

A MULTILINEAR SINGULAR VALUE DECOMPOSITION*

LIEVEN DE LATHAUWER[†], BART DE MOOR[†], AND JOOS VANDEWALLE[†]

Abstract. We discuss a multilinear generalization of the singular value decomposition. There is a strong analogy between several properties of the matrix and the higher-order tensor decomposition; uniqueness, link with the matrix eigenvalue decomposition, first-order perturbation effects, etc., are analyzed. We investigate how tensor symmetries affect the decomposition and propose a multilinear generalization of the symmetric eigenvalue decomposition for pair-wise symmetric tensors.

Key words. multilinear algebra, singular value decomposition, higher-order tensor

AMS subject classifications. 15A18, 15A69

PII. S0895479896305696

1. Introduction. An increasing number of signal processing problems involve the manipulation of quantities of which the elements are addressed by more than two indices. In the literature these higher-order equivalents of vectors (first order) and matrices (second order) are called higher-order tensors, multidimensional matrices, or multiway arrays. For a lot of applications involving higher-order tensors, the existing framework of vector and matrix algebra appears to be insufficient and/or inappropriate. In this paper we present a proper generalization of the singular value decomposition (SVD).

To a large extent, higher-order tensors are currently gaining importance due to the developments in the field of higher-order statistics (HOS): for multivariate stochastic variables the basic HOS (higher-order moments, cumulants, spectra and cepstra) are highly symmetric higher-order tensors, in the same way as, e.g., the covariance of a stochastic vector is a symmetric (Hermitean) matrix. A brief enumeration of some opportunities offered by HOS gives an idea of the promising role of higher-order tensors on the signal processing scene. It is clear that statistical descriptions of random processes are more accurate when, in addition to first- and second-order statistics, they take HOS into account; from a mathematical point of view this is reflected by the fact that moments and cumulants correspond, within multiplication with a fixed scalar, to the subsequent coefficients of a Taylor series expansion of the first, resp., second, characteristic function of the stochastic variable. In statistical nonlinear system theory HOS are even unavoidable (e.g., the autocorrelation of x^2 is a fourth-order moment).

*Received by the editors June 24, 1996; accepted for publication (in revised form) by B. Kagstrom January 4, 1999; published electronically April 18, 2000. This research was partially supported by the Flemish Government (the Concerted Research Action GOA-MIPS, the FWO (Fund for Scientific Research - Flanders) projects G.0292.95, G.0262.97, and G.0240.99, the FWO Research Communities ICCoS (Identification and Control of Complex Systems) and Advanced Numerical Methods for Mathematical Modelling, the IWT Action Programme on Information Technology (ITA/GB0/T23)-IT-Generic Basic Research), the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs (the Interuniversity Poles of Attraction Programmes IUAP P4-02 and IUAP P4-24) and the European Commission (Human Capital and Mobility Network SIMONET, SC1-CT92-0779). The first author is a post-doctoral Research Assistant with the F.W.O. The second author is a senior Research Associate with the F.W.O. and an Associate Professor at the K.U. Leuven. The third author is a Full Professor at the K.U. Leuven. The scientific responsibility is assumed by the authors.

<http://www.siam.org/journals/simax/21-4/30569.html>

[†]K.U. Leuven, E.E. Dept., ESAT-SISTA/COSIC, Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium (Lieven.DeLathauwer@esat.kuleuven.ac.be, Bart.DeMoor@esat.kuleuven.ac.be, Joos.Vandewalle@esat.kuleuven.ac.be).

Moreover, higher-order cumulants and spectra of a random variable are insensitive to an additive Gaussian perturbation of that variable, which is exploited in higher-order-only techniques, conceptually blind for Gaussian noise. HOS make it possible to solve the source separation (SS) problem by mere exploitation of the statistical independence of the sources, without knowledge of the array manifold; they also contain sufficient information to allow a blind identification of linear filters, without making assumptions on the minimum/nonminimum phase character, etc. (for general aspects of HOS, the interested reader is referred to the tutorial papers [34, 37, 38, 30] and the bibliography [41]; for the use of tensor decompositions as a basic tool in HOS-based signal processing, we refer to [11, 12, 9, 19]).

Higher-order tensors do not merely play an important role in HOS. As a matter of fact they seem to be used in the most various disciplines, like chemometrics, psychometrics, econometrics, image processing, biomedical signal processing, etc. Most often they appear in factor analysis-type problems of multiway datasets [13]. Another, more trivial, use is as a formalism to describe linear vector-matrix, matrix-vector, matrix-matrix, ... transformations, in the same way as matrices describe linear transformations between vector spaces. Interesting also is the fact that higher-order terms in the Taylor series expansion of a multivariate function and higher-order Volterra filter kernels have a tensor form.

On the other hand, one of the most fruitful developments in the world of linear algebra and linear algebra-based signal processing is the concept of the SVD of matrices [21, 35, 44]. In this paper we will discuss a multilinear generalization that has also been investigated in psychometrics as the Tucker model, originally developed to obtain a “method for searching for relations in a body of data,” for the case “each person in a group of individuals is measured on each trait in a group of traits by each of a number of methods,” or “when individuals are measured by a battery of measures on a number of occasions” [42, 43]. For three-way data, the Tucker model consists of decomposing a real $(I_1 \times I_2 \times I_3)$ -tensor \mathcal{A} according to

$$(1) \quad a_{i_1 i_2 i_3} = \sum_{j_1}^{I_1} \sum_{j_2}^{I_2} \sum_{j_3}^{I_3} s_{j_1 j_2 j_3} u_{i_1 j_1}^{(1)} u_{i_2 j_2}^{(2)} u_{i_3 j_3}^{(3)},$$

in which $u_{i_1 j_1}^{(1)}, u_{i_2 j_2}^{(2)}, u_{i_3 j_3}^{(3)}$ are entries of orthogonal matrices, and \mathcal{S} is a real $(I_1 \times I_2 \times I_3)$ -tensor with the property of “all-orthogonality,” i.e., $\sum_{i_1 i_2} s_{i_1 i_2 \alpha} s_{i_1 i_2 \beta} = \sum_{i_1 i_3} s_{i_1 \alpha i_3} s_{i_1 \beta i_3} = \sum_{i_2 i_3} s_{\alpha i_2 i_3} s_{\beta i_2 i_3} = 0$ whenever $\alpha \neq \beta$. This decomposition is virtually unknown in the communities of numerical algebra and signal processing; on the other hand, the viewpoint and language in psychometrics is somewhat different from what is common in our field. It is the aim of this paper to derive the tensor decomposition in an SVD terminology, using a notation that is a natural extension of the notation used for matrices. As already mentioned, we will show that the Tucker model is actually a convincing multilinear generalization of the SVD concept itself. From our algebraic point of view, we will ask a number of inevitable questions such as what the geometric link is between the generalized singular vectors/values and the column, row, ... vectors (oriented energy), how the concept of rank lies to the structure of the decomposition, whether the best reduced-rank approximation property carries over, and so on. Our derivations are valid for tensors of arbitrary order and hold for the complex-valued case too. In view of the many striking analogies between the matrix SVD and its multilinear generalization, we use the term *higher-order singular value decomposition* (HOSVD) in this paper; note at this point that the existence of

different multilinear SVD extensions may not be excluded—as a matter of fact, focusing on different properties of the matrix SVD does lead to the definition of different (formally less striking) multilinear generalizations, as we will explain later on.

In our own research the HOSVD has already proved its value. In [14] we showed that the decomposition is fundamentally related to the problem of blind source separation, also known as independent component analysis (ICA). In [18] we used the decomposition to compute an initial value for a tensorial equivalent of the power method, aiming at the computation of the best rank-1 approximation of a given tensor; a high-performant ICA-algorithm was based on this technique. In [19] the HOSVD was used in a dimensionality reduction for higher-order factor analysis-type problems, reducing the computational complexity. The current paper however is the first systematic, elaborated presentation of the HOSVD concept.

The paper is organized as follows. In section 2 we introduce some preliminary material on multilinear algebra, needed for the further developments. The HOSVD model is presented in section 3 and compared to its matrix counterpart. In section 4 we discuss some well-known SVD properties and demonstrate that they have striking higher-order counterparts. In section 5 we investigate how the HOSVD is influenced by symmetry properties of the higher-order tensor; in analogy with the eigenvalue decomposition (EVD) of symmetric (Hermitean) matrices we define a *higher-order eigenvalue decomposition* (HOEVD) for “pair-wise symmetric” higher-order tensors. Section 6 contains a first-order perturbation analysis and section 7 quotes some alternative ways to generalize the SVD.

Before starting with the next section, we would like to add a comment on the notation that is used. To facilitate the distinction between scalars, vectors, matrices, and higher-order tensors, the type of a given quantity will be reflected by its representation: scalars are denoted by lower-case letters ($a, b, \dots; \alpha, \beta, \dots$), vectors are written as capitals (A, B, \dots) (italic shaped), matrices correspond to bold-face capitals ($\mathbf{A}, \mathbf{B}, \dots$), and tensors are written as calligraphic letters ($\mathcal{A}, \mathcal{B}, \dots$). This notation is consistently used for lower-order parts of a given structure. For example, the entry with row index i and column index j in a matrix \mathbf{A} , i.e., $(\mathbf{A})_{ij}$, is symbolized by a_{ij} (also $(A)_i = a_i$ and $(\mathcal{A})_{i_1 i_2 \dots i_N} = a_{i_1 i_2 \dots i_N}$); furthermore, the i th column vector of a matrix \mathbf{A} is denoted as A_i , i.e., $\mathbf{A} = (A_1 A_2 \dots)$. To enhance the overall readability, we have made one exception to this rule: as we frequently use the characters i, j, r , and n in the meaning of indices (counters), I, J, R , and N will be reserved to denote the index upper bounds, unless stated otherwise.

2. Basic definitions.

2.1. Matrix representation of a higher-order tensor. The starting point of our derivation of a multilinear SVD will be to consider an appropriate generalization of the link between the column (row) vectors and the left (right) singular vectors of a matrix. To be able to formalize this idea, we define “matrix unfoldings” of a given tensor, i.e., matrix representations of that tensor in which all the column (row, ...) vectors are stacked one after the other. To avoid confusion, we will stick to one particular ordering of the column (row, ...) vectors; for order three, these unfolding procedures are visualized in Figure 1. Notice that the definitions of the matrix unfoldings involve the tensor dimensions I_1, I_2, I_3 in a cyclic way and that, when dealing with an unfolding of dimensionality $I_c \times I_a I_b$, we formally assume that the index i_a varies more slowly than i_b . In general, we have the following definition.

DEFINITION 1. Assume an N th-order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$. The matrix unfolding $\mathbf{A}_{(n)} \in \mathbb{C}^{I_n \times (I_{n+1} I_{n+2} \dots I_N I_1 I_2 \dots I_{n-1})}$ contains the element $a_{i_1 i_2 \dots i_N}$ at the

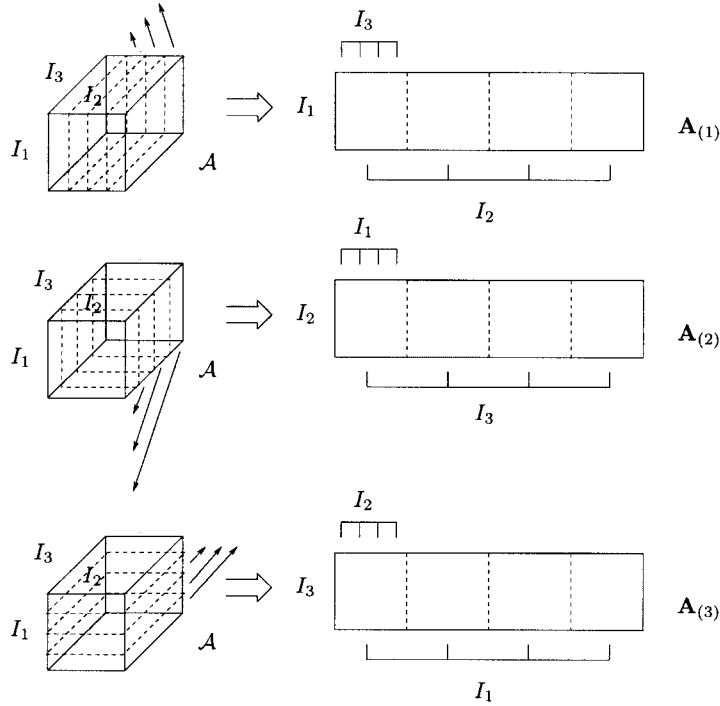


FIG. 1. Unfolding of the $(I_1 \times I_2 \times I_3)$ -tensor \mathcal{A} to the $(I_1 \times I_2 I_3)$ -matrix $\mathbf{A}_{(1)}$, the $(I_2 \times I_3 I_1)$ -matrix $\mathbf{A}_{(2)}$ and the $(I_3 \times I_1 I_2)$ -matrix $\mathbf{A}_{(3)}$ ($I_1 = I_2 = I_3 = 4$).

position with row number i_n and column number equal to

$$(i_{n+1} - 1)I_{n+2}I_{n+3} \dots I_N I_1 I_2 \dots I_{n-1} + (i_{n+2} - 1)I_{n+3}I_{n+4} \dots I_N I_1 I_2 \dots I_{n-1} + \dots \\ + (i_N - 1)I_1 I_2 \dots I_{n-1} + (i_1 - 1)I_2 I_3 \dots I_{n-1} + (i_2 - 1)I_3 I_4 \dots I_{n-1} + \dots + i_{n-1}.$$

Example 1. Define a tensor $\mathcal{A} \in \mathbb{R}^{3 \times 2 \times 3}$ by $a_{111} = a_{112} = a_{211} = -a_{212} = 1$, $a_{213} = a_{311} = a_{313} = a_{121} = a_{122} = a_{221} = -a_{222} = 2$, $a_{223} = a_{321} = a_{323} = 4$, $a_{113} = a_{312} = a_{123} = a_{322} = 0$. The matrix unfolding $\mathbf{A}_{(1)}$ is given by

$$\mathbf{A}_{(1)} = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & 2 & 0 \\ 1 & -1 & 2 & 2 & -2 & 4 \\ 2 & 0 & 2 & 4 & 0 & 4 \end{array} \right).$$

2.2. Rank properties of a higher-order tensor. There are major differences between matrices and higher-order tensors when rank properties are concerned. As we will explain in section 3, these differences directly affect the way an SVD generalization could look. As a matter of fact, there is not a unique way to generalize the rank concept.

First, it is easy to generalize the notion of column and row rank. If we refer in general to the column, row, \dots vectors of an N th-order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ as its “ n -mode vectors,” defined as the I_n -dimensional vectors obtained from \mathcal{A} by varying the index i_n and keeping the other indices fixed, then we have the following definition.

DEFINITION 2. The n -rank of \mathcal{A} , denoted by $R_n = \text{rank}_n(\mathcal{A})$, is the dimension of the vector space spanned by the n -mode vectors.

The n -rank of a given tensor can be analyzed by means of matrix techniques.

PROPERTY 1. The n -mode vectors of \mathcal{A} are the column vectors of the matrix unfolding $\mathbf{A}_{(n)}$ and

$$\text{rank}_n(\mathcal{A}) = \text{rank}(\mathbf{A}_{(n)}).$$

A major difference with the matrix case, however, is the fact that the different n -ranks of a higher-order tensor are not necessarily the same, as can easily be verified by checking some examples (see further).

The rank of a higher-order tensor is usually defined in analogy with the fact that a rank- R matrix can be decomposed as a sum of R rank-1 terms [12, 29].

DEFINITION 3. An N th-order tensor \mathcal{A} has rank 1 when it equals the outer product of N vectors $U^{(1)}, U^{(2)}, \dots, U^{(N)}$, i.e.,

$$a_{i_1 i_2 \dots i_N} = u_{i_1}^{(1)} u_{i_2}^{(2)} \dots u_{i_N}^{(N)},$$

for all values of the indices.

DEFINITION 4. The rank of an arbitrary N th-order tensor \mathcal{A} , denoted by $R = \text{rank}(\mathcal{A})$, is the minimal number of rank-1 tensors that yield \mathcal{A} in a linear combination.

(With respect to the definition of a rank-1 tensor, a remark on the notation has to be made. For matrices, the vector-vector outer product of $U^{(1)}$ and $U^{(2)}$ is denoted as $U^{(1)} \cdot U^{(2)^T}$; to avoid an ad hoc notation based on “generalized transposes” (in which the fact that column vectors are transpose-free would not have an inherent meaning), we will further denote the outer product of $U^{(1)}, U^{(2)}, \dots, U^{(N)}$ by $U^{(1)} \circ U^{(2)} \circ \dots \circ U^{(N)}$.)

A second difference between matrices and higher-order tensors is the fact that the rank is not necessarily equal to an n -rank, even when all the n -ranks are the same. From the definitions it is clear that always $R_n \leq R$.

Example 2. Consider the $(2 \times 2 \times 2)$ -tensor \mathcal{A} defined by

$$\begin{cases} a_{111} = a_{221} = a_{112} = 1 \\ a_{211} = a_{121} = a_{212} = a_{122} = a_{222} = 0. \end{cases}$$

It follows that $R_1 = R_2 = 2$ but $R_3 = 1$.

Example 3. Consider the $(2 \times 2 \times 2)$ -tensor \mathcal{A} defined by

$$\begin{cases} a_{211} = a_{121} = a_{112} = 1 \\ a_{111} = a_{222} = a_{122} = a_{212} = a_{221} = 0. \end{cases}$$

The 1-rank, 2-rank, and 3-rank are equal to 2. The rank, however, equals 3, since

$$\mathcal{A} = X_2 \circ Y_1 \circ Z_1 + X_1 \circ Y_2 \circ Z_1 + X_1 \circ Y_1 \circ Z_2,$$

in which

$$X_1 = Y_1 = Z_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad X_2 = Y_2 = Z_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is a decomposition in a minimal linear combination of rank-1 tensors (a proof is given in [19]).

2.3. Scalar product, orthogonality, norm of higher-order tensors. In the HOSVD definition of section 3, the structure constraint of diagonality of the matrix of singular values in the second-order case will be replaced by a number of geometrical conditions. This requires a generalization of the well-known definitions of scalar product, orthogonality, and Frobenius-norm to tensors of arbitrary order. These generalizations can be defined in a straightforward way as follows.

DEFINITION 5. *The scalar product $\langle \mathcal{A}, \mathcal{B} \rangle$ of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ is defined as*

$$\langle \mathcal{A}, \mathcal{B} \rangle \stackrel{\text{def}}{=} \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} b_{i_1 i_2 \dots i_N}^* a_{i_1 i_2 \dots i_N},$$

in which $*$ denotes the complex conjugation.

DEFINITION 6. *Arrays of which the scalar product equals 0 are orthogonal.*

DEFINITION 7. *The Frobenius-norm of a tensor \mathcal{A} is given by*

$$\|\mathcal{A}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}.$$

2.4. Multiplication of a higher-order tensor by a matrix. Like for matrices, the HOSVD of a higher-order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ will be defined by looking for orthogonal coordinate transformations of $\mathbb{R}^{I_1}, \mathbb{R}^{I_2}, \dots, \mathbb{R}^{I_N}$ that induce a particular representation of the higher-order tensor. In this section we establish a notation for the multiplication of a higher-order tensor by a matrix. This will allow us to express the effect of basis transformations.

Let us first have a look at the matrix product $\mathbf{G} = \mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^T$, involving matrices $\mathbf{F} \in \mathbb{R}^{I_1 \times I_2}$, $\mathbf{U} \in \mathbb{R}^{J_1 \times I_1}$, $\mathbf{V} \in \mathbb{R}^{J_2 \times I_2}$, and $\mathbf{G} \in \mathbb{R}^{J_1 \times J_2}$. To avoid working with “generalized transposes” in the multilinear case, we observe that the relationship between \mathbf{U} and \mathbf{F} and the relationship between \mathbf{V} (not \mathbf{V}^T) and \mathbf{F} are in fact completely similar: in the same way as \mathbf{U} makes linear combinations of the rows of \mathbf{F} , \mathbf{V} makes linear combinations of the columns of \mathbf{F} ; in the same way as the columns of \mathbf{F} are multiplied by \mathbf{U} , the rows of \mathbf{F} are multiplied by \mathbf{V} ; in the same way as the columns of \mathbf{U} are associated with the column space of \mathbf{G} , the columns of \mathbf{V} are associated with the row space of \mathbf{G} . This typical relationship will be denoted by means of the \times_n -symbol: $\mathbf{G} = \mathbf{F} \times_1 \mathbf{U} \times_2 \mathbf{V}$. (For complex matrices the product $\mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^H$ is consequently denoted as $\mathbf{F} \times_1 \mathbf{U} \times_2 \mathbf{V}^*$.) In general, we have the following definition.

DEFINITION 8. *The n -mode product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ by a matrix $\mathbf{U} \in \mathbb{C}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{U}$, is an $(I_1 \times I_2 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N)$ -tensor of which the entries are given by*

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \dots i_{n-1} j_n i_{n+1} \dots i_N} \stackrel{\text{def}}{=} \sum_{i_n} a_{i_1 i_2 \dots i_{n-1} i_n i_{n+1} \dots i_N} u_{j_n i_n}.$$

The n -mode product satisfies the following properties.

PROPERTY 2. *Given the tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ and the matrices $\mathbf{F} \in \mathbb{C}^{J_n \times I_n}$, $\mathbf{G} \in \mathbb{C}^{J_m \times I_m}$ ($n \neq m$), one has*

$$(\mathcal{A} \times_n \mathbf{F}) \times_m \mathbf{G} = (\mathcal{A} \times_m \mathbf{G}) \times_n \mathbf{F} = \mathcal{A} \times_n \mathbf{F} \times_m \mathbf{G}.$$

PROPERTY 3. *Given the tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ and the matrices $\mathbf{F} \in \mathbb{C}^{J_n \times I_n}$, $\mathbf{G} \in \mathbb{C}^{K_n \times J_n}$, one has*

$$(\mathcal{A} \times_n \mathbf{F}) \times_n \mathbf{G} = \mathcal{A} \times_n (\mathbf{G} \cdot \mathbf{F}).$$

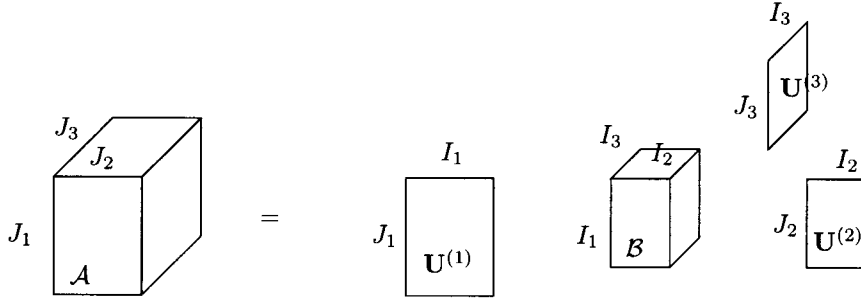


FIG. 2. Visualization of the multiplication of a third-order tensor $\mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ with matrices $\mathbf{U}^{(1)} \in \mathbb{C}^{J_1 \times I_1}$, $\mathbf{U}^{(2)} \in \mathbb{C}^{J_2 \times I_2}$, and $\mathbf{U}^{(3)} \in \mathbb{C}^{J_3 \times I_3}$.

Figure 2 visualizes the equation $\mathcal{A} = \mathcal{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$ for third-order tensors $\mathcal{A} \in \mathbb{C}^{J_1 \times J_2 \times J_3}$ and $\mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$. Unlike the conventional way to visualize second-order matrix products, $\mathbf{U}^{(2)}$ has not been transposed, for reasons of symmetry. Multiplication with $\mathbf{U}^{(1)}$ involves linear combinations of the “horizontal matrices” (index i_1 fixed) in \mathcal{B} . Stated otherwise, multiplication of \mathcal{B} with $\mathbf{U}^{(1)}$ means that every column of \mathcal{B} (indices i_2 and i_3 fixed) has to be multiplied from the left with $\mathbf{U}^{(1)}$. Similarly, multiplication with $\mathbf{U}^{(2)}$, resp., $\mathbf{U}^{(3)}$, involves linear combinations of matrices, obtained by fixing i_2 , resp., i_3 . This can be considered as a multiplication, from the left, of the vectors obtained by fixing the indices i_3 and i_1 , resp., i_1 and i_2 . Visualization schemes like Figure 2 have proven to be very useful to gain insight in tensor techniques.

The n -mode product of a tensor and a matrix is a special case of the inner product in multilinear algebra and tensor analysis [32, 26]. In the literature it often takes the form of an Einstein summation convention. Without going into details, this means that summations are written in full, but that the summation sign is dropped for the index that is repeated. This is of course the most versatile way to write down tensor equations, and in addition, a basis-independent interpretation can be given to Einstein summations. On the other hand, the formalism is only rarely used in signal processing and numerical linear algebra, whereas using the \times_n -symbol comes closer to the common way of dealing with matrix equations. It is our experience that the use of the \times_n -symbol demonstrates more clearly the analogy between matrix and tensor SVD, and that, more in general, a conceptual insight in tensor decompositions is easier induced by means of the \times_n -notation and visualizations like Figure 2 than by the use of element-wise summations.

3. A multilinear SVD. In this section an SVD model is proposed for N th-order tensors. To facilitate the comparison, we first repeat the matrix decomposition in the same notation, as follows.

THEOREM 1 (matrix SVD). *Every complex $(I_1 \times I_2)$ -matrix \mathbf{F} can be written as the product*

$$(2) \quad \mathbf{F} = \mathbf{U}^{(1)} \cdot \mathbf{S} \cdot \mathbf{V}^{(2)H} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{V}^{(2)*} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)},$$

in which

$$1. \mathbf{U}^{(1)} = \left(U_1^{(1)} U_2^{(1)} \dots U_{I_1}^{(1)} \right) \text{ is a unitary } (I_1 \times I_1)\text{-matrix,}$$

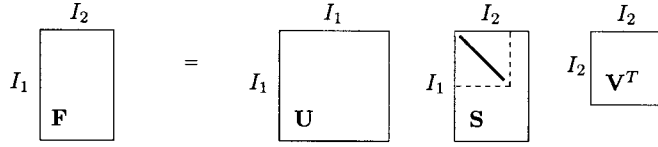


FIG. 3. Visualization of the matrix SVD.

2. $\mathbf{U}^{(2)} = \left(U_1^{(2)} U_2^{(2)} \dots U_{I_2}^{(2)} \right) (= \mathbf{V}^{(2)*})$ is a unitary $(I_2 \times I_2)$ -matrix,
3. \mathbf{S} is an $(I_1 \times I_2)$ -matrix with the properties of
 - (i) pseudodiagonality:

$$(3) \quad \mathbf{S} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{\min(I_1, I_2)}),$$

(ii) ordering:

$$(4) \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(I_1, I_2)} \geq 0.$$

The σ_i are singular values of \mathbf{F} and the vectors $U_i^{(1)}$ and $U_i^{(2)}$ are, resp., an i th left and an i th right singular vector. The decomposition is visualized in Figure 3.

Now we state the following theorem.

THEOREM 2 (N th-order SVD). *Every complex $(I_1 \times I_2 \times \dots \times I_N)$ -tensor \mathcal{A} can be written as the product*

$$(5) \quad \mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)},$$

in which

1. $\mathbf{U}^{(n)} = \left(U_1^{(n)} U_2^{(n)} \dots U_{I_n}^{(n)} \right)$ is a unitary $(I_n \times I_n)$ -matrix,
2. \mathcal{S} is a complex $(I_1 \times I_2 \times \dots \times I_N)$ -tensor of which the subtensors $\mathcal{S}_{i_n=\alpha}$, obtained by fixing the n th index to α , have the properties of
 - (i) all-orthogonality: two subtensors $\mathcal{S}_{i_n=\alpha}$ and $\mathcal{S}_{i_n=\beta}$ are orthogonal for all possible values of n , α and β subject to $\alpha \neq \beta$:

$$(6) \quad \langle \mathcal{S}_{i_n=\alpha}, \mathcal{S}_{i_n=\beta} \rangle = 0 \quad \text{when} \quad \alpha \neq \beta,$$

(ii) ordering:

$$(7) \quad \|\mathcal{S}_{i_n=1}\| \geq \|\mathcal{S}_{i_n=2}\| \geq \dots \geq \|\mathcal{S}_{i_n=I_n}\| \geq 0$$

for all possible values of n .

The Frobenius-norms $\|\mathcal{S}_{i_n=i}\|$, symbolized by $\sigma_i^{(n)}$, are n -mode singular values of \mathcal{A} and the vector $U_i^{(n)}$ is an i th n -mode singular vector. The decomposition is visualized for third-order tensors in Figure 4.

Discussion. Applied to a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, Theorem 2 says that it is always possible to find orthogonal transformations of the column, row, and 3-mode space such that $\mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^{(1)T} \times_2 \mathbf{U}^{(2)T} \times_3 \mathbf{U}^{(3)T}$ is all-orthogonal and ordered (the new basis vectors are the columns of $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, and $\mathbf{U}^{(3)}$). All-orthogonality means that the different “horizontal matrices” of \mathcal{S} (the first index i_1 is kept fixed, while the two other indices, i_2 and i_3 , are free) are mutually orthogonal with respect to the scalar product

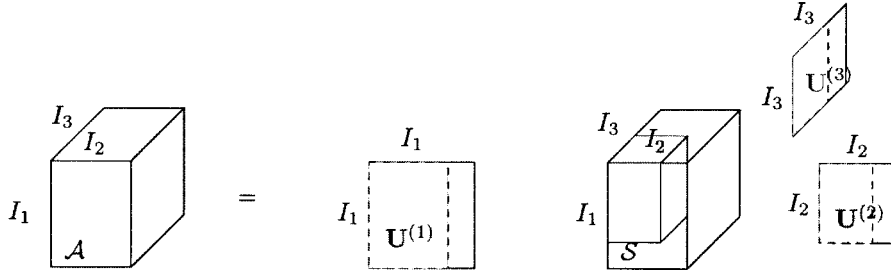


FIG. 4. Visualization of the HOSVD for a third-order tensor.

of matrices (i.e., the sum of the products of the corresponding entries vanishes); at the same time, the different “frontal” matrices (i_2 fixed) and the different “vertical” matrices (i_3 fixed) should be mutually orthogonal as well. The ordering constraint imposes that the Frobenius-norm of the horizontal (frontal, resp., vertical) matrices does not increase as the index i_1 (i_2 , resp., i_3) is increased. While the orthogonality of $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, $\mathbf{U}^{(3)}$, and the all-orthogonality of \mathcal{S} are the basic assumptions of the model, the ordering condition should be regarded as a convention, meant to fix a particular ordering of the columns of $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, and $\mathbf{U}^{(3)}$ (or the horizontal, frontal, and vertical matrices of \mathcal{S} , stated otherwise).

Comparison of the matrix and tensor theorem shows a clear analogy between the two cases. First, the left and right singular vectors of a matrix are generalized as the n -mode singular vectors. Next, the role of the singular values is taken over by the Frobenius-norms of the $(N - 1)$ th-order subtensors of the “core tensor” \mathcal{S} ; notice at this point that in the matrix case, the singular values also correspond to the Frobenius-norms of the rows and the columns of the “core matrix” \mathbf{S} . For N th-order tensors, N (possibly different) sets of n -mode singular values are defined; in this respect, remember from section 2.2 that an N th-order tensor can also have N different n -rank values. The essential difference is that \mathcal{S} is in general a full tensor, instead of being pseudodiagonal (this would mean that nonzero elements could only occur when the indices $i_1 = i_2 = \dots = i_N$). Instead, \mathcal{S} obeys the condition of all-orthogonality; here we notice that in the matrix case \mathbf{S} is all-orthogonal as well: due to the diagonal structure, the scalar product of two different rows or columns also vanishes. We also remark that, by definition, the n -mode singular values are positive and real, like in the matrix case. On the other hand the entries of \mathcal{S} are not necessarily positive in general; they can even be complex, when \mathcal{A} is a complex-valued tensor.

One could wonder whether imposing the condition of pseudodiagonality on the core tensor \mathcal{S} would not be a better way to generalize the SVD of matrices. The answer is negative: in general, it is impossible to reduce higher-order tensors to a pseudodiagonal form by means of orthogonal transformations. This is easily shown by counting degrees of freedom: pseudodiagonality of a core tensor containing I nonzero elements would imply that the decomposition would exhibit not more than $I(\sum I_n + 1 - N(I + 1)/2)$ degrees of freedom, while the original tensor contains $I_1 I_2 \dots I_N$ independent entries. Only in the second-order case both quantities are equal for $I = \min(I_1, I_2)$ —only in the second-order case, the condition of pseudodiagonality makes sense. However, we will prove that relaxation of the pseudodiagonality condition

to all-orthogonality yields a decomposition that always exists. As a matter of fact, “relaxation” is a too hard term: Property 5 in section 4 will show that the matrix SVD itself could have been defined by requiring the rows/columns of the matrix \mathbf{S} to be mutually orthogonal; apart from some trivial normalization conventions, the resulting decomposition is exactly the same as the conventional one, obtained via the pseudodiagonality condition.

Equivalent representations. A *matrix representation* of the HOSVD can be obtained by unfolding \mathcal{A} and \mathcal{S} in model equation (5):

$$\mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{S}_{(n)} \cdot \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \dots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \dots \otimes \mathbf{U}^{(n-1)} \right)^T, \quad (8)$$

in which \otimes denotes the Kronecker product [2, 40]. (The Kronecker product of two matrices $\mathbf{F} \in \mathbb{C}^{I_1 \times I_2}$ and $\mathbf{G} \in \mathbb{C}^{J_1 \times J_2}$ is defined according to

$$\mathbf{F} \otimes \mathbf{G} \stackrel{\text{def}}{=} (f_{i_1 i_2} \mathbf{G})_{1 \leq i_1 \leq I_1; 1 \leq i_2 \leq I_2} \cdot)$$

Notice that the conditions (6) and (7) imply that $\mathbf{S}_{(n)}$ has mutually orthogonal rows, having Frobenius-norms equal to $\sigma_1^{(n)}, \sigma_2^{(n)}, \dots, \sigma_{I_n}^{(n)}$. Let us define a diagonal matrix $\mathbf{\Sigma}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ and a column-wise orthonormal matrix $\mathbf{V}^{(n)} \in \mathbb{C}^{I_{n+1} I_{n+2} \dots I_N I_1 I_2 \dots I_{n-1} \times I_n}$ according to

$$(9) \quad \mathbf{\Sigma}^{(n)} \stackrel{\text{def}}{=} \text{diag}(\sigma_1^{(n)}, \sigma_2^{(n)}, \dots, \sigma_{I_n}^{(n)}),$$

$$(10) \quad \mathbf{V}^{(n)H} \stackrel{\text{def}}{=} \tilde{\mathbf{S}}_{(n)} \cdot \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \dots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \dots \otimes \mathbf{U}^{(n-1)} \right),$$

in which $\tilde{\mathbf{S}}_{(n)}$ is a normalized version of $\mathbf{S}_{(n)}$, with the rows scaled to unit-length

$$(11) \quad \mathbf{S}_{(n)} = \mathbf{\Sigma}^{(n)} \cdot \tilde{\mathbf{S}}_{(n)}.$$

Expressing (8) in terms of $\mathbf{\Sigma}^{(n)}$ and $\mathbf{V}^{(n)}$ shows that, at a matrix level, the HOSVD conditions lead to an SVD of the matrix unfoldings

$$(12) \quad \mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{\Sigma}^{(n)} \cdot \mathbf{V}^{(n)H}$$

($1 \leq n \leq N$). Below we will show that, on the other hand, the left singular matrices of the different matrix unfoldings of \mathcal{A} correspond to unitary transformations that induce the HOSVD structure. This strong link ensures that the HOSVD inherits all the classical column/row space properties from the matrix SVD (see section 4).

The dyadic decomposition could be generalized by expressing the HOSVD model as an *expansion of mutually orthogonal rank-1 tensors*,

$$(13) \quad \mathcal{A} = \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} s_{i_1 i_2 \dots i_N} U_{i_1}^{(1)} \circ U_{i_2}^{(2)} \circ \dots \circ U_{i_N}^{(N)},$$

in which the coefficients $s_{i_1 i_2 \dots i_N}$ are the entries of an ordered all-orthogonal tensor \mathcal{S} . The orthogonality of the rank-1 terms follows from the orthogonality of the n -mode singular vectors. In connection with the discussion on pseudodiagonality versus all-orthogonality, we remark that the summation generally involves $r_1 r_2 \dots r_N$ terms (instead of $\min(I_1, I_2, \dots, I_N)$), in which r_n is the highest index for which $\|\mathcal{S}_{i_n=r_n}\| >$

$$\mathcal{A} = \dots + s_{i_1 i_2 i_3} \begin{array}{c} U_{i_3}^{(3)} \diagup \\ U_{i_2}^{(2)} \text{---} \\ U_{i_1}^{(1)} \text{---} \end{array} + \dots$$

FIG. 5. Visualization of a triadic decomposition.

0 in (7). In Figure 5 this decomposition is visualized for third-order tensors. Where the dyadic decomposition expresses a matrix in an essentially unique way as a minimal linear combination of products of column and row vectors that are each mutually orthonormal, the meaning of (13) is less outspoken. For example, the matrix $\mathbf{U}'^{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{Q}$, in which \mathbf{Q} is a unitary matrix, together with the tensor $\mathcal{S}' = \mathcal{S} \times_n \mathbf{Q}^H$ still leads to an expansion in $r_1 r_2 \dots r_N$ mutually orthogonal rank-1 tensors (however, \mathcal{S}' is not all-orthogonal in general). The unitary matrix \mathbf{Q} could even be chosen in such a way that it induces zero entries in \mathcal{S}' , thereby decreasing the number of terms in the rank-1 expansion (e.g., the unitary factor of a QR -decomposition of $\mathbf{S}_{(n)}$ induces $r_n(r_n - 1)/2$ zeros).

Proof. The derivation of Theorem 2 establishes the connection between the HOSVD of a tensor \mathcal{A} and the matrix SVD of its matrix unfoldings. It is given in terms of real-valued tensors; the complex case is completely analogous but more cumbersome from a notational point of view.

Consider two $(I_1 \times I_2 \times \dots \times I_N)$ tensors \mathcal{A} and \mathcal{S} , related by

$$(14) \quad \mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^{(1)T} \times_2 \mathbf{U}^{(2)T} \dots \times_N \mathbf{U}^{(N)T},$$

in which $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(N)}$ are orthogonal matrices. Equation (14) can be expressed in a matrix format as

$$(15) \quad \mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{S}_{(n)} \cdot \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \dots \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \dots \mathbf{U}^{(n-1)} \right)^T.$$

Now consider the particular case where $\mathbf{U}^{(n)}$ is obtained from the SVD of $\mathbf{A}_{(n)}$ as

$$(16) \quad \mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{\Sigma}^{(n)} \cdot \mathbf{V}^{(n)T},$$

in which $\mathbf{V}^{(n)}$ is orthogonal and $\mathbf{\Sigma}^{(n)} = \text{diag}(\sigma_1^{(n)}, \sigma_2^{(n)}, \dots, \sigma_{I_n}^{(n)})$, where

$$(17) \quad \sigma_1^{(n)} \geq \sigma_2^{(n)} \geq \dots \geq \sigma_{I_n}^{(n)} \geq 0.$$

We call r_n the highest index for which $\sigma_{r_n}^{(n)} > 0$. Taking into account that the Kronecker factor in (15) is orthogonal, comparison of (15) and (16) shows that

$$(18) \quad \mathbf{S}_{(n)} = \mathbf{\Sigma}^{(n)} \cdot \mathbf{V}^{(n)T} \cdot \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \dots \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \dots \mathbf{U}^{(n-1)} \right).$$

This equation implies, for arbitrary orthogonal matrices $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \dots, \mathbf{U}^{(N)}$, that

$$(19) \quad \langle \mathcal{S}_{i_n=\alpha}, \mathcal{S}_{i_n=\beta} \rangle = 0 \quad \text{when} \quad \alpha \neq \beta,$$

and

$$(20) \quad \|\mathcal{S}_{i_n=1}\| = \sigma_1^{(n)} \geq \|\mathcal{S}_{i_n=2}\| = \sigma_2^{(n)} \geq \dots \geq \|\mathcal{S}_{i_n=I_n}\| = \sigma_{I_n}^{(n)} \geq 0,$$

and, if $r_n < I_n$,

$$(21) \quad \|\mathcal{S}_{i_n=r_n+1}\| = \sigma_{r_n+1}^{(n)} = \dots = \|\mathcal{S}_{i_n=I_n}\| = \sigma_{I_n}^{(n)} = 0.$$

By constructing the matrices $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \dots, \mathbf{U}^{(N)}$ in a similar way as $\mathbf{U}^{(n)}$, \mathcal{S} can be made to satisfy all the conditions of the HOSVD theorem. On the other hand, as can be seen from (12), all the matrices $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(N)}$ and tensors \mathcal{S} satisfying the HOSVD theorem can be found by means of the SVD of $\mathbf{A}_{(1)}, \mathbf{A}_{(2)}, \dots, \mathbf{A}_{(N)}$, where \mathcal{S} follows from (14). \square

Computation. Equation (12) and the preceding proof actually indicate how the HOSVD of a given tensor \mathcal{A} can be computed: the n -mode singular matrix $\mathbf{U}^{(n)}$ (and the n -mode singular values) can directly be found as the left singular matrix (and the singular values) of an n -mode matrix unfolding of \mathcal{A} (any matrix of which the columns are given by the n -mode vectors can be resorted to, as the column ordering is of no importance). Hence computing the HOSVD of an N th-order tensor leads to the computation of N different matrix SVDs of matrices with size $(I_n \times I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N) (1 \leq n \leq N)$.

Afterwards, the core tensor \mathcal{S} can be computed by bringing the matrices of singular vectors to the left side of (5):

$$(22) \quad \mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^{(1)H} \times_2 \mathbf{U}^{(2)H} \dots \times_N \mathbf{U}^{(N)H}.$$

This can be computed in a matrix format, e.g.,

$$(23) \quad \mathbf{S}_{(n)} = \mathbf{U}^{(1)H} \cdot \mathbf{A}_{(n)} \cdot \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \dots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \dots \otimes \mathbf{U}^{(n-1)} \right).$$

Equations (12) and (23) essentially form a square-root version of the “operational procedures,” discussed in [43]. As such they are numerically more reliable, especially for ill-conditioned tensors, i.e., tensors for which $\sigma_1^{(n)} \gg \sigma_{R_n}^{(n)}$, for one or more values of n [22].

Example 4. Consider the $(3 \times 3 \times 3)$ tensor \mathcal{A} defined by a matrix unfolding $\mathbf{A}_{(1)}$, equal to

$$\left(\begin{array}{ccc|ccc|ccc} 0.9073 & 0.7158 & -0.3698 & 1.7842 & 1.6970 & 0.0151 & 2.1236 & -0.0740 & 1.4429 \\ 0.8924 & -0.4898 & 2.4288 & 1.7753 & -1.5077 & 4.0337 & -0.6631 & 1.9103 & -1.7495 \\ 2.1488 & 0.3054 & 2.3753 & 4.2495 & 0.3207 & 4.7146 & 1.8260 & 2.1335 & -0.2716 \end{array} \right).$$

The 1-mode singular vectors are the columns of the left singular matrix of $\mathbf{A}_{(1)}$; in the same way, $\mathbf{U}^{(2)}$ and $\mathbf{U}^{(3)}$ can be obtained:

$$\mathbf{U}^{(1)} = \begin{pmatrix} 0.1121 & -0.7739 & -0.6233 \\ 0.5771 & 0.5613 & -0.5932 \\ 0.8090 & -0.2932 & 0.5095 \end{pmatrix},$$

$$\mathbf{U}^{(2)} = \begin{pmatrix} 0.4624 & 0.0102 & 0.8866 \\ 0.8866 & -0.0135 & -0.4623 \\ -0.0072 & -0.9999 & 0.0152 \end{pmatrix},$$

$$\mathbf{U}^{(3)} = \begin{pmatrix} 0.6208 & -0.4986 & 0.6050 \\ -0.0575 & -0.7986 & -0.5992 \\ 0.7819 & 0.3371 & -0.5244 \end{pmatrix}.$$

The core tensor of the HOSVD then follows from application of (23); its unfolding $\mathbf{S}_{(1)}$ is equal to

$$\left(\begin{array}{ccc|ccc|ccc} 8.7088 & 0.0489 & -0.2797 & 0.1066 & 3.2737 & 0.3223 & -0.0033 & -0.1797 & -0.2222 \\ -0.0256 & 3.2546 & -0.2853 & 3.1965 & -0.2130 & 0.7829 & 0.2948 & -0.0378 & -0.3704 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{array} \right).$$

The core tensor is all-orthogonal: the rows of $\mathbf{S}_{(1)}$ are mutually orthogonal, but so also are the matrices formed by columns 1/2/3, 4/5/6, and 7/8/9, as well as the three matrices formed by columns 1/4/7, 2/5/8, and 3/6/9. This boils down to orthogonality of, resp., the “horizontal,” “frontal,” and “vertical” matrices of \mathcal{A} . The core tensor is also ordered: its matrices are put in order of decreasing Frobenius-norm. The Frobenius-norms give the n -mode singular values of \mathcal{A} :

$$\begin{aligned} \text{mode 1 : } & 9.3187, 4.6664, 0, \\ \text{mode 2 : } & 9.3058, 4.6592, 0.5543, \\ \text{mode 3 : } & 9.2822, 4.6250, 1.0310. \end{aligned}$$

4. Properties. Many properties of the matrix SVD have a very clear higher-order counterpart, because of the strong link between the HOSVD of a higher-order tensor and the SVDs of its matrix unfoldings. In this section, we list the multilinear equivalents of a number of classical matrix SVD properties. The basic link itself is repeated as Property 12. The proofs are outlined at the end of the section.

PROPERTY 4 (uniqueness).

(i) The n -mode singular values are uniquely defined.

(ii) When the n -mode singular values are different, then the n -mode singular vectors are determined up to multiplication with a unit-modulus factor. When $U_{\alpha}^{(n)}$ is multiplied by $e^{j\theta}$, then $\mathcal{S}_{i_n=\alpha}$ has to be multiplied by the inverse factor $e^{-j\theta}$.

The n -mode singular vectors corresponding to the same n -mode singular value can be replaced by any unitary linear combination. The corresponding subtensors $\{\mathcal{S}_{i_n=\alpha}\}$ have to be combined in the inverse way. Formally, $\mathbf{U}^{(n)}$ can be replaced by $\mathbf{U}^{(n)} \cdot \mathbf{Q}$, in which \mathbf{Q} is a block-diagonal matrix, consisting of unitary blocks, where the block-partitioning corresponds to the partitioning of $\mathbf{U}^{(n)}$ in sets of n -mode singular vectors with identical n -mode singular value. At the same time \mathcal{S} has to be replaced by $\mathcal{S} \times_n \mathbf{Q}^H$.

For real-valued tensors uniqueness is up to the sign, resp., multiplication with an orthogonal matrix.

The first property implies that the HOSVD shows essentially the same uniqueness properties as the matrix SVD. The only difference consists of the fact that Theorem 2 contains weaker normalization conventions. The equivalent situation for matrices would be to allow that \mathbf{S} consists of diagonal blocks, corresponding to the different singular values, in which each block consists of a unitary matrix, multiplied with the singular value under consideration.

PROPERTY 5 (generalization). The tensor SVD of a second-order tensor boils down (up to the underdetermination) to its matrix SVD.

From the discussion in section 3 it is clear that the HOSVD is a formal equivalent of the matrix SVD. Moreover, according to Property 5, it is a true generalization in the sense that, when Theorem 2 is applied to matrices (second-order tensors), it leads to the classical matrix SVD. (Note however that, by convention, the 2-mode singular vectors are *defined* as the complex conjugates of the right matrix singular vectors.) As such, Theorem 2 really establishes a multilinear SVD framework, containing the matrix decomposition in the special case of second-order tensors.

PROPERTY 6 (n -rank). *Let the HOSVD of \mathcal{A} be given as in Theorem 2, and let r_n be equal to the highest index for which $\|\mathcal{S}_{i_n=r_n}\| > 0$ in (7); then one has*

$$R_n = \text{rank}_n(\mathcal{A}) = r_n.$$

The fact that the number of nonvanishing singular values of a given matrix equals its (column/row) rank carries over to the n -mode singular values and the n -rank values of a given tensor (recall from section 2.2 that the n -rank values are not necessarily the same). Like for matrices, this link even holds in a numerical sense, as will be shown by the perturbation analysis in section 6: the number of significant n -mode singular values of a given tensor equals its numerical n -rank. In a matrix context, this property is of major concern for the estimation of “problem dimensionalities,” like the estimation of the number of sources in the source separation problem, the estimation of filter lengths in identification, the estimation of the number of harmonics in the harmonic retrieval problem, etc. [45, 46]. Property 6 may play a similar role in multilinear algebra. Let us illustrate this by means of a small example. Consider the most elementary relationship in multivariate statistics, in which an I -dimensional stochastic vector X consists of a linear mixture of $J \leq I$ stochastic components. Whereas the number of components is usually estimated as the number of significant eigenvalues of the covariance matrix of X , the number of skew or kurtic components might as well be estimated as the number of significant n -mode singular values of the third-order, resp., fourth-order, cumulant tensor of X [17].

Finally, we remember from section 2.2 that knowledge of the n -rank values of a given tensor does not allow us in general to make precise statements about the rank of that tensor. With this respect, other SVD generalizations might be more interesting (see section 7).

PROPERTY 7 (structure). *Let the HOSVD of \mathcal{A} be given as in Theorem 2; then one has*

- (i) *the n -mode vector space $\mathbf{R}(\mathbf{A}_{(n)}) = \text{span}(U_1^{(n)}, \dots, U_{R_n}^{(n)})$,*
- (ii) *the orthogonal complement of $\mathbf{R}(\mathbf{A}_{(n)})$, the left n -mode null space $\mathbf{N}(\mathbf{A}_{(n)}^H) = \text{span}(U_{R_n+1}^{(n)}, \dots, U_{I_n}^{(n)})$.*

In the same way as the left and right singular vectors of a matrix give an orthonormal basis for its column and row space (and their orthogonal complements), the n -mode singular vectors of a given tensor yield an orthonormal basis for its n -mode vector space (and its orthogonal complement). For matrices, the property was the starting-point for the development of subspace algorithms [45, 46, 47]; Property 7 allows for an extension of this methodology in multilinear algebra.

PROPERTY 8 (norm). *Let the HOSVD of \mathcal{A} be given as in Theorem 2; then the following holds.*

$$\begin{aligned} \|\mathcal{A}\|^2 &= \sum_{i=1}^{R_1} \left(\sigma_i^{(1)}\right)^2 = \dots = \sum_{i=1}^{R_N} \left(\sigma_i^{(N)}\right)^2 \\ &= \|\mathcal{S}\|^2. \end{aligned}$$

In multilinear algebra as well as in matrix algebra, the Frobenius-norm is unitarily invariant. As a consequence, the well-known fact that the squared Frobenius-norm of a matrix equals the sum of its squared singular values can be generalized.

DEFINITION 9 (*n*-mode oriented energy). *The n -mode oriented energy of an N th-order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ in the direction of a unit-norm vector X , denoted by $\text{OE}_n(X, \mathcal{A})$, is the oriented energy of the set of n -mode vectors, i.e.,*

$$\text{OE}_n(X, \mathcal{A}) \stackrel{\text{def}}{=} \|X^H \mathbf{A}_{(n)}\|^2.$$

The concept of oriented energy of a given matrix and the link with the SVD of that matrix, which form the basis of SVD-based signal separation algorithms, can easily be generalized as well. Whereas Property 8 merely states that the squared Frobenius-norm of a tensor (i.e., the “energy” contained in the tensor) equals the sum of the squared n -mode singular values, the HOSVD actually gives a pretty detailed geometrical picture of the energy distribution over the n -mode vector space. Definition 9 generalizes the definition of oriented energy. We now state the following property.

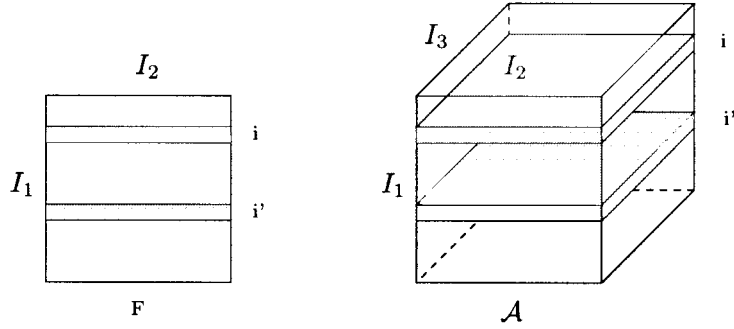
PROPERTY 9 (oriented energy). *The directions of extremal n -mode oriented energy correspond to the n -mode singular vectors, with extremal energy value equal to the corresponding squared n -mode singular value.*

This means that the n -mode vectors mainly contain contributions in the direction of $U_1^{(n)}$; this particular direction accounts for an amount of $\sigma_1^{(n)^2}$ with respect to the total amount of energy in the tensor. Next, the n -mode oriented energy reaches an extremum in the direction of $U_2^{(n)}$, perpendicular to $U_1^{(n)}$, for a value of $\sigma_2^{(n)^2}$, and so on. Discarding the components of the n -mode vectors in the direction of an n -mode singular vector $U_{i_n}^{(n)}$ (e.g., the one corresponding to the smallest n -mode singular value) to obtain a tensor $\hat{\mathcal{A}}$ introduces an error $\|\mathcal{A} - \hat{\mathcal{A}}\|^2 = \sigma_{i_n}^{(n)^2}$.

PROPERTY 10 (approximation). *Let the HOSVD of \mathcal{A} be given as in Theorem 2 and let the n -mode rank of \mathcal{A} be equal to R_n ($1 \leq n \leq N$). Define a tensor $\hat{\mathcal{A}}$ by discarding the smallest n -mode singular values $\sigma_{I'_n+1}^{(n)}, \sigma_{I'_n+2}^{(n)}, \dots, \sigma_{R_n}^{(n)}$ for given values of I'_n ($1 \leq n \leq N$), i.e., set the corresponding parts of \mathcal{S} equal to zero. Then we have*

$$(24) \quad \|\mathcal{A} - \hat{\mathcal{A}}\|^2 \leq \sum_{i_1=I'_1+1}^{R_1} \sigma_{i_1}^{(1)^2} + \sum_{i_2=I'_2+1}^{R_2} \sigma_{i_2}^{(2)^2} + \dots + \sum_{i_N=I'_N+1}^{R_N} \sigma_{i_N}^{(N)^2}.$$

This property is the higher-order equivalent of the link between the SVD of a matrix and its best approximation, in a least-squares sense, by a matrix of lower rank. However, the situation is quite different for tensors. By discarding the smallest n -mode singular values, one obtains a tensor $\hat{\mathcal{A}}$ with a column rank equal to I'_1 , row rank equal to I'_2 , etc. However, this tensor is in general not the best possible approximation under the given n -mode rank constraints (see, e.g., Example 5). Nevertheless, the ordering assumption (7) implies that the “energy” of \mathcal{A} is mainly concentrated in the part corresponding to low values of i_1, i_2, \dots, i_N . Consequently, if $\sigma_{I'_n}^{(n)} \gg \sigma_{I'_n+1}^{(n)}$ (e.g., I'_n corresponds to the numerical n -rank of \mathcal{A} ; the smaller n -mode singular values are not significant (see also Property 6)), $\hat{\mathcal{A}}$ is still to be considered as a good approximation of \mathcal{A} . The error is bounded as in (24). For procedures to enhance a given approximation, we refer to [28, 17].

FIG. 6. Construction of $\mathbf{H}^{(1)}$ for (a) a matrix and (b) a third-order tensor.

PROPERTY 11 (link between HOSVD and matrix EVD). *Let the HOSVD of \mathcal{A} be given as in Theorem 2. Define $\mathbf{H}^{(n)} \stackrel{\text{def}}{=} \mathbf{A}_{(n)} \cdot \mathbf{A}_{(n)}^H$, i.e., $\mathbf{H}^{(n)}$ contains on position (i, i') the scalar product $\langle \mathcal{A}_{i_n=i}, \mathcal{A}_{i_n=i'} \rangle$. If the EVD of $\mathbf{H}^{(n)}$ is given by*

$$\mathbf{H}^{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{D}^{(n)} \cdot \mathbf{U}^{(n)H},$$

then $\mathbf{U}^{(n)}$ contains the n -mode singular vectors of \mathcal{A} . Moreover the scalar product $\langle \mathcal{S}_{i_n=i}, \mathcal{S}_{i_n=i'} \rangle$ is the (i, i') th element of $\mathbf{D}^{(n)}$. Different subtensors of \mathcal{S} are mutually orthogonal.

In the same way as the SVD of a matrix \mathbf{F} is related to the EVD of the Hermitean (real symmetric) positive (semi)definite matrices $\mathbf{F}\mathbf{F}^H$ and $\mathbf{F}^H\mathbf{F}$, the HOSVD of a higher-order tensor \mathcal{A} is related to the EVD of Hermitean (real symmetric) positive (semi)definite matrices, constructed from \mathcal{A} . The construction is clarified in Figure 6: just like the entries of $\mathbf{F}\mathbf{F}^H$ consist of the mutual scalar products of the rows of \mathbf{F} , the matrix $\mathbf{H}^{(1)}$ in Property 11 is computed from the scalar products of the “horizontal matrices” of \mathcal{A} , in the third-order case. This property can, e.g., be useful for interpretations in a statistical context, where $\mathbf{H}^{(n)}$ corresponds to the sample correlation matrix of the n -mode vectors of \mathcal{A} .

PROPERTY 12 (link between HOSVD and matrix SVD). *Let the HOSVD of \mathcal{A} be given as in Theorem 2. Then*

$$\mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{\Sigma}^{(n)} \cdot \mathbf{V}^{(n)H}$$

is an SVD of $\mathbf{A}_{(n)}$, where the diagonal matrix $\mathbf{\Sigma}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ and the column-wise orthonormal matrix $\mathbf{V}^{(n)} \in \mathbb{C}^{I_{n+1}I_{n+2}\dots I_N I_1 I_2 \dots I_{n-1} \times I_n}$ are defined according to

$$\mathbf{\Sigma}^{(n)} \stackrel{\text{def}}{=} \text{diag}(\sigma_1^{(n)}, \sigma_2^{(n)}, \dots, \sigma_{I_n}^{(n)}),$$

$$\mathbf{V}^{(n)H} \stackrel{\text{def}}{=} \tilde{\mathbf{S}}_{(n)} \cdot \left(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \dots \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \dots \otimes \mathbf{U}^{(n-1)} \right)^T,$$

in which $\tilde{\mathbf{S}}_{(n)}$ is a normalized version of $\mathbf{S}_{(n)}$, with the rows scaled to unit-length

$$\mathbf{S}_{(n)} = \mathbf{\Sigma}^{(n)} \cdot \tilde{\mathbf{S}}_{(n)}.$$

Proof. The proofs are established by a further analysis of the derivation in section 3; the key idea behind this derivation is explained in Property 12.

Property 4 is proved by investigating the uniqueness conditions.

Property 5 follows by applying the same procedure to second-order tensors.

Property 6 follows from the combination of Property 12 and Property 1. In the same way the deduction of Property 7 and Property 9 is trivial.

Property 8: We have $\|\mathcal{A}\|^2 = \|\mathbf{A}_{(n)}\|^2$ which, in turn, equals $\sum_{i=1}^{r_n} \|\mathcal{S}_{i_n=i}\|^2 = \|\mathcal{S}\|^2$.

Property 10: We have

$$\begin{aligned}
 \|\mathcal{A} - \hat{\mathcal{A}}\|^2 &= \sum_{i_1=1}^{R_1} \sum_{i_2=1}^{R_2} \cdots \sum_{i_N=1}^{R_N} s_{i_1 i_2 \dots i_N}^2 - \sum_{i_1=1}^{I'_1} \sum_{i_2=1}^{I'_2} \cdots \sum_{i_N=1}^{I'_N} s_{i_1 i_2 \dots i_N}^2 \\
 &= \sum_{i_1=I'_1+1}^{R_1} \sum_{i_2=I'_2+1}^{R_2} \cdots \sum_{i_N=I'_N+1}^{R_N} s_{i_1 i_2 \dots i_N}^2 \\
 &\leq \sum_{i_1=I'_1+1}^{R_1} \sum_{i_2=1}^{R_2} \cdots \sum_{i_N=1}^{R_N} s_{i_1 i_2 \dots i_N}^2 + \sum_{i_1=1}^{R_1} \sum_{i_2=I'_2+1}^{R_2} \cdots \sum_{i_N=1}^{R_N} s_{i_1 i_2 \dots i_N}^2 \\
 &\quad + \cdots + \sum_{i_1=1}^{R_1} \sum_{i_2=1}^{R_2} \cdots \sum_{i_N=I'_N+1}^{R_N} s_{i_1 i_2 \dots i_N}^2 \\
 &= \sum_{i_1=I'_1+1}^{R_1} \sigma_{i_1}^{(1)^2} + \sum_{i_2=I'_2+1}^{R_2} \sigma_{i_2}^{(2)^2} + \cdots + \sum_{i_N=I'_N+1}^{R_N} \sigma_{i_N}^{(N)^2}.
 \end{aligned}$$

Property 11 follows from the link between matrix SVD and HOSVD, combined with the relation between matrix SVD and EVD. \square

Example 5. Consider the tensor \mathcal{A} and its HOSVD components, which are given in Example 4.

From the n -mode singular values, we see that $R_2 = R_3 = 3$, while $R_1 = 2$. Clearly, the column space of \mathcal{A} is only two-dimensional. The first two columns of $\mathbf{U}^{(1)}$ form an orthonormal basis for this vector space; the third 1-mode singular vector is orthogonal to the column space of \mathcal{A} .

The sum of the squared n -mode singular values is equal to 108.6136 for all three modes; 108.6136 is the squared Frobenius-norm of \mathcal{A} .

Discarding $\sigma_3^{(2)}$ and $\sigma_3^{(3)}$, i.e., replacing \mathcal{S} by a tensor $\hat{\mathcal{S}}$, having a matrix unfolding $\hat{\mathbf{S}}_{(1)}$ equal to

$$\left(\begin{array}{ccc|ccc|ccc} 8.7088 & 0.0489 & 0.0000 & 0.1066 & 3.2737 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.0256 & 3.2546 & 0.0000 & 3.1965 & -0.2130 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{array} \right)$$

gives an approximation $\hat{\mathcal{A}}$ for which $\|\mathcal{A} - \hat{\mathcal{A}}\| = 1.0880$. On the other hand, the tensor \mathcal{A}' that best matches \mathcal{A} while having three n -ranks equal to 2 is defined by the unfolding

$$\left(\begin{array}{ccc|ccc|ccc} 0.8188 & 0.8886 & -0.0784 & 1.7051 & 1.7320 & -0.0274 & 1.7849 & 0.2672 & 1.7454 \\ 1.0134 & -0.8544 & 2.1455 & 1.9333 & -1.5390 & 3.9886 & -0.2877 & 1.5266 & -2.0826 \\ 2.1815 & 0.0924 & 2.4019 & 4.3367 & 0.3272 & 4.6102 & 1.8487 & 2.1042 & -0.2894 \end{array} \right).$$

For this tensor, we have that $\|\mathcal{A} - \mathcal{A}'\| = 1.0848$.

5. A multilinear symmetric EVD. Many applications show highly symmetric higher-order tensors. As an example, higher-order moments and cumulants of a real random vector are invariant under arbitrary index permutations. The same holds, e.g., for the symmetric tensorial representation of homogeneous polynomials [12]. For matrices, this symmetry is reflected by the fact that the left and right singular vectors are, up to the sign, identical; this leads to a particular form of the EVD, in which the eigenmatrix is an orthogonal (unitary) matrix. In this section, we will investigate if similar properties hold for higher-order tensors as well.

First, we prove that, if a tensor is invariant to (or mapped onto its complex conjugate by) a permutation of its indices, this symmetry is also reflected by the HOSVD components.

THEOREM 3. *Let the HOSVD of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ be denoted as in Theorem 2. Consider a permutation \mathbf{P} of its indices and the decomposition of this permutation in a sequence of permutation cycles $\mathbf{C}_1, \mathbf{C}_2, \dots$. We assume, without losing generality (possibly by redefining the ordering of the indices of \mathcal{A}), that*

$$\begin{aligned} \mathbf{P}(i_1, i_2, \dots, i_N) &= (\mathbf{C}_1(i_1, i_2, \dots, i_{n_1}), \mathbf{C}_2(i_{n_1+1}, i_{n_1+2}, \dots, i_{n_2}), \dots) \\ &= (i_2, i_3, \dots, i_{n_1}, i_1; i_{n_1+2}, \dots, i_{n_2}, i_{n_1+1}; \dots). \end{aligned}$$

If $a_{i_1 i_2 \dots i_N} = a_{\mathbf{P}(i_1 i_2 \dots i_N)}$, then one has

$$\mathbf{U}^{(1)} = \mathbf{U}^{(2)} = \dots = \mathbf{U}^{(n_1)}; \mathbf{U}^{(n_1+1)} = \mathbf{U}^{(n_1+2)} = \dots = \mathbf{U}^{(n_2)}; \dots$$

If $a_{i_1 i_2 \dots i_N} = a_{\mathbf{P}(i_1 i_2 \dots i_N)}^*$, then one has

$$\mathbf{U}^{(1)} = \mathbf{U}^{(2)*} = \mathbf{U}^{(3)} = \dots; \mathbf{U}^{(n_1+1)} = \mathbf{U}^{(n_1+2)*} = \mathbf{U}^{(n_1+3)} = \dots; \dots$$

For permutation cycles with an odd number of elements, the matrix of singular vectors is real.

The core tensor \mathcal{S} exhibits the same permutation symmetry as \mathcal{A} .

Proof. Construct the higher-order tensors $\mathcal{A}', \mathcal{A}'', \dots$ by repeatedly permuting the indices by \mathbf{P} :

$$a'_{i_1, i_2, \dots, i_N} \stackrel{\text{def}}{=} a_{\mathbf{P}(i_1, i_2, \dots, i_N)}; \quad a''_{i_1, i_2, \dots, i_N} \stackrel{\text{def}}{=} a_{\mathbf{P}(\mathbf{P}(i_1, i_2, \dots, i_N))}; \quad \dots$$

For real symmetry, we have that $\mathcal{A} = \mathcal{A}' = \mathcal{A}'' = \dots$; for Hermitean symmetry, we have that $\mathcal{A} = (\mathcal{A}')^* = \mathcal{A}'' = \dots$. The same equalities hold for the matrix unfoldings $\mathbf{A}_{(n_1)}, \mathbf{A}'_{(n_1)}, \mathbf{A}''_{(n_1)}, \dots$, as well as for $\mathbf{A}_{(n_2)}, \mathbf{A}'_{(n_2)}, \mathbf{A}''_{(n_2)}, \dots$, etc. Hence the same symmetry is shown by the left singular matrices of these matrix unfoldings, which correspond to $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots$, resp., $\mathbf{U}^{(n_1+1)}, \mathbf{U}^{(n_1+2)}, \dots$, etc.

If \mathbf{P} corresponds to Hermitean symmetry, and, e.g., \mathbf{C}_1 is a cycle with an odd number of elements, then it can be shown that $\mathbf{U}^{(1)} = \mathbf{U}^{(1)*}$ by repeating the construction above n_1 times.

For an analysis of the symmetry of \mathcal{S} , we write down an element-wise form of (22) (we take the example of Hermitean symmetry):

$$s_{j_1 j_2 \dots j_N} = \sum_{i_1 i_2 \dots i_N} a_{i_1 i_2 \dots i_N} u_{i_1 j_1}^{(n_1)*} u_{i_2 j_2}^{(n_1)} u_{i_3 j_3}^{(n_1)*} \dots u_{i_{n_1+1} j_{n_1+1}}^{(n_2)*} u_{i_{n_1+2} j_{n_1+2}}^{(n_2)} u_{i_{n_1+3} j_{n_1+3}}^{(n_2)*} \dots \quad (25)$$

Permutation of the indices by \mathbf{P} yields the following element:

$$s_{\mathbf{P}(j_1 j_2 \dots j_N)} = \sum_{i_1 i_2 \dots i_N} a_{i_1 i_2 \dots i_N} u_{i_1 j_2}^{(n_1)*} u_{i_2 j_3}^{(n_1)} u_{i_3 j_4}^{(n_1)*} \dots u_{i_{n_1+1} j_{n_1+2}}^{(n_2)*} u_{i_{n_1+2} j_{n_1+3}}^{(n_2)} u_{i_{n_1+3} j_{n_1+4}}^{(n_2)*} \dots \quad (26)$$

Invoking the Hermitean symmetry of \mathcal{A} , and comparing (26) to (25), it follows that $s_{\mathbf{P}(j_1 j_2 \dots j_N)} = s_{j_1 j_2 \dots j_N}^*$. \square

A consequence of Theorem 3 is that the i -mode singular vectors, for different i , are fully related under a condition that we define as *pair-wise symmetry*.

DEFINITION 10. A higher-order tensor $\mathcal{A} \in \mathbb{C}^{I \times I \times \dots \times I}$ is called *pair-wise symmetric* when, for every pair of indices (i_{n_1}, i_{n_2}) , there exists a permutation $\mathbf{P}_{n_1 n_2}$ such that $i_{n_1} = \mathbf{P}_{n_1 n_2}(i_{n_2})$ and either $a_{\mathbf{P}_{n_1 n_2}(i_1 i_2 \dots i_N)} = a_{i_1 i_2 \dots i_N}$ or $a_{\mathbf{P}_{n_1 n_2}(i_1 i_2 \dots i_N)} = a_{i_1 i_2 \dots i_N}^*$.

Pair-wise symmetric higher-order tensors are a tensorial equivalent of symmetric and Hermitean matrices. Theorem 3 and Definition 10 lead in a very natural way to a generalization of the orthogonal (unitary) EVD of symmetric (Hermitean) matrices. An HOEVD for pair-wise symmetric higher-order tensors can be defined as follows.

THEOREM 4 (N th-order pair-wise symmetric EVD). Every pair-wise symmetric $(I \times I \times \dots \times I)$ -tensor \mathcal{A} can be written as the product

$$(27) \quad \mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)},$$

in which the following hold.

1. \mathcal{S} is an all-orthogonal $(I \times I \times \dots \times I)$ -tensor with the same pair-wise symmetry as \mathcal{A} .

2. $\mathbf{U}^{(1)} = \mathbf{U} = (U_1 U_2 \dots U_I)$ is a unitary $(I \times I)$ -matrix, equal to $\mathbf{U}^{(n)}$ or $\mathbf{U}^{(n)*}$ (for all n , $1 \leq n \leq N$), depending on the symmetry of \mathcal{A} as stated in Theorem 3.

The Frobenius-norms of the subtensors of \mathcal{S} , obtained by fixing one index to i , are the N th-order eigenvalues and are symbolized by λ_i . The vectors U_i are the N th-order eigenvectors.

Example 6. Consider a complex third-order tensor \mathcal{A} that satisfies $a_{\mathbf{P}(i_1 i_2 i_3)} = a_{i_2 i_3 i_1} = a_{i_1 i_2 i_3}^*$. According to Theorem 3, and taking into account that \mathbf{P} is a cycle with three elements, the HOSVD of \mathcal{A} is given by

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \times_3 \mathbf{U},$$

where \mathbf{U} is real and \mathcal{S} satisfies the same symmetries as \mathcal{A} . In fact, it is pretty obvious in this example that \mathbf{U} is real: $a_{i_1 i_2 i_3} = a_{i_2 i_3 i_1}^* = a_{i_3 i_1 i_2} = a_{i_1 i_2 i_3}^*$ implies that \mathcal{A} itself, and hence the whole decomposition, is real.

Example 7. A common way to deal with the complex nature of a random vector X in the definition of its fourth-order cumulant tensor \mathcal{C} is the following element-wise definition:

$$(28) \quad \begin{aligned} c_{i_1 i_2 i_3 i_4} &\stackrel{\text{def}}{=} \text{cum}(x_{i_1}, x_{i_2}^*, x_{i_3}^*, x_{i_4}) \\ &\stackrel{\text{def}}{=} \mathbf{E}\{\tilde{x}_{i_1} \tilde{x}_{i_2}^* \tilde{x}_{i_3}^* \tilde{x}_{i_4}\} - \mathbf{E}\{\tilde{x}_{i_1} \tilde{x}_{i_2}^*\} \mathbf{E}\{\tilde{x}_{i_3}^* \tilde{x}_{i_4}\} \\ &\quad - \mathbf{E}\{\tilde{x}_{i_1} \tilde{x}_{i_3}^*\} \mathbf{E}\{\tilde{x}_{i_2}^* \tilde{x}_{i_4}\} - \mathbf{E}\{\tilde{x}_{i_1} \tilde{x}_{i_4}\} \mathbf{E}\{\tilde{x}_{i_2}^* \tilde{x}_{i_3}^*\}, \end{aligned}$$

in which \mathbf{E} denotes the expectation operator and \tilde{x} equals $x - \mathbf{E}\{x\}$.

The permutation symmetries of \mathcal{C} can be generated by the permutation pair \mathbf{P}_1 and \mathbf{P}_2 , satisfying

$$(29) \quad c_{\mathbf{P}_1(i_1 i_2 i_3 i_4)} = c_{i_4 i_2 i_3 i_1} = c_{i_1 i_2 i_3 i_4},$$

$$(30) \quad c_{\mathbf{P}_2(i_1 i_2 i_3 i_4)} = c_{i_2 i_1 i_4 i_3} = c_{i_1 i_2 i_3 i_4}^*.$$

Equation (30) implies that $\mathbf{U}^{(1)} = \mathbf{U}^{(2)*}$ and that $\mathbf{U}^{(3)} = \mathbf{U}^{(4)*}$; on the other hand (29) implies that $\mathbf{U}^{(1)} = \mathbf{U}^{(4)}$. Hence the HOSVD of \mathcal{C} reveals the structure of (28),

$$\mathcal{C} = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U}^* \times_3 \mathbf{U}^* \times_4 \mathbf{U},$$

with \mathcal{S} satisfying (29) and (30) as well.

6. First-order perturbation analysis. In this section, the first-order derivatives of the n -mode singular values and vectors with respect to the elements of the original tensor are computed. The expressions can be used to analyze the numerical sensitivity and the statistical performance of HOSVD-based signal processing algorithms (e.g., [14]), especially for high signal-to-noise ratios: if the output of these algorithms can be considered as a function of the n -mode singular values and vectors, then the derivatives of the output could be computed via the chain rule.

An important application area is linked with the discussion on pseudodiagonality versus all-orthogonality (see section 3). In several applications, one has the a priori knowledge that a particular tensor can be made pseudoorthogonal by means of unitary transformations (e.g., cumulants of a stochastic vector with statistically independent components, after unitary transformation, like in a typical source separation context; cumulants of a stochastic vector with statistically independent components are pseudodiagonal). As pseudodiagonality is a special case of all-orthogonality, these unitary transformations coincide with the matrices of n -mode singular vectors. However, due to an imperfect knowledge of the tensor under consideration, caused by measurement errors, etc., the HOSVD of the estimated tensor will generally yield a core tensor that is all-orthogonal yet not exactly pseudodiagonal. To investigate such effects, first-order perturbation analysis results are required.

Due to the relationship between the matrix and tensor SVD, a perturbation analysis of the HOSVD can be based on existing results that have been developed in numerical linear algebra. Let \mathcal{A} be a real or complex higher-order tensor, the elements of which are analytic functions of a real parameter p . Like for matrices, an n -mode singular value and vector are analytic functions of p in the neighborhood of a nominal parameter p_0 where that singular value is isolated and nonzero. In case of a zero n -mode singular value the n -mode singular value function need not be differentiable, let alone analytic, due to the definition of an n -mode singular value as a norm being a nonnegative number: the differentiability of a smooth function that becomes negative in a neighborhood of p_0 may be lost by flipping the negative parts around the zero axis. A similar effect appears when several n -mode singular value functions cross each other, due to the ordering constraint on the HOSVD: the differentiability of the smooth functions that cross each other may be lost by combining the function parts in order of magnitude. However, one can prove that analyticity can be guaranteed by dropping the sign and ordering convention. With this goal we define, in analogy to the “unordered-unsigned singular value decomposition” (USVD) of [20] an *unordered-unsigned higher-order singular value decomposition* (UHOSVD).

THEOREM 5 (UHOSVD). *If the elements of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ are analytic functions of a real parameter p , then there exist real analytic functions $f_i^{(n)} : \mathbb{R} \rightarrow \mathbb{R}$ ($1 \leq i \leq I_n$) such that, for all $p \in \mathbb{R}$,*

$$\{\sigma_i^{(n)}(p) \mid (1 \leq i \leq I_n)\} = \{|f_i^{(n)}| \mid (1 \leq i \leq I_n)\}.$$

The functions $f_i^{(n)}$ are called unordered, unsigned n -mode singular value functions.

The analytic continuation of a set of n -mode singular vectors at a nominal parameter value, where they correspond to a multiple n -mode singular value, will be denoted by the term *preferred n -mode singular vectors*. They are given by the following theorem.

THEOREM 6 (preferred n -mode singular vectors). *Let the HOSVD of \mathcal{A} be given as in Theorem 2, the ordering of the n -mode singular vectors being of no importance.*

Let $\sigma^{(n)}$ be an n -mode singular value of multiplicity m and let an SVD of $\mathbf{A}_{(n)}$ be given by

$$(31) \quad \mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{\Sigma}^{(n)} \cdot \mathbf{V}^{(n)H} = \left(\mathbf{U}_1^{(n)} \mathbf{U}_2^{(n)} \right) \cdot \begin{pmatrix} \sigma^{(n)} \mathbf{I}_m & 0 \\ 0 & \mathbf{\Sigma}_2 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{V}_1^{(n)H} \\ \mathbf{V}_2^{(n)H} \end{pmatrix},$$

in which $\mathbf{U}^{(n)}$, $\mathbf{\Sigma}^{(n)}$, and $\mathbf{V}^{(n)}$ are defined in accordance to Property 12. In the partitioning, $\mathbf{U}_1^{(n)}$ and $\mathbf{V}_1^{(n)}$ correspond to the singular value $\sigma^{(n)}$.

Consider a first-order perturbation of \mathcal{A} as $\mathcal{A} + \epsilon \mathcal{B}$, $\epsilon \in \mathbb{R}$, and define the matrices $\mathbf{B}_{j_1 j_2}$, $j_1, j_2 = 1, 2$, as

$$\mathbf{B}_{j_1 j_2} \stackrel{\text{def}}{=} \mathbf{U}_{j_1}^{(n)H} \cdot \mathbf{B}_{(n)} \cdot \mathbf{V}_{j_2}^{(n)}.$$

There are two cases.

1. $\sigma^{(n)} \neq 0$: consider the eigenvalue problem

$$\frac{\mathbf{B}_{11} + \mathbf{B}_{11}^H}{2} = \mathbf{X} \cdot \mathbf{\Lambda} \cdot \mathbf{X}^H.$$

Then the m preferred n -mode singular vectors are given by the columns of $\mathbf{U}_1^{(n)} \cdot \mathbf{X}$. (The m preferred right singular vectors in (31) are given by the columns of $\mathbf{V}_1^{(n)} \cdot \mathbf{X}$.)

2. $\sigma^{(n)} = 0$: consider an SVD of \mathbf{B}_{11} :

$$\mathbf{B}_{11} = \mathbf{X} \cdot \mathbf{\Lambda} \cdot \mathbf{Y}^H.$$

Then the m preferred n -mode singular vectors are given by the columns of $\mathbf{U}_1^{(n)} \cdot \mathbf{X}$. (The m preferred right singular vectors in (31) are given by the columns of $\mathbf{V}_1^{(n)} \cdot \mathbf{Y}$.)

In analogy with [27, 20], we state now the following perturbation theorem.

THEOREM 7 (first-order perturbation of the HOSVD). *Consider a higher-order tensor $\mathcal{A}(\epsilon)$, the elements of which are analytic functions of a real parameter ϵ . Represent the first-order approximation of the Taylor series expansion of $\mathcal{A}(\epsilon)$ around $\epsilon = 0$ by $\mathcal{A} + \epsilon \mathcal{B}$. Let the HOSVD of \mathcal{A} be given as in Theorem 2. Let an SVD of $\mathbf{A}_{(n)}$ be given by*

$$\mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{\Sigma}^{(n)} \cdot \mathbf{V}^{(n)H},$$

in which $\mathbf{U}^{(n)}$, $\mathbf{\Sigma}^{(n)}$, and $\mathbf{V}^{(n)}$ are defined in accordance to Property 12. In case of multiple n -mode singular values, $\mathbf{U}^{(n)}$ and $\mathbf{V}^{(n)}$ contain preferred basis vectors. Define the matrix \mathbf{B} as

$$\mathbf{B} \stackrel{\text{def}}{=} \mathbf{U}^{(n)H} \cdot \mathbf{B}_{(n)} \cdot \mathbf{V}^{(n)}.$$

Then one has the following.

1. The slopes of the n -mode unordered-unsigned singular value functions are given by the diagonal elements of \mathbf{B} .
2. The first-order approximation of the Taylor series expansion of the n -mode singular vector functions is given by $\mathbf{U}^{(n)} + \epsilon \mathbf{U}^{(n)} \cdot \mathbf{\Omega}$, in which $\mathbf{\Omega}$ is a skew-Hermitian matrix defined by

$$\begin{aligned}
\omega_{ij} &\stackrel{\text{def}}{=} 0 \\
&\quad \text{if } i = j, \\
\omega_{ij} &\stackrel{\text{def}}{=} (\sigma_i^{(n)} b_{ij}^* + \sigma_j^{(n)} b_{ji}) / (\sigma_i^{(n)^2} - \sigma_j^{(n)^2}) \\
&\quad \text{if } \sigma_i^{(n)} \neq \sigma_j^{(n)}, \\
\omega_{ij} &\stackrel{\text{def}}{=} \lim_{\epsilon \rightarrow 0} \left[(\sigma_i^{(n)}(\epsilon) b_{ij}^* + \sigma_j^{(n)}(\epsilon) b_{ji}) / (\sigma_i^{(n)^2}(\epsilon) - \sigma_j^{(n)^2}(\epsilon)) \right] \\
&\quad \text{if } \sigma_i^{(n)}(\epsilon) = \sigma_j^{(n)}(\epsilon) \text{ for the isolated value } \epsilon = 0, \\
\omega_{ij} &\stackrel{\text{def}}{=} -\omega_{ji} \stackrel{\text{def}}{=} \text{arbitrary} \\
&\quad \text{if } \sigma_i^{(n)}(\epsilon) = \sigma_j^{(n)}(\epsilon) \text{ in a neighborhood of } \epsilon = 0 \text{ (} i \neq j \text{)}.
\end{aligned}$$

7. Related tensor decompositions. The SVD generalization we present in this paper is so clearly analogous to the SVD of matrices since the i -mode vectors play exactly the same role in the tensor decomposition as column and row vectors do for the matrix SVD. However, it is still possible to focus on other properties of the matrix SVD when looking for an equivalent tensor decomposition. It appears that for higher-order tensors different properties may raise different decompositions. Here we list three alternative approaches. The analogy with the matrix case is less striking than for the HOSVD. Depending on the context, one can choose the appropriate approach to address a certain problem.

7.1. Linear mapping between higher-order spaces. In the same way as there exists an isomorphic link between matrix algebra and the algebra of linear vector mappings, a higher-order tensor can be regarded as a formal representation of a linear mapping between a matrix and a vector space, a matrix and a matrix space, a matrix and a higher-order tensor space, etc. For example, assuming a basis in the space of N_1 th-order and N_2 th-order tensors, a linear transformation of $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \times \dots \times J_{N_1}}$ to $\mathcal{C} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_{N_2}}$ can be represented by $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_{N_2} \times J_1 \times \dots \times J_{N_1}}$, according to the element-wise equation

$$c_{i_1 i_2 \dots i_{N_2}} = \sum_{j_1 j_2 \dots j_{N_1}} a_{i_1 i_2 \dots i_{N_2} j_1 j_2 \dots j_{N_1}} b_{j_1 j_2 \dots j_{N_1}}$$

for a particular choice of N_1 summation indices.

Obviously, a tensorial SVD equivalent is the SVD of this linear mapping. We notice a link with the HOSVD: $\mathbf{U}^{(n)}$, $\mathbf{\Sigma}^{(n)}$, and $\mathbf{V}^{(n)}$ in (12) are the SVD components of \mathcal{A} , interpreted as a linear mapping from $\mathbb{C}^{I_{n+1} \times I_{n+2} \times \dots \times I_N \times I_1 \times I_2 \times \dots \times I_{n-1}}$ to \mathbb{C}^{I_n} , defined by summation over the indices $i_{n+1}, i_{n+2}, \dots, i_N, i_1, i_2, \dots, i_{n-1}$.

7.2. Optimal tensor diagonalization by unitary transformations. The fact that a generic higher-order tensor cannot be diagonalized by unitary transformations could be remedied by weakening the condition of the diagonality of the core tensor to a condition of “maximal diagonality.” Formally, given a higher-order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$, this new problem consists of the determination of unitary matrices $\mathbf{U}^{(1)} \in \mathbb{C}^{I_1 \times I_1}$, $\mathbf{U}^{(2)} \in \mathbb{C}^{I_2 \times I_2}$, \dots , $\mathbf{U}^{(N)} \in \mathbb{C}^{I_N \times I_N}$ such that the least-squares diagonality criterion $\sum_i |s_{ii\dots i}|^2$ is optimized, where the core tensor \mathcal{S} is still defined by (5).

For problems dealing with tensors that can theoretically be diagonalized (see also section 6), forcing the diagonal structure may be more robust (at the expense of a

higher computational cost) than the computation of the HOSVD, in which the deviation from diagonality is simply considered as a perturbation effect. For algorithms, we refer to [11, 16], exploring Jacobi-type procedures.

7.3. Canonical decomposition. A major difference between matrices and higher-order tensors is that the HOSVD does not provide precise information about rank-related issues. With this respect, the “canonical decomposition” (CANDECOMP), or “parallel factors” model (PARAFAC) may be more informative [10, 24, 1, 31, 39].

DEFINITION 11 (CANDECOMP). *A canonical decomposition or parallel factors decomposition of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is a decomposition of \mathcal{A} as a linear combination of a minimal number of rank-1 terms:*

$$(32) \quad \mathcal{A} = \sum_r^R \lambda_r U_r^{(1)} \circ U_r^{(2)} \circ \cdots \circ U_r^{(N)}.$$

This decomposition is important for applications, as the different rank-1 terms can often be related to different “mechanisms” that have contributed to the higher-order tensor; in addition, sufficiently mild uniqueness conditions enable the actual computation of these components (without imposing orthogonality constraints, as in the matrix case).

However, the practical determination of a tensor rank is a much harder problem than the determination of its n -ranks, since it involves the tensor as a global quantity, rather than as a collection of n -mode vectors. A fortiori, the computation of the CANDECOMP is much harder than the computation of the HOSVD, and many problems remain to be solved. By way of illustration, the determination of the maximal rank value over the set of $(I_1 \times I_2 \times \cdots \times I_N)$ -tensors is still an open problem in the literature (it is not bounded by $\min(I_1, I_2, \dots, I_N)$). We refer to [4] for a tutorial on the current state-of-the-art. In addition, [12] gives an overview of some partial results obtained for symmetric tensors.

8. Conclusion. In this paper the problem of generalizing the SVD of matrices to higher-order tensors is investigated. The point of departure is that the multilinear equivalent should address the so-called n -mode vectors in a similar way as the SVD does for column and row vectors. It is shown that this implies that the role of the matrix of singular values has to be assumed by a tensor with the property of all-orthogonality. The resulting decomposition, which is referred to as the HOSVD, is always possible for real or complex N th-order tensors, and when applied to matrices, it reduces—up to some trivial indeterminacies—to the matrix SVD. In the psychometric literature, the decomposition is known as the Tucker model.

From the fact that n -mode vectors and column (row) vectors play the same role in HOSVD, resp., SVD, it follows that the matrices of singular vectors can be computed from the sets of n -mode vectors in the same way as in the second-order case. In other words, the HOSVD of an N th-order tensor boils down to N matrix SVDs. As a consequence, several matrix properties have a very clear higher-order counterpart. Uniqueness, link with EVD, directions and values of extremal oriented energy, etc. are investigated and explicit expressions for a first-order perturbation analysis are presented.

The link between tensor symmetry and HOSVD is also investigated: if a higher-order tensor is transformed into itself or into its complex conjugate by a permutation of the indices, then this symmetry is reflected by the HOSVD components. It turns

out that the most adequate way to analyze this relation is the decomposition of the permutation in a sequence of permutation cycles. When every pair of indices is connected by some cycle, a property which is denoted as pair-wise symmetry, then the different matrices of higher-order singular vectors are either equal or complex conjugated. In analogy to the HOEVD of a real symmetric or Hermitean matrix the HOEVD of a pair-wise symmetric higher-order tensor is defined. Its computation, for a pair-wise symmetric N th-order $(I \times I \times \cdots \times I)$ -tensor, reduces to the SVD of an $(I \times I^{N-1})$ -matrix unfolding. It remains an open problem if and how the high symmetry of this matrix can be exploited to speed up the calculation of its SVD.

The link with the SVD of matrices makes the HOSVD the proper tool for the analysis of n -mode vector space properties. However, the fact that the core tensor is in general a full tensor, instead of being pseudodiagonal, results in two main differences with the matrix case. First, the HOSVD of a given tensor allows us to estimate its n -ranks, but these values give only a rough lower-bound on the rank; more in general, the HOSVD does not allow for an interpretation in terms of minimal expansions in rank-1 terms. This shortcoming is inherent in the structure of multilinear algebra; with respect to rank-related issues, the CANDECOMP may be informative. Second, discarding the smallest n -mode singular values does not necessarily lead to the best possible approximation with reduced n -rank values ($1 \leq n \leq N$). However, if there is a large gap between the n -mode singular values, then the approximation can still be qualified as accurate; if necessary, it can be used as a good starting value for an additional optimization algorithm. Finally, we remark that in some applications one knows a priori that the core tensor is pseudodiagonal. As pseudodiagonality is a special case of all-orthogonality, analysis procedures may then be based on the HOSVD, up to perturbation effects. If a higher accuracy is mandatory, it might be preferable to determine unitary transformations that explicitly make the higher-order tensor as diagonal as possible, rather than all-orthogonal.

REFERENCES

- [1] C.J. APPELLOF AND E.R. DAVIDSON, *Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents*, Analytical Chemistry, 53 (1981), pp. 2053–2056.
- [2] R.E. BELLMAN, *Matrix Analysis*, McGraw-Hill, NY, 1978.
- [3] R.M. BOWEN AND C.-C. WANG, *Introduction to Vectors and Tensors. Linear and Multilinear Algebra*, Plenum Press, NY, 1980.
- [4] R. BRO, *PARAFAC. Tutorial and applications*, Chemom. Intell. Lab. Syst., Special Issue 2nd Internet Conf. in Chemometrics (INCINC'96), 38 (1997), pp. 149–171.
- [5] J.-F. CARDOSO, *Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Albuquerque, NM, 1990, pp. 2655–2658.
- [6] J.-F. CARDOSO, *Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Toronto, Canada, 1991, pp. 3109–3112.
- [7] J.-F. CARDOSO, *Fourth-order cumulant structure forcing. Application to blind array processing*, in Proceedings of the IEEE Signal Processing Workshop on Statistical Signal and Array Processing, 1992, pp. 136–139.
- [8] J.-F. CARDOSO AND P. COMON, *Tensor-based independent component analysis*, in Signal Processing V: Theories and Applications, Proc. EUSIPCO-90, L. Torres, ed., Elsevier, Amsterdam, 1990, pp. 673–676.
- [9] J.-F. CARDOSO AND P. COMON, *Independent component analysis, a survey of some algebraic methods*, in Proceedings of the IEEE International Symposium on Circuits and Systems, Atlanta, GA, 1996, pp. 93–96.

- [10] J.D. CARROLL AND S. PRUZANSKY, *The CANDECOMP-CANDELINC family of models and methods for multidimensional data analysis*, in Research Methods for Multimode Data Analysis, H.G. Law, C.W. Snyder, J.A. Hattie, and R.P. McDonald, eds., Praeger, NY, 1984, pp. 372–402.
- [11] P. COMON, *Independent component analysis, a new concept?*, Signal Process., Special Issue on Higher Order Statistics, 36 (1994), pp. 287–314.
- [12] P. COMON AND B. MOURRAIN, *Decomposition of quantics in sums of powers of linear forms*, Signal Process., Special Issue on Higher Order Statistics, 53 (1996), pp. 93–108.
- [13] R. COPPI AND S. BOLASCO, EDS., *Multiway Data Analysis*, Elsevier, Amsterdam, 1989.
- [14] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *Blind source separation by higher-order singular value decomposition*, in Signal Processing VII: Theories and Applications, Proc. EUSIPCO-94, Edinburgh, UK, 1994, pp. 175–178.
- [15] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *Independent component analysis based on higher-order statistics only*, in Proceedings of the IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Corfu, Greece, 1996, pp. 356–359.
- [16] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *Blind source separation by simultaneous third-order tensor diagonalization*, in Signal Processing VIII: Theories and Applications, Proc. EUSIPCO-96, Trieste, Italy, 1996, pp. 2089–2092.
- [17] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *Dimensionality reduction in higher-order-only ICA*, in Proceedings of the IEEE Signal Processing Workshop on HOS, Banff, Alberta, Canada, 1997, pp. 316–320.
- [18] L. DE LATHAUWER, P. COMON, B. DE MOOR, AND J. VANDEWALLE, *Higher-order power method—Application in independent component analysis*, in Proceedings of the International Symposium on Nonlinear Theory and Its Applications, Las Vegas, UT, 1995, pp. 91–96.
- [19] L. DE LATHAUWER, *Signal Processing Based on Multilinear Algebra*, Ph.D. thesis, K.U. Leuven, E.E. Dept.-ESAT, Belgium, 1997.
- [20] B. DE MOOR AND S. BOYD, *Analytic Properties of Singular Values and Vectors*, Tech. Report 1989-28, SISTA, E.E. Dept.-ESAT, K.U. Leuven, 1989.
- [21] E.F. DEPRETTERE, ED., *SVD and Signal Processing. Algorithms, Applications and Architectures*, North-Holland, Amsterdam, 1988.
- [22] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [23] W. GREUB, *Multilinear Algebra*, Springer-Verlag, NY, 1978.
- [24] R.A. HARSHMAN AND M.E. LUNDY, *The PARAFAC model for three-way factor analysis and multidimensional scaling*, in Research Methods for Multimode Data Analysis, H.G. Law, C.W. Snyder, J.A. Hattie, and R.P. McDonald, eds., Praeger, NY, 1984, pp. 122–215.
- [25] J. HASTAD, *Tensor rank is NP-complete*, J. Algorithms, 11 (1990), pp. 644–654.
- [26] D.C. KAY, *Theory and Problems of Tensor Calculus*, McGraw-Hill, New York, 1988.
- [27] T. KATO, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1982.
- [28] P.M. KROONENBERG, *Three-Mode Principal Component Analysis*, DSWO Press, Leiden, The Netherlands, 1983.
- [29] J.B. KRUSKAL, *Rank, decomposition, and uniqueness for 3-way and N-way arrays*, in Multiway Data Analysis, R. Coppi and S. Bolasco, eds., North-Holland, Amsterdam, 1989, pp. 7–18.
- [30] J.L. LACOUME, M. GAETA, AND P.O. AMBLARD, *From order 2 to HOS: New tools and applications*, in Signal Processing VI: Theories and Applications, Proc. EUSIPCO-92, J. Vandewalle and R. Boite, eds., Elsevier, Amsterdam, 1992, pp. 91–98.
- [31] S.E. LEURGANS, R.T. ROSS, AND R.B. ABEL, *A decomposition for three-way arrays*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1064–1083.
- [32] M. MARCUS, *Finite Dimensional Multilinear Algebra*, Dekker, New York, 1975.
- [33] P. McCULLAGH, *Tensor Methods in Statistics*, Chapman and Hall, London, UK, 1987.
- [34] J.M. MENDEL, *Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications*, Proc. IEEE, 79 (1991), pp. 278–305.
- [35] M. MOONEN AND B. DE MOOR, EDS., *SVD and Signal Processing, III. Algorithms, Applications and Architectures*, Elsevier, Amsterdam, 1995.
- [36] D.G. NATHCOTT, *Multilinear Algebra*, Cambridge University Press, Cambridge, UK, 1984.
- [37] C.L. NIKIAS, *Higher order spectra in signal processing*, in Signal Processing V: Theories and Applications, Proc. EUSIPCO-90, L. Torres, ed., Elsevier, Amsterdam, 1990, pp. 35–41.
- [38] C.L. NIKIAS AND J.M. MENDEL, *Signal processing with higher-order spectra*, IEEE Signal Process. Mag., 10 (1993), pp. 10–37.
- [39] P. PAATERO, *A weighted non-negative least squares algorithm for three-way “PARAFAC” factor*

- analysis*, Chemom. Intell. Lab. Syst., Special Issue, 2nd Internet Conf. in Chemometrics (INCINC'96), 38 (1997), pp. 223–242.
- [40] P.A. REGALIA AND S.K. MITRA, *Kronecker products, unitary matrices and signal processing applications*, SIAM Rev., 31 (1989), pp. 586–613.
 - [41] A. SWAMI AND G. GIANNAKIS, *Bibliography on HOS*, Signal Process., 60 (1997), pp. 65–126.
 - [42] L.R. TUCKER, *The extension of factor analysis to three-dimensional matrices*, in Contributions to Mathematical Psychology, H. Gulliksen and N. Frederiksen, eds., Holt, Rinehart & Winston, New York, 1964, pp. 109–127.
 - [43] L.R. TUCKER, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311.
 - [44] R. VACCARO, ED., *SVD and Signal Processing, II. Algorithms, Applications and Architectures*, Elsevier, Amsterdam, 1991.
 - [45] J. VANDEWALLE AND D. CALLAERTS, *Singular value decomposition: A powerful concept and tool in signal processing*, in Mathematics in Signal Processing II, Clarendon Press, Oxford, 1990, pp. 539–559.
 - [46] J. VANDEWALLE AND B. DE MOOR, *On the use of the singular value decomposition in identification and signal processing*, in NATO ASI Series: Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms, F70 (1991), pp. 321–360.
 - [47] P. VAN OVERSCHEE AND B. DE MOOR, *Subspace Identification for Linear Systems. Theory, Implementation, Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.