REVIEW　　　　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Overview of constrained PARAFAC models

Gérard Favier[1*†] and André LF de Almeida[2†]

Abstract

In this paper, we present an overview of constrained parallel factor (PARAFAC) models where the constraints model linear dependencies among columns of the factor matrices of the tensor decomposition or, alternatively, the pattern of interactions between different modes of the tensor which are captured by the equivalent core tensor. Some tensor prerequisites with a particular emphasis on mode combination using Kronecker products of canonical vectors that makes easier matricization operations, are first introduced. This Kronecker product-based approach is also formulated in terms of an index notation, which provides an original and concise formalism for both matricizing tensors and writing tensor models. Then, after a brief reminder of PARAFAC and Tucker models, two families of constrained tensor models, the co-called PARALIND/CONFAC and PARATUCK models, are described in a unified framework, for Nth-order tensors. New tensor models, called nested Tucker models and block PARALIND/CONFAC models, are also introduced. A link between PARATUCK models and constrained PARAFAC models is then established. Finally, new uniqueness properties of PARATUCK models are deduced from sufficient conditions for essential uniqueness of their associated constrained PARAFAC models.

Keywords: Constrained PARAFAC; PARALIND/CONFAC; PARATUCK; Tensor models; Tucker models

1 Review

1.1 Introduction

Tensor calculus was introduced in differential geometry, at the end of the nineteenth century, and then tensor analysis was developed in the context of Einstein's theory of general relativity, with the introduction of index notation, the so-called Einstein summation convention, at the beginning of the twentieth century, which allows to simplify and shorten physics equations involving tensors. Index notation is also useful for simplifying multivariate statistical calculations, particularly those involving cumulant tensors [1]. Generally speaking, tensors are used in physics and differential geometry for characterizing the properties of a physical system, representing fundamental laws of physics, and defining geometrical objects whose components are functions. When these functions are defined over a continuum of points of a mathematical space, the tensor forms what is called a tensor field, a generalization of vector field used to solve problems involving curved surfaces or spaces, as it is the case of

curved space-time in general relativity. From a mathematical point of view, two other approaches are possible for defining tensors, in terms of tensor products of vector spaces, or multilinear maps. Symmetric tensors can also be linked with homogeneous polynomials [2].

After the first tensor developments by mathematicians and physicists, the need of analyzing collections of data matrices that can be seen as three-way data arrays gave rise to three-way models for data analysis, with the pioneering works of Tucker in psychometrics [3], and Harshman in phonetics [4], who proposed what is now referred to as the Tucker and parallel factor (PARAFAC) decompositions, respectively. The PARAFAC decomposition was independently proposed by Carroll and Chang [5] under the name canonical decomposition (CANDECOMP) and then called CANDECOMP/PARAFAC (CP) in [6]. For a history of the development of multi-way models in the context of data analysis, see [7]. Since the 1990s, multi-way analysis has known a growing success in chemistry and especially in chemometrics (see Bro's thesis [8] and the book by Smilde et al. [9] for a description of various chemical applications of three-way models, with a pedagogical presentation of these models and of various algorithms for estimating their parameters). At the same

*Correspondence: favier@i3s.unice.fr
†Equal contributors
[1]I3S Laboratory, University of Nice-Sophia Antipolis, CNRS, 2000 route des Lucioles, Les Algorithmes-B, 06903 Sophia Antipolis, France
Full list of author information is available at the end of the article

period, tensor tools were developed for signal processing applications, more particularly for solving the so-called blind source separation (BSS) problem using cumulant tensors (see [10-12] and De Lathauwer's thesis [13] where the concept of high-order singular value decomposition (HOSVD) is introduced, a tensor tool generalizing the standard matrix SVD to arrays of order higher than two). A recent overview of BSS approaches and applications can be found in the handbook co-edited by Comon and Jutten [14].

Nowadays, (high-order) tensors, also called multi-way arrays in the data analysis community, play an important role in many fields of application for representing and analyzing multidimensional data, as in psychometrics, chemometrics, food industry, environmental sciences, signal/image processing, computer vision, neuroscience, information sciences, data mining, pattern recognition, among many others. Then, they are simply considered as multidimensional arrays of numbers, constituting a generalization of vectors and matrices that are first- and second-order tensors, respectively, to orders higher than two. Tensor decompositions, also called tensor models, are very useful for analyzing multidimensional data under the form of signals, images, speech, music sequences, or texts and also for designing new systems as it is the case of wireless communication systems since the publication of the seminal paper by Sidiropoulos et al. [15]. Besides the references already cited, overviews of tensor tools, models, algorithms, and applications can be found in [16-19].

Tensor models incorporating constraints (sparsity; non-negativity; smoothness; symmetry; column orthonormality of factor matrices; Hankel, Toeplitz, and Vandermonde structured matrix factors; allocation constraints...) have been the object of intensive works, during the last years. Such constraints can be inherent to the problem under study or the result of a system design. An overview of constraints on components of tensor models most often encountered in multi-way data analysis can be found in [7]. Incorporation of constraints in tensor models may facilitate physical interpretability of matrix factors. Moreover, imposing constraints may allow to relax uniqueness conditions and to develop specialized parameter estimation algorithms with improved performance both in terms of accuracy and computational cost, as it is the case of CP models with a column-wise orthonormal factor matrix [20]. One can classify the constraints into three main categories: i) sparsity/non-negativity, ii) structural, and iii) linear dependencies/mode interactions. It is worth noting that the three categories of constraints involve specific parameter estimation algorithms, the first two generally inducing an improvement of uniqueness property of the tensor decomposition, while the third category implies a reduction of uniqueness, named partial uniqueness. We

briefly review the main results concerning the first two types of constraints, Section 1.3 of this paper being dedicated to the third category.

Sparse and non-negative tensor models have recently been the subject of many works in various fields of applications like computer vision [21,22], image compression [23], hyperspectral imaging [24], music genre classification [25] and audio source separation [26], multi-channel EEG (electroencephalography) and network traffic analysis [27], fluorescence analysis [28], data denoising and image classification [29], among many others. Two non-negative tensor models have been more particularly studied in the literature, the so-called non-negative tensor factorization (NTF), i.e., PARAFAC models with non-negativity constraints on the matrix factors, and non-negative Tucker decomposition (NTD), i.e., Tucker models with non-negativity constraints on the core tensor and/or the matrix factors. The crucial importance of NTF/NTD for multi-way data analysis applications results from the very large volume of real-world data to be analyzed under constraints of sparseness and non-negativity of factors to be estimated, when only non-negative parameters are physically interpretable. Many NTF/NTD algorithms are now available. Most of them can be viewed as high-order extensions of non-negative matrix factorization (NMF) methods, in the sense that they are based on an alternating minimization of cost functions incorporating sparsity measures (also named distances or divergences) with application of NMF methods to matricized or vectorized forms of the tensor to be decomposed (see for instance [16,23,28,30] for NTF and [29,31] for NTD). An overview of NMF and NTF/NTD algorithms can be found in [16].

The second category of constraints concerns the case where the core tensor and/or some matrix factors of the tensor model have a special structure. For instance, we recently proposed a nonlinear CDMA scheme for multiuser SIMO communication systems that is based on a constrained block-Tucker2 model whose core tensor, composed of the information symbols to be transmitted and their powers up to a certain degree, is characterized by matrix slices having a Vandermonde or a Hankel structure [32,33]. We also developed Volterra-PARAFAC models for nonlinear system modeling and identification. These models are obtained by expanding high-order Volterra kernels, viewed as symmetric tensors, by means of symmetric or doubly symmetric PARAFAC decompositions [34,35]. Block structured nonlinear systems like Wiener, Hammerstein, and parallel-cascade Wiener systems can be identified from their associated Volterra kernels that admit symmetric PARAFAC decompositions with Toeplitz factors [36,37]. Symmetric PARAFAC models with Hankel factors and symmetric block PARAFAC models with block Hankel factors are encountered for

blind identification of multiple-input multiple-output (MIMO) linear channels using fourth-order cumulant tensors, in the cases of memoryless and convolutive channels, respectively [38,39]. In the presence of structural constraints, specific estimation algorithms can be derived as it is the case for symmetric CP decompositions [40], CP decompositions with Toeplitz factors (in [41], an iterative solution was proposed, whereas in [42], a non-iterative algorithm was developed), Vandermonde factors [43], circulant factors [44], banded and/or structured matrix factors [45,46], and also for Hankel and Vandermonde structured core tensors [33].

The rest of this paper is organized as follows: In Section 1.2, we present some tensor prerequisites with a particular emphasis on mode combination using Kronecker products of canonical vectors that makes easier the matricization operations, especially to derive matrix representations of tensor models. This Kronecker product-based approach is also formulated in terms of an index notation, which provides an original and concise formalism for both matricizing tensors and writing tensor models. Then, we present the two most common tensor models, the so-called Tucker and PARAFAC models, in a general framework, i.e., for Nth-order tensors. In Section 1.3, two families of constrained tensor models, the co-called PARALIND/CONFAC and PARATUCK models, are described in a unified way, with a generalization to Nth-order tensors. New tensor models, called nested Tucker models and block PARALIND/CONFAC models, are also introduced. A link between PARATUCK models and constrained PARAFAC models is also established. In Section 1.4, uniqueness properties of PARATUCK models are deduced using this link. The paper is concluded in Section 2.

Notations and definitions. $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and complex numbers, respectively. Scalars, column vectors, matrices, and high-order tensors are denoted by lowercase, boldface lowercase, boldface uppercase, and calligraphic uppercase letters, e.g., a, $\mathbf{a}$, $\mathbf{A}$, and $\mathcal{A}$, respectively. The vector $\mathbf{A}_{i.}$ (resp. $\mathbf{A}_{.j}$) represents the ith row (resp. jth column) of $\mathbf{A}$.

$\mathbf{I}_N$, $\mathbf{1}_N^T$, and $\mathbf{e}_n^{(N)}$ stand for the identity matrix of order N, the all-ones row vector of dimensions $1 \times N$, and the nth canonical vector of the Euclidean space $\mathbb{R}^N$, respectively.

$\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$, $\mathrm{tr}(\mathbf{A})$, and $r_\mathbf{A}$ denote the transpose, the conjugate (Hermitian) transpose, the Moore-Penrose pseudo-inverse, the trace, and the rank of $\mathbf{A}$, respectively. $D_i(\mathbf{A}) = \mathrm{diag}(\mathbf{A}_{i.})$ represents the diagonal matrix having the elements of the ith row of $\mathbf{A}$ on its diagonal. The operator bdiag(.) forms a block diagonal matrix from its matrix arguments, while the operator vec(.) transforms a matrix into a column vector by stacking the columns of its matrix argument one on top of the other one. In case of a tensor $\mathcal{X}$, the vec(.) operation is defined in (6).

The outer product (also called tensor product), and the matrix Kronecker, Khatri-Rao (column-wise Kronecker), and Hadamard (element-wise) products are denoted by $\circ$, $\otimes$, $\diamond$, and $\odot$, respectively.

Let us consider the set $\mathbb{S} = \{n_1, \ldots, n_N\}$ obtained by permuting the elements of the set $\{1, \ldots, N\}$. For $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ and $\mathbf{u}^{(n)} \in \mathbb{C}^{I_n \times 1}$, $n = 1, \ldots, N$, we define

$$\underset{n \in \mathbb{S}}{\otimes} \mathbf{A}^{(n)} = \mathbf{A}^{(n_1)} \otimes \mathbf{A}^{(n_2)} \otimes \cdots \otimes \mathbf{A}^{(n_N)} \in \mathbb{C}^{I_{n_1} \cdots I_{n_N} \times R_{n_1} \cdots R_{n_N}};$$
(1)

$$\underset{n \in \mathbb{S}}{\diamond} \mathbf{A}^{(n)} = \mathbf{A}^{(n_1)} \diamond \mathbf{A}^{(n_2)} \diamond \cdots \diamond \mathbf{A}^{(n_N)} \in \mathbb{C}^{I_{n_1} \cdots I_{n_N} \times R},$$

$$\text{when } R_n = R, \forall n = 1, \ldots, N;$$
(2)

$$\underset{n \in \mathbb{S}}{\odot} \mathbf{A}^{(n)} = \mathbf{A}^{(n_1)} \odot \mathbf{A}^{(n_2)} \odot \cdots \odot \mathbf{A}^{(n_N)} \in \mathbb{C}^{I \times R},$$

$$\text{when } I_n = I, \text{ and } R_n = R, \forall n = 1, \ldots, N;$$

$$\underset{n \in \mathbb{S}}{\circ} \mathbf{u}^{(n)} = \mathbf{u}^{(n_1)} \circ \mathbf{u}^{(n_2)} \circ \cdots \circ \mathbf{u}^{(n_N)} \in \mathbb{C}^{I_{n_1} \times \cdots \times I_{n_N}}.$$

The outer product of N non-zero vectors defines a rank-one tensor of order N.

By convention, the order of dimensions is directly related to the order of variation of the associated indices. For instance, in (1) and (2), the product $I_{n_1} I_{n_2} \cdots I_{n_N}$ of dimensions means that n_1 is the index varying the most slowly while n_N is the index varying the most fastly in the Kronecker products computation.

For $\mathbb{S} = \{1, \ldots, N\}$, we have the following identities:

$$\left(\underset{n \in \mathbb{S}}{\circ} \mathbf{u}^{(n)} \right)_{i_1, \ldots, i_N} = \left(\underset{n=1}{\overset{N}{\circ}} \mathbf{u}^{(n)} \right)_{i_1, \ldots, i_N} = \prod_{n=1}^{N} u_{i_n}^{(n)},$$

$$\left(\underset{n \in \mathbb{S}}{\otimes} \mathbf{u}^{(n)} \right)_i = \left(\underset{n=1}{\overset{N}{\otimes}} \mathbf{u}^{(n)} \right)_i = \prod_{n=1}^{N} u_{i_n}^{(n)}$$

$$\text{with} \quad i = i_N + \sum_{n=1}^{N-1} (i_n - 1) \prod_{j=n+1}^{N} I_j. \quad (3)$$

In particular, for $\mathbf{u} \in \mathbb{C}^{I \times 1}, \mathbf{v} \in \mathbb{C}^{J \times 1}, \mathbf{w} \in \mathbb{C}^{K \times 1}$

$$\mathcal{X} = \mathbf{u} \circ \mathbf{v} \circ \mathbf{w} \in \mathbb{C}^{I \times J \times K} \Leftrightarrow x_{ijk} = u_i v_j w_k,$$

$$\mathbf{x} = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} \in \mathbb{C}^{IJK \times 1} \Leftrightarrow x_{k+(j-1)K+(i-1)JK} = u_i v_j w_k.$$

Some useful matrix formulae are recalled in Appendix 1.

1.2 Tensor prerequisites

In this paper, a tensor is simply viewed as a multidimensional array of measurements. Depending that these measurements are real- or complex-valued, we have a real- or complex-valued tensor, respectively. The order N of a tensor refers to the number of indices that characterize its elements $x_{i_1, \ldots, i_N}$, each index i_n ($i_n = 1, \ldots, I_n$, for $n = 1, \ldots, N$) being associated with a dimension, also called a way, or a mode, and I_n denoting the mode-n dimension.

An Nth-order complex-valued tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, also called an N-way array, of dimensions $I_1 \times \cdots \times I_N$, can be written as

$$\mathcal{X} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\ldots,i_N} \overset{N}{\underset{n=1}{\circ}} \mathbf{e}_{i_n}^{(I_n)}. \tag{4}$$

The coefficients $x_{i_1,\ldots,i_N}$ represent the coordinates of $\mathcal{X}$ in the canonical basis $\left\{ \overset{N}{\underset{n=1}{\circ}} \mathbf{e}_{i_n}^{(I_n)}, i_n = 1, \ldots, I_n; n = 1, \ldots, N \right\}$ of the space $\mathbb{C}^{I_1 \times \cdots \times I_N}$.

The identity tensor of order N and dimensions $I \times \cdots \times I$, denoted by $\mathcal{I}_{N,I}$ or simply $\mathcal{I}$, is a diagonal hypercubic tensor whose elements $\delta_{i_1,\ldots,i_N}$ are defined by means of the generalized Kronecker delta, i.e., $\delta_{i_1,\ldots,i_N} = \begin{cases} 1 & \text{if } i_1 = \cdots = i_N \\ 0 & \text{otherwise} \end{cases}$, and $I_n = I, \forall n = 1, \ldots, N$. It can be written as

$$\mathcal{I}_{N,I} = \sum_{i=1}^{I} \underbrace{\mathbf{e}_i^{(I)} \circ \cdots \circ \mathbf{e}_i^{(I)}}_{N \text{ terms}}.$$

Different reduced order tensors can be obtained by slicing the tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ along one mode or p modes, i.e., by fixing one index i_n or a set of p indices $\{i_{n_1}, \ldots, i_{n_p}\}$, which gives a tensor of order $N-1$ or $N-p$, respectively. For instance, by slicing $\mathcal{X}$ along its mode-n, we get the i_n^{th} mode-n slice of $\mathcal{X}$, denoted by $\mathcal{X}_{\ldots i_n \ldots}$, that can be written as

$$\mathcal{X}_{\ldots i_n \ldots} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_{n-1}=1}^{I_{n-1}} \sum_{i_{n+1}=1}^{I_{n+1}} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\ldots,i_n,\ldots,i_N}$$
$$\times \mathbf{e}_{i_{n+1}}^{(I_{n+1})} \circ \cdots \circ \mathbf{e}_{i_N}^{(I_N)} \circ \mathbf{e}_{i_1}^{(I_1)} \circ \cdots$$
$$\cdots \circ \mathbf{e}_{i_{n-1}}^{(I_{n-1})} \in \mathbb{C}^{I_{n+1} \times \cdots \times I_N \times I_1 \times \cdots \times I_{n-1}}.$$

For instance, by slicing the third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ along each mode, we get three types of matrix slices, respectively called horizontal, lateral, and frontal slices:

$$\mathbf{X}_{i..} \in \mathbb{C}^{J \times K}, \mathbf{X}_{.j.} \in \mathbb{C}^{K \times I} \text{ and } \mathbf{X}_{..k} \in \mathbb{C}^{I \times J},$$
with $i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K.$

1.2.1 Tensor Hadamard product

Consider $\mathcal{A} \in \mathbb{C}^{R_1 \times \cdots \times R_N \times I_1 \times \cdots \times I_{P_1}}$ and $\mathcal{B} \in \mathbb{C}^{R_1 \times \cdots \times R_N \times I_{P_1+1} \times \cdots \times I_P}$, where $\{i_1, \ldots, i_{P_1}\}$ and $\{i_{P_1+1}, \ldots, i_P\}$ are two disjoint ordered subsets of the set of indices $\{i_1, \ldots, i_P\}$ and $\mathbb{R} = \{r_1, \ldots, r_N\}$.

We define the Hadamard product of $\mathcal{A}$ with $\mathcal{B}$ along their common modes, as the tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times \cdots \times R_N \times I_1 \times \cdots \times I_P}$ such that

$$\mathcal{C} = \mathcal{A} \underset{\mathbb{R}}{\odot} \mathcal{B} \Leftrightarrow c_{r_1,\ldots,r_N,i_1,\ldots,i_P} = a_{r_1,\ldots,r_N,i_1,\ldots,i_{P_1}} b_{r_1,\ldots,r_N,i_{P_1+1},\ldots,i_P}$$

For instance, given two third-order tensors $\mathcal{A} \in \mathbb{C}^{R_1 \times R_2 \times I_1}$ and $\mathcal{B} \in \mathbb{C}^{R_1 \times R_2 \times I_2}$, the Hadamard product $\mathcal{A} \underset{\{r_1,r_2\}}{\odot} \mathcal{B}$ gives a fourth-order tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times R_2 \times I_1 \times I_2}$ such that

$$c_{r_1,r_2,i_1,i_2} = a_{r_1,r_2,i_1} b_{r_1,r_2,i_2}.$$

Such a tensor Hadamard product can be calculated by means of the matrix Hadamard product of matrix unfoldings of extended tensors, as defined in (21) and (22) (see also (94) to (96) in Appendix 2). For the example above, we have

$$\mathbf{C}_{R_1 R_2 \times I_1 I_2} = \mathbf{A}_{R_1 R_2 \times I_1} \left(\mathbf{I}_{I_1} \otimes \mathbf{1}_{I_2}^T \right) \odot \mathbf{B}_{R_1 R_2 \times I_2} \left(\mathbf{1}_{I_1}^T \otimes \mathbf{I}_{I_2} \right)$$

Example: For $\mathbf{A}_{R \times I_1} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \mathbf{B}_{R \times I_2} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$, and the tensor $\mathcal{C}$ such as $c_{r,i_1,i_2} = a_{r,i_1} b_{r,i_2}$, a mode-1 flat matrix unfolding of $\mathcal{C}$ is given by

$$\mathbf{C}_{R \times I_1 I_2} = \mathbf{A}_{R \times I_1} \left(\mathbf{I}_2 \otimes \mathbf{1}_2^T \right) \odot \mathbf{B}_{R \times I_2} \left(\mathbf{1}_2^T \otimes \mathbf{I}_2 \right)$$
$$= \begin{bmatrix} a_1 & a_1 & a_2 & a_2 \\ a_3 & a_3 & a_4 & a_4 \end{bmatrix} \odot \begin{bmatrix} b_1 & b_2 & b_1 & b_2 \\ b_3 & b_4 & b_3 & b_4 \end{bmatrix}$$
$$= \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_2 b_1 & a_2 b_2 \\ a_3 b_3 & a_3 b_4 & a_4 b_3 & a_4 b_4 \end{bmatrix}$$

1.2.2 Mode combination

Different contraction operations can be defined depending on the way according to which the modes are combined. Let us partition the set $\{1, \ldots, N\}$ in N_1 ordered subsets $\mathbb{S}_{n_1}$, constituted of $p(n_1)$ elements with $\sum_{n_1=1}^{N_1} p(n_1) = N$. Each subset $\mathbb{S}_{n_1}$ is associated with a combined mode of dimension $J_{n_1} = \prod_{n \in \mathbb{S}_{n_1}} I_n$. These mode combinations allow to rewrite the Nth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ under the form of an N_1^{th}-order tensor $\mathcal{Y} \in \mathcal{C}^{J_1 \times \cdots \times J_{N_1}}$ as follows

$$\mathcal{Y} = \sum_{j_1=1}^{J_1} \cdots \sum_{j_{N_1}=1}^{J_{N_1}} x_{j_1,\ldots,j_{N_1}} \overset{N_1}{\underset{n_1=1}{\circ}} \mathbf{e}_{j_{n_1}}^{(J_{n_1})} \text{ with } \mathbf{e}_{j_{n_1}}^{(J_{n_1})} = \underset{n \in \mathbb{S}_{n_1}}{\otimes} \mathbf{e}_{i_n}^{(I_n)}. \tag{5}$$

Two particular mode combinations corresponding to the vectorization and matricization operations are now detailed.

1.2.3 Vectorization

The vectorization of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is associated with a combination of the N modes into a unique mode of

dimension $J = \prod_{n=1}^{N} I_n$, which amounts to replace the outer product in (4) by the Kronecker product

$$\text{vec}(\mathcal{X}) = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\ldots,i_N} \overset{N}{\underset{n=1}{\otimes}} \mathbf{e}_{i_n}^{(I_n)} \in \mathbb{C}^{I_1 \cdots I_N \times 1} \quad (6)$$

the element $x_{i_1,\ldots,i_N}$ of $\mathcal{X}$ being the ith entry of $\text{vec}(\mathcal{X})$ with i defined as in (3).

The vectorization can also be carried out after a permutation of indices $\pi(i_n), n = 1, \ldots, N$.

1.2.4 Matricization or unfolding

There are different ways of matricizing the tensor $\mathcal{X}$ according to the partitioning of the set $\{1,\ldots,N\}$ into two ordered subsets $\mathbb{S}_1$ and $\mathbb{S}_2$, constituted of p and $N - p$ indices, respectively. A general formula for the matricization, for $p \in [1, N-1]$, is

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\ldots,i_N} \left(\underset{n \in \mathbb{S}_1}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right) \left(\underset{n \in \mathbb{S}_2}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right)^T \in \mathbb{C}^{J_1 \times J_2} \quad (7)$$

with $J_{n_1} = \prod_{n \in \mathbb{S}_{n_1}} I_n$, for $n_1 = 1$ and 2. From (7), we can deduce the following expression of the element $x_{i_1,\ldots,i_N}$ in terms of the matrix unfolding $\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2}$

$$x_{i_1,\ldots,i_N} = \left(\underset{n \in \mathbb{S}_1}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right)^T \mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} \left(\underset{n \in \mathbb{S}_2}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right). \quad (8)$$

1.2.5 Particular case: mode-n matrix unfoldings $\mathbf{X}_n$

A flat mode-n matrix unfolding of the tensor $\mathcal{X}$ corresponds to an unfolding under the form $\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2}$ with $\mathbb{S}_1 = \{n\}$ and $\mathbb{S}_2 = \{n+1,\ldots,N,1,\ldots,n-1\}$, which gives

$$\mathbf{X}_{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}} = \mathbf{X}_n$$
$$= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1,\ldots,i_N} \mathbf{e}_{i_n}^{(I_n)} \left(\underset{n \in \mathbb{S}_2}{\otimes} \mathbf{e}_{i_n}^{(I_n)} \right)^T \in \mathbb{C}^{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}}. \quad (9)$$

We can also define a tall mode-n matrix unfolding of $\mathcal{X}$, by choosing $\mathbb{S}_1 = \{n+1,\ldots,N,1,\ldots,n-1\}$ and $\mathbb{S}_2 = \{n\}$. Then, we have $\mathbf{X}_{I_{n+1} \cdots I_N I_1 \cdots I_{n-1} \times I_n} = \mathbf{X}_n^T \in \mathbb{C}^{I_{n+1} \cdots I_N I_1 \cdots I_{n-1} \times I_n}$.

The column vectors of a flat mode-n matrix unfolding $\mathbf{X}_n$ are the mode-n vectors of $\mathcal{X}$, and the rank of $\mathbf{X}_n$, i.e., the dimension of the mode-n linear space spanned by the mode-n vectors, is called mode-n rank of $\mathcal{X}$, denoted by $\text{rank}_n(\mathcal{X})$.

In the case of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, there are six different flat unfoldings, denoted $\mathbf{X}_{I \times JK}, \mathbf{X}_{I \times KJ}, \mathbf{X}_{J \times KI}, \mathbf{X}_{J \times IK}, \mathbf{X}_{K \times IJ}$, and $\mathbf{X}_{K \times JI}$. For instance, we have

$$\mathbf{X}_{I \times JK} = \mathbf{X}_{\{1\};\{2,3\}} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} x_{i,j,k} \mathbf{e}_i^{(I)} \left(\mathbf{e}_j^{(J)} \otimes \mathbf{e}_k^{(K)} \right)^T. \quad (10)$$

Using the properties (84), (85), and (87) of the Kronecker product gives

$$\mathbf{X}_{I \times JK} = \sum_{j=1}^{J} \left(\mathbf{e}_j^{(J)} \right)^T \otimes \sum_{i=1}^{I} \sum_{k=1}^{K} x_{i,j,k} \mathbf{e}_i^{(I)} \left(\mathbf{e}_k^{(K)} \right)^T$$
$$= \sum_{j=1}^{J} \left(\mathbf{e}_j^{(J)} \right)^T \otimes (\mathbf{X}_{j.})^T = [\mathbf{X}_{.1.}^T \cdots \mathbf{X}_{.J.}^T] \in \mathbb{C}^{I \times JK}.$$

Similarly, there are six tall matrix unfoldings, denoted $\mathbf{X}_{JK \times I}, \mathbf{X}_{KJ \times I}, \mathbf{X}_{KI \times J}, \mathbf{X}_{IK \times J}, \mathbf{X}_{IJ \times K}, \mathbf{X}_{JI \times K}$, like for instance

$$\mathbf{X}_{JK \times I} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} x_{i,j,k} \left(\mathbf{e}_j^{(J)} \otimes \mathbf{e}_k^{(K)} \right) \mathbf{e}_i^{(I)T} = \mathbf{X}_{I \times JK}^T \in \mathbb{C}^{JK \times I}. \quad (11)$$

Applying (8) to (10) gives

$$x_{i,j,k} = \left(\mathbf{e}_i^{(I)} \right)^T \mathbf{X}_{I \times JK} \left(\mathbf{e}_j^{(J)} \otimes \mathbf{e}_k^{(K)} \right) = [\mathbf{X}_{I \times JK}]_{i,(j-1)K+k}.$$

1.2.6 Mode-n product of a tensor with a matrix or a vector

The mode-n product of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots I_N}$ with $\mathbf{A} \in \mathbb{C}^{J_n \times I_n}$ along the nth mode, denoted by $\mathcal{X} \times_n \mathbf{A}$, gives the tensor $\mathcal{Y}$ of order N and dimensions $I_1 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N$, such as [47]

$$y_{i_1,\ldots,i_{n-1},j_n,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} a_{j_n,i_n} x_{i_1,\ldots,i_{n-1},i_n,i_{n+1},\ldots,i_N} \quad (12)$$

which can be expressed in terms of mode-n matrix unfoldings of $\mathcal{X}$ and $\mathcal{Y}$

$$\mathbf{Y}_n = \mathbf{A}\mathbf{X}_n.$$

This operation can be interpreted as the linear map from the mode-n space of $\mathcal{X}$ to the mode-n space of $\mathcal{Y}$, associated with the matrix $\mathbf{A}$.

The mode-n product of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with the row vector $\mathbf{u}^T \in \mathbb{C}^{1 \times I_n}$ along the nth mode, denoted by $\mathcal{X} \times_n \mathbf{u}^T$, gives a tensor $\mathcal{Y}$ of order $N-1$ and dimensions $I_1 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N$, such as

$$y_{i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} u_{i_n} x_{i_1,\ldots,i_{n-1},i_n,i_{n+1},\ldots,i_N}$$

that can be written in vectorized form as $\text{vec}^T(\mathcal{Y}) = \mathbf{u}^T \mathbf{X}_n \in \mathbb{C}^{1 \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}}$.

When multiplying a Nth-order tensor by row vectors along p different modes, we get a tensor of order $N-p$. For instance, for a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, we have

$$\mathbf{x}_{ij.} = \mathcal{X} \times_1 \mathbf{e}_i^{(I)^T} \times_2 \mathbf{e}_j^{(J)^T}, \quad x_{ijk} = \mathcal{X} \times_1 \mathbf{e}_i^{(I)^T} \times_2 \mathbf{e}_j^{(J)^T} \times_3 \mathbf{e}_k^{(K)^T}.$$

Considering an ordered subset $\mathbb{S} = \{m_1, \ldots, m_P\}$ of the set $\{1, \ldots, N\}$, a series of mode-m_p products of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with $\mathbf{A}^{(m_p)} \in \mathbb{C}^{J_{m_p} \times I_{m_p}}$, $p \in \{1, \ldots, P\}$, $P \leq N$, will be concisely noted as

$$\mathcal{X} \times_{m_1} \mathbf{A}^{(m_1)} \cdots \times_{m_P} \mathbf{A}^{(m_P)} = \mathcal{X} \times_{m=m_1}^{m_P} \mathbf{A}^{(m)}.$$

Properties

- For any permutation $\pi(.)$ of P distinct indices $m_p \in \{1, \ldots, N\}$ such as $q_p = \pi(m_p)$, $p \in \{1, \ldots, P\}$, with $P \leq N$, we have

$$\mathcal{X} \times_{q=q_1}^{q_P} \mathbf{A}^{(q)} = \mathcal{X} \times_{m=m_1}^{m_P} \mathbf{A}^{(m)}$$

  which means that the order of the mode-m_p products is irrelevant when the indices m_p are all distinct.
- For two products of $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ along the same mode-n, with $\mathbf{A} \in \mathbb{C}^{J_n \times I_n}$ and $\mathbf{B} \in \mathbb{C}^{K_n \times J_n}$, we have [13]

$$\begin{aligned} \mathcal{Y} &= \mathcal{X} \times_n \mathbf{A} \times_n \mathbf{B} \\ &= \mathcal{X} \times_n (\mathbf{BA}) \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times K_n \times I_{n+1} \times \cdots \times I_N}. \end{aligned}$$

(13)

1.2.7 Kronecker product-based approach using index notation

In this subsection, we propose to reformulate our Kronecker product-based approach for tensor matricization in terms of an index notation introduced in [48]. Using this index notation, for $\mathbf{u} \in \mathbb{C}^{I \times 1}$, $\mathbf{v}^T \in \mathbb{C}^{1 \times J}$, and $\mathbf{X} \in \mathbb{C}^{I \times J}$, we can write

$$\mathbf{u} = \sum_{i=1}^{I} u_i \mathbf{e}_i^{(I)} = u_i \mathbf{e}_i$$

$$\mathbf{v}^T = \sum_{j=1}^{J} v_j \left(\mathbf{e}_j^{(J)}\right)^T = v_j \mathbf{e}^j$$

$$\mathbf{X} = \sum_{i=1}^{I} \sum_{j=1}^{j} x_{ij} \left(\mathbf{e}_i^{(I)} \otimes \left(\mathbf{e}_j^{(J)}\right)^T\right) = x_{ij} \mathbf{e}_i^j$$

$$\mathbf{X}^T = x_{ij} \mathbf{e}_j^i$$

$$\text{vec}(\mathbf{X}) = x_{ij} \mathbf{e}_{ji}$$

As with Einstein summation convention, the index notation allows to drop summation signs. If an index $i \in [1, I]$ is repeated in an expression (or more generally in a term of an equation), it means that this expression (or this term)

must be summed over that index from 1 to I. However, it is worth noting the two differences between the index notation used in this paper and Einstein summation convention: (*i*) each index can be repeated more than twice in any expression and (*ii*) the index notation can be used with ordered sets of indices. We have to notice that the index notation can be interpreted in terms of two separate combinations of indices, one associated with the column (superscript) indices and the other one with the row (subscript) indices, with the following rules:

- the ordering of the column indices is independent of that of the row indices;
- the ordering of the column and row indices cannot be changed.

Considering the set $\mathbb{S} = \{n_1, \ldots, n_N\}$ obtained by permuting the elements of $\{1, \ldots, N\}$ and defining the ordered set of indices $\mathbb{I} = \{i_{n_1}, \ldots, i_{n_N}\}$ associated with $\mathbb{S}$, we denote by $\mathbf{e}_{\mathbb{I}}$ and $\mathbf{e}^{\mathbb{I}}$ the Kronecker products $\underset{n \in \mathbb{S}}{\otimes} \mathbf{e}_{i_n}$ and $\underset{n \in \mathbb{S}}{\otimes} \mathbf{e}_{i_n}^T$, respectively. So, we have

$$\underset{n \in \mathbb{S}}{\diamond} \mathbf{u}^{(n)} = \left(\prod_{n \in \mathbb{S}} u_{i_n}^{(n)}\right) \mathbf{e}_{\mathbb{I}}$$

(14)

Partitioning two ordered sets of indices $\mathbb{I}$ and $\mathbb{J}$ into two subsets $(\mathbb{I}_1, \mathbb{I}_2)$ and $(\mathbb{J}_1, \mathbb{J}_2)$, respectively, the rules enounced previously imply the following identities:

$$\begin{aligned} \mathbf{e}_{\mathbb{I}}^{\mathbb{J}} &= \mathbf{e}_{\mathbb{I}} \otimes \mathbf{e}^{\mathbb{J}} = \mathbf{e}^{\mathbb{J}} \otimes \mathbf{e}_{\mathbb{I}} = \mathbf{e}_{\mathbb{I}_1 \mathbb{I}_2}^{\mathbb{J}_1 \mathbb{J}_2} \\ &= \mathbf{e}_{\mathbb{I}_1} \otimes \mathbf{e}_{\mathbb{I}_2} \otimes \mathbf{e}^{\mathbb{J}_1} \otimes \mathbf{e}^{\mathbb{J}_2} = \mathbf{e}_{\mathbb{I}_1} \otimes \mathbf{e}^{\mathbb{J}_1} \otimes \mathbf{e}_{\mathbb{I}_2} \otimes \mathbf{e}^{\mathbb{J}_2} \\ &= \mathbf{e}_{\mathbb{I}_1} \otimes \mathbf{e}^{\mathbb{J}_1} \otimes \mathbf{e}^{\mathbb{J}_2} \otimes \mathbf{e}_{\mathbb{I}_2} \\ &= \mathbf{e}^{\mathbb{J}_1} \otimes \mathbf{e}^{\mathbb{J}_2} \otimes \mathbf{e}_{\mathbb{I}_1} \otimes \mathbf{e}_{\mathbb{I}_2} = \mathbf{e}^{\mathbb{J}_1} \otimes \mathbf{e}_{\mathbb{I}_1} \otimes \mathbf{e}^{\mathbb{J}_2} \otimes \mathbf{e}_{\mathbb{I}_2} \\ &= \mathbf{e}^{\mathbb{J}_1} \otimes \mathbf{e}_{\mathbb{I}_1} \otimes \mathbf{e}_{\mathbb{I}_2} \otimes \mathbf{e}^{\mathbb{J}_2} \end{aligned}$$

These identities directly result from the property that the Kronecker product of a column vector with a row vector is independent of the order of the vectors ($\mathbf{u} \otimes \mathbf{v}^T = \mathbf{v}^T \otimes \mathbf{u}$), which implies that in a sequence of Kronecker products of column and row vectors, a column vector can be permuted with a row vector without altering the final result, if the proper ordering of the column vectors and of the row vectors is not changed in the sequence ($\mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \mathbf{v}^T = \mathbf{u}_1 \otimes \mathbf{v}^T \otimes \mathbf{u}_2 = \mathbf{v}^T \otimes \mathbf{u}_1 \otimes \mathbf{u}_2 \neq \mathbf{v}^T \otimes \mathbf{u}_2 \otimes \mathbf{u}_1$ if $\mathbf{u}_1 \neq \mathbf{u}_2$).

Using the index notation, the horizontal, lateral, and frontal slices of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ can be written as

$$\mathbf{X}_{i..} = x_{ijk} \mathbf{e}_j^k \; ; \; \mathbf{X}_{.j.} = x_{ijk} \mathbf{e}_k^i \; ; \; \mathbf{X}_{..k} = x_{ijk} \mathbf{e}_i^j.$$

The Kronecker products of vectors ($\mathbf{u} \in \mathbb{C}^{I \times 1}, \mathbf{v} \in \mathbb{C}^{J \times 1}$) and matrices ($\mathbf{A} \in \mathbb{C}^{I \times J}, \mathbf{B} \in \mathbb{C}^{K \times L}$) can be concisely written as

$$
\mathbf{u} \otimes \mathbf{v} = (u_i \mathbf{e}_i) \otimes (v_j \mathbf{e}_j) = u_i v_j \mathbf{e}_{ij}
$$
$$
\mathbf{u}^T \otimes \mathbf{v}^T = u_i v_j \mathbf{e}^{ij}
$$
$$
\mathbf{u} \otimes \mathbf{v}^T = u_i v_j \mathbf{e}_i^j
$$
$$
\mathbf{A} \otimes \mathbf{B} = \left(a_{ij} \mathbf{e}_i^j \right) \otimes \left(b_{kl} \mathbf{e}_k^l \right) = a_{ij} b_{kl} \mathbf{e}_{ik}^{jl}
$$
$$
\mathbf{A}^T \otimes \mathbf{B}^T = a_{ij} b_{kl} \mathbf{e}_{jl}^{ik}
$$

For $\mathbf{U} = \left[\mathbf{u}^{(1)} \cdots \mathbf{u}^{(N)} \right] \in \mathbb{C}^{I \times N}$ and $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_N] \in \mathbb{C}^{J \times N}$, we have

$$
\mathbf{U}\mathbf{V}^T = \sum_{n=1}^{N} \mathbf{u}^{(n)} (\mathbf{v}^{(n)})^T = u_i^{(n)} v_j^{(n)} \mathbf{e}_i^j \tag{15}
$$

where the summation over n is to be done after the matricization $\mathbf{u}^{(n)} (\mathbf{v}^{(n)})^T$.

Using the index notation, the Khatri-Rao product can be written as follows:

$$
\mathbf{A} \diamond \mathbf{B} = a_{ik} b_{jk} \mathbf{e}_{ij}^{k}
$$
$$
(\mathbf{A} \diamond \mathbf{B})^T = a_{ik} b_{jk} \mathbf{e}_k^{ij} \tag{16}
$$

The Kronecker and Khatri-Rao products defined in (1) and (2), with $a_{i_n,r_n}^{(n)}$ as entry of $\mathbf{A}^{(n)}$, can then be defined as

$$
\underset{n \in \mathbb{S}}{\otimes} \mathbf{A}^{(n)} = \left(\prod_{n \in \mathbb{S}} a_{i_n,r_n}^{(n)} \right) \mathbf{e}_{i_{n_1}, \ldots, i_{n_N}}^{r_{n_1}, \ldots, r_{n_N}} = \left(\prod_{n \in \mathbb{S}} a_{i_n,r_n}^{(n)} \right) \mathbf{e}_{\mathbb{I}}^{\mathbb{R}} \tag{17}
$$

$$
\underset{n \in \mathbb{S}}{\diamond} \mathbf{A}^{(n)} = \left(\prod_{n \in \mathbb{S}} a_{i_n,r}^{(n)} \right) \mathbf{e}_{i_{n_1}, \ldots, i_{n_N}}^{r} = \left(\prod_{n \in \mathbb{S}} a_{i_n,r}^{(n)} \right) \mathbf{e}_{\mathbb{I}}^{r} \tag{18}
$$

where $\mathbb{R} = \{r_{n_1}, \ldots, r_{n_N}\}$.

Applying these results, the unfoldings (7), (10), and (11) and the formula (8) can be rewritten respectively as

$$
\mathbf{X}_{\mathbb{S}_1; \mathbb{S}_2} = x_{i_1, \ldots, i_N} \mathbf{e}_{\mathbb{I}_1}^{\mathbb{I}_2} \tag{19}
$$

$$
\mathbf{X}_{I \times JK} = x_{i,j,k} \mathbf{e}_i^{jk}
$$
$$
\mathbf{X}_{JK \times I} = x_{i,j,k} \mathbf{e}_{jk}^{i}
$$
$$
x_{i_1, \ldots, i_N} = \mathbf{e}^{\mathbb{I}_1} \mathbf{X}_{\mathbb{S}_1; \mathbb{S}_2} \mathbf{e}_{\mathbb{I}_2} \tag{20}
$$

where $\mathbb{I}_1$ and $\mathbb{I}_2$ represent the sets of indices i_n associated with the sets $\mathbb{S}_1$ and $\mathbb{S}_2$ of index n, respectively.

We can also use the index notation for deriving matrix unfoldings of tensor extensions of a matrix $\mathbf{B} \in \mathbb{C}^{I \times J}$. For instance, if we define the tensor $\mathcal{A} \in \mathbb{C}^{I \times J \times K}$ such as

$a_{i,j,k} = b_{i,j}$ for $k = 1, \ldots, K$, mode-1 flat unfoldings of $\mathcal{A}$ are given by

$$
\mathbf{A}_{I \times JK} = \sum_{i,j,k} a_{i,j,k} \mathbf{e}_i \otimes \mathbf{e}^j \otimes \mathbf{e}^k = a_{i,j,k} \mathbf{e}_i^{jk}
$$
$$
= \sum_{i,j} b_{i,j} \mathbf{e}_i^j \otimes \sum_{k=1}^{K} \mathbf{e}^k = \mathbf{B} \otimes \mathbf{1}_K^T = \mathbf{B} \left(\mathbf{I}_J \otimes \mathbf{1}_K^T \right) \tag{21}
$$

$$
\mathbf{A}_{I \times KJ} = a_{i,j,k} \mathbf{e}_i^{kj} = \sum_{k=1}^{K} \mathbf{e}^k \otimes b_{i,j} \mathbf{e}_i^j
$$
$$
= \mathbf{1}_K^T \otimes \mathbf{B} = \mathbf{B} \left(\mathbf{1}_K^T \otimes \mathbf{I}_J \right) \tag{22}
$$

These two formulae will be used later for establishing the link between PARATUCK-(2,4) models and constrained PARAFAC-4 models.

1.2.8 Basic tensor models

We now present the two most common tensor models, i.e., the Tucker [3] and PARAFAC [4] models. In [7], these models are introduced in a constructive way, in the context of a three-way data analysis. The Tucker models are presented as extensions of the matrix singular value decomposition (SVD) to three-way arrays, which gave rise to the generalization as HOSVD [13,49], whereas the PARAFAC model is introduced by emphasizing Cattell's principle of parallel proportional profiles [50] that underlies this model, so explaining the acronym PARAFAC. In the following, we adopt a more general presentation for multi-way arrays, i.e., tensors of any order N.

Tucker models For a Nth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, a Tucker model is defined in an element-wise form as

$$
x_{i_1, \ldots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{r_1, \ldots, r_N} \prod_{n=1}^{N} a_{i_n, r_n}^{(n)} \tag{23}
$$

with $i_n = 1, \ldots, I_n$ for $n = 1, \ldots, N$, where $g_{r_1, \ldots, r_N}$ is an element of the core tensor $\mathcal{G} \in \mathbb{C}^{R_1 \times \cdots \times R_N}$ and $a_{i_n, r_n}^{(n)}$ is an element of the matrix factor $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$.

Using the index notation and defining the set of indices $\mathbb{R} = \{r_1, \ldots, r_N\}$, the Tucker model can also be written simply as

$$
x_{i_1, \ldots, i_N} = g_{r_1, \ldots, r_N} \prod_{\mathbb{R}} a_{i_n, r_n}^{(n)} \tag{24}
$$

Taking the definition (4) into account and noting that $\sum_{i_n=1}^{I_n} a_{i_n, r_n}^{(n)} \mathbf{e}_{i_n}^{(I_n)} = \mathbf{A}_{.r_n}^{(n)}$, this model can be written as a

weighted sum of $\prod_{n=1}^{N} R_n$ outer products, i.e., rank-one tensors

$$
\mathcal{X} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{r_1,\ldots,r_N} \overset{N}{\underset{n=1}{\circ}} \mathbf{A}_{.,r_n}^{(n)}
$$
$$
= g_{r_1,\ldots,r_N} \underset{\mathbb{R}}{\circ} \mathbf{A}_{.,r_n}^{(n)} \quad \text{(with the index notation)} \tag{25}
$$

Using the definition (12) allows to write (23) in terms of mode-n products as

$$
\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \cdots \times_N \mathbf{A}^{(N)}
$$
$$
= \mathcal{G} \times_{n=1}^{N} \mathbf{A}^{(n)}. \tag{26}
$$

This expression evidences that the Tucker model can be viewed as the transformation of the core tensor resulting from its multiplication by the factor matrix $\mathbf{A}^{(n)}$ along its mode-n, which corresponds to a linear map applied to the mode-n space of $\mathcal{G}$, for $n = 1, \ldots, N$, i.e., a multilinear map applied to $\mathcal{G}$. From a transformation point of view, $\mathcal{G}$ and $\mathcal{X}$ can be interpreted as the input tensor and the transformed tensor, or output tensor, respectively.

Matrix representations of the Tucker model. A matrix representation of a Tucker model is directly linked with a matricization of tensor like (7), corresponding to the combination of two sets of modes $\mathbb{S}_1$ and $\mathbb{S}_2$. These combinations must be applied both to the tensor $\mathcal{X}$ and its core tensor $\mathcal{G}$.

The matrix representation (7) of the Tucker model (23) is given by

$$
\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n \in \mathbb{S}_1}{\otimes} \mathbf{A}^{(n)} \right) \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \left(\underset{n \in \mathbb{S}_2}{\otimes} \mathbf{A}^{(n)} \right)^T \tag{27}
$$

with $\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \in \mathbb{C}^{J_1 \times J_2}$, and $J_{n_1} = \prod_{n \in \mathbb{S}_{n_1}} R_n$, for $n_1 = 1$ and 2.

Proof. See Appendix 3.   □

For the flat mode-n unfolding, defined in (9), the formula (27) gives

$$
\mathbf{X}_n = \mathbf{A}^{(n)} \mathbf{G}_n \left(\mathbf{A}^{(n+1)} \otimes \cdots \otimes \mathbf{A}^{(N)} \otimes \mathbf{A}^{(1)} \otimes \cdots \otimes \mathbf{A}^{(n-1)} \right)^T. \tag{28}
$$

Applying the vec formula (92) to the right-hand side of (28), we obtain the vectorized form of $\mathcal{X}$ associated with its mode-n unfolding $\mathbf{X}_n$

$$
\text{vec}(\mathcal{X}) = \text{vec}(\mathbf{X}_n)
$$
$$
= \left(\mathbf{A}^{(n+1)} \otimes \cdots \otimes \mathbf{A}^{(N)} \otimes \mathbf{A}^{(1)} \otimes \cdots \otimes \mathbf{A}^{(n)} \right) \text{vec}(\mathbf{G}_n).
$$

Tucker-(N_1, N) models A Tucker-(N_1, N) model for a Nth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, with $N \geq N_1$, corresponds to the case where $N - N_1$ factor matrices are equal

to identity matrices. For instance, assuming that $\mathbf{A}^{(n)} = \mathbf{I}_{I_n}$, which implies $R_n = I_n$, for $n = N_1 + 1, \ldots, N$, (23) and (26) become

$$
x_{i_1,\ldots,i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} g_{r_1,\ldots,r_{N_1},i_{N_1+1},\ldots,i_N} \prod_{n=1}^{N_1} a_{i_n,r_n}^{(n)}
$$
$$
\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \cdots \times_{N_1} \mathbf{A}^{(N_1)} \times_{N_1+1} \mathbf{I}_{I_{N_1+1}} \cdots \times_N \mathbf{I}_{I_N} \tag{29}
$$
$$
= \mathcal{G} \times_{n=1}^{N_1} \mathbf{A}^{(n)}. \tag{30}
$$

One such model that is currently used in applications is the Tucker-(2,3) model, usually denoted Tucker2, for third-order tensors $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$. Assuming $\mathbf{A}^{(1)} = \mathbf{A} \in \mathbb{C}^{I \times P}$, $\mathbf{A}^{(2)} = \mathbf{B} \in \mathbb{C}^{J \times Q}$, and $\mathbf{A}^{(3)} = \mathbf{I}_K$, such a model is defined by the following equations:

$$
x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} g_{pqk} a_{ip} b_{jq} \tag{31}
$$
$$
\mathcal{X} = \mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \tag{32}
$$

with the core tensor $\mathcal{G} \in \mathbb{C}^{P \times Q \times K}$.

PARAFAC models A PARAFAC model for a Nth-order tensor corresponds to the particular case of a Tucker model with an identity core tensor of order N and dimensions $R \times \cdots \times R$

$$
\mathcal{G} = \mathcal{I}_{N,R} = \mathcal{I} \quad \Leftrightarrow \quad g_{r_1,\ldots,r_N} = \delta_{r_1,\ldots,r_N}
$$

(23) to (26) then become, respectively

$$
x_{i_1,\ldots,i_N} = \sum_{r=1}^{R} \prod_{n=1}^{N} a_{i_n,r}^{(n)} \tag{33}
$$
$$
= \prod_{n=1}^{N} a_{i_n,r}^{(n)} \quad \text{(with the index notation)} \tag{34}
$$
$$
\mathcal{X} = \sum_{r=1}^{R} \left(\overset{N}{\underset{n=1}{\circ}} \mathbf{A}_{.r}^{(n)} \right)
$$
$$
\mathcal{X} = \mathcal{I}_{N,R} \times_{n=1}^{N} \mathbf{A}^{(n)} \tag{35}
$$

with the factor matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}, n = 1, \ldots, N$.

Remarks

- The expression (33) as a sum of polyads is called a polyadic form of $\mathcal{X}$ by Hitchcock [51].
- The PARAFAC model (33, 34 and 35) amounts to decomposing the tensor $\mathcal{X}$ into a sum of R components, each component being a rank-one tensor. When R is minimal in (33), it is called the rank of $\mathcal{X}$ [52]. This rank is related to the mode-n ranks by the following inequalities
  $\text{rank}_n(\mathcal{X}) \leq R, \forall n = 1, \ldots, N$. Furthermore, contrary

to the matrices for which the rank is always at most equal to the smallest of the dimensions, for higher-order tensors, the rank can exceed any mode-n dimension I_n.

There exists different definitions of rank for tensors, like typical and generic ranks, or also symmetric rank for a symmetric tensor (see [53,54] for more details).

- In telecommunication applications, the structure parameters (rank, mode dimensions, and core tensor dimensions) of a PARAFAC or Tucker model are design parameters that are chosen in function of the performance desired for the communication system. However, in most of the applications, as for instance in multi-way data analysis, the structure parameters are generally unknown and must be determined *a priori.* Several techniques have been proposed for determining these parameters (see [55-58] and references therein).

- The PARAFAC model is also sometimes defined by the following equation

$$x_{i_1,\dots,i_N} = \sum_{r=1}^{R} g_r \prod_{n=1}^{N} a_{i_n,r}^{(n)} \quad \text{with} \quad g_r > 0. \qquad (36)$$

In this case, the identity tensor $\mathcal{I}_{N,R}$ in (35) is replaced by the diagonal tensor $\mathcal{G} \in \mathbb{C}^{R \times \cdots \times R}$ whose diagonal elements are equal to scaling factors g_r, i.e.

$$g_{r_1,\dots,r_N} = \begin{cases} g_r & \text{if } r_1 = \cdots = r_N = r \\ 0 & \text{otherwise} \end{cases}$$

and all the column vectors $\mathbf{A}_{.r}^{(n)}$ are normalized, i.e., with a unit norm, for $1 \le n \le N$.

- It is important to notice that the PARAFAC model (33) is multilinear (more precisely N-linear) in its parameters in the sense that it is linear with respect to each matrix factor. This multilinearity property is exploited for parameter estimation using the standard alternating least squares (ALS) algorithm [4,5] that consists in alternately estimating each matrix factor by minimizing a least squares error criterion conditionally to the knowledge of the other matrix factors that are fixed with their previously estimated values.

Matrix representations of the PARAFAC model. The matrix representation (7) of the PARAFAC model (33)-(35) is given by

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\diamond} \mathbf{A}^{(n)} \right) \left(\underset{n\in\mathbb{S}_2}{\diamond} \mathbf{A}^{(n)} \right)^T. \qquad (37)$$

Proof. See Appendix 4.   □

Remarks

- From (37), we can deduce that a mode combination results in a Khatri-Rao product of the corresponding factor matrices. Consequently, the tensor contraction (5) associated with the PARAFAC-N model (35) gives a PARAFAC-N_1 model whose factor matrices are equal to $\underset{n\in\mathbb{S}_{n_1}}{\diamond} \mathbf{A}^{(n)} \in \mathbb{C}^{J_{n_1} \times R}$, $n_1 = 1,\dots,N_1$, with $J_{n_1} = \prod_{n\in\mathbb{S}_{n_1}} I_n$.

- For the PARAFAC model, the flat mode-n unfolding, defined in (9), is given by

$$\mathbf{X}_n = \mathbf{A}^{(n)} \left(\mathbf{A}^{(n+1)} \diamond \cdots \diamond \mathbf{A}^{(N)} \diamond \mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(n-1)} \right)^T, \qquad (38)$$

and the associated vectorized form is obtained in applying the vec formula (93) to the right-hand side of the above equation, with $\mathbf{I}_R = \text{diag}(\mathbf{1}_R)$

$$\text{vec}(\mathcal{X}) = \text{vec}(\mathbf{X}_n) = \left(\mathbf{A}^{(n+1)} \diamond \cdots \diamond \mathbf{A}^{(N)} \diamond \mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(n)} \right) \mathbf{1}_R \qquad (39)$$

- In the case of the normalized PARAFAC model (36), (37) and (39) become, respectively,

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\diamond} \mathbf{A}^{(n)} \right) \text{diag}(\mathbf{g}) \left(\underset{n\in\mathbb{S}_2}{\diamond} \mathbf{A}^{(n)} \right)^T$$

$$\text{vec}(\mathcal{X}) = \text{vec}(\mathbf{X}_n)$$
$$= \left(\mathbf{A}^{(n+1)} \diamond \cdots \diamond \mathbf{A}^{(N)} \diamond \mathbf{A}^{(1)} \diamond \cdots \diamond \mathbf{A}^{(n)} \right) \mathbf{g}$$

where $\mathbf{g} = \begin{bmatrix} g_1 \cdots g_R \end{bmatrix}^T \in \mathbb{C}^{R \times 1}$.

- For the PARAFAC model of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ with factor matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C})$, the formula (37) gives for $\mathbb{S}_1 = \{i, j\}$ and $\mathbb{S}_2 = \{k\}$

$$\mathbf{X}_{IJ \times K} = \begin{bmatrix} \mathbf{X}_{1..} \\ \vdots \\ \mathbf{X}_{I..} \end{bmatrix} = (\mathbf{A} \diamond \mathbf{B})\mathbf{C}^T \in \mathbb{C}^{IJ \times K}.$$

Noting that $\mathbf{A} \diamond \mathbf{B} = \begin{bmatrix} \mathbf{B}D_1(\mathbf{A}) \\ \vdots \\ \mathbf{B}D_I(\mathbf{A}) \end{bmatrix}$, we deduce the following expression for mode-1 matrix slices

$$\mathbf{X}_{i..} = \mathbf{B}D_i(\mathbf{A})\mathbf{C}^T.$$

Similarly, we have

$$\mathbf{X}_{JK \times I} = (\mathbf{B} \diamond \mathbf{C})\mathbf{A}^T, \quad \mathbf{X}_{KI \times J} = (\mathbf{C} \diamond \mathbf{A})\mathbf{B}^T,$$
$$\mathbf{X}_{.j.} = \mathbf{C}D_j(\mathbf{B})\mathbf{A}^T, \quad \mathbf{X}_{..k} = \mathbf{A}D_k(\mathbf{C})\mathbf{B}^T.$$

- For the PARAFAC model of a fourth-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K \times L}$ with factor matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$, we obtain

$$
\begin{aligned}
\mathbf{X}_{IJK \times L} &= (\mathbf{A} \diamond \mathbf{B} \diamond \mathbf{C})\mathbf{D}^T \\
&= \begin{bmatrix} (\mathbf{B} \diamond \mathbf{C})D_1(\mathbf{A}) \\ \vdots \\ (\mathbf{B} \diamond \mathbf{C})D_I(\mathbf{A}) \end{bmatrix} \mathbf{D}^T \\
&= \begin{bmatrix} \mathbf{C}D_1(\mathbf{B})D_1(\mathbf{A}) \\ \vdots \\ \mathbf{C}D_J(\mathbf{B})D_I(\mathbf{A}) \end{bmatrix} \mathbf{D}^T \in \mathbb{C}^{IJK \times L}
\end{aligned}
$$
$$
\mathbf{X}_{ij..} = \mathbf{C}D_j(\mathbf{B})D_i(\mathbf{A})\mathbf{D}^T \in \mathbb{C}^{K \times L} \tag{40}
$$

  Other matrix slices can be deduced from (40) by simple permutations of the matrix factors.

In the next section, we introduce two constrained PARAFAC models, the so-called PARALIND and CONFAC models, and then PARATUCK models.

1.3 Constrained PARAFAC models

The introduction of constraints in tensor models can result from the system itself that is under study or from a system design. In the first case, the constraints are often interpreted as interactions or linear dependencies between the PARAFAC factors. Examples of such dependencies are encountered in psychometric and chemometric applications that gave origin, respectively, to the PARATUCK-2 model [59] and the parallel profiles with linear dependencies (PARALIND) model [60,61], introduced in [47] under the name canonical decomposition with linear constraints (CANDELINC), for the multiway case. A first application of the PARATUCK-2 model in signal processing was made in [62] for blind joint identification and equalization of Wiener-Hammerstein communication channels. The PARALIND model was applied for identifiability and propagation parameter estimation purposes in a context of array signal processing [63,64].

In the second case, the constraints are used as design parameters. For instance, in a telecommunications context, we proposed two constrained tensor models: the CONFAC (constrained factor) model [65] and the PARATUCK-(N_1, N) model [66,67]. The PARATUCK-2 model was also applied for designing space-time spreading-multiplexing MIMO systems [68]. For these telecommunication applications of constrained tensor models, the constraints are used for resource allocation. We are now going to describe these various constrained PARAFAC models.

1.3.1 PARALIND models

Let us define the core tensor of the Tucker model (26) as follows:

$$
\mathcal{G} = \mathcal{I}_{N,R} \times_{n=1}^{N} \mathbf{\Phi}^{(n)} \tag{41}
$$

where $\mathbf{\Phi}^{(n)} \in \mathbb{R}^{R_n \times R}$, $n = 1, \ldots, N$, with $R \geq \max_{n}(R_n)$, are constraint matrices. In this case, $\mathcal{G}$ will be called the 'interaction tensor', or 'constraint tensor'.

The PARALIND model is obtained by substituting (41) into (26) and applying the property (13), which gives

$$
\mathcal{X} = \mathcal{G} \times_{n=1}^{N} \mathbf{A}^{(n)} = \mathcal{I}_{N,R} \times_{n=1}^{N} \left(\mathbf{A}^{(n)} \mathbf{\Phi}^{(n)} \right). \tag{42}
$$

This equation leads to two different interpretations of the PARALIND model, as a constrained Tucker model whose core tensor admits a PARAFAC decomposition with factor matrices $\mathbf{\Phi}^{(n)}$, called 'interaction matrices,' and as a constrained PARAFAC model with constrained factor matrices $\bar{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \mathbf{\Phi}^{(n)}$.

The interaction matrix $\mathbf{\Phi}^{(n)}$ allows taking into account linear dependencies between the columns of $\mathbf{A}^{(n)}$, implying a rank deficiency for this factor matrix. When the columns of $\mathbf{\Phi}^{(n)}$ are formed with 0's and 1's, the dependencies simply consist in a repetition or an addition of certain columns of $\mathbf{A}^{(n)}$. In this particular case, the diagonal element $\xi_{r,r}^{(n)} \geq 1$ of the matrix $\mathbf{\Xi}^{(n)} = \mathbf{\Phi}^{(n)T} \mathbf{\Phi}^{(n)} \in \mathbb{R}^{R \times R}$ represents the number of columns of $\mathbf{A}^{(n)}$ that are added to form the rth column of the constrained factor $\mathbf{A}^{(n)} \mathbf{\Phi}^{(n)}$. The choice $\mathbf{\Phi}^{(n)} = \mathbf{I}_R$ means that there is no such dependency among the columns of $\mathbf{A}^{(n)}$.

Note that (42) can be written element-wise as

$$
\begin{aligned}
x_{i_1,\ldots,i_N} &= \sum_{r_1=1}^{R_1} \cdots \sum_{r_N=1}^{R_N} g_{r_1,\ldots,r_N} \prod_{n=1}^{N} a_{i_n,r_n}^{(n)} \\
&= \sum_{r=1}^{R} \prod_{n=1}^{N} \bar{a}_{i_n,r}^{(n)} \text{ with } \bar{a}_{i_n,r}^{(n)} = \sum_{r_n=1}^{R_n} a_{i_n,r_n}^{(n)} \phi_{r_n,r}^{(n)}.
\end{aligned}
$$
$$
\text{with } g_{r_1,\ldots,r_N} = \sum_{r=1}^{R} \prod_{n=1}^{N} \phi_{r_n,r}^{(n)} \tag{43}
$$

This constrained PARAFAC model constitutes an N-way form of the three-way PARALIND model, used for chemometric applications in [60,61].

1.3.2 CONFAC models

When the constraint matrices $\mathbf{\Phi}^{(n)} \in \mathbb{R}^{R_n \times R}$ are full row rank and their columns are chosen as canonical vectors of the Euclidean space $\mathbb{R}^{R_n}$, for $n = 1, \ldots, N$, the constrained PARAFAC model (42) constitutes a generalization to Nth-order of the third-order CONFAC model, introduced in [65] for designing MIMO communication systems with resource allocation. This CONFAC model was used in [69]

for solving the problem of blind identification of under-determined mixtures based on cumulant generating function of the observations. In a telecommunications context where $\mathcal{X}$ represents the tensor of received signals, such a constraint matrix $\boldsymbol{\Phi}^{(n)}$ can be interpreted as an 'allocation matrix' allowing to allocate resources, like data streams, codes, and transmit antennas, to the R components of the signal to be transmitted. In this case, the core tensor $\mathcal{G}$ will be called the 'allocation tensor.' By assumption, each column of the allocation matrix $\boldsymbol{\Phi}^{(n)}$ is a canonical vector of $\mathbb{R}^{R_n}$, which means that there is only one value of r_n such that $\phi_{r_n,r}^{(n)} = 1$, and this value of r_n corresponds to the nth resource allocated to the rth component.

Each element $x_{i_1,\dots,i_N}$ of the received signal tensor $\mathcal{X}$ is equal to the sum of R components, each component r resulting from the combination of N resources, each resource being associated with a column of the matrix factor $\mathbf{A}^{(n)}$, $n = 1,\dots,N$. This combination, determined by the allocation matrices, is defined by a set of N indices $\{r_1,\dots,r_N\}$ such that $\prod_{n=1}^{N} \phi_{r_n,r}^{(n)} = 1$. As for any $r \in [1,R]$, there is one and only one N-uplet $(r_1,\dots,r_N)$ such as $\prod_{n=1}^{N} \phi_{r_n,r}^{(n)} = 1$, we can deduce that each component r of $x_{i_1,\dots,i_N}$ in (43) is the result of one and only one combination of the N resources under the form of the product $\prod_{n=1}^{N} a_{i_n,r_n}^{(n)}$. For the CONFAC model, we have

$$\sum_{r_n=1}^{R_n} D_{r_n}\left(\boldsymbol{\Phi}^{(n)}\right) = \mathbf{I}_R, \ \forall n = 1,\dots,N$$

meaning that each resource r_n is allocated at least once, and the diagonal element of $\boldsymbol{\Xi}^{(n)} = \boldsymbol{\Phi}^{(n)T}\boldsymbol{\Phi}^{(n)}$ is such as $\xi_{r,r}^{(n)} = 1, \forall n = 1,\dots,N$, because only one resource r_n is allocated to each component r. Moreover, we have to notice that the assumption $R \geq \max_n(R_n)$ implies that each resource can be allocated several times, i.e., to several components. Defining the interaction matrices

$$\boldsymbol{\Gamma}^{(n)} = \boldsymbol{\Phi}^{(n)}\boldsymbol{\Phi}^{(n)T} \in \mathbb{R}^{R_n \times R_n}, \ \boldsymbol{\Gamma}^{(n_1,n_2)} = \boldsymbol{\Phi}^{(n_1)}\boldsymbol{\Phi}^{(n_2)T} \in \mathbb{R}^{R_{n_1} \times R_{n_2}}$$

the diagonal element $\gamma_{r_n,r_n}^{(n)} \in [1, R - R_n + 1]$ represents the number of times that the r_n^{th} column of $\mathbf{A}^{(n)}$ is repeated, i.e., the number of times that the r_n^{th} resource is allocated to the R components, whereas $\gamma_{r_{n_1},r_{n_2}}^{(n_1,n_2)}$ determines the number of interactions between the $r_{n_1}^{\text{th}}$ column of $\mathbf{A}^{(n_1)}$ and the $r_{n_2}^{\text{th}}$ column of $\mathbf{A}^{(n_2)}$, i.e., the number of times that the $r_{n_1}^{\text{th}}$ and $r_{n_2}^{\text{th}}$ resources are combined in the R components. If we choose $R_n = R$ and $\boldsymbol{\Phi}^{(n)} = \mathbf{I}_R, \forall n = 1,\dots,N$, the PARALIND/CONFAC model (42) becomes identical to the PARAFAC one (35).

The matrix representation (7) of the PARALIND/CONFAC model can be deduced from (37) in replacing $\mathbf{A}^{(n)}$ by $\mathbf{A}^{(n)}\boldsymbol{\Phi}^{(n)}$

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\diamond} \mathbf{A}^{(n)}\boldsymbol{\Phi}^{(n)}\right)\left(\underset{n\in\mathbb{S}_2}{\diamond} \mathbf{A}^{(n)}\boldsymbol{\Phi}^{(n)}\right)^T.$$

Using the identity, (86) gives

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\otimes} \mathbf{A}^{(n)}\right)\left(\underset{n\in\mathbb{S}_1}{\diamond} \boldsymbol{\Phi}^{(n)}\right)\left(\underset{n\in\mathbb{S}_2}{\diamond} \boldsymbol{\Phi}^{(n)}\right)^T\left(\underset{n\in\mathbb{S}_2}{\otimes} \mathbf{A}^{(n)}\right)^T,$$

(44)

or, equivalently,

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\otimes} \mathbf{A}^{(n)}\right)\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2}\left(\underset{n\in\mathbb{S}_2}{\otimes} \mathbf{A}^{(n)}\right)^T,$$

where the matrix representation $\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2}$ of the constraint/allocation tensor $\mathcal{G}$, defined by means of its PARAFAC model (41), can also be deduced from (37) as

$$\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\diamond} \boldsymbol{\Phi}^{(n)}\right)\left(\underset{n\in\mathbb{S}_2}{\diamond} \boldsymbol{\Phi}^{(n)}\right)^T.$$

1.3.3 Nested Tucker models

The PARALIND/CONFAC models can be viewed as particular cases of a new family of tensor models that we shall call nested Tucker models, defined by means of the following recursive equation:

$$\mathcal{X}^{(p)} = \mathcal{X}^{(p-1)} \times_{n=1}^{N} \mathbf{A}^{(p,n)} \ \text{for} \ p = 1,\dots,P$$

$$= \mathcal{G} \times_{n=1}^{N} \prod_{q=P}^{1} \mathbf{A}^{(q,n)}$$

with the factor matrices $\mathbf{A}^{(p,n)} \in \mathbb{C}^{R^{(p,n)} \times R^{(p-1,n)}}$ for $p = 1,\dots,P$, such as $R^{(0,n)} = R_n$ and $R^{(P,n)} = I_n$, for $n = 1,\dots,N$, the core tensor $\mathcal{X}^{(0)} = \mathcal{G} \in \mathbb{C}^{R_1 \times \dots \times R_N}$, and $\mathcal{X}^{(P)} \in \mathbb{C}^{I_1 \times \dots \times I_N}$. This equation can be interpreted as P successive linear transformations applied to each mode-n space of the core tensor $\mathcal{G}$. So, P nested Tucker models can then be interpreted as a Tucker model for which the factor matrices are products of P matrices. When $\mathcal{G} = \mathcal{I}_{N,R}$, which implies $R^{(0,n)} = R_n = R$ for $n = 1,\dots,N$, we obtain nested PARAFAC models. The PARALIND/CONFAC models correspond to two nested PARAFAC models ($P = 2$), with $\mathbf{A}^{(1,n)} = \boldsymbol{\Phi}^{(n)}$, $\mathbf{A}^{(2,n)} = \mathbf{A}^{(n)}$, $R^{(0,n)} = R$, $R^{(1,n)} = R_n$, and $R^{(2,n)} = I_n$, for $n = 1,\dots,N$.

By considering nested PARAFAC models with $P = 3$, $\mathbf{A}^{(1,n)} = \boldsymbol{\Phi}^{(n)} \in \mathbb{C}^{K_n \times R}$, $\mathbf{A}^{(2,n)} = \mathbf{A}^{(n)} \in \mathbb{C}^{J_n \times K_n}$, and $\mathbf{A}^{(3,n)} = \boldsymbol{\Psi}^{(n)} \in \mathbb{C}^{I_n \times J_n}$, for $n = 1,\dots,N$, we deduce doubly PARALIND/CONFAC models described by the following equation:

$$\mathcal{X} = \mathcal{I}_{N,R} \times_{n=1}^{N} \left(\boldsymbol{\Psi}^{(n)}\mathbf{A}^{(n)}\boldsymbol{\Phi}^{(n)}\right).$$

Such a model can be viewed as a doubly constrained PARAFAC model, with factor matrices $\boldsymbol{\Psi}^{(n)}\mathbf{A}^{(n)}\boldsymbol{\Phi}^{(n)}$, the constraint matrix $\boldsymbol{\Psi}^{(n)}$, assumed to be full column rank, allowing to take into account linear dependencies between the rows of $\mathbf{A}^{(n)}$. A third-order nested Tucker model is visualized in Figure 1.

1.3.4 Block PARALIND/CONFAC models

In some applications, the data tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is written as a sum of P sub-tensors $\mathcal{X}^{(p)}$, each sub-tensor admitting a tensor model with a possibly different structure. So, we can define a block-PARALIND/CONFAC model as

$$\mathcal{X} = \sum_{p=1}^{P} \mathcal{X}^{(p)}, \tag{45}$$

$$\mathcal{X}^{(p)} = \mathcal{G}^{(p)} \times_{n=1}^{N} \mathbf{A}^{(p,n)}, \tag{46}$$

$$\mathcal{G}^{(p)} = \mathcal{I}_{N,R^{(p)}} \times_{n=1}^{N} \boldsymbol{\Phi}^{(p,n)},$$

where $\mathbf{A}^{(p,n)} \in \mathbb{C}^{I_n \times R^{(p,n)}}$, $\boldsymbol{\Phi}^{(p,n)} \in \mathbb{C}^{R^{(p,n)} \times R^{(p)}}$, and $\mathcal{G}^{(p)} \in \mathbb{C}^{R^{(p,1)} \times \cdots \times R^{(p,N)}}$ are the mode-n factor matrix, the mode-n constraint/allocation matrix, and the core tensor of the PARALIND/CONFAC model of the pth sub-tensor, respectively. The matrix representation (44) then becomes

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \sum_{p=1}^{P} \left(\underset{n\in\mathbb{S}_1}{\otimes} \mathbf{A}^{(p,n)} \right) \left(\underset{n\in\mathbb{S}_1}{\diamond} \boldsymbol{\Phi}^{(p,n)} \right)$$
$$\times \left(\underset{n\in\mathbb{S}_2}{\diamond} \boldsymbol{\Phi}^{(p,n)} \right)^T \left(\underset{n\in\mathbb{S}_2}{\otimes} \mathbf{A}^{(p,n)} \right)^T. \tag{47}$$

Defining the following block partitioned matrices

$$\mathbf{A}^{(n)} = \left[\mathbf{A}^{(1,n)} \cdots \mathbf{A}^{(P,n)} \right] \in \mathbb{C}^{I_n \times R^{(n)}} \tag{48}$$

where $R^{(n)} = \sum_{p=1}^{P} R^{(p,n)}$, (47) can be rewritten in the following more compact form

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\otimes_b} \mathbf{A}^{(n)} \right) \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \left(\underset{n\in\mathbb{S}_2}{\otimes_b} \mathbf{A}^{(n)} \right)^T$$

where $\otimes_b$ denotes the block-wise Kronecker product defined as

$$\mathbf{A}^{(n)} \otimes_b \mathbf{A}^{(q)} = \left[\mathbf{A}^{(1,n)} \otimes \mathbf{A}^{(1,q)} \cdots \mathbf{A}^{(P,n)} \otimes \mathbf{A}^{(P,q)} \right]$$

$\mathbf{A}^{(q)}$ being partitioned in P blocks as in (48), and

$$\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} = \text{bdiag}(\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2}^{(1)} \cdots \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2}^{(P)}) \in \mathbb{C}^{J_1 \times J_2}$$

$$\mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2}^{(p)} = \left(\underset{n\in\mathbb{S}_1}{\diamond_b} \boldsymbol{\Phi}^{(p,n)} \right) \left(\underset{n\in\mathbb{S}_2}{\diamond_b} \boldsymbol{\Phi}^{(p,n)} \right)^T \in \mathbb{C}^{J_1^{(p)} \times J_2^{(p)}}$$

where $\diamond_b$ denotes the block-wise Khatri-Rao product defined in the same way as the block-wise Kronecker product, with $J_{n_1} = \sum_{p=1}^{P} J_{n_1}^{(p)}$ and $J_{n_1}^{(p)} = \prod_{n\in\mathbb{S}_{n_1}} R^{(p,n)}$ for $n_1 = 1$ and 2.

In the case of a block PARAFAC model, (46) is replaced by

$$\mathcal{X}^{(p)} = \mathcal{I}_{N,R^{(p)}} \times_{n=1}^{N} \mathbf{A}^{(p,n)} \text{ with } \mathbf{A}^{(p,n)} \in \mathbb{C}^{I_n \times R^{(p)}}$$

and the matrix representation (37) then becomes

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n\in\mathbb{S}_1}{\diamond_b} \mathbf{A}^{(n)} \right) \left(\underset{n\in\mathbb{S}_2}{\diamond_b} \mathbf{A}^{(n)} \right)^T$$

with $\mathbf{A}^{(n)} = \left[\mathbf{A}^{(1,n)} \cdots \mathbf{A}^{(P,n)} \right] \in \mathbb{C}^{I_n \times R}$, and $R = \sum_{p=1}^{P} R^{(p)}$.

Block constrained PARAFAC models were used in [70-72] for modeling different types of multiuser wireless communication systems. Block constrained Tucker models were used for space-time multiplexing MIMO-OFDM systems [73] and for blind beamforming [74]. In these applications, the symbol matrix factor is in Toeplitz or block-Toeplitz form.

The block tensor model defined by (45) and (46) can be viewed as a generalization of the block term decomposition introduced in [75] for third-order tensors $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ that are decomposed into a sum of P Tucker
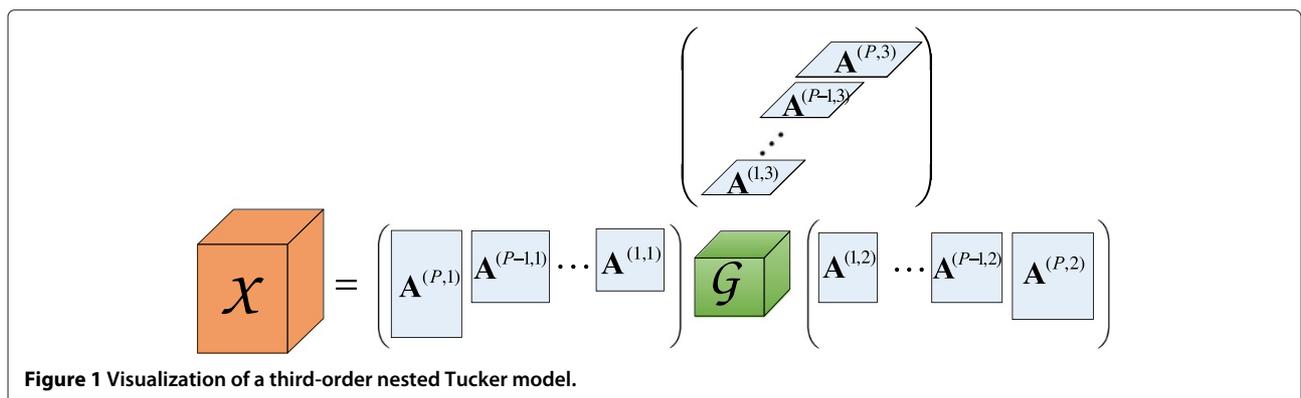

Figure 1 Visualization of a third-order nested Tucker model.

models of rank-(L, M, N), which corresponds to the particular case where all the factor matrices are full column rank, with $\mathbf{A}^{(p,1)} \in \mathbb{C}^{I \times L}$, $\mathbf{A}^{(p,2)} \in \mathbb{C}^{J \times M}$, $\mathbf{A}^{(p,3)} \in \mathbb{C}^{K \times N}$, for $p = 1, \ldots, P$, $\mathcal{G} \in \mathbb{C}^{L \times M \times N}$, and each sub-tensor $\mathcal{X}^{(p)}$ is decomposed by means of its HOSVD.

A third-order block PARALIND/CONFAC model is visualized in Figure 2. This figure is to be compared with Figure five in [76] representing a block term decomposition of a third-order tensor into rank-(L_p, M_p, N_p) terms, when each term has a PARALIND/CONFAC structure.

1.3.5 PARALIND/CONFAC-(N_1, N) models

Now, we introduce a variant of PARALIND/CONFAC models that we shall call PARALIND/CONFAC-(N_1, N) models. This variant corresponds to PARALIND/CONFAC models (42) with only N_1 constrained matrix factors, which implies $R_n = R$ and $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ for $n = N_1 + 1, \ldots, N$

$$\mathcal{X} = \mathcal{I}_{N,R} \times_{n=1}^{N_1} \left(\mathbf{A}^{(n)} \mathbf{\Phi}^{(n)} \right) \times_{n=N_1+1}^{N} \mathbf{A}^{(n)}. \tag{49}$$

In [77], a block PARALIND/CONFAC-(2,3) model that can be deduced from (49) was used for modeling uplink multiple-antenna code division multiple access (CDMA) multiuser systems.

The block term decomposition (BTD) in rank-$(1, L_p, L_p)$ terms of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, which is compared to a third-order PARATREE model in [78], can also be viewed as a particular CONFAC-(1,3) model. Indeed, such a decomposition can be written as [79]

$$\mathcal{X} = \sum_{p=1}^{P} \mathbf{a}_p \circ (\mathbf{B}_p \mathbf{C}_p^T) \tag{50}$$

where the matrices $\mathbf{B}_p \in \mathbb{C}^{J \times L_p}$ and $\mathbf{C}_p \in \mathbb{C}^{K \times L_p}$ are rank-L_p, and $\mathbf{a}_p \in \mathbb{C}^{I \times 1}$. Defining $\mathbf{B} = [\mathbf{B}_1 \cdots \mathbf{B}_P] \in \mathbb{C}^{J \times R}$, $\mathbf{C} = [\mathbf{C}_1 \cdots \mathbf{C}_P] \in \mathbb{C}^{K \times R}$, and $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_P] \in \mathbb{C}^{I \times P}$, with $R = \sum_{p=1}^{P} L_p$, it is easy to verify that the BTD (50) can be rewritten as the following CONFAC-(1,3) model:

$$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A}\mathbf{\Phi} \times_2 \mathbf{B} \times_3 \mathbf{C} \tag{51}$$

with the constraint matrix $\mathbf{\Phi} = \begin{bmatrix} \mathbf{1}_{L_1}^T & & \\ & \ddots & \\ & & \mathbf{1}_{L_P}^T \end{bmatrix} \in \mathbb{C}^{P \times R}$.

1.3.6 PARATUCK models

A PARATUCK-(N_1, N) model for a Nth-order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, with $N > N_1$, is defined in scalar form as follows [66,67]:

$$x_{i_1, \ldots, i_{N_1+1}, \ldots, i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} c_{r_1, \ldots, r_{N_1}, i_{N_1+2}, \ldots, i_N} \\ \times \prod_{n=1}^{N_1} a_{i_n, r_n}^{(n)} \phi_{r_n, i_{N_1+1}}^{(n)} \tag{52}$$

where $a_{i_n, r_n}^{(n)}$ and $\phi_{r_n, i_{N_1+1}}^{(n)}$ are entries of the factor matrix $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$ and of the interaction/allocation matrix $\mathbf{\Phi}^{(n)} \in \mathbb{C}^{R_n \times I_{N_1+1}}$, $\forall n = 1, \ldots, N_1$, respectively, and $\mathcal{C} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_{N_1+2} \times \cdots \times I_N}$ is the $(N-1)$th-order input tensor. Defining the core tensor $\mathcal{G} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_{N_1+1} \times \cdots \times I_N}$ element-wise as

$$g_{r_1, \ldots, r_{N_1}, i_{N_1+1}, \ldots, i_N} = c_{r_1, \ldots, r_{N_1}, i_{N_1+2}, \ldots, i_N} \prod_{n=1}^{N_1} \phi_{r_n, i_{N_1+1}}^{(n)},$$

the PARATUCK-(N_1, N) model can be rewritten as a Tucker-(N_1, N) model (29)-(30).

Defining the allocation/interaction tensor $\mathcal{F} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_{N_1+1}}$ of order $N_1 + 1$, such as

$$f_{r_1, \ldots, r_{N_1}, i_{N_1+1}} = \prod_{n=1}^{N_1} \phi_{r_n, i_{N_1+1}}^{(n)}, \tag{53}$$

the core tensor $\mathcal{G}$ can then be written as the Hadamard product of the tensors $\mathcal{C}$ and $\mathcal{F}$ along their first N_1 modes

$$\mathcal{G} = \mathcal{C} \underset{\{r_1, \ldots, r_{N_1}\}}{\odot} \mathcal{F}. \tag{54}$$

Remarks

- The PARATUCK-($N1$, N) model can be interpreted as the transformation of the input tensor $\mathcal{C}$ via its multiplication by the factor matrices $\mathbf{A}^{(n)}, n = 1, \ldots, N_1$, along its first N_1 modes,
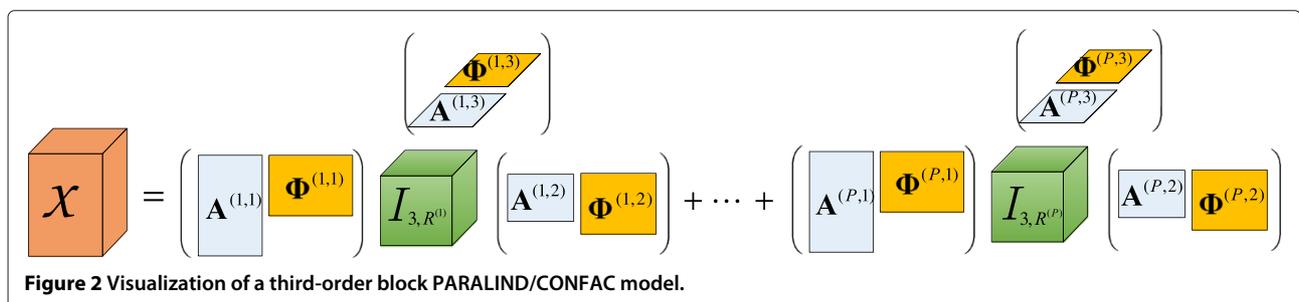

Figure 2 Visualization of a third-order block PARALIND/CONFAC model.

combined with a mode-n resource allocation
($n = 1, \ldots, N_1$) relatively to the mode-$(N_1 + 1)$ of
the transformed tensor $\mathcal{X}$, by means of the allocation
matrices $\mathbf{\Phi}^{(n)}$.

- In telecommunications applications, the output
  modes will be called diversity modes because they
  correspond to time, space, and frequency diversities,
  whereas the input modes are associated with
  resources like transmit antennas, codes, and data
  streams. For these applications, the matrices $\mathbf{\Phi}^{(n)}$ are
  formed with 0's and 1's, and they can be interpreted
  as allocation matrices used for allocating some
  resources r_n to the output mode-$(N_1 + 1)$. Another
  way to take resource allocations into account consists
  in replacing the N_1 allocation matrices $\mathbf{\Phi}^{(n)}$ by the
  $(N_1 + 1)$th-order allocation tensor
  $\mathcal{F} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_{N_1+1}}$ defined in (53).
- Special cases:

  - For $N_1 = 2$ and $N = 3$, we obtain the
    standard PARATUCK-2 model introduced in
    [59]. (52) then becomes

    $$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} c_{r_1,r_2} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)} \phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)}$$

    (55)

    The allocation tensor $\mathcal{F}$ defined in (53) can be
    rewritten as

    $$f_{r_1,r_2,i_3} = \phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)} = \sum_{j=1}^{I_3} \phi_{r_1,j}^{(1)} \phi_{r_2,j}^{(2)} \delta_{i_3,j}$$

    (56)

    which corresponds to a PARAFAC model
    with matrix factors $(\mathbf{\Phi}^{(1)}, \mathbf{\Phi}^{(2)}, \mathbf{I}_{I_3})$. The
    PARATUCK-2 model (55) can then be viewed
    as a Tucker-2 model $\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)}$
    with the core tensor $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times I_3}$ given by
    the Hadamard product of $\mathbf{C} \in \mathbb{C}^{R_1 \times R_2}$ and
    $\mathcal{F} \in \mathbb{C}^{R_1 \times R_2 \times I_3}$ along their common modes
    $\{r_1, r_2\}$

    $$\mathcal{G} = \mathbf{C} \underset{\{r_1,r_2\}}{\odot} \mathcal{F}$$

    This combination of a Tucker-2 model for $\mathcal{X}$
    with a PARAFAC model for $\mathcal{F}$ gave rise to the
    name PARATUCK-2. The constraint matrices
    $(\mathbf{\Phi}^{(1)}, \mathbf{\Phi}^{(2)})$ define interactions between
    columns of the factor matrices $(\mathbf{A}^{(1)}, \mathbf{A}^{(2)})$,
    along the mode-3 of $\mathcal{X}$, while the matrix $\mathbf{C}$
    contains the weights of these interactions.

  - For $N_1 = 2$ and $N = 4$, we obtain the
    PARATUCK-(2,4) model introduced in [66]

    $$x_{i_1,i_2,i_3,i_4} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} c_{r_1,r_2,i_4} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)} \phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)}$$

    (57)

    As for the PARATUCK-2 model, the
    PARATUCK-(2,4) can be viewed as a
    combination of a Tucker-(2,4) model
    $\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ with a
    core tensor $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times I_3 \times I_4}$ given by the
    Hadamard product of the tensors
    $\mathcal{C} \in \mathbb{C}^{R_1 \times R_2 \times I_4}$ and $\mathcal{F} \in \mathbb{C}^{R_1 \times R_2 \times I_3}$ along their
    common modes $\{r_1, r_2\}$

    $$\mathcal{G} = \mathcal{C} \underset{\{r_1,r_2\}}{\odot} \mathcal{F}$$

    with the same allocation tensor $\mathcal{F}$ defined in
    (56).

1.3.7 Rewriting of PARATUCK models as constrained PARAFAC models

This rewriting of PARATUCK models as constrained
PARAFAC models can be used to deduce both matrix
unfoldings by means of the general formula (37) and
sufficient conditions for essential uniqueness of such
PARATUCK models, as will be shown in Section 1.4.

**Link between PARATUCK-(2,4) and constrained
PARAFAC-4 models** We now establish the link between
the PARATUCK-(2,4) model (57) and the fourth-order
constrained PARAFAC model

$$x_{i_1,i_2,i_3,i_4} = \sum_{r=1}^{R} a_{i_1,r} b_{i_2,r} f_{i_3,r} d_{i_4,r} \quad \text{with} \quad R = R_1 R_2$$

(58)

whose matrix factors $(\mathbf{A} \in \mathbb{C}^{I_1 \times R}, \mathbf{B} \in \mathbb{C}^{I_2 \times R}, \mathbf{F} \in \mathbb{C}^{I_3 \times R}, \mathbf{D} \in \mathbb{C}^{I_4 \times R})$, and constraint matrices $(\mathbf{\Psi}^{(1)}, \mathbf{\Psi}^{(2)})$ acting on
the original factors $(\mathbf{A}^{(1)}, \mathbf{A}^{(2)})$, are given by

$$\mathbf{A} = \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)}, \quad \mathbf{B} = \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)}, \quad \mathbf{F} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T,$$
$$\mathbf{D} = \mathbf{C}_{I_4 \times R_1 R_2}$$

(59)

$$\mathbf{\Psi}^{(1)} = \mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^T \in \mathbb{C}^{R_1 \times R_1 R_2},$$
$$\mathbf{\Psi}^{(2)} = \mathbf{1}_{R_1}^T \otimes \mathbf{I}_{R_2} \in \mathbb{C}^{R_2 \times R_1 R_2}$$

(60)

where $\mathbf{C}_{I_4 \times R_1 R_2} \in \mathbb{C}^{I_4 \times R_1 R_2}$ is a mode-3 unfolded matrix
of the tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times R_2 \times I_4}$.

Proof. See Appendix 5. □

Remarks

- Application of the formula (38) to the constrained PARAFAC model (58), with the matrix factors $(\mathbf{A}, \mathbf{B}, \mathbf{F}, \mathbf{D}) = (\mathbf{A}^{(1)} \mathbf{\Psi}^{(1)}, \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)}, (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T, \mathbf{C}_{I_4 \times R_1 R_2})$, gives the following flat mode-1 and mode-2 matrix unfoldings for the PARATUCK-(2,4) model (57)

$$\mathbf{X}_{I_1 \times I_2 I_3 I_4} = \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \left(\mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} \diamond \mathbf{F} \diamond \mathbf{D} \right)^T \in \mathbb{C}^{I_1 \times I_2 I_3 I_4},$$

$$\mathbf{X}_{I_2 \times I_3 I_4 I_1} = \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} \left(\mathbf{F} \diamond \mathbf{D} \diamond \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \right)^T \in \mathbb{C}^{I_2 \times I_3 I_4 I_1}.$$

- The constrained PARAFAC-4 model (58)-(60) can be written in mode-n product notation as

$$\mathcal{X} = \mathcal{I}_{4,R} \times_1 \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \times_2 \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} \times_3 \mathbf{F} \times_4 \mathbf{D}. \quad (61)$$

Defining the core tensor $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times I_3 \times I_4}$ as

$$\mathcal{G} = \mathcal{I}_{4,R} \times_1 \mathbf{\Psi}^{(1)} \times_2 \mathbf{\Psi}^{(2)} \times_3 \mathbf{F} \times_4 \mathbf{D} \quad (62)$$

the constrained PARAFAC-4 model can also be viewed as the following Tucker-(2,4) model

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)}. \quad (63)$$

It can also be viewed as a CONFAC-(2,4) model with matrix factors $(\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{F}, \mathbf{D})$, and constraint matrices $\mathbf{\Psi}^{(1)}$ and $\mathbf{\Psi}^{(2)}$ defined in (60).

- Choosing $\mathbb{S}_1 = \{i_1, i_2\}$ and $\mathbb{S}_2 = \{i_3, i_4\}$, the matrix unfolding (37) of the PARAFAC model (61) is given by

$$\mathbf{X}_{I_1 I_2 \times I_3 I_4} = \left(\mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \diamond \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} \right) (\mathbf{F} \diamond \mathbf{D})^T$$

$$= \left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)} \right) (\mathbf{F} \diamond \mathbf{D})^T \in \mathbb{C}^{I_1 I_2 \times I_3 I_4}$$

$$(64)$$

Proof. Using the identity (90) gives

$$\mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \diamond \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} = \left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)} \right) \left(\mathbf{\Psi}^{(1)} \diamond \mathbf{\Psi}^{(2)} \right) \quad (65)$$

Replacing $\mathbf{\Psi}^{(1)}$ and $\mathbf{\Psi}^{(2)}$ by their expressions (102) and (103) leads to

$$\mathbf{\Psi}^{(1)} \diamond \mathbf{\Psi}^{(2)} = \left(\mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^T \right) \diamond \left(\mathbf{1}_{R_1}^T \otimes \mathbf{I}_{R_2} \right)$$

$$= \underbrace{\begin{bmatrix} \mathbf{I}_{R_2} & & \\ & \ddots & \\ & & \mathbf{I}_{R_2} \end{bmatrix}}_{R_1 \text{ blocks}} = \mathbf{I}_{R_1 R_2} \quad (66)$$

which implies

$$\mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \diamond \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} = \mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}, \quad (67)$$

and consequently (64) can be deduced.

This equation can also be obtained from the equivalent Tucker-(2,4) model (62)-(63) as

$$\mathbf{X}_{I_1 I_2 \times I_3 I_4} = \left(\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)} \right) \mathbf{G}_{R_1 R_2 \times I_3 I_4} \quad (68)$$

with

$$\mathbf{G}_{R_1 R_2 \times I_3 I_4} = \left(\mathbf{\Psi}^{(1)} \diamond \mathbf{\Psi}^{(2)} \right) (\mathbf{F} \diamond \mathbf{D})^T$$

Using the identity (66), we obtain

$$\mathbf{G}_{R_1 R_2 \times I_3 I_4} = (\mathbf{F} \diamond \mathbf{D})^T \quad (69)$$

and replacing $\mathbf{G}_{R_1 R_2 \times I_3 I_4}$ by its expression (69) into (68) gives (64). $\square$

When the allocation matrices $\left(\mathbf{\Phi}^{(1)}, \mathbf{\Phi}^{(2)} \right)$ and the input tensor $\mathcal{C}$ are known, the matrix factors $\left(\mathbf{A}^{(1)}, \mathbf{A}^{(2)} \right)$ can be estimated through the LS estimation of their Kronecker product using the matrix unfolding (64).

- The product $\phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)}$ in (57) can be replaced by f_{i_3,r_1,r_2}, which amounts to replace the allocation matrices $\mathbf{\Phi}^{(1)}$ and $\mathbf{\Phi}^{(2)}$ by the third-order allocation tensor $\mathcal{F} \in \mathbb{C}^{I_3 \times R_1 \times R_2}$, the matrix $\mathbf{F} = \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)} \right)^T \in \mathbb{C}^{I_3 \times R_1 R_2}$ being equivalent to $\mathbf{F}_{I_3 \times R_1 R_2} \in \mathbb{C}^{I_3 \times R_1 R_2}$, i.e., a mode-1 flat matrix unfolding of the allocation tensor $\mathcal{F}$.

Link between PARATUCK-2 and constrained PARAFAC-3 models By proceeding in the same way as for the PARATUCK-(2,4) model, it is easy to show that the PARATUCK-2 model (55) is equivalent to a third-order constrained PARAFAC model whose matrix factors $\mathbf{A} \in \mathbb{C}^{I_1 \times R}$, $\mathbf{B} \in \mathbb{C}^{I_2 \times R}$, and $\mathbf{F} \in \mathbb{C}^{I_3 \times R}$, with $R = R_1 R_2$, are given by

$$\mathbf{A} = \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)}, \quad \mathbf{B} = \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)},$$

$$\mathbf{F} = \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)} \right)^T \text{diag} \left(\text{vec} \left(\mathbf{C}^T \right) \right) \quad (70)$$

with the same constraint matrices $\mathbf{\Psi}^{(1)}$ and $\mathbf{\Psi}^{(2)}$ defined in (60).

By analogy with the PARATUCK-(2,4) model, (61), (63), and (64) become for the PARATUCK-2 model

$$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)} \times_2 \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} \times_3 \mathbf{F}$$

$$= \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \quad (71)$$

with the core tensor $\mathcal{G} \in \mathbb{C}^{R_1 \times R_2 \times I_3}$ defined as

$$\mathcal{G} = \mathcal{I}_{3,R} \times_1 \mathbf{\Psi}^{(1)} \times_2 \mathbf{\Psi}^{(2)} \times_3 \mathbf{F}, \quad (72)$$

and

$$\mathbf{X}_{I_1 I_2 \times I_3} = (\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}) \mathbf{F}^T \in \mathbb{C}^{I_1 I_2 \times I_3}.$$

Remarks

- Note that (71) and (72) allow interpreting the PARATUCK-2 model as a Tucker-(2,3) model, defined in (31)-(32). If we choose $c_{r_1,r_2} = 1, \forall r_k = 1, \ldots, R_k$, for $k = 1$ and 2 and define the allocation tensor $\mathcal{F} \in \mathbb{C}^{R_1 \times R_2 \times I_3}$ such as $f_{r_1,r_2,i_3} = \phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)}$, the PARATUCK-2 model (55) becomes the following Tucker-(2,3) model:

$$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} f_{r_1,r_2,i_3} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)}$$

and the associated constrained PARAFAC-3 model can be deduced from (70)

$$\mathbf{A} = \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)}, \quad \mathbf{B} = \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)},$$
$$\mathbf{F} = \mathbf{F}_{I_3 \times R_1 R_2} = \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)} \right)^T$$

with the same constraint matrices $\mathbf{\Psi}^{(1)}$ and $\mathbf{\Psi}^{(2)}$ as those defined in (60). A block Tucker-(2,3) model transformed into a block constrained PARAFAC-3 model was used in [72] for modeling in a unified way three multiuser wireless communication systems.

- Now, we show the equivalence of the expressions (72) and (54) of the core tensor. Applying the formula (38) to the PARAFAC model (72) gives

$$\mathbf{G}_{I_3 \times R_1 R_2} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T \mathrm{diag}(\mathrm{vec}(\mathbf{C}^T))(\mathbf{\Psi}^{(1)} \diamond \mathbf{\Psi}^{(2)})^T. \tag{73}$$

Using the identity (66) in (73) gives
$\mathbf{G}_{I_3 \times R_1 R_2} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T \mathrm{diag}(\mathrm{vec}(\mathbf{C}^T))$.
For the formula (54), with $N = 3$ and $N_1 = 2$, we have

$$\mathcal{G} = \mathcal{F} \underset{\{r_1,r_2\}}{\odot} \mathbf{C}$$

or equivalently in terms of matrix Hadamard product

$$\mathbf{G}_{I_3 \times R_1 R_2} = \mathbf{F}_{I_3 \times R_1 R_2} \odot \mathbf{1}_{I_3} \mathbf{c}^T$$

with $\mathbf{F}_{I_3 \times R_1 R_2} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T$, and $\mathbf{c} = \mathrm{vec}(\mathbf{C}^T) \in \mathbb{C}^{R_1 R_2 \times 1}$, which gives

$$\mathbf{G}_{I_3 \times R_1 R_2} = \mathbf{F}_{I_3 \times R_1 R_2} \odot \left. \begin{bmatrix} \mathbf{c}^T \\ \vdots \\ \mathbf{c}^T \end{bmatrix} \right\} I_3 \text{ rows}$$

and consequently $\mathbf{G}_{I_3 \times R_1 R_2} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T \mathrm{diag}(\mathbf{c})$, showing the equivalence of the two core tensor expressions (72) and (54).

Link between PARATUCK-$(N-2, N)$ and constrained PARAFAC-N models Let us consider the PARATUCK-(N_1, N) model (52) in the case $N_1 = N - 2$

$$x_{i_1,\ldots,i_{N_1+1},\ldots,i_N} = \sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} c_{r_1,\ldots,r_{N_1},i_N} \prod_{n=1}^{N_1} a_{i_n,r_n}^{(n)} \phi_{r_n,i_{N_1+1}}^{(n)} \tag{74}$$

and let us define the change of variables $r = r_{N_1} + \sum_{n=1}^{N_1-1} (r_n - 1) \prod_{i=n+1}^{N_1} R_i$ corresponding to a combination of the N_1 modes associated with the constraints/allocations. Then, (74) can be written as the following constrained PARAFAC-N model:

$$x_{i_1,\ldots,i_N} = \sum_{r=1}^{R} \prod_{n=1}^{N} \bar{a}_{i_n,r}^{(n)}, \quad R = \prod_{i=1}^{N_1} R_i \tag{75}$$

with the following matrix factors

$$\bar{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \mathbf{\Psi}^{(n)}, n = 1, \ldots, N_1; \quad \mathbf{F} = \left(\underset{n=1}{\overset{N}{\diamond}} \mathbf{\Phi}^{(n)} \right)^T;$$
$$\mathbf{D} = \mathbf{C}_{I_N \times R_1 \cdots R_{N_1}},$$

where $\mathbf{C}_{I_N \times R_1 \cdots R_{N_1}} \in \mathbb{C}^{I_N \times R_1 \cdots R_{N_1}}$ is a mode-$(N_1 + 1)$ unfolded matrix of the tensor $\mathcal{C} \in \mathbb{C}^{R_1 \times \cdots \times R_{N_1} \times I_N}$, and the constraint matrices are given in (94) as

$$\mathbf{\Psi}^{(n)} = \mathbf{1}_{R_1}^T \otimes \cdots \otimes \mathbf{1}_{R_{n-1}}^T \otimes \mathbf{I}_{R_n} \otimes \mathbf{1}_{R_{n+1}}^T \otimes \cdots \otimes \mathbf{1}_{R_{N_1}}^T$$
$$\in \mathbb{C}^{R_n \times R}, n = 1, \ldots, N_1.$$

The constrained PARAFAC model (75) can also be written as a Tucker-(N_1, N) model (30) with the core tensor defined in (54) or, equivalently,

$$\mathcal{G} = \mathcal{I}_{N,R} \times_{n=1}^{N-2} \mathbf{\Psi}^{(n)} \times_{N-1} \mathbf{F} \times_N \mathbf{D}.$$

1.3.8 Comparison of constrained tensor models
To conclude this presentation, we compare the so-called CONFAC-(N_1, N) and PARATUCK-(N_1, N) constrained tensor models, introduced in this paper with a resource allocation point of view. Due to the PARAFAC structure (41) of the core tensor of CONFAC models, each element $x_{i_1,\ldots,i_N}$ of the output tensor $\mathcal{X}$ is the sum of R components

as shown in (43). Moreover, due to the special structure of the allocation matrices $\mathbf{\Phi}^{(n)}$ whose columns are unit vectors, each component r is the result of a combination of N resources, under the form of the product $\prod_{n=1}^{N} a_{i_n,r_n}^{(n)}$, the N resources being fixed by the allocation matrices $\mathbf{\Phi}^{(n)} \in \mathbb{C}^{R_n \times R}$.

With the CONFAC-(N_1, N) model (49), each component r is a combination of N_1 resources $(r_1, \ldots, r_{N_1})$ determined by the allocation matrices $\mathbf{\Phi}^{(n)} \in \mathbb{C}^{R_n \times R}$ for $n = 1, \ldots, N_1$.

There are two main differences between the PARATUCK-(N_1, N) models (52) and the CONFAC models (42). The first one is that the allocation matrices of PARATUCK models, formed with 0's and 1's, have not necessarily unit vectors as column vectors, which means that it is possible to allocate $\gamma_n = \sum_{r_n=1}^{R_n} \phi_{r_n,i_{N_1+1}}^{(n)}$ resources r_n to the $(N_1 + 1)$th mode of the output tensor $\mathcal{X}$. The second one results from the interpretation of PARATUCK-(N_1, N) models as Tucker-(N_1, N) models, implying that each element $x_{i_1,\ldots,i_N}$ of $\mathcal{X}$ is equal to the sum of $\sum_{r_1=1}^{R_1} \cdots \sum_{r_{N_1}=1}^{R_{N_1}} f_{r_1,\ldots,r_{N_1},i_{N_1+1}}$ terms, where $f_{r_1,\ldots,r_{N_1},i_{N_1+1}}$ is an entry of the allocation tensor $\mathcal{F}$ defined in (53), each term being a combination of resources under the form of products $\prod_{n=1}^{N_1} a_{i_n,r_n}^{(n)}$. Moreover, in

telecommunication applications, the input tensor $\mathcal{C}$ can be used as a code tensor.

Another way to compare PARALIND/CONFAC and PARATUCK models is in terms of dependencies/interactions between their factor matrices. In the case of PARALIND/CONFAC models, as pointed out by (42), the constraint matrices act independently on each factor matrix, expliciting linear dependencies between columns of these matrices. For PARATUCK models, their writing as Tucker-(N_1, N) models with the core tensor defined in (54) allows to interpret the tensor $\mathcal{F}$ as an interaction tensor which defines interactions between N_1 factor matrices, the tensor $\mathcal{C}$ providing the strength of these interactions.

The main constrained PARAFAC models are summarized in Tables 1 and 2.

1.4 Uniqueness issue

Several results exist for essential uniqueness of PARAFAC models, i.e., uniqueness of factor matrices up to column permutation and scaling. These results concern both deterministic and generic uniqueness, i.e., uniqueness for a particular PARAFAC model or uniqueness with probability one in the case where the entries of the factor matrices are drawn from continuous distributions. An overview of main uniqueness conditions of PARAFAC models of third-order tensors can be found in [80] for the deterministic case and in [81] for the generic case. Hereafter, we briefly summarize some basic results on

Table 1 Main tensor models

Models	Scalar writings	Mode-n product-based writings
PARAFAC-3	$x_{i_1,i_2,i_3} = \sum_{r}^{R} a_{i_1,r}^{(1)} a_{i_2,r}^{(2)} a_{i_3,r}^{(3)}$	$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)}$
Tucker-3	$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} g_{r_1,r_2,r_3} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)} a_{i_3,r_3}^{(3)}$	$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)}$
Tucker-(2,3)	$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_3} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)}$	$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)}$
PARALIND/CONFAC-3	$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} g_{r_1,r_2,r_3} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)} a_{i_3,r_3}^{(3)}$	$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)}$
	$g_{r_1,r_2,r_3} = \sum_{r}^{R} \varphi_{r_1,r}^{(1)} \varphi_{r_2,r}^{(2)} \varphi_{r_3,r}^{(3)}$	$\mathcal{G} = \mathcal{I}_{3,R} \times_1 \mathbf{\Phi}^{(1)} \times_2 \mathbf{\Phi}^{(2)} \times_3 \mathbf{\Phi}^{(3)}$
PARATUCK-2	$x_{i_1,i_2,i_3} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_3} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)}$	$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)}$
	$g_{r_1,r_2,i_3} = c_{r_1,r_2} \varphi_{r_1,i_3}^{(1)} \varphi_{r_2,i_3}^{(2)}$	$\mathcal{G} = \mathcal{I}_{3,R} \times_1 \mathbf{\Psi}^{(1)} \times_2 \mathbf{\Psi}^{(2)} \times_3 \mathbf{F}, \quad R = R_1 R_2$
		$\mathbf{\Psi}^{(1)} = \mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^{T}, \quad \mathbf{\Psi}^{(2)} = \mathbf{1}_{R_1}^{T} \otimes \mathbf{I}_{R_2}$
		$\mathbf{F} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^{T} \text{diag}(\text{vec}(\mathbf{C}^{T}))$
PARATUCK-(2,4)	$x_{i_1,i_2,i_3,i_4} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} g_{r_1,r_2,i_3,i_4} a_{i_1,r_1}^{(1)} a_{i_2,r_2}^{(2)}$	$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)}$
	$g_{r_1,r_2,i_3,i_4} = c_{r_1,r_2,i_4} \varphi_{r_1,i_3}^{(1)} \varphi_{r_2,i_3}^{(2)}$	$\mathcal{G} = \mathcal{I}_{4,R} \times_1 \mathbf{\Psi}^{(1)} \times_2 \mathbf{\Psi}^{(2)} \times_3 \mathbf{F} \times_4 \mathbf{D}, \quad R = R_1 R_2$
		$\mathbf{\Psi}^{(1)} = \mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^{T}, \quad \mathbf{\Psi}^{(2)} = \mathbf{1}_{R_1}^{T} \otimes \mathbf{I}_{R_2}$
		$\mathbf{F} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^{T}, \quad \mathbf{D} = \mathbf{C}_{I_4 \times R_1 R_2}$

Table 2 Equivalent constrained PARAFAC models

Models	Equivalent constrained PARAFAC model	Matrix unfoldings
PARAFAC-3		$\mathbf{X}_{I_1 \times I_2 I_3} = \mathbf{A}^{(1)}(\mathbf{A}^{(2)} \diamond \mathbf{A}^{(3)})^T$
Tucker-3		$\mathbf{X}_{I_1 \times I_2 I_3} = \mathbf{A}^{(1)}\mathbf{G}_{R_1 \times R_2 R_3}(\mathbf{A}^{(2)} \otimes \mathbf{A}^{(3)})^T$
Tucker-(2,3)	$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A}^{(1)}\mathbf{\Psi}^{(1)} \times_2 \mathbf{A}^{(2)}\mathbf{\Psi}^{(2)} \times_3 \mathbf{G}_{I_3 \times R_1 R_2}$	$\mathbf{X}_{I_1 \times I_2 I_3} = \mathbf{A}^{(1)}\mathbf{\Psi}^{(1)}(\mathbf{A}^{(2)}\mathbf{\Psi}^{(2)} \diamond \mathbf{G}_{I_3 \times R_1 R_2})^T$
	$\mathbf{\Psi}^{(1)} = \mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^T, \quad \mathbf{\Psi}^{(2)} = \mathbf{1}_{R_1}^T \otimes \mathbf{I}_{R_2}$	
PARALIND/CONFAC-3	$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A}^{(1)}\mathbf{\Phi}^{(1)} \times_2 \mathbf{A}^{(2)}\mathbf{\Phi}^{(2)} \times_3 \mathbf{A}^{(3)}\mathbf{\Phi}^{(3)}$	$\mathbf{X}_{I_1 \times I_2 I_3} = \mathbf{A}^{(1)}\mathbf{\Phi}^{(1)}(\mathbf{A}^{(2)}\mathbf{\Phi}^{(2)} \diamond \mathbf{A}^{(3)}\mathbf{\Phi}^{(3)})^T$
		$= \mathbf{A}^{(1)}\mathbf{\Phi}^{(1)}(\mathbf{\Phi}^{(2)} \diamond \mathbf{\Phi}^{(3)})^T(\mathbf{A}^{(2)} \otimes \mathbf{A}^{(3)})^T$
PARATUCK-2	$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A}^{(1)}\mathbf{\Psi}^{(1)} \times_2 \mathbf{A}^{(2)}\mathbf{\Psi}^{(2)} \times_3 \mathbf{F}$	
	$\mathbf{\Psi}^{(1)} = \mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^T, \quad \mathbf{\Psi}^{(2)} = \mathbf{1}_{R_1}^T \otimes \mathbf{I}_{R_2}$	$\mathbf{X}_{I_1 \times I_2 I_3} = \mathbf{A}^{(1)}\mathbf{\Psi}^{(1)}(\mathbf{A}^{(2)}\mathbf{\Psi}^{(2)} \diamond \mathbf{F})^T$
	$\mathbf{F} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T \text{diag}(\text{vec}(\mathbf{C}^T))$	
PARATUCK-(2,4)	$\mathcal{X} = \mathcal{I}_{4,R} \times_1 \mathbf{A}^{(1)}\mathbf{\Psi}^{(1)} \times_2 \mathbf{A}^{(2)}\mathbf{\Psi}^{(2)} \times_3 \mathbf{F} \times_4 \mathbf{D}$	$\mathbf{X}_{I_1 \times I_2 I_3 I_4} = \mathbf{A}^{(1)}\mathbf{\Psi}^{(1)}(\mathbf{A}^{(2)}\mathbf{\Psi}^{(2)} \diamond \mathbf{F} \diamond \mathbf{C}_{I_4 \times R_1 R_2})^T$
	$\mathbf{\Psi}^{(1)} = \mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^T, \quad \mathbf{\Psi}^{(2)} = \mathbf{1}_{R_1}^T \otimes \mathbf{I}_{R_2}$	$\mathbf{X}_{I_1 I_2 \times I_3 I_4} = (\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)})$
	$\mathbf{F} = (\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T, \quad \mathbf{D} = \mathbf{C}_{I_4 \times R_1 R_2}$	$\left((\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)})^T \diamond \mathbf{C}_{I_4 \times R_1 R_2}\right)^T$

uniqueness of PARAFAC models. The case with linearly dependent loadings is also discussed. Then, we present new results concerning the uniqueness of PARATUCK models. These results are directly deduced from sufficient conditions for essential uniqueness of their associated constrained PARAFAC models, as established in the previous section. As these conditions involve the notion of k-rank of a matrix, we first recall the definition of k-rank.

Definition of k-rank

The k-rank (also called Kruskal's rank) of a matrix $\mathbf{A} \in \mathbb{C}^{I \times R}$, denoted by $k_\mathbf{A}$, is the largest integer such that any set of $k_\mathbf{A}$ columns of $\mathbf{A}$ is linearly independent.

It is obvious that $k_\mathbf{A} \leq r_\mathbf{A}$.

1.4.1 Uniqueness of PARAFAC-N models [82]

The PARAFAC-N model (33)-(35) is essentially unique, i.e., its factor matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}, n = 1, \ldots, N$, are unique up to column permutation and scaling, if

$$\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + N - 1 \tag{76}$$

Essential uniqueness means that two sets of factor matrices are linked by the following relations $\hat{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \mathbf{\Pi} \mathbf{\Lambda}^{(n)}$, for $n = 1, \ldots, N$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}^{(n)}$ are non-singular diagonal matrices such as $\prod_{n=1}^{N} \mathbf{\Lambda}^{(n)} = \mathbf{I}_R$.

In the generic case, the factor matrices are full rank, with $k_{\mathbf{A}^{(n)}} = r_{\mathbf{A}^{(n)}} = \min(I_n, R)$, and Kruskal's condition (76) becomes

$$\sum_{n=1}^{N} \min(I_n, R) \geq 2R + N - 1 \tag{77}$$

Case of third-order PARAFAC models. Consider a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ of rank R, satisfying a PARAFAC model with matrix factors $(\mathbf{A}, \mathbf{B}, \mathbf{C})$. Kruskal's condition (76) becomes

$$k_\mathbf{A} + k_\mathbf{B} + k_\mathbf{C} \geq 2R + 2 \tag{78}$$

Remarks

- The condition (76) is sufficient but not necessary for essential uniqueness. This condition does not hold when $R = 1$. It is also necessary for $R = 2$ and $R = 3$ but not for $R > 3$ (see [83]).

- The first sufficient condition for essential uniqueness of third-order PARAFAC models was established by Harshman [84] and then generalized by Kruskal [52] using the concept of k-rank. A more accessible proof of Kruskal's condition is provided in [85]. Kruskal's condition was extended to complex-valued tensors in [15] and to N-way arrays, with $N > 3$, in [82].

- Necessary and sufficient uniqueness conditions more relaxed than the Kruskal's one were established for third- and fourth-order tensors, under the assumption that at least one matrix factor is full column rank [86,87]. These conditions are complicated to apply. Other more relaxed conditions have been derived independently by Stegeman [88] and Guo et al. [89], for third-order PARAFAC models with a full column rank matrix factor.

- From the condition (78), we can conclude that if two matrix factors ($\mathbf{A}$ and $\mathbf{B}$) are full column rank ($k_\mathbf{A} = k_\mathbf{B} = R$), then the PARAFAC model is essentially unique if the third matrix factor ($\mathbf{C}$) has no proportional columns ($k_\mathbf{C} > 1$).

- If one matrix factor (**C** for instance) is full column rank, then (78) gives

$$k_\mathbf{A} + k_\mathbf{B} \geq R + 2 \tag{79}$$

In [88] and [89], it is shown that the PARAFAC model (**A**, **B**, **C**), with **C** of full column rank, is essentially unique if the other two matrix factors **A** and **B** satisfy the following conditions:

1) $k_\mathbf{A}, k_\mathbf{B} \geq 2$
2) $r_\mathbf{A} + k_\mathbf{B} \geq R + 2$ or $r_\mathbf{B} + k_\mathbf{A} \geq R + 2$

$$\tag{80}$$

Conditions (80) are more relaxed than (79). Indeed, if for instance $k_\mathbf{A} = 2$ and $r_\mathbf{A} = k_\mathbf{A} + \delta$ with $\delta > 0$, application of (79) implies $k_\mathbf{B} = R$, i.e., **B** must be full column rank, whereas (80) gives $k_\mathbf{B} \geq R - \delta$ which does not require that **B** be full column rank.

- When one matrix factor (**C** for instance) is known and Kruskal's condition (78) is satisfied, as it is often the case in telecommunication applications, essential uniqueness is ensured without permutation ambiguity and with only scaling ambiguities ($\mathbf{\Lambda_A}, \mathbf{\Lambda_B}$) such as $\mathbf{\Lambda_A}\mathbf{\Lambda_B} = \mathbf{I}_R$.

1.4.2 Uniqueness of PARAFAC models with linearly dependent loadings

If one matrix factor contains at least two proportional columns, i.e., its *k*-rank is equal to one, then Kruskal's condition (78) cannot be satisfied. In this case, partial uniqueness can be ensured, i.e., some columns of some matrix factors are essentially unique while the others are unique up to multiplication by a non-singular matrix [90]. To illustrate this result, let us consider the case of the PARAFAC model of a fourth-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K \times L}$ with factor matrices (**A**, **B**, **C**, **D**) whose two of them have two identical columns at the same position

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{a} & \mathbf{a} \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{b} & \mathbf{b} \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix},$$
$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{D}_2 \end{bmatrix}$$

with $\mathbf{A}_1 \in \mathbb{C}^{I \times (R-2)}, \mathbf{a} \in \mathbb{C}^{I \times 1}, \mathbf{B}_1 \in \mathbb{C}^{J \times (R-2)}, \mathbf{b} \in \mathbb{C}^{J \times 1}, \mathbf{C}_1 \in \mathbb{C}^{K \times (R-2)}, \mathbf{C}_2 \in \mathbb{C}^{K \times 2}, \mathbf{D}_1 \in \mathbb{C}^{L \times (R-2)}, \mathbf{D}_2 \in \mathbb{C}^{L \times 2}$. We have $k_\mathbf{A} = k_\mathbf{B} = 1$, and consequently, the uniqueness condition (76) for $N = 4$ becomes $k_\mathbf{C} + k_\mathbf{D} \geq 2R + 1$, which cannot be satisfied. In this case, we have partial uniqueness. Indeed, the matrix slices (40) can be developed as follows:

$$\mathbf{X}_{ij..} = \mathbf{C}D_j(\mathbf{B})D_i(\mathbf{A})\mathbf{D}^T$$
$$= \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} D_j(\mathbf{B}_1)D_i(\mathbf{A}_1) & \mathbf{0}_{(R-2)\times 2} \\ \mathbf{0}_{2\times(R-2)} & a_i b_j \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{D}_1^T \\ \mathbf{D}_2^T \end{bmatrix}$$
$$= \mathbf{C}_1 D_j(\mathbf{B}_1)D_i(\mathbf{A}_1)\mathbf{D}_1^T + a_i b_j \mathbf{C}_2 \mathbf{D}_2^T.$$

From this expression, it is easy to conclude that the last two columns of **C** and **D** are unique up to a rotational indeterminacy. Indeed, if one replaces the matrices ($\mathbf{C}_2, \mathbf{D}_2$) by ($\mathbf{C}_2\mathbf{T}, \mathbf{D}_2\mathbf{T}^{-T}$), where $\mathbf{T} \in \mathbb{C}^{2\times 2}$ is a non-singular matrix, the matrix slices $\mathbf{X}_{ij..}$ remain unchanged. So, the PARAFAC model is said partially unique in the sense that only the blocks ($\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1, \mathbf{D}_1$) are essentially unique, the blocks $\mathbf{C}_2$ and $\mathbf{D}_2$ being unique up to a non-singular matrix. Essential uniqueness means that any alternative blocks ($\hat{\mathbf{A}}_1, \hat{\mathbf{B}}_1, \hat{\mathbf{C}}_1, \hat{\mathbf{D}}_1$) are such as $\hat{\mathbf{A}}_1 = \mathbf{A}_1\mathbf{\Pi}\mathbf{\Delta}_a, \hat{\mathbf{B}}_1 = \mathbf{B}_1\mathbf{\Pi}\mathbf{\Delta}_b, \hat{\mathbf{C}}_1 = \mathbf{C}_1\mathbf{\Pi}\mathbf{\Delta}_c, \hat{\mathbf{D}}_1 = \mathbf{D}_1\mathbf{\Pi}\mathbf{\Delta}_d$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Delta}_a$, $\mathbf{\Delta}_b$, $\mathbf{\Delta}_c$, and $\mathbf{\Delta}_d$ are diagonal matrices such as $\mathbf{\Delta}_a\mathbf{\Delta}_b\mathbf{\Delta}_c\mathbf{\Delta}_d = \mathbf{I}_{R-2}$. In [91], sufficient conditions are provided for essential uniqueness of fourth-order PARAFAC models with one full column rank factor matrix and at most three collinear factor matrices, i.e., having one (or more) column(s) proportional to another column. Note that this type of model can be interpreted as a fourth-order CONFAC model with constraints on at most three matrix factors. Uniqueness is ensured if any pair of proportional columns cannot be common to two collinear factors, which is not the case of the example above due to the fact that the two equal columns of **A** and **B** are in the same position.

The PARALIND and CONFAC models represent a class of constrained PARAFAC models where the columns of one or more matrix factors are linearly dependent or collinear. In the case of CONFAC models, such a collinearity takes the form of repeated columns, the repetitions being explicitly modeled by means of constraint matrices. The work [92] derived both essential uniqueness conditions and partial uniqueness conditions for PARALIND/CONFAC models of third-order tensors. Therein, the relation with uniqueness of constrained Tucker3 models and the block decomposition in rank-($L, L, 1$) terms is also discussed. The essential uniqueness condition for a given matrix factor in PARALIND models makes use of Kruskal's permutation lemma [52,86].

Consider a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ satisfying a PARALIND model with matrix factors (**A**, **B**, **C**) and constraint matrices $\mathbf{\Phi}^{(i)}$, $i = 1, 2, 3$. Suppose $(\mathbf{B} \otimes \mathbf{C})\mathbf{G}_{R_2 R_3 \times R_1}$ and **A** have full column rank and let $\omega(\cdot)$ denote the number of nonzero elements of its vector argument. Define $N_i = \text{rank}(\mathbf{\Phi}^{(2)}\text{diag}(\mathbf{\Phi}_{i,.}^{(1)})\mathbf{\Phi}^{(3)T})$, $i = 1, \ldots, R_1$. If for any vector **d**

$$\text{rank}\Big[\mathbf{B}\mathbf{\Phi}^{(2)}\text{diag}(\mathbf{d}^T\mathbf{\Phi}^{(1)})(\mathbf{C}\mathbf{\Phi}^{(3)})^T\Big] \leq \max(N_1, \ldots, N_{R_1})$$
$$\text{implies} \quad \omega(\mathbf{d}) \leq 1, \tag{81}$$

then **A** is essentially unique [92]. The uniqueness condition for **B** and **C** is analogous to condition (81) by interchanging the roles of $\mathbf{\Phi}^{(1)}$, $\mathbf{\Phi}^{(2)}$, and $\mathbf{\Phi}^{(3)}$.

When PARALIND model reduces to PARAFAC model, condition (81) is identical to Condition B of [86] for the

essential uniqueness of the PARAFAC model in the case of a full column rank matrix factor. More recently in [93], improved versions of the main uniqueness conditions of PARALIND/CONFAC models have been derived. The results presented therein involve simpler proofs than those of [92]. Moreover, the associated uniqueness conditions are easy to check in comparison with the ones presented earlier in [92].

In [94], a 'uni-mode' uniqueness condition is derived for a PARAFAC model with linearly dependent (proportional/identical) columns in one matrix factor. This condition is particularly useful for a subclass of PARALIND/CONFAC models with $\boldsymbol{\Phi}^{(2)} = \boldsymbol{\Phi}^{(3)} = \mathbf{I}_R$, i.e., when collinearity is confined within the first matrix factor. Let $\bar{\mathbf{A}} = \mathbf{A}\boldsymbol{\Phi}^{(1)}$, where $\bar{\mathbf{A}} \in \mathbb{C}^{I_1 \times R}$ contains collinear columns, the collinearity pattern being captured by $\boldsymbol{\Phi}^{(1)}$. Assuming that $\bar{\mathbf{A}}$ does not contain an all-zero column, if

$$r_{\bar{\mathbf{A}}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \tag{82}$$

then $\mathbf{A}$ is essentially unique. Generalizations of this condition can be obtained by imposing additional constraints on the ranks and k-ranks of the matrix factors (see [94] for details).

1.4.3 Uniqueness of Tucker models

Contrary to PARAFAC models, the Tucker ones are generally not essentially unique. Indeed, the parameters of Tucker models can be only estimated up to non-singular transformations characterized by non-singular matrices $\mathbf{T}^{(n)}$ that act on the mode-n matrix factors $\mathbf{A}^{(n)}$ and can be cancelled in replacing the core tensor by $\mathcal{G} \times_{n=1}^N [\mathbf{T}^{(n)}]^{-1}$. This result is easy to verify by applying the property (13) of mode-n product

$$\mathcal{G} \times_{n=1}^N \left[\mathbf{T}^{(n)}\right]^{-1} \times_{n=1}^N \mathbf{A}^{(n)}\mathbf{T}^{(n)} = \mathcal{G} \times_{n=1}^N \mathbf{A}^{(n)}\mathbf{T}^{(n)} \left[\mathbf{T}^{(n)}\right]^{-1}$$
$$= \mathcal{G} \times_{n=1}^N \mathbf{A}^{(n)}.$$

Uniqueness can be obtained by imposing some constraints on the core tensor or the matrix factors (see [9] for a review of main results concerning uniqueness of Tucker models, with discussion of three different approaches for simplifying core tensors so that uniqueness is ensured). Uniqueness can also result from a core with information redundancy and structure constraints as in [33] where the core is characterized by matrix slices in Hankel and Vandermonde forms.

1.4.4 Uniqueness of the PARATUCK-(2,4) model

Let us consider the PARATUCK-(2,4) model defined by (57), with matrix factors $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$, constraint matrices $\boldsymbol{\Phi}^{(1)}$ and $\boldsymbol{\Phi}^{(2)}$ and core tensor $\mathcal{C}$. As previously shown, this model is equivalent to the constrained PARAFAC model (58) whose matrix factors are

$$\mathbf{A} = \mathbf{A}^{(1)}\boldsymbol{\Psi}^{(1)}, \quad \mathbf{B} = \mathbf{A}^{(2)}\boldsymbol{\Psi}^{(2)}, \quad \mathbf{F} = (\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T,$$
$$\mathbf{D} = \mathbf{C}_{I_4 \times R_1 R_2}$$

with $\boldsymbol{\Psi}^{(1)}$ and $\boldsymbol{\Psi}^{(2)}$ defined in (60). Due to the repetition of some columns of $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ and assuming that these matrices do not contain an all-zero column, we have $k_{\mathbf{A}} = k_{\mathbf{B}} = 1$, and application of Kruskal's condition (76), with $N = 4$, gives

$$k_{\mathbf{A}}+k_{\mathbf{B}}+k_{\mathbf{F}}+k_{\mathbf{D}} \geq 2R_1R_2+3 \quad \Rightarrow \quad k_{\mathbf{F}}+k_{\mathbf{D}} \geq 2R_1R_2+1,$$

which can never be satisfied. However, more relaxed sufficient conditions can be established for essential uniqueness of the PARATUCK-(2,4) model. For that purpose, we consider the contracted constrained PARAFAC model obtained by combining the first two modes and using (67), which leads to a third-order PARAFAC model with matrix factors

$$(\mathbf{A} \diamond \mathbf{B}, \mathbf{F}, \mathbf{D}) = (\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}, (\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T, \mathbf{C}_{I_4 \times R_1 R_2}) \tag{83}$$

Note that uniqueness of the matrix factors of the contracted PARAFAC model (83) implies uniqueness of the matrix factors $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ of the original PARATUCK-(2,4) model. This comes from the fact that $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ can be recovered (up to a scaling factor) from their Kronecker product [95]. Application of the conditions (80) to the contracted PARAFAC model (83) allows deriving the following theorem.

Theorem: The PARATUCK-(2,4) model defined by (57) is essentially unique

- 1) When $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are full column rank
  $(r_{\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}} = R_1R_2 \Rightarrow k_{\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}} = R_1R_2)$
  If $\begin{cases} k_{(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T} \geq 2 \\ k_{\mathbf{C}_{I_4 \times R_1 R_2}} \geq 2 \end{cases}$ and
  $\begin{cases} r_{(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T} + k_{\mathbf{C}_{I_4 \times R_1 R_2}} \geq R_1R_2 + 2 \\ \text{or} \\ r_{\mathbf{C}_{I_4 \times R_1 R_2}} + k_{(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T} \geq R_1R_2 + 2 \end{cases}$
- 2) When $(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T$ is full column rank
  If $\begin{cases} k_{\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}} \geq 2 \\ k_{\mathbf{C}_{I_4 \times R_1 R_2}} \geq 2 \end{cases}$ and
  $\begin{cases} r_{\mathbf{A}^{(1)}} r_{\mathbf{A}^{(2)}} + k_{\mathbf{C}_{I_4 \times R_1 R_2}} \geq R_1R_2 + 2 \\ \text{or} \\ r_{\mathbf{C}_{I_4 \times R_1 R_2}} + k_{\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}} \geq R_1R_2 + 2 \end{cases}$
- 3) When $\mathbf{C}_{I_4 \times R_1 R_2}$ is full column rank
  If $\begin{cases} k_{\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}} \geq 2 \\ k_{(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T} \geq 2 \end{cases}$ and
  $\begin{cases} r_{\mathbf{A}^{(1)}} r_{\mathbf{A}^{(2)}} + k_{(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T} \geq R_1R_2 + 2 \\ \text{or} \\ r_{(\boldsymbol{\Phi}^{(1)} \diamond \boldsymbol{\Phi}^{(2)})^T} + k_{\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}} \geq R_1R_2 + 2 \end{cases}$

In [67], an application of the PARATUCK-(2,4) model to tensor space-time (TST) coding is considered. Therein, the matrix factors $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ represent the symbol and channel matrices to be estimated, while the constraint matrices $\mathbf{\Phi}^{(1)}$ and $\mathbf{\Phi}^{(2)}$ play the role of allocation matrices of the transmission system, and the tensor $\mathcal{C}$ is the coding tensor. In this context, $\mathbf{\Phi}^{(1)}$, $\mathbf{\Phi}^{(2)}$, and $\mathcal{C}$ can be properly designed to satisfy the sufficient conditions of item 1) of the theorem.

The sufficient conditions of this theorem can easily be extended to the case of PARATUCK-(N_1, N) models in replacing $\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}$, $\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)}$, $\mathbf{C}_{I_4 \times R_1 R_2}$, and $R_1 R_2$, by $\overset{N_1}{\underset{n=1}{\otimes}} \mathbf{A}^{(n)}$, $\overset{N_1}{\underset{n=1}{\diamond}} \mathbf{\Phi}^{(n)}$, $\mathbf{C}_{I_{N_1+2}...I_N \times R}$, and $R = \prod_{n=1}^{N_1} R_n$, respectively.

2 Conclusions

Several tensor models among which some are new have been presented in a general and unified framework. The use of the index notation for mode combination based on Kronecker products provides an original and concise way to derive vectorized and matricized forms of tensor models. A particular focus on constrained tensor models has been made with a perspective of designing MIMO communication systems with resource allocation. A link between PARATUCK models and constrained PARAFAC models has been established, which allows to apply results concerning PARAFAC models to derive uniqueness properties and parameter estimation algorithms for PARATUCK models. In a companion paper, several tensor-based MIMO systems are presented in a unified way based on constrained PARAFAC models, and a new tensor-based space-time-frequency (TSTF) MIMO transmission system with a blind receiver is proposed using a generalized PARATUCK model [96]. Even if this presentation of constrained tensor models has been made with the aim of designing MIMO transmission systems, we believe that such tensor models can be applied to other areas than telecommunications, like for instance biomedical signal processing, and more particularly for ECG and EEG signals modeling, with spatial constraints allowing to take into account the relative weight of the contributions of different areas of surface to electrodes. The considered constrained tensor models allow to take constraints into account either independently on each matrix factor of a PARAFAC decomposition, in the case of PARALIND/CONFAC models, or between factors, in the case of PARATUCK models. A perspective of this work is to consider constraints into tensor networks which decompose high-order tensors into lower-order tensors for big data processing [97]. In this case, the constraints could act either separately on each tensor component to facilitate their physical interpretability or between tensor components to explicit their interactions.

Appendices

Appendix 1

Some matrix formulae

For $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R_n}$, $\mathbf{B}^{(n)} \in \mathbb{C}^{R_n \times J_n}$, $\mathbf{\Phi}^{(n)} \in \mathbb{C}^{R_n \times R}$, and $\mathbf{\Psi}^{(n)} \in \mathbb{C}^{R_n \times Q}$, $n = 1, \ldots, N$

$$\left(\overset{N}{\underset{n=1}{\otimes}} \mathbf{A}^{(n)} \right)^T = \overset{N}{\underset{n=1}{\otimes}} \mathbf{A}^{(n)T} \in \mathbb{C}^{R_1 \cdots R_N \times I_1 \cdots I_N} \tag{84}$$

(Associative property)

$$\left(\overset{N}{\underset{n=1}{\otimes}} \mathbf{A}^{(n)} \right) \left(\overset{N}{\underset{n=1}{\otimes}} \mathbf{B}^{(n)} \right) = \overset{N}{\underset{n=1}{\otimes}} \mathbf{A}^{(n)} \mathbf{B}^{(n)} \in \mathbb{C}^{I_1 \cdots I_N \times J_1 \cdots J_N} \tag{85}$$

$$\left(\overset{N}{\underset{n=1}{\otimes}} \mathbf{A}^{(n)} \right) \left(\overset{N}{\underset{n=1}{\diamond}} \mathbf{\Phi}^{(n)} \right) = \overset{N}{\underset{n=1}{\diamond}} \mathbf{A}^{(n)} \mathbf{\Phi}^{(n)} \in \mathbb{C}^{I_1 \cdots I_N \times R} \tag{86}$$

$$\left(\overset{N}{\underset{n=1}{\diamond}} \mathbf{\Psi}^{(n)} \right)^T \left(\overset{N}{\underset{n=1}{\diamond}} \mathbf{\Phi}^{(n)} \right) = \overset{N}{\underset{n=1}{\odot}} \mathbf{\Psi}^{(n)T} \mathbf{\Phi}^{(n)} \in \mathbb{C}^{Q \times R}.$$

For $\mathbf{A}^{(n)} \in \mathbb{C}^{I \times J}$, $n = 1, \ldots, N$, and $\mathbf{B}^{(p)} \in \mathbb{C}^{K \times L}$, $p = 1, \ldots, P$

(Distributive property)

$$\left(\sum_{n=1}^{N} \mathbf{A}^{(n)} \right) \otimes \left(\sum_{p=1}^{P} \mathbf{B}^{(p)} \right) = \sum_{n=1}^{N} \sum_{p=1}^{P} \left(\mathbf{A}^{(n)} \otimes \mathbf{B}^{(p)} \right) \in \mathbb{C}^{IK \times JL} \tag{87}$$

In particular, for $\mathbf{A} \in \mathbb{C}^{I \times M}$, $\mathbf{B} \in \mathbb{C}^{J \times N}$, $\mathbf{C} \in \mathbb{C}^{M \times P}$, $\mathbf{D} \in \mathbb{C}^{N \times Q}$, $\mathbf{E} \in \mathbb{C}^{P \times J}$, $\mathbf{\Phi} \in \mathbb{C}^{M \times R}$, $\mathbf{\Psi} \in \mathbb{C}^{N \times R}$, $\mathbf{\Omega} \in \mathbb{C}^{M \times Q}$, $\mathbf{\Xi} \in \mathbb{C}^{N \times Q}$, and $\mathbf{x} \in \mathbb{C}^{M \times 1}$, we have

$$(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T, \tag{88}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}, \tag{89}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{\Phi} \diamond \mathbf{\Psi}) = \mathbf{A\Phi} \diamond \mathbf{B\Psi}, \tag{90}$$

$$(\mathbf{\Omega} \diamond \mathbf{\Xi})^T (\mathbf{\Phi} \diamond \mathbf{\Psi}) = \mathbf{\Omega}^T \mathbf{\Phi} \odot \mathbf{\Xi}^T \mathbf{\Psi}, \tag{91}$$

$$\text{vec}(\mathbf{ACE}) = (\mathbf{E}^T \otimes \mathbf{A})\text{vec}(\mathbf{C}), \tag{92}$$

$$\text{vec}\left(\mathbf{A}\text{diag}(\mathbf{x})\mathbf{C}\right) = (\mathbf{C}^T \diamond \mathbf{A})\mathbf{x}. \tag{93}$$

Appendix 2

Tensor extension of a matrix

Following the same demonstration as for (21) and (22), it is easy to deduce the following more general formula for the extension of $\mathbf{B} \in \mathbb{C}^{I \times R_n}$ into a tensor $\mathcal{A} \in \mathbb{C}^{I \times R_1 \times \cdots \times R_N}$ such as $a_{i,r_1,\ldots,r_n,\ldots,r_N} = b_{i,r_n} \, \forall r_k = 1, \ldots, R_k$, for $k = 1, \ldots, n-1, n+1, \ldots, N$. Defining $R = \prod_{n=1}^{N} R_n$, we have

$$\mathbf{A}_{I \times R} = \mathbf{B}\left(\mathbf{1}_{R_1}^T \otimes \cdots \otimes \mathbf{1}_{R_{n-1}}^T \otimes \mathbf{I}_{R_n} \otimes \mathbf{1}_{R_{n+1}}^T \otimes \cdots \otimes \mathbf{1}_{R_N}^T\right) \in \mathbb{C}^{I \times R}. \tag{94}$$

Similarly, for the extension of $\mathbf{B} \in \mathbb{C}^{I_n \times R}$ into a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times R}$ such as $a_{i_1,\ldots,i_n,\ldots,i_N,r} = b_{i_n,r} \; \forall i_k = 1,\ldots,I_k$, for $k = 1,\ldots,n-1,n+1,\ldots,N$, we have

$$\mathbf{A}_{I \times R} = (\mathbf{1}_{I_1} \otimes \cdots \otimes \mathbf{1}_{I_{n-1}} \otimes \mathbf{I}_{I_n} \otimes \mathbf{1}_{I_{n+1}} \otimes \cdots \otimes \mathbf{1}_{I_N}) \mathbf{B} \in \mathbb{C}^{I \times R}. \tag{95}$$

where $I = \prod_{n=1}^{N} I_n$.

For instance, if we consider the following tensor extension of $\mathbf{B} \in \mathbb{C}^{I \times J}$

$$a_{m,n,i,j,k,l} = b_{i,j} \; \forall m = 1,\ldots,M, \forall n = 1,\ldots,N, \\ \forall k = 1,\ldots,K, \forall l = 1,\ldots,L$$

the combination of formulae (94) and (95) gives

$$\mathbf{A}_{MNI \times JKL} = (\mathbf{1}_{MN} \otimes \mathbf{I}_I) \mathbf{B} \left(\mathbf{I}_J \otimes \mathbf{1}_{KL}^T \right) \tag{96}$$

which can be written as

$$\mathbf{A}_{MNI \times JKL} = \mathbf{B} \times_1 \mathbf{\Psi}_1 \times_2 (\mathbf{\Psi}_2)^T$$

with $\mathbf{\Psi}_1 = \mathbf{1}_{MN} \otimes \mathbf{I}_I$ and $\mathbf{\Psi}_2 = \mathbf{I}_J \otimes \mathbf{1}_{KL}^T$.

Appendix 3
Proof of (27)
Defining $(\mathbb{I}_1, \mathbb{I}_2)$ and $(\mathbb{R}_1, \mathbb{R}_2)$ as the sets of indices i_n and r_n associated respectively with the sets $(\mathbb{S}_1, \mathbb{S}_2)$ of index n, the formula (20) allows writing the element $g_{r_1,\ldots,r_N}$ of the core tensor as

$$g_{r_1,\ldots,r_N} = \mathbf{e}^{\mathbb{R}_1} \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \mathbf{e}_{\mathbb{R}_2}. \tag{97}$$

where $\mathbb{R}_1 = \{r_n, n \in \mathbb{S}_1\}$ and $\mathbb{R}_2 = \{r_n, n \in \mathbb{S}_2\}$.
Substituting $x_{i_1,\ldots,i_N}$ and $g_{r_1,\ldots,r_N}$ by their expressions (24) and (97) into (19) gives

$$\begin{aligned}
\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} &= x_{i_1,\ldots,i_N} \mathbf{e}_{\mathbb{I}_1}^{\mathbb{I}_2} = \mathbf{e}_{\mathbb{I}_1} x_{i_1,\ldots,i_N} \mathbf{e}^{\mathbb{I}_2} \\
&= \mathbf{e}_{\mathbb{I}_1} g_{r_1,\ldots,r_N} \left(\prod_{n=1}^{N} a_{i_n,r_n}^{(n)} \right) \mathbf{e}^{\mathbb{I}_2} \\
&= \left(\prod_{n=1}^{N} a_{i_n,r_n}^{(n)} \right) \mathbf{e}_{\mathbb{I}_1} \mathbf{e}^{\mathbb{R}_1} \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \mathbf{e}_{\mathbb{R}_2} \mathbf{e}^{\mathbb{I}_2} \\
&= \left(\prod_{n \in \mathbb{S}_1} a_{i_n,r_n}^{(n)} \right) \mathbf{e}_{\mathbb{I}_1}^{\mathbb{R}_1} \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \left(\prod_{n \in \mathbb{S}_2} a_{i_n,r_n}^{(n)} \right) \mathbf{e}_{\mathbb{R}_2}^{\mathbb{I}_2}
\end{aligned} \tag{98}$$

Applying the general Kronecker formula (17) in terms of the index notation allows to rewrite this matrix unfolding as

$$\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} = \left(\underset{n \in \mathbb{S}_1}{\otimes} \mathbf{A}^{(n)} \right) \mathbf{G}_{\mathbb{S}_1;\mathbb{S}_2} \left(\underset{n \in \mathbb{S}_2}{\otimes} \mathbf{A}^{(n)} \right)^T.$$

Appendix 4
Proof of (37)
Substituting the expression (34) of $x_{i_1,\ldots,i_N}$ into (19) and using the identities (14) and (15) give

$$\begin{aligned}
\mathbf{X}_{\mathbb{S}_1;\mathbb{S}_2} &= x_{i_1,\ldots,i_N} \mathbf{e}_{\mathbb{I}_1}^{\mathbb{I}_2} \\
&= \left(\prod_{n \in \mathbb{S}_1} a_{i_n,r}^{(n)} \right) \mathbf{e}_{\mathbb{I}_1} \left(\prod_{n \in \mathbb{S}_2} a_{i_n,r}^{(n)} \right) \mathbf{e}^{\mathbb{I}_2} \\
&= \left(\underset{n \in \mathbb{S}_1}{\diamond} \mathbf{A}_{\cdot r}^{(n)} \right) \left(\underset{n \in \mathbb{S}_2}{\diamond} \mathbf{A}_{\cdot r}^{(n)} \right)^T \\
&= \left(\underset{n \in \mathbb{S}_1}{\diamond} \mathbf{A}^{(n)} \right) \left(\underset{n \in \mathbb{S}_2}{\diamond} \mathbf{A}^{(n)} \right)^T
\end{aligned} \tag{99}$$

which ends the proof of (37).

Appendix 5
Proof of (59) and (60)
Let us define the third-order tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times R_1 \times R_2}$, $\mathcal{B} \in \mathbb{C}^{I_2 \times R_1 \times R_2}$, $\mathcal{F} \in \mathbb{C}^{I_3 \times R_1 \times R_2}$, and $\mathcal{D} \in \mathbb{C}^{I_4 \times R_1 \times R_2}$ such as

$$\begin{aligned}
a_{i_1,r_1,r_2} &= a_{i_1,r_1}^{(1)} \; \forall r_2 = 1,\ldots,R_2 \quad ; \\
b_{i_2,r_1,r_2} &= a_{i_2,r_2}^{(2)} \; \forall r_1 = 1,\ldots,R_1 \quad ; \\
f_{i_3,r_1,r_2} &= \phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)} \quad ; \quad d_{i_4,r_1,r_2} = c_{r_1,r_2,i_4}.
\end{aligned} \tag{100}$$

The tensor model (57) can be rewritten as

$$x_{i_1,i_2,i_3,i_4} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{i_1,r_1,r_2} b_{i_2,r_1,r_2} f_{i_3,r_1,r_2} d_{i_4,r_1,r_2}. \tag{101}$$

Defining the change of variables $r = (r_1 - 1)R_2 + r_2$ that corresponds to a combination of the last two modes of the tensors $\mathcal{A}, \mathcal{B}, \mathcal{F}$, and $\mathcal{D}$, (101) can be rewritten as the constrained PARAFAC-4 model (58), where $a_{i_1,r}, b_{i_2,r}, f_{i_3,r}$, and $d_{i_4,r}$ are entries of mode-1 matrix unfoldings of the tensors $\mathcal{A}, \mathcal{B}, \mathcal{F}$, and $\mathcal{D}$, i.e., entries of $\mathbf{A} \triangleq \mathbf{A}_{I_1 \times R_1 R_2}$, $\mathbf{B} \triangleq \mathbf{B}_{I_2 \times R_1 R_2}$, $\mathbf{F} \triangleq \mathbf{F}_{I_3 \times R_1 R_2}$, and $\mathbf{D} \triangleq \mathbf{D}_{I_4 \times R_1 R_2}$, respectively. Using the formulae (21) and (22), we can directly deduce the following expressions of $\mathbf{A}$ and $\mathbf{B}$:

$$\mathbf{A} = \mathbf{A}^{(1)} \otimes \mathbf{1}_{R_2}^T = \mathbf{A}^{(1)} \left(\mathbf{I}_{R_1} \otimes \mathbf{1}_{R_2}^T \right) = \mathbf{A}^{(1)} \mathbf{\Psi}^{(1)}. \tag{102}$$

$$\mathbf{B} = \mathbf{1}_{R_1}^T \otimes \mathbf{A}^{(2)} = \mathbf{A}^{(2)} \left(\mathbf{1}_{R_1}^T \otimes \mathbf{I}_{R_2} \right) = \mathbf{A}^{(2)} \mathbf{\Psi}^{(2)} \tag{103}$$

For the matrix $\mathbf{F}$, using the index notation with the definition (100) gives

$$\mathbf{F} = \left(f_{i_3,r_1,r_2} \mathbf{e}_{i_3}^{r_1 r_2} \right) = \left(\phi_{r_1,i_3}^{(1)} \phi_{r_2,i_3}^{(2)} \mathbf{e}_{i_3}^{r_1 r_2} \right)$$

Applying the formula (16), we directly obtain

$$\mathbf{F} = \left(\mathbf{\Phi}^{(1)} \diamond \mathbf{\Phi}^{(2)} \right)^T.$$

Author details
[1]I3S Laboratory, University of Nice-Sophia Antipolis, CNRS, 2000 route des
Lucioles, Les Algorithmes-B, 06903 Sophia Antipolis, France. [2]Department of
Teleinformatics Engineering, Federal University of Ceará, Campus do Pici, B.
725, 60455-970 Fortaleza, Brazil.

References
1. P McCullagh, *Tensor Methods in Statistics*. (Chapman and Hall, New York, 1987)
2. P Comon, ed. by JG McWhirter, IK Proudler, Tensor decompositions: state of the art and applications, in *Mathematics in Signal Processing V* (Clarendon Press, Oxford, 2002), pp. 1–24
3. LR Tucker, Some mathematical notes on three-mode factor analysis. Psychometrika **31**, 279–311 (1966)
4. RA Harshman, Foundations of the PARAFAC procedure: model and conditions for an "explanatory" multimodal factor analysis. UCLA Working Pap. Phon. **16**, 1–84 (1970)
5. JD Carroll, J Chang, Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika **35**(3), 283–319 (1970)
6. HAL Kiers, Towards a standardized notation and terminology in multiway analysis J. Chemometrics **14**(2), 105–122 (2000)
7. PM Kroonenberg, *Applied Multiway Data Analysis*. (Wiley, Hoboken, 2008)
8. R Bro, Multi-way analysis in the food industry: models, algorithms and applications. Ph.D. dissertation, University of Amsterdam, Amsterdam (1998)
9. A Smilde, R Bro, P Geladi, *Multi-way Analysis: Applications in the Chemical Sciences*. (Wiley, Chichester, 2004)
10. J-F Cardoso, Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem, in *Proc. of IEEE ICASSP'90* (Albuquerque, 1990), pp. 2655–2658
11. J-F Cardoso, P Comon, Tensor-based independent component analysis, in *Proc. of EUSIPCO'90* (Barcelona, 1990), pp. 673–676
12. J-F Cardoso, Super-symmetric decomposition of the fourth-order cumulant tensor: blind identification of more sources than sensors, in *Proc. of IEEE ICASSP'91* (Toronto, 1991), pp. 3109–3112
13. L De Lathauwer, Signal processing based on multilinear algebra. Ph.D. dissertation, KU Leuven, Leuven (1997)
14. P Comon, C Jutten, *Handbook of Blind Source Separation. Independent Component Analysis and Applications*. (Elsevier, Oxford, 2010)
15. ND Sidiropoulos, GB Giannakis, R Bro, Blind PARAFAC receivers for DS-CDMA systems. IEEE Trans. Signal Process. **48**(3), 810–823 (2000)
16. A Cichocki, R Zdunek, AH Phan, S-I Amari, *Nonnegative Matrix and Tensor Factorizations. Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. (Wiley, Chichester, 2009)
17. TG Kolda, BW Bader, Tensor decompositions and applications. SIAM J. Matrix Anal. Appl. **51**(3), 455–500 (2009)
18. E Acar, B Yener, Unsupervised multiway data analysis: a literature survey. IEEE Trans. Knowledge Data Eng. **21**(1), 6–20 (2009)
19. M Morup, Applications of tensor (multiway array) factorizations and decompositions in data mining, in *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, issue 1 (Wiley, Chichester, 2011), pp. 24–40
20. M Sorensen, L De Lathauwer, P Comon, S Icart, L Deneire, Canonical polyadic decomposition with a columnwise orthonormal factor matrix. SIAM J. Matrix Anal. Appl. **33**(4), 1190–1213 (2012)
21. A Shashua, T Hazan, Non-negative tensor factorization with applications to statistics and computer vision, in *Proc. of 22nd Int. Conf. on Machine Learning* (Bonn, 2005), pp. 792–799
22. S Hazan, S Polak, A Shashua, Sparse image coding using a 3D non-negative tensor factorization, in *Proc. of 10th IEEE Int. Conf. on Computer Vision (ICCV' 2005)* (Beijing, 2005), pp. 50–57
23. MP Friedlander, K Hatz, Computing nonnegative tensor factorizations. Optim. Meth. Software. **23**(4), 631–647 (2008)
24. Q Zhang, H Wang, R Plemmons, P Pauca, Tensor methods for hyperspectral data processing: a space object identification study. J. Opt. Soc. Am. A **25**(12), 3001–3012 (2008)
25. E Benetos, C Kotropoulos, Non-negative tensor factorization applied to music genre classification. IEEE Trans. Audio, Speech, Language Proc. **18**(8), 1955–1967 (2010)
26. A Ozerov, C Févote, R Blouet, G Durrieu, Multichannel nonnegative tensor factorization with structured constraints for user-guided audio source separation, in *Proc. of ICASSP' 2011* (Prague, 2011)
27. E Acar, DM Dunlavy, TG Kolda, M Morup, Scalable tensor factorizations with missing data, in *Proc. of 10th SIAM Int. Conf. on Data Mining* (Columbus, 2010), pp. 701–712
28. J-P Royer, N Thirion-Moreau, P Comon, Computing the polyadic decomposition of nonnegative third order tensors. Signal Process. **91**, 2159–2171 (2011)
29. A-H Phan, A Cichocki, Extended HALS algorithm for nonnegative Tucker decomposition and its applications for multiway analysis and classification. Neurocomputing **74**, 1956–1969 (2011)
30. M Welling, M Weber, Positive tensor factorization. Pattern Recogn. Lett. **22**(12), 1255–1261 (2001)
31. M Morup, LK Hansen, Algorithms for sparse non-negative Tucker decompositions. Neural Comput. **20**, 2112–2131 (2008)
32. G Favier, T Bouilloc, A constrained tensor based approach for MIMO NL-CDMA systems, in *Proc. of EUSIPCO' 2010* (Aalborg, 2010)
33. G Favier, T Bouilloc, ALF de Almeida, Blind constrained block-Tucker2 receiver for multiuser SIMO NL-CDMA communication systems. Signal Process. **92**(7), 1624–1636 (2012)
34. G Favier, AY Kibangou, T Bouilloc, Nonlinear system modeling and identification using Volterra-PARAFAC models. Int. J. of Adaptive Control and Sig. Proc. **26**, 30–53 (2012)
35. T Bouilloc, G Favier, Nonlinear channel modeling and identification using bandpass Volterra-PARAFAC models. Signal Process. **92**(6), 1492–1498 (2012)
36. AY Kibangou, G Favier, Identification of parallel-cascade Wiener systems using joint diagonalization of third-order Volterra kernel slices. IEEE Signal Proc. Lett. **16**(3), 188–191 (2009)
37. G Favier, Nonlinear system modeling and identification using tensor approaches, in *Proc. of 10th Int. Conf. on Sciences and Techniques of Automatic Control and Computer Engineering (STA' 2009)* (Hammamet, 2009)
38. CER Fernandes, G Favier, JCM Mota, Blind channel identification algorithms based on the Parafac decomposition of cumulant tensors: the single and multiuser cases. Signal Process. **88**, 1382–1401 (2008)
39. CER Fernandes, G Favier, JCM Mota, Parafac-based blind identification of convolutive MIMO linear systems, in *Proc. of 15th IFAC Symp. on System Identification (SYSID' 2009)* (Saint-Malo, 2009)
40. J Brachat, P Comon, B Mourrain, E Tsigaridas, Symmetric tensor decomposition. Lin. Algebra Appl. **433**(11–12), 1851–1872 (2010)
41. D Nion, L De Lathauwer, A block component model-based blind DS-CDMA receiver. IEEE Trans. Signal Proc. **56**(11), 5567–5579 (2008)
42. AY Kibangou, G Favier, Non-iterative solution for PARAFAC with a Toeplitz matrix factor, in *Proc. of EUSIPCO' 2009* (Glasgow, 2009), pp. 691–695
43. M Sorensen, L De Lathauwer, Blind signal separation via tensor decomposition with Vandermonde factor: canonical polyadic decomposition. IEEE Trans. Signal Process. **61**(22), 5507–5519 (2013)
44. JH Goulart, G Favier, An algebraic solution for the Candecomp/PARAFAC decomposition with circulant factors. SIAM J. Matrix Analysis and Appl (2014). http://hal.archives-ouvertes.fr/docs/00/96/72/63/PDF/RR-2014-02_I3S.pdf
45. P Comon, M Sorensen, E Tsigaridas, Decomposing tensors with structured matrix factors reduces to rank-1 approximations, in *Proc. of IEEE ICASSP' 2010* (Dallas, 2010), pp. 14–19
46. M Sorensen, P Comon, Tensor decompositions with banded matrix factors. Lin. Algebra Appl. **438**, 919–941 (2013)

47. JD Carroll, S Pruzansky, JB Kruskal, Candelinc: a general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. Psychometrika **45**(1), 3–24 (1980)

48. DSG Pollock, On Kronecker products, tensor products and matrix differential calculus, Working paper 11/34. Univ.of Leicester, Dept. of Economics, UK (2011). http://www.le.ac.uk/ec/research/RePEc/lec/leecon/dp11-34.pdf

49. L De Lathauwer, B De Moor, J Vandewalle, A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. **21**(4), 1253–1278 (2000)

50. RB Cattell, "Parallel proportional profiles" and other principles for determining the choice of factors by rotation. Psychometrika **9**, 267–283 (1944)

51. FL Hitchcock, The expression of a tensor or a polyadic as a sum of products. J. Math. Phys. **6**(3), 164–189 (1927)

52. JB Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Lin. Algebra Appl. **18**(2), 95–138 (1977)

53. P Comon, JMF ten Berge, L De Lathauwer, J Castaing, Generic and typical ranks of multi-way arrays. Lin. Algebra Appl. **430**(11), 2997–3007 (2009)

54. P Comon, G Golub, L-H Lim, B Mourrain, Symmetric tensors and symmetric tensor rank. SIAM J. Matrix Anal. Appl. **30**(3), 1254–1279 (2008)

55. R Bro, HAL Kiers, A new efficient method for determining the number of components in PARAFAC models. J. Chemometrics **17**(5), 274–286 (2003)

56. JPCL da Costa, M Haardt, F Roemer, Robust methods based on HOSVD for estimating the model order in PARAFAC models, in *Proc. of 5th IEEE Sensor Array and Multich. Signal Proc. Workshop (SAM 2008)* (Darmstadt, 2008), pp. 510–514

57. JPCL da Costa, F Roemer, M Weis, M Haardt, Robust *R*-D parameter estimation via closed-form PARAFAC, in *Proc. of ITG Workshop on Smart Antennas (WSA 2010)* (Bremen, 2010), pp. 99–106

58. JPCL da Costa, F Roemer, M Haardt, RT de Sousa, Multi-dimensional model order selection. EURASIP J. Adv. Signal Process. **26** (2011)

59. RA Harshman, ME Lundy, Uniqueness proof for a family of models sharing features of Tucker's three-mode factor analysis and PARAFAC/CANDECOMP. Psychometrika **61**, 133–154 (1996)

60. R Bro, RA Harshman, ND Sidiropoulos, Modeling multi-way data with linearly dependent loadings, KVL tech. report 176 (2005)

61. R Bro, RA Harshman, ND Sidiropoulos, ME Lundy, Modeling multi-way data with linearly dependent loadings. Chemometrics **23**(7–8), 324–340 (2009)

62. AY Kibangou, G Favier, Blind joint identification and equalization of Wiener-Hammerstein communication channels using PARATUCK-2 tensor decomposition, in *Proc. EUSIPCO' 2007* (Poznan, Sept 2007)

63. L Xu, J Ting, Y Longxiang, Z Hongbo, PARALIND-based identifiability results for parameter estimation via uniform linear array. EURASIP J. Adv. Sig. Proc. **2012**, 154 (2012)

64. L Xu, G Liang, Y Longxiang, Z Hongbo, PARALIND-based blind joint angle and delay estimation for multipath signals with uniform linear array. EURASIP J. Adv, Sig. Proc. **2012**, 130 (2012)

65. ALF de Almeida, G Favier, JCM Mota, A constrained factor decomposition with application to MIMO antenna systems. IEEE Trans. Signal Process. **56**(6), 2429–2442 (2008)

66. G Favier, MN da Costa, ALF de Almeida, JMT Romano, Tensor coding for CDMA-MIMO wireless communication systems, in *Proc. of EUSIPCO' 2011* (Barcelona, 29 Aug–2 Sept 2011)

67. G Favier, MN da Costa, ALF de Almeida, JMT Romano, Tensor space-time (TST) coding for MIMO wireless communication systems. Signal Process. **92**(4), 1079–1092 (2012)

68. ALF de Almeida, G Favier, JCM Mota, Space-time spreading-multiplexing for MIMO wireless communication systems using the PARATUCK-2 tensor model. Signal Process. **89**(11), 2103–2116 (2009)

69. ALF de Almeida, X Luciani, A Stegeman, P Comon, CONFAC decomposition approach to blind identification of underdetermined mixtures based on generating function derivatives. IEEE Trans. Signal Process. **60**(11), 5698–5713 (2012)

70. ALF de Almeida, G Favier, JCM Mota, Generalized PARAFAC model for multidimensional wireless communications with application to blind multiuser equalization, in *Asilomar Conf. Sig. Syst. Comp.* (Pacific Grove, Nov 2005)

71. ALF de Almeida, G Favier, JCM Mota, PARAFAC models for wireless communication systems, in *Int. Conf. on Physics in Signal and Image Processing (PSIP)* (Toulouse, 31 Jan–2 Feb 2005)

72. ALF de Almeida, G Favier, JCM Mota, PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization. Signal Process. **87**, 337–351 (2007)

73. ALF de Almeida, G Favier, JCM Mota, Tensor-based space-time multiplexing codes for MIMO-OFDM systems with blind detection, in *Proc. of 17th IEEE Symp. Pers. Ind. Mob. Radio Com. (PIMRC' 2006)* (Helsinki, Sept 2006)

74. ALF de Almeida, G Favier, JCM Mota, Constrained Tucker-3 model for blind beamforming. Signal Process. **89**, 1240–1244 (2009)

75. L De Lathauwer, Decompositions of a higher-order tensor in block terms-part II: definitions and uniqueness. SIAM J. Matrix Anal. Appl. **30**(3), 1033–1066 (2008)

76. A Cichocki, D Mandic, A-H Phan, C Caiafa, G Zhou, Q Zhao, L De Lathauwer, Tensor decompositions for signal processing applications. From two-way to multiway component analysis. IEEE Signal Process. Mag. (2014). arXiv:1403.4462v1

77. ALF de Almeida, G Favier, JCM Mota, Constrained tensor modeling approach to blind multiple-antenna CDMA schemes. IEEE Trans. Signal Process. **56**(6), 2417–2428 (2008)

78. J Salmi, A Richter, V Koivunen, Sequential unfolding SVD for tensors with applications in array signal processing. IEEE Trans. Signal Process. **57**(12), 4719–4733 (2009)

79. 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)

80. I Domanov, 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). arXiv:1301.4602v1

81. I Domanov, L De Lathauwer, Generic uniqueness conditions for the canonical polyadic decomposition and INDSCAL. arXiv:1405.6238v1, (KU Leuven, Belgium, 2014)

82. ND Sidiropoulos, R Bro, On the uniqueness of multilinear decomposition of N-way arrays. J. Chemometrics **14**, 229–239 (2000)

83. JMF ten Berge, ND Sidiropoulos, On uniqueness in CANDECOMP/PARAFAC. Psychometrika **67**(3), 399–409 (2002)

84. RA Harshman, Determination and proof of minimum uniqueness conditions for PARAFAC1. UCLA Working Pap. Phon. **22**, 111–117 (1972)

85. A Stegeman, ND Sidiropoulos, On Kruskal's uniqueness condition for the CANDECOMP/PARAFAC decomposition. Lin. Algebra Appl. **420**, 540–552 (2007)

86. T Jiang, ND 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 (2004)

87. L De Lathauwer, A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. SIAM J. Matrix Anal. Appl. **28**(3), 642–666 (2006)

88. A Stegeman, On uniqueness conditions for CANDECOMP/PARAFAC and INDSCAL with full column rank in one mode. Lin. Algebra Appl. **431**(1–2), 211–227 (2008)

89. X Guo, S Miron, D Brie, S Zhu, X Liao, A CANDECOMP/PARAFAC perspective on uniqueness of DOA estimation using a vector sensor array. IEEE Trans. Signal Process. **59**(7), 3475–3481 (2011)

90. JMF Ten Berge, Partial uniqueness in CANDECOMP/PARAFAC. J. Chemometrics **18**, 12–16 (2004)

91. D Brie, S Miron, F Caland, C Mustin, An uniqueness condition for the 4-way CANDECOMP/PARAFAC model with collinear loadings in three modes, in *Proc. of IEEE ICASSP' 2011* (Prague, May 2011)

92. A Stegeman, ALF de Almeida, Uniqueness conditions for constrained three-way factor decompositions with linearly dependent loadings. SIAM J. Matrix Anal. Appl. **31**(3), 1469–1490 (2009)

93. A Stegeman, TTT Lam, Improved uniqueness conditions for canonical tensor decompositions with linearly dependent loadings. SIAM J. Matrix Anal. Appl. **33**(4), 1250–1271 (2012)

94. X Guo, S Miron, D Brie, 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**(1), 111–129 (2012)

95. CF Van Loan, N Pitsianis, ed. by MS Moonen, GH Golub, and BLR de Moor, in *Linear Algebra for Large Scale and Real-Time Applications* (Kluwer, The Netherlands, 1993), p. 293

96. G Favier, ALF de Almeida, Tensor space-time-frequency coding with semi-blind receivers for MIMO wireless communication systems. IEEE Trans. Signal Process (2014). in press

97. A Cichocki, Era of big data processing: a new approach via tensor networks and tensor decompositions, in *International Workshop on Smart Info-Media Systems in Asia (SISA-2013)* (Nagoya, 30 Sept–2 Oct 2013)