dimensions, uniqueness has been proven under a condition that is an order of magnitude more relaxed than Kruskal's [15, 4]. For that case the decomposition may be computed by means of conventional linear algebra if

it is exact [4]. In Chapter 4 we propose an algorithmic approach that covers the intermediate cases. Roughly speaking, our algorithms compute the CPD of an $I \times J \times K$ tensor of rank R, where the bound on R is given in (1.8). The result is based on the reduction to the Generalized Eigenvalue Decomposition and generalizes the approach of [4]. The result may be used to initialize iterative algorithms.

Figure 1.5 gives an overview of the different chapters in this dissertation.

1.4 Guide for the user

In this subsection we provide a few guidelines to help the reader, interested in using the various uniqueness theorems in practical applications, find his way through the material.

1.4.1 Unstructured CPD

Chapters 2 and 3 contain both deterministic and generic uniqueness results. Deterministic conditions concern one particular PD $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. On the other hand, generic uniqueness means uniqueness that holds with probability one. Besides uniqueness of the overall decomposition, we also give results for the uniqueness of one factor matrix. The latter are useful in cases where all CPDs of a tensor have the same factor matrix but the overall CPD is not unique (see Example 2.4.11).

In the thesis we give answers to the following questions:

Q1: Does the rank of $\mathcal{T}$ coincide with R? (Or, is the PD $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ canonical?)

Q2: Is the third (resp. first or second) factor matrix of $\mathcal{T}$ unique?

Q3: Is the CPD of $\mathcal{T}$ unique?

Q4: Is the CPD of a generic $I \times J \times K$ tensor of rank R unique?

The answers to Q1–Q2 are summarized in scheme (2.12). For the convenience of the reader we present direct references to the relevant results for all questions. The results are ordered in priority of ease-of-use.

Q1–Q2: Corollaries 2.4.5 and 2.4.4, Proposition 2.4.3, Corollary 2.4.10, Proposition 2.4.9.

Q3: Corollaries 3.1.30, 3.1.29, 3.1.25, 3.1.28, 3.1.24, 3.1.27, 3.1.23, Propositions 3.1.26 and 3.1.22.

Q4: Proposition 3.1.31 (see also Theorems 3.1.16–3.1.19 for results obtained in algebraic geometry).

1.4.2 Structured CPD

In the structured PD the entries of factor matrices are subject to constraints. In the thesis we consider the case when the entries depend analytically on some parameters. The related subsections are:

Subsections 3.4 and 3.6.2: contain result on tensors with symmetric frontal slices. We study uniqueness of PDs of which the rank-1 terms have the same symmetry (also known as INDSCAL decomposition).

Subsection 3.6.3: many examples demonstrating how to use the results for different structured CPDs: generic uniqueness of the structured rank-1 perturbation of the "identity" $4 \times 4 \times 4$ tensor, uniqueness of PDs in which one or several factor matrices have structure (Toeplitz, Hankel, Vandermonde, etc.).

1.4.3 Tensors of order higher than 3

The thesis contains results on third-order tensors. Nevertheless, many results on the uniqueness of tensors of order higher than 3 can be obtained by the standard reduction to the third-order case: if the N-th-order tensor $\mathcal{T}$ has factor matrices $\mathbf{A}_1, \ldots, \mathbf{A}_N$, then the uniqueness of the CPD of $\mathcal{T}$ follows from the uniqueness of the CPD of the third-order tensor with factor matrices $\mathbf{A}_{i_1} \odot \cdots \odot \mathbf{A}_{i_k}$, $\mathbf{A}_{i_{k+1}} \odot \cdots \odot \mathbf{A}_{i_l}$, and $\mathbf{A}_{i_{l+1}} \odot \cdots \odot \mathbf{A}_{i_N}$, where $i_1, \ldots, i_N$ is an arbitrary permutation of $1, \ldots, N$, if one ignores the Khatri-Rao structure of the factors.

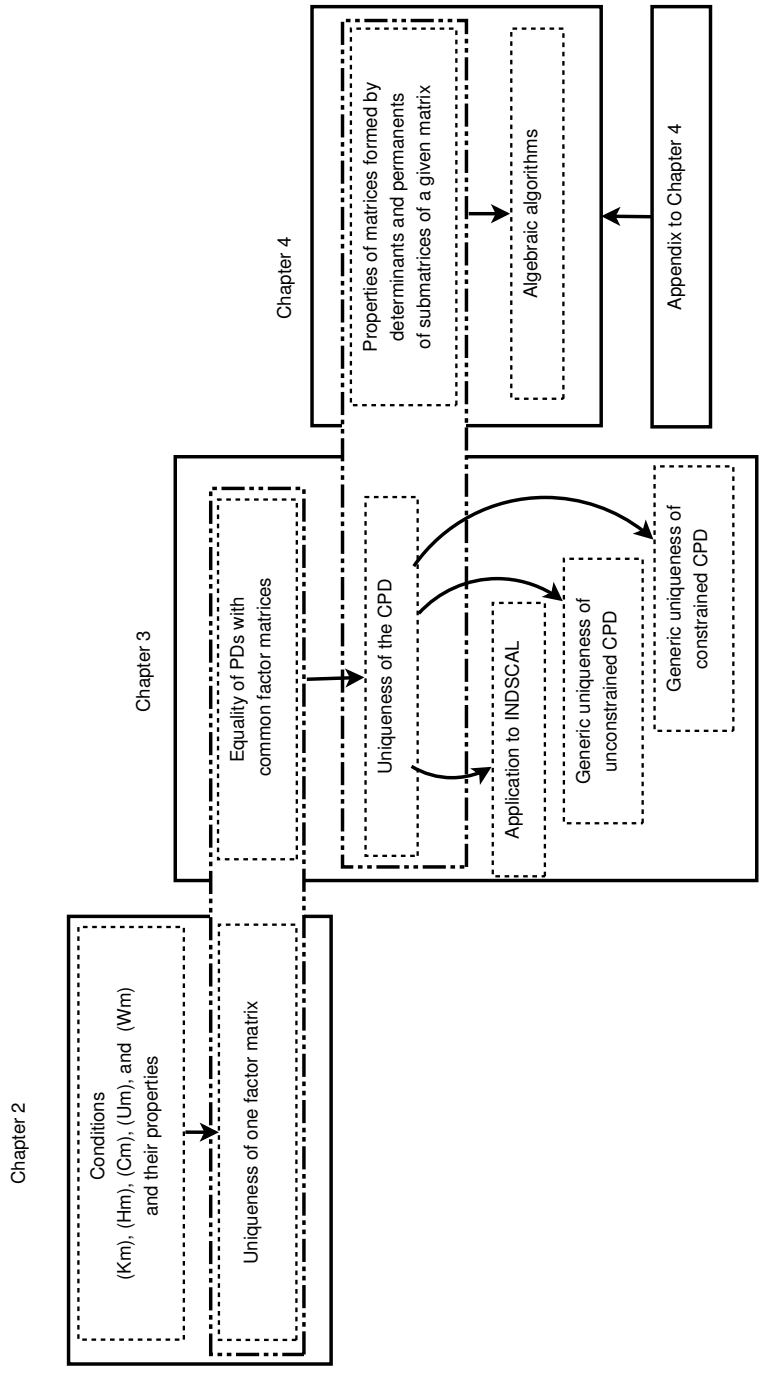


Figure 1.5: Links between Chapters 2-4.

Bibliography

[1] G. Bergqvist. Exact probabilities for typical ranks of $2 \times 2 \times 2$ and $3 \times 3 \times 2$ tensors. *Linear Algebra Appl.*, 438(2):663–667, 2013.

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

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

[4] L. De Lathauwer. A Link Between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization. *SIAM J. Matrix Anal. Appl.*, 28:642–666, August 2006.

[5] L. De Lathauwer and J. Castaing. Blind Identification of Underdetermined Mixtures by Simultaneous Matrix Diagonalization. *IEEE Trans. Signal Process.*, 56:1096–1105, 2008.

[6] I. Domanov and L. De Lathauwer. Enhanced line search for blind channel identification based on the parafac decomposition of cumulant tensors. *in Proc. of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010), Budapest, Hungary*, pages 1001–1002, 2010.

[7] I. Domanov and L. De Lathauwer. Blind channel identification of MISO systems based on the CP decomposition of cumulant tensors. *Proc. of the 2011 European Signal Processing Conference (EUSIPCO 2011),Barcelona, Spain*, pages 2215–2218, 2011.

[8] C. Fernandes, P. Comon, and G. Favier. Blind identification of MISO-FIR channels. *Signal Processing*, 90(2):490–503, 2010.

[9] C. Fernandes, G. Favier, and J. Mota. Blind channel identification algorithms based on the Parafac decomposition of cumulant tensors: The single and multiuser cases. *Signal Processing*, 88(6):1382–1401, 2008.

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

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

[12] Johan Håstad. Tensor rank is NP-complete. *J. Algorithms*, 11(4):644–654, December 1990.

[13] C. Hillar and L.-H. Lim. Most tensor problems are NP-hard. *arXiv:0911.1393v4*.

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

[15] T. Jiang and N. D. Sidiropoulos. Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints. *IEEE Trans. Signal Process.*, 52(9):2625–2636, September 2004.

[16] T. G. Kolda and B. W. Bader. Tensor Decompositions and Applications. *SIAM Review*, 51(3):455–500, September 2009.

[17] J. B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.*, 18(2):95–138, 1977.

[18] 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*, pages 7–18, 1989.

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

[20] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Trans. Signal Process.*, 48(3):810–823, 2000.

[21] A.K. Smilde, R. Bro, and P. Geladi. *Multi-way analysis with applications in the chemical sciences*. J. Wiley, 2004.

[22] L. Sorber, M. Van Barel, and L. De Lathauwer. Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-(L_r,L_r,1) terms and a new generalization. *SIAM J. Optim.*, 23(2):695–720, 2013.

[23] L. Xiangqian and N. D. Sidiropoulos. Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays. *IEEE Trans. Signal Process.*, 49(9):2074–2086, sep 2001.

Chapter 2

On the Uniqueness of the Canonical Polyadic Decomposition of third-order tensors — Part I: Basic Results and Uniqueness of One Factor Matrix

This chapter is based on
Domanov, I., De Lathauwer, L. On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors—Part I: Basic Results and Uniqueness of One Factor Matrix. SIAM Journal on Matrix Analysis and Applications, 34-3 (2013), pp. 855-875.

2.1 Introduction

2.1.1 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 2.1.1. *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 2.1.2. *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:*

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

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

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 (2.1) as $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$.

Definition 2.1.3. *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 2.1.4. *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, (2.1) can be identified either with

$$\mathbf{t} = \sum_{r=1}^{R} \mathbf{a}_r\otimes\mathbf{b}_r\otimes\mathbf{c}_r, \tag{2.2}$$

or with the matrix decomposition

$$\mathbf{T} = \sum_{r=1}^{R} (\mathbf{a}_r\otimes\mathbf{b}_r)\mathbf{c}_r^T. \tag{2.3}$$

Further, (2.3) can be rewritten as a factorization of $\mathbf{T}$,

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

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 (2.1)–(2.3) 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.

2.1.2 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 [17, 4, 7] and the references therein for background, applications, and algorithms. We also refer to [30] for a discussion of optimization-based algorithms.

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

$$r_{\mathbf{A}} = r_{\mathbf{B}} = r_{\mathbf{C}} = R. \tag{2.5}$$

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

Condition (2.5) may be relaxed as follows.

Theorem 2.1.6. *[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.*

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 2.1.7. *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 2.1.8. *[20] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2. \tag{2.6}$$

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 2.1.8 is given.

Before Kruskal arrived at Theorem 2.1.8 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 2.1.9. *[9, Theorem 2.3] (for original formulation see [20, Theorems 3a,b]) Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and suppose*

$$\begin{cases} k_{\mathbf{C}} \geq 1, \\ r_{\mathbf{C}} + \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq R + 2, \\ r_{\mathbf{C}} + k_{\mathbf{A}} + k_{\mathbf{B}} + \max(r_{\mathbf{A}} - k_{\mathbf{A}}, r_{\mathbf{B}} - k_{\mathbf{B}}) \geq 2R + 2. \end{cases} \tag{2.7}$$

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 2.1.10. *(see §2.4, for original formulation see [20, Theorem 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 2.1.8. These results will be discussed in Part II [8].

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 2.2.1 and Example 2.2.2).

Theorem 2.1.11. *[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*

$$(\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}))\begin{bmatrix} d_1d_2 & d_1d_3 & \dots & d_1d_R & d_2d_3 & \dots & d_{R-1}d_R \end{bmatrix}^T = \mathbf{0},$$

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 2.1.12. *[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 (2.6) is more restrictive than condition (C2).

Theorem 2.1.13. *[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 2.1.13 is due to A. Stegeman [31, Proposition 3.2, p. 215 and Lemma 4.4, p. 221]. Recently, another proof of Theorem 2.1.13 has been obtained in [10, Theorem 1, p. 3477].

Assuming $r_{\mathbf{C}} = R$, the conditions of Theorems 2.1.8 through 2.1.13 are related by

$$\begin{aligned} k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2 \;&\Rightarrow (\mathrm{K}_2) \;\Rightarrow (\mathrm{C}_2) \;\Rightarrow (\mathrm{U}_2) \\ &\Leftrightarrow r_{\mathcal{T}} = R \text{ and the CPD of } \mathcal{T} \text{ is unique.} \end{aligned} \tag{2.8}$$

Necessary conditions for uniqueness of the CPD. Results concerning rank and k-rank of Khatri-Rao product

It was shown in [35] that condition (2.6) is not only sufficient but also necessary for the uniqueness of the CPD if $R = 2$ or $R = 3$. Moreover, it was proved in [35] 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 (2.6) holds. Passing to higher values of R we have the following theorems.

Theorem 2.1.14. *[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 2.1.15. *[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 2.1.15 gives more restrictive uniqueness conditions than Theorem 2.1.14 and generalizes the implication (iii)⇒(ii) of Theorem 2.1.11 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 2.1.16. *[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 2.1.17. *[27, Lemma 1, p. 2382] If* $k_{\mathbf{A}} + k_{\mathbf{B}} \geq R + 1$, *then* (C1) *holds.*

Corollary 2.1.18. *[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 2.1.18 in [28] was based on Corollary 2.1.17. 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 [34].) All mentioned proofs are based on the Sylvester rank inequality.

2.1.3 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

$$\widehat{\mathbf{d}}^m := \begin{bmatrix} d_1 \cdots d_m & d_1 \cdots d_{m-1} d_{m+1} & \dots & d_{R-m+1} \cdots d_R \end{bmatrix}^T \in \mathbb{F}^{C_R^m}, \quad (2.9)$$

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:

$$\begin{cases} r_{\mathbf{A}} + k_{\mathbf{B}} \geq R + m, \\ k_{\mathbf{A}} \geq m \end{cases} \quad \text{or} \quad \begin{cases} r_{\mathbf{B}} + k_{\mathbf{A}} \geq R + m, \\ k_{\mathbf{B}} \geq m \end{cases}; \quad (\text{K}_m)$$

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

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

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

In §2.2 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 §2.3 we establish the following implications:

$$\begin{array}{lccccccccc}
 & (\mathrm{W}_m) & & (\mathrm{W}_{m-1}) & & \dots & & (\mathrm{W}_2) & & (\mathrm{W}_1) \\
\text{(Lemma 2.3.3)} & \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Uparrow \\
\text{(Lemma 2.3.7)} & (\mathrm{U}_m) & \Rightarrow & (\mathrm{U}_{m-1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{U}_2) & \Rightarrow & (\mathrm{U}_1) \\
\text{(Lemma 2.3.1)} & \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Updownarrow \\
\text{(Lemma 2.3.6)} & (\mathrm{C}_m) & \Rightarrow & (\mathrm{C}_{m-1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{C}_2) & \Rightarrow & (\mathrm{C}_1) \\
\text{(Lemma 2.3.8)} & \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Uparrow \\
\text{(Lemma 2.3.4)} & (\mathrm{K}_m) & \Rightarrow & (\mathrm{K}_{m-1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{K}_2) & \Rightarrow & (\mathrm{K}_1)
\end{array} \tag{2.10}$$

as well as (Lemma 2.3.12)

$$\text{if } \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1, \text{ then } (\mathrm{W}_m) \Rightarrow (\mathrm{W}_{m-1}) \Rightarrow \dots \Rightarrow (\mathrm{W}_2) \Rightarrow (\mathrm{W}_1). \tag{2.11}$$

We also show in Lemmas 2.3.5, 2.3.9–2.3.10 that (2.10) remains valid after replacing conditions $(\mathrm{C}_m),\dots,(\mathrm{C}_1)$ and equivalence $(\mathrm{C}_1) \Leftrightarrow (\mathrm{U}_1)$ by conditions $(\mathrm{H}_m),\dots,(\mathrm{H}_1)$ and implication $(\mathrm{H}_1) \Rightarrow (\mathrm{U}_1)$, respectively.

Equivalence of (C_1) and (U_1) is trivial, since the two conditions are the same. The implications $(\mathrm{K}_2) \Rightarrow (\mathrm{C}_2) \Rightarrow (\mathrm{U}_2)$ already appeared in (2.8). The implication $(\mathrm{K}_1) \Rightarrow (\mathrm{C}_1)$ was given in Lemma 2.1.16, and the implications $(\mathrm{K}_m) \Rightarrow (\mathrm{H}_m) \Rightarrow (\mathrm{U}_m)$ 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 §2.4 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 follows: if $k_{\mathbf{C}} \geq 1$ and $m = m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2$, then

$$(2.7) \;\Leftrightarrow\; \begin{array}{ccc} & (\mathrm{C}_m) & \\ \nearrow & & \searrow \\ (\mathrm{K}_m) & & (\mathrm{U}_m) \\ \searrow & & \nearrow \\ & (\mathrm{H}_m) & \end{array} \;\Rightarrow\; \begin{cases} \mathbf{A} \odot \mathbf{B} \text{ has full column rank,} \\ (\mathrm{W}_m), \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1 \end{cases}$$

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

$$\Rightarrow \quad r_{\mathcal{T}} = R \;\text{ and the third factor matrix of } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \;\text{ is unique.} \tag{2.12}$$

Thus, Theorems 2.1.9–2.1.10 are implied by the more general statement (2.12), which therefore provides new, more relaxed sufficient conditions for uniqueness of one factor matrix.

Further, compare (2.12) to (2.8). For the case $r_{\mathbf{C}} = R$, i.e., $m = 2$, uniqueness of the overall CPD has been established in Theorem 2.1.11. Actually, in this case overall CPD uniqueness follows easily from uniqueness of $\mathbf{C}$.

In §2.5 we simplify the proof of Theorem 2.1.11 using the material we have developed so far. In Part II [8] we will use (2.12) to generalize (2.8) to cases where possibly $r_{\mathbf{C}} < R$, i.e., $m > 2$.

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

$$S_n^k := \{(i_1, \ldots, i_k) : 1 \leq i_1 < \cdots < i_k \leq n\} \tag{2.13}$$

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

$$\sigma_{n,k} : \{1, 2, \ldots, C_n^k\} \to S_n^k = \{S_n^k(1), S_n^k(2), \ldots, S_n^k(C_n^k)\}. \tag{2.14}$$

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 (2.14).

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

$$\mathbf{d} = \begin{bmatrix} d_1 & d_2 & \ldots & d_R \end{bmatrix} \in \mathbb{F}^R,$$

$$\mathbf{d}_{S_R^m} = \begin{bmatrix} d_{(1,\ldots,m)} & \ldots & d_{(j_1,\ldots,j_m)} & \ldots & d_{(R-m+1,\ldots,R)} \end{bmatrix}^T \in \mathbb{F}^{C_R^m}. \tag{2.15}$$

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 (2.9).

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

Definition 2.2.1. *[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 2.2.2. *Let*

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

Then

$$\mathcal{C}_2(\mathbf{A}) = \begin{bmatrix} \mathcal{C}_2(\mathbf{A})_1 & \mathcal{C}_2(\mathbf{A})_2 & \mathcal{C}_2(\mathbf{A})_3 & \mathcal{C}_2(\mathbf{A})_4 & \mathcal{C}_2(\mathbf{A})_5 & \mathcal{C}_2(\mathbf{A})_6 \end{bmatrix}$$

$$= \begin{bmatrix} \mathcal{C}_2(\mathbf{A})_{(1,2)} & \mathcal{C}_2(\mathbf{A})_{(1,3)} & \mathcal{C}_2(\mathbf{A})_{(1,4)} & \mathcal{C}_2(\mathbf{A})_{(2,3)} & \mathcal{C}_2(\mathbf{A})_{(2,4)} & \mathcal{C}_2(\mathbf{A})_{(3,4)} \end{bmatrix}$$

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

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

Definition 2.2.1 immediately implies the following lemma.

Lemma 2.2.3. *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 2.2.4. *[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 2.2.5. *Let* $\mathbf{d} \in \mathbb{F}^R$*,* $k \leq R$*, and let* $\widehat{\mathbf{d}}^k$ *be defined by* (2.9)*. 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 2.2.6. *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:

$$\mathrm{vec}(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T) = \mathrm{vec}\left(\sum_{r=1}^{R} \mathbf{b}_r\mathbf{a}_r^T d_r\right) = \sum_{r=1}^{R} (\mathbf{a}_r \otimes \mathbf{b}_r)d_r = (\mathbf{A} \odot \mathbf{B})\mathbf{d}, \tag{2.16}$$

$$\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T = \mathbf{O} \ \Leftrightarrow \ \mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T = \mathbf{O} \ \Leftrightarrow \ (\mathbf{A} \odot \mathbf{B})\mathbf{d} = \mathbf{0}. \tag{2.17}$$

The following generalization of property (2.16) will be used throughout the paper.

Lemma 2.2.7. *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*

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

where $\widehat{\mathbf{d}}^k \in \mathbb{F}^{C_R^k}$ *is defined by* (2.9).

Proof. From Lemma 2.2.4 (1),(3) and Lemma 2.2.5 (3) it follows that

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

By (2.16),

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

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

Lemma 2.2.8. *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* $r_{\mathbf{A}\operatorname{Diag}(\mathbf{d})\mathbf{B}^T} \leq m-1$, *then* $\omega(\mathbf{d}) \leq m-1$;

(ii) (Um) *holds.*

Proof. From the definition of the m-th compound matrix and Lemma 2.2.7 it follows that

$$r_{\mathbf{A}\operatorname{Diag}(\mathbf{d})\mathbf{B}^T} = r_{\mathbf{B}\operatorname{Diag}(\mathbf{d})\mathbf{A}^T} \leq m-1 \quad \Leftrightarrow \quad \mathcal{C}_m(\mathbf{B}\operatorname{Diag}(\mathbf{d})\mathbf{A}^T) = \mathbf{O}$$

$$\Leftrightarrow \quad \operatorname{vec}(\mathcal{C}_m(\mathbf{B}\operatorname{Diag}(\mathbf{d})\mathbf{A}^T)) = \mathbf{0} \quad \Leftrightarrow \quad [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\widehat{\mathbf{d}}^m = \mathbf{0}.$$

Now the result follows from Lemma 2.2.5 (1). $\square$

The following three auxiliary lemmas will be used in §2.3.

Lemma 2.2.9. *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 (2.9). 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

$$(\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))\widehat{\mathbf{d}}^m =$$

$$\mathcal{C}_m(\begin{bmatrix}\mathbf{a}_{i_1} & \dots & \mathbf{a}_{i_m}\end{bmatrix}) \odot \mathcal{C}_m(\begin{bmatrix}\mathbf{b}_{i_1} & \dots & \mathbf{b}_{i_m}\end{bmatrix})d_{i_1} \cdots d_{i_m} = \mathbf{0},$$

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

Lemma 2.2.10. *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*

$$\mathcal{C}_m([\mathbf{A}\ \mathbf{x}]) = \Phi^{I,m}(\mathbf{x})\mathcal{C}_{m-1}(\mathbf{A}) \quad \textit{for all } \mathbf{A} \in \mathbb{F}^{I\times(m-1)} \textit{ and for all } \mathbf{x} \in \mathbb{F}^I. \tag{2.18}$$

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 (2.18), 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.

In more 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 2.2.10 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 a 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

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

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 (2.14). Then by (i),

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

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 2.2.11. *Let us illustrate Lemma 2.2.10 for $m = 2$ and $I = 4$. If $\mathbf{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} \end{bmatrix}^T$, then*

$$\mathcal{C}_2(\begin{bmatrix} \mathbf{A} & \mathbf{x} \end{bmatrix}) = \mathcal{C}_2(\begin{bmatrix} a_{11} & x_1 \\ a_{21} & x_2 \\ a_{31} & x_3 \\ a_{41} & x_4 \end{bmatrix}) = \begin{bmatrix} x_2a_{11} - x_1a_{21} \\ x_3a_{11} - x_1a_{31} \\ x_4a_{11} - x_1a_{41} \\ x_3a_{21} - x_2a_{31} \\ x_4a_{21} - x_2a_{41} \\ x_4a_{31} - x_3a_{41} \end{bmatrix}$$

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

2.3 Basic implications

In this section we derive the implications of (2.10) and (2.11). We first establish scheme (2.10) by means of Lemmas 2.3.1, 2.3.2, 2.3.3, 2.3.4, 2.3.6, 2.3.7, and 2.3.8.

Lemma 2.3.1. *Let $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, and $2 \le m \le \min(I, J)$. 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 fortiori not have another vector in its kernel with the structure specified in (Um). □

Lemma 2.3.2. *Let* $\mathbf{A} \in \mathbb{F}^{I \times R}$ *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 2.2.3.1, since $\widehat{\mathbf{d}}^1 = \mathbf{d}$. □

Lemma 2.3.3. *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 2.3.4. *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 2.3.5. *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 2.3.6. *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 (2.17), is equivalent to

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

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 2.2.10, 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 (2.17), is equivalent with

$$\left[\left(\Phi^{I,k}(\mathbf{a}_r)\mathcal{C}_{k-1}(\mathbf{A})\right) \odot \left(\Phi^{J,k}(\mathbf{b}_r)\mathcal{C}_{k-1}(\mathbf{B})\right)\right] \mathbf{d}_{S_R^{k-1}} = \mathbf{0}, \qquad r = 1, \ldots, R. \tag{2.19}$$

By (2.18),

$$\Phi^{I,k}(\mathbf{a}_r)\mathcal{C}_{k-1}([\mathbf{a}_{i_1} \ \ldots \ \mathbf{a}_{i_{k-1}}]) = \mathcal{C}_k([\mathbf{a}_{i_1} \ \ldots \ \mathbf{a}_{i_{k-1}} \ \mathbf{a}_r]) = \begin{cases} \mathbf{0}, & \text{if } r \in \{i_1, \ldots, i_{k-1}\}; \\ \pm\mathcal{C}_l(\mathbf{A})_{[i_1,i_2,\ldots,i_{k-1},r]}, & \text{if } r \notin \{i_1, \ldots, i_{k-1}\}, \end{cases} \tag{2.20}$$

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 (2.14), the columns of $\mathcal{C}_k(\mathbf{A})$ can be enumerated with S_R^k.) Similarly,

$$\Phi^{J,k}(\mathbf{b}_r)\mathcal{C}_{k-1}([\mathbf{b}_{i_1} \ \ldots \ \mathbf{b}_{i_{k-1}}]) = \begin{cases} \mathbf{0}, & \text{if } r \in \{i_1, \ldots, i_{k-1}\}; \\ \pm\mathcal{C}_k(\mathbf{B})_{[i_1,i_2,\ldots,i_{k-1},r]}, & \text{if } r \notin \{i_1, \ldots, i_{k-1}\}. \end{cases} \tag{2.21}$$

Now, equations (2.19)–(2.21) yield

$$\sum_{\substack{1 \le i_1 < \cdots < i_{k-1} \le R \\ i_1, \ldots, i_{k-1} \neq r}} d_{(i_1,\ldots,i_{k-1})} \left(\mathcal{C}_k(\mathbf{A})_{[i_1,i_2,\ldots,i_{k-1},r]} \otimes \mathcal{C}_k(\mathbf{B})_{[i_1,i_2,\ldots,i_{k-1},r]}\right) = \mathbf{0}, \qquad r = 1, \ldots, R. \tag{2.22}$$

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 2.3.7. *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* (Um) *implies conditions* (Um-1),. . . ,(U1).

Proof. It is sufficient to prove that (Uk) implies (Uk-1) for $k \in \{m, m-1, \ldots, 2\}$. Assume to the contrary that (Uk-1) does not hold. Then there exists a nonzero vector $\widehat{\mathbf{d}}^{k-1}$ such that $[\mathcal{C}_{k-1}(\mathbf{A}) \odot \mathcal{C}_{k-1}(\mathbf{B})]\, \widehat{\mathbf{d}}^{k-1} = \mathbf{0}$. Analogously to the proof of Lemma 2.3.6 we obtain that (2.22) holds with

$$d_{(i_1,\ldots,i_{k-1})} = d_{i_1} \cdots d_{i_{k-1}}, \qquad (i_1, \ldots, i_{k-1}) \in S_R^{k-1}. \tag{2.23}$$

Thus, multiplying the r-th equation from (2.22) by d_r, for $1 \leq r \leq R$, we obtain

$$\sum_{\substack{1 \leq i_1 < \cdots < i_{k-1} \leq R \\ i_1, \ldots, i_{k-1} \neq r}} d_{i_1} \cdots d_{i_{k-1}} d_r \left(\mathcal{C}_k(\mathbf{A})_{[i_1, i_2, \ldots, i_{k-1}, r]} \otimes \mathcal{C}_k(\mathbf{B})_{[i_1, i_2, \ldots, i_{k-1}, r]} \right) = \mathbf{0}. \tag{2.24}$$

Summation of (2.24) over r yields

$$k \left[\mathcal{C}_k(\mathbf{A}) \odot \mathcal{C}_k(\mathbf{B}) \right] \widehat{\mathbf{d}}^k = \mathbf{0}. \tag{2.25}$$

Since (Uk) holds, (2.25) 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, $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \leq k-2$. On the other hand, Lemma 2.2.9 implies that $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq k$, which is a contradiction. □

The following lemma completes scheme (2.10).

Lemma 2.3.8. *Let* $\mathbf{A} \in \mathbb{F}^{I \times R}$, $\mathbf{B} \in \mathbb{F}^{J \times R}$, *and* $m \leq \min(I, J)$. *Then condition* (Km) *implies condition* (Cm).

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 2.2.4 (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 (2.17),

$$\mathcal{C}_m(\mathbf{A}) \mathrm{Diag}(\mathbf{d}_{S_R^m}) \mathcal{C}_m(\mathbf{B})^T = \mathbf{O}_{C_I^m \times C_J^m}. \tag{2.26}$$

(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 $\mathbf{X} := \begin{bmatrix}\mathbf{b}_1 & \dots & \mathbf{b}_{k_{\mathbf{B}}}\end{bmatrix}^T$ 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

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

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

$$\mathcal{C}_m(\mathbf{B}^T\mathbf{Y}) = \begin{bmatrix} \mathbf{0}_{C_R^m - C_{R-k_{\mathbf{B}}+m}^m} \\ 1 \\ \boxplus_{(C_{R-k_{\mathbf{B}}+m}^m - 1)\times 1} \end{bmatrix}. \tag{2.27}$$

We now have

$$\begin{aligned}\mathbf{0}_{C_I^m} &= \mathbf{O}_{C_I^m \times C_J^m} \cdot \mathcal{C}_m(\mathbf{Y}) \\ &\overset{(2.26)}{=} \mathcal{C}_m(\mathbf{A})\mathrm{Diag}(\mathbf{d}_{S_R^m})\mathcal{C}_m(\mathbf{B}^T) \cdot \mathcal{C}_m(\mathbf{Y}) \\ &= \mathcal{C}_m(\mathbf{A})\mathrm{Diag}(\mathbf{d}_{S_R^m})\mathcal{C}_m(\mathbf{B}^T\mathbf{Y}) \\ &\overset{(2.27)}{=} \mathcal{C}_m(\mathbf{A}) \begin{bmatrix} \mathbf{0}_{C_R^m - C_{R-k_{\mathbf{B}}+m}^m} \\ d_{(k_{\mathbf{B}}-m+1,\dots,k_{\mathbf{B}})} \\ \boxplus_{(C_{R-k_{\mathbf{B}}+m}^m - 1)\times 1} \end{bmatrix}.\end{aligned} \tag{2.28}$$

Since the last $r_{\mathbf{A}}$ columns of $\mathbf{A}$ are linearly independent, Lemma 2.2.4 (5) implies that the $C_I^m \times C_{r_{\mathbf{A}}}^m$ matrix $\mathbf{M} = \mathcal{C}_m(\begin{bmatrix}\mathbf{a}_{R-r_{\mathbf{A}}+1} & \dots & \mathbf{a}_R\end{bmatrix})$ 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 (2.28) is zero. In particular, $d_{(k_{\mathbf{B}-m+1},\dots,k_{\mathbf{B}})} = 0$.

(iii) We show that $d_{(j_1,\dots,j_m)} = 0$ for any choice of $j_1, j_2, \dots, j_m$, $1 \leq j_1 < \dots < j_m \leq R$.

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

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

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 $k_{\mathbf{B}} \geq m$.

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

(ii) Suppose that $\delta \geq m$ and $\delta \geq k_{\mathbf{B}}$. Then $r_{\tilde{\mathbf{B}}} \geq k_{\mathbf{B}}$. 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 $r_{\tilde{\mathbf{A}}} \geq r_{\mathbf{A}} - r_{\tilde{\mathbf{A}}^c}$. Hence, $r_{\tilde{\mathbf{A}}} + r_{\tilde{\mathbf{B}}} - \delta \geq r_{\mathbf{A}} - r_{\tilde{\mathbf{A}}^c} + r_{\tilde{\mathbf{B}}} - \delta \geq r_{\mathbf{A}} - (R - \delta) + k_{\mathbf{B}} - \delta \geq m$. Thus, $H(\delta) \geq m$.

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

□

Lemma 2.3.10. *Let* $\mathbf{A} \in \mathbb{F}^{I\times R}$, $\mathbf{B} \in \mathbb{F}^{J\times R}$, *and* $m \leq \min(I, J)$. *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 2.2.7, $\mathcal{C}_m(\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T) = \mathcal{C}_m(\mathbf{B})\mathrm{Diag}(\widehat{\mathbf{d}}^m)\mathcal{C}_m(\mathbf{A})^T = \mathbf{O}$. Hence, $r_{\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T} \leq m - 1$. 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

$$\min(\delta, m) \le H(\delta) \le r_{\tilde{\mathbf{B}}} + r_{\tilde{\mathbf{A}}} - \delta = r_{\tilde{\mathbf{B}}\mathrm{Diag}(\tilde{\mathbf{d}})} + r_{\mathrm{Diag}(\tilde{\mathbf{d}})\tilde{\mathbf{A}}} - r_{\mathrm{Diag}(\tilde{\mathbf{d}})}$$

$$\le r_{\tilde{\mathbf{B}}\mathrm{Diag}(\tilde{\mathbf{d}})\tilde{\mathbf{A}}^T} = r_{\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{A}^T} \le m - 1.$$

Hence, $\delta \le m - 1$. From Lemma 2.2.5 1) it follows that $\widehat{\mathbf{d}}^m = \mathbf{0}$. □

The remaining part of this section concerns (2.11).

Lemma 2.3.7 and Lemma 2.2.9 can be summarized as follows:

$$(\mathrm{U_m}) \Rightarrow \begin{cases} (\mathrm{U_k}),\ k \le m; \\ \mathbf{A} \odot \mathbf{B} \ \text{ has full column rank } (\Leftrightarrow (\mathrm{U_1})); \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m. \end{cases}$$

The following example demonstrates that similar implications do not necessarily hold for $(\mathrm{W_m})$. Namely, in general, $(\mathrm{W_m})$ does not imply any of the following conditions:

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

$$\mathbf{A} \odot \mathbf{B} \ \text{ has full column rank,}$$

$$\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m - 1.$$

Example 2.3.11. *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 $(\mathrm{W_2})$ *holds but condition* $(\mathrm{W_1})$ *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 $\mathbf{d} \in \mathrm{range}(\mathbf{C}^T)$. *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*

$$(\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}))\widehat{\mathbf{d}}^2 = \mathbf{0} \Leftrightarrow x_2 = 0 \Rightarrow \omega(\mathbf{d}) \le 1 = 2 - 1 \Rightarrow (\mathrm{W_2}) \text{ holds},$$

$$(\mathbf{A} \odot \mathbf{B})\mathbf{e}_3^4 = \mathbf{0}, \quad \mathbf{e}_3^4 \in \mathrm{range}(\mathbf{C}^T), \quad \omega(\mathbf{e}_3^4) = 1 > 1 - 1 \Rightarrow (\mathrm{W_1}) \text{ does not hold},$$

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 (2.11).

Lemma 2.3.12. *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* (W$_m$) *implies conditions* (W$_{m-1}$),...,(W$_1$).

Proof. The proof is the same as the proof of Lemma 2.3.7, with the difference that instead of Lemma 2.2.9 we use the condition $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m - 1$. □

2.4 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 2.4.1. *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* $\mathbf{\Lambda}_{\mathbf{A}}$ *(resp.* $\mathbf{\Lambda}_{\mathbf{B}}$ *or* $\mathbf{\Lambda}_{\mathbf{C}}$*) such that* $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}}$ *(resp.* $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{B}}$ *or* $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}$).

2.4.1 Conditions based on (U$_m$), (C$_m$), (H$_m$), and (K$_m$)

First, we recall Kruskal's permutation lemma, which we will use in the proof of Proposition 2.4.3.

Lemma 2.4.2. *[20, 16, 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* $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$.

We start the derivation of (2.12) with the proof of the following proposition.

Proposition 2.4.3. *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* (U$_m$) *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 2.4.2 are satisfied. From Lemma 2.3.7 it follows that conditions (U$_{m\text{-}1}$), ..., (U$_2$), (U$_1$) hold. The fact that (U$_1$) holds, means that $\mathbf{A} \odot \mathbf{B}$ has full column rank. Hence,

$$r_{\mathbf{C}} = r_{\mathbf{C}^T} = r_{(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T} = r_{(\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T} \leq r_{\bar{\mathbf{C}}^T} = r_{\bar{\mathbf{C}}}. \tag{2.29}$$

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 2.4.2. Then $r_{\bar{\mathbf{A}}\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{x})\bar{\mathbf{B}}^T} \leq k - 1$ and, by (2.29),

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

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 2.2.4(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 2.2.5(1). Hence, by Lemma 2.2.7,

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

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 2.4.2, $\bar{R} = R$ and the matrices $\mathbf{C}$ and $\bar{\mathbf{C}}$ are the same up to permutation and column scaling. □

The implications (C$_m$) $\Rightarrow$ (U$_m$) and (H$_m$) $\Rightarrow$ (U$_m$) in scheme (2.10) lead to Corollary 2.4.4 and to Theorem 2.1.10, respectively. The implication (K$_m$) $\Rightarrow$ (C$_m$) in turn leads to Corollary 2.4.5. 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 2.4.4. *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*

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

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

Corollary 2.4.5. *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*

$$\begin{cases} r_{\mathbf{A}} + k_{\mathbf{B}} & \geq R + m, \\ k_{\mathbf{A}} & \geq m, \\ k_{\mathbf{C}} & \geq 1 \end{cases} \quad \text{or} \quad \begin{cases} r_{\mathbf{B}} + k_{\mathbf{A}} & \geq R + m, \\ k_{\mathbf{B}} & \geq m, \\ k_{\mathbf{C}} & \geq 1, \end{cases} \tag{2.30}$$

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

Remark 2.4.6. *Condition* (ii) *in Proposition 2.4.3 and Corollary 2.4.4 guarantees that the matrices* $\mathcal{C}_m(\mathbf{A})$ *and* $\mathcal{C}_m(\mathbf{B})$ *are defined. In Corollary 2.4.5,* (Km) *cannot hold if* $m = R - r_{\mathbf{C}} + 2 > \min(I, J)$.

Corollary 2.4.7. *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*

$$\begin{cases} k_{\mathbf{C}} \geq 1, \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m, \\ \max(r_{\mathbf{A}} + k_{\mathbf{B}}, r_{\mathbf{B}} + k_{\mathbf{A}}) \geq R + m, \end{cases} \tag{2.31}$$

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

Proof. It can easily be checked that (2.31) and (2.30) are equivalent. □

Remark 2.4.8. *It is easy to see that Corollaries 2.4.5 and 2.4.7 are equivalent to Theorem 2.1.9. Indeed, if* $m = R - r_{\mathbf{C}} + 2$, *then*

$$(2.7) \Leftrightarrow \begin{cases} k_{\mathbf{C}} \geq 1, \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m, \\ k_{\mathbf{A}} + k_{\mathbf{B}} + \max(r_{\mathbf{A}} - k_{\mathbf{A}}, r_{\mathbf{B}} - k_{\mathbf{B}}) \geq R + m \end{cases} \Leftrightarrow (2.31) \Leftrightarrow (2.30).$$

2.4.2 Conditions based on (Wm)

In this subsection we deal with condition (Wm). Similar to condition (Um) in Proposition 2.4.3, condition (Wm) in Proposition 2.4.9 will imply the uniqueness of one factor matrix. However, condition (Wm) is more relaxed than condition (Um). Like condition (Um), condition (Wm) may be hard to check. We give an example in which the uniqueness of one factor matrix can nevertheless be demonstrated using condition (Wm).

Proposition 2.4.9. *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* (Wm), ..., (W1) *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 2.4.3, with two points of difference. Namely, the fact that $\mathbf{A} \odot \mathbf{B}$ has full column rank does not follow from (W1) but is assumed in condition (iii). Second, $\omega(\mathbf{C}^T\mathbf{x}) \leq \omega(\bar{\mathbf{C}}^T\mathbf{x})$ follows from (Wk) instead of (Uk). □

From Lemmas 2.3.7 and 2.3.3 it follows that Proposition 2.4.9 is more relaxed than Proposition 2.4.3.

Combining Proposition 2.4.9 and Lemma 2.3.12 we obtain the following result, which completes the derivation of scheme (2.12).

Corollary 2.4.10. *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* (Wm) *holds.*

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

Example 2.4.11. *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 2.2.9 that condition ($\mathrm{U_m}$) *does not hold. We show that, on the other hand, condition* ($\mathrm{W_m}$) *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*

$$\begin{cases} d_2 = d_3 = 0, \\ d_1d_4d_5d_6d_7 \neq 0 \end{cases} \quad or \quad \begin{cases} d_6 = d_7 = 0, \\ d_1d_2d_3d_4d_5 \neq 0. \end{cases} \tag{2.32}$$

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

$$\mathbf{d} = \mathbf{C}^T\mathbf{x} = \begin{bmatrix} x_1 + x_4 & x_2 & x_3 & x_1 & x_2 & x_3 & x_4 \end{bmatrix}. \tag{2.33}$$

One can easily check that set (2.32) *does not have solutions of the form* (2.33). *Thus, condition* ($\mathrm{W_5}$) *holds.*

Corollary 2.1.18 implies that $\mathbf{A} \odot \mathbf{B}$ has full column rank. Thus, by Corollary 2.4.10, $r_{\mathcal{T}} = 7$ and the third factor matrix of $\mathcal{T}$ is unique.

Note that, since $k_{\mathbf{C}} = 1$, it follows from Theorem 2.1.14 that the CPD $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_7$ is not unique.

2.5 Overall CPD uniqueness

2.5.1 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 (2.8). In particular, Theorems 2.1.11, 2.1.12 and 2.1.13 present (U2), (C2) and (K2), respectively, as sufficient conditions for CPD uniqueness.

The implications (i)⇔(ii)⇔(iii) in Theorem 2.1.11 were proved in [16]. The core implication is (ii)⇒(iii). This implication follows almost immediately from Proposition 2.4.3, which establishes uniqueness of $\mathbf{C}$, as we show below. Implication (iii)⇒(i) follows from Theorem 2.1.15. Together with an explanation of the equivalence (i)⇔(ii) we obtain a short proof of Theorem 2.1.11.

Next, Theorem 2.1.12 follows immediately from Theorem 2.1.11. Theorem 2.1.13 in turn follows immediately from Theorem 2.1.12, cf. scheme (2.10).

Proof of Theorem 2.1.11. (i)⇔(ii): Follows from Lemma 2.2.8 for $m = 2$.

(ii)⇒(iii): By Proposition 2.4.3, $r_{\mathcal{T}} = R$ 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 (2.4), $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\bar{\mathbf{C}}^T = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\mathbf{\Lambda_C}\mathbf{\Pi}_{\mathbf{C}}^T\mathbf{C}^T$. Since the matrix $\mathbf{C}$ has full column rank, it follows that $\mathbf{A} \odot \mathbf{B} = (\bar{\mathbf{A}} \odot \bar{\mathbf{B}})\mathbf{\Lambda_C}\mathbf{\Pi}_{\mathbf{C}}^T$. 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 2.1.15. □

Proof of Theorem 2.1.12: Condition (C2) in Theorem 2.1.12 trivially implies condition (U2) in Theorem 2.1.11. Hence, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Proof of Theorem 2.1.13: By Lemma 2.3.8, condition (K2) in Theorem 2.1.13 implies condition (C2). Hence, by Theorem 2.1.12, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Remark 2.5.1. *The results obtained in [35] (see the beginning of subsection 2.1.2) 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 a low tensor rank, we now have 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.*

2.5.2 No factor matrix is required to have full column rank

Kruskal's original proof (also the simplified version in [32]) of Theorem 2.1.8 consists of three main steps. The first step is the proof of the permutation lemma (Lemma 2.4.2). The second and third steps concern the following two implications:

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

$$\Rightarrow \begin{cases} k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \\ r_{\mathcal{T}} = R, \\ \text{every factor matrix in } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is by itself unique} \end{cases} \tag{2.34}$$

$$\Rightarrow \text{the overall CPD } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is unique.}$$

That is, the proof goes via demonstrating that individual factor matrices are unique. Similarly, the proof of uniqueness result (ii)⇔(iii) in Theorem 2.1.11 for the case $r_{\mathbf{C}} = R$, goes, via Proposition 2.4.3, in two steps, which correspond to the proofs of the following two equivalences:

$$(\text{U2}) \Leftrightarrow \begin{cases} r_{\mathcal{T}} = R, \\ \text{the third factor matrix of } \mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \text{ is unique} \end{cases} \tag{2.35}$$

$$\Leftrightarrow \text{ the CPD of } \mathcal{T} \text{ is unique.}$$

Again the proof goes via demonstrating that one factor matrix is unique. Note that the second equivalence in (2.35) is almost immediate since $\mathbf{C}$ has full column rank. In contrast, the proof of the second implication in (2.34) is not trivial.

Scheme (2.12) generalizes the first implications in (2.34) and (2.35). 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.

2.6 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 (2.10) and (2.11). For the case $r_{\mathbf{C}} = R$ we have demonstrated the overall CPD uniqueness results in (2.8) using second compound matrices. Using (2.10) and (2.11) 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 the 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.

2.7 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).

Bibliography

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

[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, editors. *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(7-8):393–405, 2009.

[5] L. De Lathauwer. A Link Between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization. *SIAM J. Matrix Anal. Appl.*, 28:642–666, August 2006.

[6] L. De Lathauwer. 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(4):1451–1474, 2011.

[7] L. De Lathauwer. 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*, pages 558–563, 2011.

[8] I. Domanov and L. De Lathauwer. On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors— Part II: Overall Uniqueness. *SIAM J. Matrix Anal. Appl.*, 34(3):876–903, 2013.

[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.*, 33:111–129, 2012.

[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(7):3475–3481, September 2011.

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

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

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

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

[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(9):2625–2636, September 2004.

[17] T. G. Kolda and B. W. Bader. Tensor Decompositions and Applications. *SIAM Review*, 51(3):455–500, September 2009.

[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(2):95–138, 1977.

[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*, pages 7–18, 1989.

[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(6):311–320, 2010.

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

[25] J. A. Rhodes. A concise proof of Kruskal's theorem on tensor decomposition. *Linear Algebra Appl.*, 432(7):1818–1824, 2010.

[26] N. D. Sidiropoulos and R. Bro. On the uniqueness of multilinear decomposition of N-way arrays. *J. Chemometrics*, 14(3):229–239, 2000.

[27] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis. Parallel Factor Analysis in Sensor Array Processing. *IEEE Trans. Signal Process.*, 48:2377–2388, 2000.

[28] N. D. Sidiropoulos and L. Xiangqian. Identifiability results for blind beamforming in incoherent multipath with small delay spread. *IEEE Trans. Signal Process.*, 49(1):228–236, January 2001.

[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. *SIAM J. Optim.*, 23(2):695–720, 2013.

[31] A. Stegeman. On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode. *Linear Algebra Appl.*, 431(1-2):211–227, 2009.

[32] A. Stegeman and N. D. Sidiropoulos. On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition. *Linear Algebra Appl.*, 420(2-3):540–552, 2007.

[33] V. Strassen. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52–53(0):645–685, 1983.

[34] J. Ten Berge. The k-rank of a Khatri-Rao product. *Unpublished Note, Heijmans Institute of Psychological Research, University of Groningen, The Netherlands*, 2000.

[35] J. Ten Berge and N. D. Sidiropoulos. On uniqueness in Candecomp/Parafac. *Psychometrika*, 67:399–409, 2002.

[36] L. Xiangqian and N. D. Sidiropoulos. Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays. *IEEE Trans. Signal Process.*, 49(9):2074–2086, sep 2001.

Chapter 3

On the Uniqueness of the Canonical Polyadic Decomposition of third-order tensors — Part II: Uniqueness of the overall decomposition

This chapter is based on
Domanov, I., De Lathauwer, L. On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors — Part II: Uniqueness of the Overall Decomposition. SIAM Journal on Matrix Analysis and Applications, 34-3 (2013), pp. 876-903.

3.1 Introduction

3.1.1 Problem statement

Throughout the paper $\mathbb{F}$ denotes the field of real or complex numbers; $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote conjugate, transpose, and conjugate transpose, respectively; $r_{\mathbf{A}}$, range($\mathbf{A}$), and ker($\mathbf{A}$) denote the rank, the range, and the null space of a matrix $\mathbf{A}$, respectively; Diag($\mathbf{d}$) denotes a square diagonal matrix with the

elements of a vector $\mathbf{d}$ on the main diagonal; $\text{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ denotes the linear span of the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_k$; $\mathbf{e}_r^R$ denotes the r-th vector of the canonical basis of $\mathbb{F}^R$; 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. A third-order tensor $\mathcal{T} = (t_{ijk}) \in \mathbb{F}^{I\times J\times K}$ is *rank*-1 if there exist three nonzero vectors $\mathbf{a} \in \mathbb{F}^I$, $\mathbf{b} \in \mathbb{F}^J$ and $\mathbf{c} \in \mathbb{F}^K$, such that $\mathcal{T} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$, in which "$\circ$" denotes the *outer product*. That is, $t_{ijk} = a_i b_j c_k$ for all values of the indices.

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:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{3.1}$$

where $\mathbf{a}_r \in \mathbb{F}^I$, $\mathbf{b}_r \in \mathbb{F}^J$, $\mathbf{c}_r \in \mathbb{F}^K$ are nonzero vectors.

We call the matrices $\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \ldots & \mathbf{a}_R\end{bmatrix} \in \mathbb{F}^{I\times R}$, $\mathbf{B} = \begin{bmatrix}\mathbf{b}_1 & \ldots & \mathbf{b}_R\end{bmatrix} \in \mathbb{F}^{J\times R}$ and $\mathbf{C} = \begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_R\end{bmatrix} \in \mathbb{F}^{K\times R}$ the *first, second* and *third factor matrix* of $\mathcal{T}$, respectively. We also write (3.1) as $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$.

Definition 3.1.1. *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 $r_{\mathcal{T}}$.*

Definition 3.1.2. *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 matrix $\mathbf{T} \in \mathbb{F}^{IJ\times K}$ as follows: the (i, j, k)-th entry of $\mathcal{T}$ corresponds to the $((i-1)J + j, k)$-th entry of $\mathbf{T}$. In particular, the rank-1 tensor $\mathbf{a}\circ\mathbf{b}\circ\mathbf{c}$ corresponds to the rank-1 matrix $(\mathbf{a}\otimes\mathbf{b})\mathbf{c}^T$, in which "$\otimes$" denotes the *Kronecker product*. Thus, (3.1) can be identified with

$$\mathbf{T}^{(1)} := \mathbf{T} = \sum_{r=1}^{R} (\mathbf{a}_r \otimes \mathbf{b}_r)\mathbf{c}_r^T = [\mathbf{a}_1 \otimes \mathbf{b}_1 \ \cdots \ \mathbf{a}_R \otimes \mathbf{b}_R]\mathbf{C}^T = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T, \tag{3.2}$$

in which "$\odot$" denotes the *Khatri-Rao product* or column-wise Kronecker product. Similarly, one can reshape $\mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$ into any of the matrices

$$(\mathbf{b} \otimes \mathbf{c})\mathbf{a}^T, \quad (\mathbf{c} \otimes \mathbf{a})\mathbf{b}^T, \quad (\mathbf{a} \otimes \mathbf{c})\mathbf{b}^T, \quad (\mathbf{b} \otimes \mathbf{a})\mathbf{c}^T, \quad (\mathbf{c} \otimes \mathbf{b})\mathbf{a}^T$$

and obtain the factorizations

$$\mathbf{T}^{(2)} = (\mathbf{B} \odot \mathbf{C})\mathbf{A}^T, \qquad \mathbf{T}^{(3)} = (\mathbf{C} \odot \mathbf{A})\mathbf{B}^T, \qquad \mathbf{T}^{(4)} = (\mathbf{A} \odot \mathbf{C})\mathbf{B}^T, \qquad \text{etc.} \tag{3.3}$$

The matrices $\mathbf{T}^{(1)}$, $\mathbf{T}^{(2)}, \ldots$ are called *the matrix representations* or *matrix unfoldings* of the tensor $\mathcal{T}$.

It is clear that in (3.1)–(3.2) 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. Formally, we have the following definition.

Definition 3.1.3. *Let $\mathcal{T}$ be a tensor of rank R. The CPD of $\mathcal{T}$ is* essentially 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 $R \times R$ nonsingular diagonal matrices $\mathbf{\Lambda_A}$, $\mathbf{\Lambda_B}$, and $\mathbf{\Lambda_C}$ such that*

$$\bar{\mathbf{A}} = \mathbf{A\Pi\Lambda_A}, \quad \bar{\mathbf{B}} = \mathbf{B\Pi\Lambda_B}, \quad \bar{\mathbf{C}} = \mathbf{C\Pi\Lambda_C}, \quad \mathbf{\Lambda_A\Lambda_B\Lambda_C} = \mathbf{I}_R.$$

PDs can also be partially unique. That is, a factor matrix may be essentially unique without the overall PD being essentially unique. We will resort to the following definition.

Definition 3.1.4. *Let $\mathcal{T}$ be a tensor of rank R. The first (resp. second or third) factor matrix of $\mathcal{T}$ is* essentially 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*

$$\bar{\mathbf{A}} = \mathbf{A\Pi\Lambda_A} \quad (\textit{resp.} \quad \bar{\mathbf{B}} = \mathbf{B\Pi\Lambda_B} \quad \textit{or} \quad \bar{\mathbf{C}} = \mathbf{C\Pi\Lambda_C}).$$

For brevity, in the sequel we drop the term "essential", both when it concerns the uniqueness of the overall CPD and when it concerns the uniqueness of one factor matrix.

In this paper we present both deterministic and generic uniqueness results. Deterministic conditions concern one particular PD $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. For generic uniqueness we resort to the following definitions.

Definition 3.1.5. *Let μ be the Lebesgue measure on $\mathbb{F}^{(I+J+K)R}$. The CPD of an $I \times J \times K$ tensor of rank R is generically unique if*

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

Definition 3.1.6. *Let μ be the Lebesgue measure on $\mathbb{F}^{(I+J+K)R}$. The first (resp. second or third) factor matrix of an $I \times J \times K$ tensor of rank R is generically unique if*

$$\mu\{(\mathbf{A}, \mathbf{B}, \mathbf{C}) : \textit{ the first (resp. second or third) factor matrix of the tensor } [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R \textit{ is not unique}\} = 0.$$

Let the matrices $\mathbf{A} \in \mathbb{F}^{I \times R}$, $\mathbf{B} \in \mathbb{F}^{J \times R}$, and $\mathbf{C} \in \mathbb{F}^{K \times R}$ be randomly sampled from a continuous distribution. Generic uniqueness then means uniqueness that holds with probability one.

3.1.2 Literature overview

We refer to the overview papers [12, 3, 6] and the references therein for background, applications, and algorithms for CPD. Here, we focus on results concerning uniqueness of the CPD.

Deterministic conditions

We refer to [7, Subsection 1.2] for a detailed overview of deterministic conditions. Here we just recall three Kruskal theorems and new results from [7] that concern the uniqueness of one factor matrix. To present Kruskal's theorem we recall the definition of k-rank.

Definition 3.1.7. *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.*

Kruskal's theorem states the following.

Theorem 3.1.8. *[14, Theorem 4a, p. 123] Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and let*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2. \tag{3.4}$$

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

Kruskal also obtained the following more general results which are less known.

Theorem 3.1.9. *[14, Theorem 4b, p. 123] (see also Corollary 3.1.29 below) Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and let*

$$\begin{aligned}
&\min(k_{\mathbf{A}}, k_{\mathbf{C}}) + r_{\mathbf{B}} \geq R + 2, \\
&\min(k_{\mathbf{A}}, k_{\mathbf{B}}) + r_{\mathbf{C}} \geq R + 2, \\
&r_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R + 2 + \min(r_{\mathbf{A}} - k_{\mathbf{A}}, r_{\mathbf{B}} - k_{\mathbf{B}}), \\
&r_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R + 2 + \min(r_{\mathbf{A}} - k_{\mathbf{A}}, r_{\mathbf{C}} - k_{\mathbf{C}}).
\end{aligned}$$

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

$$H_{\mathbf{AB}}(\delta) := \min_{card(\tilde{\mathbf{A}})=\delta} [r_{\tilde{\mathbf{A}}} + r_{\tilde{\mathbf{B}}} - \delta] \quad \text{for} \quad \delta = 1, 2, \ldots, R.$$

We will say that condition (Hm) holds for the matrices $\mathbf{A}$ and $\mathbf{B}$ if

$$H_{\mathbf{AB}}(\delta) \geq \min(\delta, m) \quad \text{for} \quad \delta = 1, 2, \ldots, R. \tag{Hm}$$

The following Theorem is the strongest result about uniqueness from [14].

Theorem 3.1.10. *[14, Theorem 4e, p. 125](see also Corollary 3.1.27 below) Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and let $m_{\mathbf{B}} := R - r_{\mathbf{B}} + 2$, $m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2$. Assume that*

(i) (H1) *holds for* $\mathbf{B}$ *and* $\mathbf{C}$*;*

(ii) ($\text{Hm}_{\mathbf{B}}$) *holds for* $\mathbf{C}$ *and* $\mathbf{A}$*;*

(iii) ($\text{Hm}_{\mathbf{C}}$) *holds for* $\mathbf{A}$ *and* $\mathbf{B}$*.*

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

For the formulation of other results we recall the definition of a compound matrix.

Definition 3.1.11. *[7, Definition 2.1 and Example 2.2] The k-th compound matrix of $I \times R$ 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.*

With a vector $\mathbf{d} = \begin{bmatrix} d_1 & \ldots & d_R \end{bmatrix}^T$ we associate the vector

$$\widehat{\mathbf{d}}^m := \begin{bmatrix} d_1 \cdots d_m & d_1 \cdots d_{m-1} d_{m+1} & \ldots & d_{R-m+1} \cdots d_R \end{bmatrix}^T \in \mathbb{F}^{C_R^m}, \tag{3.5}$$

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 the sequel, we will, for instance, say that "condition (Um) holds for the matrices $\mathbf{X}$ and $\mathbf{Y}$" if condition (Um) holds for the matrices $\mathbf{A}$ and $\mathbf{B}$ replaced by the matrices $\mathbf{X}$ and $\mathbf{Y}$, respectively. We will simply write (Um) (resp. (Km),(Hm),(Cm), or (Wm)) when no confusion is possible.

It is known that conditions (K2), (C2), (U2) guarantee uniqueness of the CPD with full column rank in the third mode (see Proposition 3.1.15 below), and that condition (Km) guarantees the uniqueness of the third factor matrix [8], [7, Theorem 1.12].

In the following Proposition we gather, for later reference, properties of conditions (Km), (Cm), (Um), and (Wm) that were established in [7, §2–§3]. The proofs follow from properties of compound matrices [7, Subsection 2.1].

Proposition 3.1.12. *(1) If* (Km) *holds, then* (Cm) *and* (Hm) *hold [7, Lemmas 3.8, 3.9];*

(2) if (Cm) *or* (Hm) *holds, then* (Um) *holds [7, Lemmas 3.1, 3.10];*

(3) if (Um) *holds, then* (Wm) *holds [7, Lemma 3.3];*

(4) if (Km) *holds, then* (Kk) *holds for* $k \leq m$ *[7, Lemma 3.4];*

(5) if (Hm) *holds, then* (Hk) *holds for* $k \leq m$ *[7, Lemma 3.5];*

(6) if (Cm) *holds, then* (Ck) *holds for* $k \leq m$ *[7, Lemma 3.6];*

(7) if (Um) *holds, then* (Uk) *holds for* $k \leq m$ *[7, Lemma 3.7];*

(8) if (Wm) *holds and* $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1$*, then* (Wk) *holds for* $k \leq m$ *[7, Lemma 3.12];*

(9) if (Um) *holds, then* $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m$ *[7, Lemma 2.8].*

The following schemes illustrate Proposition 3.1.12:

$$\begin{cases} k_{\mathbf{A}} \geq m, \\ k_{\mathbf{B}} \geq m \end{cases} \quad \Leftarrow \quad \begin{array}{ccccccccc} (\mathrm{W}_m) & & (\mathrm{W}_{m\text{-}1}) & & \dots & & (\mathrm{W}_2) & & (\mathrm{W}_1) \\ \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Uparrow \\ (\mathrm{U}_m) & \Rightarrow & (\mathrm{U}_{m\text{-}1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{U}_2) & \Rightarrow & (\mathrm{U}_1) \\ \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Updownarrow \\ (\mathrm{C}_m) & \Rightarrow & (\mathrm{C}_{m\text{-}1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{C}_2) & \Rightarrow & (\mathrm{C}_1) \\ \Uparrow & & \Uparrow & & \dots & & \Uparrow & & \Uparrow \\ (\mathrm{K}_m) & \Rightarrow & (\mathrm{K}_{m\text{-}1}) & \Rightarrow & \dots & \Rightarrow & (\mathrm{K}_2) & \Rightarrow & (\mathrm{K}_1) \end{array}, \tag{3.6}$$

and

$$\text{if } \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1, \text{ then } \quad (\mathrm{W}_m) \Rightarrow (\mathrm{W}_{m\text{-}1}) \Rightarrow \dots \Rightarrow (\mathrm{W}_2) \Rightarrow (\mathrm{W}_1). \tag{3.7}$$

Scheme (3.6) also remains valid after replacing conditions $(\mathrm{C}_m),\dots,(\mathrm{C}_1)$ and equivalence $(\mathrm{C}_1) \Leftrightarrow (\mathrm{U}_1)$ by conditions $(\mathrm{H}_m),\dots,(\mathrm{H}_1)$ and implication $(\mathrm{H}_1) \Rightarrow (\mathrm{U}_1)$, respectively. One can easily construct examples where (C_m) holds but (H_m) does not hold. We do not know examples where (H_m) is more relaxed than (C_m).

Deterministic results concerning the uniqueness of one particular factor matrix were presented in [7, §4]. We first have the following proposition.

Proposition 3.1.13. *[7, 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) *the triplet of matrices* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ *satisfies conditions* $(\mathrm{W}_m), \dots, (\mathrm{W}_1)$.

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

Combining Propositions 3.1.12 and 3.1.13 we obtained the following result.

Proposition 3.1.14. *[7, Proposition 4.3, Corollaries 4.4 and 4.5] Let* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, *and* $\mathcal{T}$ *be as in Proposition 3.1.13. Assume that* $k_{\mathbf{C}} \geq 1$ *and* $m = m_{\mathbf{C}} :=$

$R - r_{\mathbf{C}} + 2$. *Then*

$$(3.4) \xRightarrow{\textit{trivial}} \begin{array}{ccc} & (\mathrm{C}_m) & \\ {\scriptstyle (3.6)}\nearrow & & \searrow{\scriptstyle (3.6)} \\ (\mathrm{K}_m) & & (\mathrm{U}_m) \\ {\scriptstyle (3.6)}\searrow & & \nearrow{\scriptstyle (3.6)} \\ & (\mathrm{H}_m) & \end{array} \xRightarrow{(3.6)} \begin{cases} (\mathrm{C}_1) \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m-1, \\ (\mathrm{W}_m) \end{cases}$$

$$\xRightarrow{(3.7)} \begin{cases} (\mathrm{C}_1) \\ (\mathrm{W}_1), \ldots, (\mathrm{W}_m) \end{cases} \Rightarrow \begin{cases} r_{\mathcal{T}} = R, \\ \textit{the third factor matrix of } \mathcal{T} \textit{ is unique.} \end{cases} \tag{3.8}$$

Note that for $r_{\mathbf{C}} = R$, we have $m = 2$ and (U_2) is equivalent to (W_2). Moreover, in this case (U_2) is necessary for uniqueness. We obtain the following counterpart of Proposition 3.1.14.

Proposition 3.1.15. *[4, 10, 15] Let* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, *and* $\mathcal{T}$ *be as in Proposition 3.1.13. Assume that* $r_{\mathbf{C}} = R$. *Then*

$$(3.4) \Rightarrow \begin{array}{ccc} & (\mathrm{C}_2) & \\ \nearrow & & \searrow \\ (\mathrm{K}_2) & & (\mathrm{U}_2) \\ \searrow & & \nearrow \\ & (\mathrm{H}_2) & \end{array} \Leftrightarrow \begin{cases} r_{\mathcal{T}} = R, \\ \textit{the CPD of } \mathcal{T} \textit{ is unique.} \end{cases} \tag{3.9}$$

Generic conditions

Let the matrices $\mathbf{A} \in \mathbb{F}^{I \times R}$, $\mathbf{B} \in \mathbb{F}^{J \times R}$, and $\mathbf{C} \in \mathbb{F}^{K \times R}$ be randomly sampled from a continuous distribution. It can be easily checked that the equations

$$k_{\mathbf{A}} = r_{\mathbf{A}} = \min(I, R), \quad k_{\mathbf{B}} = r_{\mathbf{B}} = \min(J, R), \quad k_{\mathbf{C}} = r_{\mathbf{C}} = \min(K, R)$$

hold generically. Thus, by (3.4), the CPD of an $I \times J \times K$ tensor of rank R is generically unique if

$$\min(I, R) + \min(J, R) + \min(K, R) \geq 2R + 2. \tag{3.10}$$

The generic uniqueness of one factor matrix has not yet been studied as such. It can be easily seen that in (3.8) the generic version of (K_m) for $m = R - K + 2$ is also given by (3.10).

Let us additionally assume that $K \geq R$. Under this assumption, (3.10) reduces to

$$\min(I, R) + \min(J, R) \geq R + 2.$$

The generic version of condition (C2) was given in [4, 16]. It was indicated that the $C_I^2 C_J^2 \times C_R^2$ matrix $\mathbf{U} = \mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ generically has full column rank whenever the number of columns of $\mathbf{U}$ does not exceed the number of rows. By Proposition 3.1.15 the CPD of an $I \times J \times K$ tensor of rank R is then generically unique if

$$K \geq R \qquad \text{and} \quad I(I-1)J(J-1)/4 = C_I^2 C_J^2 \geq C_R^2 = R(R-1)/2. \tag{3.11}$$

The four following results have been obtained in algebraic geometry.

Theorem 3.1.16. *[18, Corollary 3.7] Let $3 \leq I \leq J \leq K$, $K-1 \leq (I-1)(J-1)$, and let K be odd. Then the CPD of an $I \times J \times K$ tensor of rank R is generically unique if $R \leq IJK/(I+J+K-2) - K$.*

Theorem 3.1.17. *[2, Theorem 1.1] Let $I \leq J \leq K$. Let α, β be maximal such that $2^\alpha \leq I$ and $2^\beta \leq J$. Then the CPD of an $I \times J \times K$ tensor of rank R is generically unique if $R \leq 2^{\alpha+\beta-2}$.*

Theorem 3.1.18. *[2, Proposition 5.2],[18, Theorem 2.7] Let $R \leq (I-1)(J-1) \leq K$. Then the CPD of an $I \times J \times K$ tensor of rank R is generically unique.*

Theorem 3.1.19. *[2, Theorem 1.2] The CPD of an $I \times I \times I$ tensor of rank R is generically unique if $R \leq k(I)$, where $k(I)$ is given in Table 3.1.*

Table 3.1: Upper bound $k(I)$ on R under which generic uniqueness of the CPD of an $I \times I \times I$ tensor is guaranteed by Theorem 3.1.19.

I	2	3	4	5	6	7	8	9	10
$k(I)$	2	3	5	9	13	18	22	27	32

Finally, for a number of specific cases of dimensions and rank, generic uniqueness results have been obtained in [19].

3.1.3 Results and organization

In this paper we use the conditions in (3.8) to establish CPD uniqueness in cases where $r_{\mathbf{C}} < R$.

In §3.2 we assume that a tensor admits two PDs that have one or two factor matrices in common. We establish conditions under which both decompositions are the same. We obtain the following results.

Proposition 3.1.20. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_\mathbf{C}]_R$*, where* $\mathbf{\Pi}$ *is an* $R \times R$ *permutation matrix and* $\mathbf{\Lambda}_\mathbf{C}$ *is a nonsingular diagonal matrix. Let the matrices* $\mathbf{A}$*,* $\mathbf{B}$*, and* $\mathbf{C}$ *satisfy the following condition*

$$\max(\min(k_\mathbf{A}, k_\mathbf{B} - 1),\ \min(k_\mathbf{A} - 1, k_\mathbf{B})) + k_\mathbf{C} \geq R + 1. \tag{3.12}$$

Then there exist nonsingular diagonal matrices $\mathbf{\Lambda}_\mathbf{A}$ *and* $\mathbf{\Lambda}_\mathbf{B}$ *such that*

$$\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_\mathbf{A}, \qquad \bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_\mathbf{B}, \qquad \mathbf{\Lambda}_\mathbf{A}\mathbf{\Lambda}_\mathbf{B}\mathbf{\Lambda}_\mathbf{C} = \mathbf{I}_R.$$

Proposition 3.1.21. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R = [\mathbf{A}\mathbf{\Pi}_\mathbf{A}\mathbf{\Lambda}_\mathbf{A}, \bar{\mathbf{B}}, \mathbf{C}\mathbf{\Pi}_\mathbf{C}\mathbf{\Lambda}_\mathbf{C}]_R$*, where* $\mathbf{\Pi}_\mathbf{A}$ *and* $\mathbf{\Pi}_\mathbf{C}$ *are* $R \times R$ *permutation matrices and where* $\mathbf{\Lambda}_\mathbf{A}$ *and* $\mathbf{\Lambda}_\mathbf{C}$ *are nonsingular diagonal matrices. Let the matrices* $\mathbf{A}$*,* $\mathbf{B}$*, and* $\mathbf{C}$ *satisfy at least one of the following conditions*

$$\begin{aligned} &k_\mathbf{C} \geq 2 \quad \textit{and} \quad \max(\min(k_\mathbf{A}, k_\mathbf{B} - 1),\ \min(k_\mathbf{A} - 1, k_\mathbf{B})) + r_\mathbf{C} \geq R + 1, \\ &k_\mathbf{A} \geq 2 \quad \textit{and} \quad \max(\min(k_\mathbf{B}, k_\mathbf{C} - 1),\ \min(k_\mathbf{B} - 1, k_\mathbf{C})) + r_\mathbf{A} \geq R + 1. \end{aligned} \tag{3.13}$$

Then $\mathbf{\Pi}_\mathbf{A} = \mathbf{\Pi}_\mathbf{C}$ *and* $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}_\mathbf{A}\mathbf{\Lambda}_\mathbf{A}^{-1}\mathbf{\Lambda}_\mathbf{C}^{-1}$.

Note that in Propositions 3.1.20 and 3.1.21 we do not assume that R is minimal. Neither do we assume in Proposition 3.1.21 that $\mathbf{\Pi}_\mathbf{A}$ and $\mathbf{\Pi}_\mathbf{C}$ are the same.

In §3.3 we obtain new results concerning the uniqueness of the overall CPD by combining (3.8) with results from §3.2.

Combining (3.8) with Proposition 3.1.20 we prove the following statements.

Proposition 3.1.22. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m_\mathbf{C} := R - r_\mathbf{C} + 2$*. Assume that*

(i) *condition* (3.12) *holds;*

(ii) *condition* ($\mathrm{W}_{\mathrm{m}_\mathbf{C}}$) *holds for* $\mathbf{A}$*,* $\mathbf{B}$*, and* $\mathbf{C}$*;*

(iii) $\mathbf{A} \odot \mathbf{B}$ *has full column rank.* (C1)

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

Corollary 3.1.23. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m_\mathbf{C} := R - r_\mathbf{C} + 2$*. Assume that*

(i) *condition* (3.12) *holds;*

(ii) *condition* ($\mathrm{Um}_{\mathbf{C}}$) *holds for* $\mathbf{A}$ *and* $\mathbf{B}$.

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

Corollary 3.1.24. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2$. *Assume that*

(i) *condition* (3.12) *holds;*

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

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

Corollary 3.1.25. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2$. *Assume that*

(i) *condition* (3.12) *holds;*

(ii) $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B})$ *has full column rank.*

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

Note that Proposition 3.1.15 is a special case of the results in Proposition 3.1.22, Corollaries 3.1.23–3.1.25, and Kruskal's Theorem 3.1.8. In the former, one factor matrix is assumed to have full column rank ($r_{\mathbf{C}} = R$) while in the latter this is not necessary ($r_{\mathbf{C}} = R - m_{\mathbf{C}} + 2$ with $m_{\mathbf{C}} \geq 2$). The condition on $\mathbf{C}$ is relaxed by tightening the conditions on $\mathbf{A}$ and $\mathbf{B}$. For instance, Corollary 3.1.23 allows $r_{\mathbf{C}} = R - m_{\mathbf{C}} + 2$ with $m := m_{\mathbf{C}} \geq 2$ by imposing (3.12) and (Cm). From scheme (3.6) we have that (Cm) implies (C2), and hence (Cm) is more restrictive than (C2). Scheme (3.6) further shows that Corollary 3.1.23 is more general than Corollaries 3.1.24 and 3.1.25. In turn, Proposition 3.1.22 is more general than Corollary 3.1.23. Note that we did not formulate a combination of implication (Km) $\Rightarrow$ (Cm) (or (Hm)) from scheme (3.8) with Proposition 3.1.20. Such a combination leads to a result that is equivalent to Corollary 3.1.29 below.

Combining (3.8) with Proposition 3.1.21 we prove the following results.

Proposition 3.1.26. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let*

$$m_{\mathbf{A}} := R - r_{\mathbf{A}} + 2, \quad m_{\mathbf{B}} := R - r_{\mathbf{B}} + 2, \quad m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2. \tag{3.14}$$

Assume that at least two of the following conditions hold

(i) *condition* ($\mathrm{Um}_{\mathbf{A}}$) *holds for* $\mathbf{B}$ *and* $\mathbf{C}$*;*

(ii) *condition* $(\mathrm{Um}_{\mathbf{B}})$ *holds for* $\mathbf{C}$ *and* $\mathbf{A}$*;*

(iii) *condition* $(\mathrm{Um}_{\mathbf{C}})$ *holds for* $\mathbf{A}$ *and* $\mathbf{B}$*.*

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

Corollary 3.1.27. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and consider* $m_{\mathbf{A}}$*,* $m_{\mathbf{B}}$*, and* $m_{\mathbf{C}}$ *defined in* (3.14)*. Assume that at least two of the following conditions hold*

(i) *condition* $(\mathrm{Hm}_{\mathbf{A}})$ *holds for* $\mathbf{B}$ *and* $\mathbf{C}$*;*

(ii) *condition* $(\mathrm{Hm}_{\mathbf{B}})$ *holds for* $\mathbf{C}$ *and* $\mathbf{A}$*;*

(iii) *condition* $(\mathrm{Hm}_{\mathbf{C}})$ *holds for* $\mathbf{A}$ *and* $\mathbf{B}$*.*

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

Corollary 3.1.28. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and consider* $m_{\mathbf{A}}$*,* $m_{\mathbf{B}}$*, and* $m_{\mathbf{C}}$ *defined in* (3.14)*. Let at least two of the matrices*

$$\mathcal{C}_{m_{\mathbf{A}}}(\mathbf{B}) \odot \mathcal{C}_{m_{\mathbf{A}}}(\mathbf{C}), \quad \mathcal{C}_{m_{\mathbf{B}}}(\mathbf{C}) \odot \mathcal{C}_{m_{\mathbf{B}}}(\mathbf{A}), \quad \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B}) \tag{3.15}$$

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

Corollary 3.1.29. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let* $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ *coincide with* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$*,* $(\mathbf{B}, \mathbf{C}, \mathbf{A})$*, or* $(\mathbf{C}, \mathbf{A}, \mathbf{B})$*. If*

$$\begin{cases} k_{\mathbf{X}} + r_{\mathbf{Y}} + r_{\mathbf{Z}} & \geq 2R + 2, \\ \min(r_{\mathbf{Z}} + k_{\mathbf{Y}}, k_{\mathbf{Z}} + r_{\mathbf{Y}}) & \geq R + 2, \end{cases} \tag{3.16}$$

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

Corollary 3.1.30. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let the following conditions hold*

$$\begin{cases} k_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R + 2, \\ r_{\mathbf{A}} + k_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R + 2, \\ r_{\mathbf{A}} + r_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2. \end{cases} \tag{3.17}$$

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

Let us compare Kruskal's Theorems 3.1.8–3.1.10 with Corollaries 3.1.24, 3.1.27, 3.1.29, and 3.1.30. Elementary algebra yields that Theorem 3.1.9 is equivalent to Corollary 3.1.29. From Corollary 3.1.27 it follows that assumption (i) of Theorem 3.1.10 is redundant. We will demonstrate in Examples 3.3.2 and 3.3.3 that it is not possible to state in general which of the Corollaries 3.1.24 or 3.1.27

is more relaxed. Thus, Corollary 3.1.24 (obtained by combining implication (Hm) ⇒ (Um) from scheme (3.8) with Proposition 3.1.21) is an (Hm)–type result on uniqueness that was not in [14]. Corollary 3.1.30 is a special case of Corollary 3.1.29, which is obviously more relaxed than Kruskal's well-known Theorem 3.1.8. Finally we note that if condition (Hm) holds, then $r_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R+2$. Thus, neither Kruskal's Theorems 3.1.8–3.1.10 nor Corollaries 3.1.24, 3.1.27, 3.1.29, 3.1.30 can be used for demonstrating the uniqueness of a PD $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ when $r_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} < 2R + 2$.

We did not present a result based on a combination of (Wm)-type implications from scheme (3.8) with Proposition 3.1.21 because we do not have examples of cases where such conditions are more relaxed than those in Proposition 3.1.26.

In §3.4 we indicate how our results can be adapted in the case of PD symmetries.

Well-known necessary conditions for the uniqueness of the CPD are [21, p. 2079, Theorem 2], [13, p. 28],[18, p. 651]

$$\min(k_{\mathbf{A}}, k_{\mathbf{B}}, k_{\mathbf{C}}) \geq 2, \tag{3.18}$$

$$\mathbf{A} \odot \mathbf{B}, \quad \mathbf{B} \odot \mathbf{C}, \quad \mathbf{C} \odot \mathbf{A} \quad \text{have full column rank.} \tag{3.19}$$

Further, the following necessary condition was obtained in [5, Theorem 2.3]

$$(\mathrm{U}2) \text{ holds for pairs } \quad (\mathbf{A}, \mathbf{B}), \quad (\mathbf{B}, \mathbf{C}), \quad \text{and} \quad (\mathbf{C}, \mathbf{A}). \tag{3.20}$$

It follows from scheme (3.6) that (3.20) is more restrictive than (3.18) and (3.19). Our most general condition concerning uniqueness of one factor matrix is given in Proposition 3.1.13. Note that in Proposition 3.1.13, condition (i) is more relaxed than (3.18) and condition (iii) coincides with (3.19). One may wonder whether condition (iv) in Proposition 3.1.13 is necessary for the uniqueness of at least one factor matrix. In §3.5 we show that this is not the case. We actually study an example in which CPD uniqueness can be established without (Wm) being satisfied.

In §3.6 we study generic uniqueness of one factor matrix and generic CPD uniqueness. Our result on overall CPD uniqueness is the following.

Proposition 3.1.31. *The CPD of an $I \times J \times K$ tensor of rank R is generically unique if there exist matrices $\mathbf{A}_0 \in \mathbb{F}^{I \times R}$, $\mathbf{B}_0 \in \mathbb{F}^{J \times R}$, and $\mathbf{C}_0 \in \mathbb{F}^{K \times R}$ such that at least one of the following conditions holds:*

(i) $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}_0) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B}_0)$ *has full column rank, where* $m_{\mathbf{C}} = R - \min(K, R) + 2$*;*

(ii) $\mathcal{C}_{m_{\mathbf{A}}}(\mathbf{B}_0) \odot \mathcal{C}_{m_{\mathbf{A}}}(\mathbf{C}_0)$ *has full column rank, where* $m_{\mathbf{A}} = R - \min(I, R) + 2$*;*

(iii) $\mathcal{C}_{m_{\mathbf{B}}}(\mathbf{C}_0) \odot \mathcal{C}_{m_{\mathbf{B}}}(\mathbf{A}_0)$ *has full column rank, where* $m_{\mathbf{B}} = R - \min(J, R) + 2$.

We give several examples that illustrate the uniqueness results in the generic case.

3.2 Equality of PDs with common factor matrices

In this section we assume that a tensor admits two not necessarily canonical PDs that have one or two factor matrices in common. In the latter case, the two PDs may have the columns of the common factor matrices permuted differently. We establish conditions that guarantee that the two PDs are the same.

3.2.1 One factor matrix in common

In this subsection we assume that two PDs have the factor matrix $\mathbf{C}$ in common. The result that we are concerned with is Proposition 3.1.20. The proof is based on the following three lemmas.

Lemma 3.2.1. *For matrices $\mathbf{A}, \bar{\mathbf{A}} \in \mathbb{F}^{I \times R}$, and indices $r_1, \ldots, r_n \in \{1, \ldots, R\}$ define the subspaces $E_{r_1 \ldots r_n}$ and $\bar{E}_{r_1 \ldots r_n}$ as follows*

$$E_{r_1 \ldots r_n} := \operatorname{span}\{\mathbf{a}_{r_1}, \ldots, \mathbf{a}_{r_n}\}, \qquad \bar{E}_{r_1 \ldots r_n} := \operatorname{span}\{\bar{\mathbf{a}}_{r_1}, \ldots, \bar{\mathbf{a}}_{r_n}\}.$$

Assume that $k_{\mathbf{A}} \geq 2$ and that there exists $m \in \{2, \ldots, k_{\mathbf{A}}\}$ such that

$$E_{r_1 \ldots r_{m-1}} \subseteq \bar{E}_{r_1 \ldots r_{m-1}} \qquad \text{for all} \qquad 1 \leq r_1 < r_2 < \cdots < r_{m-1} \leq R. \tag{3.21}$$

Then there exists a nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\mathbf{A} = \bar{\mathbf{A}}\mathbf{\Lambda}$.

Proof. For $m = 2$ we have

$$\operatorname{span}\{\mathbf{a}_{r_1}\} = E_{r_1} \subseteq \bar{E}_{r_1} = \operatorname{span}\{\bar{\mathbf{a}}_{r_1}\}, \qquad \text{for all } 1 \leq r_1 \leq R, \tag{3.22}$$

such that the Lemma trivially holds. For $m \geq 3$ we arrive at (3.22) by downward induction on $l = m, m-1, \ldots, 3$. Assuming that

$$E_{r_1 \ldots r_{l-1}} \subseteq \bar{E}_{r_1 \ldots r_{l-1}} \qquad \text{for all} \qquad 1 \leq r_1 < r_2 < \cdots < r_{l-1} \leq R, \tag{3.23}$$

we show that

$$E_{r_1 \ldots r_{l-2}} \subseteq \bar{E}_{r_1 \ldots r_{l-2}} \qquad \text{for all} \qquad 1 \leq r_1 < r_2 < \cdots < r_{l-2} \leq R.$$

Assume $r_1, r_2, \ldots, r_{l-2}$ fixed and let $i, j \in \{1, \ldots, R\} \setminus \{r_1, \ldots, r_{l-2}\}$, with $i \neq j$. Since $l \leq m \leq k_{\mathbf{A}}$, we have that $\dim E_{r_1,\ldots,r_{l-2},i,j} = l$. Because

$$l = \dim E_{r_1,\ldots,r_{l-2},i,j} \leq \dim \operatorname{span}\{E_{r_1,\ldots,r_{l-2},i},\ E_{r_1,\ldots,r_{l-2},j}\}$$
$$\overset{(3.23)}{\leq} \dim \operatorname{span}\{\bar{E}_{r_1,\ldots,r_{l-2},i},\ \bar{E}_{r_1,\ldots,r_{l-2},j}\}$$

we have

$$\bar{E}_{r_1,\ldots,r_{l-2},i} \neq \bar{E}_{r_1,\ldots,r_{l-2},j}. \tag{3.24}$$

Therefore,

$$E_{r_1,\ldots,r_{l-2}} \subseteq \left(E_{r_1,\ldots,r_{l-2},i} \cap E_{r_1,\ldots,r_{l-2},j}\right)$$
$$\overset{(3.23)}{\subseteq} \left(\bar{E}_{r_1,\ldots,r_{l-2},i} \cap \bar{E}_{r_1,\ldots,r_{l-2},j}\right) \overset{(3.24)}{=} \bar{E}_{r_1,\ldots,r_{l-2}}.$$

The induction follows. To conclude the proof, we note that $\mathbf{\Lambda}$ is nonsingular since $k_{\mathbf{A}} \geq 2$. ∎

Lemma 3.2.2. *Let $\mathbf{C} \in \mathbb{F}^{K \times R}$ and consider m such that $m \leq k_{\mathbf{C}}$. Then for any set of distinct indices $\mathcal{I} = \{i_1, \ldots, i_{m-1}\} \subseteq \{1, \ldots, R\}$ there exists a vector $\mathbf{x} \in \mathbb{F}^K$ such that*

$$\mathbf{x}^T \mathbf{c}_i = 0 \ \text{ for } \ i \in \mathcal{I} \ \text{ and } \ \mathbf{x}^T \mathbf{c}_i \neq 0 \ \text{for } i \in \mathcal{I}^{\mathrm{c}} := \{1, \ldots, R\} \setminus \mathcal{I}. \tag{3.25}$$

Proof. Let $\mathbf{C}_{\mathcal{I}} \in \mathbb{F}^{K \times (m-1)}$ and $\mathbf{C}_{\mathcal{I}^{\mathrm{c}}} \in \mathbb{F}^{K \times (R-m+1)}$ contain the columns of $\mathbf{C}$ indexed by $\mathcal{I}$ and $\mathcal{I}^{\mathrm{c}}$, respectively, and let the columns of $\mathbf{C}_{\mathcal{I}}^{\perp} \in \mathbb{F}^{K \times (K-m+1)}$ form a basis for the orthogonal complement of range$(\mathbf{C}_{\mathcal{I}})$. The matrix $(\mathbf{C}_{\mathcal{I}}^{\perp})^H \mathbf{C}_{\mathcal{I}^{\mathrm{c}}}$ cannot have a zero column, otherwise the corresponding column of $\mathbf{C}_{\mathcal{I}^{\mathrm{c}}}$ would be in range$(\mathbf{C}_{\mathcal{I}})$, which would be a contradiction with $k_{\mathbf{C}} \geq m$. We conclude that (3.25) holds for $\mathbf{x} = (\mathbf{C}_{\mathcal{I}}^{\perp} \mathbf{y})^*$, with $\mathbf{y} \in \mathbb{F}^{K-m+1}$ generic. ∎

Lemma 3.2.3. *Let $\mathbf{P}$ be an $R \times R$ permutation matrix. Then for any vector $\lambda \in \mathbb{F}^R$,*

$$\operatorname{Diag}(\mathbf{\Pi}\lambda)\mathbf{\Pi} = \mathbf{\Pi}\operatorname{Diag}(\lambda). \tag{3.26}$$

Proof. The lemma follows directly from the definition of permutation matrix. ∎

We are now ready to prove Proposition 3.1.20.

Proof. Let $\widehat{\mathbf{A}} := \bar{\mathbf{A}}\mathbf{\Pi}^T$ and $\widehat{\mathbf{B}} := \bar{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{C}}^{-1}\mathbf{\Pi}^T$. Then

$$\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}]_R = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \mathbf{C}]_R. \tag{3.27}$$

We show that the columns of $\mathbf{A}$ and $\mathbf{B}$ coincide up to scaling with the corresponding columns of $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{B}}$, respectively. Consider indices $i_1, \ldots,$

$i_{R-k_{\mathbf{C}}+1}$ such that $1 \le i_1 < \cdots < i_{R-k_{\mathbf{C}}+1} \le R$. Let $m := k_{\mathbf{C}}$ and let $\mathcal{I} := \{1, \ldots, R\} \setminus \{i_1, \ldots, i_{R-k_{\mathbf{C}}+1}\}$. From Lemma 3.2.2 it follows that there exists a vector $\mathbf{x} \in \mathbb{F}^K$ such that

$$\mathbf{x}^T \mathbf{c}_i = 0 \ \text{ for } \ i \in \mathcal{I} \ \text{ and } \mathbf{x}^T \mathbf{c}_i \neq 0 \ \text{ for } \ i \in \mathcal{I}^{\mathrm{c}} = \{i_1, \ldots, i_{R-k_{\mathbf{C}}+1}\}.$$

Let $\mathbf{d} = \begin{bmatrix} \mathbf{x}^T \mathbf{c}_{i_1} & \ldots & \mathbf{x}^T \mathbf{c}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix}^T$. Then $(\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{x} = (\widehat{\mathbf{A}} \odot \widehat{\mathbf{B}})\mathbf{C}^T\mathbf{x}$ is equivalent to

$$\left(\begin{bmatrix} \mathbf{a}_{i_1} & \ldots & \mathbf{a}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix} \odot \begin{bmatrix} \mathbf{b}_{i_1} & \ldots & \mathbf{b}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix} \right) \mathbf{d} =$$

$$\left(\begin{bmatrix} \widehat{\mathbf{a}}_{i_1} & \ldots & \widehat{\mathbf{a}}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix} \odot \begin{bmatrix} \widehat{\mathbf{b}}_{i_1} & \ldots & \widehat{\mathbf{b}}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix} \right) \mathbf{d},$$

which may be expressed as

$$\begin{bmatrix} \mathbf{a}_{i_1} & \ldots & \mathbf{a}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix} \mathrm{Diag}(\mathbf{d}) \begin{bmatrix} \mathbf{b}_{i_1} & \ldots & \mathbf{b}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix}^T$$

$$= \begin{bmatrix} \widehat{\mathbf{a}}_{i_1} & \ldots & \widehat{\mathbf{a}}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix} \mathrm{Diag}(\mathbf{d}) \begin{bmatrix} \widehat{\mathbf{b}}_{i_1} & \ldots & \widehat{\mathbf{b}}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix}^T.$$

By (3.12), $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge R - k_{\mathbf{C}} + 1$. Hence, the matrices $\begin{bmatrix} \mathbf{a}_{i_1} & \ldots & \mathbf{a}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{b}_{i_1} & \ldots & \mathbf{b}_{i_{R-k_{\mathbf{C}}+1}} \end{bmatrix}$ have full column rank. Since by construction the vector $\mathbf{d}$ has only nonzero components, it follows that

$$\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_{R-k_{\mathbf{C}}+1}} \in \mathrm{span}\{\widehat{\mathbf{a}}_{i_1}, \ldots, \widehat{\mathbf{a}}_{i_{R-k_{\mathbf{C}}+1}}\},$$

$$\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_{R-k_{\mathbf{C}}+1}} \in \mathrm{span}\{\widehat{\mathbf{b}}_{i_1}, \ldots, \widehat{\mathbf{b}}_{i_{R-k_{\mathbf{C}}+1}}\}.$$

By (3.12), $\max(k_{\mathbf{A}}, k_{\mathbf{B}}) \ge m := R - k_{\mathbf{C}} + 2 \ge 2$. Without loss of generality we confine ourselves to the case $k_{\mathbf{A}} \ge m$. Then, by Lemma 3.2.1, there exists a nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\mathbf{A} = \widehat{\mathbf{A}}\mathbf{\Lambda}$. Denoting $\lambda_{\mathbf{A}} := \mathbf{\Pi}^T \mathrm{diag}(\mathbf{\Lambda}^{-1})$ and $\mathbf{\Lambda}_{\mathbf{A}} = \mathrm{Diag}(\lambda_{\mathbf{A}})$ and applying Lemma 3.2.3, we have

$$\bar{\mathbf{A}} = \widehat{\mathbf{A}}\mathbf{\Pi} = \mathbf{A}\mathbf{\Lambda}^{-1}\mathbf{\Pi} = \mathbf{A}\mathrm{Diag}(\mathbf{\Pi}\lambda_{\mathbf{A}})\mathbf{\Pi} = \mathbf{A}\mathbf{\Pi}\mathrm{Diag}(\lambda_{\mathbf{A}}) = \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}}.$$

It follows from (3.27) and (3.2) that

$$(\mathbf{C} \odot \mathbf{A})\mathbf{B}^T = (\mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}} \odot \bar{\mathbf{A}})\bar{\mathbf{B}}^T = (\mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}} \odot \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}})\bar{\mathbf{B}}^T$$

$$= (\mathbf{C} \odot \mathbf{A})\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}\mathbf{\Lambda}_{\mathbf{A}}\bar{\mathbf{B}}^T.$$

Since $k_{\mathbf{A}} \ge R - k_{\mathbf{C}} + 2$, it follows that condition (K1) holds for the matrices $\mathbf{A}$ and $\mathbf{C}$. From Proposition 3.1.12 (1) it follows that the matrix $\mathbf{C} \odot \mathbf{A}$ has full column rank. Hence, $\mathbf{B}^T = \mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{C}}\mathbf{\Lambda}_{\mathbf{A}}\bar{\mathbf{B}}^T$, i.e., $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{A}}^{-1}\mathbf{\Lambda}_{\mathbf{C}}^{-1} =: \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_{\mathbf{B}}$. □

Example 3.2.4. *Consider the* $2 \times 3 \times 3$ *tensor given by* $\mathcal{T} = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_3$*, where*

$$\widehat{\mathbf{A}} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 3 \end{bmatrix}, \quad \widehat{\mathbf{B}} = \begin{bmatrix} 6 & 12 & 2 \\ 3 & 4 & -1 \\ 4 & 6 & -4 \end{bmatrix}, \quad \widehat{\mathbf{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since $k_{\widehat{\mathbf{A}}} + k_{\widehat{\mathbf{B}}} + k_{\widehat{\mathbf{C}}} = 2 + 3 + 3 \geq 2 \times 3 + 2$*, it follows from Theorem 3.1.8 that* $r_{\mathcal{T}} = 3$ *and that the CPD of* $\mathcal{T}$ *is unique.*

Increasing the number of terms, we also have $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_4$ *for*

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

$$\mathbf{C} = \begin{bmatrix} 6 & -6 & -3 & -2 \\ 12 & -24 & -8 & -6 \\ 2 & 6 & -3 & -6 \end{bmatrix}.$$

Since $k_{\mathbf{A}} = 2$ *and* $k_{\mathbf{B}} = k_{\mathbf{C}} = 3$*, condition* (3.12) *holds. Hence, by Proposition 3.1.20, if* $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_4$ *and* $\bar{\mathbf{C}} = \mathbf{C}$*, then there exists a nonsingular diagonal matrix* $\mathbf{\Lambda}$ *such that* $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Lambda}$ *and* $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Lambda}^{-1}$*.*

The following condition is also satisfied:

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

By symmetry, we have from Proposition 3.1.20 that, if $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_4$ *and* $\bar{\mathbf{B}} = \mathbf{B}$*, then there exists a nonsingular diagonal matrix* $\mathbf{\Lambda}$ *such that* $\bar{\mathbf{A}} = \mathbf{A}\mathbf{\Lambda}$ *and* $\bar{\mathbf{C}} = \mathbf{C}\mathbf{\Lambda}^{-1}$*.*

Finally, we show that the inequality of condition (3.12) *is sharp. We have*

$$\max(\min(k_{\mathbf{B}}, k_{\mathbf{C}} - 1),\ \min(k_{\mathbf{B}} - 1, k_{\mathbf{C}})) + k_{\mathbf{A}} = R < R + 1.$$

One can verify that $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_4$ *with* $\bar{\mathbf{A}} = \mathbf{A}$ *and with* $\bar{\mathbf{B}}$ *and* $\bar{\mathbf{C}}$ *given by*

$$\bar{\mathbf{B}} = \begin{bmatrix} 6 & 12 & 2 \\ 3 & 4 & -1 \\ 4 & 6 & -4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4/3 & 3/2 \\ 1 & -3 & 3 & 9 \end{bmatrix},$$

$$\bar{\mathbf{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\alpha & 0 \\ 0 & 0 & 1/\beta \end{bmatrix} \begin{bmatrix} 6 & -6 & -3 & -2 \\ -24/5 & 48/5 & 16/5 & 12/5 \\ 2/15 & 2/5 & -1/5 & -2/5 \end{bmatrix}$$

for arbitrary nonzero α *and* β*. Hence, there exist infinitely many PDs* $\mathcal{T} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_4$ *with* $\bar{\mathbf{A}} = \mathbf{A}$*; the columns of* $\bar{\mathbf{B}}$ *and* $\bar{\mathbf{C}}$ *are only proportional to the columns of* $\mathbf{B}$ *and* $\mathbf{C}$*, respectively, for* $\alpha = -2/5$ *and* $\beta = 1/15$*. We conclude that the inequality of condition* (3.12) *is sharp.*

3.2.2 Two factor matrices in common

In this subsection we assume that two PDs have the factor matrices $\mathbf{A}$ and $\mathbf{C}$ in common. We do not assume however that in the two PDs the columns of these matrices are permuted in the same manner. The result that we are concerned with is Proposition 3.1.21.

Proof. Without loss of generality, we confine ourselves to the case

$$k_{\mathbf{C}} \geq 2 \quad \text{and} \quad \min(k_{\mathbf{A}} - 1, k_{\mathbf{B}}) + r_{\mathbf{C}} \geq R + 1. \tag{3.28}$$

We set for brevity $r := r_{\mathbf{C}}$. Denoting $\mathbf{\Pi} = \mathbf{\Pi}_{\mathbf{A}}\mathbf{\Pi}_{\mathbf{C}}^T$ and $\widehat{\mathbf{B}} = \bar{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{C}}\mathbf{\Pi}_{\mathbf{C}}^T$, we have $[\mathbf{A}\mathbf{\Pi}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{A}}, \bar{\mathbf{B}}, \mathbf{C}\mathbf{\Pi}_{\mathbf{C}}\mathbf{\Lambda}_{\mathbf{C}}]_R = [\mathbf{A}\mathbf{\Pi}_{\mathbf{A}}\mathbf{\Pi}_{\mathbf{C}}^T, \bar{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{C}}\mathbf{\Pi}_{\mathbf{C}}^T, \mathbf{C}]_R = [\mathbf{A}\mathbf{\Pi}, \widehat{\mathbf{B}}, \mathbf{C}]_R$. We will show that, under (3.28), $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_R = [\mathbf{A}\mathbf{\Pi}, \widehat{\mathbf{B}}, \mathbf{C}]_R$ implies that $\mathbf{\Pi} = \mathbf{I}_R$. This, in turn, immediately implies that $\mathbf{\Pi}_{\mathbf{A}} = \mathbf{\Pi}_{\mathbf{C}}$ and $\bar{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{A}}^{-1}\mathbf{\Lambda}_{\mathbf{C}}^{-1}$.

(i) Let us fix integers $i_1, \ldots, i_r$ such that the columns $\mathbf{c}_{i_1}, \ldots, \mathbf{c}_{i_r}$ form a basis of range($\mathbf{C}$) and let us set $\{j_1, \ldots, j_{R-r}\} := \{1, \ldots, R\} \setminus \{i_1, \ldots, i_r\}$. Let $\mathbf{X} \in \mathbb{F}^{K \times r}$ denote a right inverse of $\begin{bmatrix}\mathbf{c}_{i_1} & \ldots & \mathbf{c}_{i_r}\end{bmatrix}^T$, i.e., $\begin{bmatrix}\mathbf{c}_{i_1} & \ldots & \mathbf{c}_{i_r}\end{bmatrix}^T \mathbf{X} = \mathbf{I}_r$. Define the subspaces $E, E_{i_k} \subseteq \mathbb{F}^R$ as follows:

$$E = \text{span}\{\mathbf{e}_{j_1}^R, \ldots \mathbf{e}_{j_{R-r}}^R\},$$

$$E_{i_k} = \text{span}\{\mathbf{e}_l^R : \ \mathbf{c}_l^T \mathbf{x}_k \neq 0, \ l \in \{j_1, \ldots, j_{R-r}\}\}, \qquad k \in \{1, \ldots, r\}.$$

By construction, $E_{i_k} \subseteq E$ and $\mathbf{e}_{i_l}^R \notin E_{i_k}$, $k, l \in \{1, \ldots, r\}$.

(ii) Let us show that $\mathbf{\Pi}\text{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\} = \text{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\}$ for all $k \in \{1, \ldots, r\}$. Let us fix $k \in \{1, \ldots, r\}$. Assume that $\mathbf{C}^T\mathbf{x}_k$ has nonzero entries at positions $k_1, \ldots, k_L$. Denote these entries by $\alpha_1, \ldots, \alpha_L$. From the definition of $\mathbf{X}$ and E_{i_k} it follows that $L \leq R - r + 1$ and $\text{span}\{\mathbf{e}_{k_1}^R, \ldots, \mathbf{e}_{k_L}^R\} = \text{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\}$. Define $\mathbf{P}_k = \begin{bmatrix}\mathbf{e}_{k_1}^R & \ldots & \mathbf{e}_{k_L}^R\end{bmatrix}$. Then we have

$$\mathbf{P}_k\mathbf{P}_k^T\text{Diag}(\mathbf{C}^T\mathbf{x}_k)\mathbf{P}_k\mathbf{P}_k^T = \text{Diag}(\mathbf{C}^T\mathbf{x}_k), \tag{3.29}$$

$$\mathbf{P}_k^T\text{Diag}(\mathbf{C}^T\mathbf{x}_k)\mathbf{P}_k = \text{Diag}(\begin{bmatrix}\alpha_1 & \ldots & \alpha_L\end{bmatrix}). \tag{3.30}$$

Further, $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_R = [\mathbf{A}\mathbf{\Pi}, \widehat{\mathbf{B}}, \mathbf{C}]_R$ implies that

$$\mathbf{A}\text{Diag}(\mathbf{C}^T\mathbf{x}_k)\mathbf{B}^T = \mathbf{A}\mathbf{\Pi}\text{Diag}(\mathbf{C}^T\mathbf{x}_k)\widehat{\mathbf{B}}^T. \tag{3.31}$$

Using (3.29)–(3.31), we obtain

$$\begin{aligned}\mathbf{A}\mathbf{P}_k\mathrm{Diag}(\begin{bmatrix}\alpha_1 & \dots & \alpha_L\end{bmatrix})\mathbf{P}_k^T\mathbf{B}^T &= \mathbf{A}\mathbf{P}_k\mathbf{P}_k^T\mathrm{Diag}(\mathbf{C}^T\mathbf{x}_k)\mathbf{P}_k\mathbf{P}_k^T\mathbf{B}^T \\ &= \mathbf{A}\mathrm{Diag}(\mathbf{C}^T\mathbf{x}_k)\mathbf{B}^T \\ &= \mathbf{A}\boldsymbol{\Pi}\mathrm{Diag}(\mathbf{C}^T\mathbf{x}_k)\widehat{\mathbf{B}}^T \qquad (3.32)\\ &= \mathbf{A}\boldsymbol{\Pi}\mathbf{P}_k\mathbf{P}_k^T\mathrm{Diag}(\mathbf{C}^T\mathbf{x}_k)\mathbf{P}_k\mathbf{P}_k^T\widehat{\mathbf{B}}^T \\ &= \mathbf{A}\boldsymbol{\Pi}\mathbf{P}_k\mathrm{Diag}(\begin{bmatrix}\alpha_1 & \dots & \alpha_L\end{bmatrix})\mathbf{P}_k^T\widehat{\mathbf{B}}^T.\end{aligned}$$

Note that $\mathbf{B}\mathbf{P}_k = \begin{bmatrix}\mathbf{b}_{k_1} & \dots & \mathbf{b}_{k_L}\end{bmatrix}$. Since, by (3.28), $k_{\mathbf{B}} \geq R - r + 1 \geq L$, it follows that the matrix $\mathbf{P}_k^T\widehat{\mathbf{B}}^T$ has full row rank. Further noting that $\mathbf{A}\mathbf{P}_k = \begin{bmatrix}\mathbf{a}_{k_1} & \dots & \mathbf{a}_{k_L}\end{bmatrix}$ and $\mathbf{A}\boldsymbol{\Pi}\mathbf{P}_k = \begin{bmatrix}(\mathbf{A}\boldsymbol{\Pi})_{k_1} & \dots & (\mathbf{A}\boldsymbol{\Pi})_{k_L}\end{bmatrix}$, we obtain from (3.32) that

$$\mathrm{span}\{\mathbf{a}_{k_1}, \dots, \mathbf{a}_{k_L}\} \subseteq \mathrm{span}\{(\mathbf{A}\boldsymbol{\Pi})_{k_1}, \dots, (\mathbf{A}\boldsymbol{\Pi})_{k_L}\}. \qquad (3.33)$$

Since, by (3.28), $k_{\mathbf{A}} \geq R - r + 2 \geq L + 1$, (3.33) is only possible if $\boldsymbol{\Pi}\mathrm{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\} = \mathrm{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\}$.

(iii) Let us show that $\boldsymbol{\Pi}E = E$. Let us fix $j \in \{j_1, \dots, j_{R-r}\}$. From $\mathbf{X}^T\mathbf{c}_{i_k} = \mathbf{e}_k^r$ for $k \in \{1, \dots, r\}$, the fact that the vectors $\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_r}$ form a basis of range($\mathbf{C}$), and $k_{\mathbf{C}} \geq 2$, it follows that the vector $\mathbf{X}^T\mathbf{c}_j$ has at least two nonzero components, say, the m-th and n-th component. Since $\mathbf{c}_j^T\mathbf{x}_m \neq 0$ and $\mathbf{c}_j^T\mathbf{x}_n \neq 0$, we have $\mathbf{e}_j^R \in E_{i_m} \cap E_{i_n}$. From the preceding steps we have

$$\begin{aligned}\boldsymbol{\Pi}\mathbf{e}_j^R \in \boldsymbol{\Pi}(E_{i_m} \cap E_{i_n}) &\overset{\text{(i)}}{=} \boldsymbol{\Pi}\left(\mathrm{span}\{E_{i_m}, \mathbf{e}_{i_m}^R\} \cap \mathrm{span}\{E_{i_n}, \mathbf{e}_{i_n}^R\}\right) \\ &\overset{\text{(ii)}}{\subseteq} \mathrm{span}\{E_{i_m}, \mathbf{e}_{i_m}^R\} \cap \mathrm{span}\{E_{i_n}, \mathbf{e}_{i_n}^R\} \overset{\text{(i)}}{=} E_{i_m} \cap E_{i_n} \subseteq E.\end{aligned}$$

Since this holds true for any index $j \in \{j_1, \dots, j_{R-r}\}$, it follows that $\boldsymbol{\Pi}E = E$.

(iv) Let us show that $\boldsymbol{\Pi}\mathbf{e}_{i_k}^R = \mathbf{e}_{i_k}^R$ for all $k \in \{1, \dots, r\}$. From the preceding steps we have

$$\boldsymbol{\Pi}E_{i_k} \overset{\text{(i)}}{=} \boldsymbol{\Pi}\left(\mathrm{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\} \cap E\right) \overset{\text{(ii), (iii)}}{\subseteq} \mathrm{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\} \cap E \overset{\text{(i)}}{=} E_{i_k}.$$

On the other hand, we have from step (iii) that $\boldsymbol{\Pi}\mathrm{span}\{E_{i_k}, \mathbf{e}_{i_k}^R\} = \{E_{i_k}, \mathbf{e}_{i_k}^R\}$, with, as shown in step (i), $\mathbf{e}_{i_k}^R \notin E_{i_k}$. It follows that $\boldsymbol{\Pi}\mathbf{e}_{i_k}^R = \mathbf{e}_{i_k}^R$ for all $k \in \{1, \dots, r\}$.

(v) We have so far shown that, if the columns $\mathbf{c}_{i_1}, \ldots, \mathbf{c}_{i_r}$ form a basis of range($\mathbf{C}$), then $\mathbf{\Pi} \begin{bmatrix} \mathbf{e}_{i_1}^R & \ldots \mathbf{e}_{i_r}^R \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{i_1}^R & \ldots \mathbf{e}_{i_r}^R \end{bmatrix}$. To complete the proof of the overall equality $\mathbf{\Pi} = \mathbf{I}_R$, it suffices to note that a basis of range($\mathbf{C}$) can be constructed starting from any column of $\mathbf{C}$.

□

3.3 Overall CPD uniqueness

In Proposition 3.1.22 and Corollaries 3.1.23–3.1.25 overall CPD uniqueness is derived from uniqueness of one factor matrix, where the latter is guaranteed by Proposition 3.1.20. In Proposition 3.1.26 and Corollaries 3.1.28–3.1.30 overall CPD is derived from uniqueness of two factor matrices, where the latter is guaranteed by Proposition 3.1.21. We illustrate our results with some examples.

Proof of Proposition 3.1.22. By (3.12), $k_{\mathbf{C}} \geq 1$ and $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m_{\mathbf{C}} - 1$. Hence, by Proposition 3.1.14, $r_{\mathcal{T}} = R$ and the third factor matrix of $\mathcal{T}$ is unique. The result now follows from Proposition 3.1.20. □

Proof of Corollary 3.1.23. From Proposition 3.1.12 (3) it follows that ($\mathrm{W}_{\mathrm{m}_{\mathbf{C}}}$) holds for $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. Since (U1) is equivalent to (C1), it follows from Proposition 3.1.12 (7) that $\mathbf{A} \odot \mathbf{B}$ has full column rank. The result now follows from Proposition 3.1.22. □

Proof of Corollaries 3.1.24 and 3.1.25. By Proposition 3.1.12 (2), both ($\mathrm{H}_{\mathrm{m}_{\mathbf{C}}}$) and ($\mathrm{C}_{\mathrm{m}_{\mathbf{C}}}$) imply ($\mathrm{U}_{\mathrm{m}_{\mathbf{C}}}$). The result now follows from Corollary 3.1.23. □

Proof of Proposition 3.1.26. Without loss of generality we assume that (i) and (iii) hold. By Proposition 3.1.12 (9),

$$\min(k_{\mathbf{B}}, k_{\mathbf{C}}) \geq m_{\mathbf{A}} \geq 2, \qquad \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m_{\mathbf{C}} \geq 2. \tag{3.34}$$

It follows from Proposition 3.1.14 that $r_{\mathcal{T}} = R$ and that the first and third factor matrices of the tensor $\mathcal{T}$ are unique. One can easily check that (3.34) implies (3.13). Hence, by Proposition 3.1.21, the CPD of $\mathcal{T}$ is unique. □

Proof of Corollary 3.1.27. Without loss of generality we assume that (ii) and (iii) hold. From Proposition 3.1.12 (2) it follows that (ii) and (iii) in Proposition 3.1.26 also hold. Hence, by Proposition 3.1.26, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Proof of Corollary 3.1.28. By Proposition 3.1.12 (2), if two of the matrices in (3.15) have full column rank, then at least two of conditions (i)–(iii) in Proposition 3.1.26 hold. Hence, by Proposition 3.1.26, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Proof of Corollary 3.1.29. Without loss of generality we assume that $(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) = (\mathbf{B}, \mathbf{C}, \mathbf{A})$. Then,

$$\begin{cases} \begin{cases} k_{\mathbf{B}} + r_{\mathbf{A}} + r_{\mathbf{C}} & \geq 2R+2, \\ k_{\mathbf{A}} + r_{\mathbf{C}} & \geq R+2, \end{cases} \\ \begin{cases} k_{\mathbf{B}} + r_{\mathbf{A}} + r_{\mathbf{C}} & \geq 2R+2, \\ r_{\mathbf{A}} + k_{\mathbf{C}} & \geq R+2, \end{cases} \end{cases} \Rightarrow \begin{cases} (\mathrm{Km}_{\mathbf{A}}) \text{ holds for } \mathbf{B} \text{ and } \mathbf{C}, \\ (\mathrm{Km}_{\mathbf{C}}) \text{ holds for } \mathbf{A} \text{ and } \mathbf{B}, \end{cases}$$

where $m_{\mathbf{A}} = R - r_{\mathbf{A}} + 2$ and $m_{\mathbf{C}} = R - r_{\mathbf{C}} + 2$. From Proposition 3.1.12 (1) it follows that the matrices $\mathcal{C}_{m_{\mathbf{A}}}(\mathbf{B}) \odot \mathcal{C}_{m_{\mathbf{A}}}(\mathbf{C})$ and $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B})$ have full column rank. Hence, by Corollary 3.1.28, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Proof of Corollary 3.1.30. It can be easily checked that all conditions of Corollary 3.1.29 hold. Hence, $r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique. □

Example 3.3.1. *Consider a* $5 \times 5 \times 5$ *tensor given by the PD* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_6$, *where the matrices* $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{C}^{5\times 6}$ *satisfy*

$$r_{\mathbf{A}} = r_{\mathbf{B}} = r_{\mathbf{C}} = 5, \qquad k_{\mathbf{A}} = k_{\mathbf{B}} = k_{\mathbf{C}} = 4.$$

For instance, consider

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

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

where $*$ *denotes arbitrary nonzero entries. Then Kruskal's condition* (3.4) *does not hold. On the other hand, the conditions of Corollary 3.1.29 are satisfied. Hence, the PD of* $\mathcal{T}$ *is canonical and unique.*

Example 3.3.2. *Consider the* $4 \times 4 \times 4$ *tensor given by the PD* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_5$, *where*

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

We have

$$r_{\mathbf{A}} = r_{\mathbf{B}} = r_{\mathbf{C}} = 4, \qquad k_{\mathbf{A}} = k_{\mathbf{B}} = k_{\mathbf{C}} = 3, \qquad m_{\mathbf{A}} = m_{\mathbf{B}} = m_{\mathbf{C}} = 3.$$

Hence, Kruskal's condition (3.4) *does not hold. Moreover, condition* (K3) *does not hold for* $(\mathbf{A}, \mathbf{B})$*,* $(\mathbf{C}, \mathbf{A})$*, nor* $(\mathbf{B}, \mathbf{C})$*. Hence, the conditions of Corollary 3.1.29 are not satisfied. On the other hand, we have*

$$\begin{aligned}
&\mathcal{C}_3(\mathbf{A}) \odot \mathcal{C}_3(\mathbf{B}) \\
&= \begin{bmatrix} \mathbf{e}_1^{16} & \mathbf{e}_6^{16} & \mathbf{e}_2^{16} & \mathbf{e}_{11}^{16} & \mathbf{e}_{1,-3}^{16} & \mathbf{e}_{6,10}^{16} & \mathbf{e}_{16}^{16} & \mathbf{e}_{1,4}^{16} & \mathbf{e}_{6,-14}^{16} & \mathbf{e}_{11,12,15,16}^{16} \end{bmatrix}, \\
&\mathcal{C}_3(\mathbf{C}) \odot \mathcal{C}_3(\mathbf{A}) \\
&= \begin{bmatrix} \mathbf{e}_1^{16} & \mathbf{e}_6^{16} & \mathbf{e}_{1,5}^{16} & \mathbf{e}_{11}^{16} & -\mathbf{e}_9^{16} & \mathbf{e}_{10,11}^{16} & \mathbf{e}_{16}^{16} & \mathbf{e}_{1,13}^{16} & \mathbf{e}_{6,16,-8,-14}^{16} & \mathbf{e}_{11,12}^{16} \end{bmatrix}, \\
&\mathcal{C}_3(\mathbf{B}) \odot \mathcal{C}_3(\mathbf{C}) \\
&= \begin{bmatrix} \mathbf{e}_1^{16} & \mathbf{e}_6^{16} & \mathbf{e}_{5,6}^{16} & \mathbf{e}_{11}^{16} & \mathbf{e}_{11,-3}^{16} & \mathbf{e}_7^{16} & \mathbf{e}_{16}^{16} & \mathbf{e}_{1,4,13,16}^{16} & \mathbf{e}_{6,-8}^{16} & \mathbf{e}_{11,15}^{16} \end{bmatrix},
\end{aligned}$$

where

$$\mathbf{e}_{i,\pm j}^{16} := \mathbf{e}_i^{16} \pm \mathbf{e}_j^{16}, \quad \mathbf{e}_{i,j,\pm k,\pm l}^{16} := \mathbf{e}_i^{16} + \mathbf{e}_j^{16} \pm \mathbf{e}_k^{16} \pm \mathbf{e}_l^{16}, \quad i,j,k,l \in \{1,\dots,16\}.$$

It is easy to check that the matrices $\mathcal{C}_3(\mathbf{A}) \odot \mathcal{C}_3(\mathbf{B})$*,* $\mathcal{C}_3(\mathbf{C}) \odot \mathcal{C}_3(\mathbf{A})$*, and* $\mathcal{C}_3(\mathbf{B}) \odot \mathcal{C}_3(\mathbf{C})$ *have full column rank. Hence, by Corollary 3.1.28, the PD is canonical and unique.*

One can easily verify that $H_{\mathbf{AB}}(\delta) = H_{\mathbf{BC}}(\delta) = H_{\mathbf{CA}}(\delta) = \min(\delta, 3)$*. Hence the uniqueness of the CPD follows also from Corollary 3.1.27.*

Note that, since condition (3.12) *does not hold, the result does not follow from Proposition 3.1.22 and its Corollaries 3.1.23–3.1.25.*

Example 3.3.3. *Consider the* $5 \times 5 \times 8$ *tensor given by the PD* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_8$*, where*

$$\mathbf{A} = \begin{bmatrix} \widehat{\mathbf{A}} \\ (\mathbf{e}_1^8)^T \end{bmatrix} \in \mathbb{F}^{5\times 8}, \qquad \mathbf{B} = \begin{bmatrix} \widehat{\mathbf{B}} \\ (\mathbf{e}_8^8)^T \end{bmatrix} \in \mathbb{F}^{5\times 8}, \qquad \mathbf{C} = \mathbf{I}_8,$$

and $\widehat{\mathbf{A}}$ *and* $\widehat{\mathbf{B}}$ *are* 4×8 *matrices such that* $k_{\widehat{\mathbf{A}}} = k_{\widehat{\mathbf{B}}} = 4$*. We have* $r_{\mathbf{A}} = r_{\mathbf{B}} = 5$*,* $k_{\mathbf{A}} = k_{\mathbf{B}} = 4$*, and* $r_{\mathbf{C}} = k_{\mathbf{C}} = 8$*. One can easily check that*

$$H_{\mathbf{AB}}(\delta) = \begin{cases} \delta, & 1 \le \delta \le 4, \\ 3, & \delta = 5, \\ 2, & 6 \le \delta \le 8 \end{cases} \qquad \ge \min(\delta, 8 - 8 + 2)$$

and that condition (3.12) *holds. Hence, by Corollary 3.1.24, the PD is canonical and unique. On the other hand,* $H_{\mathbf{BC}}(\delta) = H_{\mathbf{CA}}(\delta) = 4 < \min(\delta, 8-5+2)$ *for* $\delta = 5$*. Hence, the result does not follow from Corollary 3.1.27.*

Example 3.3.4. *Let*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 3 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 & 5 \end{bmatrix}, \qquad \mathbf{C} = \mathbf{I}_5.$$

It has already been shown in [17] that the CPD of the tensor $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_5$ *is unique. We give a shorter proof, based on Corollary 3.1.23. It is easy to verify that*

$$\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}) = \begin{bmatrix} 1 & 0 & 1 & 6 & 0 & 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 10 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & -1 & -5 & 0 & 0 & 2 \\ 0 & 0 & 1 & 9 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 1 & 1 & 15 & 0 & 0 & 0 & 1 & 1 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & -1 & -3 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 & 1 & 15 & 1 & 6 & 2 \end{bmatrix},$$

$$\ker(\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})) = \operatorname{span}\{\begin{bmatrix} 0 & 0 & -4 & 0 & 0 & 2 & 0 & -4 & 0 & -1 \end{bmatrix}^T\}.$$

If $\mathbf{d} \in \mathbb{C}^5$ *is such that* $\operatorname{diag}(\mathcal{C}_2(\operatorname{Diag}(\mathbf{d}))) \in \ker(\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B}))$*, we have*

$$\begin{aligned} &d_1 d_2 = 0, && d_2 d_3 = 0, && d_3 d_4 = -4c, && d_4 d_5 = -c, \\ &d_1 d_3 = 0, && d_2 d_4 = 2c, && d_3 d_5 = 0, \\ &d_1 d_4 = -4c, && d_2 d_5 = 0, \\ &d_1 d_5 = 0. \end{aligned}$$

One can check that this set of equations only has a solution if $c = 0$*, in which case* $\mathbf{d} = \mathbf{0}$*. Hence, by Corollary 3.1.23, the PD is canonical and unique. Note that, since* $m_{\mathbf{A}} = m_{\mathbf{B}} = 5 - 3 + 2 = 4$*, the* $m_{\mathbf{A}}$*-th compound matrix of* $\mathbf{A}$ *and the* $m_{\mathbf{B}}$*-th compound matrix of* $\mathbf{B}$ *are not defined. Hence, the uniqueness of the matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *does not follow from Proposition 3.1.26.*

Example 3.3.5. *Experiments indicate that for random* 7×10 *matrices* $\mathbf{A}$ *and* $\mathbf{B}$*, the matrix* $\mathbf{A} \odot \mathbf{B}$ *has full column rank and that condition* (U5) *does not hold. Namely, the kernel of the* 441×252 *matrix* $\mathcal{C}_5(\mathbf{A}) \odot \mathcal{C}_5(\mathbf{B})$ *is spanned*

by a vector $\widehat{\mathbf{d}}^5$ *associated with some* $\mathbf{d} \in \mathbb{F}^{10}$. *Let* $\mathbf{C}$ *be a* 7×10 *matrix such that* $\mathbf{d} \notin \operatorname{range}(\mathbf{C}^T)$. *Then* (W5) *holds for the triplet* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$. *If additionally* $k_{\mathbf{C}} \geq 5$, *then* (3.12) *holds. Hence, by Proposition 3.1.22,* $r_{\mathcal{T}} = 10$ *and the CPD of* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_{10}$ *is unique.*

The same situation occurs for tensors with other dimensions (see Table 3.2).

Table 3.2: Some cases where the rank and the uniqueness of the CPD of $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ may be easily obtained from Proposition 3.1.22 or its Corollary 3.1.23 (see Example 3.3.5). Matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are generated randomly. Simulations indicate that the dimensions of $\mathbf{A}$ and $\mathbf{B}$ cause the dimension of $\ker(\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}))$ to be equal to 1. Thus, (Um) and (Wm) may be easily checked.

dimensions of $\mathcal{T}$, $I \times J \times K$	$r_{\mathcal{T}} = R$	m=R-K+2	dimensions of $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$	(Um)	(Wm)
$4 \times 5 \times 6$	7	3	40×35	does not hold	holds
$4 \times 6 \times 14$	14	2	90×91	holds	holds
$5 \times 7 \times 7$	9	4	175×216	does not hold	holds
$6 \times 9 \times 8$	11	5	756×462	does not hold	holds
$7 \times 7 \times 7$	10	5	441×252	does not hold	holds

3.4 Application to tensors with symmetric frontal slices and Indscal

In this section we consider tensors with symmetric frontal slices (SFS), which we will briefly call SFS-tensors. We are interested in PDs of which the rank-1 terms have the same symmetry. Such decompositions correspond to the INDSCAL model, as introduced by Carroll and Chang [1]. A similar approach may be followed in the case of full symmetry.

We start with definitions of SFS-rank, SFS-PD, and SFS-CPD.

Definition 3.4.1. *A third-order SFS-tensor* $\mathcal{T} \in \mathbb{F}^{I \times I \times K}$ *is* SFS-rank-1 *if it equals the outer product of three nonzero vectors* $\mathbf{a} \in \mathbb{F}^I$, $\mathbf{a} \in \mathbb{F}^I$, *and* $\mathbf{c} \in \mathbb{F}^K$.

Definition 3.4.2. *A* SFS-PD *of a third-order SFS-tensor* $\mathcal{T} \in \mathbb{F}^{I \times I \times K}$ *expresses* $\mathcal{T}$ *as a sum of SFS-rank-*1 *terms:*

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{a}_r \circ \mathbf{c}_r, \tag{3.35}$$

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

Definition 3.4.3. *The* SFS-rank *of a SFS-tensor* $\mathcal{T} \in \mathbb{F}^{I\times I\times K}$ *is defined as the minimum number of SFS-rank-*1 *tensors in a PD of* $\mathcal{T}$ *and is denoted by* $r_{SFS,\mathcal{T}}$.

Definition 3.4.4. *A* SFS-CPD *of a third-order SFS-tensor* $\mathcal{T}$ *expresses* $\mathcal{T}$ *as a minimal sum of SFS-rank-*1 *terms.*

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

Now we can define uniqueness of the SFS-CPD.

Definition 3.4.5. *Let* $\mathcal{T}$ *be a SFS-tensor of SFS-rank* R. *The SFS-CPD of* $\mathcal{T}$ *is* unique *if* $\mathcal{T} = [\mathbf{A}, \mathbf{A}, \mathbf{C}]_R = [\bar{\mathbf{A}}, \bar{\mathbf{A}}, \bar{\mathbf{C}}]_R$ *implies that there exist an* $R \times R$ *permutation matrix* $\mathbf{\Pi}$ *and* $R \times R$ *nonsingular diagonal matrices* $\mathbf{\Lambda_A}$ *and* $\mathbf{\Lambda_C}$ *such that*

$$\bar{\mathbf{A}} = \mathbf{A\Pi\Lambda_A}, \quad \bar{\mathbf{C}} = \mathbf{C\Pi\Lambda_C}, \quad \mathbf{\Lambda_A^2\Lambda_C} = \mathbf{I}_R.$$

Example 3.4.6. *Some SFS-tensors admit both SFS-CPDs and CPDs of which the terms are not partially symmetric. For instance, consider the SFS-tensor* $\mathcal{T} \in \mathbb{R}^{I\times I\times K}$ *in which* $\mathbf{I}_I$ *is stacked* K *times. Let* $\mathbf{E}$ *denote the* $K \times I$ *matrix of which all entries are equal to one. Then* $\mathcal{T} = [\mathbf{X}, (\mathbf{X}^{-1})^T, \mathbf{E}]_I$, *is a CPD of* $\mathcal{T}$ *for any nonsingular* $I \times I$ *matrix* $\mathbf{X}$. *On the other hand,* $\mathcal{T} = [\mathbf{A}, \mathbf{A}, \mathbf{E}]_I$, *is a SFS-CPD of* $\mathcal{T}$ *for any orthogonal* $I \times I$ *matrix* $\mathbf{A}$.

The following result was obtained in [20]. We present the proof for completeness.

Lemma 3.4.7. *Let* $\mathcal{T}$ *be a SFS-tensor of rank* R *and let the CPD of* $\mathcal{T}$ *be unique. Then* $r_{SFS,\mathcal{T}} = r_{\mathcal{T}}$, *and the SFS-CPD of* $\mathcal{T}$ *is also unique.*

Proof. Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ be a CPD of the SFS-tensor $\mathcal{T}$. Because of the symmetry we also have $\mathcal{T} = [\mathbf{B}, \mathbf{A}, \mathbf{C}]_R$. Since the CPD of $\mathcal{T}$ is unique, there exist an $R\times R$ permutation matrix $\mathbf{\Pi}$ and $R \times R$ nonsingular diagonal matrices $\mathbf{\Lambda_A}$, $\mathbf{\Lambda_B}$, and $\mathbf{\Lambda_C}$ such that $\mathbf{B} = \mathbf{A\Pi\Lambda_A}$, $\mathbf{A} = \mathbf{B\Pi\Lambda_B}$, $\mathbf{C} = \mathbf{C\Pi\Lambda_C}$, and $\Lambda_\mathbf{A}\Lambda_\mathbf{B}\mathbf{\Lambda_C} = \mathbf{I}_R$. Since the CPD is unique, by (3.18), we have $k_\mathbf{C} \geq 2$. Hence, $\mathbf{\Pi} = \mathbf{\Lambda_C} = \mathbf{I}_R$ and $\mathbf{B} = \mathbf{A\Lambda_A}$. Thus, any CPD of $\mathcal{T}$ is in fact a SFS-CPD. Hence, $r_{SFS,\mathcal{T}} = r_{\mathcal{T}}$, and the SFS-CPD of $\mathcal{T}$ is unique. □

Remark 3.4.8. *To the authors' knowledge, it is still an open question whether there exist SFS-tensors with unique SFS-CPD but non-unique CPD.*

Lemma 3.4.7 implies that conditions guaranteeing uniqueness of SFS-CPD may be obtained from conditions guaranteeing uniqueness of CPD by just ignoring

the SFS-structure. To illustrate this, we present SFS-variants of Corollaries 3.1.25 and 3.1.28.

Proposition 3.4.9. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{A}, \mathbf{C}]_R$ and $m_{\mathbf{C}} := R - r_{\mathbf{C}} + 2$. Assume that*

(i) $k_{\mathbf{A}} + k_{\mathbf{C}} \geq R + 2$;

(ii) $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A})$ *has full column rank.*

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

Proof. From Corollary 3.1.25 it follows that $r_{\mathcal{T}} = R$ and that the CPD of tensor $\mathcal{T}$ is unique. The proof now follows from Lemma 3.4.7. □

Remark 3.4.10. *Under the additional assumption $r_{\mathbf{C}} = R$, Proposition 3.4.9 was proved in [15].*

Proposition 3.4.11. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{A}, \mathbf{C}]_R$ and $m_{\mathbf{A}} := R - r_{\mathbf{A}} + 2$. Assume that*

(i) $k_{\mathbf{A}} + \max(\min(k_{\mathbf{C}} - 1, k_{\mathbf{A}}), \min(k_{\mathbf{C}}, k_{\mathbf{A}} - 1)) \geq R + 1$;

(ii) $\mathcal{C}_{m_{\mathbf{A}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{A}}}(\mathbf{C})$ *has full column rank.*

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

Proof. By Lemma 3.4.7 it is sufficient to show that $r_{\mathcal{T}} = R$ and that the CPD of tensor $\mathcal{T}$ is unique. Both these statements follow from Corollary 3.1.25 applied to the tensor $[\mathbf{A}, \mathbf{C}, \mathbf{A}]_R$. □

Proposition 3.4.12. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{A}, \mathbf{C}]_R$, $m_{\mathbf{A}} = R - r_{\mathbf{A}} + 2$, and $m_{\mathbf{C}} = R - r_{\mathbf{C}} + 2$. Assume that the matrices*

$$\mathcal{C}_{m_{\mathbf{A}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{A}}}(\mathbf{C}), \tag{3.36}$$

$$\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}) \tag{3.37}$$

have full column rank. Then $r_{SFS,\mathcal{T}} = R$ and the SFS-CPD of tensor $\mathcal{T}$ is unique.

Proof. From Corollary 3.1.28 it follows that $r_{\mathcal{T}} = R$ and that the CPD of tensor $\mathcal{T}$ is unique. The proof now follows from Lemma 3.4.7. □

3.5 Uniqueness beyond (W_m)

In this section we discuss an example in which even condition (W_m) is not satisfied. Hence, CPD uniqueness does not follow from Proposition 3.1.13 or Proposition 3.1.14. A fortiori, it does not follow from Proposition 3.1.22, Corollaries 3.1.23–3.1.25, Proposition 3.1.26, and Corollaries 3.1.28–3.1.30. We show that uniqueness of the CPD can nevertheless be demonstrated by combining subresults.

In this section we will denote by $\omega(\mathbf{d})$ the number of nonzero components of $\mathbf{d}$ and we will write $\mathbf{a} \parallel \mathbf{b}$ if the vectors $\mathbf{a}$ and $\mathbf{b}$ are collinear, that is there exists a nonzero number $c \in \mathbb{F}$ such that $\mathbf{a} = c\mathbf{b}$.

For easy reference we include the following lemma concerning second compound matrices.

Lemma 3.5.1. *[7, Lemma 2.4 (1) and Lemma 2.5]*

(1) *Let the product* $\mathbf{XYZ}$ *be defined. Then the product* $\mathcal{C}_2(\mathbf{X})\mathcal{C}_2(\mathbf{Y})\mathcal{C}_2(\mathbf{Z})$ *is also defined and*
$$\mathcal{C}_2(\mathbf{XYZ}) = \mathcal{C}_2(\mathbf{X})\mathcal{C}_2(\mathbf{Y})\mathcal{C}_2(\mathbf{Z}).$$

(2) *Let* $\mathbf{d} = \begin{bmatrix} d_1 & d_2 & \dots & d_R \end{bmatrix} \in \mathbb{F}^R$. *Then* $\mathcal{C}_2(\mathrm{Diag}(\mathbf{d})) = \mathrm{Diag}(\widehat{\mathbf{d}}^2)$.
In particular, $\omega(\mathbf{d}) \leq 1$ *if and only if* $\widehat{\mathbf{d}}^2 = \mathbf{0}$ *if and only if* $\mathcal{C}_2(\mathrm{Diag}(\mathbf{d})) = \mathbf{0}$.

Example 3.5.2. *Let* $\mathcal{T}_\alpha = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_5$, *where*
$$\mathbf{A} = \begin{bmatrix} 0 & \alpha & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix},$$
and $\alpha \neq 0$. *Then* $r_{\mathbf{A}} = r_{\mathbf{B}} = r_{\mathbf{C}} = 4$, $k_{\mathbf{A}} = k_{\mathbf{B}} = k_{\mathbf{C}} = 2$, *and* $m := m_{\mathbf{A}} = m_{\mathbf{B}} = m_{\mathbf{C}} = 5 - 4 + 2 = 3$. *One can check that none of the triplets* $(\mathbf{A}, \mathbf{B}, \mathbf{C})$, $(\mathbf{B}, \mathbf{C}, \mathbf{A})$, $(\mathbf{C}, \mathbf{A}, \mathbf{B})$ *satisfies condition* (W_m). *Hence, the rank and the uniqueness of the factor matrices of* $\mathcal{T}_\alpha$ *do not follow from Proposition 3.1.13 or Proposition 3.1.14. We prove that* $r_{\mathcal{T}_\alpha} = 5$ *and that the CPD* $\mathcal{T}_\alpha = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_5$ *is unique.*

(i) *A trivial verification shows that*

$$\mathbf{A}\odot\mathbf{B},\quad \mathbf{B}\odot\mathbf{C},\quad \mathbf{C}\odot\mathbf{A} \qquad \textit{have full column rank,} \tag{3.38}$$

$$\mathcal{C}_2(\mathbf{A})\odot\mathcal{C}_2(\mathbf{B}),\quad \mathcal{C}_2(\mathbf{B})\odot\mathcal{C}_2(\mathbf{C}),\quad \mathcal{C}_2(\mathbf{C})\odot\mathcal{C}_2(\mathbf{A}) \quad \textit{have full column rank.} \tag{3.39}$$

Elementary algebra yields

$$\omega(\mathbf{A}^T\mathbf{x})=1 \Leftrightarrow \ \mathbf{x}\| \ \mathbf{e}_1^4 \ \ \textit{or} \ \ \mathbf{x}\parallel \mathbf{e}_4^4, \tag{3.40}$$

$$\omega(\mathbf{B}^T\mathbf{y})=1 \Leftrightarrow \ \mathbf{y}\| \ \mathbf{e}_1^4 \ \ \textit{or} \ \ \mathbf{y}\parallel \mathbf{e}_3^4, \tag{3.41}$$

$$\omega(\mathbf{C}^T\mathbf{z})=1 \Leftrightarrow \ \mathbf{z}\| \ \mathbf{e}_2^4 \ \ \textit{or} \ \ \mathbf{z}\parallel \mathbf{e}_3^4. \tag{3.42}$$

(ii) *Consider a CPD* $\mathcal{T}_\alpha=[\bar{\mathbf{A}},\bar{\mathbf{B}},\bar{\mathbf{C}}]_{\bar{R}}$, *i.e.* $\bar{R}=r_{\mathcal{T}_\alpha}$ *is minimal. We have* $\bar{R}\le 5$. *For later use we show that any three solutions of the equation* $\omega(\bar{\mathbf{C}}^T\mathbf{z})=1$ *are linearly dependent. Indeed, assume that there exist three vectors* $\mathbf{z}_1,\ \mathbf{z}_2,\ \mathbf{z}_3\in\mathbb{F}^4$ *such that* $\omega(\bar{\mathbf{C}}^T\mathbf{z}_1)=\omega(\bar{\mathbf{C}}^T\mathbf{z}_2)=\omega(\bar{\mathbf{C}}^T\mathbf{z}_3)=1$. *By* (3.2)–(3.3),

$$\mathbf{T}^{(1)}=(\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T=(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T, \tag{3.43}$$

$$\mathbf{T}^{(2)}=(\bar{\mathbf{B}}\odot\bar{\mathbf{C}})\bar{\mathbf{A}}^T=(\mathbf{B}\odot\mathbf{C})\mathbf{A}^T, \tag{3.44}$$

$$\mathbf{T}^{(3)}=(\bar{\mathbf{C}}\odot\bar{\mathbf{A}})\bar{\mathbf{B}}^T=(\mathbf{C}\odot\mathbf{A})\mathbf{B}^T. \tag{3.45}$$

From (3.43) *it follows that* $\mathbf{A}\mathrm{Diag}(\mathbf{C}^T\mathbf{z}_i)\mathbf{B}^T=\bar{\mathbf{A}}\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{z}_i)\bar{\mathbf{B}}^T$ *and hence, by Lemma 3.5.1 (1),*

$$\begin{aligned}&\mathcal{C}_2(\mathbf{A})\mathcal{C}_2(\mathrm{Diag}(\mathbf{C}^T\mathbf{z}_i))\mathcal{C}_2(\mathbf{B}^T)\\ &\qquad=\mathcal{C}_2(\bar{\mathbf{A}})\mathcal{C}_2(\mathrm{Diag}(\bar{\mathbf{C}}^T\mathbf{z}_i))\mathcal{C}_2(\bar{\mathbf{B}}^T)=\mathbf{O},\ i\in\{1,2,3\},\end{aligned}$$

which can also be expressed as

$$[\mathcal{C}_2(\mathbf{A})\odot\mathcal{C}_2(\mathbf{B})]\,\widehat{\mathbf{d}}_i^2=\mathbf{0},\qquad \mathbf{d}_i:=\mathbf{C}^T\mathbf{z}_i,\qquad i\in\{1,2,3\}.$$

By (3.39), $\widehat{\mathbf{d}}_i^2=\mathbf{0}$ *for* $i\in\{1,2,3\}$. *Since* $\mathbf{C}^T$ *has full column rank, Lemma 3.5.1 (2) implies that* $\omega(\mathbf{C}^T\mathbf{z}_1)=\omega(\mathbf{C}^T\mathbf{z}_2)=\omega(\mathbf{C}^T\mathbf{z}_3)=1$. *From* (3.42) *it follows that at least two of the vectors* $\mathbf{z}_1$, $\mathbf{z}_2$, *and* $\mathbf{z}_3$ *are collinear. Hence, the vectors* $\mathbf{z}_1$, $\mathbf{z}_2$, *and* $\mathbf{z}_3$ *are linearly dependent.*

(iii) *Since* $\mathbf{A}\odot\mathbf{B}$ *and* $\mathbf{C}^T$ *have full column rank, from* (3.43) *and Sylvester's rank inequality it follows that*

$$r_{\bar{\mathbf{C}}^T}\ge r_{(\bar{\mathbf{A}}\odot\bar{\mathbf{B}})\bar{\mathbf{C}}^T}=r_{(\mathbf{A}\odot\mathbf{B})\mathbf{C}^T}\ge r_{\mathbf{A}\odot\mathbf{B}}+r_{\mathbf{C}^T}-5=5+4-5=4.$$

In a similar fashion, from (3.44) *and* (3.45) *we obtain* $r_{\bar{\mathbf{A}}^T} \geq 4$ *and* $r_{\bar{\mathbf{B}}^T} \geq 4$, *respectively. We conclude that* $\bar{R} \geq r_{\bar{\mathbf{A}}} = r_{\bar{\mathbf{B}}} = r_{\bar{\mathbf{C}}} = 4$.

Let us show that $\bar{R} = 5$. *To obtain a contradiction, assume that* $\bar{R} = 4$. *In this case, since* $r_{\bar{\mathbf{C}}^T} = 4$, $\bar{\mathbf{C}}^T$ *is a nonsingular square matrix. Then the columns of* $\mathbf{Z} := (\bar{\mathbf{C}}^T)^{-1}$ *are linearly independent solutions of the equation* $\omega(\bar{\mathbf{C}}^T\mathbf{z}) = 1$, *which is a contradiction with* (ii). *Hence,* $\bar{R} = 5$.

(iv) *Let us show that* $k_{\bar{\mathbf{C}}} \geq 2$. *Conversely, assume that* $k_{\bar{\mathbf{C}}} = 1$. *Since* $r_{\bar{\mathbf{C}}} = 4$, *it follows that there exists exactly one pair of proportional columns of* $\bar{\mathbf{C}}$. *Without loss of generality we will assume that* $\bar{\mathbf{C}}_4 \parallel \bar{\mathbf{C}}_5$. *Hence,* $\mathbb{F}^4 = \text{range}(\bar{\mathbf{C}}) = \text{span}\{\bar{\mathbf{C}}_1, \bar{\mathbf{C}}_2, \bar{\mathbf{C}}_3, \bar{\mathbf{C}}_4\}$. *Let* $\begin{bmatrix}\mathbf{z}_1 & \mathbf{z}_2 & \mathbf{z}_3 & \mathbf{z}_4\end{bmatrix} := (\begin{bmatrix}\bar{\mathbf{C}}_1 & \bar{\mathbf{C}}_2 & \bar{\mathbf{C}}_3 & \bar{\mathbf{C}}_4\end{bmatrix}^T)^{-1}$. *Then* $\omega(\bar{\mathbf{C}}^T\mathbf{z}_1) = \omega(\bar{\mathbf{C}}^T\mathbf{z}_2) = \omega(\bar{\mathbf{C}}^T\mathbf{z}_3) = 1$, *which is a contradiction with* (ii).

In a similar fashion we can prove that $k_{\bar{\mathbf{A}}} \geq 2$ *and* $k_{\bar{\mathbf{B}}} \geq 2$. *Thus,* $\min(k_{\bar{\mathbf{A}}}, k_{\bar{\mathbf{B}}}, k_{\bar{\mathbf{C}}}) \geq 2$.

(v) *Assume that there exist indices* i, j, k, l *and nonzero values* t_1, t_2, t_3, t_4 *such that*

$$(\bar{\mathbf{A}}^T)_1 = t_1\mathbf{e}_i^5, \quad (\bar{\mathbf{A}}^T)_4 = t_2\mathbf{e}_j^5, \quad (\bar{\mathbf{B}}^T)_1 = t_3\mathbf{e}_k^5, \quad (\bar{\mathbf{B}}^T)_3 = t_4\mathbf{e}_l^5. \tag{3.46}$$

Here we show that (3.46) *implies the uniqueness of the CPD of* $\mathcal{T}_\alpha$ *and a fortiori the uniqueness of the third factor matrix. The latter implication will as such be instrumental in the proof of* (vi). *That assumption* (3.46) *really holds, and thus implies CPD uniqueness, will be demonstrated in* (vii).

Combination of (3.44), (3.45), *and* (3.46) *yields*

$$\alpha\mathbf{b}_2 \otimes \mathbf{c}_2 = t_1\bar{\mathbf{b}}_i \otimes \bar{\mathbf{c}}_i, \qquad \mathbf{c}_2 \otimes \mathbf{a}_2 = t_3\bar{\mathbf{c}}_k \otimes \bar{\mathbf{a}}_k,$$

$$\mathbf{b}_5 \otimes \mathbf{c}_5 = t_2\bar{\mathbf{b}}_j \otimes \bar{\mathbf{c}}_j, \qquad \mathbf{c}_4 \otimes \mathbf{a}_4 = t_4\bar{\mathbf{c}}_l \otimes \bar{\mathbf{a}}_l.$$

We see that $\mathbf{b}_2 \parallel \bar{\mathbf{b}}_i$. *Also,* $\mathbf{c}_2 \parallel \bar{\mathbf{c}}_i$ *and* $\mathbf{c}_2 \parallel \bar{\mathbf{c}}_k$. *Since* $k_{\bar{\mathbf{C}}} \geq 2$, *it follows that* $i = k$. *Therefore, also* $\mathbf{a}_2 \parallel \bar{\mathbf{a}}_i$. *It is now clear that* $[\mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2]_1 - [\bar{\mathbf{a}}_i, \bar{\mathbf{b}}_i, \bar{\mathbf{c}}_i]_1 = \beta[\mathbf{e}_1^4, \mathbf{e}_1^4, \mathbf{e}_1^4]$ *for some* $\beta \in \mathbb{F}$. *Let*

$$\mathcal{T}_\beta := \mathcal{T}_\alpha - [\mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2]_1 + \beta[\mathbf{e}_1^4, \mathbf{e}_1^4, \mathbf{e}_1^4] = [\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}]_5 - [\bar{\mathbf{a}}_i, \bar{\mathbf{b}}_i, \bar{\mathbf{c}}_i]_1.$$

Obviously, $\mathcal{T}_\beta$ *is rank-*4. *We claim that* $\beta = 0$. *Indeed, if* $\beta \neq 0$, *then repeating steps* (i)–(iii) *for* $\mathcal{T}_\alpha$ *replaced by* $\mathcal{T}_\beta$ *we obtain that* $\mathcal{T}_\beta$ *is rank-*5, *which is a contradiction. Hence,* $[\mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2]_1 = [\bar{\mathbf{a}}_i, \bar{\mathbf{b}}_i, \bar{\mathbf{c}}_i]_1$.

*What is left to show, is that the CPD of the rank-*4 *tensor* $\mathcal{T}_\alpha - [\mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2]_1$ *is unique. Note that the matrix* $\begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_3 & \mathbf{c}_4 & \mathbf{c}_5\end{bmatrix}$ *has full column rank. From* (3.39) *it follows that* $\mathcal{C}_2(\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5\end{bmatrix}) \odot \mathcal{C}_2(\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_3 & \mathbf{b}_4 & \mathbf{b}_5\end{bmatrix})$ *also has full column rank. Hence, by Proposition 3.1.15, the CPD of* $\mathcal{T}_\alpha - [\mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2]_1$ *is unique.*

(vi) *Let us show that* $k_{\bar{\mathbf{A}}} = k_{\bar{\mathbf{B}}} = k_{\bar{\mathbf{C}}} = 2$. *Conversely, assume that* $k_{\bar{\mathbf{C}}} \geq 3$. *Then* $r_{\bar{\mathbf{B}}} + k_{\bar{\mathbf{C}}} \geq 4 + 3 \geq R + 2$. *Recall from (iv) that* $k_{\bar{\mathbf{B}}} \geq 2$. *Hence, condition (K2) holds for* $\bar{\mathbf{B}}, \bar{\mathbf{C}}$. *By Proposition 3.1.12 (1),*

$$\mathcal{C}_2(\bar{\mathbf{B}}) \odot \mathcal{C}_2(\bar{\mathbf{C}}) \text{ has full column rank.} \tag{3.47}$$

Let $\mathbf{x} \in \mathbb{F}^4$. *From* (3.44) *it follows that*

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

Hence, by Lemma 3.5.1 (1),

$$\mathcal{C}_2(\mathbf{B})\mathcal{C}_2(\mathrm{Diag}(\mathbf{A}^T\mathbf{x}))\mathcal{C}_2(\mathbf{C}^T) = \mathcal{C}_2(\bar{\mathbf{B}})\mathcal{C}_2(\mathrm{Diag}(\bar{\mathbf{A}}^T\mathbf{x}))\mathcal{C}_2(\bar{\mathbf{C}}^T),$$

which can also be expressed as

$$\left[\mathcal{C}_2(\mathbf{B}) \odot \mathcal{C}_2(\mathbf{C})\right] \widehat{\mathbf{d}}^2_{\mathbf{A}} = \left[\mathcal{C}_2(\bar{\mathbf{B}}) \odot \mathcal{C}_2(\bar{\mathbf{C}})\right] \widehat{\mathbf{d}}^2_{\bar{\mathbf{A}}}, \tag{3.48}$$

where $\mathbf{d}_{\mathbf{A}} = \mathbf{A}^T\mathbf{x}$ *and* $\mathbf{d}_{\bar{\mathbf{A}}} = \bar{\mathbf{A}}^T\mathbf{x}$. *From* (3.39), (3.47), *and Lemma 3.5.1 (2) it follows that*

$$\begin{aligned} \omega(\mathbf{A}^T\mathbf{x}) = 1 &\overset{\text{Lemma 3.5.1 (2)}}{\Longleftrightarrow} \widehat{\mathbf{d}}^2_{\mathbf{A}} = \mathbf{0} \overset{(3.39),(3.47),(3.48)}{\Longleftrightarrow} \widehat{\mathbf{d}}^2_{\bar{\mathbf{A}}} = \mathbf{0} \\ &\overset{\text{Lemma 3.5.1 (2)}}{\Longleftrightarrow} \omega(\bar{\mathbf{A}}^T\mathbf{x}) = 1. \end{aligned} \tag{3.49}$$

In a similar fashion we can prove that for $\mathbf{y} \in \mathbb{F}^4$,

$$\omega(\mathbf{B}^T\mathbf{y}) = 1 \Leftrightarrow \omega(\bar{\mathbf{B}}^T\mathbf{y}) = 1. \tag{3.50}$$

Therefore, by (i), *there exist indices* i, j, k, l *and nonzero values* t_1, t_2, t_3, t_4 *such that* (3.46) *holds. It follows from step* (v) *that the matrices* $\mathbf{C}$ *and* $\bar{\mathbf{C}}$ *are the same up to permutation and column scaling. Hence,* $k_{\bar{\mathbf{C}}} = k_{\mathbf{C}} = 2$, *which is a contradiction with* $k_{\bar{\mathbf{C}}} \geq 3$. *We conclude that* $k_{\bar{\mathbf{C}}} < 3$. *On the other hand, we have from (iv) that* $k_{\bar{\mathbf{C}}} \geq 2$. *Hence,* $k_{\bar{\mathbf{C}}} = 2$.

In a similar fashion we can prove that $k_{\bar{\mathbf{A}}} = k_{\bar{\mathbf{B}}} = 2$.

(vii) *Since* $k_{\bar{\mathbf{A}}} = k_{\bar{\mathbf{B}}} = 2$, *both* $\bar{\mathbf{A}}$ *and* $\bar{\mathbf{B}}$ *have a rank-deficient* 4×3 *submatrix. Since* $r_{\bar{\mathbf{A}}} = r_{\bar{\mathbf{B}}} = 4$, *it follows that there exist vectors* $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{y}_1$, $\mathbf{y}_2$ *such that*

$$\omega(\bar{\mathbf{A}}^T\mathbf{x}_1) = \omega(\bar{\mathbf{A}}^T\mathbf{x}_2) = \omega(\bar{\mathbf{B}}^T\mathbf{y}_1) = \omega(\bar{\mathbf{B}}^T\mathbf{y}_2) = 1, \qquad \mathbf{x}_1 \nparallel \mathbf{x}_2, \qquad \mathbf{y}_1 \nparallel \mathbf{y}_2.$$

From (3.48)–(3.50) *it follows that* $\omega(\mathbf{A}^T\mathbf{x}_1) = \omega(\mathbf{A}^T\mathbf{x}_2) = \omega(\mathbf{B}^T\mathbf{y}_1) = \omega(\mathbf{B}^T\mathbf{y}_2) = 1$. *By* (3.40)–(3.41) *there exist indices* i, j, k, l *and nonzero values* t_1, t_2, t_3, t_4 *such that* (3.46) *holds. Hence, by* (v), *the CPD of* $\mathcal{T}_\alpha$ *is unique.*

3.6 Generic uniqueness

3.6.1 Generic uniqueness of unconstrained CPD

It was explained in [4, 16] that the conditions $r_{\mathbf{C}} = R$ and (C2) in Proposition 3.1.15 hold generically when they hold for one particular choice of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. It was indicated that this implies that the CPD of an $I \times J \times K$ tensor $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ is generically unique whenever $K \geq R$ and $C_I^2 C_J^2 \geq C_R^2$. These conditions guarantee that the matrix $\mathbf{C}$ generically has full column rank and that the number of columns of the $C_I^2 C_J^2 \times C_R^2$ matrix $\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ does not exceed its number of rows. In this subsection we draw conclusions for the generic case from the more general Proposition 3.1.14 and Corollary 3.1.25.

As in [4, 16], our proofs are based on the following lemma.

Lemma 3.6.1. *Let $f(\mathbf{x})$ be an analytic function of $\mathbf{x} \in \mathbb{F}^n$ and let μ_n be the Lebesgue measure on $\mathbb{F}^n$. If $\mu_n\{\mathbf{x} : \ f(\mathbf{x}) = 0\} > 0$, then $f \equiv 0$.*

Proof. The result easily follows from the uniqueness theorem for analytic functions (see, for instance, [11, Lemma 2, p. 1855]). □

The following corollary trivially follows from Lemma 3.6.1.

Corollary 3.6.2. *Let $f(\mathbf{x})$ be an analytic function of $\mathbf{x} \in \mathbb{F}^n$ and let μ_n be the Lebesgue measure on $\mathbb{F}^n$. Assume that there exists a point $\mathbf{x}_0$ such that $f(\mathbf{x}_0) \neq 0$. Then $\mu_n\{\mathbf{x} : \ f(\mathbf{x}) = 0\} = 0$.*

We will use the following matrix analogue of Corollary 3.6.2.

Lemma 3.6.3. *Let $\mathbf{F}(\mathbf{x}) = (f_{pq}(\mathbf{x}))_{p,q=1}^{P,Q}$, with $P \geq Q$, be an analytic matrix-valued function of $\mathbf{x} \in \mathbb{F}^n$ (that is, each entry $f_{pq}(\mathbf{x})$ is an analytic function of $\mathbf{x}$) and let μ_n be the Lebesgue measure on $\mathbb{F}^n$. Assume that there exists a point $\mathbf{x}_0$ such that $\mathbf{F}(\mathbf{x}_0)$ has full column rank. Then*

$$\mu_n\{\mathbf{x} : \ \mathbf{F}(\mathbf{x}) \text{ does not have full column rank}\} = 0.$$

Proof. Let $\mathbf{f}(\mathbf{x}) := \mathcal{C}_Q(\mathbf{F}(\mathbf{x}))$ and $L := C_P^Q$. Then $\mathbf{f} : \ \mathbb{F}^n \to \mathbb{F}^L : \mathbf{x} \to \mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) & \dots & f_L(\mathbf{x}) \end{bmatrix}^T$ is a vector-valued analytic function. Note that $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ if and only if the matrix $\mathbf{F}(\mathbf{x})$ does not have full column rank. Since $\mathbf{f}(\mathbf{x}_0) \neq \mathbf{0}$, there exists $l_0 \in \{1, \dots, L\}$ such that $f_{l_0}(\mathbf{x}) \neq 0$. Hence, by Corollary 3.6.2,

$\mu_n\{x:\ f_{l_0}(\mathbf{x})=0\}=0$. Therefore,

$$\mu_n\{\mathbf{x}:\ \mathbf{F}(\mathbf{x}) \text{ does not have full column rank}\}=\mu_n\{\mathbf{x}:\ \mathbf{f}(\mathbf{x})=\mathbf{0}\}$$

$$=\mu_n\left\{\bigcap_{l=1}^{L}\{x:\ f_l(\mathbf{x})=0\}\right\}\leq \mu_n\{x:\ f_{l_0}(\mathbf{x})=0\}=0. \qquad \square$$

The following lemma implies that, if $k_{\mathbf{C}}=r_{\mathbf{C}}$, then (3.12) in Proposition 3.1.20 holds generically, provided there exist matrices $\mathbf{A}_0\in\mathbb{F}^{I\times R}$ and $\mathbf{B}_0\in\mathbb{F}^{J\times R}$ for which $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}_0)\odot\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B}_0)$ has full column rank.

Lemma 3.6.4. *Suppose the matrices $\mathbf{A}_0\in\mathbb{F}^{I\times R}$, $\mathbf{B}_0\in\mathbb{F}^{J\times R}$, and $\mathbf{C}\in\mathbb{F}^{K\times R}$ satisfy the following conditions:*

$$k_{\mathbf{A}_0}=\min(I,R),\quad k_{\mathbf{B}_0}=\min(J,R),\quad k_{\mathbf{C}}=r_{\mathbf{C}}.$$

Suppose further the matrix $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}_0)\odot\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B}_0)$ has full column rank, where $m=R-r_{\mathbf{C}}+2$. Then

$$\max(\min(I,J-1,R-1),\min(I-1,J,R-1))+k_{\mathbf{C}}\geq R+1. \tag{3.51}$$

Proof. By Proposition 3.1.12 (2) and (9), $\min(k_{\mathbf{A}_0},k_{\mathbf{B}_0})\geq m_{\mathbf{C}}$. Hence,

$$\min(I,J,R)\geq\min(k_{\mathbf{A}_0},k_{\mathbf{B}_0})\geq m_{\mathbf{C}}=R-r_{\mathbf{C}}+2=R-k_{\mathbf{C}}+2.$$

Therefore,

$$\max(\min(I,J-1,R-1),\min(I-1,J,R-1))+k_{\mathbf{C}}$$

$$\geq\min(I-1,J-1)+k_{\mathbf{C}}\geq R-k_{\mathbf{C}}+2-1+k_{\mathbf{C}}=R+1.$$

Hence, (3.51) holds. $\square$

The following proposition is the main result of this section.

Proposition 3.6.5. *Let the matrix $\mathbf{C}\in\mathbb{F}^{K\times R}$ be fixed and suppose $k_{\mathbf{C}}\geq 1$. Assume that there exist matrices $\mathbf{A}_0\in\mathbb{F}^{I\times R}$ and $\mathbf{B}_0\in\mathbb{F}^{J\times R}$ such that $\mathcal{C}_m(\mathbf{A}_0)\odot\mathcal{C}_m(\mathbf{B}_0)$ has full column rank, where $m=R-r_{\mathbf{C}}+2$. Set $n=(I+J)R$. Then*

(i)

$$\mu_n\{(\mathbf{A},\mathbf{B}):\mathcal{T}:=[\mathbf{A},\mathbf{B},\mathbf{C}]_R \textit{ has rank less than } R \textit{ or}$$

$$\textit{the third factor matrix of } \mathcal{T} \textit{ is not unique}\}=0.$$

(ii) *If additionally,* $k_{\mathbf{C}} = r_{\mathbf{C}}$*, or* (3.51) *holds, then*

$$\mu_n\{(\mathbf{A},\mathbf{B}) :\ \mathcal{T} := [\mathbf{A},\mathbf{B},\mathbf{C}]_R \text{ has rank less than } R \text{ or } \text{the CPD of } \mathcal{T} \text{ is not unique}\} = 0. \tag{3.52}$$

Proof. (i) Let $P := C_I^m C_J^m$, $Q := C_R^m$, $n := (I+J)R$, $\mathbf{x} := (\mathbf{A},\mathbf{B})$, $\mathbf{x}_0 := (\mathbf{A}_0,\mathbf{B}_0)$, and $\mathbf{F}(\mathbf{x}) := \mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$. Since $k_{\mathbf{C}} \geq 1$, from Proposition 3.1.14 and Lemma 3.6.3 it follows that

$$\begin{aligned}&\mu_n\{(\mathbf{A},\mathbf{B}) : \mathcal{T} := [\mathbf{A},\mathbf{B},\mathbf{C}]_R \text{ has rank less than } R \text{ or}\\ &\qquad \text{the third factor matrix of } \mathcal{T} \text{ is not unique}\}\\ &\leq \mu_n\{(\mathbf{A},\mathbf{B}) : \mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}) \text{ does not have full column rank}\} = 0.\end{aligned}$$

(ii) By Lemma 3.6.4, we can assume that (3.51) holds. We obviously have

$$\mu_n\{(\mathbf{A},\mathbf{B}) :\ k_{\mathbf{A}} < \min(I,R) \ \text{ or } k_{\mathbf{B}} < \min(J,R)\} = 0.$$

Hence, by (3.51),

$$\mu_n\{(\mathbf{A},\mathbf{B}) :\ (3.12) \text{ does not hold}\} = 0.$$

From Proposition 3.1.20 and (i) it follows that

$$\begin{aligned}&\mu_n\{(\mathbf{A},\mathbf{B}) :\ \mathcal{T} := [\mathbf{A},\mathbf{B},\mathbf{C}]_R \text{ has rank less than } R \text{ or}\\ &\qquad \text{the CPD of } \mathcal{T} \text{ is not unique}\}\\ &\leq \mu_n\{(\mathbf{A},\mathbf{B}) :\ \mathcal{T} := [\mathbf{A},\mathbf{B},\mathbf{C}]_R \text{ has rank less than } R \text{ or}\\ &\qquad \text{the third factor matrix of } \mathcal{T} \text{ is not unique or}\\ &\qquad (3.12) \text{ does not hold}\} = 0. \qquad \square\end{aligned}$$

Proposition 3.6.6. *The CPD of an* $I \times J \times K$ *tensor of rank* R *is generically unique if there exist matrices* $\mathbf{A}_0 \in \mathbb{F}^{I\times R}$ *and* $\mathbf{B}_0 \in \mathbb{F}^{J\times R}$ *such that* $\mathcal{C}_m(\mathbf{A}_0) \odot \mathcal{C}_m(\mathbf{B}_0)$ *has full column rank, where* $m = R - \min(K,R) + 2$*.*

Proof. Generically we have $r_{\mathbf{C}} = \min(K,R)$. Let $N = (I+J+K)R$, $n = (I+J)R$, and let $\Omega = \{\mathbf{C} :\ k_{\mathbf{C}} < r_{\mathbf{C}}\} \subset \mathbb{F}^{KR}$. By application of Lemma 3.6.3, one obtains that $\mu_{KR}(\Omega) = 0$. From Proposition 3.6.5 it follows that (3.52) holds for $\mathbf{C} \notin \Omega$. Now

$$\begin{aligned}&\mu_N\{(\mathbf{A},\mathbf{B},\mathbf{C}) :\ \mathcal{T} := [\mathbf{A},\mathbf{B},\mathbf{C}]_R \text{ has rank less than } R \text{ or}\\ &\qquad \text{the CPD of } \mathcal{T} \text{ is not unique}\} = 0\end{aligned}$$

follows from Fubini's theorem [9, Theorem C, p. 148]. □

Proof of Proposition 3.1.31. Proposition 3.1.31 follows from Proposition 3.6.6 by permuting factors. □

3.6.2 Generic uniqueness of SFS-CPD

For generic uniqueness of the SFS-CPD we resort to the following definition.

Definition 3.6.7. *Let μ be the Lebesgue measure on $\mathbb{F}^{(2I+K)R}$. The SFS-CPD of an $I \times I \times K$ tensor of SFS-rank R is generically unique if*

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

We have the following counterpart of Proposition 3.1.31.

Proposition 3.6.8. *The SFS-CPD of an $I \times I \times K$ SFS-tensor of SFS-rank R is generically unique if there exist matrices $\mathbf{A}_0 \in \mathbb{F}^{I \times R}$ and $\mathbf{C}_0 \in \mathbb{F}^{K \times R}$ such that $\mathcal{C}_{m_\mathbf{C}}(\mathbf{A}_0) \odot \mathcal{C}_{m_\mathbf{C}}(\mathbf{A}_0)$ or $\mathcal{C}_{m_\mathbf{A}}(\mathbf{A}_0) \odot \mathcal{C}_{m_\mathbf{A}}(\mathbf{C}_0)$ has full column rank, where $m_\mathbf{C} = R - \min(K, R) + 2$ and $m_\mathbf{A} = R - \min(I, R) + 2$.*

Proof. The proof is obtained by combining Proposition 3.1.31 and Lemma 3.4.7. □

3.6.3 Examples

Example 3.6.9. *This example illustrates how one may adapt the approach in subsections 3.6.1 and 3.6.2 to particular types of structured factor matrices.*

Let $\mathcal{I}_4$ be the $4 \times 4 \times 4$ tensor with ones on the main diagonal and zero off-diagonal entries and let $\mathcal{T} = \mathcal{I}_4 + \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$ be a generic rank-1 perturbation of $\mathcal{I}_4$. Then $\mathcal{T} = [[\mathbf{I}_4\ \mathbf{a}], [\mathbf{I}_4\ \mathbf{b}], [\mathbf{I}_4\ \mathbf{c}]]_5$. Since the k-ranks of all factor matrices of $\mathcal{T}$ are equal to 4, it follows from Kruskal's Theorem 3.1.8 that $r_\mathcal{T} = 5$ and that the CPD of $\mathcal{T}$ is unique.

Let us now consider structured rank-1 perturbations $\bar{\mathbf{a}} \circ \bar{\mathbf{b}} \circ \bar{\mathbf{c}}$ that do not change the fourth vertical, third horizontal, and second frontal slice of $\mathcal{I}_4$. The vectors $\bar{\mathbf{a}}$, $\bar{\mathbf{b}}$, and $\bar{\mathbf{c}}$ admit the following parameterizations

$$\bar{\mathbf{a}} = \begin{bmatrix} a_1 & a_2 & a_3 & 0 \end{bmatrix}, \quad \bar{\mathbf{b}} = \begin{bmatrix} b_1 & b_2 & 0 & b_4 \end{bmatrix}, \quad \bar{\mathbf{c}} = \begin{bmatrix} c_1 & 0 & c_3 & c_4 \end{bmatrix},$$

with $a_i, b_j, c_k \in \mathbb{F}$.

Now the k-ranks of all factor matrices of $\bar{\mathcal{T}} := \mathcal{I}_4 + \bar{\mathbf{a}} \circ \bar{\mathbf{b}} \circ \bar{\mathbf{c}}$ *are equal to* 3*, and (generic) uniqueness of the CPD of* $\bar{\mathcal{T}}$ *does not follow from Kruskal's Theorem 3.1.8.*

We show that

$$\mu_9\{(\bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}) : \textit{the CPD of } \bar{\mathcal{T}} := \mathcal{I}_4 + \bar{\mathbf{a}} \circ \bar{\mathbf{b}} \circ \bar{\mathbf{c}} \textit{ is not unique or } r_{\bar{\mathcal{T}}} < 5\} = 0,$$

*that is, the CPD of rank-*1 *structured generic perturbation of* $\mathcal{I}_4$ *is again unique.*

Let the matrices $\mathbf{A}_0$*,* $\mathbf{B}_0$*, and* $\mathbf{C}_0$ *be given by the matrices* $\mathbf{A}$*,* $\mathbf{B}$*, and* $\mathbf{C}$*, respectively, in Example 3.3.2. As in Example 3.3.2 the matrix pairs* $(\mathbf{A}_0, \mathbf{B}_0)$*,* $(\mathbf{B}_0, \mathbf{C}_0)$*, and* $(\mathbf{C}_0, \mathbf{A}_0)$ *satisfy condition* (C3)*. Then, by Lemma 3.6.3,*

$$\mu_6\{(\bar{\mathbf{a}}, \bar{\mathbf{b}}) : \mathcal{C}_3(\mathbf{A}) \odot \mathcal{C}_3(\mathbf{B}) \textit{ does not have full column rank}\} = 0,$$

$$\mu_6\{(\bar{\mathbf{b}}, \bar{\mathbf{c}}) : \mathcal{C}_3(\mathbf{B}) \odot \mathcal{C}_3(\mathbf{C}) \textit{ does not have full column rank}\} = 0.$$

By Fubini's theorem [9, Theorem C, p. 148],

$$\begin{aligned}\mu_9\{(\bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}) :& \mathcal{C}_3(\mathbf{A}) \odot \mathcal{C}_3(\mathbf{B}) \textit{ has not full column rank or } k_{\mathbf{C}} < 3 \textit{ or } r_{\mathbf{C}} < 4\} \\ &= 0,\end{aligned}$$

$$\begin{aligned}\mu_9\{(\bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}) :& \mathcal{C}_3(\mathbf{B}) \odot \mathcal{C}_3(\mathbf{C}) \textit{ has not full column rank or } k_{\mathbf{A}} < 3 \textit{ or } r_{\mathbf{A}} < 4\} \\ &= 0.\end{aligned}$$

*Now generic uniqueness of the structured rank-*1 *perturbation of* $\mathcal{I}_4$ *follows from Proposition 3.1.21.*

Example 3.6.10. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *denote a PD of an* $I \times I \times (2I-1)$ *tensor, where* $I \geq 4$*. Generically,* $k_{\mathbf{A}} = k_{\mathbf{B}} = I$ *and* $k_{\mathbf{C}} = 2I - 1$*. Then Kruskal's condition* (3.4) *guarantees generic uniqueness for* $R \leq \lfloor \frac{I+I+2I-1-2}{2} \rfloor = 2I - 1$*.*

On the other hand, (3.11) guarantees generic uniqueness of the CPD under the conditions $R \leq 2I - 1$ *and* $C_R^2 \leq (C_I^2)^2$*. The maximum value of* R *that satisfies these bounds is shown in the column corresponding to* $m_{\mathbf{C}} = 2$ *in Table 3.3. The condition in Theorem 3.1.18 is even more relaxed.*

We now move to cases where $R > 2I - 1$*, where Theorem 3.1.18 no longer applies. By Proposition 3.6.6, the CPD of* $\mathcal{T}$ *of an* $I \times I \times (2I-1)$ *tensor of rank* R *is generically unique if there exist matrices* $\mathbf{A}_0 \in \mathbb{F}^{I \times R}$ *and* $\mathbf{B}_0 \in \mathbb{F}^{I \times R}$ *such that* $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}_0) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B}_0)$ *has full column rank, where* $m_{\mathbf{C}} = R - (2I-1) + 2 = R - 2I + 3$*. The proof of Proposition 3.6.6 shows that, if there exist* $\mathbf{A}_0$ *and* $\mathbf{B}_0$ *such that* $\mathcal{C}_{m_{\mathbf{C}}}(\mathbf{A}_0) \odot \mathcal{C}_{m_{\mathbf{C}}}(\mathbf{B}_0)$ *has full column rank, then actually*

$\mathcal{C}_{m_\mathbf{C}}(\mathbf{A}_0) \odot \mathcal{C}_{m_\mathbf{C}}(\mathbf{B}_0)$ has full column rank with probability one when $\mathbf{A}_0$ and $\mathbf{B}_0$ are drawn from continuous distributions. Hence, we generate random $\mathbf{A}_0$ and $\mathbf{B}_0$ and check up to which value of R the matrix $\mathcal{C}_{m_\mathbf{C}}(\mathbf{A}_0) \odot \mathcal{C}_{m_\mathbf{C}}(\mathbf{B}_0)$ has full column rank. Table 3.3 shows the results for $4 \leq I \leq 9$. For instance, we obtain that the CPD of a $9 \times 9 \times 17$ tensor of rank R is generically unique if $R \leq 20$. (By of comparison, Theorem 3.1.18 only guarantees uniqueness up to $R = 17$.)

Proposition 3.6.6 corresponds to condition (i) in Proposition 3.1.31. Note that, for $R \geq 2I - 1$, we generically have $m_\mathbf{A} = m_\mathbf{B} = R - I + 2 \geq I + 1$ such that the $m_\mathbf{B}$-th compound matrix of $\mathbf{A}$ and the $m_\mathbf{A}$-th compound matrix of $\mathbf{B}$ are not defined. Hence, we cannot resort to condition (ii) or (iii) in Proposition 3.1.31.

Table 3.3: Upper bounds on R under which generic uniqueness of the CPD of an $I \times I \times (2I - 1)$ tensor is guaranteed by Proposition 3.6.6.

dimensions of $\mathcal{T}$	$m = R - 2I + 3$			
$I \times I \times (2I-1)$	2	3	4	5
$4 \times 4 \times 7$	7			
$5 \times 5 \times 9$	9			
$6 \times 6 \times 11$	11	12		
$7 \times 7 \times 13$	13	14		
$8 \times 8 \times 15$	15	16	17	
$9 \times 9 \times 17$	17	18	19	20

Remark 3.6.11. *For $I = 3$ and $R = 2I - 1 = 5$, the CPD of an $I \times I \times (2I - 1)$ tensor $\mathcal{T}$ is not generically unique [19],[17]. This is the reason why in Table 3.3 we start from $I = 4$.*

Remark 3.6.12. *It was shown in [11, Corollary 1, p.1852] that the matrix $\mathcal{C}_1(\mathbf{A}) \odot \mathcal{C}_1(\mathbf{B}) = \mathbf{A} \odot \mathbf{B}$ has full column rank with probability one when the number of rows of $\mathbf{A} \odot \mathbf{B}$ does not exceed its number of columns. The same statement was made for the matrix $\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ in [4, 16]. However, the statement does not hold for compound matrices of arbitrary order. For instance, it does not hold for $\mathcal{C}_5(\mathbf{A}) \odot \mathcal{C}_5(\mathbf{B})$, where $\mathbf{A} \in \mathbb{F}^{6 \times 9}$ and $\mathbf{B} \in \mathbb{F}^{7 \times 9}$.*

Example 3.6.13. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_9$ denote a generic PD in 9 terms in the $6 \times 7 \times 6$ case. Then $m_\mathbf{A} = m_\mathbf{C} = 9 - 6 + 2 = 5$ and $m_\mathbf{B} = 9 - 7 + 2 = 4$. The matrices $\mathbf{M_C} := \mathcal{C}_5(\mathbf{A}) \odot \mathcal{C}_5(\mathbf{B})$ and $\mathbf{M_A} := \mathcal{C}_5(\mathbf{B}) \odot \mathcal{C}_5(\mathbf{C})$ have $C_6^5 C_7^5 = C_9^5 = 126$ rows and columns. Numerical experiments indicate that $\dim \ker(\mathbf{M_C}) = \dim \ker(\mathbf{M_A}) = 15$ with probability one. Hence, we cannot use Proposition 3.1.31* (i) *or* (ii) *for proving uniqueness of the CPD. On the other hand, the $C_6^4 C_6^4 \times C_9^4$ (225×126) matrix $\mathbf{M_B} := \mathcal{C}_4(\mathbf{C}) \odot \mathcal{C}_4(\mathbf{A})$ turns out to have full column rank for a random choice of $\mathbf{A}$ and $\mathbf{C}$. Hence, by Proposition 3.1.31* (iii)*, the CPD is generically unique.*

Example 3.6.14. *Here we consider $I \times I \times K$ tensors with $I \in \{4, \ldots, 9\}$ and $K \in \{2, \ldots, 33\}$, which is more general than Example 3.6.10.*

We check up to which value of R one of the conditions in Proposition 3.1.31 holds for a random choice of the factor matrices. Up to this value the CPD is generically unique. The results are shown as the left-most values in Table 3.4. We also check up to which value of R one of the conditions in Proposition 3.6.8 holds for a random choice of the factor matrices. Up to this value the SFS-CPD is generically unique. The results are shown as the middle values in Table 3.4.

The right-most values correspond to the maximum value of R for which generic uniqueness is guaranteed by Kruskal's Theorems 3.1.8–3.1.10, i.e., the largest value of R that satisfies $2\min(I, R) + \min(K, R) \geq 2R + 2$. Note that Kruskal's bound is the same for CPD and SFS-CPD. The bold values in the table correspond to the results that were not yet covered by Kruskal's Theorems 3.1.8–3.1.10 or Proposition 3.1.15 ($m = 2$).

Remark 3.6.15. *Most of the improved left-most values in Table 3.4 also follow from Theorems 3.1.16, 3.1.18–3.1.19. (Concerning the latter, if the CPD of an $I \times I \times I$ tensor of rank R is generically unique for $R \leq k(I)$, then a forteriori the CPD of a rank-R $I \times I \times K$ tensor with $K > I$ is generically unique for $R \leq k(I)$.) An important difference is that our bounds remain valid for many constrained CPDs. We briefly give two examples. Rather than going into details, let us suffice by mentioning that (generic) uniqueness in these examples may be defined and studied in the same way as it was done in Subsections 3.6.1 and 3.6.2 for unsymmetric CPD and SFS-CPD, respectively.*

1. *Let the third factor matrix of $I \times I \times K$ tensor $\mathcal{T}$ belong to a class of structured matrices Ω such that the condition $I + k_{\mathbf{C}} \geq R + 2$ is valid for generic $\mathbf{C} \in \Omega$. An example of a class for which this may be true, is the class of $K \times R$ Hankel matrices. In Subsection 3.6.1, Proposition 3.6.5 leads to Proposition 3.1.31 for unconstrained CPD. Similarly, Proposition 3.6.5 with condition* (3.51) *replaced by condition $I + k_{\mathbf{C}} \geq R + 2$ leads to an analogue of Proposition 3.1.31 that guarantees that a CPD with the third factor matrix belonging to Ω is generically unique for R bounded by the values in Table 3.4 (left values for unconstrained first and second factor matrices, and middle values in the case of partial symmetry).*

2. *Let us now assume that the third factor matrix is unstructured and that the first two matrices have Toeplitz structure. Random Toeplitz matrices also yield the values in Table 3.4. Hence, such a constrained CPD is again generically unique for R bounded by the values in Table 3.4.*

Remark 3.6.16. *In the case $r_{\mathbf{C}} = R$, both* (C2) *and* (U2) *are sufficient for overall CPD uniqueness, see* (3.9). *In the case of* (C2), *we generically have*

Table 3.4: Upper bounds on R under which generic uniqueness of the CPD (left and right value) and SFS-CPD (middle and right value) of an $I \times I \times K$ tensor is guaranteed by Proposition 3.1.31 (left), Proposition 3.6.8 (middle), and Kruskal's Theorems 3.1.8–3.1.10 (right). The values shown in bold correspond to the results that were not yet covered by Kruskal's Theorems 3.1.8–3.1.10 or Proposition 3.1.15 ($m = 2$).

		I					
		4	5	6	7	8	9
K	2	4, 4, 4	5, 5, 5	6, 6, 6	7, 7, 7	8, 8, 8	9, 9, 9
	3	4, 4, 4	5, 5, 5	6, 6, 6	7, 7, 7	8, 8, 8	9, 9, 9
	4	5, 5, 5	6, 6, 6	7, 7, 7	8, 8, 8	9, 9, 9	10, 10, 10
	5	5, 5, 5	6, 6, 6	7, 7, 7	8, 8, 8	**10**, **10**, 9	**11**, **11**, 10
	6	6, 6, 6	7, 7, 7	8, 8, 8	9, 9, 9	10, 10, 10	11, 11, 11
	7	7, 6, 6	**8**, 7, 7	**9**, 8, 8	9, 9, 9	**11**, **11**, 10	**12**, **12**, 11
	8	8, 6, 6	**9**, 8, 8	9, 9, 9	10, 10, 10	11, 11, 11	12, 12, 12
	9	9, 6, 6	9, 9, 8	**10**, **10**, 9	**11**, 10, 10	**12**, 11, 11	**13**, **13**, 12
	10	9, 6, 6	10, 10, 8	**11**, 10, 10	**12**, 11, 11	**13**, 12, 12	**14**, 13, 13
	11	9, 6, 6	11, 10, 8	**12**, 11, 10	**13**, **12**, 11	**14**, **13**, 12	**15**, **14**, 13
	12	9, 6, 6	12, 10, 8	**13**, 12, 10	**14**, **13**, 12	**15**, **14**, 13	**15**, **15**, 14
	13	9, 6, 6	13, 10, 8	**14**, 13, 10	**14**, **14**, 12	**15**, **15**, 13	**16**, **15**, 14
	14	9, 6, 6	14, 10, 8	14, 14, 10	**15**, **15**, 12	**16**, **15**, 14	**17**, **16**, 15
	15	9, 6, 6	14, 10, 8	15, 15, 10	**16**, 15, 12	**17**, **16**, 14	**18**, **17**, 15
	16	9, 6, 6	14, 10, 8	16, 15, 10	**17**, 16, 12	**18**, **17**, 14	**19**, **18**, 16
	17	9, 6, 6	14, 10, 8	17, 15, 10	**18**, 17, 12	**19**, **18**, 14	**20**, **19**, 16
	18	9, 6, 6	14, 10, 8	18, 15, 10	**19**, 18, 12	**20**, **19**, 14	**20**, **20**, 16
	19	9, 6, 6	14, 10, 8	19, 15, 10	**20**, 19, 12	**20**, **20**, 14	**21**, **20**, 16
	20	9, 6, 6	14, 10, 8	20, 15, 10	20, 20, 12	**21**, 20, 14	**22**, **21**, 16
	21	9, 6, 6	14, 10, 8	21, 15, 10	21, 20, 12	**22**, 21, 14	**23**, **22**, 16
	22	9, 6, 6	14, 10, 8	21, 15, 10	22, 20, 12	**23**, 22, 14	**24**, **23**, 16
	23	9, 6, 6	14, 10, 8	21, 15, 10	23, 20, 12	**24**, 23, 14	**25**, **24**, 16
	24	9, 6, 6	14, 10, 8	21, 15, 10	24, 20, 12	**25**, 24, 14	**26**, **25**, 16
	25	9, 6, 6	14, 10, 8	21, 15, 10	25, 20, 12	**26**, 25, 14	**26**, 25, 16
	26	9, 6, 6	14, 10, 8	21, 15, 10	26, 20, 12	**27**, 26, 14	**27**, 26, 16
	27	9, 6, 6	14, 10, 8	21, 15, 10	27, 20, 12	27, 26, 14	**28**, 27, 16
	28	9, 6, 6	14, 10, 8	21, 15, 10	28, 20, 12	28, 26, 14	**29**, 28, 16
	29	9, 6, 6	14, 10, 8	21, 15, 10	29, 20, 12	29, 26, 14	**30**, 29, 16
	30	9, 6, 6	14, 10, 8	21, 15, 10	30, 20, 12	30, 26, 14	**31**, 30, 16
	31	9, 6, 6	14, 10, 8	21, 15, 10	30, 20, 12	31, 26, 14	**32**, 31, 16
	32	9, 6, 6	14, 10, 8	21, 15, 10	30, 20, 12	32, 26, 14	**33**, 32, 16
	33	9, 6, 6	14, 10, 8	21, 15, 10	30, 20, 12	33, 26, 14	**34**, 33, 16

condition (3.11). *The more relaxed generic condition derived from* (U2) *is given in Theorem 3.1.18. For the case* $r_{\mathbf{C}} < R$ *we have obtained the deterministic result in Corollary 3.1.25 and its its generic version Proposition 3.1.31, both based on condition* (Cm). *This suggests that by starting from Corollary 3.1.23, based on* (Um), *more relaxed generic uniqueness results may be obtained.*

On the other hand, in Example 3.3.5 we have studied CPD of a rank-10 ($7 \times 7 \times 7$*) tensor. Simulations along the lines of Example 3.3.5 suggest that condition* (W5) *holds for random factor matrices, which then implies generic overall CPD uniqueness for* $R = 10$. *Starting from* (C5) *we have only demonstrated generic uniqueness up to* $R = 9$, *see the entry for* $I = K = 7$ *in Table 3.4. This suggests that by starting from Proposition 3.1.22, based on* (Wm), *further relaxed generic uniqueness results may be obtained.*

3.7 Conclusion

Using results obtained in Part I [7], we have obtained new conditions guaranteeing uniqueness of a CPD. In the framework of the new uniqueness theorems, Kruskal's theorem and the existing uniqueness theorems for the case $R = r_{\mathbf{C}}$ are special cases. We have derived both deterministic and generic conditions.

The results can be easily adapted to the case of PDs in which one or several factor matrices are equal, such as INDSCAL. In the deterministic conditions the equalities can simply be substituted. In the generic setting one checks the same rank constraints as in the unconstrained case for a random example. The difference is that there are fewer independent entries to draw randomly. This may decrease the value of R up to which uniqueness is guaranteed. However, the procedure for determining this maximal value is completely analogous. The same holds true for PDs in which one or several factor matrices have structure (Toeplitz, Hankel, Vandermonde, etc.).

3.8 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).

Bibliography

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

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

[3] P. Comon, X. Luciani, and A. L. F. de Almeida. Tensor decompositions, alternating least squares and other tales. *J. Chemometrics*, 23(7-8):393–405, 2009.

[4] L. De Lathauwer. A Link Between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization. *SIAM J. Matrix Anal. Appl.*, 28:642–666, August 2006.

[5] L. De Lathauwer. 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(4):1451–1474, 2011.

[6] L. De Lathauwer. 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*, pages 558–563, 2011.

[7] I. Domanov and L. De Lathauwer. On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors— Part I: Basic Results and Uniqueness of One Factor Matrix. *SIAM J. Matrix Anal. Appl.*, 34(3):855–875, 2013.

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

[9] P. R. Halmos. *Measure theory.* Springer-Verlag, New-York, 1974.

[10] T. Jiang and N. D. Sidiropoulos. Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints. *IEEE Trans. Signal Process.*, 52(9):2625–2636, September 2004.

[11] T. Jiang, N. D. Sidiropoulos, and J. M. F. Ten Berge. Almost-Sure Identifiability of Multidimensional Harmonic Retrieval. *IEEE Trans. Signal Process.*, 49(9):1849–1859, 2001.

[12] T. G. Kolda and B. W. Bader. Tensor Decompositions and Applications. *SIAM Review*, 51(3):455–500, September 2009.

[13] W. P. Krijnen. *The analysis of three-way arrays by constrained Parafac methods*. DSWO Press, Leiden, 1991.

[14] J. B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.*, 18(2):95–138, 1977.

[15] A. Stegeman. On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode. *Linear Algebra Appl.*, 431(1-2):211–227, 2009.

[16] A. Stegeman, J. Ten Berge, and L. De Lathauwer. Sufficient conditions for uniqueness in Candecomp/Parafac and Indscal with random component matrices. *Psychometrika*, 71(2):219–229, June 2006.

[17] A. Stegeman and J. M. F. Ten Berge. Kruskal's condition for uniqueness in Candecomp/Parafac when ranks and k-ranks coincide. *Comput. Stat. Data Anal.*, 50(1):210–220, 2006.

[18] V. Strassen. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52–53(0):645–685, 1983.

[19] J. Ten Berge. Partial uniqueness in CANDECOMP/PARAFAC. *J. Chemometrics*, 18(1):12–16, 2004.

[20] J. Ten Berge, N. D. Sidiropoulos, and R. Rocci. Typical rank and indscal dimensionality for symmetric three-way arrays of order $I \times 2 \times 2$ or $I \times 3 \times 3$. *Linear Algebra Appl.*, 388(0):363 – 377, 2004.

[21] L. Xiangqian and N. D. Sidiropoulos. Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays. *IEEE Trans. Signal Process.*, 49(9):2074–2086, sep 2001.

Chapter 4

Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition

This chapter is based on
Domanov, I., De Lathauwer, L. Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition, submitted to SIAM Journal on Matrix Analysis and Applications (accepted with major revision).

4.1 Introduction

4.1.1 Basic notations and terminology

Throughout the paper $\mathbb{R}$ denotes the field of real numbers and $\mathcal{T} = (t_{ijk}) \in \mathbb{R}^{I\times J\times K}$ denotes a third-order tensor with frontal slices $\mathbf{T}_1, \dots, \mathbf{T}_K \in \mathbb{R}^{I\times J}$; $r_{\mathbf{A}}$, range($\mathbf{A}$), and ker($\mathbf{A}$) denote the rank, the range, and the null space of a matrix $\mathbf{A}$, respectively; $k_{\mathbf{A}}$ (the k-rank of $\mathbf{A}$) is the largest number such that every subset of $k_{\mathbf{A}}$ columns of the matrix $\mathbf{A}$ is linearly independent; $\omega(\mathbf{d})$ denotes the

number of nonzero components of a vector $\mathbf{d}$; $\operatorname{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ denotes the linear span of the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_k$; $\mathbf{O}_{m\times n}$, $\mathbf{0}_m$, and $\mathbf{I}_n$ are the zero $m \times n$ matrix, the zero $m \times 1$ vector, and the $n \times n$ identity matrix, respectively; C_n^k denotes the binomial coefficient, $C_n^k = \frac{n!}{k!(n-k)!}$; $\mathcal{C}_m(\mathbf{A})$ (the m-th compound matrix of $\mathbf{A}$) is the matrix containing the determinants of all $m \times m$ submatrices of $\mathbf{A}$, arranged with the submatrix index sets in lexicographic order (see §4.2 for details).

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

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

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{4.1}$$

where $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$, $\mathbf{c}_r \in \mathbb{R}^K$, $1 \le r \le R$. If the number R of rank-1 terms in (4.1) is minimal, then (4.1) is called the *Canonical Polyadic Decomposition* (CPD) of $\mathcal{T}$ and R is called the rank of the tensor $\mathcal{T}$ (denoted by $r_{\mathcal{T}}$).

We write (4.1) as $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, where the matrices $\mathbf{A} := \begin{bmatrix}\mathbf{a}_1 & \ldots & \mathbf{a}_R\end{bmatrix} \in \mathbb{R}^{I\times R}$, $\mathbf{B} := \begin{bmatrix}\mathbf{b}_1 & \ldots & \mathbf{b}_R\end{bmatrix} \in \mathbb{R}^{J\times R}$ and $\mathbf{C} := \begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_R\end{bmatrix} \in \mathbb{R}^{K\times R}$ are called the *first*, *second* and *third factor matrix* of $\mathcal{T}$, respectively.

Obviously, $\mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$ has frontal slices $\mathbf{a}\mathbf{b}^T c_1, \ldots, \mathbf{a}\mathbf{b}^T c_K \in \mathbb{R}^{I\times J}$. Hence, (4.1) is equivalent to the system of matrix identities

$$\mathbf{T}_k = \sum_{r=1}^{R} \mathbf{a}_r \mathbf{b}_r^T c_{kr} = \mathbf{A}\operatorname{Diag}(\mathbf{c}^k)\mathbf{B}^T, \qquad 1 \le k \le K, \tag{4.2}$$

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

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

$$\operatorname{vec}(\mathbf{A}\operatorname{Diag}(\mathbf{d})\mathbf{B}^T) = (\mathbf{B} \odot \mathbf{A})\mathbf{d}, \qquad \mathbf{d} \in \mathbb{R}^R \tag{4.3}$$

it follows that

$$\operatorname{Matr}(\mathcal{T}) := \begin{bmatrix}(\mathbf{A} \odot \mathbf{B})\mathbf{c}^1 & \ldots & (\mathbf{A} \odot \mathbf{B})\mathbf{c}^K\end{bmatrix} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T, \tag{4.4}$$

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

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

and "$\otimes$" denotes the *Kronecker product*: $\mathbf{a} \otimes \mathbf{b} = [a_1 b_1 \dots a_1 b_J \ \dots \ a_I b_1 \dots a_I b_J]^T$.

It is clear that in (4.1) the rank-1 terms can be arbitrarily permuted and that vectors within the same rank-1 term can be arbitrarily scaled provided the overall rank-1 term remains the same. The CPD of a tensor *is unique* when it is only subject to these trivial indeterminacies.

4.1.2 Problem statement

The CPD was introduced by F. Hitchcock in [13] and was later referred to as *Canonical Decomposition* (Candecomp) [3], *Parallel Factor Model* (Parafac) [10, 12], and *Topographic Components Model* [25]. We refer to the overview papers [16, 4, 6], the books [17, 33] and the references therein for background and applications in Signal Processing, Data Analysis, Chemometrics, and Psychometrics.

Note that in applications one most often deals with a perturbed version of (4.1):

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

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

$$\min \|\widehat{\mathcal{T}} - [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R\|, \qquad \text{s.t.} \quad \mathbf{A} \in \mathbb{R}^{I \times R}, \ \mathbf{B} \in \mathbb{R}^{J \times R}, \ \mathbf{C} \in \mathbb{R}^{K \times R}, \quad (4.5)$$

where $\|\cdot\|$ denotes a suitable (usually Frobenius) norm [35].

In this paper we limit ourselves to the noiseless case. We show that under mild conditions on factor matrices the CPD is unique and can be found algebraically in the following sense: the CPD can be computed by using basic operations on matrices, by computing compound matrices, by taking the orthogonal complement of a subspace, and by computing generalized eigenvalue decomposition. We make connections with concepts like permanents, mixed discriminants, and compound matrices, which have so far received little attention in applied linear algebra but are of interest. Our presentation is in terms of real-valued tensors for notational convenience. Complex variants are easily obtained by taking into account complex conjugations.

The heart of the algebraic approach is the following straightforward connection between CPD of a two-slice tensor and Generalized Eigenvalue Decomposition

(GEVD) of a matrix pencil. Consider an $R \times R \times 2$ tensor $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, where $\mathbf{A}$ and $\mathbf{B}$ are nonsingular matrices and the matrix $\mathrm{Diag}(\mathbf{d}) := \mathrm{Diag}(\mathbf{c}^1)\mathrm{Diag}(\mathbf{c}^2)^{-1}$ is defined and has distinct diagonal entries. From the equations $\mathbf{T}_k = \mathbf{A}\mathrm{Diag}(\mathbf{c}^k)\mathbf{B}^T$, $k = 1, 2$ it follows easily that $\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{A}^{-1} = \mathbf{T}_1\mathbf{T}_2^{-1}$ and $\mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{B}^{-1} = (\mathbf{T}_2^{-1}\mathbf{T}_1)^T$. Hence, the matrix $\mathrm{Diag}(\mathbf{d})$ can be found (up to permutation of its diagonal entries) from the eigenvalue decomposition of $\mathbf{T}_1\mathbf{T}_2^{-1}$ or $(\mathbf{T}_2^{-1}\mathbf{T}_1)^T$ and the columns of $\mathbf{A}$ (resp. $\mathbf{B}$) are the eigenvectors of $\mathbf{T}_1\mathbf{T}_2^{-1}$ (resp. $(\mathbf{T}_2^{-1}\mathbf{T}_1)^T$) corresponding to the R distinct eigenvalues $d_1, \ldots, d_R$. Since the matrices $\mathbf{A}$ and $\mathbf{B}$ are nonsingular, the matrix $\mathbf{C}$ can be easily found from (4.4). More generally, when $\mathbf{A}$ and $\mathbf{B}$ have full column rank and $\mathbf{C}$ does not have collinear columns, $\mathbf{A}$ and $\mathbf{B}$ follow from the GEVD of the matrix pencil $(\mathbf{T}_1, \mathbf{T}_2)$.

4.1.3 Previous results on uniqueness and algebraic algorithms

We say that an $I \times R$ matrix *has full column rank* if its column rank is R, which implies $I \geq R$. The following theorem generalizes the result discussed at the end of the previous subsection. Several variants of this theorem have appeared in the literature [11, 39, 6, 20, 31, 30]. The proof is essentially obtained by picking two slices (or two mixtures of slices) from $\mathcal{T}$ and computing their GEVD.

Theorem 4.1.1. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and suppose that $\mathbf{A}$ and $\mathbf{B}$ have full column rank and that $k_{\mathbf{C}} \geq 2$. Then*

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

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

In Theorem 4.1.1 the third factor matrix plays a different role than the first and the second factor matrices. Obviously, the theorem still holds when $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are permuted. In the sequel we will present only one version of results. Taking this into account, we may say that the following result is stronger than Theorem 4.1.1.

Theorem 4.1.2. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $r_{\mathbf{C}} = R$, and suppose that $\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ has full column rank. Then*

(i) *$r_{\mathcal{T}} = R$ and the CPD of $\mathcal{T}$ is unique [5, 15];*

(ii) *the CPD of $\mathcal{T}$ can be found algebraically [5].*

Computationally, we may obtain from $\mathcal{T}$ a partially symmetric tensor $\mathcal{W}$ that has CPD $\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T}, \mathbf{M}]_R$ in which both $\mathbf{C}^{-T}$ and $\mathbf{M}$ have full column

rank and work as in Theorem 4.1.1 to obtain $\mathbf{C}^{-T}$. The matrices $\mathbf{A}$ and $\mathbf{B}$ are subsequently easily obtained from (4.4).

Also, some algorithms for symmetric CPD have been obtained in the context of algebraic geometry. We refer to [29, 19] and references therein. Further, algebraic algorithms have been obtained for CPDs in which factor matrices are subject to constraints (such as orthogonality and Vandermonde) [38, 36].

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

In the remaining part of this subsection we present some results on the uniqueness of the CPD. These results will guarantee CPD uniqueness under the conditions for which we will derive algebraic algorithms. For more general results on uniqueness we refer to [7, 8]. The following result was obtained by J. Kruskal, which is little known. We present the compact version from [8]. Corollary 4.1.4 presents what is widely known as "Kruskal's condition" for CPD uniqueness.

Theorem 4.1.3. *[18, Theorem 4b, p. 123],[8, Corollary 1.29] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Suppose that*

$$k_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} \geq 2R + 2 \quad \textit{and} \quad \min(r_{\mathbf{C}} + k_{\mathbf{B}}, k_{\mathbf{C}} + r_{\mathbf{B}}) \geq R + 2. \tag{4.6}$$

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

Corollary 4.1.4. *[18, Theorem 4a, p. 123] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2. \tag{4.7}$$

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

In [7, 8] the authors obtained new sufficient conditions expressed in terms of compound matrices. We will use the following result.

Theorem 4.1.5. *[8, Corollary 1.25] Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and* $m := R - r_{\mathbf{C}} + 2$. *Suppose that*

$$\max(\min(k_{\mathbf{A}}, k_{\mathbf{B}} - 1),\ \min(k_{\mathbf{A}} - 1, k_{\mathbf{B}})) + k_{\mathbf{C}} \geq R + 1, \tag{4.8}$$

$$\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}) \textit{ has full column rank.} \tag{4.9}$$

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

Since the k-rank of a matrix cannot exceed its rank (and a fortiori not its number of columns), condition (4.7) immediately implies conditions (4.6) and

(4.8). It was shown in [8] that (4.6) implies (4.9) for $m = R - r_{\mathbf{C}} + 2$. Thus, Theorem 4.1.5 guarantees the uniqueness of the CPD under milder conditions than Theorem 4.1.3. Note also that statement (i) of Theorem 4.1.2 is the special case of Theorem 4.1.5 obtained for $r_{\mathbf{C}} = R$, i.e., when one of the factor matrices has full column rank.

4.1.4 New results and organization

To simplify the presentation and without loss of generality we will assume throughout the paper that the third dimension of the tensor $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ coincides with $r_{\mathbf{C}}$ (this can always be achieved in a "dimensionality reduction" step: if the columns of a matrix $\mathbf{V}$ form an orthonormal basis of range($\mathbf{C}$), then $r_{\mathbf{C}} = r_{\mathbf{V}^T\mathbf{C}}$, and by (4.4), the matrix $\mathrm{Matr}(\mathcal{T})\mathbf{V} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{V}$ has $r_{\mathbf{C}}$ columns, which means that the third dimension of the tensor $\mathcal{T}_{\mathbf{V}} := \mathrm{Tens}(\mathrm{Matr}(\mathcal{T})\mathbf{V}, I, J)$ is equal to $r_{\mathbf{C}}$; if the CPD $\mathcal{T}_{\mathbf{V}} = [\mathbf{A}, \mathbf{B}, \mathbf{V}^T\mathbf{C}]_R$ has been computed, then the matrix $\mathbf{C}$ can be recovered as $\mathbf{C} = \mathbf{V}(\mathbf{V}^T\mathbf{C})$).

The following theorems are the main results of the paper. In all cases we will reduce the computation to the situation as in Theorem 4.1.1.

Theorem 4.1.6. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $m := R - r_{\mathbf{C}} + 2$. Suppose that $k_{\mathbf{C}} = r_{\mathbf{C}}$ and that* (4.9) *holds. Then*

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

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

Theorem 4.1.7 generalizes Theorem 4.1.6 to case where possibly $k_{\mathbf{C}} < r_{\mathbf{C}}$. The more general situation for $\mathbf{C}$ is accomodated by tightening the condition on $\mathbf{A}$ and $\mathbf{B}$. (Indeed, (4.10) is more restricitive than (4.9) when $n > m$.) The proof of Theorem 4.1.7 is simple; we essentially consider a $k_{\mathbf{C}}$-slice subtensor $\bar{\mathcal{T}} = [\mathbf{A}, \mathbf{B}, \bar{\mathbf{C}}]_R$ for which $k_{\bar{\mathbf{C}}} = r_{\bar{\mathbf{C}}}$, so that Theorem 4.1.6 applies. (Actually, to guarantee that $k_{\bar{\mathbf{C}}} = r_{\bar{\mathbf{C}}}$, we will consider a random slice-mixture.)

Theorem 4.1.7. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $n := R - k_{\mathbf{C}} + 2$. Suppose that*

$$\mathcal{C}_n(\mathbf{A}) \odot \mathcal{C}_n(\mathbf{B}) \text{ has full column rank.} \tag{4.10}$$

Then

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

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

We also obtain the following corollaries.

Corollary 4.1.8. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. Suppose that*

$$k_{\mathbf{A}} + r_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \quad \textit{and} \quad k_{\mathbf{B}} + k_{\mathbf{C}} \geq R + 2. \tag{4.11}$$

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

Corollary 4.1.9. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ and let $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2$. Then the CPD of $\mathcal{T}$ is unique and can be found algebraically.*

Let us further explain how the theorems that we have formulated so far, relate to one another. First, we have obviously that $n = R - k_{\mathbf{C}} + 2 \geq R - r_{\mathbf{C}} + 2 = m$. Next the following implications were proved in [7]:

$$\begin{array}{ccccc} (4.11) & \Longrightarrow & (4.10) & \Longrightarrow & \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq n \overset{\text{trivial}}{\Longrightarrow} (4.8) \\ \Big\Downarrow{\scriptstyle \text{trivial}} & & \Big\Downarrow & & \Big\Downarrow \quad \Big\Uparrow{\scriptstyle \text{if } k_{\mathbf{C}} = r_{\mathbf{C}} \text{ (trivial)}} \\ (4.6) & \Longrightarrow & (4.9) & \Longrightarrow & \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m \end{array} \tag{4.12}$$

The first thing that follows from scheme (4.12) is that Theorem 4.1.7 is indeed more general than Corollary 4.1.8. Corollary 4.1.9 follows trivially from Corollary 4.1.8. Next, it appears that the conditions of Theorems 4.1.6–4.1.7 are more restrictive than the conditions of Theorem 4.1.5. Also, the conditions of Corollary 4.1.8 are more restrictive than the conditions of Theorem 4.1.3. Hence, we immediately obtain the uniqueness of the CPD in Theorems 4.1.6–4.1.7 and Corollary 4.1.8. Consequently, we can limit ourselves to the derivation of the algebraic algorithms.

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

$$\mathcal{B}(\mathbf{C}) := \mathbf{L}\mathcal{C}_{K-1}(\mathbf{C}), \tag{4.13}$$

where

$$\mathbf{L} := \begin{bmatrix} 0 & 0 & \dots & (-1)^{K-1} \\ \vdots & \vdots & \iddots & \vdots \\ 0 & -1 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{bmatrix}. \tag{4.14}$$

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

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

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

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

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

Recall that every column of the classical Moore-Penrose pseudo-inverse $\mathbf{C}^\dagger \in \mathbb{R}^{R\times K}$ is orthogonal to exactly $K-1$ *rows* of $\mathbf{C}$ and vice-versa. The equality $\mathbf{C}\mathbf{C}^\dagger = \mathbf{I}_K$ works along the "long" dimension of $\mathbf{C}$. If $\mathbf{C}^\dagger$ is known, then $\mathbf{C}$ may easily be found by pseudo-inverting again, $\mathbf{C} = (\mathbf{C}^\dagger)^\dagger$. The interaction with $\mathcal{B}(\mathbf{C})$ takes place along the "short" dimension of $\mathbf{C}$, and this complicates things. Nevertheless, it is also possible to reconstruct $\mathbf{C}$ from $\mathcal{B}(\mathbf{C})$. The definition of $\mathcal{B}(\mathbf{C})$, its properties and the reconstruction of $\mathbf{C}$ from $\mathcal{B}(\mathbf{C})$ are discussed in Subsection 4.2.1.

In the second and third phase of our algorithms we use $\mathcal{B}(\mathbf{C})$ to compute CPD. The following two properties of $\mathcal{B}(\mathbf{C})$ will be crucial for our derivation.

Proposition 4.1.10. *Let $\mathbf{C} \in \mathbb{R}^{K\times R}$ and $k_{\mathbf{C}} = K$. Then*

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

(ii) *the matrices*

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

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

The rest of the paper is organized as follows. Sections 4.2–4.3 contain auxiliary results. In Subsection 4.2.1 we recall the properties of compound matrices and provide an intuitive understanding of properties (P1)–(P4) and Propositions 4.1.10 (since the proofs of properties (P1)-(P4) and Proposition 4.1.10 are rather long and technical, they are included in the supplementary materials). In Subsections 4.2.2–4.2.3 we study variants of permanental compound matrices.

Let the columns of the K^m-by-C_R^m matrix $\mathcal{R}_m(\mathbf{C})$ be equal to the vectorized symmetric parts of the tensors $\mathbf{c}_{i_1} \circ \cdots \circ \mathbf{c}_{i_m}$, $1 \leq i_1 < \cdots < i_m \leq R$. We prove

$$\text{Proposition 4.2.13 (iii) : } \ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\text{range}(\pi_S)}\right) = \text{range}(\mathcal{B}(\mathbf{C})^{(m)}), \quad (4.15)$$

where the notation $\mathcal{R}_m(\mathbf{C})^T \restriction_{\text{range}(\pi_S)}$ means that we let the matrix $\mathcal{R}_m(\mathbf{C})^T$ act only on $K^m \times 1$ vectorized versions of $K \times \cdots \times K$ symmetric tensors.

In §4.3 we introduce polarized compound matrices — a notion closely related to the rank detection mappings in [5, 28]. The entries of polarized compound matrices are mixed discriminants [21, 2, 1]. Using polarized compound matrices we construct a $C_I^m C_J^m \times K^m$ matrix $\mathbf{R}_m(\mathcal{T})$ from the given tensor $\mathcal{T}$ such that

$$\mathbf{R}_m(\mathcal{T}) = [\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})]\, \mathcal{R}_m(\mathbf{C})^T. \quad (4.16)$$

Assuming that $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank and combining (4.15) with (4.16) we find the space generated by the columns of the matrix $\mathcal{B}(\mathbf{C})^{(m)}$:

$$\ker\left(\mathbf{R}_m(\mathcal{T}) \restriction_{\text{range}(\pi_S)}\right) = \ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\text{range}(\pi_S)}\right) = \text{range}(\mathcal{B}(\mathbf{C})^{(m)}). \quad (4.17)$$

In §4.4 we combine all results to obtain Theorems 4.1.6–4.1.7 and we present two algebraic CPD algorithms. Both algorithms rely on the key formula (4.17) which makes a link between the known matrix $\mathbf{R}_m(\mathcal{T})$ constructed from $\mathcal{T}$ and unknown matrix $\mathcal{B}(\mathbf{C})$.

The overall derivation generalizes ideas from [5] ($K = R$). Before discussing the new algorithms, let us first recall the CPD algorithm from [5] using our notations. We have $K = R$, which implies $m = 2$. By (P3)–(P4), the columns of $\mathcal{B}(\mathbf{C})$ are proportional to the columns of $\mathbf{C}^{-T}$, i.e., $\mathcal{B}(\mathbf{C})^T$ is equal to the inverse of $\mathbf{C}$ up to column permutation and scaling. Let $\mathbf{w}_1, \ldots, \mathbf{w}_R \in \mathbb{R}^{R^2}$ be a basis of $\ker\left(\mathbf{R}_2(\mathcal{T}) \restriction_{\text{range}(\pi_S)}\right)$. By (4.17), $\text{range}(\mathbf{W}) = \text{range}(\mathbf{C}^{-T} \odot \mathbf{C}^{-T})$. Hence, there exists a nonsingular matrix $\mathbf{M}$ such that $\mathbf{W} = [\mathbf{w}_1 \ \ldots \ \mathbf{w}_R] = \left(\mathbf{C}^{-T} \odot \mathbf{C}^{-T}\right)\mathbf{M}^T$. Therefore, by (4.4), $\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T}, \mathbf{M}]_R$, where $\mathcal{W}$ denotes the $R \times R \times R$ tensor such that $\mathbf{W} = \text{Matr}(\mathcal{W})$. Since all factor matrices of $\mathcal{W}$ have full column rank, the CPD of $\mathcal{W}$ can be computed algebraically. Thus, we can find $\mathbf{C}^{-T}$ (and hence, $\mathbf{C}$) up to column permutation and scaling. The matrices $\mathbf{A}$ and $\mathbf{B}$ can now be easily found from $\text{Matr}(\mathcal{T})\mathbf{C}^{-T} := \mathbf{A} \odot \mathbf{B}$ using the fact that the columns of $\mathbf{A} \odot \mathbf{B}$ are vectorized rank-1 matrices.

Both new algorithms contain the same first phase in which we find a matrix $\mathbf{F}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. This is done as follows. We construct the matrix $\mathbf{R}_m(\mathcal{T})$ and construct the $K \times K^{m-1} \times C_R^{K-1}$ tensor $\mathcal{W}$ such that the columns of $\text{Matr}(\mathcal{W}) \in \mathbb{R}^{K^m \times C_R^{K-1}}$ form a basis of $\ker\left(\mathbf{R}_m(\mathcal{T}) \restriction_{\text{range}(\pi_S)}\right)$. From Proposition 4.1.10 and Theorem 4.1.1 it follows

that the CPD $\mathcal{W} = [\mathcal{B}(\mathbf{C}), \mathcal{B}(\mathbf{C})^{(m-1)}, \mathbf{M}]_{C_R^{K-1}}$ can be found algebraically. This allows us to find a matrix $\mathbf{F}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. In the second and third phase of the first algorithm we find the matrix $\mathbf{C}$ and the matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. For finding $\mathbf{C}$, we resort to properties (P3)–(P4). Full exploitation of the structure has combinatorial complexity and is infeasible unless the dimensions of the tensor are relatively small. As an alternative, in the second algorithm we first find the matrices $\mathbf{A}$ and $\mathbf{B}$ and then we find the matrix $\mathbf{C}$. This is done as follows. We construct the new $I \times J \times C_R^{K-1}$ tensor $\mathcal{V}$ with the matrix unfolding $\mathrm{Matr}(\mathcal{V}) := \mathrm{Matr}(\mathcal{T})\mathbf{F} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{F}$. We find subtensors of $\mathcal{V}$ such that each subtensor has dimensions $I \times J \times 2$ and its CPD can be found algebraically. Full exploitation of the structure yields $C_R^m C_m^2$ subtensors. From the CPD of the subtensors we simultaneously obtain the columns of $\mathbf{A}$ and $\mathbf{B}$, and finally we set $\mathbf{C} = \left((\mathbf{A} \odot \mathbf{B})^\dagger \mathrm{Matr}(\mathcal{T})\right)^T$.

We conclude the paper with two examples. In the first example we demonstrate how the algorithms work for a $4 \times 4 \times 4$ tensor of rank 5 for which $k_{\mathbf{A}} = k_{\mathbf{B}} = 3$. In the second example we consider a generic $6 \times 6 \times 7$ tensor of rank 9 and compare the complexity of algorithms. Note that in both cases the uniqueness of the CPDs do not follow from Kruskal's Theorem 4.1.3.

4.2 Matrices formed by determinants and permanents of submatrices of a given matrix

Throughout the paper we will use the following multi-index notations. Let $i_1, \ldots, i_k$ be integers. Then $\{i_1, \ldots, i_k\}$ denotes the set with elements $i_1, \ldots, i_k$ (the order does not matter) and $(i_1, \ldots, i_k)$ denotes a k-tuple (the order is important). Let

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

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

Q_n^k or R_n^k). For example,

$$S_2^2 = \{(1,2)\}, \quad Q_2^2 = \{(1,1),(1,2),(2,2)\}, \quad R_2^2 = \{(1,1),(1,2),(2,1),(2,2)\},$$

$$S_2^2(1) = Q_2^2(2) = R_2^2(2), \quad Q_2^2(3) = R_2^2(4).$$

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

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

4.2.1 Matrices whose entries are determinants

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

Definition 4.2.1. *[14] Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $k \le \min(m,n)$. The C_m^k-by-C_n^k matrix whose (i,j)-th entry is $\det \mathbf{A}(S_m^k(i), S_n^k(j))$ is called the k-th compound matrix of $\mathbf{A}$ and is denoted by $\mathcal{C}_k(\mathbf{A})$.*

Example 4.2.2. *Let $\mathbf{A} = [\mathbf{I}_3\ \mathbf{a}]$, where $\mathbf{a} = [a_1\ a_2\ a_3]^T$. Then*

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

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

Definition 4.2.1 immediately implies the following lemma.

Lemma 4.2.3. *Let* $\mathbf{A} \in \mathbb{R}^{I\times R}$ *and* $k \leq \min(I, R)$. *Then*

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

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

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

PD representation (4.2) will make us need compound matrices of diagonal matrices.

Lemma 4.2.4. *Let* $\mathbf{d} \in \mathbb{R}^R$, $k \leq R$, *and let* $\widehat{\mathbf{d}}^k := [d_1 \cdots d_k \quad d_1 \cdots d_{k-1}d_{k+1} \quad \ldots \quad d_{R-k+1} \cdots d_R]^T \in \mathbb{R}^{C_R^k}$. *Then*

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

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

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

The following result is known as Binet-Cauchy formula.

Lemma 4.2.5. *[14, p. 19–22] Let* k *be a positive integer and let* $\mathbf{A}$ *and* $\mathbf{B}$ *be matrices such that* $\mathcal{C}_k(\mathbf{A})$ *and* $\mathcal{C}_k(\mathbf{B})$, *are defined. Then* $\mathcal{C}_k(\mathbf{A}\mathbf{B}^T) = \mathcal{C}_k(\mathbf{A})\mathcal{C}_k(\mathbf{B}^T)$. *If additionally* $\mathbf{d}$ *is a vector such that* $\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T$ *is defined, then* $\mathcal{C}_k(\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{B}^T) = \mathcal{C}_k(\mathbf{A})\mathrm{Diag}(\widehat{\mathbf{d}}^k)\mathcal{C}_k(\mathbf{B})^T$.

The goal of the remaining part of this subsection is to provide an intuitive understanding of properties (P1)–(P4) and Proposition 4.1.10.

Let $K \geq 2$, and let $\mathbf{C}$ be a $K \times K$ nonsingular matrix. By Cramer's rule and (4.13), the matrices $\det(\mathbf{C})\mathbf{C}^{-1}$ and $\mathcal{B}(\mathbf{C})$ are formed by $(K-1)\times(K-1)$ minors (also known as cofactors) of $\mathbf{C}$. It is easy to show that $\mathcal{B}(\mathbf{C}) = (\det(\mathbf{C})\mathbf{C}^{-1})^T\mathbf{L}$, where $\mathbf{L}$ is given by (4.14). It now trivially follows that every column of $\mathcal{B}(\mathbf{C})$ is a nonzero vector orthogonal to exactly $K-1$ columns of $\mathbf{C}$. Indeed,

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

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

$$\omega(\mathbf{x}^T\mathcal{B}(\mathbf{C})) = \omega(\mathbf{x}^T \det(\mathbf{C})\mathbf{C}^{-T}\mathbf{L}) = \omega(\mathbf{x}^T\mathbf{C}^{-T}) = \omega(\mathbf{C}^{-1}\mathbf{x}) \leq 1 \Leftrightarrow \quad \mathbf{x} \text{ is proportional to a column of } \mathbf{C}. \tag{4.18}$$

Properties (P3)-(P4) generalize (4.18) for rectangular matrices and imply that, if we know $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling, then we know $\mathbf{C}$ up to column permutation and scaling. This result will be directly used in Algorithm 1 further: we will first estimate $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling and then obtain $\mathbf{C}$ up to column permutation and scaling. Statements (P1)–(P3) are easy to show. Statement (P4) is more difficult. Since the proofs are technical, they are given in the supplementary materials.

Let us illustrate properties (P1)–(P4) and Proposition 4.1.10 for a rectangular matrix $\mathbf{C}$ ($K < R$).

Example 4.2.6. *Let*

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

implying $k_{\mathbf{C}} = K = 3$ *and* $R = 4$. *From* (4.13) *and Example 4.2.2 it follows that*

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

One can easily check the statements of properties (P1)–(P4) *and Proposition 4.1.10. Note in particular that exactly* 4 *sets of* 3 *columns of* $\mathcal{B}(\mathbf{C})$ *are linearly dependent. The vectors that are orthogonal to these sets are proportional to the columns of* $\mathbf{C}$.

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

4.2.2 Matrices whose entries are permanents

Definition 4.2.7. *Let* $\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \dots & \mathbf{a}_n\end{bmatrix} \in \mathbb{R}^{n\times n}$. *Then the permanent of* $\mathbf{A}$ *is defined as*

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

The definition of the permanent of $\mathbf{A}$ differs from that of the determinant of $\mathbf{A}$ in that the signatures of the permutations are not taken into account. This makes the permanent invariant for column permutations of $\mathbf{A}$. The notations perm $\mathbf{A}$ and $\overset{+}{|}\mathbf{A}\overset{+}{|}$ are due to Minc [26] and Muir [27], respectively.

We have the following permanental variant of compound matrix.

Definition 4.2.8. *[24] Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$. *The* C_K^m*-by-*C_R^m *matrix whose* (i,j)*-th entry is* perm $\mathbf{C}(S_K^m(i), S_R^m(j))$ *is called the* m*-th permanental compound matrix of* $\mathbf{C}$ *and is denoted by* $\mathcal{PC}_m(\mathbf{C})$.

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

Definition 4.2.9. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$. *The* C_{K+m-1}^m*-by-*C_R^m *matrix whose* (i,j)*-th entry is* perm $\mathbf{C}(Q_K^m(i), S_R^m(j))$ *is denoted by* $\mathcal{Q}_m(\mathbf{C})$.

Definition 4.2.10. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$. *The* K^m*-by-*C_R^m *matrix whose* (i,j)*-th entry is* perm $\mathbf{C}(R_K^m(i), S_R^m(j))$ *is denoted by* $\mathcal{R}_m(\mathbf{C})$.

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

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

Lemma 4.2.11. *Let* $\mathbf{C} = [\mathbf{c}^1 \quad \dots \quad \mathbf{c}^K]^T \in \mathbb{R}^{K\times R}$. *Then* $\mathcal{Q}_m(\mathbf{C})^T$ *(resp.* $\mathcal{R}_m(\mathbf{C})^T$*) has columns* $\mathcal{PC}_m([\mathbf{c}^{j_1} \quad \dots \quad \mathbf{c}^{j_m}])$, *where* $(j_1,\dots,j_m) \in Q_K^m$ *(resp.* R_K^m*).*

Example 4.2.12. *Let* $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. *Then*

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

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

4.2.3 Links between matrix $\mathcal{R}_m(\mathrm{C})$, matrix $\mathcal{B}(\mathrm{C})$ and symmetrizer

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

$$\frac{1}{m!} \sum_{(l_1,\ldots,l_m) \in P_{\{j_1,\ldots,j_m\}}} (\mathcal{T})_{(l_1,\ldots,l_m)} \tag{4.19}$$

(that is, to get $\pi_S(\mathcal{T})$ we should take the average of $m!$ tensors obtained from $\mathcal{T}$ by all possible permutations of the indices). The mapping π_S is called symmetrizer (also known as symmetrization map [23] or completely symmetric operator [22]; in [32] a matrix representation of π_S was called Kronecker product permutation matrix).

It is well known that mth-order $K \times \cdots \times K$ tensors can be vectorized into vectors of $\mathbb{R}^{K^m}$ in such a way that for any vectors $\mathbf{t}_1, \ldots, \mathbf{t}_m \in \mathbb{R}^K$ the rank-1 tensor $\mathbf{t}_1 \circ \cdots \circ \mathbf{t}_m$ corresponds to the vector $\mathbf{t}_1 \otimes \cdots \otimes \mathbf{t}_m$. This allows us to

consider the symmetrizer π_S on the space $\mathbb{R}^{K^m}$. In particular, by (4.19),

$$\pi_S(\mathbf{t}_1 \otimes \cdots \otimes \mathbf{t}_m) = \frac{1}{m!} \sum_{(l_1,\ldots,l_m)\in P_m} \mathbf{t}_{l_1} \otimes \cdots \otimes \mathbf{t}_{l_m}. \tag{4.20}$$

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

Proposition 4.2.13. *Let* $\mathbf{C} \in \mathbb{R}^{K\times R}$, $K \leq R$, $m = R-K+2$, *and* $k_{\mathbf{C}} = K$. *Let also* $\mathcal{B}(\mathbf{C})$ *be defined by* (4.13) *and let* $\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}$ *denote the restriction of the mapping* $\mathcal{R}_m(\mathbf{C})^T : \mathbb{R}^{K^m} \to \mathbb{R}^{C_R^m}$ *onto* $\mathrm{range}(\pi_S)$. *Then*

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

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

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

In the remaining part of this subsection we prove Proposition 4.2.13. We need auxiliary results and notations that we will also use in Subsection 4.3.3.

Let $\{\mathbf{e}_j^K\}_{j=1}^K$ denote the canonical basis of $\mathbb{R}^K$. Then $\{\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K\}_{(j_1,\ldots,j_m)\in R_K^m}$ is the canonical basis of $\mathbb{R}^{K^m}$ and by (4.20),

$$\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K) = \frac{1}{m!} \sum_{(l_1,\ldots,l_m)\in P_{\{j_1,\ldots,j_m\}}} \mathbf{e}_{l_1}^K \otimes \cdots \otimes \mathbf{e}_{l_m}^K. \tag{4.21}$$

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

$$\mathbf{G} \text{ has columns } \{\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K) : \ (j_1,\ldots,j_m) \in Q_K^m\}. \tag{4.22}$$

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

Lemma 4.2.14. *[32] Let* π_S *and* $\mathbf{G}$ *be defined by* (4.21)–(4.22). *Then the columns of the matrix* $\mathbf{G}$ *form an orthogonal basis of* $\mathrm{range}(\pi_S)$; *in particular,* $\dim \mathrm{range}(\pi_S) = C_{K+m-1}^m$.

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

Lemma 4.2.15. *Let* $\mathbf{C} = \begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_R\end{bmatrix} \in \mathbb{R}^{K\times R}$. *Then*

$$\mathcal{R}_m(\mathbf{C}) = m!\begin{bmatrix}\pi_S(\mathbf{c}_1\otimes\cdots\otimes\mathbf{c}_m) & \dots & \pi_S(\mathbf{c}_{R-m+1}\otimes\cdots\otimes\mathbf{c}_R)\end{bmatrix}. \tag{4.23}$$

Proof. By (4.20), the $(i_1,\dots,i_m)$-th entry of the vector $m!\pi_S(\mathbf{c}_{j_1}\otimes\cdots\otimes\mathbf{c}_{j_m})$ is equal to

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

Hence, (4.23) follows from Definition 4.2.10. □

Example 4.2.16. *Let the matrix* $\mathbf{C}$ *be as in Example 4.2.12. Then*

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

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

$$\mathbf{H} \text{ has columns } \{\mathbf{e}^{C^m_{K+m-1}}_{[j_1,\dots,j_m]} :\ (j_1,\dots,j_m)\in R^m_K\}, \tag{4.24}$$

in which $[j_1,\dots,j_m]$ denotes the ordered version of $(j_1,\dots,j_m)$. For all K^m entries of a symmetric m-th order $K\times\cdots\times K$ tensor, the corresponding column of $\mathbf{H}$ contains a "1" at the first index combination (in lexicographic ordering) where that entry can be found. The matrix $\mathbf{H}$ can be used to "compress" symmetric $K\times\cdots\times K$ tensors by removing redundancies. The matrix $\mathbf{G}$ above does the opposite thing, so $\mathbf{G}$ and $\mathbf{H}$ act as each other's inverse. It is easy to prove that indeed $\mathbf{HG} = \mathbf{I}_{C^m_{K+m-1}}$. The relations in the following lemma reflect the same relationship and will be used in Subsection 4.3.3.

Lemma 4.2.17. *Let* $\mathbf{C}\in\mathbb{R}^{K\times R}$ *and let the matrices* $\mathbf{G}$ *and* $\mathbf{H}$ *be defined by* (4.22) *and* (4.24)*, respectively. Then*

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

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

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

Proof of Proposition 4.2.13. (i) Assume that there exists $\widehat{\mathbf{t}} = \begin{bmatrix} t_{(1,\dots,m)} & \dots & t_{(R-m+1,\dots,R)} \end{bmatrix}^T \in \mathbb{R}^{C_R^m}$ such that $\mathcal{R}_m(\mathbf{C})\widehat{\mathbf{t}} = \mathbf{0}$. Then, by Lemma 4.2.15,

$$\sum_{(p_1,\dots,p_m)\in S_R^m} t_{(p_1,\dots,p_m)}\pi_S(\mathbf{c}_{p_1}\otimes\cdots\otimes\mathbf{c}_{p_m}) = \mathbf{0}. \tag{4.25}$$

Let us fix $(i_1,\dots,i_m) \in S_R^m$ and set $\{j_1,\dots,j_{K-1}\} := \{1,\dots,R\} \setminus \{i_1,\dots,i_{m-1}\}$. Then $i_m \in \{j_1,\dots,j_{K-1}\}$. Without loss of generality we can assume that $j_{K-1} = i_m$. Let the vector $\mathbf{x}$ be orthogonal to the vectors $\mathbf{c}_{j_1},\dots,\mathbf{c}_{j_{K-1}}$. Since $k_{\mathbf{C}} = K$, it follows that $\mathbf{x}$ is not orthogonal to any of $\mathbf{c}_{i_1},\dots,\mathbf{c}_{i_{m-1}}$. Similarly, there exists a vector $\mathbf{y}$ such that $\mathbf{y}$ is orthogonal to the vectors $\mathbf{c}_{j_1},\dots,\mathbf{c}_{j_{K-2}}$, and $\mathbf{y}$ is not orthogonal to any of $\mathbf{c}_{i_1},\dots,\mathbf{c}_{i_m}$. We define $\alpha_{(p_1,\dots,p_m)}$ by

$$\alpha_{(p_1,\dots,p_m)} := \frac{1}{m!}\sum_{(l_1,\dots,l_m)\in P_{\{p_1,\dots,p_m\}}} (\mathbf{c}_{l_1}^T\mathbf{x})\cdots(\mathbf{c}_{l_{m-1}}^T\mathbf{x})(\mathbf{c}_{l_m}^T\mathbf{y}),$$

$$(p_1,\dots,p_m) \in S_R^m.$$

Note that $\mathbf{c}_{l_m}^T\mathbf{y} \neq 0$ for $l_m \in \{i_1,\dots,i_m\}$ and that the product $(\mathbf{c}_{l_1}^T\mathbf{x})\cdots(\mathbf{c}_{l_{m-1}}^T\mathbf{x})$ is nonzero if and only if $\{l_1,\dots,l_{m-1}\} = \{i_1,\dots,i_{m-1}\}$. Hence, the overall product $(\mathbf{c}_{l_1}^T\mathbf{x})\cdots(\mathbf{c}_{l_{m-1}}^T\mathbf{x})(\mathbf{c}_{l_m}^T\mathbf{y})$ is nonzero if and only if $\{l_1,\dots,l_{m-1}\} = \{i_1,\dots,i_{m-1}\}$ and $l_m = i_m$. Thus, $\alpha_{(p_1,\dots,p_m)} = 0$ for $(p_1,\dots,p_m) \neq (i_1,\dots,i_m)$ and

$$\alpha_{(i_1,\dots,i_m)} = \frac{1}{m!}\sum_{\substack{(l_1,\dots,l_{m-1})\in\\ P_{\{i_1,\dots,i_{m-1}\}}}} (\mathbf{c}_{l_1}^T\mathbf{x})\cdots(\mathbf{c}_{l_{m-1}}^T\mathbf{x})(\mathbf{c}_{i_m}^T\mathbf{y}) =$$

$$\frac{1}{m}(\mathbf{c}_{i_1}^T\mathbf{x})\cdots(\mathbf{c}_{i_{m-1}}^T\mathbf{x})(\mathbf{c}_{i_m}^T\mathbf{y}).$$

Then, by (4.25),

$$0 = \sum_{(p_1,\dots,p_m)\in S_R^m} t_{(p_1,\dots,p_m)}\pi_S(\mathbf{c}_{p_1}\otimes\cdots\otimes\mathbf{c}_{p_m})^T(\underbrace{\mathbf{x}\otimes\cdots\otimes\mathbf{x}}_{m-1}\otimes\mathbf{y}) =$$

$$\sum_{(p_1,\dots,p_m)\in S_R^m} t_{(p_1,\dots,p_m)}\left[\frac{1}{m!}\sum_{(l_1,\dots,l_m)\in P_{\{p_1,\dots,p_m\}}} (\mathbf{c}_{l_1}^T\mathbf{x})\cdots(\mathbf{c}_{l_{m-1}}^T\mathbf{x})(\mathbf{c}_{l_m}^T\mathbf{y})\right] =$$

$$\sum_{(p_1,\dots,p_m)\in S_R^m} t_{(p_1,\dots,p_m)}\alpha_{(p_1,\dots,p_m)} = t_{(i_1,\dots,i_m)}\frac{1}{m}(\mathbf{c}_{i_1}^T\mathbf{x})\cdots(\mathbf{c}_{i_{m-1}}^T\mathbf{x})(\mathbf{c}_{i_m}^T\mathbf{y}).$$

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

(ii) From step (i), Lemma 4.2.14, and Lemma 4.2.17 (i),(ii) it follows that

$$C_R^m = \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T) \geq \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) =$$

$$\dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T\mathbf{G}) = \dim \operatorname{range}(\mathcal{Q}_m(\mathbf{C})^T) \geq$$

$$\dim \operatorname{range}(\mathcal{Q}_m(\mathbf{C})^T\mathbf{H}) = \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T) = C_R^m.$$

Hence, $\dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) = C_R^m$. By the rank–nullity theorem,

$$\dim \ker (\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) =$$

$$\dim \operatorname{range}(\pi_S) - \dim \operatorname{range}(\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) =$$

$$C_{K+m-1}^m - C_R^m = C_{R+1}^{R-K+2} - C_R^{R-K+2} = C_R^{K-1}.$$

(iii) Let $\mathbf{t}$ denote the $(j_1,\ldots,j_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$. It is clear that the vector $\mathbf{t}^{(m)} := \underbrace{\mathbf{t} \otimes \cdots \otimes \mathbf{t}}_{m}$ is contained in $\operatorname{range}(\pi_S)$. Hence, $\operatorname{range}\left(\mathcal{B}(\mathbf{C})^{(m)}\right) \subseteq \operatorname{range}(\pi_S)$. By step (ii) and Proposition 4.1.10 (ii), $\dim \ker (\mathcal{R}_m(\mathbf{C})^T \restriction_{\operatorname{range}(\pi_S)}) = C_R^{K-1} = \dim \operatorname{range}\left(\mathcal{B}(\mathbf{C})^{(m)}\right)$. To complete the proof we must check that $\mathcal{R}_m(\mathbf{C})^T\mathbf{t}^{(m)} = \mathbf{0}$ for all $(j_1,\ldots,j_{K-1}) \in S_R^{K-1}$. From the construction of the matrix $\mathcal{B}(\mathbf{C})$ it follows that $\mathbf{t}$ is orthogonal to the vectors $\mathbf{c}_{j_1},\ldots,\mathbf{c}_{j_{K-1}}$. Since $(K-1)+m = R+1 > R$, it follows that $(\mathbf{c}_{l_1}^T\mathbf{t})\cdots(\mathbf{c}_{l_m}^T\mathbf{t}) = 0$ for all $(l_1,\ldots,l_m) \in S_R^m$. Hence, by Lemma 4.2.15, the $(i_1,\ldots,i_m)$-th entry of the vector $\mathcal{R}_m(\mathbf{C})^T\mathbf{t}^{(m)}$ is equal to

$$\pi_S(\mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_m})^T\mathbf{t}^{(m)} = \frac{1}{m!} \sum_{(l_1,\ldots,l_m)\in P_{\{i_1,\ldots,i_m\}}} (\mathbf{c}_{l_1}^T\mathbf{t})\cdots(\mathbf{c}_{l_m}^T\mathbf{t}) = 0.$$

This completes the proof of (iii). □

4.3 Transformation of the CPD using polarized compound matrices

In this section we derive the crucial expression (4.16). The matrix $\mathbf{R}_m(\mathcal{T})$ is constructed from polarized compound matrices of the slices of the given tensor $\mathcal{T}$. The entries of polarized compound matrices are mixed discriminants. The notions of mixed discriminants and polarized compound matrices are introduced in the first two subsections.

4.3.1 Mixed discriminants

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

Definition 4.3.1. *[1] Let* $\mathbf{T}_1, \ldots, \mathbf{T}_m \in \mathbb{R}^{m \times m}$. *The mixed discriminant, denoted by* $\mathcal{D}(\mathbf{T}_1, \ldots, \mathbf{T}_m)$, *is defined as the coefficient of* $x_1 \cdots x_m$ *in* $\det(x_1\mathbf{T}_1 + \cdots + x_m\mathbf{T}_m)$, *that is*

$$\mathcal{D}(\mathbf{T}_1, \ldots, \mathbf{T}_m) = \left. \frac{\partial^m \left(\det(x_1\mathbf{T}_1 + \cdots + x_m\mathbf{T}_m)\right)}{\partial x_1 \ldots \partial x_m} \right|_{x_1 = \cdots = x_m = 0}. \tag{4.26}$$

For convenience, we have dropped the factor $1/m!$ before the fraction in (4.26). Definition 4.3.1 implies the following lemmas.

Lemma 4.3.2. *[1] The mapping* $(\mathbf{T}_1, \ldots, \mathbf{T}_m) \to \mathcal{D}(\mathbf{T}_1, \ldots, \mathbf{T}_m)$ *is multilinear and symmetric in its arguments.*

Lemma 4.3.3. *[9] Let* $\mathbf{d}_1, \ldots, \mathbf{d}_m \in \mathbb{R}^m$. *Then* $\mathcal{D}\left(\mathrm{Diag}(\mathbf{d}_1), \ldots, \mathrm{Diag}(\mathbf{d}_m)\right) = \mathrm{perm}\begin{bmatrix}\mathbf{d}_1 & \ldots & \mathbf{d}_m\end{bmatrix}$.

Proof.

$$\mathcal{D}(\mathrm{Diag}(\begin{bmatrix}d_{11} & \ldots & d_{m1}\end{bmatrix}), \ldots, \mathrm{Diag}(\begin{bmatrix}d_{1m} & \ldots & d_{mm}\end{bmatrix})) =$$

$$\left. \frac{\partial^m \left((x_1 d_{11} + \cdots + x_m d_{1m}) \cdots (x_1 d_{m1} + \cdots + x_m d_{mm})\right)}{\partial x_1 \ldots \partial x_m} \right|_{x_1 = \cdots = x_m = 0} =$$

$$\sum_{(l_1, \ldots, l_m) \in P_m} d_{1l_1} \cdots d_{ml_m} = \mathrm{perm}\begin{bmatrix}\mathbf{d}_1 & \ldots & \mathbf{d}_m\end{bmatrix}. \qquad \square$$

Mixed discriminants may be computed numerically from (4.26). A direct expression in terms of determinants is given in the following lemma.

Lemma 4.3.4. *[21, 2] Let* $\mathbf{T}_1, \ldots, \mathbf{T}_m \in \mathbb{R}^{m \times m}$. *Then*

$$\mathcal{D}(\mathbf{T}_1, \ldots, \mathbf{T}_m) = \sum_{k=1}^{m} (-1)^{m-k} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le m} \det(\mathbf{T}_{i_1} + \cdots + \mathbf{T}_{i_k}). \tag{4.27}$$

The way in which (4.27) obtains the mixed discriminant from the determinant is an instance of a technique called polarization [19].

4.3.2 Polarized compound matrices

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

Definition 4.3.5. *Let* $\min(I, J) \geq m \geq 2$ *and let* $\mathbf{T}_1, \ldots, \mathbf{T}_m \in \mathbb{R}^{I \times J}$. *The* C_I^m*-by-*C_J^m *matrix* $\mathcal{F}_{m-1}(\mathbf{T}_1, \ldots, \mathbf{T}_m)$ *is defined by*

$$\mathcal{F}_{m-1}(\mathbf{T}_1, \ldots, \mathbf{T}_m) = \left. \frac{\partial^m \left(\mathcal{C}_m(x_1\mathbf{T}_1 + \cdots + x_m\mathbf{T}_m) \right)}{\partial x_1 \ldots \partial x_m} \right|_{x_1 = \cdots = x_m = 0}. \quad (4.28)$$

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

Lemma 4.3.6. *Let* $\mathbf{T} \in \mathbb{R}^{I \times J}$ *and* $\mathbf{d}_1, \ldots, \mathbf{d}_m \in \mathbb{R}^R$. *Then*

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

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

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

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

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

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

Proof. From Definitions 4.2.1 and 4.3.5 it follows that the (i, j)-th entry of the matrix $\mathcal{F}_{m-1}(\mathbf{T}_1, \ldots, \mathbf{T}_m)$ is equal to $\mathcal{D}(\mathbf{T}_1(S_I^m(i), S_J^m(j)), \ldots, \mathbf{T}_m(S_I^m(i), S_J^m(j)))$. Hence, statements (i) and (ii) follow from Lemma 4.3.2 and Lemma 4.3.4, respectively. Statement (iii) follows from (4.28). Statement (iv) follows from (iii) and Lemma 4.2.3 (2). Finally, (v) follows from Lemma 4.2.4, statement (ii), and Lemma 4.3.3. □

Example 4.3.7.

$$\mathcal{F}_2(\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3) = \mathcal{C}_3(\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3) -$$
$$\mathcal{C}_3(\mathbf{T}_1 + \mathbf{T}_2) - \mathcal{C}_3(\mathbf{T}_1 + \mathbf{T}_3) - \mathcal{C}_3(\mathbf{T}_2 + \mathbf{T}_3) + \mathcal{C}_3(\mathbf{T}_1) + \mathcal{C}_3(\mathbf{T}_2) + \mathcal{C}_3(\mathbf{T}_3). \quad (4.29)$$

Remark 4.3.8. *The polarized compound matrix is a matrix representation of the higher-order tensor obtained by the low-rank detection mapping in [5, 28]. More specifically, in [5] a rank-*1 *detection mapping (*$m = 2$*) was used to compute the CPD and in [28] a rank-*$(L, L, 1)$ *detection mapping (*m *arbitrary) was used to compute the decomposition in rank-*$(L, L, 1)$ *terms. Statement* (iv) *of Lemma 4.3.6 explains the terminology.*

The following counterpart of Lemma 4.2.5 holds for polarized compound matrices.

Lemma 4.3.9. *Let* $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, $\mathbf{d}_1, \ldots, \mathbf{d}_m \in \mathbb{R}^R$, *and* $m \leq \min(I, J, R)$. *Then*

$$\begin{aligned}\mathcal{F}_{m-1}\left(\mathbf{A}\mathrm{Diag}(\mathbf{d}_1)\mathbf{B}^T, \ldots, \mathbf{A}\mathrm{Diag}(\mathbf{d}_m)\mathbf{B}^T\right) = \\ \mathcal{C}_m(\mathbf{A})\mathrm{Diag}\left(\mathcal{P}\mathcal{C}_m(\begin{bmatrix}\mathbf{d}_1 & \ldots & \mathbf{d}_m\end{bmatrix})\right)\mathcal{C}_m(\mathbf{B})^T.\end{aligned} \tag{4.30}$$

Proof. From Lemma 4.3.6 (ii) and Lemma 4.2.5 we have

$$\begin{aligned}\mathcal{F}_{m-1}\left(\mathbf{A}\mathrm{Diag}(\mathbf{d}_1)\mathbf{B}^T, \ldots, \mathbf{A}\mathrm{Diag}(\mathbf{d}_m)\mathbf{B}^T\right) = \\ \mathcal{C}_m(\mathbf{A})\mathcal{F}_{m-1}\left(\mathrm{Diag}(\mathbf{d}_1), \ldots, \mathrm{Diag}(\mathbf{d}_m)\right)\mathcal{C}_m(\mathbf{B})^T.\end{aligned} \tag{4.31}$$

Now (4.30) follows from (4.31) and Lemma 4.3.6 (v). □

4.3.3 Transformation of the tensor

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

Definition 4.3.10. *Let* $\mathcal{T}$ *be an* $I \times J \times K$ *tensor with frontal slices* $\mathbf{T}_1, \ldots, \mathbf{T}_K \in \mathbb{R}^{I\times J}$. *The* $(j_1, \ldots, j_m)$*-th column of the* $C_I^m C_J^m$*-by-*K^m *(resp.* $C_I^m C_J^m$*-by-*C_{K+m-1}^m*) matrix* $\mathbf{R}_m(\mathcal{T})$ *(resp.* $\mathbf{Q}_m(\mathcal{T})$*) equals* $\mathrm{vec}\left(\mathcal{F}_{m-1}(\mathbf{T}_{j_1}, \ldots, \mathbf{T}_{j_m})\right)$, *where* $(j_1, \ldots, j_m) \in R_K^m$ *(resp.* Q_K^m*).*

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

Lemma 4.3.11. *Let* $\mathbf{A} \in \mathbb{R}^{I\times R}$, $\mathbf{B} \in \mathbb{R}^{J\times R}$, $\mathbf{C} \in \mathbb{R}^{K\times R}$, *and* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Then for* $m \leq \min(I, J, K, R)$,

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

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

(iii) $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}) = \mathbf{G} \ker(\mathbf{Q}_m(\mathcal{T}))$, *where* $\mathbf{G}$ *is defined in* (4.22).

Proof. (i) Let $\mathbf{c}^1, \ldots, \mathbf{c}^K$ be the columns of the matrix $\mathbf{C}^T$. Recall that the frontal slices of $\mathcal{T}$ can be expressed as in (4.2). Then, by Lemma 4.3.9 and identity (4.3),

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

Now (i) and (ii) follow from Definition 4.3.10 and Lemma 4.2.11.

(iii) From (i), (ii), and Lemma 4.2.17 (ii) it follows that $\mathbf{R}_m(\mathcal{T})\mathbf{G} = \mathbf{Q}_m(\mathcal{T})$. Since, by Lemma 4.2.14, $\mathrm{range}(\pi_S) = \mathrm{range}(\mathbf{G})$ we obtain (iii). □

4.4 Overall results and algorithms

The overall algorithm goes now as follows. We first compute $\mathbf{Q}_m(\mathcal{T})$ from $\mathcal{T}$, determine its null space, which, after symmetrization, yields $\ker\left(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}\right)$, as explained in Lemma 4.3.11 (iii). The following lemma makes now, for a particular choice of m, a connection with $\mathcal{B}(\mathbf{C})$.

Lemma 4.4.1. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $m := R - K + 2$, *and* $\mathbf{R}_m(\mathcal{T})$ *be defined in Definition 4.3.10. Assume that* $k_{\mathbf{C}} = K$ *and that* $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ *has full column rank. Then*

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

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

Proof. Since $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank, it follows from Lemma 4.3.11 (i) that $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}) = \ker\left(\mathcal{R}_m(\mathbf{C})^T \restriction_{\mathrm{range}(\pi_S)}\right)$. Statements (i) and (ii) now follow from Proposition 4.2.13 (iii) and (ii), respectively. □

So far, we have obtained from $\mathcal{T}$ a basis for the column space of $\mathcal{B}(\mathbf{C})^{(m)}$. The following lemma explains that the basis vectors may be stacked in a tensor that has $\mathcal{B}(\mathbf{C})$ as factor matrix. Moreover, the CPD may be computed by a GEVD as in Theorem 4.1.1.

Lemma 4.4.2. *Suppose that the conditions of Lemma 4.4.1 hold. Let* $\mathbf{W}$ *be a* $K^m \times C_R^{K-1}$ *matrix such that*

$$\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)}) = \mathrm{range}\ (\mathbf{W}) \tag{4.32}$$

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

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

$$\mathcal{W} = \left[\mathcal{B}(\mathbf{C}), \mathcal{B}(\mathbf{C})^{(m-1)}, \mathbf{M}\right]_{C_R^{K-1}}; \tag{4.33}$$

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

Proof. (i) From Lemma 4.4.1 (ii) and (4.32) it follows that there exists a nonsingular $C_R^{K-1} \times C_R^{K-1}$ matrix $\mathbf{M}$, such that $\mathbf{W} = \mathcal{B}(\mathbf{C})^{(m)}\mathbf{M}^T = \left(\mathcal{B}(\mathbf{C}) \odot \mathcal{B}(\mathbf{C})^{(m-1)}\right)\mathbf{M}^T$. Hence, by (4.4), (4.33) holds.

(ii) From Proposition 4.1.10 it follows that $k_{\mathcal{B}(\mathbf{C})} \geq 2$ and that the matrix $\mathcal{B}(\mathbf{C})^{(m-1)}$ has rank C_R^{K-1}. The statement now follows from Theorem 4.1.1. □

After finding $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling, we may define $\mathbf{C}$ as explained in Subsection 4.2.1. The following Lemma completes the proof of Theorem 4.1.6 (ii). Its proof shows how the other factor matrices may be determined once $\mathbf{C}$ has been obtained. The computation involves another CPD of the form in Theorem 4.1.1. The result is a variant of [37, Theorem 3.8]; in this step of the derivation we do not assume that the decomposition is canonical.

Lemma 4.4.3. *Let* $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and the* $K \times R$ *matrix* $\mathbf{C}$ *be known. Assume that* $k_{\mathbf{C}} = K \geq 2$, *and that* $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) + k_{\mathbf{C}} \geq R + 2$. *Then the matrices* $\mathbf{A}$, $\mathbf{B}$ *can be found algebraically up to column scaling.*

Proof. We obviously have $k_{\mathbf{C}} = r_{\mathbf{C}} = K$. Let $\mathbf{X} = \begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_K\end{bmatrix}$. By multiplying with $\mathbf{X}^{-1}$ we will create a CPD with $K-2$ terms less than R. It is clear that the matrix formed by the first two rows of $\mathbf{X}^{-1}\mathbf{C}$ has the form $\begin{bmatrix}\mathbf{I}_2 & \mathbf{O}_{2\times(K-2)} & \mathbf{Y}\end{bmatrix}$, where $\mathbf{Y}$ is a $2 \times (R-K)$ matrix. Define $\widetilde{\mathbf{A}} := \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_{K+1} & \dots & \mathbf{a}_R\end{bmatrix}$, $\widetilde{\mathbf{B}} := \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_{K+1} & \dots & \mathbf{b}_R\end{bmatrix}$, and $\widetilde{\mathbf{C}} = \begin{bmatrix}\mathbf{I}_2 & \mathbf{Y}\end{bmatrix}$. Let also $\widetilde{\mathcal{T}}$ denote the $I \times J \times 2$ tensor such that $\mathrm{Matr}(\widetilde{\mathcal{T}})$ coincides with the first two columns of the matrix $\mathrm{Matr}(\mathcal{T})\mathbf{X}^{-T}$. From (4.4) it

follows that $\text{Matr}(\mathcal{T})\mathbf{X}^{-T} = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{X}^{-T}$. Hence, $\text{Matr}(\widetilde{\mathcal{T}}) = (\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}})\widetilde{\mathbf{C}}^T$ or $\widetilde{\mathcal{T}} = [\widetilde{\mathbf{A}}, \widetilde{\mathbf{B}}, \widetilde{\mathbf{C}}]_{R-K+2}$, which of the desired form.

It is easy to show that $\widetilde{\mathcal{T}}$ satisfies the conditions of Theorem 4.1.1, which means that its rank is $R - K + 2$, that its CPD is unique and that the factor matrices may be found algebraically. The indeterminacies in $\widetilde{\mathcal{T}} = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_{R-K+2}$ are limited to the existence of a permutation matrix $\mathbf{P}$ and a nonsingular diagonal matrix $\mathbf{\Lambda}$ such that $\widetilde{\mathbf{C}} = \widehat{\mathbf{C}}\mathbf{P}\mathbf{\Lambda}$ and $\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}} = (\widehat{\mathbf{A}} \odot \widehat{\mathbf{B}})\mathbf{P}\mathbf{\Lambda}^{-1}$.

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

The overall procedure that constitutes the proof of Theorem 4.1.6 (ii) is summarized in Algorithm 1. Phase 2 is formulated in the way that has combinatorial complexity and quickly becomes computationally infeasible. The amount of work may be reduced by exploiting the dependencies in $\mathbf{F}$ only partially.

We derive an algorithmic variant that further reduces the computational cost. This algorithm is given in Algorithm 2 below. While Algorithm 1 first determines $\mathbf{C}$ and then finds $\mathbf{A}$ and $\mathbf{B}$, Algorithm 2 works the other way around. The basic idea is as follows. Like in Algorithm 1, we first find a matrix $\mathbf{F}$ that is equal to $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. If $\mathbf{C}$ is square, we have from Subsection 4.2.1 that $\mathcal{B}(\mathbf{C}) = \det(\mathbf{C})\mathbf{C}^{-T}\mathbf{L}$ and multiplication of $\mathcal{T}$ with $\mathbf{F}^T$ in the third mode yields a tensor of which every frontal slice is a rank-1 matrix, proportional to $\mathbf{a}_r\mathbf{b}_r^T$ for some $r \in \{1, \ldots, R\}$. On the other hand, if $\mathbf{C}$ is rectangular ($K < R$), then multiplication with $\mathbf{F}^T$ yields a tensor of which all slices are rank-$(R - K + 1)$ matrices, generated by $R - K + 1$ rank-1 matrices $\mathbf{a}_r\mathbf{b}_r^T$. If we choose slices that have all but one rank-1 matrix in common, these form a tensor that is as in Theorem 4.1.1 and of which the CPD yields $R-K+2$ columns of $\mathbf{A}$ and $\mathbf{B}$. The result is formalized in the following lemma. The second statement implies that we do not have to compute the CPD to verify whether a slice combination is suitable.

Lemma 4.4.4. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, the matrix $\mathbf{F} \in \mathbb{R}^{K \times C_R^{K-1}}$ coincide with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling, $\mathcal{V} = [\mathbf{A}, \mathbf{B}, \mathbf{F}^T\mathbf{C}]_R$, $k_{\mathbf{C}} = K$, $m := R - K + 2$, and $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ have full column rank. Let also $\mathbf{y}_1, \ldots, \mathbf{y}_{C_R^{K-1}}$*

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

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

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

Phase 1 (based on Lemma 4.4.2): Find the matrix $\mathbf{F} \in \mathbb{R}^{K\times C_R^{K-1}}$ such that $\mathbf{F}$ coincides with $\mathcal{B}(\mathbf{C})$ up to (unknown) column permutation and scaling

Apply Lemma 4.3.11 (iii) to find the $K^m \times C_R^{K-1}$ *matrix* $\mathbf{W}$ *such that* (4.32) *holds*

1: Construct the $C_I^m C_J^m$-by-C_{K+m-1}^m matrix $\mathbf{Q}_m(\mathcal{T})$ by Definition 4.3.10
2: Find $\bar{\mathbf{w}}_1, \ldots, \bar{\mathbf{w}}_{C_R^{K-1}}$ which form a basis of $\ker(\mathbf{Q}_m(\mathcal{T}))$
3: $\mathbf{W} \leftarrow \mathbf{G}\begin{bmatrix}\bar{\mathbf{w}}_1 & \ldots & \bar{\mathbf{w}}_{C_R^{K-1}}\end{bmatrix}$, where $\mathbf{G} \in \mathbb{R}^{K^m \times C_{K+m-1}^m}$ is defined by (4.22)

Apply Theorem 4.1.1 (ii) *to find* $\mathbf{F}$

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

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

6: Compute R subsets of C_{R-1}^{K-2} columns of $\mathbf{F}$ that are linearly dependent
7: Compute $\mathbf{c}_1, \ldots, \mathbf{c}_R$ as orthogonal complements to sets found in step 6

Phase 3 (based on Lemma 4.4.3): Find the matrices $\mathbf{A}$ and $\mathbf{B}$

Apply Theorem 4.1.1 (ii) *to find the columns of* $\mathbf{S} := \mathbf{A} \odot \mathbf{B}$ *with indices in* $\{1, 2, K+1, \ldots, R\}$

8: $\mathbf{Z} = \begin{bmatrix}\mathbf{z}_1 & \ldots & \mathbf{z}_K\end{bmatrix} \leftarrow \mathrm{Matr}(\mathcal{T}) \begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_K\end{bmatrix}^{-T}$
9: $\widetilde{\mathcal{T}} \leftarrow \mathrm{Tens}(\begin{bmatrix}\mathbf{z}_1 & \mathbf{z}_2\end{bmatrix}, I, J)$
10: Compute the CPD $\widetilde{\mathcal{T}} = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_{R-K+2}$ (GEVD)
11: $\widetilde{\mathbf{C}} \leftarrow \begin{bmatrix}\mathbf{I}_2 & \mathbf{Y}\end{bmatrix}$, where $\mathbf{Y}$ is the $2 \times R$ submatrix in upper right-hand corner of $\begin{bmatrix}\mathbf{c}_1 & \ldots & \mathbf{c}_K\end{bmatrix}^{-1} \mathbf{C}$
12: Compute permutation matrix $\mathbf{P}$ and diagonal matrix $\mathbf{\Lambda}$ such that $\widetilde{\mathbf{C}} = \widehat{\mathbf{C}}\mathbf{P}\mathbf{\Lambda}$
13: $\begin{bmatrix}\mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_{K+1} \ldots & \mathbf{s}_R\end{bmatrix} \leftarrow (\widehat{\mathbf{A}} \odot \widehat{\mathbf{B}})\mathbf{P}\mathbf{\Lambda}^{-1}$

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

14: $\mathbf{Z} \leftarrow \mathrm{Matr}(\mathcal{T}) - (\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}}) \begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_{K+1} & \ldots & \mathbf{c}_R\end{bmatrix}^T$
15: $\begin{bmatrix}\mathbf{s}_3 & \ldots & \mathbf{s}_K\end{bmatrix} \leftarrow \mathbf{Z} \begin{bmatrix}\mathbf{c}_3 & \ldots & \mathbf{c}_K\end{bmatrix}^\dagger$
16: Find the columns of $\mathbf{A}$ and $\mathbf{B}$ from the equations $\mathbf{a}_r \otimes \mathbf{b}_r = \mathbf{s}_r$, $r = 1, \ldots, R$

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

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

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

Phase 1 (based on Lemma 4.4.2): Find the matrix $\mathbf{F} \in \mathbb{R}^{K \times C_R^{K-1}}$ such that $\mathbf{F}$ coincides with $\mathcal{B}(\mathbf{C})$ up to (unknown) column permutation and scaling

1–5: identical to Algorithm 1

Phase 2 (based on Lemma 4.4.4): Find the matrices $\mathbf{A}$ and $\mathbf{B}$

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

7: $\mathcal{J} \leftarrow \{(i,j) : \ r_{[\mathbf{V}_i \ \mathbf{V}_j]} = r_{[\mathbf{V}_i^T \ \mathbf{V}_j^T]} = m, \ 1 \leq i < j \leq C_R^{K-1}\}$
Apply Theorem 4.1.1 (ii) *to find* $C_R^m C_m^2$ *sets of* m *columns of* $\mathbf{A}$ *and* $\mathbf{B}$ *each*

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

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

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

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

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

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

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

denote the columns of $\mathbf{C}^T\mathbf{F}$ *and* $\mathbf{V}_1, \dots, \mathbf{V}_{C_R^{K-1}}$ *denote the frontal slices of* $\mathcal{V}$. *Then the following statements are equivalent.*

(i) *The matrix* $[\mathbf{y}_i\ \mathbf{y}_j]$ *has exactly* $K-2$ *zero rows.*

(ii) *The matrices* $[\mathbf{V}_i\ \mathbf{V}_j]$ *and* $[\mathbf{V}_i^T\ \mathbf{V}_j^T]$ *have rank* m.

(iii) *The tensor* $\mathcal{V}_{ij}$ *formed by the frontal slices* $\mathbf{V}_i$ *and* $\mathbf{V}_j$ *has rank* m.

The CPD $\mathcal{V}_{ij} = [\begin{bmatrix}\mathbf{a}_{p_1} & \dots & \mathbf{a}_{p_m}\end{bmatrix}, \begin{bmatrix}\mathbf{b}_{p_1} & \dots & \mathbf{b}_{p_m}\end{bmatrix}, \widehat{\mathbf{C}}]_m$ *can be found algebraically by Theorem 4.1.1* (ii) *and the indices* $p_1, \dots, p_m$ *are uniquely defined by the pair* (i, j).

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

To summarize, we first find a matrix $\mathbf{F} \in \mathbb{R}^{K \times C_R^{K-1}}$ that coincides with $\mathcal{B}(\mathbf{C})$ up to column permutation and scaling. We construct a tensor $\mathcal{V}$ from the tensor $\mathcal{T}$ and the matrix $\mathbf{F}$ as in Lemma 4.4.4 and choose slice combinations for which $[\mathbf{V}_i\ \mathbf{V}_j]$ and $[\mathbf{V}_i^T\ \mathbf{V}_j^T]$ have rank m. For each such slice combination the CPD of the corresponding tensor yields m columns of $\mathbf{A}$ and $\mathbf{B}$. In this way we obtain all columns of $\mathbf{A}$ and $\mathbf{B}$. Overall, there exist exactly $C_R^m C_m^2$ pairs (i, j) such that (i)–(iii) hold. The amount of work can be reduced by finding enough, instead of all, tensors $\mathcal{V}_{ij}$ that yield columns of $\mathbf{A}$ and $\mathbf{B}$. The matrix $\mathbf{C}$ is finally obtained by $\mathbf{C} = \big((\mathbf{A} \odot \mathbf{B})^\dagger \mathrm{Matr}(\mathcal{T})\big)^T$.

Proof of Theorem 4.1.7 (ii). Let the matrix $\mathbf{C}$ have K rows, $\mathbf{X}$ be a $k_{\mathbf{C}} \times K$ matrix, and $\bar{\mathcal{T}} := [\mathbf{A}, \mathbf{B}, \mathbf{X}\mathbf{C}]_R$. Then $\mathbf{X}\mathbf{C} \in \mathbb{R}^{k_{\mathbf{C}} \times R}$ and by (4.4), $\mathrm{Matr}(\mathcal{T})\mathbf{X}^T := (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T\mathbf{X}^T = (\mathbf{A} \odot \mathbf{B})(\mathbf{X}\mathbf{C})^T = \mathrm{Matr}(\bar{\mathcal{T}})$. Thus, the multiplication of the third factor matrix of $\mathcal{T}$ by $\mathbf{X}$ from the left is equivalent to the multiplication of the matrix unfolding $\mathrm{Matr}(\mathcal{T})$ by $\mathbf{X}^T$ from the right.

(i) Assume that $\mathbf{X}$ is such that $r_{\mathbf{X}\mathbf{C}} = k_{\mathbf{X}\mathbf{C}} = k_{\mathbf{C}}$. Then by Theorem 4.1.6, the CPD of $\bar{\mathcal{T}}$ is unique and can be found algebraically. In particular, the matrix $\mathbf{A} \odot \mathbf{B}$ has full column rank and can be found up to column permutation and scaling. Hence, $\mathbf{C} = \big((\mathbf{A} \odot \mathbf{B})^\dagger \mathrm{Matr}(\mathcal{T})\big)^T$, and the proof is completed.

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

$$\mu\{\mathrm{vec}(\mathbf{X}) :\ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K},\ \ k_{\mathbf{X}\mathbf{C}} < k_{\mathbf{C}}\} = 0, \tag{4.34}$$

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

$$\mu\{\mathrm{vec}(\mathbf{X}) :\ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K},\ \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathbf{f} = 0\} = 0. \tag{4.35}$$

From Lemma 4.2.3 (1) it follows that the matrix $\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C})$ has all columns nonzero. By Lemma 4.2.5, $k_{\mathbf{XC}} < k_{\mathbf{C}}$ if and only if the vector $\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C}) = \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{XC})$ has a zero entry. Hence, by (4.35),

$$\begin{aligned}
&\{\mathrm{vec}(\mathbf{X}) :\ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K},\ \ k_{\mathbf{XC}} < k_{\mathbf{C}}\} = \\
&\{\mathrm{vec}(\mathbf{X}) :\ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K},\ \ \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C}) \text{ has a zero entry}\} = \\
&\bigcup_{\mathbf{f} \text{ is a column of } \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{C})} \{\mathrm{vec}(\mathbf{X}) :\ \mathbf{X} \in \mathbb{R}^{k_{\mathbf{C}} \times K},\ \ \mathcal{C}_{k_{\mathbf{C}}}(\mathbf{X})\mathbf{f} = 0\}.
\end{aligned} \tag{4.36}$$

Now (4.34) follows from (4.36) and (4.35). □

Remark 4.4.5. *Let the conditions of Theorem 4.1.7, Corollary 4.1.8 or Corollary 4.1.9 hold. The following procedure for computing the CPD follows from the proof of Theorem 4.1.7 (ii). First, we generate a random $k_{\mathbf{C}} \times K$ matrix $\mathbf{X}$ and set $\bar{\mathbf{T}} = \mathrm{Matr}(\mathcal{T})\mathbf{X}^T$. Then the $I \times J \times k_{\mathbf{C}}$ tensor $\bar{\mathcal{T}} := \mathrm{Tens}(\bar{\mathbf{T}}, I, J)$ satisfies the conditions of Theorem 4.1.6 (ii) with K replaced by $k_{\mathbf{C}}$. Hence, the CPD $\bar{\mathcal{T}} = [\mathbf{A}, \mathbf{B}, \mathbf{XC}]_R$ can be found by Algorithm 1 or Algorithm 2. Finally, the matrix $\mathbf{C}$ is obtained by $\mathbf{C} = \left((\mathbf{A} \odot \mathbf{B})^{\dagger}\mathrm{Matr}(\mathcal{T})\right)^T$.*

Example 4.4.6. *Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_5$ with*

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

Since condition (4.6) *does not hold, the rank and uniqueness of the CPD do not follow from Kruskal's Theorem 4.1.3. One can easily check that the conditions of Theorem 4.1.6 hold for $m = 5 - 4 + 2 = 3$. Hence, the factor matrices of $\mathcal{T}$ can be found by Algorithms 1 and 2.*

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

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

We construct the $C_4^3C_4^3$-by-C_6^3 (or 16*-by-*20*) matrix $\mathbf{Q}_3(\mathcal{T})$ by Definition 4.3.10. For instance, the* $(1, 2, 3)$*-rd (or the* 6*-th) column of $\mathbf{Q}_3(\mathcal{T})$ is equal to* $\mathrm{vec}(\mathcal{F}_2(\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3))$*, where $\mathcal{F}_2(\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3)$ is computed by* (4.29) *and equals*

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

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

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

and $\mathbf{e}_{i,-j}^{20} := \mathbf{e}_i^{20} - \mathbf{e}_j^{20}$*. Let* $\mathbf{G}$ *be the* 64×20 *matrix defined by* (4.22)*. We denote by* $\mathcal{W}$ *the* $4 \times 16 \times 10$ *tensor such that* $\operatorname{Matr}(\mathcal{W}) = \mathbf{G}\overline{\mathbf{W}}$*. We find algebraically the CPD* $\mathcal{W} = [\mathbf{F}, \mathbf{F}_2, \mathbf{F}_3]_{10}$ *with*

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

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

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

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

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

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

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

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

slices of $\mathcal{V}$ *and let* $\mathcal{V}_{ij}$ *denote the* $4 \times 4 \times 2$ *tensor with frontal slices* $\mathbf{V}_i$ *and* $\mathbf{V}_j$. *We construct the set*

$$\mathcal{J} := \{(i,j) : \text{ the matrices } [\mathbf{V}_i \ \mathbf{V}_j] \text{ and } [\mathbf{V}_i^T \ \mathbf{V}_j^T] \text{ have rank } 3, \\ 1 \le i < j \le 10\} =$$

$$\{(1,2),\ (1,3),\ (1,4),\ (1,8),\ (1,9),\ (1,10),\ (2,4),\ (2,5),\ (2,6),\ (2,8),$$

$$(2,10),\ (3,5),\ (3,7),\ (3,8),\ (3,9),\ (3,10),\ (4,6),\ (4,7),\ (4,8),\ (4,9),$$

$$(5,6),\ (5,7),\ (5,8),\ (5,10),\ (6,7),\ (6,9),\ (6,10),\ (7,8),\ (7,9),\ (9,10)\}.$$

For $(i,j) \in \mathcal{J}$, $\mathcal{V}_{ij}$ *has rank* 3 *and the CPD can be computed algebraically. For instance,* $\mathrm{Matr}(\mathcal{V}_{12}) = \mathrm{Matr}(\mathcal{T})[\mathbf{f}_1 \ \mathbf{f}_2] = (\mathbf{A} \odot \mathbf{B})\mathbf{C}^T[\mathbf{f}_1 \ \mathbf{f}_2]$. *Since* $\mathbf{C}^T[\mathbf{f}_1 \ \mathbf{f}_2] = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 \end{bmatrix}^T$, *we have*

$$\mathcal{V}_{12} = \left[[\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3], [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3], \begin{bmatrix} 2 & 1 & 0 \\ 0 & -1 & 2 \end{bmatrix} \right]_3.$$

In this way for each pair $(i,j) \in \mathcal{J}$ *we estimate up to column scaling three columns of* $\mathbf{A}$ *and the corresponding columns of* $\mathbf{B}$. *If we store all the estimates of columns of* $\mathbf{A}$ *and* $\mathbf{B}$ *in* 4×90 *matrices* $\widetilde{\mathbf{A}}$ *and* $\widetilde{\mathbf{B}}$ *then* $\widetilde{\mathbf{A}} \odot \widetilde{\mathbf{B}}$ *will contain* 5 *clusters of* 18 *collinear columns. Taking the cluster centers we get a matrix* $\mathbf{Z}$ *which coincides with* $\mathbf{A} \odot \mathbf{B}$ *up to column scaling and permutation. Finally, the matrix* $\left(\mathbf{Z}^{\dagger}\mathrm{Matr}(\mathcal{T})\right)^T$ *coincides with* $\mathbf{C}$ *up to column scaling and the same permutation.*

Example 4.4.7. *It was shown in [8] that the conditions of Theorem 4.1.6 hold for a generic* $6 \times 6 \times 7$ *tensor of rank* 9. *This case is beyond Kruskal's bound. Let* $\mathbf{F}$ *be the* 7×84 *matrix produced by Phase 1 of Algorithms 1 and 2. Each column of the third factor matrix of the tensor is orthogonal to exactly* 42 *columns of the matrix* $\mathbf{F}$. *Since* C_{84}^{42} *is of order* 10^{24}, *Phase 2 as presented in Algorithm 1 is computationally infeasible. On the other hand, in Phase 2 of Algorithm 2 we check the rank of* $2C_{84}^2 = 3486$ *matrices of the size* 6×12 *each. Then we have to find algebraically the CPD of* $C_9^4 C_4^2 = 756$ *rank-*4 *tensors with dimensions* $6 \times 6 \times 2$, *which is equivalent with the computation of the GEVD of the associated matrix pencils. Moreover, one may further limit the amount of work by only determining subsets of* $\mathcal{J}$. *We implemented Algorithm 2 in MATLAB 2008a and we did experiments on a computer with Intel(R) Core(TM) T9600 Duo 2.80GHz CPU and 4GB memory running Windows Vista. The simulations demonstrate that with a suboptimal implementation, it takes less than* 9 *seconds to compute the CPD of a generic* $6 \times 6 \times 7$ *tensor of rank* 9.

4.5 Conclusion

We have proposed two algorithms to compute CPD. Both algorithms are algebraic in the sense that they rely only on standard linear algebra and reduce the problem to the computation of GEVD. The reduction exploits properties of (polarized) compound matrices and permanents. The derivation spans the possibilities from [18] to [5, 15] and covers cases beyond Kruskal's bound.

In this paper we have limited ourselves to exact CPD. In applications, CPD most often only approximates the given (noisy) tensor. A first observation is that the "exact result" could be used to initialize iterative algorithms for problem (4.5). We also note that (4.33) may be interpreted as the CPD of a partially symmetric tensor of order $m+1$ of which the first m factor matrices are equal and parameterized by $\mathbf{C}$. This is a structure that can be handled by current algorithms in Tensorlab [34]. These algorithms are optimization-based and are not formally guaranteed to find the solution. However, they show excellent performance in practice. So far, we have computed $\ker(\mathbf{R}_m(\mathcal{T}) \restriction_{\mathrm{range}(\pi_S)})$ and then we have fitted the CPD structure to the result. Numerically, we could go a step further and take the Khatri-Rao structure into account in the computation of the kernel itself, with the kernel vectors parameterized by $\mathbf{C}$ and $\mathbf{M}$. One may also investigate whether the Khatri-Rao structure and the structure of $\mathbf{R}_m(\mathcal{T})$ may be exploited to avoid the computation of the mixed discriminants, so that one obtains an algorithm that works directly on $\mathcal{T}$. Since numerical aspects lead to a different type of study, we choose to defer them to an other paper.

Bibliography

[1] A. D. Aleksandrov. Zur Theorie der gemischten Volumina von konvexen Körpern. IV. Die gemischten Diskriminanten und die gemischten Volumina (Russian) (German summary). *Mat. Sb.*, 3 (45), no. 2:227–251., 1938.

[2] R.B. Bapat. Mixed discriminants of positive semidefinite matrices. *Linear Algebra Appl.*, 126(0):107–124, 1989.

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

[4] P. Comon, X. Luciani, and A. L. F. de Almeida. Tensor decompositions, alternating least squares and other tales. *J. Chemometrics*, 23(7-8):393–405, 2009.

[5] L. De Lathauwer. A Link Between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization. *SIAM J. Matrix Anal. Appl.*, 28:642–666, August 2006.

[6] L. De Lathauwer. 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*, pages 558–563, 2011.

[7] I. Domanov and L. De Lathauwer. On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors— Part I: Basic Results and Uniqueness of One Factor Matrix. *SIAM J. Matrix Anal. Appl.*, 34(3):855–875, 2013.

[8] I. Domanov and L. De Lathauwer. On the Uniqueness of the Canonical Polyadic Decomposition of Third-Order Tensors— Part II: Overall Uniqueness. *SIAM J. Matrix Anal. Appl.*, 34(3):876–903, 2013.

[9] G. P. Egorychev. Proof of the van der Waerden conjecture for permanents. *Siberian Math. J.*, 22:854–859, 1981.

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

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

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

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

[14] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[15] T. Jiang and N. D. Sidiropoulos. Kruskal's Permutation Lemma and the Identification of CANDECOMP/PARAFAC and Bilinear Models with Constant Modulus Constraints. *IEEE Trans. Signal Process.*, 52(9):2625–2636, September 2004.

[16] T. G. Kolda and B. W. Bader. Tensor Decompositions and Applications. *SIAM Review*, 51(3):455–500, September 2009.

[17] P. M Kroonenberg. *Applied Multiway Data Analysis*. Hoboken, NJ: Wiley, 2008.

[18] J. B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.*, 18(2):95–138, 1977.

[19] J. M. Landsberg. *Tensors: Geometry and Applications*. AMS, Providence, Rhode Island, 2012.

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

[21] J. S. Lomont and M. S. Cheena. A multilinearity property of determinant functions. *Linear and Multilinear Algebra*, 14(3):199–223, 1983.

[22] M. Marcus. The Hadamard Theorem for Permanents. *Proc. Amer. Math. Soc.*, 15(6):pp. 967–973, 1964.

[23] M. Marcus. *Finite dimensional multilinear algebra*. Number pt. 1 in Pure and applied mathematics. M. Dekker, 1973.

[24] M. Marcus and H. Minc. On the relation between the determinant and the permanent. *Illinois J. Math.*, 5(3):376–381, 1961.

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

[26] H. Minc. *Permanents*. Encyclopedia of Mathematics and its Applications. Addison-Wesley Publishing Company, Advanced Book Program, 1978.

[27] T. Muir. *A treatise on the theory of determinants: with graduated sets of exercises for use in colleges and schools*. Macmillan and Co., 1882.

[28] D. Nion and L. De Lathauwer. A Study of the Decomposition of a Third-Order Tensor in Rank-$(L, L, 1)$ Terms. *ESAT-SISTA Internal Report, 11-239, Leuven, Belgium: Department of Electrical Engineering (ESAT), KU Leuven*, 2011.

[29] L. Oeding and G. Ottaviani. Eigenvectors of tensors and algorithms for Waring decomposition. *Journal of Symbolic Computation*, 54(0):9–35, 2013.

[30] E. Sanchez and B.. Kowalski. Tensorial resolution: A direct trilinear decomposition. *J. Chemometrics*, 4(1):29–45, 1990.

[31] R: Sands and F. Young. Component models for three-way data: An alternating least squares algorithm with optimal scaling features. *Psychometrika*, 45(1):39–67, 1980.

[32] J. R. Schott. Kronecker product permutation matrices and their application to moment matrices of the normal distribution. *Journal of Multivariate Analysis*, 87(1):177 – 190, 2003.

[33] A.K. Smilde, R. Bro, and P. Geladi. *Multi-way analysis with applications in the chemical sciences*. J. Wiley, 2004.

[34] L. Sorber, M. Van Barel, and L. De Lathauwer. Tensorlab v1.0. *Available online, February 2013. URL: http://esat.kuleuven.be/sista/tensorlab/.*

[35] L. Sorber, M. Van Barel, and L. De Lathauwer. Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-(L_r,L_r,1) terms and a new generalization. *SIAM J. Optim.*, 23(2):695–720, 2013.

[36] M. Sørensen and L. De Lathauwer. Blind Signal Separation via Tensor Decomposition with Vandermonde Factor Part I: Canonical Polyadic Decomposition. *IEEE Trans. Signal Process.*, (accepted for publication), 2013.

[37] M. Sørensen and L. De Lathauwer. New Uniqueness Conditions for the Canonical Polyadic Decomposition of Third-Order Tensors. *ESAT-SISTA Internal Report, 13-05, Leuven, Belgium: Department of Electrical Engineering (ESAT), KU Leuven*, 2013.

[38] M. Sørensen, L. De Lathauwer, P. Comon, S. Icart, and L. Deneire. Canonical Polyadic Decomposition with a Columnwise Orthonormal Factor Matrix. *SIAM J. Matrix Anal. Appl.*, 33(4):1190–1213, 2012.

[39] J. Ten Berge and Jorge N. Tendeiro. The link between sufficient conditions by Harshman and by Kruskal for uniqueness in Candecomp/Parafac. *Journal of Chemometrics*, 23(7-8):321–323, 2009.

Chapter 5

Conclusion

5.1 Conclusion

Since 1970, CPD finds more and more applications in Chemometrics, Psychometrics, Signal Processing, Data Analysis, etc. The most attractive property of the CPD is its uniqueness under mild (realistic) constraints on factor matrices. In applied multilinear algebra, only three results (except cases of particular tensors) on uniqueness have been obtained till recently:

1970: two factor matrices have full column rank, and the third factor matrix has no proportional columns (see Theorem 2.1.6);

1977: factor matrices satisfy Kruskal's type conditions (see Theorems 3.1.8 and 3.1.9);

2004: one factor matrix has full column rank, and the remaining factor matrices satisfy condition (U_2) or (C_2) (see Theorems 2.1.11 and 2.1.12).

In Chapter 2 we considered the case where at least one factor matrix is unique. The situation when one factor matrix is unique but the overall CPD is not unique is typical for tensors with collinear loading vectors in some mode (see Example 2.4.11). We employed linear algebra tools to obtain new results on compound matrices. Then we obtained new conditions guaranteeing rank and uniqueness of one factor matrix in CPD.

In Chapter 3 we studied the overall uniqueness of the CPD. Our results cover the deterministic results mentioned at the beginning of this chapter.

We also obtain new results on generic uniqueness of the structured CPD, the case when factor matrices analytically depend on some parameters. This includes the cases of partially symmetric tensors, tensors with Hankel, Toeplitz or Vandermonde factor matrices and cases when some entries of factor matrices are not allowed to change.

Quite strong results on generic uniqueness have been obtained in algebraic geometry (see Theorems 3.1.16–3.1.19). However, for structured (like Hankel or Vandermonde) factor matrices the use of algebraic geometry results is not trivial, if possible at all. We explained how to modify our method for particular structure. In particular, we obtained new bounds on rank for which the Indscal decomposition is unique.

We also considered the case when two (in general non-canonical) polyadic decompositions of the same tensor share one or two factor matrices. We found conditions that guarantee that the polyadic decompositions coincide.

In Chapter 4 we presented two algorithms for computation of the CPD. Both algorithms work under mild conditions on factor matrices (for instance, under famous Kruskal's condition) and reduce the problem to the generalized eigenvalue decomposition.

The results obtained in the thesis have applications in signal separation and data analysis, of which the general goal can be formulated as the unmixing of signals that are observed together and/or the splitting of available data into meaningful, interpretable components. The uniqueness properties of tensor decompositions make multilinear algebra the proper framework for solving such problems. Tensor data are nowadays ubiquitous (e.g. in array processing, wireless communication, machine learning, image processing, biomedical applications, etc.).

5.2 Future work

Algorithms. The computational aspects of results that have been obtained in thesis may be further developed. In particular, one may develop algorithms for the computation of CPD under conditions that are more general than Kruskal type conditions and conditions that take the form of the Khatri-Rao product of compound matrices having full column rank. It seems that here it is not possible anymore to derive algebraic algorithms. One could derive optimization-based algorithms, with an efficient handling of the Jacobian and (approximate) Hessian, as in [4], and with exact line/plane search, as in [2, 3]. The algorithms may be broadly useful, as similar equations are found in system identification, machine learning, image analysis, telecommunication,... Along

the same lines, optimization variants of the algorithms in Chapter 4 may be developed. The optimization-based approach is more convenient for fitting decompositions to (noisy) data; the latter rarely admit an exact CPD or BTD. The solution obtained in Chapter 4 may nevertheless be used as a good starting value. This work should allow computation of CPD under much more general conditions than that can currently be dealt with.

Block Term Decomposition (BTD). Matrix-based thinking has so far led one to think almost exclusively in terms of rank-1 components. BTD has recently been proposed as a profoundly new concept [1]. Block terms are only required to have low "multilinear rank". So far only Kruskal-type conditions have been obtained. Generalizing results from Chapters 2 and 3, one may obtain conditions that are an order of magnitude more relaxed. One may also find ways to explicitly compute decompositions when they are exact. BTD allows a fundamentally new approach to signal separation. An analogy is that CPD splits data in "atoms" (rank-1 terms) while BTD splits data in "molecules" (made of atoms). CPD owes much of its success to rank-1 terms that capture the essence of components that are in fact more complex and better represented by a block term. That is, it is unlikely that real-life data components (e.g., of natural images) have all columns proportional, all rows proportional, and so on, like a rank-1 term; low (multilinear) rank, as in BTD, is more likely. In addition, block terms may model multidimensional sources, variations around mean activity, mildly nonlinear phenomena, drifts of setting points, frequency shifts, mildly convolutive mixtures, and so on. Efficient variants may become a true alternative to current mainstream algorithms for signal separation.

Bibliography

[1] L. De Lathauwer. Decompositions of a Higher-Order Tensor in Block Terms—Part I: Lemmas for Partitioned Matrices. *SIAM J. Matrix Anal. Appl.*, 30(3):1022–1032, 2008.

[2] I. Domanov and L. De Lathauwer. Enhanced line search for blind channel identification based on the parafac decomposition of cumulant tensors. *in Proc. of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010), Budapest, Hungary*, pages 1001–1002, 2010.

[3] L. Sorber, I. Domanov, M. Van Barel, and L. De Lathauwer. Exact Line and Plane Search for Tensor Optimization by Global Minimization of Bivariate Polynomials and Rational Functions. *In preparation.*

[4] L. Sorber, M. Van Barel, and L. De Lathauwer. Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-(L_r,L_r,1) terms and a new generalization. *SIAM Journal on Optimization (accepted)*, 2013.

Appendix to Chapter 4

This Appendix is based on
Domanov, I., De Lathauwer, L. Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition, Supplementary materials, submitted to SIAM Journal on Matrix Analysis and Applications (accepted with major revision).

A.1 Supplementary material related to Proposition 4.1.10

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

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

where

$$\mathbf{L} := \begin{bmatrix} 0 & 0 & \dots & (-1)^{K-1} \\ \vdots & \vdots & \iddots & \vdots \\ 0 & -1 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{bmatrix}. \tag{A.2}$$

We start with a trivial Lemma.

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

(i)

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

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

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

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

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

$$\mathbf{y}^T\mathbf{L}\mathcal{C}_{K-1}\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}}\end{bmatrix} = \det\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}} & \mathbf{y}\end{bmatrix}, \qquad \mathbf{y} \in \mathbb{R}^K. \quad \text{(A.3)}$$

Now, the statement (i) follows from (A.3) by setting $\mathbf{y} = \mathbf{x}$.

(ii) Since the vectors $\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_{K-1}}$ are linearly independent in $\mathbb{R}^K$, it follows that there exists a vector $\mathbf{y}$ such that $\det\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}} & \mathbf{y}\end{bmatrix} \neq 0$. Hence, by (A.3), the $(i_1, \dots, i_{K-1})$-th column of $\mathcal{B}(\mathbf{C})$ is nonzero.

(iii) follows from (A.3) and the fact that $\det\begin{bmatrix}\mathbf{c}_{i_1} & \dots & \mathbf{c}_{i_{K-1}} & \mathbf{y}\end{bmatrix} = 0$ if and only if $\mathbf{y} \in \operatorname{span}\{\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_{K-1}}\}$.

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

Lemma A.1.2. *Let* $\mathbf{C} \in \mathbb{R}^{K \times R}$*,* $K \leq R$*, and* $k_{\mathbf{C}} = K$*. Let also* $\mathcal{B}(\mathbf{C})$ *be defined by* (A.1)*. Then*

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

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

(iii) *the matrix*

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

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

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

$$\mathbf{C}^T\mathcal{B}(\mathbf{C}) = \mathbf{C}^T\mathbf{L}\mathcal{C}_{K-1}(\mathbf{C}) =$$

$$\begin{bmatrix} \det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{c}_1\end{bmatrix} & \dots & \det\begin{bmatrix}\mathbf{c}_{R-K+2} & \dots & \mathbf{c}_R & \mathbf{c}_1\end{bmatrix} \\ \vdots & & \vdots & & \vdots \\ \det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{c}_R\end{bmatrix} & \dots & \det\begin{bmatrix}\mathbf{c}_{R-K+2} & \dots & \mathbf{c}_R & \mathbf{c}_R\end{bmatrix} \end{bmatrix}. \tag{A.4}$$

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

$$\det\begin{bmatrix}\mathbf{c}_{j_1} & \dots & \mathbf{c}_{j_{K-1}} & \mathbf{c}_r\end{bmatrix} = 0 \Leftrightarrow r \in \{j_1, \dots, j_{K-1}\}. \tag{A.5}$$

The results now easily follow from (A.4)-(A.5).

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

Suppose that $\mathcal{B}(\mathbf{C})^{(R-K+1)}\widehat{\mathbf{t}} = \mathbf{0}$ for $\widehat{\mathbf{t}} = \begin{bmatrix}t_{(1,\dots,K-1)} & \dots & t_{(R-K+2,\dots,R)}\end{bmatrix} \in \mathbb{R}^{C_R^{K-1}}$. We show that $t_{(j_1,\dots,j_{K-1})} = 0$ for all $(j_1, \dots, j_{K-1}) \in S_R^{K-1}$. We fix $(j_1, \dots, j_{K-1}) \in S_R^{K-1}$ and set $\{i_1, \dots, i_{R-K+1}\} = \{1, \dots, R\} \setminus \{j_1, \dots, j_{K-1}\}$. Since $k_{\mathbf{C}} = K$, (A.5) holds. In particular,

$$\alpha_{(j_1,\dots,j_{K-1})} := \prod_{l=1}^{R-K+1} \det\begin{bmatrix}\mathbf{c}_{j_1} & \dots & \mathbf{c}_{j_{K-1}} & \mathbf{c}_{i_l}\end{bmatrix} \neq 0. \tag{A.6}$$

Let $\mathbf{f} := \mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_{R-K+1}}$. Then by (A.5), we have

$$0 = \mathbf{f}^T\mathbf{0} = \mathbf{f}^T\left(\mathcal{B}(\mathbf{C})^{(R-K+1)}\widehat{\mathbf{t}}\right) =$$

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

$$\Bigg[\prod_{l=1}^{R-K+1} \det\begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_{K-1} & \mathbf{c}_{i_l}\end{bmatrix} \dots$$

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

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

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

Hence, by (A.6), $t_{(j_1,\dots,j_{K-1})} = 0$. Thus, $\widehat{\mathbf{t}} = \mathbf{0}$. Therefore, the matrix $\mathcal{B}(\mathbf{C})^{(R-K+1)}$ has full column rank. □

Proof of Proposition 4.1.10.

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

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

A.2 Supplementary material related to properties (P1)–(P4)

We will say that condition (K,R) holds if for any $K \times R$ matrix $\mathbf{C}$ with $k_{\mathbf{C}} = K$ and for any nonzero vector $\mathbf{x} \in \mathbb{R}^K$, the implication in the following scheme holds

$$\omega(\mathbf{x}^T\mathcal{B}(\mathbf{C})) \le C_R^{K-1} - C_{R-1}^{K-2} = C_{R-1}^{K-1} \Leftrightarrow$$

$$\mathbf{x} \text{ is orthogonal to at least } C_{R-1}^{K-2} \text{ columns of } B(\mathbf{C}) \Rightarrow \quad \text{(A.7)}$$

$$\mathbf{x} \text{ is proportional to a column of } \mathbf{C}.$$

Note that the equivalence "⇔" and the implication opposite to "⇒" in (A.7) follow from the definition of $\omega(\cdot)$ and Lemma A.1.2 (ii), respectively. It can be easily checked that (2,R) holds for $R \ge 2$ and that (K,K) holds. Our goal is to show that (K,R) holds for $R \ge K$. We need the following lemma.

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

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

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

$$\mathcal{J}_1 := \{(1, j_2,\dots,j_{K-1}) :\ (1, j_2,\dots,j_{K-1}) \in \mathcal{J}\}, \qquad \overline{\mathcal{J}}_1 := \mathcal{J} \setminus \mathcal{J}_1.$$

Then by Lemma A.1.1 (i) and (A.7), card $\mathcal{J} \ge C_{R-1}^{K-2}$. We consider two cases: card $\overline{\mathcal{J}}_1 \ge C_{R-2}^{K-2}$ and card $\overline{\mathcal{J}}_1 < C_{R-2}^{K-2}$. If card $\overline{\mathcal{J}}_1 \ge C_{R-2}^{K-2}$, then

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

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

$$\text{card } \mathcal{J}_1 = \text{card } \mathcal{J} - \text{card } \overline{\mathcal{J}}_1 > C_{R-1}^{K-2} - C_{R-2}^{K-2} = C_{R-2}^{K-3}. \tag{A.8}$$

By Lemma A.1.1 (i) and (A.7), there exist numbers $\alpha_{p_1}, \alpha_{p_{j_2}}, \dots, \alpha_{p_{j_{K-1}}}$ such that

$$\mathbf{x} = \alpha_{p_1}\mathbf{c}_1 + \sum_{q=2}^{K-1} \alpha_{p_{j_q}}\mathbf{c}_{j_q}, \qquad p \in \{1, \dots, \text{card } \mathcal{J}_1\}, \qquad (1, j_2, \dots, j_{K-1}) \in \mathcal{J}_1. \tag{A.9}$$

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

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

If $\mathbf{T}\mathbf{x} = \mathbf{0}$, then either $\mathbf{x} = \mathbf{0}$, or $\mathbf{x}$ is proportional to $\mathbf{c}_1$. Hence, we can assume that $\widetilde{\mathbf{x}}$ is a nonzero vector. Since $k_{\mathbf{C}} = K$ and $\ker(\mathbf{T}) = \text{span}\,\{\mathbf{c}_1\}$, it follows that $k_{\widetilde{\mathbf{C}}} = K - 1$. Let us apply $\mathbf{T}$ to (A.9)

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

Hence,

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

By Lemma A.1.1 (i) and (A.8),

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

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

Lemma A.2.2. *Let $R \ge K$. Then condition* (K,R) *holds.*

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

conditions (k,k+1) and (k+1,k+1) imply condition (k+1,k+2),

conditions (k,k+2) and (k+1,k+2) imply condition (k+1,k+3),

$$\vdots$$

conditions (k,R-1) and (k+1,R-1) imply condition (k+1,R). □

Proof of properties (P1)–(P4).

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

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

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

(P4) We prove that

$$\mathbf{x} \text{ is orthogonal to } C_{R-1}^{K-2} \text{ columns of } B(\mathbf{C}) \Leftrightarrow$$

$$\mathbf{x} \text{ is proportional to a column of } \mathbf{C}.$$

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

□

A.3 Supplementary material related to Lemma 4.2.17

Proof of Lemma 4.2.17.
(i) From the definitions of the matrices $\mathcal{Q}_m(\mathbf{C})^T$ and $\mathcal{R}_m(\mathbf{C})^T$ (see Definitions 4.2.9 and 4.2.10) it follows that the matrix $\mathcal{Q}_m(\mathbf{C})^T$ is obtained from $\mathcal{R}_m(\mathbf{C})^T$ be removing columns that are repeated. Thus, (i) just expresses the following fact: $(j_1, \ldots, j_m)$-th column of the matrix $\mathcal{Q}_m(\mathbf{C})^T$ coincides with the $(l_1, \ldots, l_m)$-th column of the matrix $\mathcal{R}_m(\mathbf{C})^T$ whenever $(l_1, \ldots, l_m) \in P_{\{j_1,\ldots,j_m\}}$.

(ii) From Lemma 4.2.15 and definition of the matrix $\mathbf{G}$ (4.22) it follows that the $((i_1, \ldots, i_m), (j_1, \ldots, j_m))$-th entry of the matrix $\mathcal{R}_m(\mathbf{C})^T\mathbf{G}$ is equal to

$$m! \left(\pi_S(\mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_m})\right)^T \left(\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K)\right) =$$

$$\frac{1}{m!} \sum_{(q_1,\ldots,q_m)\in P_{\{i_1,\ldots,i_m\}}} \sum_{(l_1,\ldots,l_m)\in P_{\{j_1,\ldots,j_m\}}} (\mathbf{c}_{q_1}^T \mathbf{e}_{l_1}^K) \cdot \cdots \cdot (\mathbf{c}_{q_m}^T \mathbf{e}_{l_m}^K) =$$

$$\frac{1}{m!} \sum_{(q_1,\ldots,q_m)\in P_{\{i_1,\ldots,i_m\}}} \text{perm } \mathbf{C}((j_1, \ldots, j_m), (q_1, \ldots, q_m)).$$

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

$$\left(\pi_S(\mathbf{c}_{i_1} \otimes \cdots \otimes \mathbf{c}_{i_m})\right)^T \left(\pi_S(\mathbf{e}_{j_1}^K \otimes \cdots \otimes \mathbf{e}_{j_m}^K)\right) =$$

$$\text{perm } \mathbf{C}((j_1, \ldots, j_m), (i_1, \ldots, i_m)).$$

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

A.4 Supplementary material related to Lemma 4.4.4

We need the following Lemma.

Lemma A.4.1. *Let $\widehat{\mathbf{A}}$ be any set of $\widehat{m}$ columns of $\mathbf{A}$, let $\widehat{\mathbf{B}}$ be the corresponding set of columns of $\mathbf{B}$. Assume that the matrix $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank. Then*

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

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

Proof. Since $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ has full column rank it follows that all columns of $\mathcal{C}_m(\mathbf{A})$ and $\mathcal{C}_m(\mathbf{B})$ are nonzero. Hence, (i) follows from Lemma 4.2.3 (1). If $\widehat{m} \leq m$, then (ii) follows from (i). If $\widehat{m} > m$, then, by Lemma 4.2.5, $r_{\mathcal{C}_m(\widehat{\mathbf{A}})} = C^m_{r_{\widehat{\mathbf{A}}}}$ and $r_{\mathcal{C}_m(\widehat{\mathbf{B}})} = C^m_{r_{\widehat{\mathbf{B}}}}$. Hence,

$$C^m_{r_{\widehat{\mathbf{A}}}} C^m_{r_{\widehat{\mathbf{B}}}} = r_{\mathcal{C}_m(\widehat{\mathbf{A}})} r_{\mathcal{C}_m(\widehat{\mathbf{B}})} = r_{\mathcal{C}_m(\widehat{\mathbf{A}}) \otimes \mathcal{C}_m(\widehat{\mathbf{B}})} \geq r_{\mathcal{C}_m(\widehat{\mathbf{A}}) \odot \mathcal{C}_m(\widehat{\mathbf{B}})} = C^m_{\widehat{m}} > 1, \quad \text{(A.10)}$$

where the last equality holds since the matrix $\mathcal{C}_m(\widehat{\mathbf{A}}) \odot \mathcal{C}_m(\widehat{\mathbf{B}})$ has full column rank. The statement (ii) for $\widehat{m} > m$ now follows from (A.10). □

Proof of Lemma 4.4.4
Without loss of generality we may assume that $\mathbf{F}$ coincides with $\mathcal{B}(\mathbf{C})$. In the proof we will associate indices $i, j \in \{1, \ldots, C_R^{K-1}\}$ with their multi-index analogues $(i_1, \ldots, i_{K-1}), (j_1, \ldots, j_{K-1}) \in S_R^{K-1}$. By definition set

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

Since

$$\operatorname{card}\{\{i_1, \ldots, i_{K-1}\} \cap \{j_1, \ldots, j_{K-1}\}\} \leq K-2,$$

$$\{p_1, \ldots, p_{\widehat{m}}\} \subset \{\{1, \ldots, R\} \setminus \{i_1, \ldots, i_{K-1}\}\} \cup \{\{1, \ldots, R\} \setminus \{j_1, \ldots, j_{K-1}\}\}$$

it follows that

$$m = R - (K-2) \leq \widehat{m}, \quad \text{(A.11)}$$

$$\widehat{m} \leq \min(R, 2(R-(K-1))) = \min(R, 2m-2). \quad \text{(A.12)}$$

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

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

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

$$\mathbf{V}_i = \mathbf{A}\mathrm{Diag}(\mathbf{y}_i)\mathbf{B}^T = \widehat{\mathbf{A}}\mathrm{Diag}(\widehat{\mathbf{y}}_i)\widehat{\mathbf{B}}^T, \quad \mathbf{V}_j = \mathbf{A}\mathrm{Diag}(\mathbf{y}_j)\mathbf{B}^T = \widehat{\mathbf{A}}\mathrm{Diag}(\widehat{\mathbf{y}}_j)\widehat{\mathbf{B}}^T. \tag{A.13}$$

Hence,

$$[\mathbf{V}_i \ \mathbf{V}_j] = \widehat{\mathbf{A}}[\mathrm{Diag}(\widehat{\mathbf{y}}_i)\widehat{\mathbf{B}}^T \ \mathrm{Diag}(\widehat{\mathbf{y}}_j)\widehat{\mathbf{B}}^T] = \widehat{\mathbf{A}}(\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}})^T, \tag{A.14}$$

$$[\mathbf{V}_i^T \ \mathbf{V}_j^T] = \widehat{\mathbf{B}}[\mathrm{Diag}(\widehat{\mathbf{y}}_i)\widehat{\mathbf{A}}^T \ \mathrm{Diag}(\widehat{\mathbf{y}}_j)\widehat{\mathbf{A}}^T] = \widehat{\mathbf{B}}(\widehat{\mathbf{C}} \odot \widehat{\mathbf{A}})^T, \tag{A.15}$$

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

$$\text{the matrices } \ \widehat{\mathbf{C}} \odot \widehat{\mathbf{B}} \ \text{ and } \ \widehat{\mathbf{C}} \odot \widehat{\mathbf{A}} \ \text{ have full column rank.} \tag{A.16}$$

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

$$(\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}})\mathbf{P} = (\widehat{\mathbf{C}}\mathbf{P}) \odot (\widehat{\mathbf{B}}\mathbf{P}) = \begin{bmatrix} * & \ldots & * & * & \ldots & * & 0 & \ldots & 0 \\ 0 & \ldots & 0 & * & \ldots & * & * & \ldots & * \end{bmatrix} \odot (\widehat{\mathbf{B}}\mathbf{P}) =$$

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

where $*$ denotes a nonzero value, $\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_{\widetilde{m}}$ denote the columns of the matrix $\widehat{\mathbf{B}}\mathbf{P}$, and where we use the dimensionality constraints in (A.11)–(A.12). By Lemma A.4.1 (i), $k_{\widehat{\mathbf{B}}} \geq k_{\mathbf{B}} \geq m$. Hence, the matrix $(\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}})\mathbf{P}$ has full column rank. Since the matrix $\mathbf{P}$ is nonsingular, it follows that the matrix $\widehat{\mathbf{C}} \odot \widehat{\mathbf{B}}$ also has full column rank. In a similar fashion one can prove that the matrix $\widehat{\mathbf{C}} \odot \widehat{\mathbf{A}}$ has full column rank. From (A.14)–(A.16) it follows that $r_{[\mathbf{V}_i \ \mathbf{V}_j]} = r_{\widehat{\mathbf{A}}}$ and $r_{[\mathbf{V}_i^T \ \mathbf{V}_j^T]} = r_{\widehat{\mathbf{B}}}$. Using Lemma A.4.1 we have

$$\widehat{m} \geq \max(r_{[\mathbf{V}_i \ \mathbf{V}_j]}, r_{[\mathbf{V}_i^T \ \mathbf{V}_j^T]}) = \max(r_{\widehat{\mathbf{A}}}, r_{\widehat{\mathbf{B}}}) \geq \min(\widehat{m}, m+1), \tag{A.17}$$

$$\min(r_{[\mathbf{V}_i \ \mathbf{V}_j]}, r_{[\mathbf{V}_i^T \ \mathbf{V}_j^T]}) = \min(r_{\widehat{\mathbf{A}}}, r_{\widehat{\mathbf{B}}}) \geq \min(k_{\widehat{\mathbf{A}}}, k_{\widehat{\mathbf{B}}}) \geq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq m. \tag{A.18}$$

The equivalence (ii) $\Leftrightarrow \widehat{m} = m$ now easily follows from (A.17)–(A.18).

(iii) $\Leftrightarrow \widehat{m} = m$: by (A.13), $\mathcal{V}_{ij} = [\mathbf{A}, \mathbf{B}, [\mathbf{y}_i\ \mathbf{y}_j]^T]_R = [\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}]_{\widehat{m}}$. Hence, $\widehat{m} \geq r_{\mathcal{V}_{ij}}$. On the other hand, from (A.14)-(A.15) it follows that $r_{\mathcal{V}_{ij}} \geq \max(r_{[\mathbf{V}_i\ \mathbf{V}_j]}, r_{[\mathbf{V}_i^T\ \mathbf{V}_j^T]})$. Hence, by (A.17), $r_{\mathcal{V}_{ij}} \geq \min(\widehat{m}, m+1)$. The equivalence (iii) $\Leftrightarrow \widehat{m} = m$ now follows from the inequalities $\widehat{m} \geq r_{\mathcal{V}_{ij}} \geq \min(\widehat{m}, m+1)$. □

A.5 Supplementary material related to Example 4.4.6

The full matrix $\mathbf{Q}_3(\mathcal{T})$ is given by

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

List of publications

Articles in internationally reviewed journals

[1] Domanov, I., De Lathauwer, L. On the uniqueness of the canonical polyadic decomposition of third-order tensors - Part I: Basic results and uniqueness of one factor matrix. SIAM Journal on Matrix Analysis and Applications, 34-3 (2013), pp. 855-875.

[2] Domanov, I., De Lathauwer, L. On the uniqueness of the canonical polyadic decomposition of third-order tensors - Part II: Uniqueness of the overall decomposition. SIAM Journal on Matrix Analysis and Applications, 34-3 (2013), pp. 876-903.

[3] Domanov, I., De Lathauwer, L. Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition, submitted to SIAM Journal on Matrix Analysis and Applications (accepted with major revision).

[4] Domanov, I. (2010). On invariant subspaces of matrices : a new proof of a theorem of Halmos. Linear Algebra and Its Applications, 433(11-12), 2255-2256.

Papers at international conferences and symposia, published in full in proceedings

[1] Domanov, I., De Lathauwer, L. Blind channel identification of MISO systems based on the CP decomposition of cumulant tensors. In : Proc. of the 2011 European Signal Processing Conference (EUSIPCO 2011), Barcelona, Spain, Aug-Sept, 2011, pp. 2215–2218.

[2] Domanov, I., De Lathauwer, L. Enhanced line search for blind channel identification based on the PARAFAC decomposition of cumulant tensors.

In : Proc. of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010), Budapest, Hungary, Jul. 2010, pp. 1001–1002.

Meeting abstracts, presented at international conferences and symposia, published or not published in proceedings or journals

[1] Sorber, L., Domanov, I., Van Barel, M., De Lathauwer, L. (2013). Exact Line and Plane Search for Tensor Optimization by Global Minimization of Bivariate Polynomials and Rational Functions. SIAM Conference on Computational Science and Engineering. Boston, Massachusetts, USA, February 25-March 1 2013.

[2] Domanov, I., De Lathauwer, L. (2012). On the uniqueness of the canonical polyadic decomposition : new sufficient conditions based on properties of Khatri-Rao Products of Compound Matrices. TRICAP 2012. Bruges, Belgium, Jun. 2012.

[3] De Lathauwer, L., Domanov, I. (2012). Canonical polyadic decomposition and block term decompositions : uniqueness and signal separation. Int. Conf. Spectral Theory of Tensors. Tianjin, China, Jun. 2012.

[4] Domanov, I., De Lathauwer, L. (2011). On the (generic) uniqueness of the canonical polyadic decomposition of third-order tensors. Workshop on matrix equations and tensor techniques. Aachen, Germany, Nov. 2011.

[5] Domanov, I., De Lathauwer, L. (2011). On the uniqueness of the canonical polyadic decomposition of third-order tensors : new sufficient conditions. Int. Conf. on Matrix Methods in Mathematics and applications. Moscow, Russia, Jun. 2011.

[6] Domanov, I., De Lathauwer, L. (2011). On the uniqueness of the canonical polyadic decomposition (CPD) and the link between CPD and generalized Schur and Oppenheim inequalities. Pure and Applied Linear Algebra : the new generation. Braunschweig, Germany, Aug. 2011.

[7] Domanov, I., De Lathauwer, L. (2010). An enhanced plane search scheme for complex-valued tensor decompositions. 16th Conference of the International Linear Algebra Society (ILAS). Pisa, Italy, Jun. 2010.

[8] Domanov, I., De Lathauwer, L. (2010). Enhanced line search for blind channel identification based on the parafac decompositioin of cumulant tensors. 16th Conference of the International Linear Algebra Society (ILAS). Pisa, Italy, Jun. 2010.

[9] Domanov, I., De Lathauwer, L. (2010). Higher order cumulant based blind channel identification : Enhanced line search solutions. In : Book of Abstracts. 29th Benelux meeting on Systems and Control. Heeze, The Netherlands, March 30 - Apr. 1, 2010 (164).

[10] Domanov, I., De Lathauwer, L. (2010). On fixed point algorithms for the blind estimation of a single-input single-output fintie impulse response system. Workshop on tensor decompositions and applications. Monopoli, Bari, Italy, Sep. 2010.

[11] Domanov, I., De Lathauwer, L. (2009). Enhanced line search for blind channel identification based on the parafac decompositioin of cumulant tensors. PAI VI/4 Dysco study Day. Leuven, Belgium, Nov. 2009.



KU LEUVEN - KULAK
GROUP SCIENCE, ENGINEERING AND TECHNOLOGY
ELECTRICAL ENGINEERING (ESAT), KULAK
E. Sabbelaan 53
B-8500 Kortrijk