Tensor similarity in chemometrics

Frederik Van Eeghem, Lieven De Lathauwer

May 2020

Author information

- **Frederik Van Eeghem**
  Group of Science, Engineering and Technology, KU Leuven Kulak
  E. Sabbelaan 53, 8500 Kortrijk, Belgium
  and
  Department of Electrical Engineering (ESAT), KU Leuven
  Kasteelpark Arenberg 10, B-3001 Leuven, Belgium.
  frederik.vaneeghem@kuleuven.be

- **Lieven De Lathauwer**
  Group of Science, Engineering and Technology, KU Leuven Kulak
  E. Sabbelaan 53, 8500 Kortrijk, Belgium
  and
  Department of Electrical Engineering (ESAT), KU Leuven
  Kasteelpark Arenberg 10, B-3001 Leuven, Belgium.
  lieven.delathauwer@kuleuven.be

Abstract

Tensors and their decompositions are widespread in various domains such as signal processing, biomedical data processing and chemometrics. Because of the uniqueness properties of tensor decompositions, tensors have become popular tools to resolve mixtures into their pure components. In some applications however, one wishes to compare two tensors rather than to decompose them. More specifically, one wishes to assess whether two tensors consist of the same rank-1 terms or even block terms in two or more dimensions. Recent advances in tensor analysis have shown that this is possible without having to compute any tensor decompositions, which results in faster and numerically stable algorithms. This chapter aims to introduce this new approach in chemometrics. This is done by explaining the initial insights in tensor similarity, followed by more practical step-by-step approaches which are illustrated with a chemometrics example.

Keywords: tensor, tensors, similarity, CPD, BTD, decomposition, PARAFAC, CANDECOMP, canonical polyadic decomposition, block term decomposition, comparison, tensor comparison, fluorescence, tensorization, tensor similarity.

1 Introduction

Chemometrics has a rich history of extracting information from chemical systems using various statistical or mathematical techniques. Since chemical data is often multi-way by nature, tensors are well-known tools within the community (see [1] and references therein). One popular technique is blind source separation, which tries to resolve spectra of mixtures into the spectra of the pure components. There is a whole body of literature on blind source separation, which is also encountered frequently in domains as biomedical data analysis, digital communications, seismic monitoring and more [2–4].

Several algorithms for blind source separation exist, including tensor-based techniques, which are the focus of this chapter. Tensor-based blind source separation relies on tensor decompositions to extract the underlying pure components. Again, this approach is strongly embedded within chemometrics community [5], which featured the CP decomposition (also known as PARAFAC) two decades ago.

Recently, new tensor techniques have been developed that go beyond decompositions [6–8]. It has been shown that the decompositions of two tensors can be compared to each other without having to compute the full decompositions. For instance, one can show that two tensors consist of the same rank-1 terms, possibly differently scaled, without having to compute these rank-1 terms. These new insights and algorithms open up opportunities in all fields where tensor decompositions are used. The goal of this chapter is to introduce these new developments in the field of tensors to a broader audience in chemometrics.

The remainder of this introduction introduces the mathematical notations that are used throughout this chapter. Next, Section 2 briefly introduces tensors and their main properties. In Section 3, the relevant tensor decompositions

are discussed. Next, tensor similarity is introduced and discussed at length in Section 4, together with some initial insights. Finally, Section 5 illustrates some of the concepts with applications in chemometrics.

Notations Scalars are represented by italic lowercase letters (e.g., a), column vectors by bold lowercase letters (e.g., $\mathbf{a}$), matrices by bold uppercase letters (e.g., $\mathbf{A}$) and tensors by uppercase calligraphic letters (e.g., $\mathcal{A}$). Subscripts are used to denote subsets of vectors, matrices and tensors. For instance, the rth column vector of the matrix $\mathbf{A}$ is denoted by $\mathbf{a}_r$. Analogously, t_{ijk} represents the entry with indices (i, j, k) of the tensor $\mathcal{T}$. Slices of a tensor will be denoted using Matlab-style notation. For instance, the kth frontal slice of a tensor $\mathcal{T}$ is denoted by $\mathbf{T}_{::k}$. Column-wise vectorization of a matrix is denoted by $\mathrm{Vec}\,(\mathbf{A})$. Consider following matrix for instance:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix},$$

then $\mathrm{Vec}\,(\mathbf{A}) = [1, 4, 2, 5, 3, 6]^{\mathrm{T}}$.

Consider the matrix $\mathbf{A} \in \mathbb{C}^{I \times J}$ given by

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1J} \\ a_{21} & a_{22} & \dots & a_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1} & a_{I2} & \dots & a_{IJ} \end{bmatrix}.$$

The Kronecker product of $\mathbf{A} \in \mathbb{C}^{I \times J}$ and $\mathbf{B} \in \mathbb{C}^{K \times L}$ is then denoted by $\mathbf{A} \otimes \mathbf{B}$ and given by:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2J}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \dots & a_{IJ}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{IK \times JL}$$

The Khatri–Rao product of partitioned matrices $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_N]$ and $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_N]$ is given by

$$\mathbf{A} \odot \mathbf{B} = [\mathbf{A}_1 \otimes \mathbf{B}_1, \dots, \mathbf{A}_N \otimes \mathbf{B}_N].$$

Unless specified otherwise, matrices in a Khatri–Rao product are assumed to be partitioned columnwise, i.e. $\mathbf{A} \odot \mathbf{B} = [\mathbf{a}_1 \otimes \mathbf{b}_1, \dots, \mathbf{a}_N \otimes \mathbf{b}_N]$. The outer product is denoted by $\circ$ and is defined as follows: consider the outer product of N vectors $\mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \dots \circ \mathbf{u}^{(N)}$, this yields an Nth-order tensor $\mathcal{A}$ of which each entry is given by

$$a_{i_1 i_2 \dots i_N} = u^{(1)}_{i_1} u^{(2)}_{i_2} \dots u^{(N)}_{i_N}.$$

The transpose of a matrix is denoted by $\cdot^{\mathrm{T}}$, the Hermitian transpose by $\cdot^{\mathrm{H}}$ and the column space of a matrix is denoted by $\mathrm{span}\,\{\mathbf{A}\}$. The tensor-matrix

product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ and a matrix $\mathbf{M} \in \mathbb{C}^{J_k \times I_k}$ in the kth-mode yields a tensor $\mathcal{B} \in \mathbb{C}^{I_1 \times \ldots \times I_{k-1} \times J_k \times I_{k+1} \times \ldots \times I_N}$ is denoted by $\mathcal{B} = \mathcal{A} \cdot_k \mathbf{M}$ and is given by

$$b_{i_1,i_2,\ldots,i_{k-1},j_k,i_{k+1},\ldots,i_N} = \sum_{i_k=1}^{I_k} a_{i_1,i_2,\ldots,i_k,\ldots,i_N} m_{j,i_k}$$

The intersection of two sets A and B is denoted by $A \cap B$. The notation $\mathbf{a} \perp \mathbf{b}$ is used for two vectors $\mathbf{a} \in C^K$ and $\mathbf{b} \in C^K$ that are perpendicular to each other.

2 Tensors and matrices

Tensors are generalizations of vectors and matrices and are sometimes denoted by multi-way arrays. There are two main reasons why tensors have gained popularity in the past decades.

First, they allow a natural representation of multi-way data. For instance, the temperature in a room monitored over time is naturally represented in a four-dimensional tensor, with three spatial dimensions and one temporal dimension. A chemical example of a third-order tensor could be fluorescence emission spectra measured at several excitation wavelengths for several samples. Another example is any kind of spectrum measured chromatographically for several samples [5]. This multi-way representation simplifies further computations and interpretation of the data. If some specific computations require matrices rather than tensors, the tensor entries can be rearranged in a matrix. This process is called matricization or unfolding of a tensor. Unfolding a third-order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$ to a matrix can be done in various ways, depending on which modes are combined in one matrix dimensions. Here, following definitions are used:

$$\begin{aligned}
\left(\mathbf{X}_{[1,2;3]}\right)_{j+(i-1)J,k} &= x_{ijk}, \\
\left(\mathbf{X}_{[2,3;1]}\right)_{k+(j-1)K,i} &= x_{ijk}, \\
\left(\mathbf{X}_{[1,3;2]}\right)_{k+(i-1)K,j} &= x_{ijk}.
\end{aligned}$$

The second strength of tensors lies in their decompositions, which are unique under mild conditions. This last property does not exist in matrix decompositions, which can be illustrated by considering a 3×3 matrix and its rank-2 decomposition:

$$\begin{bmatrix} 21 & 31 & 6 \\ 27 & 41 & 9 \\ 33 & 51 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 5 & 7 & 1 \end{bmatrix}. \tag{1}$$

In such a matrix decomposition, it is possible to post- and premultiply the factors with a regular matrix and its inverse without changing the outcome,

e.g.:

$$\begin{bmatrix} 21 & 31 & 6 \\ 27 & 41 & 9 \\ 33 & 51 & 12 \end{bmatrix} = \left(\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \right) \left(\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 3 & 2 \\ 5 & 7 & 1 \end{bmatrix} \right),$$
$$= \begin{bmatrix} 5 & 8 \\ 7 & 10 \\ 9 & 12 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 2 & 2 & -0.5 \end{bmatrix}. \quad (2)$$

The example shows two different rank-2 decompositions (1) and (2) of the same matrix, which immediately illustrates that matrix decompositions are not unique. The only way to obtain unique matrix decompositions is by imposing additional constraints, such as orthogonality in the singular value decomposition. The next section will show that this is in stark contrast with tensor decompositions, which do not need additional constraints and are unique under mild conditions.

3 Tensor decompositions

Various tensor decompositions exist, which are used for different objectives such as compression, blind signal separation and classification [9–11]. This chapter focuses on the multilinear singular value decomposition (MLSVD), the canonical polyadic decomposition (CPD) and several variations of block term decompositions (BTD), which will be used further in the chapter.

3.1 Multilinear singular value decomposition

The multilinear singular value decomposition (MLSVD) of a third-order tensor is a multilinear generalization of the well-known singular value decomposition of a matrix [12–15]. Other terms sometimes used for this decompositions are the higher-order singular value decomposition or the Tucker decomposition. Note that the latter is often used when no additional constraints are imposed on the core tensor or factor matrices of the decomposition.

Consider a third-order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$. The MLSVD of $\mathcal{A}$ decomposes the tensor as

$$\mathcal{A} = \mathcal{W} \cdot_1 \mathbf{U}^{(1)} \cdot_2 \mathbf{U}^{(2)} \cdot_3 \mathbf{U}^{(3)}, \quad (3)$$

in which the factor matrices $\mathbf{U}^{(1)} \in \mathbb{C}^{I_1 \times J_1}$, $\mathbf{U}^{(2)} \in \mathbb{C}^{I_2 \times J_2}$ and $\mathbf{U}^{(3)} \in \mathbb{C}^{I_3 \times J_3}$ are columnwise orthogonal. The core tensor $\mathcal{W} \in \mathbb{C}^{J_1 \times J_2 \times J_3}$ is all-orthogonal and ordered [16]. All-orthogonality means that the different matrix slices in each mode of $\mathcal{W}$ are orthogonal whereas ordered means that the slices in each mode of $\mathcal{W}$ are ordered by descending Frobenius norm. The MLSVD is illustrated graphically in Figure 1.

The MLSVD is closely linked with the concept of mode-n rank and multilinear rank. The mode-n rank of a tensor $\mathcal{A}$ is the dimension spanned by its mode-n vectors, or equivalently, the rank of its mode-n matrix unfolding

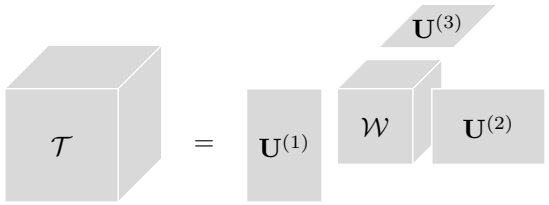


Figure 1: Graphical illustration of a MLSVD.

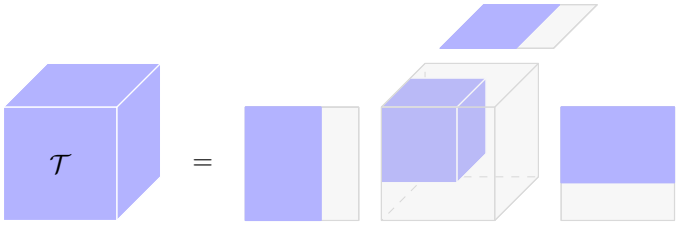


Figure 2: Illustration of the MLSVD of a tensor with low multilinear rank (assuming the blue core tensor has full multilinear rank). The blue parts of the factor matrices denote the vectors that span the mode-n spaces of $\mathcal{T}$.

$\mathbf{A}_{[k,m;n]}$. The multilinear rank of a third-order tensor $\mathcal{A}$ then equals (I, J, K) if its mode-1, mode-2 and mode-3 rank equal I, J and K, respectively [16].

In a matrix SVD, where $\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}$, the factor matrices $\mathbf{U}$ and $\mathbf{V}$ contain orthogonal bases for the column space and row space, respectively. The MLSVD generalizes this property to tensors, where the factor matrices $\mathbf{U}^{(n)} \in \mathbb{C}^{I_n \times J_n}$ contain an orthonormal basis for the space spanned by the mode-n vectors of $\mathcal{A}$. More specifically, if $\mathcal{A}$ has multilinear rank (L_1, L_2, L_3) smaller than (J_1, J_2, J_3), then the ordered property of the core tensor $\mathcal{W}$ ensures that the first L_n column vectors of $\mathbf{U}^{(n)}$ form a basis for the space spanned by the mode-n vectors of $\mathcal{A}$, analogous to the SVD in the matrix case. This property is illustrated graphically in Figure 2.

If the tensor $\widetilde{\mathcal{W}} \in \mathbb{C}^{J_1 \times J_2 \times I_3}$ is defined as $\widetilde{\mathcal{W}} = \mathcal{W} \cdot_3 \mathbf{U}^{(3)}$, then (3) can be written as

$$\mathcal{A} = \widetilde{\mathcal{W}} \cdot_1 \mathbf{U}^{(1)} \cdot_2 \mathbf{U}^{(2)}. \tag{4}$$

Ignoring the structure of $\widetilde{\mathcal{W}}$, equation (4) represents a normalized Tucker-2 decomposition of $\mathcal{A}$ [15].

3.2 Canonical polyadic decomposition

The canonical polyadic decomposition or CANDECOMP/PARAFAC is defined as the decomposition of a tensor in a minimal number of rank-1 terms. To understand this definition, the notions of a rank-1 term and a "minimal number" need to be explained. A rank-1 tensor $\mathcal{A}$ of order N is a tensor that can be

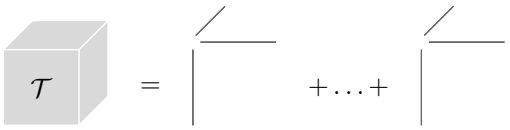


Figure 3: Graphical illustration of a CPD.

written as the outer product of N nonzero vectors. Mathematically, this can be written as

$$\mathcal{A} = \mathbf{u}^{(1)} \circ \cdots \circ \mathbf{u}^{(N)},$$

with $\mathbf{u}^{(n)}$ nonzero vectors. Each entry $a_{i_1 i_2 \cdots i_N}$ of the tensor $\mathcal{A}$ is thus given by $a_{i_1 i_2 \cdots i_N} = u^{(1)}_{i_1} u^{(2)}_{i_2} \cdots u^{(N)}_{i_N}$.

The rank of a tensor is then defined as the minimum number of rank-1 terms that make up the tensor in a linear combination. Consequently, the CPD of a rank-R tensor decomposes the tensor as a linear combination of R rank-1 tensors. Mathematically, we can write the CPD of a tensor $\mathcal{T} \in \mathbb{C}^{I_1 \times I_2 \times I_N}$ as

$$\mathcal{T} = \sum_{r=1}^{R} \lambda_r \mathbf{u}^{(1)}_r \circ \cdots \circ \mathbf{u}^{(N)}_r, \tag{5}$$

in which λ_r is the scaling factor of the rth term. The factor vectors $\mathbf{u}^{(n)}_r \in \mathbb{C}^{I_n}$ are often concatenated in factor matrices $\mathbf{U}^{(n)} = \left[\mathbf{u}^{(n)}_1, \ldots, \mathbf{u}^{(n)}_R\right] \in \mathbb{C}^{I_n \times R}$. Following the notation used in [9], (5) can be concisely written as

$$\mathcal{T} = \left[\!\left[\boldsymbol{\lambda}; \mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(N)} \right]\!\right],$$

with $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_R] \in \mathbb{C}^R$. A graphical illustration of a CPD is given in Figure 3.

One important property of the CPD is its essential uniqueness under mild conditions. Essential uniqueness denotes that the decomposition is unique up to two trivial indeterminacies. First, the different rank-1 terms can be arbitrarily permuted. Second, each vector within a rank-1 term can be scaled as long as the other vectors of the rank-1 term are counterscaled. The mathematical conditions ensuring this essential uniqueness for CPDs have been extensively studied in the literature, see e.g. [10, 17–19]. Various algorithms to compute (constrained) CPDs have been developed as well, see e.g. [10, 20, 21] and references therein.

Fluorescence spectrography example: Consider five samples containing different amounts of the same three unknown pure analytes, which we wish to retrieve. Denote the excitation-emission matrix of the rth sample as $\mathbf{E}_r \in \mathbb{C}^{K \times N}$, with K excitation wavelengths and N wavelengths at which the emitted light is measured. Let us now stack these five matrices to obtain a tensor $\mathcal{E} \in \mathbb{C}^{K \times N \times 5}$. In [5, 22, 23], the theoretical multilinearity of this tensor was shown for sample

with low absorbance and without interactions between chromophores. The paper showed that this tensor admits a rank-R CPD with R the number of pure analytes. The first factor matrix of this CPD contains excitation information, the second one emission information and the third contains the concentrations of the pure analytes in each mixture. To apply this theory to the data at hand, decompose the tensor $\mathcal{E}$ in three rank-1 terms as

$$\mathcal{E} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!] = \sum_{r=1}^{3} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r,$$

with $\mathbf{A} \in \mathbb{C}^{K\times 3}, \mathbf{B} \in \mathbb{C}^{N\times 3}$ and $\mathbf{C} \in \mathbb{C}^{5\times 3}$. Any scaling of the different terms was incorporated in the factor matrix $\mathbf{C}$. Because the CPD is essentially unique under mild conditions, which we assume are fulfilled in this case, we know that the factor matrices are unique up to scaling and permutation. Consequently, we can reconstruct the excitation-emission matrices of the pure analytes by computing the outer product of the columns of the first two factor matrices ($\mathbf{A}$ and $\mathbf{B}$) as follows:

$$\begin{aligned}
&\text{Excitation-emission matrix of Analyte 1} = \mathbf{a}_1 \circ \mathbf{b}_1 \coloneqq \mathbf{a}_1 {\mathbf{b}_1}^{\mathrm{T}},\\
&\text{Excitation-emission matrix of Analyte 2} = \mathbf{a}_2 \circ \mathbf{b}_2 \coloneqq \mathbf{a}_2 {\mathbf{b}_2}^{\mathrm{T}},\\
&\text{Excitation-emission matrix of Analyte 3} = \mathbf{a}_3 \circ \mathbf{b}_3 \coloneqq \mathbf{a}_3 {\mathbf{b}_3}^{\mathrm{T}}.
\end{aligned}$$

By comparing these reconstructed excitation-emission matrices with the known spectra of pure analytes, the components of the mixtures can be identified. Note that due to the scaling ambiguity, the excitation-emission matrices could be scaled differently, so normalization of the matrices may be required before comparing them with known spectra of pure analytes.

3.3 Block term decompositions

In several applications, the rank-1 assumption on each term in a CPD may be too strict to model reality. Because of this, a series of more general decompositions was developed in [16], which are denoted by block-term decompositions (BTD). In this section, a few relevant block-term decompositions are briefly introduced and discussed.

3.4 Block term decomposition in rank-$(L, L, 1)$ terms

A BTD in multilinear rank-$(L, L, 1)$ terms is a generalization of a CPD that allows low-rank contributions in the first two modes of each term [16]. This allows for more flexible modeling in applications where the rank-1 assumption is too strict. A graphical representation of a BTD in multilinear rank-$(L, L, 1)$ terms is given in Figure 4. Mathematically, the BTD of a third-order tensor $\mathcal{T} \in \mathbb{C}^{I\times J\times K}$ in rank-$(L, L, 1)$ terms can be written as

$$\mathcal{T} = \sum_{r=1}^{R} (\mathbf{A}_r \mathbf{B}_r^{\mathrm{T}}) \circ \mathbf{c}_r,$$



Figure 4: Graphical illustration of a BTD in multilinear rank-$(L, L, 1)$ terms.

in which $\mathbf{A}_r \in \mathbb{C}^{I \times L}$ and $\mathbf{B}_r \in \mathbb{C}^{J \times L}$ have full column rank and $\mathbf{c}_r \in \mathbb{C}^K$. Essential uniqueness conditions for this decomposition can be found in [16, 24–26]. Just as in the CPD case, the factor matrices are unique under mild conditions up to trivial indeterminacies. More specifically, the terms can again be arbitrarily permuted, but the scaling indeterminacy is slightly more complex than before. Within each term, one can postmultiply $\mathbf{A}_r$ by any nonsingular $\mathbf{G}_r \in \mathbb{C}^{L \times L}$ as long as $\mathbf{B}_r$ is postmultiplied by $\left(\mathbf{G}_r^{-1}\right)^{\mathrm{T}}$. Moreover, the factors within each multilinear rank-$(L, L, 1)$ term can be scaled and counterscaled as long as their product remains the same. Various algorithms for the computation of this type of decomposition have been developed and can be found in [20, 21, 26–28].

Fluorescence spectrography example: Let us return to the same problem statement as in the CPD example, but now only one sample is available containing different amounts of three unknown pure analytes, which we want to retrieve. The introduction showed that matrix decompositions are not unique, so tensorization is used to transform the excitation-emission matrix $\mathbf{E}$ of the sample to a tensor [29, 30]. Many tensorization techniques exist, and choosing the right one is an art on its own. In the context of blind source separation, the suitable tensorization technique depends on the assumptions on the source signals [30]. For example, in ICA the sources are assumed to be statistically independent. It has been shown that this independence assumption can be exploited by computing higher-order statistics of the output signals. If the sources are indeed independent, this yields a tensor that admits a CPD of which the factor matrices contain the mixing information [2]. Analogously, if the sources are exponential polynomials, it has been shown that Hankelization leads to a tensor that allows unmixing [24]. In this fluorescence spectroscopy example, the sources are the emission spectra of the pure analytes. These sources are not statistically independent, but they can be approximated using rational functions, which in turn can be exploited using a tensorization technique called Löwnerization [31]. A Löwner matrix $\mathbf{L}$ has entries

$$l_{i,j} = \frac{f(x_i) - f(y_i)}{x_i - y_i},$$

in which $f(t)$ is a function sampled on N different points t_i and this point set $T = t_1, t_2, \ldots, t_N$ is partitioned in two point sets $X = x_1, x_2, \ldots, x_\alpha$ and $Y = y_1, y_2, \ldots, y_{N-\alpha}$. Note that if N is even and $\alpha = N/2$, then $\mathbf{L}$ is square. The parameter α is typically chosen to be $\lceil N/2 \rceil$. Further background information

and details on Löwner matrices can be found in [31] and references therein. To apply Löwnerization to the chemometrics example, each data row along the emission mode is transformed to a Löwner matrix as detailed in [31]. By stacking these matrices, a tensor $\mathcal{E}$ is obtained. It has been shown in [31] that this tensor admits a BTD in three rank-$(L, L, 1)$-terms, with L depending on the spectra of the pure analytes. For a mixture of Tyrosine, Phenylalanine and Tryptophan, a value of $L = 2$ was used in [31]. The number of components R (or more precisely, a lower bound on R) can be estimated as the number of dominant singular values of $\mathbf{E}_{[1,2;3]}$. In this example, $R = 3$ by construction. Similarly, (a lower bound on) the value of RL can be estimated as the number of dominant singular values of $\mathbf{E}_{[2,3;1]}$. Indeed, in this example $\mathbf{E}_{[2,3;1]}$ has 6 dominant singular values. Dividing by R then yields an estimate $L = 2$.

Computing a BTD in three rank-$(2, 2, 1)$-terms of $\mathcal{E}$ yields one block-term per pure analyte:

$$\mathcal{E} = \sum_{r=1}^{3} \left(\widehat{\mathbf{A}}_r \widehat{\mathbf{B}}_r^{\mathrm{T}} \right) \circ \widehat{\mathbf{c}}_r \in \mathbb{C}^{101 \times 100 \times 61},$$

in which the factor matrix $\widehat{\mathbf{C}}$ contains the estimated excitation spectra. If this decomposition is unique up to trivial indeterminacies, the mixture is effectively separated. Several approaches to retrieve the original emission spectra are given in [31]. The most straightforward technique is to invert the factor matrix $\widehat{\mathbf{C}}$ and compute its matrix product with the excitation-emission matrix $\mathbf{E}$:

$$\widehat{\mathbf{S}} = \widehat{\mathbf{C}}^{\dagger} \mathbf{E}$$

The matrix $\widehat{\mathbf{S}}$ then contains the emission spectra of the three underlying pure analytes. Note that this approach is only valid for overdetermined mixtures with a full column rank mixing matrix, see [31] for further details and alternatives.

3.5 Block term decomposition in rank-$(L, L, \cdot)$ terms

A BTD in multilinear rank-$(L, L, \cdot)$ is similar to the previous block-term decomposition in the first two modes but has no factor matrix in the third mode. Mathematically, the BTD of a third-order tensor $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ in rank-$(L, L, \cdot)$ terms can be written as

$$\mathcal{T} = \sum_{r=1}^{R} \mathcal{C}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r,$$

in which $\mathbf{A}_r \in \mathbb{C}^{I \times L}$ and $\mathbf{B}_r \in \mathbb{C}^{J \times L}$ have full column rank and $\mathcal{C}_r \in \mathbb{C}^{L \times L \times K}$ has mode-1 rank and mode-2 rank equal to L [16]. A graphical representation of a BTD in multilinear rank-$(L, L, \cdot)$ terms is given in Figure 5. Essential uniqueness conditions for this decomposition can be found in [16, 32]. The trivial ambiguities in the decomposition are the following: the different terms can be arbitrarily permuted and within each term, one can postmultiply $\mathbf{A}_r$ by any nonsingular $\mathbf{G}_r \in \mathbb{C}^{L \times L}$ and $\mathbf{B}_r$ by any nonsingular $\mathbf{H}_r \in \mathbb{C}^{L \times L}$ as

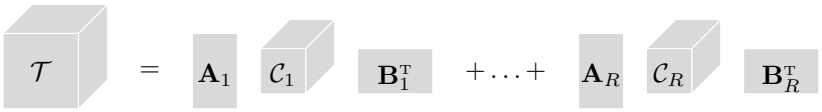


Figure 5: Graphical illustration of a BTD in multilinear rank-$(L, L, \cdot)$ terms.

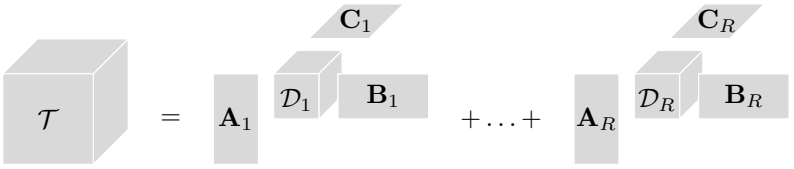


Figure 6: Graphical illustration of a BTD in multilinear rank-(L, M, N) terms.

long as the core tensor $\mathcal{C}_r$ is replaced by $\mathcal{C}_r \bullet_1 (\mathbf{G}_r)^{-1} \bullet_2 (\mathbf{H}_r)^{-1}$. Algorithms for the computation of this type of decomposition can be found in [27, 32] and toolboxes that implement these algorithms are available as well [33, 34].

3.6 Block term decomposition in rank-(L, M, N) terms

A block term decomposition (BTD) in multilinear rank-(L, M, N) is the most general block term decomposition and writes a tensor as a sum of core tensors, each multiplied by a factor matrix in every mode. Mathematically, the BTD of a third-order tensor $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$ in rank-(L, M, N) terms can be written as

$$\mathcal{T} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r,$$

in which $\mathcal{D}_r \in \mathbb{C}^{L \times M \times N}$ are full rank-(L, M, N) and $\mathbf{A}_r \in \mathbb{C}^{I \times L}, \mathbf{B}_r \in \mathbb{C}^{J \times M}$ and $\mathbf{C}_r \in \mathbb{C}^{K \times N}$ have full column rank [16]. A tensor $\mathcal{D}_r \in \mathbb{C}^{L \times M \times N}$ has full rank-(L, M, N) if its matrix unfoldings in the first, second and third mode have rank L, M and N, respectively. A graphical representation of a BTD in multilinear rank-(L, M, N) terms is given in Figure 6. Essential uniqueness conditions for this decomposition can be found in [16]. The trivial ambiguities in the decomposition are the following: the different terms can be arbitrarily permuted and within each term, one can postmultiply $\mathbf{A}_r$ by any nonsingular $\mathbf{F}_r \in \mathbb{C}^{L \times L}$, $\mathbf{B}_r$ by any nonsingular $\mathbf{G}_r \in \mathbb{C}^{M \times M}$ and $\mathbf{C}_r$ by any nonsingular $\mathbf{H}_r \in \mathbb{C}^{N \times N}$ as long as $\mathcal{D}_r$ is replaced by $\mathcal{D}_r \bullet_1 (\mathbf{F}_r)^{-1} \bullet_2 (\mathbf{G}_r)^{-1} \bullet_3 (\mathbf{H}_r)^{-1}$. Algorithms for the computation of this type of decomposition can be found in [27].

4 Tensor similarity

Rather than computing the underlying factors of a tensor, one may wish to compare the factors of two tensors. This can be done by decomposing both

tensors and subsequently comparing their factor matrices. However, recent advances in tensor analysis have shown that it is possible to retrieve information on the decompositions of tensors without explicitly computing them [6–8]. This approach is computationally more efficient and is numerically more stable [7].

This chapter explains and illustrates how information on tensor decompositions and their interrelations can be obtained without computing the actual tensor decompositions. Section 4.1 first defines the problem statement in more detail, after which the approach and underlying theory for several cases is treated in sections 4.2 to 4.3.2.

4.1 Problem statement

The concept of tensor similarity was introduced in [6]. First note that if two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{N \times N}$ are available with $\mathbf{A}$ having full column rank, knowledge that $\mathbf{B}$ has the same column and row spaces as $\mathbf{A}$ does not provide any information on the structure of $\mathbf{B}$ apart from the obvious result that $\mathbf{B}$ must also have full column rank. However, the situation is different when higher-order tensors are involved, where insights can be derived from the column and the row spaces of various matrix unfoldings.

Consider two third-order tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ for which it holds that

$$\text{span}\left\{\mathbf{A}_{[2,3;1]}\right\} = \text{span}\left\{\mathbf{B}_{[2,3;1]}\right\} \quad \text{and} \quad \text{span}\left\{\mathbf{A}_{[1,3;2]}\right\} = \text{span}\left\{\mathbf{B}_{[1,3;2]}\right\} \tag{6}$$

and possibly also that

$$\text{span}\left\{\mathbf{A}_{[1,2;3]}\right\} = \text{span}\left\{\mathbf{B}_{[1,2;3]}\right\}. \tag{7}$$

This chapter aims to explain what can be learned about $\mathcal{A}$ and $\mathcal{B}$ and their decompositions from (6) and (7).

Section 4.2 shows how to determine which decompositions $\mathcal{A}$ and $\mathcal{B}$ admit and how these decompositions are related based solely on equations (6) and (7). In particular, we examine if we can infer "essential equality" of the decompositions of $\mathcal{A}$ and $\mathcal{B}$ from (6) and (7). Essential equality means that both decompositions have the same factor matrices up to the trivial indeterminacies associated with the overall decomposition. Note that if the decomposition is not unique, essential equality means that there exist a pair of decompositions of $\mathcal{A}$ and $\mathcal{B}$ that are essentially equal. Essential equality in two modes means that only the factor matrices in two modes are equal up to trivial indeterminacies, as illustrated graphically for CPDs in Figure 7.

In section 4.3 prior information of the type of tensor decomposition is taken into account as well. If it is known that $\mathcal{A}$ and $\mathcal{B}$ admit CPDs or BTDs in multilinear rank-$(L, L, 1)$ terms, and the number of terms in known, the problem can be further simplified.

Note that the spaces spanned by the matrix unfoldings above are very different from the spaces spanned by the mode-n fibers. The spaces in (6) and (7) are spanned by the combination of two tensor modes. This is a key element in our analysis of similarity.

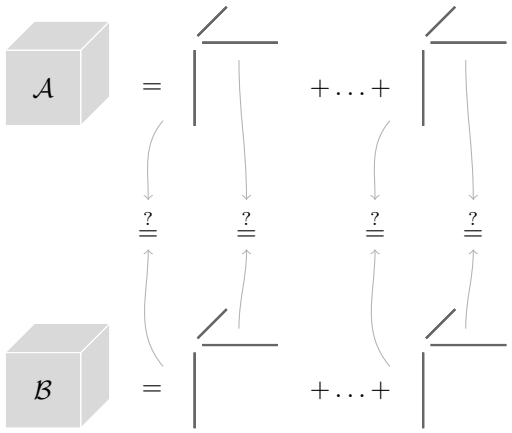


Figure 7: Graphical representation of comparing the factor matrices of $\mathcal{A}$ and $\mathcal{B}$ to investigate essential equality in the first two modes for the case in which both tensors admit a CPD. (Adopted with changes from [7])

4.2 Tensor similarity without assumptions on decompositions

Assume two tensors have coinciding spaces spanned by their matrix unfoldings in two dimensions as in equation (6), possibly even in three dimensions if (7) holds as well. In [6], it is shown how tensor structures and interrelations can be derived from this information in two cases for third-order tensors.

The first case considers two third-order tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ which both have mode-3 rank at least 2 and which both can be simultaneously compressed to tensors of dimensions $(N \times N \times I_3)$ such that the compressed version of $\mathcal{A}$ has full mode-1 and mode-2 rank. Additionally, assume that

$$\operatorname{span}\left\{\mathbf{A}_{[2,3;1]}\right\} = \operatorname{span}\left\{\mathbf{B}_{[2,3;1]}\right\} \quad \text{and} \quad \operatorname{span}\left\{\mathbf{A}_{[1,3;2]}\right\} = \operatorname{span}\left\{\mathbf{B}_{[1,3;2]}\right\}. \tag{8}$$

It follows from (8) that after compression, there must exist exactly two nonsingular matrices $\mathbf{M}_1 \in \mathbb{C}^{N \times N}$ and $\mathbf{M}_2 \in \mathbb{C}^{N \times N}$ such that

$$\mathcal{A} \cdot_1 \mathbf{M}_1 = \mathcal{B} \quad \text{and} \quad \mathcal{A} \cdot_2 \mathbf{M}_2 = \mathcal{B}.$$

Because of the interrelations between these first two modes, it is shown in [6] that the structures of $\mathcal{A}$ and $\mathcal{B}$ are strongly related and depend on the eigenvalues of $\mathbf{M}_2$ (or $\mathbf{M}_1$, which has the same eigenvalues [8]) and their multiplicities. More specifically, following scenarios are identified in [6]:

- If all eigenvalues of $\mathbf{M}_2$ are identical and equal to α and $\mathbf{M}_2$ can be diagonalized, then $\mathcal{A}$ and $\mathcal{B}$ are scaled versions of each other with scaling factor α. It thus holds that $\mathcal{B} = \alpha\mathcal{A}$.

- If all eigenvalues of $\mathbf{M}_2$ are distinct and equal to $\alpha_1, \alpha_2, \ldots, \alpha_N$, then $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-1 terms, but scaled by the eigenvalues of $\mathbf{M}_2$. Both tensors can be written as

$$\mathcal{A} = \sum_{n=1}^{N} \mathbf{x}_n \circ \mathbf{y}_n \circ \mathbf{z}_n,$$
$$\mathcal{B} = \sum_{n=1}^{N} \alpha_n \mathbf{x}_n \circ \mathbf{y}_n \circ \mathbf{z}_n.$$

- If $\mathbf{M}_2$ can be diagonalized and has R eigenvalues α_r with multiplicity L_r, then $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-$(L_r, L_r, \cdot)$ terms, again scaled by the eigenvalues of $\mathbf{M}_2$. In this case, both tensors can be written as

$$\mathcal{A} = \sum_{r=1}^{R} \mathcal{Z}_r \bullet_1 \mathbf{X}_r \bullet_2 \mathbf{Y}_r,$$
$$\mathcal{B} = \sum_{r=1}^{R} \alpha_r \mathcal{Z}_r \bullet_1 \mathbf{X}_r \bullet_2 \mathbf{Y}_r.$$

- Assume $\mathbf{M}_2$ has R eigenvalues α_r with multiplicity L_r. For $\alpha_2, \ldots, \alpha_R$ the geometric multiplicity is equal to the algebraic multiplicity and for α_1 the geometric multiplicity is equal to 1. In this case, $\mathcal{A}$ and $\mathcal{B}$ consist of the same (possibly differently scaled) rank-$(L_r, L_r, \cdot)$ terms, apart from one block term of which the core is multiplied by a Jordan cell, which has the form

$$\begin{bmatrix} \alpha_1 & 1 & 0 & \cdots & 0 \\ 0 & \alpha_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \alpha_1 & 1 \\ 0 & 0 & 0 & 0 & \alpha_1 \end{bmatrix} \in \mathbb{C}^{L_1 \times L_1}.$$

  If this Jordan cell is denoted by $\mathbf{J}(\alpha_1)$, then both tensors can be written as

$$\mathcal{A} = \sum_{r=1}^{R} \mathcal{Z}_r \bullet_1 \mathbf{X}_r \bullet_2 \mathbf{Y}_r,$$
$$\mathcal{B} = \mathcal{Z}_1 \bullet_1 \mathbf{X}_1 \bullet_2 (\mathbf{Y}_1 \mathbf{J}^{\mathrm{T}}(\alpha_1)) + \sum_{r=2}^{R} \alpha_r \mathcal{Z}_r \bullet_1 \mathbf{X}_r \bullet_2 \mathbf{Y}_r.$$

- More scenarios are possible based on the eigenvalues of $\mathbf{M}_2$ and their multiplicities.

These scenarios show that if (8) holds and $\mathcal{A}$ has full mode-1 and mode-2 rank after compression, the user can first compute $\mathbf{M}_2$ and subsequently find

its eigenvalues to determine the structure and interrelations of two tensors $\mathcal{A}$ and $\mathcal{B}$.

In a second case, the matrix unfoldings of both tensors should span the same space for all three dimensions. Mathematically, two third-order tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ are considered such that they can both be simultaneously compressed to tensors of size $(N \times N \times N)$ with the compressed version of $\mathcal{A}$ having full multilinear rank. If following conditions hold:

$$\begin{aligned} \operatorname{span}\left\{\mathbf{A}_{[2,3;1]}\right\} = \operatorname{span}\left\{\mathbf{B}_{[2,3;1]}\right\}, \quad \operatorname{span}\left\{\mathbf{A}_{[1,3;2]}\right\} = \operatorname{span}\left\{\mathbf{B}_{[1,3;2]}\right\} \\ \text{and} \quad \operatorname{span}\left\{\mathbf{A}_{[1,2;3]}\right\} = \operatorname{span}\left\{\mathbf{B}_{[1,2;3]}\right\}, \end{aligned} \tag{9}$$

it follows as before that after compression, there must exist nonsingular matrices $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3 \in \mathbb{C}^{N \times N}$ such that

$$\mathcal{A} \cdot_1 \mathbf{M}_1 = \mathcal{B}, \qquad \mathcal{A} \cdot_2 \mathbf{M}_2 = \mathcal{B} \quad \text{and} \quad \mathcal{A} \cdot_3 \mathbf{M}_3 = \mathcal{B}.$$

Similarly as in the first case, the interrelations between all three modes link the structures of both tensors to each other [6]. The tensor structures depend on the eigenvalues of $\mathbf{M}_2$ (or $\mathbf{M}_1$ or $\mathbf{M}_3$) and their multiplicities. Just as in the previous case, this leads to several scenarios:

- If all eigenvalues of $\mathbf{M}_2$ are equal to α and $\mathbf{M}_2$ can be diagonalized, then $\mathcal{A}$ and $\mathcal{B}$ are scaled versions of each other with scaling factor α. It thus holds that $\mathcal{B} = \alpha\mathcal{A}$.

- If all eigenvalues of $\mathbf{M}_2$ are distinct and equal to $\alpha_1, \alpha_2, \ldots, \alpha_N$, then $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-1 terms, but scaled by the eigenvalues of $\mathbf{M}_2$. Both tensors can be written as

$$\begin{aligned} \mathcal{A} &= \sum_{n=1}^{N} \mathbf{x}_n \circ \mathbf{y}_n \circ \mathbf{z}_n, \\ \mathcal{B} &= \sum_{n=1}^{N} \alpha_n \mathbf{x}_n \circ \mathbf{y}_n \circ \mathbf{z}_n. \end{aligned}$$

- If $\mathbf{M}_2$ can be diagonalized and has R eigenvalues α_r with multiplicity L_r, then $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-(L_r, L_r, L_r) terms, scaled by the eigenvalues of $\mathbf{M}_2$. In this case, both tensors can be written as

$$\begin{aligned} \mathcal{A} &= \sum_{r=1}^{R} \mathcal{W}_r \cdot_1 \mathbf{X}_r \cdot_2 \mathbf{Y}_r \cdot_3 \mathbf{Z}_r, \\ \mathcal{B} &= \sum_{r=1}^{R} \alpha_r \mathcal{W}_r \cdot_1 \mathbf{X}_r \cdot_2 \mathbf{Y}_r \cdot_3 \mathbf{Z}_r. \end{aligned}$$

- Assume $\mathbf{M}_2$ has R eigenvalues α_r with multiplicity L_r. For $\alpha_2, \ldots, \alpha_R$ the geometric multiplicity is equal to the algebraic multiplicity and for α_1

the geometric multiplicity is equal to 1. In this case, $\mathcal{A}$ and $\mathcal{B}$ consist of the same (possibly differently scaled) rank-(L_r, L_r, L_r) terms, apart from one block term of which the core is multiplied by a Jordan cell $\mathbf{J}(\alpha_1)$ associated with α_1. Both tensors can then be written as

$$\mathcal{A} = \sum_{r=1}^{R} \mathcal{W}_r \bullet_1 \mathbf{X}_r \bullet_2 \mathbf{Y}_r \bullet_3 \mathbf{Z}_r,$$

$$\mathcal{B} = \mathcal{W}_1 \bullet_1 \mathbf{X}_1 \bullet_2 (\mathbf{Y}_1 \mathbf{J}^{\mathrm{T}}(\alpha_1)) \bullet_3 \mathbf{Z}_1 + \sum_{r=2}^{R} \alpha_r \mathcal{W}_r \bullet_1 \mathbf{X}_r \bullet_2 \mathbf{Y}_r \bullet_3 \mathbf{Z}_r.$$

- Depending on the eigenvalues of $\mathbf{M}_2$ (or $\mathbf{M}_1$ or $\mathbf{M}_3$) and their multiplicities, other scenarios are possible.

Both cases above show that coinciding spaces spanned by tensor unfoldings indicate that the decompositions of those tensors are related. The exact structure of the tensors can to a large extent be derived from the eigenvalues of the matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$. In practice, following steps may be followed for both cases to determine how two third-order tensors are related:

1. **Verify the rank conditions:** The first step is to verify whether the (multilinear) rank conditions on $\mathcal{A}$ hold. This step is split up according to the number of modes in which the spaces spanned by the tensor unfoldings should coincide:

   - **Coinciding spaces of tensor unfoldings in two modes:** The rank conditions on $\mathcal{A}$ can be checked by computing its multilinear rank and verifying whether the mode-1 and mode-2 rank are equal and the mode-3 rank is at least two.
     If the mode-3 rank of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ is only 1, this implies that all frontal slices of $\mathcal{A}$ are scaled versions of each other. Since there is no additional structure in the third mode, this situation is equivalent to the matrix case. The only conclusion one can draw is that the mode-1 and mode-2 ranks of $\mathcal{B}$ must be the same as those of $\mathcal{A}$.
     If $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ has mode-1 and mode-2 rank equal to $N \leq I_1, I_2$, then it is possible to compress both tensors such that their compressed versions do have full mode-1 and mode-2 rank. This can be done as follows: first construct an orthonormal basis for the row space of $\mathbf{A}_{[2,3;1]}$. This can be done by computing the singular value decomposition of $\mathbf{A}_{[2,3;1]}^{\mathrm{T}}$ and storing the left singular vectors corresponding to the N dominant singular values as columns in a matrix $\mathbf{U}_1 \in \mathbb{C}^{I_1 \times N}$. Similarly, an orthonormal basis for the row space of $\mathbf{A}_{[1,3;2]}$ can be constructed and stored in $\mathbf{U}_2 \in \mathbb{C}^{I_2 \times N}$. Next, both tensors can be transformed as

$$\widetilde{\mathcal{A}} = \mathcal{A} \bullet_1 \mathbf{U}_1^{\mathrm{H}} \bullet_2 \mathbf{U}_2^{\mathrm{H}}, \tag{10}$$

$$\widetilde{\mathcal{B}} = \mathcal{B} \bullet_1 \mathbf{U}_1^{\mathrm{H}} \bullet_2 \mathbf{U}_2^{\mathrm{H}}. \tag{11}$$

Because both tensors are transformed using the same matrices and (8) is assumed to hold, their interrelations are not altered [8] and it still holds that

$$\operatorname{span}\left\{\widetilde{\mathbf{A}}_{[2,3;1]}\right\} = \operatorname{span}\left\{\widetilde{\mathbf{B}}_{[2,3;1]}\right\} \quad \text{and}$$
$$\operatorname{span}\left\{\widetilde{\mathbf{A}}_{[1,3;2]}\right\} = \operatorname{span}\left\{\widetilde{\mathbf{B}}_{[1,3;2]}\right\}.$$

Moreover, the matrix unfoldings of $\widetilde{\mathcal{A}}$ have full column rank by construction and it is proven in [8] that the multilinear rank of $\widetilde{\mathcal{B}}$ is the same as the multilinear rank of $\mathcal{B}$. These transformed tensors can consequently be used throughout the remaining analysis by directly replacing both tensors and their dimensions.

The computation of the compressed tensors $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ in (10) and (11) corresponds to the computation of a normalized Tucker-2 decomposition as in (4). The compressed tensors equal the core tensors of size (N, N, I_3) of (4). The fact that the same matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ can be used for the compression of both $\mathcal{A}$ and $\mathcal{B}$ is an obvious requirement for their decompositions to be similar. Mathematically, the existence of common matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ is implied by (8).

It follows from (10) and (11) that the original tensors $\mathcal{A}$ and $\mathcal{B}$ can be retrieved as follows:

$$\mathcal{A} = \widetilde{\mathcal{A}} \cdot_1 \mathbf{U}_1 \cdot_2 \mathbf{U}_2,$$
$$\mathcal{B} = \widetilde{\mathcal{B}} \cdot_1 \mathbf{U}_1 \cdot_2 \mathbf{U}_2.$$

- **Coinciding spaces of tensor unfoldings in three modes:** In this case, $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ must be compressible to a cubic tensor having full multilinear rank. To verify this, the multilinear rank of $\mathcal{A}$ must be computed. If the mode-n rank of $\mathcal{A}$ is not the same for all three modes, then the discussed theorem cannot be applied but the broader version of [8] explained further in this section is applicable. If the multilinear rank of $\mathcal{A}$ equals (N, N, N) with $N \leq I_1, I_2, I_3$, then (9) implies that both tensors can be compressed using the same matrices such that the compressed tensor $\widetilde{\mathcal{A}}$ is cubic and has full multilinear rank.

  To do this compression, the same approach as in the case with two coinciding modes can be used, but in all three modes. This means that orthonormal bases for the row spaces of $\mathbf{A}_{[2,3;1]}, \mathbf{A}_{[1,3;2]}$ and $\mathbf{A}_{[1,2;3]}$ are computed and stored as columns in the matrices $\mathbf{U}_1 \in \mathbb{C}^{I_1 \times N}, \mathbf{U}_2 \in \mathbb{C}^{I_2 \times N}, \mathbf{U}_3 \in \mathbb{C}^{I_3 \times N}$, respectively. The compressed tensors can then be computed as

$$\widetilde{\mathcal{A}} = \mathcal{A} \cdot_1 \mathbf{U}_1^{\mathrm{H}} \cdot_2 \mathbf{U}_2^{\mathrm{H}} \cdot_3 \mathbf{U}_3^{\mathrm{H}}, \tag{12}$$
$$\widetilde{\mathcal{B}} = \mathcal{B} \cdot_1 \mathbf{U}_1^{\mathrm{H}} \cdot_2 \mathbf{U}_2^{\mathrm{H}} \cdot_3 \mathbf{U}_3^{\mathrm{H}}. \tag{13}$$

This is equivalent to computing the MLSVD of $\mathcal{A}$ with core tensor $\widetilde{\mathcal{A}} \in \mathbb{C}^{N \times N \times N}$, which has full multilinear rank since the multilinear rank of $\mathcal{A}$ is (N, N, N). Again, the interrelations of both tensors are not altered since both tensors are compressed using the same factor matrices and (9) is assumed to hold [8]. Just as in the previous case, the tensors $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ can then be used throughout the remaining analysis.

2. **Verify coinciding spaces spanned by tensor unfoldings:** One way to characterize the similarity between two subspaces uses principal angles. For R-dimensional subspaces, the principal angles form a set of R minimized angles between both subspaces $\theta_1, \theta_2, \ldots, \theta_R$, ordered from smallest to largest with each angle $\theta_r \in [0, \pi/2]$. The principal angles between two matrices $\mathbf{A}$ and $\mathbf{B}$ are defined recursively, starting with the smallest. For every angle, two vectors $\mathbf{a} \in \operatorname{span}\{\mathbf{A}\}$ and $\mathbf{b} \in \operatorname{span}\{\mathbf{B}\}$ are sought such that the angle between these vectors is minimal under the additional constraint that these vectors must be perpendicular to the vectors that were used for the previous (smaller) principal angles. Mathematically, this can be written as

$$\theta_i \stackrel{\text{def}}{=} \min \left\{ \arccos\left(\frac{|\mathbf{a}^{\mathrm{T}}\mathbf{b}|}{||\mathbf{a}||\,||\mathbf{b}||} \right) \;\middle|\; \mathbf{a} \in \operatorname{span}\{\mathbf{A}\}, \ldots \right.$$
$$\left. \mathbf{b} \in \operatorname{span}\{\mathbf{B}\}, \mathbf{a} \perp \mathbf{a}_j, \mathbf{b} \perp \mathbf{b}_j \quad \forall j \in \{1, \ldots, i-1\} \right\},$$

in which $\mathbf{a}_j$ and $\mathbf{b}_j$ are the vectors corresponding to the angle θ_j. Further mathematical details can be found in [35] and algorithms are given in [36].

Since the principal angles form a set of minimized angles, they will all be zero if the subspaces are exactly the same or will be small if the subspaces are approximately equal, as is often the case in noisy conditions. By contrast, if the subspaces are not equal at all, the principal angles will all be close to $\pi/2$. If only $L < R$ principal angles between two subspaces X and Y are small, this indicates that $X \cap Y$ is an L-dimensional subspace.

To verify the condition that the subspaces coincide, the principal angles must be sufficiently small. The application at hand and the noise level typically determine what "sufficiently small" is. Determining whether subspaces are coinciding based on the principal angles is essentially a machine learning problem, and can be tackled with machine learning algorithms (of which thresholding is a simple example).

3. **Find the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ (and $\mathbf{M}_3$ if applicable):** these matrices can be directly computed from the sets of linear equations

$$\mathbf{A}_{[2,3;1]}\mathbf{M}_1 = \mathbf{B}_{[2,3;1]}, \quad \mathbf{A}_{[1,3;2]}\mathbf{M}_2 = \mathbf{B}_{[1,3;2]},$$
$$\text{and} \quad \mathbf{A}_{[1,2;3]}\mathbf{M}_3 = \mathbf{B}_{[1,2;3]},$$

4. **Compute the eigenvalues and their multiplicities of $\mathbf{M}_1$ and $\mathbf{M}_2$ (and $\mathbf{M}_3$ if applicable):** for this step, standard algorithms can be used.

5. **Determine the structure and relation of $\mathcal{A}$ and $\mathcal{B}$:** by mapping the eigenvalues and their multiplicities from the previous step to the cases presented above, the structure and relation between $\mathcal{A}$ and $\mathcal{B}$ can be determined.

In [8], a more general study of tensor similarity is conducted that also covers the cases of nonsquare tensors and tensors of order higher than 3. This is especially relevant in cases where the results presented in [6] cannot be directly applied, such as for tensors that have different mode-n ranks in each dimension or for higher-order tensors.

For third-order tensors, the main results of [8] can be summarized as follows: consider two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ with $\mathcal{A}$ having full mode-1 and mode-2 rank after compression of both tensors. More explicitly, if $\mathcal{A}$ has mode-1 rank $J_1 < I_1$ and mode-2 rank $J_2 < I_2$, both tensors $\mathcal{A}$ and $\mathcal{B}$ can be compressed such that the transformed tensor $\widetilde{\mathcal{A}} \in \mathbb{C}^{J_1 \times J_2 \times I_3}$ does satisfy the rank conditions. This compression can be done in the same way as explained in equations (10) and (11), except that J_1 and J_2 may be different in this case. The transformed tensors $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ can then be used throughout the rest of the theorem since their interrelations are not altered by the transformation under condition (14) below [8].

Assume the spaces spanned by the matrix unfoldings of $\mathcal{A}$ and $\mathcal{B}$ in two modes are related as follows:

$$\text{span}\left\{\mathbf{B}_{[2,3;1]}\right\} \subseteq \text{span}\left\{\mathbf{A}_{[2,3;1]}\right\} \text{ and } \text{span}\left\{\mathbf{B}_{[1,3;2]}\right\} \subseteq \text{span}\left\{\mathbf{A}_{[1,3;2]}\right\}. \quad (14)$$

From (14) it follows that there must exist two matrices $\mathbf{M}_1 \in \mathbb{C}^{I_1 \times I_1}$ and $\mathbf{M}_2 \in \mathbb{C}^{I_2 \times I_2}$ such that

$$\mathbf{B}_{[2,3;1]} = \mathbf{A}_{[2,3;1]}\mathbf{M}_1 \text{ and } \mathbf{B}_{[1,3;2]} = \mathbf{A}_{[1,3;2]}\mathbf{M}_2. \quad (15)$$

Under these conditions, it is shown in [8] that the tensors $\mathcal{A}$ and $\mathcal{B}$ both admit decompositions into multilinear rank-$(L_{1r}, L_{2r}, \cdot)$ terms which have the same factor matrices but different core tensors:

$$\mathcal{A} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{X}_r^{(1)} \bullet_2 \mathbf{X}_r^{(2)}, \quad (16)$$

$$\mathcal{B} = \sum_{r=1}^{R} \lambda_r \mathcal{D}_r \bullet_1 \mathbf{X}_r^{(1)} \bullet_2 \mathbf{X}_r^{(2)}. \quad (17)$$

with $\mathcal{D}_r \in \mathbb{C}^{L_{1r} \times L_{2r} \times I_3}$, $\mathbf{X}_r^{(1)} \in \mathbb{C}^{I_1 \times L_{1r}}$ and $\mathbf{X}_r^{(2)} \in \mathbb{C}^{I_2 \times L_{2r}}$. Just as in the previous cases, the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ can be computed directly from the sets of linear equations in (15). If these equations have no solution, it means that $\mathcal{B}$ cannot be generated by terms of the decomposition of $\mathcal{A}$ [8]. The value

for R in (16) and (17) can be found as the number of distinct eigenvalues λ_r of $\mathbf{M}_1$ (or $\mathbf{M}_2$). Note that $\mathbf{M}_1$ and $\mathbf{M}_2$ will have the same eigenvalues, but not necessarily the same multiplicities. The (shared) eigenvalues represent the scaling of the different terms in (17). The multiplicities of the eigenvalues λ_r of $\mathbf{M}_x$ corresponds to L_{xr}, which is the xth entry in the multilinear rank of the rth term and consequently determines the dimensions of $\mathcal{D}_r$ and $\mathbf{X}_r^{(x)}$.

To use these theorems in practice, the approach outlined before can be used taking into account that the rank conditions are different from the previous cases since nonsquare tensors are allowed after compression. If all conditions are satisfied, both tensors admit the same decomposition with the same factor matrices as outlined in (16) and (17). The eigenvalues of $\mathbf{M}_1$ (or $\mathbf{M}_2$) can then be used to determine the number of terms and the scaling of the terms in both decompositions. Note that the procedure can also be used for approximate decompositions, for instance when noise is present. In this case, equation (15) can be solved in least-squares sense and the average of eigenvalues λ_{1r} and λ_{2r} can be used to estimate λ_r. The values L_{1r} and L_{2r} can then be estimated by assessing how close the different eigenvalues λ_{nr} are to each of the estimated eigenvalues λ_r [8].

For further generalizations, mathematical specifics and proofs we refer to [8].

4.3 Similarity for specific decompositions

4.3.1 CPD similarity in two modes

In the previous section, no prior knowledge was available on the type of decomposition $\mathcal{A}$ and $\mathcal{B}$ admit. Here, it is assumed that both tensors admit a rank-R CPD, which simplifies the approach and computations. Moreover, only two modes of third-order tensors are compared. This is motivated by the observation that in several applications, such as emitter localization in telecommunications, the information of interest is contained in just two modes of a tensor decomposition [7]. In emitter localization for example, all information on the emitter location is contained in the first two factor matrices after computing a tensor decomposition, while the factor matrix in the third mode contains information about the source signals [37].

As stated in the problem statement in Section 4.1, the aim is to determine whether $\mathcal{A}$ and $\mathcal{B}$ are essentially equal in two modes. Mathematically, this translates to following problem statement: consider two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ that both admit a rank-R CPD:

$$\begin{aligned} \mathcal{A} &= \left[\!\left[\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)} \right]\!\right], \\ \mathcal{B} &= \left[\!\left[\mathbf{V}^{(1)}, \mathbf{V}^{(2)}, \mathbf{V}^{(3)} \right]\!\right]. \end{aligned} \tag{18}$$

The goal is to compare the matrices $\mathbf{U}^{(1)}$ and $\mathbf{V}^{(1)} \in \mathbb{C}^{I_1 \times R}$, and $\mathbf{U}^{(2)}$ and $\mathbf{V}^{(2)} \in \mathbb{C}^{I_2 \times R}$ up to the trivial indeterminacies of the CPD. This comparison can be done for the full factor matrices ("Full similarity in two modes") or for a subset of the CPD terms ("Partial similarity in two modes").

In this formulation, the first two out of three factor matrices are compared. If one wishes to compare other factors, the modes of the tensors need to be permuted first such that the modes of interest are the first two modes.

Full similarity in two modes The heart of the approach for full similarity in two modes lies in the following theorem [7]:

Theorem 1. *Let two tensors $\mathcal{A}, \mathcal{B}$ admit a rank-R CPD as shown in* (18)*. If following conditions hold:*

$$(i) \qquad \mathbf{U}^{(3)} \text{ and } \mathbf{V}^{(3)} \text{ have full column rank,}$$

$$(ii) \qquad \mathcal{C}_2\left(\mathbf{U}^{(1)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(2)}\right) \text{ has full column rank,}$$

then $\mathcal{A}$ and $\mathcal{B}$ are essentially equal in the first two modes if and only if span $\left\{\mathbf{A}_{[1,2;3]}\right\}$ = span $\left\{\mathbf{B}_{[1,2;3]}\right\}$.

In this theorem, $\mathcal{C}_2(\mathbf{K})$ is defined as the second compound matrix of an $I \times R$ matrix $\mathbf{K}$, which is the $\binom{I}{2} \times \binom{R}{2}$ matrix containing the determinants of all 2×2 submatrices of $\mathbf{K}$ [38]. The determinants are arranged such that the submatrix index sets are in lexicographic order. The intuition for the theorem is as follows: the conditions (i) and (ii) actually correspond to the uniqueness conditions for the CPD of $\mathcal{A}$. This implies that there exist no other rank-1 terms in the space span $\left\{\mathbf{A}_{[1,2;3]}\right\}$ apart from the rank-1 terms of the first two modes of $\mathcal{A}$. Consequently, if span $\left\{\mathbf{A}_{[1,2;3]}\right\}$ = span $\left\{\mathbf{B}_{[1,2;3]}\right\}$, the rank-1 terms in the first two modes of $\mathcal{B}$ must be the same up to scaling and permutation. The mathematical details of the proof can be found in [7]. The practical implications of this theorem are as follows: if the conditions hold, one can simply check whether the columns of both unfolded tensors span the same space. If this is the case, the factor $\mathbf{U}^{(1)}$ equals $\mathbf{V}^{(1)}$ up to column scaling and permutation, and $\mathbf{U}^{(2)}$ equals $\mathbf{V}^{(2)}$ up to column scaling and permutation.

Though the approach is conceptually simple, the details can be challenging. Below, we outline the steps one can follow to compare the factors in two modes of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ that both admit a rank-R CPD. These steps are largely based on the approach outlined in [7].

1. **Verify the conditions:** First, the factor matrices in the third mode of both tensors ($\mathbf{U}^{(3)}$ and $\mathbf{V}^{(3)}$) need to have full column rank. Second, the matrix $\mathcal{C}_2\left(\mathbf{U}^{(1)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(2)}\right)$ needs to have full column rank as well. The challenge in verifying these conditions is that the factor matrices are not explicitly available. Indeed, the whole purpose of the algorithm is to compare tensor factors without having to explicitly compute them. Following approaches can be used to verify the conditions:

   - The full column rank condition of a factor matrix (e.g., $\mathbf{U}^{(3)} \in \mathbb{C}^{I_3 \times R}$) can be verified as follows: consider the unfolded tensor $\mathbf{A}_{[1,2;3]}$ and check whether its rank equals R. Because this unfolding can be

written as $\mathbf{A}_{[1,2;3]} = \left(\mathbf{U}^{(1)} \odot \mathbf{U}^{(2)}\right) \mathbf{U}^{(3)^{\mathrm{T}}} \in \mathbb{C}^{I_1 I_2 \times I_3}$, this will be the case if $\mathbf{U}^{(3)}$ has full column rank and $R \leq I_3$ and $R \leq I_1 I_2$.

- To verify whether $\mathcal{C}_2\left(\mathbf{U}^{(1)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(2)}\right)$ has full column rank, one can turn to [39] where it is shown that this condition is the same as requiring that the tensors $\left\{\Phi\left(\mathbf{u}_t^{(1)}\mathbf{u}_t^{(2)^{\mathrm{T}}}, \mathbf{u}_v^{(1)}\mathbf{u}_v^{(2)^{\mathrm{T}}}\right)\right\}_{1\leq t<v\leq R}$ are linearly independent. In this condition, $\Phi(\cdot,\cdot)$ is a bilinear mapping that maps two matrices $\mathbf{M}, \mathbf{N} \in \mathbb{C}^{I\times J}$ to a fourth-order tensor $\Phi(\mathbf{M},\mathbf{N}) \in \mathbb{C}^{I\times I\times J\times J}$ given by

$$(\Phi(\mathbf{M},\mathbf{N}))_{ijkl} = m_{ik}n_{jl} + m_{jl}n_{ik} - m_{il}n_{jk} - m_{jk}n_{il}. \qquad (20)$$

To verify that the tensors $\left\{\Phi\left(\mathbf{u}_t^{(1)}\mathbf{u}_t^{(2)^{\mathrm{T}}}, \mathbf{u}_v^{(1)}\mathbf{u}_v^{(2)^{\mathrm{T}}}\right)\right\}_{1\leq t<v\leq R}$ are linearly independent without knowledge of the factor matrices $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$, first construct a basis $\mathbf{E}$ for $\mathbf{A}_{[1,2;3]} = \left(\mathbf{U}^{(1)} \odot \mathbf{U}^{(2)}\right) \mathbf{U}^{(3)^{\mathrm{T}}} \in \mathbb{C}^{I_1 I_2 \times I_3}$ by selecting the left singular vectors corresponding to the R dominant singular values and storing these as columns in $\mathbf{E} \in \mathbb{C}^{I_1 I_2 \times R}$. Next, construct the matrices $\mathbf{E}_r \in \mathbb{C}^{I_1 \times I_2}$ by matricizing each column such that

$$(\mathbf{E}_r)_{i_1 i_2} = e_{(i_1-1)I_2+i_2,r} \qquad \forall i_1, i_2, r.$$

It now follows from the derivation in [39] that if the tensors $\{\Phi(\mathbf{E}_r,\mathbf{E}_s)\}_{1\leq r\leq s\leq R}$ are linearly independent, then $\left\{\Phi\left(\mathbf{u}_t^{(1)}\mathbf{u}_t^{(2)^{\mathrm{T}}}, \mathbf{u}_v^{(1)}\mathbf{u}_v^{(2)^{\mathrm{T}}}\right)\right\}_{1\leq t<v\leq R}$ are linearly independent as well and it consequently holds that $\mathcal{C}_2\left(\mathbf{U}^{(1)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(2)}\right)$ has full column rank. To verify that the tensors $\{\Phi(\mathbf{E}_r,\mathbf{E}_s)\}_{1\leq r\leq s\leq R}$ are linearly independent, first compute them via (20). Next, vectorize the results to obtain vectors $\mathbf{p}_{rs} \in \mathbb{C}^{I_1^2 I_2^2}$ and stack these vectors in a matrix $\mathbf{P} \in \mathbb{C}^{I_1^2 I_2^2 \times R(R+1)/2}$. If $\mathbf{P}$ has full column rank, the tensors $\{\Phi(\mathbf{E}_r,\mathbf{E}_s)\}_{1\leq r\leq s\leq R}$ are linearly independent and it follows that $\mathcal{C}_2\left(\mathbf{U}^{(1)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(2)}\right)$ has full column rank..

A different approach to validate the conditions is based on generic properties, which are easier to check but only hold for "almost all" cases.

- For the first condition, which states that $\mathbf{U}^{(3)} \in \mathbb{C}^{I_3\times R}$ and $\mathbf{V}^{(3)} \in \mathbb{C}^{I_3\times R}$ must have full column rank, the property can be used that states that a generic matrix has full column rank if it has at least as many rows as columns. Consequently, the first condition is satisfied if $I_3 \geq R$.
- For the second condition, it is stated in [18] that a matrix $\mathcal{C}_2(\mathbf{X}) \odot \mathcal{C}_2(\mathbf{Y})$ generically has full column when it has at least as many rows as columns. In our case, the matrix $\mathcal{C}_2\left(\mathbf{U}^{(1)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(2)}\right)$ has dimensions $\binom{I_1}{2}\binom{I_2}{2} \times \binom{R}{2}$, which means that the condition is generically satisfied if $\binom{I_1}{2}\binom{I_2}{2} \geq \binom{R}{2}$.

2. **Determine the number of components:** In many applications, the rank R of the tensors is not known in advance. Determining the rank of a tensor is an NP-hard problem and often requires decomposing the tensor trying various ranks. Since the approach presented here aims to avoid computing tensor decompositions, prior knowledge or heuristics are needed. In the simplest case, the rank is known from the application. If no prior knowledge is available, one can resort to estimates. As mentioned earlier, a lower bound on the tensor rank is the number of significant singular values of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$, capturing most of the information of both matrices. Under the assumptions of Theorem 1, these matrices will have rank R in the noiseless case. If noise is present, there will be R dominant singular values and the other singular values are expected to be significantly smaller.

3. **Find a basis for the dominant column spaces of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$:** The subspace spanned by the columns of $\mathbf{A}_{[1,2;3]}$ and the number of components R are known. In this step, the goal is to construct a basis for the R-dimensional dominant subspace of $\mathbf{A}_{[1,2;3]}$. One way to do this is by computing the left singular vectors of $\mathbf{A}_{[1,2;3]}$ corresponding to the R largest principal values. Store these left singular vectors as the columns in a matrix $\mathbf{X} \in \mathbb{C}^{I_1 I_2 \times R}$. The same approach can be used to extract a basis $\mathbf{Y} \in \mathbb{C}^{I_1 I_2 \times R}$ for the dominant subspace of $\mathbf{B}_{[1,2;3]}$.

4. **Compute the principal angles between both subspaces:** In the previous step, the bases $\mathbf{X} \in \mathbb{C}^{I_1 I_2 \times R}$ and $\mathbf{Y} \in \mathbb{C}^{I_1 I_2 \times R}$ were constructed for both the subspaces spanned by the columns of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$, respectively. The principal angles between these two bases can then be computed as explained in Section 4.2.

5. **Classify the subspaces as equal or not equal:** Based on the R principal angles, classify the subspaces must be classified as "equal" or "not equal" based on whether the principal angles are sufficiently small.

Partial similarity in two modes The theorem and procedure in the previous section aims at determining whether all terms of two CPDs are essentially equal in two modes. Here, partial similarity is considered, where only a subset of the terms of both CPDs are essentially equal in two modes. Again, the approach is based on a central theorem:

Theorem 2. *Consider two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ admitting a rank-R CPD*

as defined in (18). *If*

(i) $\mathbf{U}^{(3)}$ *and* $\mathbf{V}^{(3)}$ *have full column rank* R,

(ii) $\mathcal{C}_2\left(\mathbf{U}^{(2)}\right) \odot \mathcal{C}_2\left(\mathbf{U}^{(1)}\right)$ *has full column rank*,

(iii) $\mathcal{C}_2\left(\mathbf{V}^{(2)}\right) \odot \mathcal{C}_2\left(\mathbf{V}^{(1)}\right)$ *has full column rank*,

(iv) $\mathit{span}\left\{\mathbf{A}_{[1,2;3]}\right\} \cap \mathit{span}\left\{\mathbf{B}_{[1,2;3]}\right\}$ *is* Q*-dimensional and admits a unique (unfolded) rank-*Q *CPD*,

then there are Q *essentially equal terms in two modes if and only if* Q *principal angles between the unfolded tensors* $\mathbf{A}_{[1,2;3]}$ *and* $\mathbf{B}_{[1,2;3]}$ *are zero, with* $Q \leq R$.

The approach for partial similarity is identical to the approach outlined under "full similarity in two modes", apart from two differences:

- In the first step, the conditions of Theorem 2 must be verified. To do this, the same techniques as before can be used for the first three conditions.

  The fourth condition states that the Q-dimensional space $\left(\text{span}\left\{\mathbf{A}_{[1,2;3]}\right\} \cap \text{span}\left\{\mathbf{B}_{[1,2;3]}\right\}\right)$ must admit a rank-Q CPD. Say a basis of $\left(\text{span}\left\{\mathbf{A}_{[1,2;3]}\right\} \cap \text{span}\left\{\mathbf{B}_{[1,2;3]}\right\}\right)$ is stacked in $\mathbf{W} \in \mathbb{C}^{I_1 I_2 \times Q}$. This matrix is to be seen as the unfolding $\mathbf{W}_{[1,2;3]}$ of a third-order tensor $\mathcal{W} \in \mathbb{C}^{I_1 \times I_2 \times Q}$. Now a rank-$Q$ CPD of $\mathcal{W}$ needs to be computed to check whether it exists/admits a good approximation of $\mathcal{W}$ and whether it is unique (e.g., the uniqueness conditions in [10, 17–19]). This means that it is not possible to compare partial similarity in two modes without computing at least this decomposition, which risks issues with ill-conditioning.

- In the last step in which the subspaces are classified based on the principal angles, only the Q smallest principal angles must be considered for the classification problem.

If two tensors are partially equal in two modes, the common terms can be extracted without computing the full tensor decompositions of both tensors. This is especially advantageous when both tensors admits high-rank CPDs with only a few terms in common. The interested reader is refered to [7] for a step-by-step approach on how to do this.

4.3.2 BTD in rank-$(L, L, 1)$ terms similarity in two modes

This section assumes that both tensors $\mathcal{A}$ and $\mathcal{B}$ admit a BTD in R multilinear rank-$(L, L, 1)$ terms, which, compared to Section 4.2 simplifies the approach and computations but, compared to the CPD case in Section 4.3.1, allows tensor similarity in situations where the rank-1 assumption is too strict. The aim here is to determine whether both tensors are essentially equal in two modes. Mathematically, this translates to following problem statement. Consider two

tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ that both admit a BTD in R multilinear rank-$(L, L, 1)$ terms:

$$\mathcal{A} = \sum_{r=1}^{R} \left(\mathbf{U}_r^{(1)} \mathbf{U}_r^{(2)^{\mathrm{T}}} \right) \circ \mathbf{u}_r^{(3)}, \tag{22}$$

$$\mathcal{B} = \sum_{r=1}^{R} \left(\mathbf{V}_r^{(1)} \mathbf{V}_r^{(2)^{\mathrm{T}}} \right) \circ \mathbf{v}_r^{(3)}, \tag{23}$$

with $\mathbf{U}_r^{(1)}, \mathbf{V}_r^{(1)} \in \mathbb{C}^{I_1 \times L}$ and $\mathbf{U}_r^{(2)}, \mathbf{V}_r^{(2)} \in \mathbb{C}^{I_2 \times L}$. The goal is to assess whether the low-rank matrices $\mathbf{U}_r^{(1)} \mathbf{U}_r^{(2)^{\mathrm{T}}}$ and $\mathbf{V}_r^{(1)} \mathbf{V}_r^{(2)^{\mathrm{T}}}$ are the same up to permutation and scaling.

The heart of the approach lies in following theorem, adapted from [7]:

Theorem 3. *Consider two tensors $\mathcal{A}, \mathcal{B}$ admitting a BTD in R rank-$(L, L, 1)$ terms as defined in* (22) *and* (23)*. If following conditions hold*

(i) $\mathbf{U}^{(3)}$ *and* $\mathbf{V}^{(3)}$ *have full column rank,*

(ii) *The BTD of $\mathcal{A}$ is essentially unique,*

then $\mathcal{A}$ and $\mathcal{B}$ are essentially equal in the first two modes if and only if span $\left\{ \mathbf{A}_{[1,2;3]} \right\} =$ *span* $\left\{ \mathbf{B}_{[1,2;3]} \right\}$.

In practice, following steps can be followed [7]:

1. **Verifying the conditions:** Since the factor matrices are not readily available, alternative approaches are needed to verify the conditions. The factor matrices in the third mode of both tensors ($\mathbf{U}^{(3)}$ and $\mathbf{V}^{(3)}$) need to have full column rank, which can be checked using the same approach presented in Section 4.3.1. To determine whether the BTD of $\mathcal{A}$ is essentially unique, deterministic uniqueness conditions can be found in [26], though these require access to (some algebraic properties of) the factors of $\mathcal{A}$. In some cases, these properties can be inferred directly from $\mathcal{A}$, without computing the decomposition. For instance, if the multilinear rank of $\mathcal{A}$ equals (LR, LR, R), then the BTD in (22) is necessarily essentially unique. Generic uniqueness conditions which only use the dimensions of $\mathcal{A}$ and its factors can be found in [16, 26].

2. **Determine the number of components:** As before, (a lower bound on) the number of components is given by the number of dominant singular values of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$. If both conditions of Theorem 3 hold, these matrices will have rank R in the noiseless case. If noise is present, there will be R dominant singular values and the other singular values are expected to be significantly smaller.

3. **Find a basis for the dominant column spaces of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$:** In this step, the same techniques as in Section 4.3.1 can be used.

4. **Compute the principal angles between both subspaces:** As in Section 4.2, principal angles can be used as a similarity measure for the subspaces span $\left\{\mathbf{A}_{[1,2;3]}\right\}$ and span $\left\{\mathbf{A}_{[1,2;3]}\right\}$.

5. **Classify the subspaces as equal or not equal:** This step is identical to the approach outlined in Section 4.3.1.

5 Similarity in chemometrics

This section aims to clarify the use of the procedures and theorems given in the previous sections by applying them to a simple chemometrics example. Consider the amino acids fluorescence dataset (`http://www.models.life.ku.dk/amino_acid_fluo`), which consists of five simple laboratory-made samples. Each of these samples consists of different amounts of tyrosine, tryptophan and phenylalanine dissolved in phosphate-buffered water. Two samples are mixtures of all components, while the other three samples are measurements of the separate analytes. The samples were measured by fluorescence with excitation wavelengths of $240-300$ nm and emmision wavelengths of $250-450$ nm with 1 nm intervals. For more technical details, we refer the reader to [5] or the URL given earlier.

5.1 Comparison of sets of mixtures without prior information

In Section 4.2, a procedure was outlined to assess the relationship between two tensors without any prior knowledge on their decompositions. To illustrate this in a chemometrics context, two sets of four mixtures are artificially created using the excitation-emission spectra of the pure analytes in the amino acids fluorescence dataset (tyrosine, tryptophan and phenylalanine). The resulting excitation-emission matrices of each mixture are subsequently stacked per mixture set as frontal slices in two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{201\times 61\times 4}$. The concentrations of the analytes in each frontal slice are given in Table 1.

To compare $\mathcal{A}$ and $\mathcal{B}$ without making any assumption on their decompositions, the procedure outlined in Section 4.2 can be followed:

1. **Verify the rank conditions:** The multilinear rank of both tensors can be determined by computing the dominant singular values of their matrix unfoldings. For $\mathcal{A}$, the largest three singular values in each mode capture most of the information, as illustrated in Figure 8. This implies that the multilinear rank of $\mathcal{A}$ is $(3, 3, 3)$. Analogously, it can be shown that $\mathcal{B}$ has multilinear rank $(3, 3, 3)$ as well. Before any compression of the tensors can be done, it must be verified whether their matrix unfoldings span coinciding spaces.

2. **Verify coinciding spaces spanned by tensor unfoldings:** To compute the principal angles, standard algorithms can be used (e.g., [36]). Table 2 shows the five largest principal angles for each of the equations

Table 1: Concentrations of the analytes Tyrosine, Tryptophan and Phenylalanine in each mixture, represented by the frontal slices of $\mathcal{A}$ and $\mathcal{B}$ which contain the excitation-emission spectra of the mixtures.

	Tyr	Tryp	Phe
$\mathbf{A}_{::1}$	$0.0267 \cdot 10^{-3}$	$0.0133 \cdot 10^{-3}$	$0.1800 \cdot 10^{-3}$
$\mathbf{A}_{::2}$	$0.0267 \cdot 10^{-3}$	$0.0266 \cdot 10^{-3}$	$0.2700 \cdot 10^{-3}$
$\mathbf{A}_{::3}$	$0.0320 \cdot 10^{-3}$	$0.0067 \cdot 10^{-3}$	$0.3600 \cdot 10^{-3}$
$\mathbf{A}_{::4}$	$0.0107 \cdot 10^{-3}$	$0.0100 \cdot 10^{-3}$	$0.7200 \cdot 10^{-3}$
$\mathbf{B}_{::1}$	$0.0133 \cdot 10^{-3}$	$0.0266 \cdot 10^{-3}$	$0.1350 \cdot 10^{-3}$
$\mathbf{B}_{::2}$	$0.0133 \cdot 10^{-3}$	$0.0532 \cdot 10^{-3}$	$0.2025 \cdot 10^{-3}$
$\mathbf{B}_{::3}$	$0.0160 \cdot 10^{-3}$	$0.0133 \cdot 10^{-3}$	$0.2700 \cdot 10^{-3}$
$\mathbf{B}_{::4}$	$0.0053 \cdot 10^{-3}$	$0.0200 \cdot 10^{-3}$	$0.5400 \cdot 10^{-3}$

Table 2: Largest principal angles for the comparison of the spaces spanned by the matrix unfoldings of $\mathcal{A}$ and $\mathcal{B}$ in each mode.

	$\operatorname{span}\left\{\mathbf{A}_{[2,3;1]}\right\}$ $\|\|^?$ $\operatorname{span}\left\{\mathbf{B}_{[2,3;1]}\right\}$	$\operatorname{span}\left\{\mathbf{A}_{[1,3;2]}\right\}$ $\|\|^?$ $\operatorname{span}\left\{\mathbf{B}_{[1,3;2]}\right\}$	$\operatorname{span}\left\{\mathbf{A}_{[1,2;3]}\right\}$ $\|\|^?$ $\operatorname{span}\left\{\mathbf{B}_{[1,2;3]}\right\}$
Principal angles	$1.168 \cdot 10^{-9}$	0.4759	$0.838 \cdot 10^{-13}$
	$1.251 \cdot 10^{-9}$	0.4885	$0.948 \cdot 10^{-13}$
	$1.629 \cdot 10^{-9}$	0.5346	$7.920 \cdot 10^{-13}$
	$1.917 \cdot 10^{-9}$	0.5753	
	$2.567 \cdot 10^{-9}$	0.6478	

in (9). The principal angles are extremely low for the comparison in two modes. For the comparison of $\operatorname{span}\left\{\mathbf{A}_{[1,3;2]}\right\}$ and $\operatorname{span}\left\{\mathbf{B}_{[1,3;2]}\right\}$, the principal angles are significantly larger. However, noise must be taken into account since this example is based on actual excitation-emission spectroscopy measurements of analytes. With this in mind, the largest principal angle (0.6478) is still small compared to the value for perpendicular spaces ($\pi/2$). Consequently, it is reasonable to conclude that the spaces spanned by the matrix unfoldings in three modes can be considered equal, i.e., (9) holds.

Since it has been shown that $\mathcal{A}$ is compressible to a $(3,3,3)$-tensor with full multilinear rank and (9) holds, tensors $\mathcal{A}$ and $\mathcal{B}$ can be compressed to $\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}} \in \mathbb{R}^{3\times3\times3}$ as in (12) and (13). It can be verified that the principal angles between the spaces spanned by the matrix unfoldings of $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are still sufficiently small to be considered equal.

3. **Find the matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$:** This step is trivial.

4. **Compute the eigenvalues and their multiplicities of $\mathbf{M}_1$, $\mathbf{M}_2$ and**

Table 3: Largest principal angles for the comparison of the spaces spanned by the matrix unfoldings of $\mathcal{A}$ and $\mathcal{B}$ in each mode.

$\mathbf{M}_1$	$\mathbf{M}_2$	$\mathbf{M}_3$
0.5518	1.8455	0.7418
0.7647	0.5557	0.5524
1.6818	0.7026	1.7090

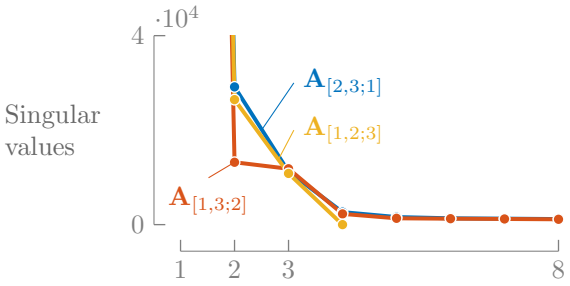


Figure 8: The first three singular values of the matrix unfoldings $\mathbf{A}_{[2,3;1]}$, $\mathbf{A}_{[1,3;2]}$ and $\mathbf{A}_{[1,2;3]}$ contain most of the information.

$\mathbf{M}_3$: The eigenvalues of all three matrices are given in Table 3. The table shows that the matrices have clearly distinct eigenvalues. Note that the eigenvalues of the different matrices lie close to eachother but are not exactly the same since noise is present in the measurement data and the compressed tensors are not perfect mixtures.

5. **Determine the structure and relation of $\mathcal{A}$ and $\mathcal{B}$:** Since all eigenvalues of $\mathbf{M}_1$ (or $\mathbf{M}_2$ or $\mathbf{M}_3$) are distinct, it follows that $\mathcal{A}$ and $\mathcal{B}$ admit a rank-3 CPD consisting of the same (possibly differently scaled) terms. Each of the terms represent one of the pure analytes in this example. More specifically, the factor vector in the first mode of each term contains the excitation information, the second the emission information and the third the concentrations. Since both tenors consist of the same terms, the concentrations of each analyte in the mixtures of $\mathcal{B}$ should be a scaled version of the concentrations in the mixtures of $\mathcal{A}$. Indeed, it can be verified from Table 1 that the mixtures in the frontal slices of $\mathcal{B}$ contain 0.5 times the concentration of Tyrosine in $\mathcal{A}$, 2 times the concentration of Tryptophan and 0.75 times the concentration of Phenylalanine. These scaling factors are retrieved as the eigenvalues of $\mathbf{M}_1$ (or $\mathbf{M}_2$ or $\mathbf{M}_3$), but in this case only approximately because of the influence of noise.

5.2 Comparison of analytes in mixtures (with prior information)

In this example, the approach of Section 4.3.2 is illustrated in two chemometrics applications as was done in [7], though more details and a step-by-step outline are provided in this chapter. Specifically, the goal is to determine whether two amino acid fluorescence mixture samples contain the same analytes.

In a first experiment, the two samples of the amino acids fluorescence dataset containing mixtures are selected. Each of these samples is represented by a matrix of size 61×201. To transform these matrices to tensors, Löwnerization is used along the emission mode as explained in Section 3.4. This yields two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{101\times100\times61}$, which both admit a BTD in rank-$(2, 2, 1)$ terms, which means that each of the pure analytes contributes one multilinear rank-$(2, 2, 1)$ term to the tensors [31]. In section 3.4 it was explained how the number of components and the multilinear rank of the terms can be determined based on the tensors $\mathcal{A}$ and $\mathcal{B}$.

To compare the underlying analytes of both mixtures, the steps outlined in Section 4.3.2 are followed:

1. **Verifying the conditions:** Because the tensors admit BTDs in multilinear rank-$(2, 2, 1)$ terms, the conditions of Theorem 3 need to be verified.

   - The first condition states that the factors in the third mode must have full column rank. To verify this, the dominant singular values of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$ are computed. Figure 9 shows the eight largest principal angles for both matrices. Most of the information seems to be captured by the first three singular vectors. For this reason, the number of components is estimated to be equal to $R = 3$ and it immediately follows that the factors in the third mode must have full column rank.
   - The second condition says that the BTD of $\mathcal{A}$ must be essentially unique. For this, Theorem 4.1 in [16] can be used, which states that this BTD is generically unique if $\min(I, J) \geq RL$ and the third factor does not have proportional columns. The number of components was computed to be $R = 3$ based on the dominant singular values of $\mathbf{A}_{[1,2;3]}$, from which it follows that the third factor does not have proportional columns. To determine L, one can compute the number of dominant singular values of $\mathbf{A}_{[2,3;1]}$, which equals RL if $\min(I, J) \geq RL$. In this experiment, it can be verified that $\mathbf{A}_{[2,3;1]}$ has 6 dominant singular values. Since the other dimensions are known: $I = 101$, $J = 100$ and $R = 3$, it follows that $L = 2$ and that $\min(I, J) \geq RL$ effectively holds. It thus follows that the BTD of $\mathcal{A}$ is essentially unique.

2. **Determine the number of components:** The number of components was already computed to be $R = 3$ when verifying the conditions. This

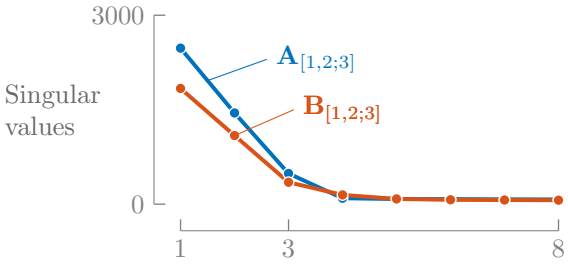


Figure 9: The first three singular values of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$ contain most of the information. Adopted with changes from [7].

value means that the considered mixtures are made up out of three pure analytes, each contributing one rank-$(2, 2, 1)$ term to the tensor.

3. **Find a basis for the dominant column spaces of $\mathbf{A}_{[1,2;3]}$ and $\mathbf{B}_{[1,2;3]}$:** This can be done by taking the left singular vectors that correspond to the three largest singular values and storing them in the matrices $\mathbf{T}^{(1)}$ and $\mathbf{T}^{(2)} \in \mathbb{C}^{10100\times 3}$.

4. **Compute the principal angles between both subspaces:** For this step, an implementation of the algorithm described in [36] was used. The computed principal values are 0.037, 0.064 and 0.328.

5. **Classify the subspaces as equal or not equal:** To classify whether the subspaces are equal or not, the principal angles are examined in more detail. If the subspaces are perpendicular, all principal angles are $\pi/2$. In this application, noise is present and the largest principal angle (0.328) is still reasonably low. For our example application, it can be concluded that the subspaces can be considered equal.

Because the conditions in Theorem 3 are valid and the (dominant) subspaces of the unfolded tensors are equal, it follows that $\mathcal{A}$ and $\mathcal{B}$ are essentially equal in the first two modes. This means that the block terms making up both tensors are the same up to scaling and permutation. Since each of these block terms represents a pure analyte, it can be concluded that both samples are mixtures of the same pure analytes. The description of the data set confirms this because both samples have been obtained by mixing different amounts of the same three amino acids: Tyrosine, Tryptophan and Phenylalanine.

Apart from the mixtures above, the dataset also contains measurements of the three pure analytes. In a final experiment, these measurements were used to artificially create several mixtures consisting of two analytes. The approach

Table 4: The principal angles clearly indicate which mixtures consist of the same components. Adopted from [7].

Mixtures to compare	Principal angles (rad)	
Mix1 and Mix2	0.0664	1.3032
Mix1 and Mix3	0.0217	0.0674
Mix2 and Mix3	0.1198	1.2892

of [7] is followed and following mixtures have been constructed:

$$\mathrm{EE}\,(\mathrm{Mix1}) = \mathrm{EE}\,(\mathrm{Trp}) + \mathrm{EE}\,(\mathrm{Tyr})\,, \tag{25}$$

$$\mathrm{EE}\,(\mathrm{Mix2}) = 0.4\ \mathrm{EE}\,(\mathrm{Tyr}) + 0.8\ \mathrm{EE}\,(\mathrm{Phe})\,,$$

$$\mathrm{EE}\,(\mathrm{Mix3}) = 0.8\ \mathrm{EE}\,(\mathrm{Trp}) + 0.4\ \mathrm{EE}\,(\mathrm{Tyr})\,, \tag{26}$$

in which $\mathrm{EE}\,(\mathrm{x})$ is a shorthand to denote the excitation-emission matrix of x. The construction of the mixtures shows that Mix1 and Mix3 contain the same pure analytes. Let us now use tensor similarity to verify this statement. A step-by-step outline is given below:

First, the data needs to be tensorized since only three separate excitation-emission matrices are available. For this, Löwnerization is used as in the previous experiment to obtain tensors $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, which all have dimensions $101 \times 100 \times 61$. Next, the procedure to compare tensors that admit a BTD in rank-$(2, 2, 1)$ terms can be used, just as before:

1. **Verifying the conditions:** Because the tensors admit BTDs in rank-$(2, 2, 1)$ terms, the conditions of Theorem 3 must be verified. It can be verified in exactly the same way as in the previous experiment that both conditions hold for any two tensors of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ one wishes to compare.

2. **Determine the number of components:** By computing the number of dominant singular values of $\mathbf{A}_{[1,2;3]}$, $\mathbf{B}_{[1,2;3]}$ and $\mathbf{C}_{[1,2;3]}$, it can be verified that all three tensors lead to $R = 2$ as the number of components. This is in line with the construction of the mixtures out of two pure analytes.

3. **Find a basis for the dominant column spaces of $\mathbf{A}_{[1,2;3]}$, $\mathbf{B}_{[1,2;3]}$ and $\mathbf{C}_{[1,2;3]}$:** This is done by taking the left singular vectors that correspond to the two largest singular values.

4. **Compute the principal angles between the subspaces:** To compute the principal angles an implementation of the algorithm described in [36] is used. Table 4 shows the principal angles between all mixtures.

5. **Classify the subspaces as equal or not equal:** Table 4 shows that all principal angles between Mix1 and Mix3 are small. The other combinations have one principal angle that is significantly larger. From these observations, it follows that the subspaces related to Mix1 and Mix3 can be considered equal.

Because the conditions of Theorem 3 are verified and the subspaces of the first and third mixture can be considered equal, it follows that Mix1 and Mix3 consist of the same pure analytes. This is indeed confirmed by the construction of the mixtures in (25) - (26), which shows that the same pure analytes have been used to create Mix1 and Mix3. By contrast, the principal angles clearly showed that Mix2 does not consist of the same pure analytes. Though it is not part of any of the theorems presented in this chapter, note that one of the principal angles between Mix2 and the other mixtures is small since one of the pure analytes is shared.

6 Conclusion

Tensors have been widespread in chemometrics for decades. In this chapter, state-of-the-art results on tensor similarity have been presented. The specific properties of tensors allow one to extract information on the decompositions of two tensors, without having to compute these decompositions. Results and algorithms have been discussed for tensors admitting CPDs and BTDs in multilinear rank-$(L, L, 1)$ terms, and for tensors of which the decomposition is not known a priori. The procedure and results have been illustrated on a simple chemometrics example as well.

References

(1) Smilde, A.; Bro, R.; Geladi, P., *Multi-way analysis: applications in the chemical sciences*; John Wiley & Sons: 2005.

(2) *Handbook of blind source separation, independent component analysis and applications*; Comon, P., Jutten, C., Eds.; Academic Press: 2010.

(3) De Lathauwer, L.; De Moor, B.; Vandewalle, J. Fetal electrocardiogram extraction by blind source subspace separation. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 567–572.

(4) Hyvärinen, A.; Oja, E. Independent component analysis: algorithms and applications. *Neural networks* **2000**, *13*, 411–430.

(5) Bro, R. PARAFAC. Tutorial and applications. *Chemometr. Intell. Lab.* **1997**, *38*, 149–171.

(6) De Lathauwer, L. Characterizing higher-order tensors by means of subspaces., tech. report 11–32, ESAT–STADIUS, KU Leuven, Belgium, 2011.

(7) Van Eeghem, F.; Debals, O.; De Lathauwer, L. Tensor Similarity in Two Modes. *IEEE Trans. Signal Process.* **2018**, *66*, 1273–1285.

(8) Domanov, I.; De Lathauwer, L. From computation to comparison of tensor decompositions. *arXiv:1912.04694*.

(9) Kolda, T. G.; Bader, B. W. Tensor decompositions and applications. *SIAM Rev.* **2009**, *51*, 455–500.

(10) Sidiropoulos, N. D.; De Lathauwer, L.; Fu, X.; Huang, K.; Papalexakis, E. E.; Faloutsos, C. Tensor Decomposition for Signal Processing and Machine Learning. *IEEE Trans. Signal Process.* **2017**, *65*, 3551–3582.

(11) Cichocki, A.; Mandic, C.; Phan, A. H.; Caifa, C.; Zhou, G.; Zhao, Q.; De Lathauwer, L. Tensor Decompositions for Signal Processing Applications: From two-way to multiway component analysis. *IEEE Signal Process. Mag.* **2015**, *32*, 145–163.

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

(13) De Lathauwer, L.; De Moor, B.; Vandewalle, J. On the Best Rank-1 and Rank-(R_1, R_2, ..., R_N) Approximation of Higher Order Tensors. *SIAM J. Matrix Anal. Appl.* **2000**, *21*, 1324–1342.

(14) Tucker, L. R. et al. The extension of factor analysis to three-dimensional matrices. *Contributions to mathematical psychology* **1964**, *110119*.

(15) Tucker, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika* **1966**, *31*, 279–311.

(16) De Lathauwer, L. Decomposition of a Higher-Order Tensor in Block Terms — Part II: Definitions and Uniqueness. *SIAM J. Matrix Anal. Appl.* **2008**, *30*, 1033–1066.

(17) Domanov, I.; De Lathauwer, L. Canonical polyadic decomposition of third-order tensors: relaxed uniqueness conditions and algebraic algorithm. *Linear Algebra Appl.* **2017**, *513*, 342–375.

(18) Domanov, I.; De Lathauwer, L. On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Overall uniqueness. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 876–903.

(19) Kruskal, J. B. Three-way arrays: Rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics. *Linear Algebra Appl.* **1977**, *18*, 95–138.

(20) Sorber, L.; Van Barel, M.; De Lathauwer, L. Optimization-Based Algorithms for Tensor Decompositions: Canonical Polyadic Decomposition, Decomposition in Rank-$(L_r, L_r, 1)$ Terms, and a New Generalization. *SIAM J. Optim.* **2013**, *23*, 695–720.

(21) Vervliet, N.; De Lathauwer, L. Numerical optimization-based algorithms for data fusion. *Data Handling in Science and Technology* **2019**, *33*, 81–128.

(22) Leurgans, S.; Ross, R. Multilinear models: applications in spectroscopy. *Statistical Science* **1992**, 289–310.

(23) Leurgans, S.; Ross, R.; Abel, R. A decomposition for three-way arrays. *SIAM J. Matrix Anal. Appl.* **1993**, *14*, 1064–1083.

(24) De Lathauwer, L. Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank-$(L_r, L_r, 1)$ Terms. *SIAM J. Matrix Anal. Appl.* **2011**, *32*, 1451–1474.

(25) Sørensen, M.; De Lathauwer, L. Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-$(L_{r,n}, L_{r,n}, 1)$ terms — Part I: Uniqueness. *SIAM J. Matrix Anal. Appl.* **2015**, *36*, 496–522.

(26) Domanov, I.; Vervliet, N.; De Lathauwer, L. *Decomposition of a tensor into multilinear rank-*$(1, L_r, L_r)$ *terms*; tech. rep. 18-52; ESAT-STADIUS, KU Leuven, Belgium, 2018.

(27) De Lathauwer, L.; Nion, D. Decompositions of a Higher-Order Tensor in Block Terms — Part III: Alternating least squares algorithms. *SIAM J. Matrix Anal. Appl.* **2008**, *30*, 1067–1083.

(28) Bro, R.; Harshman, R. A.; Sidiropoulos, N. D.; Lundy, M. E. Modeling multi-way data with linearly dependent loadings. *Journal of Chemometrics* **2009**, *23*, 324–340.

(29) Debals, O.; De Lathauwer, L. Stochastic and deterministic tensorization for blind signal separation. *International Conference on Latent Variable Analysis and Signal Separation* **2015**, 3–13.

(30) Debals, O.; De Lathauwer, L. *The concept of tensorization*; tech. rep. 17-99; ESAT-STADIUS, KU Leuven, Belgium, 2017.

(31) Debals, O.; Van Barel, M.; De Lathauwer, L. Löwner-based blind signal separation of rational functions with applications. *IEEE Trans. Signal Process.* **2016**, *64*, 1909–1918.

(32) Domanov, I.; Vervliet, N.; De Lathauwer, L. *Decomposition of a tensor into multilinear rank* $(M_r, N_r, \cdot)$ *terms*; tech. rep. 18-51; ESAT-STADIUS, KU Leuven, Belgium, 2018.

(33) Vervliet, N.; Debals, O.; Sorber, L.; Van Barel, M.; De Lathauwer, L. Tensorlab 3.0., http://www.tensorlab.net, 2016.

(34) Vervliet, N.; Vandecappelle, M.; Bousse, M.; Zink, R.; De Lathauwer, L. *Recent numerical and conceptual advances for tensor decompositions — A preview of Tensorlab 4.0*; tech. rep. 19-45; ESAT-STADIUS, KU Leuven, Belgium, 2019.

(35) Björck, A.; Golub, G. H. Numerical methods for computing angles between linear subspaces. *Mathematics of computation* **1973**, *27*, 579–594.

(36) Knyazev, A. V.; Argentati, M. E. Principal angles between subspaces in an A-based scalar product: algorithms and perturbation estimates. *SIAM Journal on Scientific Computing* **2002**, *23*, 2008–2040.

(37) Boussé, M.; Debals, O.; De Lathauwer, L. A tensor-based method for large-scale blind source separation using segmentation. *IEEE Trans. Signal Process.* **2017**, *65*, 346–358.

(38) Domanov, I.; De Lathauwer, L. On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part I: Basic results and uniqueness of one factor matrix. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 855–875.

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