Statistical Papers (2025) 66:91
https://doi.org/10.1007/s00362-025-01714-w

REGULAR ARTICLE

Deflation properties in tensor-based eye blink removal algorithm

Zuzana Rošťáková[1] · **Roman Rosipal**[1,2]

Received: 30 October 2024 / Revised: 5 March 2025

Abstract
Blind source separation refers to a set of techniques designed to uncover latent (i.e. directly unobservable) structures in data. Depending on user preferences and the chosen algorithm, latent components can be estimated either simultaneously or iteratively, one at a time. The latter approach is typically performed using component deflation. However, Camacho et al. (Chemom Intell Lab Syst 208:104212, 2021) showed that deflation can introduce spurious artefacts into the data, particularly when the latent components are estimated under constraints. This study explored the theoretical properties of deflation in the context of higher-order arrays and tensor decomposition. In certain cases, the tensor latent components may represent noise and must be removed before further decomposition to accurately reveal the underlying structure of the data. Building on the ideas presented in Camacho et al. (Chemom Intell Lab Syst 208:104212, 2021), we investigated whether specific forms of deflation can generate spurious artefacts in electroencephalogram (EEG) tensor data, particularly under nonnegativity or unimodality constraints, where orthogonality may lack a natural interpretation. Our results are demonstrated using two real EEG datasets and one simulated dataset.

Keywords Deflation · Component subtraction · Projection · Tensor decomposition · Eye blink removal · Electroencephalogram

Zuzana Rošťáková and Roman Rosipal have contributed equally to this work.

✉ Zuzana Rošťáková
zuzana.rostakova@savba.sk

Roman Rosipal
roman.rosipal@savba.sk

[1] Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 84104 Bratislava, Slovak Republic

[2] Pacific Development and Technology, LLC, 408 Hill Street, Capitola, CA 95010, USA

Published online: 08 May 2025

1 Introduction

Blind source separation (BSS) methods are powerful techniques for identifying hidden (latent) structures in both matrix data and higher-order arrays (tensors). Depending on the chosen algorithm, the latent components can be estimated simultaneously or iteratively.

The latter approach is referred to in the literature as deflation. Deflation techniques are used in the partial least squares (PLS) regression, independent component analysis, sparse principal component analysis (sPCA), and other approaches where the simultaneous estimation of all latent components is not possible. In this case, the components are estimated sequentially, one or more at a time, and then deflated from the data to create a new data matrix for the next estimation. Höskuldsson (1988) proved the advantage of deflating only one component at a time in PLS using the properties of singular value decomposition.

However, Camacho et al. (2021) showed that deflation within sPCA can introduce new directions of variability in the original data. Despite the results being formulated in the context of sparsity constraints, they can be generalised to an arbitrary situation in which the components should satisfy different restrictions, for example, nonnegativity or unimodality, assuming that the model structure and deflation step follow the same principles as in Camacho et al. (2021). Similar problems with different types of deflations were discussed by Mackey (2008) and Witten et al. (2009). In both studies, alternative ways of orthogonalization were proposed to avoid introducing possible artefacts during deflation.

Regarding higher-order arrays, latent components are preferably estimated simultaneously, for example, by the so-called alternating least squares algorithm, in the widely used CANDECOMP/PARAFAC (CP) tensor decomposition (Harshman 1970; Carroll and Chang 1970) or in the Tucker model (Tucker 1966) because successive component estimation results in a model with poorer fit (Bro 1997). An exception is the algorithm by Phan et al. (2015), who proposed a novel component estimation for CP based on deflation and tensor rank reduction.

However, deflation can also be used for purposes other than component estimation. Acar et al. (2007) used the Tucker model to visually identify artefact-related latent components, particularly eye blinks, in an electroencephalogram (EEG) tensor. Once the selected components were deflated from the tensor, a second round of tensor decomposition was performed on the artefact-free data to successfully detect the sources of epileptic activity. This study also inspired our eye blink removal algorithm, which combines nonnegative CP, component deflation, and spectrum-to-signal transformation (Rošťáková et al. 2025).

In both studies, the analyses were performed without discussing the possible impact of deflation on the data structure. This topic is addressed in this study. Because most tensor decomposition models can be expressed in matricized form, the ideas and proposed metrics from Camacho et al. (2021) can be easily adapted for tensor models.

Among the various combinations of tensor decomposition models, deflation methods, and data types, we focus on deflation following nonnegative CP decomposition applied to human EEG data. This specific focus was motivated by several factors. Compared with more general tensor decomposition approaches, the CP model offers

distinct advantages, including easier interpretability and uniqueness. Moreover, it has been widely employed in chemometrics, psychometrics, image processing, and neurophysiology (Bro 1997; Cichocki et al. 2009). Among these, the analysis of EEG signals is central to our long-term research interests. Nevertheless, the theoretical framework that we describe for deflation following CP decomposition is presented in a general form. This ensures that our results can be applied or adapted to other types of tensor data.

The tensor decomposition of EEG data produces neurophysiologically interpretable and stable results under nonnegativity and unimodality constraints (Rošťáková et al. 2020; Rosipal et al. 2022). In contrast, while orthogonality may help mitigate the effects of component deflation (Mackey 2008; Witten et al. 2009), it does not align with the natural interpretation of EEG. Similarly, nonnegativity is favoured in fields such as chemometrics, where orthogonality is typically used only for exploratory purposes (Bro 1997).

The article is organised as follows: Sect. 2 provides a concise overview of the key findings from Camacho et al. (2021) and describes the fundamentals of CP decomposition. Section 3 focuses on two deflation techniques in the context of tensor data by adapting the results from Camacho et al. (2021). Real EEG data contaminated by eye blinks and a simulated dataset, as described in Sect. 4, are used in Sect. 5 to evaluate the effects of deflation on the data structure. Finally, the results are discussed and the conclusions are presented in Sects. 6 and 7.

2 Methods

2.1 Notation and basic tensor operations

In this study, a tensor or an N-way array ($N \in \mathbb{N}$, $N \geq 3$) is denoted by the underlined uppercase letter $\underline{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. Although our primary focus is on the three-way tensor $\underline{X} \in \mathbb{R}^{I \times J \times K}$ with the ijkth element denoted as $\underline{X}_{ijk}, i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K$, the formulas presented below also hold for higher-order arrays.

The matrices are denoted by the uppercase letters $A \in \mathbb{R}^{J_1 \times J_2}$, vectors by the bold lowercase letters $\mathbf{a} \in \mathbb{R}^K$ and scalars by the lowercase letters $a \in \mathbb{R}$. The $M \times M$ identity matrix is denoted by $\mathbb{I}_M$, and $\mathbf{e}_l \in \mathbb{R}^M$ represents a unit vector with one at the lth position and zeros elsewhere. A zero vector of length $J \in \mathbb{N}$ is denoted by $\mathbf{0}_J$, whereas $\mathbf{1}_I$ represents a vector of ones with length I. The Moore-Penrose pseudoinverse and transpose of the matrix $Y \in \mathbb{R}^{N \times M}$ are denoted by Y^+ and $Y^\top$, respectively.

The definitions of the Kronecker ($A \otimes B$) and Khatri-Rao or column-wise Kronecker products ($A \odot C$) follows (Cichocki et al. 2009, pp. 35–36). The tensor-matrix product in the nth mode $\underline{X} \times_n A$ (Cichocki et al. 2009) of a tensor $\underline{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ and a matrix $A \in \mathbb{R}^{J \times I_n}$ yields a tensor $\underline{Y} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ with a general element

$$\underline{Y}_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} \underline{X}_{i_1 \ldots i_{n-1} i_n i_{n+1} \ldots i_N} A_{j i_n}.$$

The outer product $\mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$ of three vectors $\mathbf{a} \in \mathbb{R}^I$, $\mathbf{b} \in \mathbb{R}^J$, and $\mathbf{c} \in \mathbb{R}^K$ results in a three-way tensor $\underline{X} \in \mathbb{R}^{I \times J \times K}$ with a general element $\underline{X}_{ijk} = a_i b_j c_k$ (Cichocki et al. 2009).

Fixing all but one index in a tensor yields a tensor fibre (Cichocki et al. 2009). For example, a mode-1 fibre is a vector $X_{.jk} = \left(\underline{X}_{1jk} \underline{X}_{2jk} \ldots \underline{X}_{Ijk}\right)^\top \in \mathbb{R}^I$. Matricization of a tensor in the first mode refers to the transformation of the tensor $\underline{X} \in \mathbb{R}^{I \times J \times K}$ into matrix

$$X_{(1)} = (X_{.11}\ X_{.21}\ \ldots\ X_{.J1} X_{.12} \ldots X_{.J2} \ldots X_{.1K} \ldots X_{.JK}) \in \mathbb{R}^{I \times JK}$$

by concatenating the mode-1 fibres as matrix columns (Cichocki et al. 2009). Matricization in the second and third modes results in the matrices $X_{(2)} \in \mathbb{R}^{J \times IK}$ and $X_{(3)} \in \mathbb{R}^{K \times IJ}$.

The vector space spanned by the columns of the matrix $Y \in \mathbb{R}^{N \times M}$ is denoted by $C(Y)$ and is defined as

$$C(Y) = \left\{\mathbf{x} \in \mathbb{R}^N : \mathbf{x} = Y\mathbf{z}, \mathbf{z} \in \mathbb{R}^M\right\}, \quad C(Y) \subseteq \mathbb{R}^N. \tag{1}$$

Similarly, $C\left(Y^\top\right)$ denotes the vector space spanned by the rows of Y and is defined as

$$C(Y^\top) = \left\{\mathbf{c} \in \mathbb{R}^M : \mathbf{c} = Y^\top \mathbf{d}, \mathbf{d} \in \mathbb{R}^N\right\}, \quad C(Y^\top) \subseteq \mathbb{R}^M. \tag{2}$$

2.2 Deflation in sparse PCA

In this section, we consider a two-dimensional space for a while before dealing with tensors and EEG signal. Let us consider the original data matrix $Y_1 \in \mathbb{R}^{N \times M}$ which should be decomposed into sparse principal components (PCs). Several algorithms use a sequential deflation approach to estimate sparse PCs (Camacho et al. 2021). In other words, the loading $\mathbf{p}_1 \in \mathbb{R}^M$ and scores $\mathbf{t}_1 \in \mathbb{R}^N$ of the first sparse PC are removed from the data in the following step, producing a "new" data matrix $Y_2 = Y_1 - \mathbf{t}_1 \mathbf{p}_1^\top$. The second component score $\mathbf{t}_2 \in \mathbb{R}^N$ and its corresponding loading $\mathbf{p}_2 \in \mathbb{R}^M$ are then estimated from Y_2. In general, $\mathbf{t}_{n+1}, \mathbf{p}_{n+1}$ are estimated after n deflation steps from the matrix $Y_{n+1} = Y_n - \mathbf{t}_n \mathbf{p}_n^\top, n \in \mathbb{N}$.

However, Camacho et al. (2021) demonstrated that this deflation procedure can introduce new directions of variability in the data when performing sPCA. They made the following two assumptions:

- Y_1 is of full rank, so $rank(Y_1) = \min(M, N)$,

- $\mathbf{t}_n = Y_n \mathbf{q}_n$, $\mathbf{q}_n \in \mathbb{R}^M$. In other words, the vector $\mathbf{t}_n$ must be a linear combination of the Y_n columns in each deflation step. For example, $\mathbf{t}_n = Y_n \mathbf{p}_n \left(\mathbf{p}_n^\top \mathbf{p}_n\right)^{-1}$ in projection deflation with $\mathbf{q}_n = \mathbf{p}_n \left(\mathbf{p}_n^\top \mathbf{p}_n\right)^{-1}$.

Under these assumptions,

$$Y_{n+1} = Y_n - \mathbf{t}_n \mathbf{p}_n^\top = Y_n - Y_n \mathbf{q}_n \mathbf{p}_n^\top = Y_n \left(\mathbb{I}_M - \mathbf{q}_n \mathbf{p}_n^\top\right).$$

Let $P_i = \mathbb{I}_M - \mathbf{q}_i \mathbf{p}_i^\top$; then,

$$Y_{n+1} = Y_n P_n = Y_{n-1} P_{n-1} P_n = \cdots = Y_1 P_1 \ldots P_n, \tag{3}$$

and the subsequent component score can be expressed as

$$\mathbf{t}_{n+1} = Y_{n+1} \mathbf{q}_{n+1} = Y_1 \mathbf{q}_{n+1}^\star, \quad \text{where } \mathbf{q}_{n+1}^\star = P_1 P_2 \ldots P_n q_{n+1}.$$

Thus, $\mathbf{t}_{n+1} \in C(Y_1)$ for all $n \in \mathbb{N}$ because it is directly expressed as a linear combination of Y_1 columns.

However, we must distinguish between two cases for the vector $\mathbf{p}_{n+1}$:

- $M \leq N$
  In this case, $rank(Y_1) = M$ and consequently $C(Y_1^\top) = \mathbb{R}^M$ (Camacho et al. 2021). Therefore, any sparse vector $\mathbf{p}_{n+1}$ belongs to $C(Y_1^\top)$.
- $M > N$
  If Y_1 has more columns than rows, $rank(Y_1) = N < M$ and $C(Y_1^\top) \subset \mathbb{R}^M$. If $\mathbf{p}_{n+1}$ is estimated by ordinary least squares or another algorithm expressing $\mathbf{p}_{n+1}$ as a linear combination of the Y_{n+1} rows, then from the equation (3) we can conclude that $\mathbf{p}_{n+1} \in C(Y_1^\top)$. However, $\mathbf{p}_{n+1}$ is restricted to be sparse in sPCA, thus $\mathbf{p}_{n+1} \in C(Y_1^\top)$ is not guaranteed. This means that new directions of variability may be introduced in the data through deflation, potentially becoming the next component extracted from Y_{n+2}.

Camacho et al. (2021) proposed several diagnostic tools for assessing the distortion introduced by deflation, particularly when $M > N$. In this study, we focus on the percentage of artefacts contaminating the components, denoted $PercA$:

$$\begin{aligned} PercA_{n+1} &= \frac{tr(O_{n+1}^\top O_{n+1})}{tr(\mathbf{p}_{n+1} \mathbf{t}_{n+1}^\top \mathbf{t}_{n+1} \mathbf{p}_{n+1}^\top)} \times 100\ [\%], \\ O_{n+1} &= (\mathbb{I}_M - Y_1^\top (Y_1^\top)^+)(Y_n^\top (Y_n^\top)^+) \mathbf{p}_{n+1} \mathbf{t}_{n+1}^\top. \end{aligned} \tag{4}$$

In other words, $PercA$ measures the amount of variance in a component that is not related to the original data, indicating the amount of distortion introduced during n deflation steps.

It is evident that $PercA = 0$ when $M \leq N$. However, Camacho et al. (2021) demonstrated that deflation can still introduce new directions of variability in this case, especially when the component loadings overlap. Unfortunately, no efficient method has yet been proposed to evaluate the impact of deflation when $M \leq N$.

2.3 CANDECOMP/PARAFAC

The CP decomposition of a tensor $\underline{X} \in \mathbb{R}^{I \times J \times K}$ into F latent components follows the formula

$$\underline{X} = \underline{\Lambda} \times_1 A \times_2 B \times_3 C + \underline{E} = \sum_{f=1}^{F} \lambda_f \mathbf{a}_f \circ \mathbf{b}_f \circ \mathbf{c}_f + \underline{E}, \tag{5}$$

where $\underline{E} \in \mathbb{R}^{I \times J \times K}$ represents the model error, and $\underline{\Lambda} \in \mathbb{R}^{F \times F \times F}$ is a tensor with nonzero elements only along its super-diagonal - that is, $\underline{\Lambda}_{ijk} \neq 0$ if and only if $i = j = k,\ i, j, k \in \{1, \dots F\}$.

The columns of the component matrices $A = (\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_F) \in \mathbb{R}^{I \times F}$, $B = (\mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_F) \in \mathbb{R}^{J \times F}$, and $C = (\mathbf{c}_1 \mathbf{c}_2 \dots \mathbf{c}_F) \in \mathbb{R}^{K \times F}$ represent the component signatures in the first, second, and the third modes, respectively. To avoid potential scaling ambiguities, component signatures are usually normalised to the unit norm. The weight of the f^{th}, $f = 1, \dots, F$ component is stored in the corresponding fth superdiagonal element of $\underline{\Lambda}$, denoted by λ_f.

CP model (5) can also be expressed using the following matricized forms:

$$X_{(1)}^{\top} = (C \odot B)\Lambda^{\star} A^{\top} + E_{(1)}^{\top}, \tag{6}$$

$$X_{(2)}^{\top} = (A \odot C)\Lambda^{\star} B^{\top} + E_{(2)}^{\top}, \tag{7}$$

$$X_{(3)}^{\top} = (B \odot A)\Lambda^{\star} C^{\top} + E_{(3)}^{\top}, \tag{8}$$

where

$$\Lambda^{\star} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_F \end{pmatrix}.$$

The alternating least squares (ALS) algorithm is a widely used method for estimating component matrices in tensor decomposition. In each iteration, two component matrices are fixed, whereas the third is estimated using the least-squares criterion. Specifically, A is estimated from (6) with B and C fixed; subsequently, B is estimated from (7) with A and C fixed. Finally, C is estimated from (8) with A and B fixed. These steps are repeated until the stopping criteria are met.

Furthermore, the interpretability and stability of the decomposition can be enhanced by applying nonnegativity or unimodality constraints to specific modes (Rosipal et al. 2022). As employed in Acar et al. (2007), orthogonality constraints can simplify computational issues; however, they lack neurophysiological interpretation in EEG tensor decomposition.

3 Selected deflation methods for tensor data and their theoretical properties

In line with Rošťáková et al. (2025) and Acar et al. (2007), we considered two deflation techniques for tensor data: projection to the nullspace of the space spanned by the selected component signatures and component subtraction. Both deflation approaches were used for EEG data analysis. In Acar et al. (2007), latent components representing artefacts were removed from the tensor using projection deflation, and the deflated tensor formed an input for subsequent tensor decomposition. The SPECTER algorithm (Rošťáková et al. 2025) was proposed to detect and remove eye blink-related components from EEG data. Although it does not directly apply a second CP decomposition on a deflated tensor, this subsequent analysis is natural due to the interest in the "true" data latent structure after removing artefacts. However, the possible influence of component deflation has not been discussed, and this study aims to fill this gap.

3.1 Projection to the nullspace of the space spanned by selected component signatures

We begin with examining the projection to the nullspace of the space spanned by the mode-2 signatures. Projection to the nullspace of the space spanned by the mode-1 or mode-3 signatures can be conducted similarly. Although combining projections across two or more modes simultaneously is feasible, this study focused on a single-mode projection during each deflation step.

First, we considered the removal of a single component. Because the order of the components can be permuted by rearranging the corresponding columns of the component matrices, let us remove the first component with signature $\mathbf{b}_1$. The matrix $Q_B = \mathbb{I}_J - \mathbf{b}_1(\mathbf{b}_1^\top \mathbf{b}_1)^{-1}\mathbf{b}_1{}^\top$ is a projection matrix[1] to the nullspace of the space spanned by $\mathbf{b}_1$. After some algebra, the multiplication of the tensor $\underline{X}$ by Q_B in the second mode results in

$$\underline{X}_{new} = \underline{X} \times_2 Q_B = \sum_{f=2}^{F} \mathbf{a}_f \circ (Q_B \mathbf{b}_f) \circ \mathbf{c}_f + \underline{E} \times_2 Q_B. \tag{9}$$

Clearly, the first component is removed, because $Q_B \mathbf{b}_1 = \mathbf{0}_J$.

If the columns of B are mutually orthogonal,[2] then $Q_B \mathbf{b}_j = \mathbf{b}_j$, $j = 2, \ldots, J$ indicating that the remaining component signatures are unaffected. Alternatively, the variability captured by $\mathbf{b}_1$ is "removed" from the other mode-2 signatures when component signatures (the columns of component matrices) are considered nonnegative, unimodal, or simply non-orthogonal. However, such modifications may inadvertently introduce variability that does not correlate with the original data.

[1] Since $\|\mathbf{b}_1\| = 1$, Q_B can be equivalently expressed as $Q_B = \mathbb{I}_J - \mathbf{b}_1\mathbf{b}_1^\top$.

[2] In this case, it suffices that only $\mathbf{b}_1$ is orthogonal to all $\mathbf{b}_j$, $j = 2, \ldots, J$. However, in practice, the same constraints are typically assumed for all the signatures in a given mode.

In the following analysis, it is advantageous to express equation (9) in the matricized form:

$$\begin{aligned} X_{new(2)}^{\top} &= X_{(2)}^{\top} Q_B = X_{(2)}^{\top}(\mathbb{I}_J - \mathbf{b}_1(\mathbf{b}_1^{\top}\mathbf{b}_1)^{-1}\mathbf{b}_1^{\top}) \\ &= X_{(2)}^{\top} - X_{(2)}^{\top}\mathbf{b}_1(\mathbf{b}_1^{\top}\mathbf{b}_1)^{-1}\mathbf{b}_1^{\top} \\ &= Y - \mathbf{t}\mathbf{p}^{\top}, \end{aligned}$$

where $Y = X_{(2)}^{\top}$, $\mathbf{p} = \mathbf{b}_1$ and $\mathbf{t} = X_{(2)}^{\top}\mathbf{b}_1(\mathbf{b}_1^{\top}\mathbf{b}_1)^{-1}$. This formulation precisely aligns with the deflation model presented by Camacho et al. (2021). Moreover, the assumption $\mathbf{t} \in C(Y) = C(X_{(2)}^{\top})$ holds because $\mathbf{t} = X_{(2)}^{\top}\mathbf{b}_1(\mathbf{b}_1^{\top}\mathbf{b}_1)^{-1}$.

Let us now consider a situation in which multiple components are removed in a single step. Without loss of generality, we may permute the columns of the component matrices such that their first G columns represent the signatures of the components selected for removal. In this case, matrix B can be divided into the following two parts:

$$B = (B^{out}B^{in}), \quad B^{out} = (\mathbf{b}_1 \ldots \mathbf{b}_G) \in \mathbb{R}^{J\times G}, \quad B^{in} = (\mathbf{b}_{G+1} \ldots \mathbf{b}_F) \in \mathbb{R}^{J\times(F-G)}.$$

The corresponding projection matrix is given by

$$Q_B = (\mathbb{I}_J - B_{out}(B_{out}^{\top}B_{out})^{-1}B_{out}^{\top})^{\top}.$$

Therefore, we can express

$$\begin{aligned} X_{new(2)}^{\top} &= X_{(2)}^{\top} Q_B^{\top} = X_{(2)}^{\top}(\mathbb{I}_J - B_{out}(B_{out}^{\top}B_{out})^{-1}B_{out}^{\top})^{\top} \\ &= X_{(2)}^{\top} - X_{(2)}^{\top}B_{out}(B_{out}^{\top}B_{out})^{-1}B_{out}^{\top} = Y - \sum_{l=1}^{G}\mathbf{t}_l\mathbf{p}_l^{\top}, \end{aligned}$$

where $Y = X_{(2)}^{\top}$, $\mathbf{t}_l = X_{(2)}^{\top}B_{out}(B_{out}^{\top}B_{out})^{-1}\mathbf{e}_l$ and $\mathbf{p}_l$ is the lth column of B_{out}, $l = 1 \ldots, G$. This is also in line with the assumptions and deflation model of Camacho et al. (2021).

A similar situation arises when projecting to the nullspace of the space spanned by the mode-3 signatures represented by the columns of matrix C from the CP model (5). In this case, $Y = X_{(3)}^{\top}$, where $N = IJ$ and $M = K$. Finally, $Y = X_{(1)}^{\top} \in \mathbb{R}^{JK\times I}$ with $N = JK$ and $M = I$ for projection to the nullspace of the space spanned by the mode-1 signatures.

3.2 Component subtraction

Let S denotes an index set of components selected for removal. Component subtraction follows the formula:

$$\underline{X}_{new} = \underline{X} - \sum_{f\in S}\lambda_s\mathbf{a}_f \circ \mathbf{b}_f \circ \mathbf{c}_f.$$

First, we consider subtraction of a single component with the signatures $\mathbf{a}_s, \mathbf{b}_s, \mathbf{c}_s, s \in \{1, \ldots, F\}$. This operation can be expressed in the matricized form as

$$\underline{X}_{new} = \underline{X} - \lambda_s \mathbf{a}_s \circ \mathbf{b}_s \circ \mathbf{c}_s \rightarrow \left.\begin{array}{rcl} X^{\top}_{new(1)} = & X^{\top}_{(1)} - \lambda_s \left(\mathbf{c}_s \otimes \mathbf{b}_s\right) \mathbf{a}_s^{\top} \\ X^{\top}_{new(2)} = & X^{\top}_{(2)} - \lambda_s \left(\mathbf{a}_s \otimes \mathbf{c}_s\right) \mathbf{b}_s^{\top} \\ X^{\top}_{new(3)} = & X^{\top}_{(3)} - \lambda_s \left(\mathbf{b}_s \otimes \mathbf{a}_s\right) \mathbf{c}_s^{\top} \end{array}\right\} = Y - \mathbf{t}\mathbf{p}^{\top}. \tag{10}$$

That is, component subtraction adheres to the deflation scheme in Camacho et al. (2021).

In contrast to projection deflation, where the vector $\mathbf{t}$ is expressed as a linear combination of $X^{\top}_{(1)}$, $X^{\top}_{(2)}$ or $X^{\top}_{(3)}$ columns, $\mathbf{t}$ is equal to the Kronecker product of the signatures from the two modes in component subtraction. For example, consider the matricization in the second mode, where $Y = X^{\top}_{(2)}$, $\mathbf{p} = \mathbf{b}_s$, and $\mathbf{t} = \lambda_s(\mathbf{a}_s \otimes \mathbf{c}_s)$. Therefore, component subtraction does not ensure that $\mathbf{t}$ belongs to $C(X^{\top}_{(2)})$.

Theoretically, it is possible to test whether $\mathbf{t} \in C(X^{\top}_{(2)})$ and to compute the amount of spurious artefacts if $\mathbf{t} \notin C(X^{\top}_{(2)})$, using the theoretical results for $\mathbf{p}$ in a transposed version of the model. However, as demonstrated in Sect. 5, this may present significant challenges in real-world data situations. For example, when calculating $PercA$ for $\mathbf{p} = \mathbf{b}_s$, matrix $O^{\top}_{n+1} O_{n+1}$ in Eq. (4) is of size $J \times J$. By contrast, if we compute $PercA$ for $\mathbf{t} = \lambda_s \left(\mathbf{a}_s \otimes \mathbf{c}_s\right)$, then $O^{\top}_{n+1} O_{n+1} \in \mathbb{R}^{IK \times IK}$. Suppose I and K are at most in hundreds. In this case, IK may be in thousands or tens of thousands, and computing $PercA$ for $\mathbf{t}$ on a standard computer is either computationally infeasible or numerically unstable owing to its high dimensionality.

4 EEG tensor data

4.1 Eye blink corrupted EEG data

In this study, we used real EEG datasets contaminated with eye blinks from Rošťáková et al. (2025). For a detailed description, please refer to the original article and the references therein.

Dataset 1 consisted of 11 EEG recordings from three healthy male volunteers measured under eyes-open condition before, during, and after meditation. Each EEG recording lasted approximately one minute before (SubId_pre) and after (SubId_post) meditation, and several minutes during meditation (SubId_med),[3] where Id denotes volunteer identifier (1, 2, or 3) (Rošťáková et al. 2025).

Dataset 2 included data for 45 subjects from the OSF EEG eye artefact dataset, which is publicly available at https://osf.io/2qgrd/.

EEG data analysis can be divided into three parts:

[3] Subject 1 took part in three separate meditation sessions denoted here as Sub1_med1, Sub1_med2, Sub1_med3.

Tensor construction and preprocessing For each subject from both datasets, the EEG signal can be represented by a matrix $X_{EEG} \in \mathbb{R}^{L \times J}$, where J is the number of electrodes, and $L = S_f \times T$ is the number of time points at which the EEG signal is recorded. The sampling frequency in Hertz (the number of points recorded per second) is denoted by S_f, and T represents the duration of the recording time interval in seconds. The tensor construction and preprocessing were performed using the following steps in MATLAB® (The Mathworks, Inc. 2021):

1. *Tensor construction* The EEG signal is divided into overlapping time windows for each electrode separately. In line with Rošťáková et al. (2025), we used 0.5-second time windows with a 0.4-s overlap. Additionally, for each electrode and time window separately, the frequency spectrum is computed using the Fast Fourier Transform (FFT). The resulting amplitude spectra are concatenated into a tensor $\underline{X}_{raw} \in \mathbb{R}^{I \times J \times K}$ where:
   - I is the number of time windows equal to
   $$I = \left\lfloor \frac{L - winLen \times S_f}{winLen \times S_f - \lfloor winOverlap \times S_f \rfloor} \right\rfloor,$$
   where $winLen$ is the time window length, and $winOverlap$ is the length of overlap, both in seconds. For example, for a one-minute EEG signal measured with $S_f = 128$ Hz, divided into 0.5-second time windows with a 0.4-second overlap, the number of time windows is $I = \left\lfloor \frac{60 \times 128 - 0.5 \times 128}{0.5 \times 128 - \lfloor 0.4 \times 128 \rfloor} \right\rfloor = 585$.
   - K is the number of frequency points in which the spectrum was computed. According to the Nyquist theorem, we can compute the frequency spectrum up to $\frac{S_f}{2}$ Hz with a resolution of $\frac{1}{winLen}$ Hz. However, focusing only on a specific frequency range is often helpful. For example, we considered the frequency interval 0–30 Hz with a resolution of 2 Hz ($winLen = 0.5$), giving $K = \frac{30 - 0}{2} + 1 = 16$.
   - J represents the number of electrodes. Originally, the EEG signal was recorded by 64 electrodes. However, using the information from all 64 electrodes would lead to the situation $M \leq N$ for EEG recordings shorter than 1.7 min. [4] Consequently, $PercA$ equals zero by default in all the deflation approaches considered, as explained in Sect. 2.2, for half of our data. To be consistent across all datasets and subjects, we used only 19 electrodes as in Rošťáková et al. (2025), allowing us to assess the proportion of new variability introduced by deflation through $PercA$.
2. *The logarithmic transform* Once X_{raw} is constructed, the element-wise $\log_{10}$-transform is applied, resulting in the tensor
$$\underline{X} = \log_{10}(\underline{X}_{raw} + 1).$$

[4] EEG recording lasting 102 s (1.7 min) results in $M = I = 999$. If $J = 64$ and $K = 16$, then $N = JK = 64.16 = 1024$, so $M < N$.

Two main reasons exist for this transformation. First, previous studies have shown interpretational and computational advantages of using the log-transformed tensor over the raw tensor in detecting EEG oscillatory rhythms (Rošťáková et al. 2020; Rosipal et al. 2022).
The second, and more significant, reason is discussed in Rošťáková et al. (2025), and we will only touch upon it briefly here. Tensor $\underline{X}_{raw}$ is nonnegative; however, as demonstrated in Rošťáková et al. (2025), any deflation approach may introduce negative values. Consequently, the resulting tensor cannot represent the spectrum of the artefact-free EEG signal. Moreover, the algorithm used to transform the cleaned tensor back into a signal is not equipped to handle negative values (Rošťáková et al. 2025). Nevertheless, this issue can be effectively addressed by employing the log-transformed tensor instead of the original. After performing CP decomposition and deflation, we can apply an inverse transformation ($10^{\underline{X}}$, element-wise) to ensure that the resulting tensor remains nonnegative (Rošťáková et al. 2025).
Notably, in this study, the inverse transformation is not applied, as our interest lies solely in the deflated tensor itself, and transforming it back into an EEG signal is not necessary.

3. *Centring* Before applying PCA or CP, centring the data is recommended. Tensor centring can be performed in one or multiple modes simultaneously. However, Bro (1997) recommends centring across a single mode at a time. In our previous studies (Rošťáková et al. 2020; Rosipal et al. 2022), we preferred centring across the first (temporal) mode:

$$\underline{X}^c_{ijk} = \underline{X}_{ijk} - \frac{1}{I}\sum_{s=1}^{I} \underline{X}_{sjk}. \tag{11}$$

CP decomposition and deflation

1. *CP decomposition* The CP decomposition with the nonnegativity constraints in all three modes was applied to the tensor $\underline{X}^c$ for each subject separately, with the number of components F determined by the *tripleC* algorithm (Rošťáková and Rosipal 2022). In this study, CP decomposition was always performed on the centred tensor $\underline{X}^c$. The interaction between tensor centring and deflation will be further explored in Sect. 4.3.
2. *Component selection* Latent components were automatically inspected for eye blink characteristics like in Rošťáková et al. (2025), followed by manual selection if necessary. Selected components were labeled as artefactual.
3. *Deflation* For each subject, component deflation was performed using either component subtraction or projection to the nullspace of the space spanned by the (i) mode-1, (ii) mode-2, or (iii) mode-3 signatures. Moreover, either

   - a single artefactual component with the highest weight (corresponding λ from the mixing tensor $\underline{\Lambda}$) was used in the deflation step, or
   - all components labelled as artefactual were removed during deflation.

Evaluation of the deflation effect

1. *CP decomposition of deflated data* The second CP decomposition was performed on the deflated tensor, with the number of components set to

$$F_{new} = F - \text{number of removed artefactual components.}$$

   If $F_{new} = 0$, the second CP decomposition was not performed, and the corresponding subject was removed from the analysis.
2. *Proportion of new variance introduced by the deflation* The newly estimated latent components were examined to determine the percentage of variability introduced by the deflation procedures. Camacho et al. (2021) assumed that deflation significantly impacts data variability if $PercA$ is above 20%.

4.2 Simulated EEG data

In addition to real-world datasets, we analysed a simulated dataset from Rošťáková and Rosipal (2022). Although this dataset does not specifically involve eye-blink artefacts, it closely mimics the properties of real EEG signal for 50 "subjects". Furthermore, the known underlying structure allows for a direct assessment of deflation impact by evaluating whether the true remaining components are accurately revealed after deflation, or if any artefactual components are introduced.

The simulated EEG data included four oscillatory rhythms: a 5 Hz rhythm located in the frontal region, two sensorimotor rhythms oscillating at 8 Hz and 14 Hz in the central region symmetrically across both hemispheres, and an occipital alpha rhythm at 11 Hz. Additionally, the data included broadband brain activity (BBA), simulated as fractional Brownian motion with the Hurst exponent $H = 0.6$. In contrast to Rošťáková and Rosipal (2022), only simulated data with low level of BBA were used in this study.[5] The simulated EEG data were generated as if they were recorded with 64 electrodes over a one-minute time interval with a sampling frequency of 256 Hz. For more details on the data generation, see Rošťáková and Rosipal (2022).

Tensor construction and preprocessing, nonnegative CP decomposition, and deflation were performed exactly as in the case of the real data. The only difference was segmenting the EEG signal into two-second time windows with a 1.9-second overlap and focusing on the frequency interval of 4–25 Hz. In the CP decomposition, the nonnegativity in the frequency mode was combined with unimodality constraints because of the focus on narrowband simulated oscillatory rhythms instead of broadband eye blinks.

Tensor $\underline{X}$ had dimensions $I \times J \times K$, where $I = 571$, $J = 64$, and $K = 43$. All three matricized versions of the tensor satisfied the condition $M < N$ from Camacho et al. (2021):

[5] The amplitude of BBA was modulated such that the ratio between the amplitude of BBA and target oscillations was 0.2 (Rošťáková and Rosipal 2022).

$$
\begin{array}{lllll}
 & N & M & & M < N \\
X_{(1)}^{\top} \in \mathbb{R}^{JK \times I} & JK = 2752, & I = 571 & \rightarrow & I < JK, \\
X_{(2)}^{\top} \in \mathbb{R}^{IK \times J} & IK = 24553, & J = 64 & \rightarrow & J < IK, \\
X_{(2)}^{\top} \in \mathbb{R}^{IJ \times K} & IJ = 36544, & K = 43 & \rightarrow & K < IJ.
\end{array}
$$

Because of the presence of BBA, matricized versions of the tensor were of full rank, resulting in $PercA \approx 0$. This implies that the effect of component deflation is not measurable using $PercA$ in this case.

Nevertheless, we aimed to demonstrate the effect of deflation on both the uncentered and centred tensors from a different perspective. First, we applied the CP decomposition to the data and extracted four latent components. Next, the 11 Hz component (if present) was removed from both the uncentered and centred tensors by either component subtraction or projection to the nullspace of the space spanned by selected signatures in a given mode. Finally, a second CP decomposition was applied to the deflated tensor. Ideally, the three detected latent components should represent oscillatory rhythms at 5, 8, and 14 Hz, respectively. However, this was not always the case for all subjects. Therefore, we focused on the proportion of participants for whom:

- the 11 Hz rhythm remained detectable after deflation, and
- only the 5, 8, and 14 Hz oscillations were identified after the second CP.

This procedure was repeated, replacing the 11 Hz rhythm with an 8 Hz or 14 Hz oscillation in the deflation step.

4.3 Component removal from centred and uncentered tensor

In the previous sections, we introduced two component deflation procedures. However, an important question arises: should deflation be applied to the original tensor $\underline{X}$ or its centred version $\underline{X}^c$, given that the components were estimated from the centred tensor?

To address this issue, we examined tensor centring in more details. When we rewrite the centring across the first mode (11) in matrix form using $X_{(1)}^{\top}$, we arrive at the following model:

$$
{X_{(1)}^{c}}^{\top} = X_{(1)}^{\top} - \frac{1}{I} X_{(1)}^{\top} \mathbf{1}_I \mathbf{1}_I^{\top} = X_{(1)}^{\top} (\mathbb{I}_I - \frac{1}{I} \mathbf{1}_I \mathbf{1}_I^{\top}). \tag{12}
$$

The matrix $\mathbb{I}_I - \frac{1}{I} \mathbf{1}_I \mathbf{1}_I^{\top}$ is a projection matrix, implying that centring across the first mode can be understood as a projection in the first mode that removes the influence of the "baseline" vector $\mathbf{1}_I$. Furthermore, the centring process across the first mode, as expressed in Eq. (12), also follows the deflation model $Y - \mathbf{t}\mathbf{p}^{\top}$ where

$$
Y = X_{(1)}^{\top}, \quad \mathbf{t} = \frac{1}{I} X_{(1)}^{\top} \mathbf{1}_I \quad \text{and} \quad \mathbf{p} = \mathbf{1}_I.
$$

In other words, centring across the first mode is also a form of deflation. Moreover, $\mathbf{t} \in C(X_{(1)}^{\top})$, as in Camacho et al. (2021).

Therefore, we considered two cases when analysing the impact of deflation on both real and simulated EEG data:

- *Case 1:* deflation applied to the original tensor $\underline{X}$, i.e.,

$$X_{new(1)}^{\top} = X_{(1)}^{\top} - \mathbf{t}\mathbf{p}^{\top}.$$

- *Case 2:* deflation applied to the centred tensor $\underline{X}^c$, i.e.,

$$X_{new(1)}^{\top} = {X_{(1)}^c}^{\top} - \mathbf{t}\mathbf{p}^{\top} = X_{(1)}^{\top} - \frac{1}{I} X_{(1)}^{\top} \mathbf{1}_I \mathbf{1}_I^{\top} - \mathbf{t}\mathbf{p}^{\top}.$$

The vectors $\mathbf{t}$, $\mathbf{p}$ in deflation have the same meaning as those in the previous sections.

Case 2 can be viewed as deflation applied to $\underline{X}^c$. However, Camacho et al. (2021) assumed that Y_1 was of full rank. When $I < JK$ (i.e. $M \leq N$), the rank of $X_{(1)}$ is I, whereas the rank of the centred tensor $X_{(1)}^c$ is $I - 1$. Consequently, if we denote $Y_1 = {X_{(1)}^c}^{\top}$, the $PercA$ values will be nonzero even in the case $I < JK$ because $C({X_{(1)}^c}^{\top}) \subset \mathbb{R}^I$.

Therefore, we prefer the second interpretation of *Case 2* as two consecutive deflation steps applied to the original tensor $\underline{X}$. In Eq. (4), $Y_1^{\top}$ always corresponds to a matricized versions of the original, uncentered tensor $\underline{X}$.

Finally, we would like to highlight that for each subject, component deflation was performed on both the uncentered and centred versions of the tensor (*Case 1/Case 2*) using either component subtraction, or projection to the nullspace of the space spanned by the (i) mode-1, (ii) mode-2, and (iii) mode-3 signatures. As described in Sect. 4.1, either a single artefactual component with the highest weight or all artefactual components were removed during deflation. The analysis resulted in $2 \times 4 \times 2$ different scenarios.

5 Results

5.1 Eye blink corrupted EEG data

In the first step, we focused on the projection to the nullspace of the space spanned by the mode-2 signatures. In this case, $Y = X_{(2)}^{\top} \in \mathbb{R}^{IK \times J}$, where $N = IK$ and $M = J$ in the model of Camacho et al. (2021). As highlighted in Sect. 4.1, $J = 19$ is much smaller than the number of time windows I in both of our EEG datasets (and also in EEG in general). Consequently, $J < IK$ leads to the first scenario ($M \leq N$) described by Camacho et al. (2021), and the $PercA$ values are theoretically expected to be zero in this case.

A similar situation also occurs for the projection to the nullspace of the space spanned by the mode-3 signatures. Because $Y = X_{(3)}^{\top} \in \mathbb{R}^{IJ \times K}$ and $K = 16$, we are also in the scenario $M \leq N$ with $N = IJ$ and $M = K$. Again, the corresponding $PercA$ values are expected to be zero.

These theoretical expectations were satisfied for both the real-world datasets. For both mode-2 and mode-3 projection deflations, the $PercA$ values for all the remaining latent components across subjects ranged from 10^{-26} to 10^{-23}, which was effectively zero because of the numerical precision of MATLAB. We hypothesise that these negligible deviations from zero are due to numerical inaccuracies in the Moore–Penrose pseudoinverse. Near-zero values were consistently observed across all scenarios and subjects, including both *Case 1* and *Case 2*, and with the removal of a single component or multiple components.

A different situation occurred when considering projection to the nullspace of the space spanned by mode-1 signatures. In this case, $Y = X_{(1)}^{\top} \in \mathbb{R}^{JK \times I}$ and $X_{(1)}$ is of full rank due to the presence of noise in the real data. Now, we are in the situation $N < M$ because $N = JK = 19 \times 16$ and the shortest EEG recording in our datasets exceeded one minute producing $M = I > 500$. In other words, the diagnostic tools from Camacho et al. (2021) can be used to evaluate the impact of deflation.

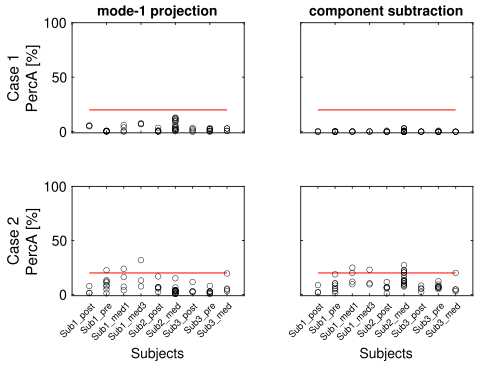
Projection to the nullspace of the space spanned by mode-1 signatures and component subtraction yielded similar results (Figs. 1 and 2). Let us first consider *Case 1*, in which the deflation step was applied to an uncentered tensor. As depicted in Fig. 1 (first rows, right), $PercA$ values were low for the majority of components and subjects, with a maximum of 3.21% in Dataset 1 and 1.62% in Dataset 2 when a single component was subtracted. In contrast, the mode-1 projection deflation of a single component (Fig. 1, first rows, left) resulted in higher maximum $PercA$ values of 12.85% in Dataset 1 and 6.33% in Dataset 2. Similar outcomes were observed when more than one component was removed during deflation (Fig. 2).

Overall, the mode-1 projection led to slightly higher $PercA$ values than the component subtraction in *Case 1* (Figs. 1 and 2, first rows). Nevertheless, for both deflation methods, the $PercA$ values remained below the threshold of 20% when either single or multiple components were removed in a single step. In accordance with Camacho et al. (2021), this indicates that the proportion of the variance introduced by deflation was negligible.

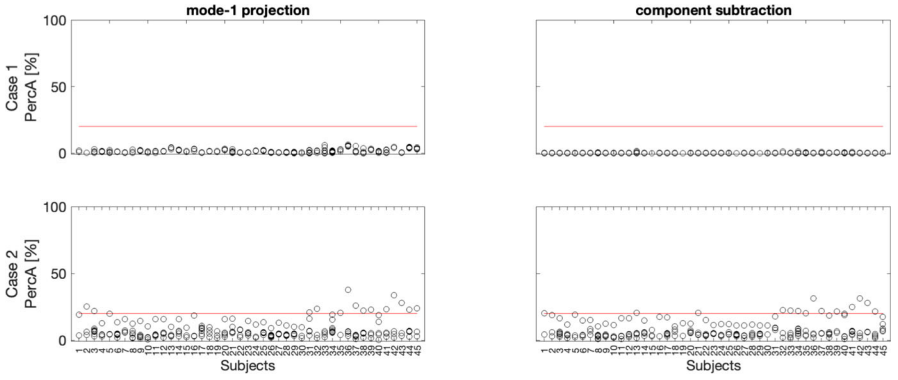
A different situation occurred in *Case 2*, in which the $PercA$ values were significantly higher for both datasets, as confirmed by the Wilcoxon signed-rank test. This was true for both mode-1 projection and component subtraction compared to *Case 1*. In *Case 2*, the maximum $PercA$ reached 27.26% in Dataset 1 and 31.35% in Dataset 2 after subtracting a single component using the component subtraction method. For the mode-1 projection, the $PercA$ values were even higher, peaking at 31.84% for Dataset 1 and 37.66% for Dataset 2 (Fig. 1, second rows). Removing multiple components in a single deflation step in *Case 2*, either by component subtraction or mode-1 projection, resulted in slightly higher $PercA$ values than when only a single component was deflated.

Moreover, $PercA$ for several components exceeded the 20% threshold in both deflation procedures in *Case 2*. Following Camacho et al. (2021), this indicates that the proportion of variance introduced by component subtraction and mode-1 projection to the centred tensor is not negligible, and that the components estimated after deflation may represent spurious artefacts rather than the original data structure.

In line with Sect. 3.2, our goal was to compute $PercA$ for vector $\mathbf{t}$ in the context of component subtraction. Owing to dimensionality constraints, we could only conduct



(a) Dataset 1



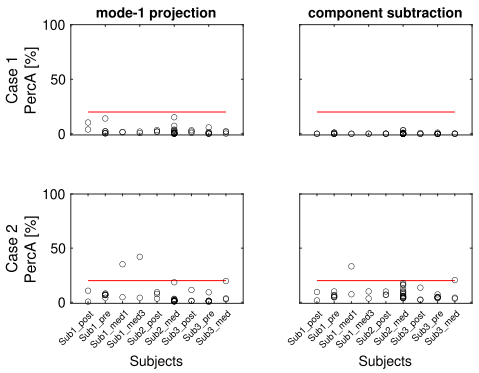
(b) Dataset 2

Fig. 1 Single component deflation. The percentage of variance ($PercA$, black circles) in the latent components that does not correspond to the original data but was introduced by (i) mode-1 projection (left), and (ii) component subtraction (right) applied to both uncentered (*Case 1*) and centred versions of the tensor (*Case 2*) in Datasets 1 and 2. During the deflation step, only the dominant eye blink-related component was removed. If multiple components remained after deflation for a given subject, their corresponding $PercA$ values are displayed stacked vertically

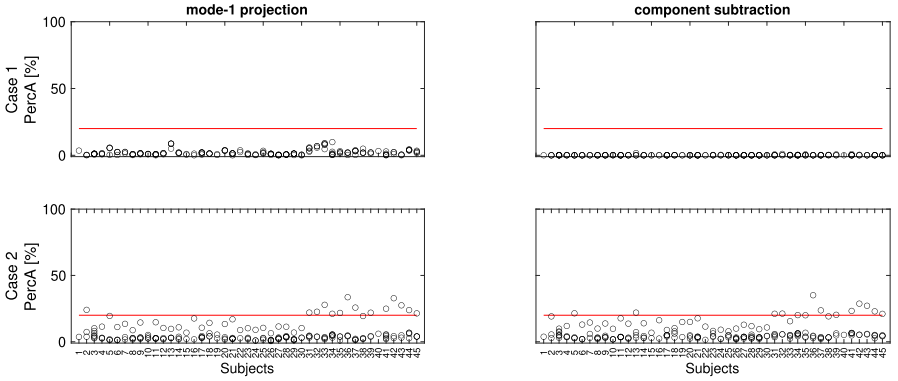
this analysis on a subset of shorter EEG recordings from Dataset 1 (6–7 subjects), utilising a standard computer (OS X, Apple M1 chip, 16 GB of memory).

When considering the component subtraction expressed by matricization in the first mode, we have $\mathbf{t} = (\mathbf{c}_s \otimes \mathbf{b}_s)$ and $Y = X_{(1)} \in \mathbb{R}^{I \times JK}$, where $N = I$ and $M = JK$. In this case, "$M \leq N$", leading to $PercA = 0$.

For matricization in the second mode, we get $\mathbf{t} = (\mathbf{a}_s \otimes \mathbf{c}_s)$ and $Y = X_{(2)} \in \mathbb{R}^{J \times IK}$ where $J < IK$ in our EEG data, indicating the scenario of "$N < M$". In this case, all the subjects exhibited $PercA$ values below 2.5% when component subtraction was performed on an uncentered tensor $\underline{X}$. By contrast, for the centred tensor $\underline{X}^c$, the $PercA$ values were generally higher, exceeding the threshold of 20% in several cases. In addition, similar $PercA$ values for $\mathbf{t} = (\mathbf{b}_s \otimes \mathbf{a}_s)$ were obtained when considering the matricization in the third mode for component subtraction.



(a) Dataset 1



(b) Dataset 2

Fig. 2 Multiple components deflation. The percentage of variance ($PercA$, black circles) in the latent components that does not correspond to the original data but was introduced by (i) mode-1 projection (left) and (ii) component subtraction (right) applied to both uncentered (*Case 1*) and centred versions of the tensor (*Case 2*) in Datasets 1 and 2. During the deflation step, all selected eye blink-related components were removed. If multiple components remained after deflation for a given subject, their corresponding $PercA$ values are displayed stacked vertically

5.2 Simulated data

As an initial step, CP decomposition with four components was applied separately to the simulated EEG data of all 50 subjects. The 11 Hz rhythm was identified in 41 cases (Table 1, upper part).

When applying component subtraction to the uncentered tensor (*Case 1*), the subsequent CP decomposition revealed only three true rhythms (5, 8, and 14 Hz) in 39 subjects, indicating that the 11 Hz rhythm was successfully eliminated. In one subject, the 11 Hz rhythm was still detected after the second CP decomposition (Table 1), whereas in another subject, the second CP decomposition revealed spurious oscillation in place of the true rhythms.

Table 1 Simulated data

	11 Hz				14 Hz				8 Hz			
Component detected (1st CP)	41 out of 50				49 out of 50				50 out of 50			
After component removal (2nd CP)	*Case 1*		*Case 2*		*Case 1*		*Case 2*		*Case 1*		*Case 2*	
Component subtraction	1	39	20	17	49	0	42	5	49	0	49	1
Mode-1 projection	0	36	0	37	0	48	0	40	0	45	1	42
Mode-2 projection	0	32	0	34	0	13	0	11	0	4	1	7
Mode-3 projection	0	39	0	38	0	45	0	40	0	41	0	42

The number of subjects in which the selected rhythm (8, 11, or 14 Hz) was detected following the first CP decomposition is shown in the upper section of the table. The selected rhythm was removed from either the original tensor (*Case 1*) or the centred tensor (*Case 2*) through component subtraction, mode-1, mode-2, or mode-3 projection deflation. The lower section of the table presents the results of the second CP decomposition applied to the deflated tensor. Each table cell contains: (i) the number of subjects in which the selected rhythm remained detectable after deflation (left), and (ii) the number of subjects in which only the three remaining underlying rhythms were detected (right)

In *Case 2*, the 11 Hz oscillation was deflated from the centred tensor. Following component subtraction, we observed that in 20 subjects (49%), the 11 Hz rhythm remained present and was once again detected by the second CP decomposition (Table 1). The true three remaining components were identified in 17 subjects (41%). In the remaining four cases, component subtraction successfully removed the 11 Hz rhythm; however, the second CP decomposition detected one or more artificial oscillations instead of the true ones.

Regarding mode-1, mode-2, or mode-3 projection deflation, the 11 Hz rhythm was completely eliminated in all 41 subjects. The number of subjects in whom the three remaining rhythms were detected by the second CP decomposition ranged from 32 (78%) to 39 (95%) (Table 1). In contrast to component subtraction, the projections applied to the centred and uncentered tensors produced qualitatively similar results.

A different scenario occurred when the 14 Hz or 8 Hz oscillatory rhythm was removed during the deflation step. The 14 Hz component was detected in 49 of the 50 subjects (Table 1). However, component subtraction did not fully eliminated the influence of the 14 Hz rhythm. The following CP decomposition detected a 14 Hz component in all 49 subjects when component subtraction was applied to the uncentered tensor (*Case 1*) and in 42 subjects when the 14 Hz rhythm was subtracted from the centred tensor (*Case 2*). Similarly, the 8 Hz rhythm was identified in all 50 subjects, yet remained detectable by the second CP decomposition in 49 subjects after its subtraction from both the centred and uncentered tensors. From this perspective, component subtraction was not able to eliminate the selected oscillatory components completely.

In contrast, projection deflation of the 14 Hz rhythm led to its successful removal because it was not identified in any subject by the second CP decomposition (Table 1). The true 5, 8, and 11 Hz components were detected in 40 subjects when the projection to the nullspace of the space spanned by mode-1 or mode-3 14 Hz signatures was applied to the centred tensor (*Case 2*, Table 1). In *Case 1*, the three true rhythms were detected in 48 subjects (mode-1 projection) and 45 subjects (mode-3 projection). As depicted in Table 1, mode-2 projection exhibits inferior results, where the three true rhythms were detected in 13 subjects in *Case 1* and 11 subjects in *Case 2*.

Projection deflation successfully removed the 8 Hz rhythm in most subjects, as the second CP decomposition identified the 8 Hz component in at most one subject (for mode-1 and mode-2 projections). The three remaining true components (5, 11, and 14 Hz) were detected in 41 or 45 subjects in *Case 1* and 42 subjects in *Case 2* when considering mode-1 and mode-3 projection deflation. Once again, inferior results were observed for mode-2 deflation, where three true rhythms were detected in only four subjects in *Case 1* and seven subjects in *Case 2*.

6 Discussion

Component deflation plays a crucial role in matrix decomposition, particularly in latent component estimation. However, only a limited number of studies have addressed component deflation in tensor data (Acar et al. 2007; Rošťáková et al. 2025), and none have examined the potential side effects of deflation, namely, the introduction

of spurious artefacts. The primary objective of this study was to address this gap by analysing the presence of spurious artefacts that may arise from deflation following CP decomposition with nonnegativity or unimodality constraints. We focused on two deflation schemes: projection to the nullspace of the space spanned by mode-1, mode-2, or mode-3 signatures and component subtraction.

In sparse PCA or other constrained matrix decomposition models, the column and row spaces of matrix Y, from which the latent components are deflated, played an important role (Camacho et al. 2021). We defined similar "tensor fibre spaces" for tensor-based models such as CP and attempted to adapt the ideas from Camacho et al. (2021). However, as described in Appendix A, the proposed tensor fibre spaces exhibit different properties compared to the row or column space of a matrix, leading to different conclusions from those in Camacho et al. (2021). Due to these obstacles, our research did not follow the path of tensor fibre spaces.

Fortunately, each tensor decomposition model can be reformulated using various matricized forms. Consequently, the ideas and conclusions presented by Camacho et al. (2021) can be directly applied or easily adapted, as demonstrated in Sect. 3. Moreover, the $PercA$ metric proposed in Camacho et al. (2021), as the proportion of artefacts introduced during the deflation step, is also applicable in the tensor case.

The theoretical results presented in Sect. 3 were applied to the real and simulated EEG data. Our focus on EEG data stems from its significance in our long-term research. Moreover, the aforementioned studies addressed deflation within the context of EEG tensor data. Nevertheless, the derived theoretical results are general and can easily be adapted to any real tensor data problem involving component deflation.

In line with the theoretical results, mode-2 and mode-3 projection deflation produced near-zero $PercA$ values in all examined real and simulated EEG data scenarios. Consequently, these two deflation methods may still introduce some artefacts. However, their influence remains undetectable by the $PercA$ measure, not only in our data but also in EEG data in general. This is because the number of electrodes or frequencies considered is often significantly lower than the product of the remaining two tensor dimensions. Consequently, the side effect of deflation can only be assessed for projection to the nullspace of the space spanned by mode-1 signatures and component subtraction.

As centring the tensor data before CP decomposition is generally recommended (Bro 1997), we analysed both situations in which deflation was applied to both centred and uncentered tensors. The proportion of the variance introduced by deflation remained below 20% when the procedure was applied to an uncentered tensor (*Case 1*). Moreover, the $PercA$ values were nearly zero when considering component subtraction from an uncentered tensor. In contrast, mode-1 projection deflation and component subtraction applied to the centred tensor (*Case 2*) resulted in significantly higher $PercA$ values above 20%.

As discussed in Sect. 4.3, the deflation applied to the centred tensor can be interpreted as two consecutive deflation steps on the uncentered tensor. Consequently, the higher $PercA$ values in *Case 2* are consistent with the observations of Camacho et al. (2021), where $PercA$ values increased with the number of deflation steps, as the variance introduced by component removal accumulates.

Based on these observations, and in line with the results of Camacho et al. (2021), we prefer performing deflation on the original uncentered tensor (when possible), as the impact of the introduced spurious artefacts is negligible compared to the deflation applied to the centred tensor.

These conclusions were further validated by analysing of simulated data. When the 11 Hz component was subtracted from the uncentered tensor, subsequent CP decomposition did not detect this rhythm in most subjects. In contrast, the 11 Hz rhythm remained detectable in nearly half of the subjects when subtracted from the centred tensor.

The inability of component subtraction to entirely remove the influence of 8 or 14 Hz rhythms in the simulated data, whether applied to the centred or uncentered tensor, could be attributed to how the data were simulated. The 8 and 14 Hz oscillatory rhythms were simulated near each other in the central region in both hemispheres. Consequently, some spatial information related to the 8 Hz rhythm may still be present in the detected 14 Hz component, preventing component subtraction from fully removing it. The results of projection deflation also support this idea. As highlighted in Sect. 3.1, projection deflation not only removes the selected component(s) but also eliminates related information from the signatures of the remaining components. Following projection deflation, the subsequent CP decomposition detected residual information about the removed 8 or 14 Hz rhythm in a single subject.

However, this property of projection deflation can be a "double-edged sword". Due to the spatial proximity of the 8 and 14 Hz rhythms, mode-2 projection of one rhythm effectively "removed" both rhythms in most subjects. Consequently, the three true remaining rhythms were detected in, at most, 13 out of 50 subjects. In other words, if the estimated components are correlated in one signature, projection deflation in that mode may eliminate not only the selected component but also a substantial amount of the remaining relevant information. This finding is consistent with the results of our previous study (Rošťáková et al. 2025), where mode-2 deflation resulted in a "clean" EEG signal with near-zero values and missing neurophysiological interpretation.

Although we identified scenarios in which the side effects of deflation were negligible in the EEG data case, there may be real data problems beyond EEG, where deflation significantly contaminates the data, as indicated by high $PercA$ values. In such cases, what should be done? First, caution must be exercised when interpreting the results of the subsequent tensor decomposition. The introduction of artefacts through deflation in the matrix decomposition can be effectively addressed by considering component orthogonalization, as demonstrated by Witten et al. (2009). As discussed in Sect. 3, the CP model can be easily represented in various matrix forms. Therefore, the results of Witten et al. (2009) can be adapted to this case provided that the nature of the tensor data allows for orthogonality.

7 Conclusion

In conclusion, one must be careful when deflating the latent components from tensor data under non-orthogonality constraints. This is a critical point because many

researchers outside the mathematics field, but dealing with real data, may not recognise this issue. Although our findings were demonstrated in the specific context of real and simulated EEG data, the derived mathematical formulas can be applied in a more general context.

Our analysis revealed that the deflation applied to the uncentered tensor yielded better results than deflation applied to the centred tensor. Consistent with the findings of Rošťáková et al. (2025), we prefer component subtraction to projection to the nullspace of the space spanned by signatures in a given mode.

The most challenging scenario occurs when attempting to remove the effect of a latent component whose signatures are highly correlated with those of the other components in at least one mode. Although component subtraction could not eliminate the selected component in our simulated data example, projection deflation removed too much relevant information along with the deflated component.

Although this study addressed several questions regarding the properties of deflation in tensor models, there is still room for further research focusing on other tensor decomposition or deflation approaches.

Appendix A: Definition of "tensor fibre spaces"

This study analysed the theoretical properties of two deflation techniques using the ideas from Camacho et al. (2021) because each tensor model can be rewritten in a matricized form. However, the question remains whether building a similar theory directly for tensors is possible.

The central concept of Camacho et al. (2021) is the analysis of matrix column and row spaces. Based on (1) and (2), a natural definition of the mode-1, mode-2 and mode-3 "fibre" spaces of a tensor $\underline{X}$ follows the formulas:

$$
\begin{aligned}
C_1(\underline{X}) &= \{\mathbf{x} \in \mathbb{R}^I : \mathbf{x} = \underline{X} \times_2 \mathbf{x}_2 \times_3 \mathbf{x}_3; \mathbf{x}_2 \in \mathbb{R}^J, \mathbf{x}_3 \in \mathbb{R}^K\}, \\
C_2(\underline{X}) &= \{\mathbf{y} \in \mathbb{R}^J : \mathbf{y} = \underline{X} \times_1 \mathbf{y}_1 \times_3 \mathbf{y}_3; \mathbf{y}_1 \in \mathbb{R}^I, \mathbf{y}_3 \in \mathbb{R}^K\}, \\
C_3(\underline{X}) &= \{\mathbf{z} \in \mathbb{R}^K : \mathbf{z} = \underline{X} \times_1 \mathbf{z}_1 \times_2 \mathbf{z}_2; \mathbf{z}_1 \in \mathbb{R}^I, \mathbf{z}_2 \in \mathbb{R}^J\}.
\end{aligned}
$$

From the first point of view, the following procedure appears similar to the matrix case: following Camacho et al. (2021), we analyse whether vectors $\mathbf{p}, \mathbf{t}$ (signatures of a component selected for removal) belong to the appropriate tensor fibre space. However, deeper analysis revealed several obstacles.

Let us focus on $C_1\left(\underline{X}\right)$; the situation with $C_2\left(\underline{X}\right)$ and $C_3\left(\underline{X}\right)$ is analogous. Using the properties of the tensor-matrix product, the formula $\mathbf{x} = \underline{X} \times_2 \mathbf{x}_2 \times_3 \mathbf{x}_3$ can be equivalently rewritten as

$$
\mathbf{x} = X_{(1)} \left(\mathbf{x}_3 \otimes \mathbf{x}_2\right)^\top, \mathbf{x}_2 \in \mathbb{R}^J, \mathbf{x}_3 \in \mathbb{R}^K.
$$

By contrast, $C\left(X_{(1)}\right)$ is a set of all vectors $\mathbf{x} \in \mathbb{R}^I$ which can be expressed as

$$
\mathbf{x} = X_{(1)}\mathbf{q}, \quad \mathbf{q} \in \mathbb{R}^{JK}.
$$

Consequently, $C_1\left(\underline{X}\right) \subset C\left(X_{(1)}\right)$ because vector $\mathbf{q}$ can be an arbitrary vector of the length JK from the definition of $C\left(X_{(1)}\right)$, but it has to have the form of a Kronecker product of two vectors in $C_1\left(\underline{X}\right)$.

Now, let us focus on the way the component matrix A is estimated. In Sect. 2.3, we described the ALS approach, in which two component matrices are considered fixed, and the third one is estimated using ordinary least squares:

$$\widehat{A} = X_{(1)}\left(C \odot B\right)\Lambda^{\star}\left(\Lambda^{\star}\left(C \odot B\right)^{\top}\left(C \odot B\right)\Lambda^{\star}\right)^{-1}.$$

However, more than the entire estimator, we are interested in the expression of the fth column of A

$$\widehat{\mathbf{a}}_f = X_{(1)}\left(C \odot B\right)\Lambda^{\star}\left(\Lambda^{\star}\left(C \odot B\right)^{\top}\left(C \odot B\right)\right)^{-1}\mathbf{e}_f = X_{(1)}\mathbf{q}_{a_f}.$$

Thus, $\mathbf{a}_f \in C\left(X_{(1)}\right)$. However, $\mathbf{a}_f$ does not belong to $C_1\left(\underline{X}\right)$ because $\mathbf{q}_{a_f}$ is generally not expressed as a Kronecker product. The only exception is when C and B have orthogonal columns. However, as highlighted in our previous studies, the orthogonality assumption misses the neurophysiological interpretation in EEG data. Moreover, if $\mathbf{a}_f \notin C_1\left(\underline{X}\right)$ under ordinary least squares, it does not belong to $C_1\left(\underline{X}\right)$ when considering nonnegativity or unimodality constraints.

The $PercA$ formula has the same structure for both $C\left(X_{(1)}\right)$ and $C_1\left(\underline{X}\right)$. Consequently, it produces values close to zero for $\mathbf{a}_f \notin C_1\left(\underline{X}\right)$, which contrasts with the use of this metric. Therefore, our research did not follow the path of tensor fibre spaces. Instead, we focused on the matricized versions of the CP model and worked directly with the corresponding matrix column and row spaces.

Acknowledgements This research was funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V04-00205 (Z.R.) and project No. 09I03-03-V04-00443 (R.R).

Author contributions Zuzana Rošťáková and Roman Rosipal have contributed equally to this work.

Funding Open access funding provided by The Ministry of Education, Science, Research and Sport of the Slovak Republic in cooperation with Centre for Scientific and Technical Information of the Slovak Republic This research was funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V04-00205 (Z.R.) and project No. 09I03-03-V04-00443 (R.R).

Data availability Data will be made available on request.

Materials availability Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Acar E, Aykut-Bingol C, Bingol H, Bro R, Yener B (2007) Multiway analysis of epilepsy tensors. Bioinformatics 23(13):10–18. https://doi.org/10.1093/bioinformatics/btm210

Bro R (1997) PARAFAC. Tutorial and applications. Chemom Intell Lab Syst 38(2):149–171. https://doi.org/10.1016/S0169-7439(97)00032-4

Camacho J, Smilde AK, Saccenti E, Westerhuis JA, Bro R (2021) All sparse PCA models are wrong, but some are useful. Part II: limitations and problems of deflation. Chemom Intell Lab Syst 208:104212. https://doi.org/10.1016/j.chemolab.2020.104212

Carroll JD, Chang J-J (1970) Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika 35(3):283–319. https://doi.org/10.1007/BF02310791

Cichocki A, Zdunek R, Phan AH, Amari S (2009) Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation. Wiley, Chichester. https://doi.org/10.1002/9780470747278

Harshman RA (1970) Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multimodal factor analysis. UCLA Work Pap Phonetics 16:1–84

Höskuldsson A (1988) PLS regression methods. J Chemom 2(3):211–228. https://doi.org/10.1002/cem.1180020306

Mackey L (2008) Deflation methods for sparse PCA. In: Advances in neural information processing systems (NIPS), vol 21, pp 1017–1024

Phan A-H, Tichavský P, Cichocki A (2015) Tensor deflation for CANDECOMP/PARAFAC -Part I: alternating subspace update algorithm. IEEE Trans Signal Process 63(22):5924–5938

Rosipal R, Rošťáková Z, Trejo LJ (2022) Tensor decomposition of human narrowband oscillatory brain activity in frequency, space and time. Biol Psychol 169:108287. https://doi.org/10.1016/j.biopsycho.2022.108287

Rošťáková Z, Rosipal R (2022) Determination of the number of components in the PARAFAC model with a nonnegative tensor structure: a simulated EEG data study. Neural Comput Appl. https://doi.org/10.1007/s00521-022-07318-x

Rošťáková Z, Rosipal R, Seifpour S, Trejo LJ (2020) A comparison of non-negative Tucker decomposition and parallel factor analysis for identification and measurement of human EEG rhythms. Meas Sci Rev 20(3):126–138. https://doi.org/10.2478/msr-2020-0015

Rošťáková Z, Rosipal R, Trejo LJ (2025) SPECTER—The Signal sPECtrum Tensor decomposition and Eye blink Removal algorithm. Biomed Signal Process Control 99:106889. https://doi.org/10.1016/j.bspc.2024.106889

The Mathworks, Inc. (2021) MATLAB Version 23.2.0.2459199 (R2023b). Mathworks, Natick

Tucker LR (1966) Some mathematical notes on three-mode factor analysis. Psychometrika 31(3):279–311. https://doi.org/10.1007/BF02289464

Witten D, Tibshirani R, Hastie T (2009) A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. Biostatistics 10:515–534. https://doi.org/10.1093/biostatistics/kxp008