NeuroImage 334 (2026) 121962

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Equivalence of modified k-means and tensor decomposition in EEG microstates: Implications for analysis and interpretation

Zuzana Rošťáková *, Roman Rosipal

Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, Bratislava, 841 04, Slovakia

ARTICLE INFO

Dataset link: https://www.mathworks.com/matlabcentral/fileexchange/183118-imscand-algorithm-for-eeg-microstate-analysis

Keywords:
Electroencephalogram EEG
EEG microstates
Modified k-means
Tensor decomposition
Parallel factor analysis
Implicit slice canonical decomposition

ABSTRACT

In the resting state, the human brain quickly transitions among a limited set of quasi-stable states known as EEG microstates, which characterize brain spatio-temporal activity with high temporal resolution. EEG microstates are typically identified using a clustering algorithm that analyzes peaks in the EEG global field power. In this study, we focus on the modified k-means algorithm (*modKM*), one of the most commonly employed methods for detecting EEG microstates. We demonstrate, both theoretically and through simulation and real EEG data, its equivalence with a different method known as the Implicit Slice Canonical Decomposition (*IMSCAND*), which is a special kind of tensor decomposition for symmetric higher order arrays. This relationship opens new avenues for EEG microstate detection, interpretation, and analysis by utilizing tensor decomposition methods and related techniques.

1. Introduction

Multichannel electroencephalogram (EEG) is nowadays a standard tool to describe brain activity, either in health or to uncover pathological processes. During the EEG signal analysis, one can focus on the signal properties in the temporal, spatial, and frequency domain or their combination. As an example, we can mention the detection of specific oscillatory rhythms in the EEG signal spectrum (Miwakeichi et al., 2004; Rosipal et al., 2022), which is a representative of the EEG temporal-spatial-frequency analysis, or artifact detection in the spatio-temporal domain by the independent component analysis (ICA).

In this study, we focus on the resting-state EEG analysis in the spatio-temporal domain through so-called EEG microstates. Historically, the roots of microstate analysis go back to 1970's in the work of Lehman (Lehmann, 1971). It characterizes the EEG signal through a limited set of quasi-stable states lasting approximately 60 - 150 ms (Michel and Koenig, 2017) and characterized by specific spatial topographies. These spatial topographies can be detected by using a clustering algorithm, for example, standard k-means (Lloyd, 1982) or its modified version (*modKM*) introduced by Pascual-Marqui et al. (1995). Other approaches include principal component analysis (PCA) (Skrandies, 1989) or topographic atomize and agglomerate hierarchical clustering (TAAHC) (Murray et al., 2008; Khanna et al., 2015).

In this study, we focus solely on the *modKM* algorithm as described in Pascual-Marqui et al. (1995). Inspired by the work of Bauckhage (2015) who derived equivalence between the standard k-means algorithm and matrix decomposition with restrictions, we aimed to detect a similar relationship also for *modKM*. However, as we will show later in the text, in this case, we have to inspect the decomposition models of higher-order arrays (tensors) instead of matrices.

The equivalence between modified k-means and higher-order array decomposition offers new perspectives for EEG microstate analysis. Since Lehman's pioneering work (Lehmann, 1971), the field has operated under a consensus that EEG microstates emerge from cluster analysis, with exactly one microstate active at any given time point. However, reinterpreting EEG microstates as latent components derived from tensor decomposition could integrate them into a broader blind-source separation framework and potentially enable their deeper analysis in a continuous space. Beyond this conceptual shift, we would like to highlight additional theoretical advantages: tensor-based approaches could guide selection of the optimal number of microstates, enhance *modKM*'s robustness to initialization, or conversely, properties of *modKM* can help to accelerate tensor decomposition algorithms that are typically more computationally demanding.

This study is organized in the following way: Section 2.1 introduces the basic notation and mathematical properties of several matrix and tensor models. Basic principles of tensor decomposition, with particular emphasis on the CP model [1] (Harshman, 1970; Carroll and Chang,

* Corresponding author.
E-mail address: zuzana.rostakova@savba.sk (Z. Rošťáková).

[1] CP abbreviates CANDECOMP (Carroll and Chang, 1970) and PARAFAC (Harshman, 1970), two equivalent tensor decomposition model

https://doi.org/10.1016/j.neuroimage.2026.121962
Received 20 November 2025; Received in revised form 27 April 2026; Accepted 28 April 2026
Available online 2 May 2026

1970) and its specialized form, Implicit Slice Canonical Decomposition (*IMSCAND*) (Selee et al., 2007), and EEG microstate detection procedure are described in Sections 2.2–2.4. The modified k-means algorithm is introduced in Section 3. The theoretical relationship between *modKM* and the tensor decomposition approach *IMSCAND* is derived in Section 4 and proved on simulated and real-world EEG data, described in Sections 5 and 6. Finally, the results are discussed and conclusions are provided in Sections 7 and 8.

2. Methods

2.1. Basic notation

In the following text, $\mathbb{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ represents a three-way array (tensor), while an uppercase letter $X \in \mathbb{R}^{I \times J}$, a bold lowercase letter $\mathbf{x} \in \mathbb{R}^I$ and a lowercase letter $x \in \mathbb{R}$ are used for matrices, vectors and scalars, respectively. The sets of all natural numbers, real numbers, and strictly positive real numbers are depicted by $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ respectively. An $F \times F$ identity matrix is denoted by $\mathbb{I}_{F\times F}$ and $\mathbb{I}_{F\times F\times F}$ represents its tensor counterpart with ones on its main super-diagonal and zeros elsewhere. A vector of ones is denoted by $\mathbf{1}_N = (1 \;\; 1 \ldots 1)^T \in \mathbb{R}^N$. Finally, $\mathbf{e}_j \in \mathbb{R}^N$ represents a standard basis unit vector with the jth element equal to 1 and zeros elsewhere.

The Frobenius norm of a vector, matrix, or tensor is denoted by $\|.\|_{Fro}$ and defined in the following way:

$$\|\mathbf{x}\|_{Fro} = \sqrt{\sum_{i=1}^{I} x_i^2} = \sqrt{\mathrm{Tr}\left(\mathbf{x}^T\mathbf{x}\right)}$$

$$\|X\|_{Fro} = \sqrt{\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}^2} = \sqrt{\mathrm{Tr}\left(X^T X\right)}$$

$$\|\mathbb{X}\|_{Fro} = \sqrt{\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\cdots\sum_{i_n=1}^{I_N} X_{i_1 i_2 \ldots i_n}^2}$$

The definition of the Kronecker product $A \otimes B$, the Khatri–Rao (column-wise Kronecker) product $A \odot C$ and Hadamard product $A \circledast D$ follows (Cichocki et al., 2009):

$$A \in \mathbb{R}^{I\times K}, B \in \mathbb{R}^{J\times L} \qquad A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1K}B \\ a_{11}B & a_{12}B & \ldots & a_{iK}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}B & a_{I2}B & \ldots & a_{IK}B \end{pmatrix} \in \mathbb{R}^{IJ\times KL}$$

$$A \in \mathbb{R}^{I\times K}, C \in \mathbb{R}^{J\times K} \qquad A \odot C = \left(\mathbf{a}_1 \otimes \mathbf{c}_1, \mathbf{a}_2 \otimes \mathbf{c}_2, \ldots, \mathbf{a}_K \otimes \mathbf{c}_K\right) \in \mathbb{R}^{IJ\times K}$$

$$A \in \mathbb{R}^{I\times K}, D \in \mathbb{R}^{I\times K} \qquad A \circledast D = \begin{pmatrix} a_{11}d_{11} & a_{12}d_{12} & \ldots & a_{1K}d_{1K} \\ a_{21}d_{21} & a_{22}d_{22} & \ldots & a_{2K}d_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}d_{I1} & a_{I2}d_{I2} & \ldots & a_{IK}d_{IK} \end{pmatrix} \in \mathbb{R}^{I\times K}$$

The outer product of three vectors $\mathbf{a} \in \mathbb{R}^I, \mathbf{b} \in \mathbb{R}^J, \mathbf{c} \in \mathbb{R}^K$ is denoted by $\mathbf{a}\circ\mathbf{b}\circ\mathbf{c}$ and results into a tensor $\mathbb{D} \in \mathbb{R}^{I\times J\times K}$ with $d_{ijk} = a_i b_j c_k$. For more details about the products' properties, see Cichocki et al. (2009, p. 33 – 38).

For a three-way tensor $\mathbb{X} \in \mathbb{R}^{I_1\times I_2\times I_3}$, let us define tensor slices and fibers. Following a "MATLAB-like" notation, an $i^{th}, i = 1, \ldots, I_1$ horizontal slice is a matrix $\mathbb{X}_{i..} \in \mathbb{R}^{I_2\times I_3}$ which we obtain by fixing the first index and varying the remaining two. Similarly, matrices $\mathbb{X}_{.i.} \in \mathbb{R}^{I_1\times I_3}, i = 1, \ldots, I_2$ and $\mathbb{X}_{..i} \in \mathbb{R}^{I_1\times I_2}, i = 1, \ldots, I_3$ represent lateral and frontal slices of a tensor and are created by fixing the second or the third index, respectively. When fixing two indices, we obtain a vector called tensor fiber (Cichocki et al., 2009).

In the following analysis, two tensor transformations will play an important role. Matricisation or unfolding transforms a tensor $\mathbb{X} \in \mathbb{R}^{I_1\times I_2\times I_3}$ into a matrix by concatenating tensor frontal or lateral slices. This can be done in three different ways:

$$\begin{aligned}
&\text{in the 1st mode:} && \mathbb{X}_{(1)} \in \mathbb{R}^{I_1\times I_2 I_3}, && \mathbb{X}_{(1)} = \left(X_{..1} \;\; X_{..2} \;\; \ldots \;\; X_{..I_3}\right), \\
&\text{in the 2nd mode:} && \mathbb{X}_{(2)} \in \mathbb{R}^{I_2\times I_1 I_3}, && \mathbb{X}_{(2)} = \left(X_{.1.}^T \;\; X_{.2.}^T \;\; \ldots \;\; X_{.I_2.}^T\right), \\
&\text{in the 3rd mode:} && \mathbb{X}_{(3)} \in \mathbb{R}^{I_3\times I_1 I_2}, && \mathbb{X}_{(3)} = \left(X_{1..}^T \;\; X_{2..}^T \;\; \ldots \;\; X_{I_2..}^T\right).
\end{aligned} \tag{1}$$

Vectorization transforms an $I_1 \times I_2 \times I_3$ tensor $\mathbb{X}$ into an $I_1 I_2 I_3 \times 1$ vector $\mathrm{vec}(\mathbb{X})$ by concatenating tensor slices above each other (Cichocki et al., 2009). Similar transformation can be defined also for matrices: for a matrix $A \in \mathbb{R}^{I_1\times I_2}$, $\mathrm{vec}(A) \in \mathbb{R}^{I_1 I_2}$ is created by stacking the matrix columns under each other. For more details see Cichocki et al. (2009, p. 30 – 33). In our analysis, the following equality plays an essential role

$$\mathrm{vec}(ABC) = \left(C^T \otimes A\right)\mathrm{vec}(B), \qquad A \in \mathbb{R}^{I\times J}, B \in \mathbb{R}^{J\times K}, C \in \mathbb{R}^{K\times L}. \tag{2}$$

The tensor-matrix product between a tensor $\mathbb{X} \in \mathbb{R}^{I_1\times I_2\times I_3}$ and matrices $A \in \mathbb{R}^{J\times I_1}, B \in \mathbb{R}^{J\times I_2}, C \in \mathbb{R}^{J\times I_3}$ in the corresponding mode follows the formulas

$$\begin{aligned}
&\mathbb{E} = \left(\mathbb{X}\times_1 A\right) \in \mathbb{R}^{J\times I_2\times I_3}, && e_{ji_2i_3} = \sum_{i_1=1}^{I_1} x_{i_1i_2i_3}a_{ji_1}, && \mathbb{E}_{(1)} = A\mathbb{X}_{(1)}, \\
&\mathbb{G} = \left(\mathbb{X}\times_2 B\right) \in \mathbb{R}^{I_1\times J\times I_3}, && g_{i_1ji_3} = \sum_{i_2=1}^{I_2} x_{i_1i_2i_3}b_{ji_2}, && \mathbb{G}_{(2)} = B\mathbb{X}_{(2)}, \\
&\mathbb{H} = \left(\mathbb{X}\times_3 C\right) \in \mathbb{R}^{I_1\times I_2\times J}, && h_{i_1i_2j} = \sum_{i_3=1}^{I_3} x_{i_1i_2i_3}c_{ji_3}, && \mathbb{H}_{(3)} = C\mathbb{X}_{(3)}.
\end{aligned} \tag{3}$$

2.2. CP decomposition and its special version IMSCAND

Before describing the EEG microstate analysis and the modified k-means algorithm, let us give a short introduction to tensor decomposition. Tensor decomposition refers to a set of blind source separation methods developed to detect latent components in higher-order arrays (tensors).

Consider a three-way tensor $\mathbb{X} \in \mathbb{R}^{I_1\times I_2\times I_3}$. A CP model[1] (Harshman, 1970; Carroll and Chang, 1970) decomposes tensor $\mathbb{X}$ into three latent matrices: $A \in \mathbb{R}^{I_1\times K}$, $B \in \mathbb{R}^{I_2\times K}$, and $C \in \mathbb{R}^{I_3\times K}$, all with K columns:

$$\mathbb{X} = \Lambda\times_1 A\times_2 B\times_3 C + \mathbb{E} = \sum_{k=1}^{K}\lambda_{kkk}\mathbf{a}_k\circ\mathbf{b}_k\circ\mathbf{c}_k + \mathbb{E}, \tag{4}$$

where $\Lambda \in \mathbb{R}^{K\times K\times K}$ is a super-diagonal tensor with $\lambda_{ijk} \neq 0$ only if $i = j = k$. The component matrices typically have unit columns to prevent scaling ambiguity, so the weights are stored in Λ.

The matricised versions of (4)

$$\begin{aligned}
\mathbb{X}_{(1)} &= A\Lambda_{(1)}(C \odot B)^T, \\
\mathbb{X}_{(2)} &= B\Lambda_{(2)}(A \odot C)^T, \\
\mathbb{X}_{(3)} &= C\Lambda_{(3)}(B \odot A)^T.
\end{aligned}$$

play an essential role during the estimation process of component matrices. This is especially true when using the alternating least squares algorithm (Bro, 1997).

2.3. Implicit slice canonical decomposition model (IMSCAND)

In addition to CP decomposition, Carroll and Chang (1970) also discussed a special version of CP symmetric in two modes in their paper. Let us consider $I_1 = I_2$ and the tensor $\mathbb{X} \in \mathbb{R}^{I_1\times I_1\times I_3}$ including

symmetric frontal slices $\mathbb{X}_{..i_3} = \mathbb{X}^T_{..i_3}$. Then the Individual Differences in Scaling (*INDSCAL*) (Carroll and Chang, 1970) follows the model

$$\mathbb{X} = \Lambda \times_1 A \times_2 A \times_3 C + \mathbb{E} = \sum_{k=1}^{K} \lambda_{kkk} \mathbf{a}_k \circ \mathbf{a}_k \circ \mathbf{c}_k + \mathbb{E}. \tag{5}$$

This model was further studied in Selee et al. (2007) under the assumption that the frontal slices are not only symmetric, but can also be expressed as a product of a matrix and its transpose: $\mathbb{X}_{..i} = Y_i Y_i^T, i = 1 \ldots, I_3$. This model is called the Implicit Slice Canonical Decomposition (*IMSCAND*) (Selee et al., 2007; Cichocki et al., 2009). It was proposed for clustering objects from datasets with multiple similarity matrices.

Although CP and *IMSCAND* may yield the same decomposition, the estimation of component matrices in *IMSCAND* can be performed more efficiently because of the special structure of the tensor frontal slices. This increased efficiency enables *IMSCAND* to be applied to larger datasets than CP (Selee et al., 2007). An iterative algorithm for fitting the *IMSCAND* model is described in Selee et al. (2007). Although the first two component matrices are equal in *IMSCAND*, Selee et al. (2007) recommend estimating them as distinct matrices, A_1 and A_2, during the main iteration process.

The overall idea of the algorithm is similar to the alternating least squares algorithm commonly used in tensor decomposition methods (Bro, 1997) and can be summarized as follows:

0. Set initial estimates for A_1, A_2, C.
1. Estimate C by considering A_1, A_2 fixed. Thanks to the structure of the tensor $\mathbb{X}$, this can be done element-wise:
$$\hat{C}_{ik} = \left(Y_i \mathbf{a}_{2_k}\right)^T \left(Y_i \mathbf{a}_{1_k}\right), \qquad i = 1, \ldots, I_3; k = 1, \ldots, K.$$
Here, $\mathbf{a}_{1_k}, \mathbf{a}_{2_k}$ represent kth columns of matrices A_1, A_2. Note, this is in contrast to the standard CP decomposition, where C should be estimated by the least squares algorithm as a whole.
2. Estimate A_2 by setting A_1, C fixed. Again, in contrast to CP, where A_2 has to be estimated as a whole, in this case, a column-wise estimation is possible:
$$\hat{\mathbf{a}}_{2_k} = \sum_{i=1}^{I_3} C_{ik} \left(Y_i \left(Y_i \mathbf{a}_{1_k}\right)\right), \qquad k = 1, \ldots, K.$$
3. Finally, A_1 is estimated in a similar way as A_2
$$\hat{\mathbf{a}}_{1_k} = \sum_{i=1}^{I_3} C_{ik} \left(Y_i \left(Y_i \mathbf{a}_{2_k}\right)\right), \qquad k = 1, \ldots, K.$$

Steps 1 to 3 are repeated until standard stopping criteria are met, such as reaching a maximal number of iterations, achieving a small relative change in the estimated matrices, or obtaining a mean squared error below a defined threshold. After the iteration process stops, Selee et al. (2007) recommend setting $A_1 = A_2$ and computing the final estimate of C using the formula from step 1.

Because all three component matrices are estimated in separate steps, it is possible to replace standard least squares optimization with its constrained variants, as is done in CP. In our situation, this approach becomes particularly relevant, as will be shown later. Specifically, matrix C will be associated with cluster assignment, and should therefore exhibit a sparse structure with only one entry equal to 1 in each row (hard clustering) or nonnegative elements that sum to 1 (soft clustering). To address the latter case, the Franz-Wolfe algorithm can be used, as described in Bauckhage (2016).

2.4. EEG microstate analysis

A multichannel EEG signal recorded by M electrodes over $N^\star$ sampling points can be represented by a matrix $Y^\star = \left(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N^\star}\right) \in \mathbb{R}^{M \times N^\star}$. The main aim of the EEG microstate analysis is to segment the signal into K microstates according to the similarity between their spatial maps. Microstate spatial maps as prototypes will be stored column-wise in a matrix $U = \left(\mu_1, \mu_2, \ldots, \mu_k\right) \in \mathbb{R}^{M \times K}$ in the following text.

In the first step, it is recommended to re-reference the EEG signal to the average reference:

$$Y_{ij} = Y^\star_{ij} - \frac{1}{M} \sum_{m=1}^{M} Y^\star_{mj}, \qquad i = 1, \ldots, M; j = 1, \ldots, N^\star.$$

Consequently, $\sum_{m=1}^{M} Y_{mj} = 0$ for each $j = 1, \ldots, N^\star$.

Then, the global field power (GFP) is computed in the following way

$$GFP(j) = \sqrt{\frac{1}{M} \sum_{m=1}^{M} Y^2_{mj}}, \qquad j = 1, \ldots, N^\star.$$

Once the GFP has been obtained, its local maxima – referred to as GFP peaks – are detected. The EEG spatial maps from these maxima then form the input for the cluster analysis. At this stage, one can typically choose between the standard k-means algorithm (Lloyd, 1982), modified k-means (Pascual-Marqui et al., 1995), topographic atomize and agglomerate hierarchical clustering (TAAHC) (Murray et al., 2008; Khanna et al., 2015), principal component analysis (Skrandies, 1989), but rarely also independent component analysis (ICA) (Yuan et al., 2012).

The result is a set of prototype spatial microstate maps. These maps are further back-fitted to GFP peaks. Following the "winner-takes-all" strategy, each GFP peak is assigned to the microstate map with the highest absolute similarity. Time points beyond GFP peaks are assigned to the same microstate as the closest GFP peak. However, noise in the EEG signal may cause very short microstate segments. This problem is usually solved by applying temporal smoothing. The algorithm inspects the microstate sequence and if the length of some microstate segments is under a given threshold – usually 20 ms – they are reassigned to the second most similar microstate (Poulsen et al., 2018). This process is repeated until the length of all microstate segments exceeds the threshold. Another smoothing technique was proposed in Pascual-Marqui et al. (1995).

2.4.1. Standard k-means algorithm and its relationship to matrix decomposition

Bauckhage (2015) showed that the standard k-means clustering tries to find an optimal cluster assignment for data $\mathbf{x}_1, \ldots, \mathbf{x}_N$ and corresponding cluster centroids $\mu_1, \ldots, \mu_K$ by minimizing the cost function

$$E(K) = \sum_{i=1}^{K} \sum_{\mathbf{x_n} \in C_i} \|\mathbf{x}_n - \mu_i\|^2_{Fro},$$
$$C_i = \{\mathbf{x}_n \quad : \quad \|\mathbf{x}_n - \mu_i\|^2_{Fro} \leq \|\mathbf{x}_n - \mu_j\|^2_{Fro} \quad \forall j \neq i\},$$
$$X = \left(\mathbf{x}_1, \ldots, \mathbf{x}_N\right) \text{ is a data matrix.}$$

This is equivalent to the constrained matrix factorization

$$E(K) = \|X - UZ\|^2_{Fro}$$

with $U = \left(\mu_1, \mu_2, \ldots, \mu_k\right) \in \mathbb{R}^{M \times K}$ and $Z \in \mathbb{R}^{K \times N}$ defined as the cluster assignment indicator matrix: each column of Z corresponds to a data point, and each row to a cluster, with $z_{in} = 1$ if the nth data point is assigned to cluster i and 0 otherwise. In some formulations, each column of Z contains only one 1 (hard assignment), while in others the entries are nonnegative and each column sums to one (soft assignment). This alternative view on the k-means clustering as constrained matrix decomposition may bring new insight into its properties and to develop new algorithms for solving either of the two equivalent problems (Bauckhage, 2016).

Building on this perspective, we now investigate whether analogous relationships can be established for the modified k-means clustering (Pascual-Marqui et al., 1995).

3. Modified k-means algorithm

In this section, we focus on the modified k-means algorithm (Pascual-Marqui et al., 1995), a standard tool for EEG microstate detection. After computing the GFP, we store the EEG signals from the GFP peaks in the matrix $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \in \mathbb{R}^{M\times N}$, where M is the number of electrodes and N is the number of GFP peaks. With a fixed number of microstates (clusters) K, the cost function in *modKM* can be defined in several ways. For our purpose, we use the formula from Pascual-Marqui et al. (1995), Poulsen et al. (2018):

$$C(\text{cluster assignment}, U; K) = \sum_{n=1}^{N} \|\mathbf{x}_n - (\mathbf{x}_n^T \mu_{l_n})\mu_{l_n}\|_{Fro}^2, \tag{6}$$

where each $l_1, \dots, l_N \in \{1, \dots, K\}$ represents a cluster assignment. The columns of $U = (\mu_1, \dots, \mu_K)$ are the cluster centroids representing microstate spatial maps with unit norm $\|\mu_k\|_{Fro} = 1$. The dot product $\mathbf{x}_n^T \mu_{l_n}$ makes this criterion polarity invariant (Poulsen et al., 2018).

Inspired by Bauckhage (2016) let us define an indicator matrix $Z = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N) \in \mathbb{R}^{K\times N}$ with $z_{kn} \in \{0, 1\}$ and $\sum_{k=1}^{K} z_{kn} = 1$ following the "winner-takes-all" assumption of microstates or so called hard clustering. Then (6) can be equivalently rewritten

$$C(Z, U; K) = \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn} \left\|\mathbf{x}_n - (\mathbf{x}_n^T \mu_k)\,\mu_k\right\|_{Fro}^2, \tag{7}$$

$$z_{kn} \in \{0, 1\} \text{ and } \sum_{k=1}^{K} z_{kn} = 1, \qquad \|\mu_k\| = 1, \qquad k = 1, \dots, K.$$

Let us denote the number of time points assigned to the kth cluster by n_k. Then $\sum_{n=1}^{N} z_{kn} = n_k$ and $ZZ^T = \operatorname{diag}(n_1 \dots n_K)$.

To achieve our main goal – detection of a possible relationship between *modKM* and some blind source separation model – we have to rewrite the cost function (7) in a matrix-like format:

$$\begin{aligned}
C(Z, U; K) &= \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn} \left\|\mathbf{x}_n - (\mathbf{x}_n^T \mu_k)\,\mu_k\right\|_{Fro}^2 = \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn} \left(\mathbf{x}_n - (\mathbf{x}_n^T \mu_k)\,\mu_k\right)^T \left(\mathbf{x}_n - (\mathbf{x}_n^T \mu_k)\,\mu_k\right) = \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K} \left(z_{kn}\mathbf{x}_n^T\mathbf{x}_n - 2z_{kn}\mathbf{x}_n^T(\mathbf{x}_n^T\mu_k)\,\mu_k\right. \\
&\quad \left. + z_{kn}\mu_k^T(\mathbf{x}_n^T\mu_k)^2\,\mu_k\right) = \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}\|\mathbf{x}_n\|^2 - 2\sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)^2 \\
&\quad + \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)^2 \underbrace{\|\mu_k\|^2}_{1} = \\
&= \sum_{n=1}^{N}\|\mathbf{x}_n\|^2 \underbrace{\left(\sum_{k=1}^{K} z_{kn}\right)}_{1} - 2\sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)^2 \\
&\quad + \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)^2 \\
&= \sum_{n=1}^{N}\|\mathbf{x}_n\|^2 - \underbrace{\sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)^2}_{T_2} = \operatorname{Tr}(X^TX) - T_2.
\end{aligned}$$

In the last row, we used the identity $\sum_{n=1}^{N}\|\mathbf{x}_n\|^2 = \operatorname{Tr}(X^TX)$.

Now let us focus on the second term T_2 in more detail:

$$\begin{aligned}
T_2 &= \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)^2 = \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}(\mathbf{x}_n^T\mu_k)(\mu_k^T\mathbf{x}_n) \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K} z_{kn}\mathbf{x}_n^T(\mu_k\mu_k^T)\mathbf{x}_n
\end{aligned}$$

since $\mathbf{x}_n^T\mu_k$ is a scalar and therefore equals to its transpose $\mu_k^T\mathbf{x}_n$.

$$T_2 = \sum_{n=1}^{N}\mathbf{x}_n^T\left(\sum_{k=1}^{K} z_{kn}\mu_k\mu_k^T\right)\mathbf{x}_n = \sum_{n=1}^{N}\mathbf{x}_n^T\left(U\operatorname{diag}(\mathbf{z}_n)U^T\right)\mathbf{x}_n,$$

where $\operatorname{diag}(\mathbf{z}_n)$ is a $K\times K$ matrix with the vector $\mathbf{z}_n$ on its diagonal. Let us define a set of $M\times M$ matrices $Y_n = U\operatorname{diag}(\mathbf{z}_n)U^T, n = 1, \dots, N$. Then

$$T_2 = \sum_{n=1}^{N}\mathbf{x}_n^T Y_n \mathbf{x}_n.$$

The elements of the above-mentioned sum are scalars and therefore they equal to their "vectorized" form $\mathbf{x}_n^T Y_n \mathbf{x}_n = \operatorname{vec}(\mathbf{x}_n^T Y_n \mathbf{x}_n)$. By applying the property (2) we obtain:

$$T_2 = \sum_{n=1}^{N}\mathbf{x}_n^T Y_n \mathbf{x}_n = \sum_{n=1}^{N}\operatorname{vec}(\mathbf{x}_n^T Y_n \mathbf{x}_n) = \sum_{n=1}^{N}(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\operatorname{vec}(Y_n). \tag{8}$$

At this step, we have to dive into the theory of tensors. Let us define a three-way array $\mathbb{Y} \in \mathbb{R}^{M\times M\times N}$ with $\mathbb{Y}_{..n} = Y_n$ as its frontal slices. Using the properties of a tensor-vector product (Cichocki et al., 2009), we can further express

$$Y_n = \mathbb{Y} \times_3 \mathbf{e}_n^T, \qquad \mathbf{e}_n \in \mathbb{R}^N. \tag{9}$$

However, we need a formula for $\operatorname{vec}(Y_n)$ in order to substitute it into (8).

For a while, let us consider $Y_n = (\mathbf{y}_{1n}, \mathbf{y}_{2n}, \dots, \mathbf{y}_{Mn})$ as a (degenerate) three-way tensor of the size $M \times M \times 1$. According to (1), its unfolding in the 3rd mode produces a structure of the size $1 \times M^2$, which is a vector:

$$Y_{n(3)} = (\mathbf{y}_{1n}^T, \mathbf{y}_{2n}^T, \dots, \mathbf{y}_{Mn}^T) = \operatorname{vec}(Y_n)^T \tag{10}$$

and equals to the transposed version of $\operatorname{vec}(Y_n)$ (Cichocki et al., 2009, p.32, Eq. (1.67)). Combining (9), (10), and the 3rd mode unfolding of a tensor-matrix product in (3) we can write:

$$\operatorname{vec}(Y_n) = Y_{n(3)}^T = (\mathbb{Y} \times_3 \mathbf{e}_n^T)_{(3)}^T = (\mathbf{e}_n^T\mathbb{Y}_{(3)})^T = \mathbb{Y}_{(3)}^T\mathbf{e}_n, \tag{11}$$

which is also in line with (Cichocki et al., 2009, p. 39, Table 1.1). Substituting (11) into T_2 we obtain

$$T_2 = \sum_{n=1}^{N}(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\operatorname{vec}(Y_n) = \sum_{n=1}^{N}(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\mathbb{Y}_{(3)}^T\mathbf{e}_n.$$

However, $(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\mathbb{Y}_{(3)}^T\mathbf{e}_n$ is a scalar for each $n = 1, \dots, N$ and therefore equals to its trace $(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\mathbb{Y}_{(3)}^T\mathbf{e}_n = \operatorname{Tr}\left((\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\mathbb{Y}_{(3)}^T\mathbf{e}_n\right)$. Then

$$\begin{aligned}
T_2 &= \sum_{n=1}^{N}\operatorname{Tr}\left((\mathbf{x}_n^T \otimes \mathbf{x}_n^T)\mathbb{Y}_{(3)}^T\mathbf{e}_n\right) = \sum_{n=1}^{N}\operatorname{Tr}\left(\mathbf{e}_n^T\mathbb{Y}_{(3)}(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)^T\right) = \\
&= \sum_{n=1}^{N}\operatorname{Tr}\left(\mathbb{Y}_{(3)}(\mathbf{x}_n^T \otimes \mathbf{x}_n^T)^T\mathbf{e}_n^T\right) = \operatorname{Tr}\left(\mathbb{Y}_{(3)}\sum_{n=1}^{N}(\mathbf{x}_n \otimes \mathbf{x}_n)\mathbf{e}_n^T\right) = \\
&= \operatorname{Tr}\left(\mathbb{Y}_{(3)}(X \odot X)\right).
\end{aligned} \tag{12}$$

In the formulas above, we used the cyclic property of trace

$$\operatorname{Tr}(ABC) = \operatorname{Tr}(BCA) = \operatorname{Tr}(CAB)$$

for vectors or matrices of appropriate sizes, and the definition of the Khatri–Rao product.

In the last step, we need to go back from the tensor $\mathbb{Y} \in \mathbb{R}^{M\times M\times N}$ to matrices U and Z, which are at the center of our interest. Based on (Cichocki et al., 2009, p. 48, Eq. (1.131)) and using the fact that

$$Y_n = U\operatorname{diag}(\mathbf{z}_n)U^T, \qquad n = 1, \dots, N$$

are frontal slices of $\mathbb{Y} \in \mathbb{R}^{M\times M\times N}$, we can conclude that $\mathbb{Y}$ follows a CP model with component matrices U, U, Z^T

$$\mathbb{Y} = \mathbb{I}_{K\times K\times K} \times_1 U \times_2 U \times_3 Z^T,$$

which can be equivalently expressed in a 3rd mode unfolded version

$$\mathbb{Y}_{(3)} = Z^T (U \odot U)^T . \tag{13}$$

Finally, substituting (13) into (12) we obtain

$$T_2 = \mathrm{Tr}\left(Z^T (U \odot U)^T (X \odot X)\right)$$

and

$$C(Z,U;K) = \mathrm{Tr}\left(X^T X\right) - \mathrm{Tr}\left(Z^T (U \odot U)^T (X \odot X)\right). \tag{14}$$

Now, the cost function of *modKM* is rewritten only using matrices X, U and Z, similarly as in Bauckhage (2015).

4. Relationship between the *modKM* and tensor decomposition

Presence of the Khatri–Rao product in (14) and the theoretical properties of the constructed tensor $\mathbb{Y}$ led us to the idea of whether it is possible to rewrite the $C(Z,U;K)$ in terms of tensor decomposition.

Using the columns of the matrix $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \in \mathbb{R}^{M\times N}$, let us construct an $M \times M \times N$ tensor $\mathbb{X}$ with symmetric frontal slices $\mathbb{X}_{..n} = \mathbf{x}_n \mathbf{x}_n^T, n = 1, \ldots, N$, which can be alternatively expressed as $\mathbb{X}_{..n} = X\,\mathrm{diag}(\mathbf{e}_n) X^T$. Following the same reasoning as in tensor $\mathbb{Y}$, $\mathbb{X}$ follows a CP model

$$\mathbb{X} = \mathbb{I}_{N\times N\times N} \times_1 X \times_2 X \times_3 I_{N\times N}.$$

Tensor $\mathbb{X}$ and tensor $\mathbb{Y}$ from the previous section are both of the same size $M\times M\times N$. Tensor $\mathbb{X}$ is fixed and includes only the data (although in a specific form), but tensor $\mathbb{Y}$ follows a model with unknown (latent) matrices U and Z. Let us approximate the data tensor $\mathbb{X}$ by the tensor $\mathbb{Y}$. Then, the cost function has the following form

$$\begin{aligned}
&S(Z,U;K) = \|\mathbb{X} - \mathbb{Y}\|^2_{Fro} = \\
&= \|\mathbb{X} - \mathbb{I}_{K\times K\times K} \times_1 U \times_2 U \times_3 Z^T\|^2_{Fro} = \|\mathbb{X}_{(3)} - Z^T (U \odot U)^T\|^2_{Fro} = \\
&= \|I_{N\times N} (X \odot X)^T - Z^T (U \odot U)^T\|^2_{Fro} = \|(X \odot X) - (U \odot U) Z\|^2_{Fro} = \\
&= \mathrm{Tr}\left((X \odot X) - (U \odot U) Z\right)^T \left((X \odot X) - (U \odot U) Z\right) = \\
&= \mathrm{Tr}\left((X \odot X)^T (X \odot X) - 2 (X \odot X)^T (U \odot U) Z\right. \\
&\left. + Z^T (U \odot U)^T (U \odot U) Z\right) = \\
&= \mathrm{Tr}\left(\left(X^T X\right) \circledast \left(X^T X\right)\right) - 2\,\mathrm{Tr}\left((X \odot X)^T (U \odot U) Z\right) \\
&+ \mathrm{Tr}\left(Z^T (U \odot U)^T (U \odot U) Z\right) = \\
&= \mathrm{Tr}\left(\left(X^T X\right) \circledast \left(X^T X\right)\right) - \underbrace{2\,\mathrm{Tr}\left(Z^T (U \odot U)^T (X \odot X)\right)}_{T_2} \\
&\underbrace{+\,\mathrm{Tr}\left((U \odot U) Z Z^T (U \odot U)^T\right)}_{T_3}.
\end{aligned}$$

The last expression T_3 is actually a constant

$$\begin{aligned}
T_3 &= \mathrm{Tr}\left((U \odot U) Z Z^T (U \odot U)^T\right) \\
&= \mathrm{Tr}\left((U \odot U)\,\mathrm{diag}\left(n_1 \ldots n_K\right) (U \odot U)^T\right) = \\
&= \mathrm{Tr}\left(\sum_{k=1}^{K} n_k \left(\mu_k \otimes \mu_k\right) \left(\mu_k \otimes \mu_k\right)^T\right) \\
&= \mathrm{Tr}\left(\underbrace{\sum_{k=1}^{K} n_k \left(\mu_k \otimes \mu_k\right)^T \left(\mu_k \otimes \mu_k\right)}_{scalar}\right) = \\
&= \sum_{k=1}^{K} n_k \underbrace{\left(\mu_k^T \mu_k\right)}_{1} \otimes \underbrace{\left(\mu_k^T \mu_k\right)}_{1} = \sum_{k=1}^{K} n_k = N.
\end{aligned}$$

Consequently

$$S(Z,U;K) = \mathrm{Tr}\left(\left(X^T X\right) \circledast \left(X^T X\right)\right) - 2T_2 + N.$$

Moreover, $S(Z,U;K)$ is a cost function for the *IMSCAND* model with K latent components and latent component matrices U, U, Z^T fitted to the data tensor $\mathbb{X}$.

Finally, let us compare the two cost functions $C(Z,U;K)$ and $S(Z,U;K)$

$$\begin{aligned}
C(Z,U;K) &= \mathrm{Tr}\left(X^T X\right) - T_2 \\
S(Z,U;K) &= \mathrm{Tr}\left(\left(X^T X\right) \circledast \left(X^T X\right)\right) - 2T_2 + N.
\end{aligned}$$

Both include the term T_2, which is the only one depending on unknown matrices U and Z. The other terms are only constants, playing no role when estimating U and Z. Therefore, we can conclude that the *modKM* clustering is equivalent to fitting an *IMSCAND* model to the data tensor $\mathbb{X}$ with symmetric frontal slices.

When computing the similarity between EEG time points in GFP peaks and cluster centroids in the *modKM* algorithm, Tait and Zhang (2022b) considered normalization of the X columns. Under the assumption $\|\mathbf{x}_n\|_{Fro} = 1, \quad n = 1, \ldots, N$ it is easy to show

$$\begin{aligned}
\mathrm{Tr}\left(X^T X\right) &= \sum_{n=1}^{N} 1 = N, \\
\mathrm{Tr}\left(\left(X^T X\right) \circledast \left(X^T X\right)\right) &= \sum_{n=1}^{N} 1^2 = N.
\end{aligned}$$

Consequently,

$$\begin{aligned}
C(Z,U;K) &= N - T_2 \\
S(Z,U;K) &= N - 2T_2 + N = 2\left(N - T_2\right) = 2C(K)
\end{aligned}$$

which makes the equivalence between *modKM* and *IMSCAND* even more obvious.

5. Data

In this section, we focus on the practical implications of the equivalence derived between *modKM* and *IMSCAND*. In an ideal scenario, both algorithms are expected to produce identical spatial microstate maps and assign the GFP peak time points to the same microstates. To illustrate this, we demonstrate the previously derived theoretical result using simulated EEG data with known properties in the subsequent section.

Additionally, real-world resting-state EEG signals from eight healthy volunteers were analyzed. In this context, the equivalence between *modKM* and *IMSCAND* may be less apparent due to noise, artifacts, and the complex structure of EEG signals. The iterative nature of both algorithms can also lead to convergence to local optima rather than the global optimum, potentially resulting in differences in outcomes. Nevertheless, we show that both approaches can yield similar results despite these challenges.

It should be emphasized that the real data analysis serves only as an illustrative example of method consistency. A more robust, population-level assessment of equivalence would require further algorithm optimization and a larger real EEG dataset, which is beyond the scope of this study.

5.1. Simulated data

In the first step, we aimed to demonstrate the equivalence between the *modKM* clustering and *IMSCAND* tensor decomposition using well-controlled simulated data with known microstate sequences and corresponding spatial maps. To achieve this, we utilized the +microstate toolbox for MATLAB (Tait and Zhang, 2022b) and the simulation procedure described in Mishra et al. (2020), which enabled the required simulations using different techniques.

Fifty trials of microstate sequences, each comprising four microstates, were generated using a random walk decision tree (Tait and Zhang, 2022a). Prototype spatial maps for these microstates were

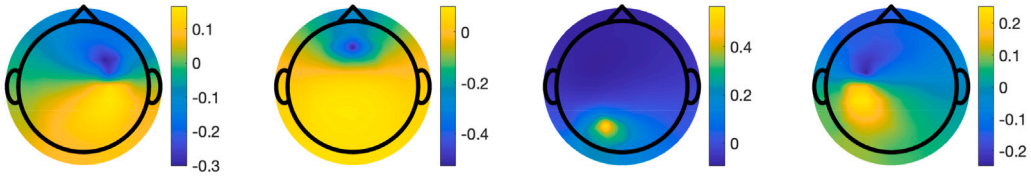


Fig. 1. Four prototype microstate spatial maps used for EEG data simulation as described in Tait and Zhang (2022a).

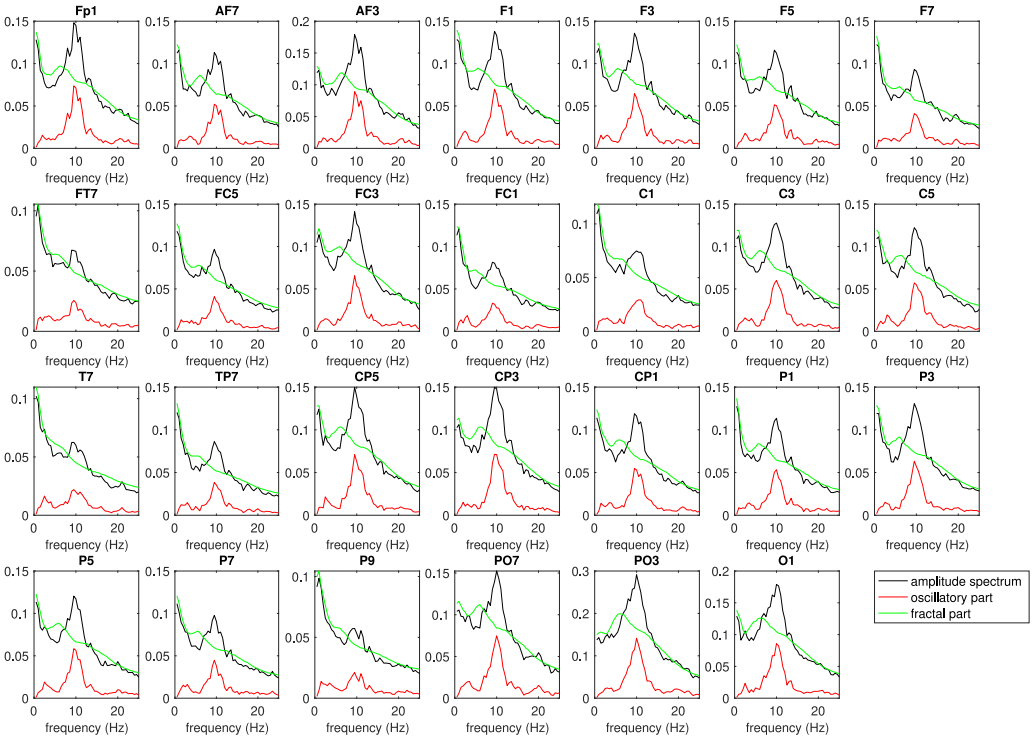


Fig. 2. Amplitude spectrum (black) of the simulated EEG signal for 27 left-hemisphere electrodes. The fractal and oscillatory parts of the spectrum are depicted in green and red, respectively.

selected to represent distinct activity in the right central, frontal, left occipital, and left central regions (Fig. 1). The resulting signals were stored in the matrices $X_{EEG} \in \mathbb{R}^{M \times N^\star}$, where $M = 64$ electrodes and $N^\star = 256 \times 60$ time points, corresponding to a sampling rate of 256 Hz. In line with Tait and Zhang (2022a), the minimum duration of an EEG microstate epoch was set to 50 ms. To verify that the simulated data retained natural neurological characteristics, we conducted a visual inspection of the corresponding amplitude spectra (Fig. 2).

5.2. Real EEG data

In the second step, we analyzed the multichannel EEG signals of eight healthy volunteers (4 males, 4 females; 22.86 ± 3.44 years). Prior to participation, all subjects provided their informed written consent and were financially compensated afterwards.

Each participant completed two sessions of the motor imagery experiment. Two minutes of resting-state, eyes-open EEG signals were recorded before and after each session. For each subject, these resting-state EEG recordings were combined, yielding approximately eight minutes of EEG data per participant for analysis.

EEG data were recorded at the sampling rate of 250 Hz using the wireless g.Nautilus PRO FLEXIBLE system with 31 Ag/AgCl wet electrodes arranged according to the international 10-20 system. Artifactual EEG segments were detected and removed semi-automatically using the BrainVision Analyzer (Brain Products, GmbH, 2013). For more details, see Evetović et al. (2025).

5.3. Basic settings and criteria for methods comparison

Consistent with the pipeline used in the literature for real EEG microstate analysis (Lehmann and Skrandies, 1980; Tait and Zhang, 2022b; Jajcay and Hlinka, 2023), both simulated and real EEG signals were average re-referenced and bandpass-filtered between 2 and 20 Hz.

GFP was computed, and its corresponding peaks were detected by the +microstate toolbox (Tait and Zhang, 2022b). GFP peak time points were then stored in a matrix $X \in \mathbb{R}^{M \times N}$, where N represents the number of GFP peaks in the dataset, and M is equal to the number of electrodes. Matrix X served as the input for the *modKM* clustering algorithm. It was also used to construct the tensor $\mathbb{X}$ for the *IMSCAND* decomposition.

In this study, we considered the *modKM* algorithm as implemented in the +microstate toolbox for MATLAB (Tait and Zhang, 2022b). Our own implementation of the *IMSCAND* algorithm[2] was inspired by the estimation process described in Selee et al. (2007). However, matrix $C = Z^T$ should follow a sparse structure, with only one nonzero element per row representing microstate assignment (hard clustering) in our case. This requirement can be relaxed by allowing the elements in each row of $C = Z^T$ to take values from the interval $[0, 1]$ and

[2] An implementation of the IMSCAND algorithm for EEG microstate detection is freely available at https://www.mathworks.com/matlabcentral/fileexchange/183118-imscand-algorithm-for-eeg-microstate-analysis

sum to one (soft clustering). This property of the matrix Z is used when deriving the equivalence between *modKM* and *IMSCAND*, and is implemented in this study.

Consequently, the standard least squares criterion is not suitable when estimating $C = Z^T$ in this context. A comparable scenario was addressed in Bauckhage (2016), where the Franz-Wolfe algorithm was employed for this purpose. Consistent with Bauckhage (2016), the Franz-Wolfe algorithm was also applied for estimating $C = Z^T$ in this study. The resulting matrix contained elements within the range $[0, 1]$. The microstate assignment can then be determined by identifying the column with the largest element in each row.

In the real data analysis, *modKM* was run for two to 10 microstates. The optimal number of microstates was determined using the *kneedle* algorithm applied to global explained variance (GEV) (Hu et al., 2022), yielding $K_{modKM\_GEV} = 5$ for all subjects. The same range of microstates was considered for *IMSCAND*, and the *kneedle* algorithm similarly identified $K_{IMSCAND\_GEV} = 5$ as optimal. Nevertheless, since *IMSCAND* is a tensor decomposition method, core consistency diagnostics (CCD) (Bro and Kiers, 2003), originally proposed for selecting the optimal number of latent components in tensor data, was also employed to determine the appropriate number of microstates. Using CCD, the optimal number of microstates was $K_{IMSCAND\_CCD} = 3$ across all eight subjects. To ensure comparability between algorithms, both methods were rerun with $K = 3$ and $K = 5$ microstates in the final step.

For the simulated data, the optimal number of microstates was set to the true value $K = 4$. Nevertheless, we also applied the *kneedle* algorithm in the case of both *modKM* and *IMSCAND* and CCD for *IMSCAND*. In all cases, the selected number of microstates was equal to the true $K = 4$.

When comparing the outcomes from *modKM* and *IMSCAND* we considered following criteria:

- *Similarity of microstate spatial maps:* In simulated data, the four true microstate prototype maps were compared to those estimated by *modKM* and *IMSCAND* using the absolute Spearman correlation coefficient ($\rho_{Spear.}$) and the global map dissimilarity (GMD) (Brunet et al., 2011). The GMD measure follows the formula

$$GMD = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{u_i-\bar{u}}{\sqrt{\frac{\sum_{j=1}^{M}(u_j-\bar{u})^2}{M}}} - \frac{v_i-\bar{v}}{\sqrt{\frac{\sum_{j=1}^{M}(v_j-\bar{v})^2}{M}}}\right)^2}, \quad (15)$$

where u and v are two microstate spatial maps and M is the number of electrodes. GMD ranges from 0 (identical maps) to 2 (maps with opposite polarity). However, in EEG microstate analysis, two spatial maps with opposite polarity are considered to represent the same microstate. Therefore, we adapted the GMD formula so that values close to 0 indicate similar maps (taking into account polarity invariance) and values close to 1 indicate dissimilar maps:

$$GMD = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{u_i-\bar{u}}{\sqrt{\frac{\sum_{j=1}^{M}(u_j-\bar{u})^2}{M}}} - \text{sign}(\rho(u,v))\frac{v_i-\bar{v}}{\sqrt{\frac{\sum_{j=1}^{M}(v_j-\bar{v})^2}{M}}}\right)^2}. \quad (16)$$

In Eq. (16), $\rho(u, v)$ denotes the correlation coefficient between two microstate spatial maps u and v. The function sign(x) is the signum function, which returns 1 if $x > 0$, -1 if $x < 0$, and 0 if $x = 0$.
For real data, where true spatial maps were unknown, we compared the spatial maps estimated by *modKM* and *IMSCAND*.

Table 1
The median and interquartile range (IQR) of the global explained variance (GEV) for *modKM* and *IMSCAND*, as well as the global map dissimilarity (GMD) and the Spearman correlation coefficient between the prototype microstate spatial maps and their versions estimated by *modKM* and *IMSCAND* when applied to simulated data consisting of four microstates.

		modKM		*IMSCAND*	
		median	IQR	median	IQR
GEV		0.673	0.008	0.673	0.008
GMD	micro. 1	0.035	0.014	0.036	0.015
	micro. 2	0.041	0.014	0.043	0.016
	micro. 3	0.063	0.021	0.068	0.023
	micro. 4	0.034	0.011	0.036	0.012
$\lvert\rho_{Spear.}\rvert$	micro. 1	0.998	0.001	0.998	0.001
	micro. 2	0.985	0.019	0.983	0.018
	micro. 3	0.958	0.023	0.956	0.020
	micro. 4	0.999	0.000	0.999	0.000

- *Global explained variance (GEV):* GEV is another measure evaluating the quality of EEG microstate detection. In this study, we follow the GEV formula from Hu et al. (2022).
- *Microstate sequences assigned to GFP peak time points:* In practice, microstates are assigned exclusively to GFP peak time points, while all other time points inherit the microstate of the nearest GFP peak time point. Consequently, comparisons between estimated and original microstate sequences are restricted to GFP peak time points or a defined subset of them.
For the simulated data, the ratio of GFP peak time points originally associated with the kth microstate ($k = 1, \ldots, K$) and subsequently assigned to the same microstate by *modKM* was calculated. This procedure was also applied to *IMSCAND*. In both real and simulated data analyses, a similar comparison was conducted between the *modKM* and *IMSCAND* microstate sequences.

6. Results

6.1. Simulated data

The initial analysis involved comparing the prototype microstate spatial maps with those estimated by the *modKM* clustering and the *IMSCAND* algorithm. Fig. 3 presents an example from five randomly selected subjects. With the exception of the opposite polarity observed in several spatial maps estimated by *IMSCAND*, which is acceptable under the assumed polarity invariance, both algorithms produced microstate maps visually equivalent to their true versions.

In addition to the visual comparison, we computed the Spearman correlation coefficient and GMD between the prototype maps and their estimated versions. Overall, both algorithms worked well. The true four microstate spatial maps were identified in all 50 trials with the absolute Spearman correlation with the original prototype maps reaching 1, as depicted in Table 1. GMD was close to zero for the majority of simulated data. But what is more important, both measures reached consistent values across both methods. The median values and the interquartile range (IQR) of GEV were equal to 0.673 and 0.008, respectively, in both *modKM* and *IMSCAND* (Table 1).

The second step involved comparing the microstate assignments produced by both algorithms with the true generated microstate sequence. Specifically, the proportion of the GFP peak time points assigned to the true microstate $k \in \{1, 2, 3, 4\}$ by *modKM* was calculated. The same procedure was applied to the microstate assignments from the *IMSCAND* algorithm. The results are depicted in Fig. 4(a) (blue and red) in a form of violin plots. We were interested also in the correspondence in microstate assignment among *modKM* and *IMSCAND*. Therefore, the proportion of GFP peak time points from the kth microstate according

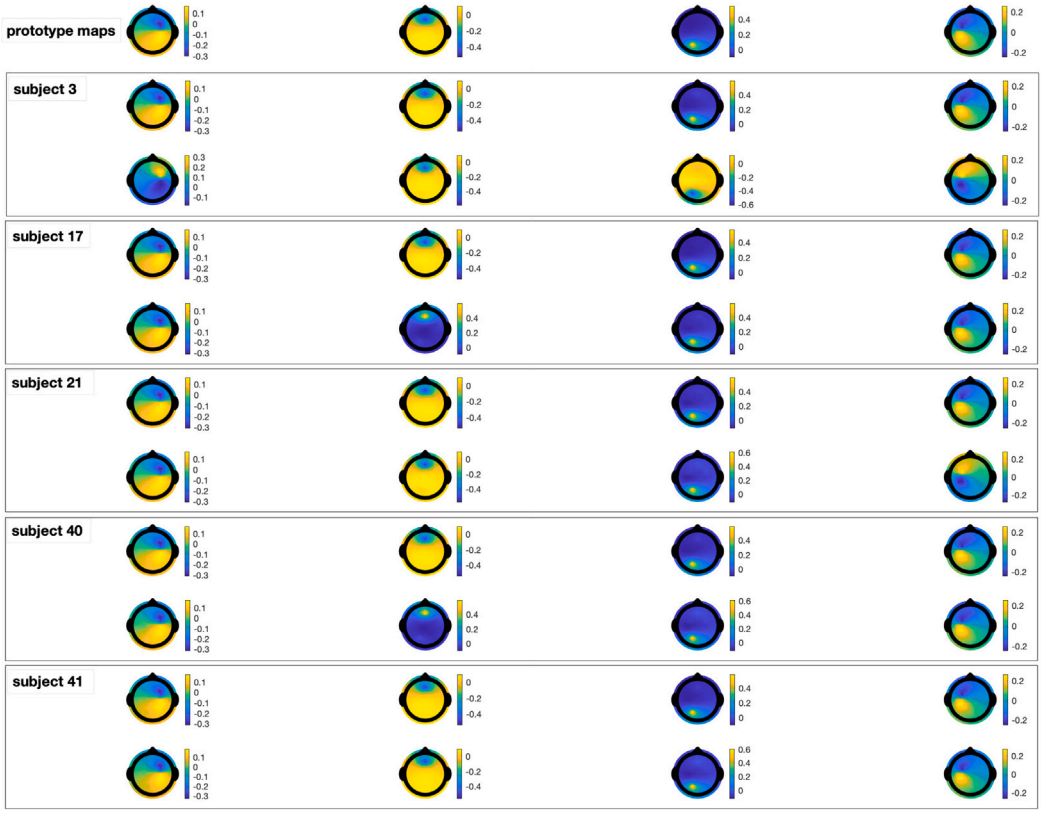


Fig. 3. First row: Four prototype microstate spatial maps used for EEG data simulation. Next rows: Microstate spatial maps estimated for five randomly selected subjects by the *modKM* clustering (top row for each subject) and the *IMSCAND* algorithm (bottom row of each subject).

to *modKM* which were assigned into the same microstate also by *IMSCAND* was inspected and is depicted in Fig. 4(a) (yellow). The analysis was repeated also by considering only the time points belonging to the top 25% GFP peaks (Fig. 4(b)).

The comparison of assignments between *modKM* and *IMSCAND* showed that the proportion of points assigned to the same microstate by both algorithms was generally high, exceeding 0.99, with complete agreement observed in at least 50% of cases. When considering only the points belonging to the 25% of the highest GFP peaks, both approaches led to exactly same microstate assignment for all 50 simulated data (Fig. 4(b), yellow). These findings support the theoretical equivalence between methods derived in the previous sections. Minor discrepancies in assignment may arise from the iterative nature of both *IMSCAND* and *modKM* algorithms, which can converge to slightly different solutions or be affected by numerical factors.

The correspondence between original and estimated microstate assignments consistently exceeded 0.88 for both algorithms. When analysis was restricted to the points corresponding to the 25% of highest GFP peaks (Fig. 4(b)), the match improved further. Both algorithms accurately assigned the points belonging to the highest GFP peaks in the true microstates, although minor discrepancies were observed for points in lower GFP peaks.

We hypothesize that this observation can be explained by the findings of Mishra et al. (2020), who proposed a continuous representation of microstates. Rather than assigning each point to a single microstate, they hypothesized that points can be characterized by a mixtures of microstates. They observed that some points associated with low GFP peaks lie "in-between" several microstates, having approximately equal distances to each of them, which makes it challenging to assign them to a single microstate under the winner-takes-all assumption.

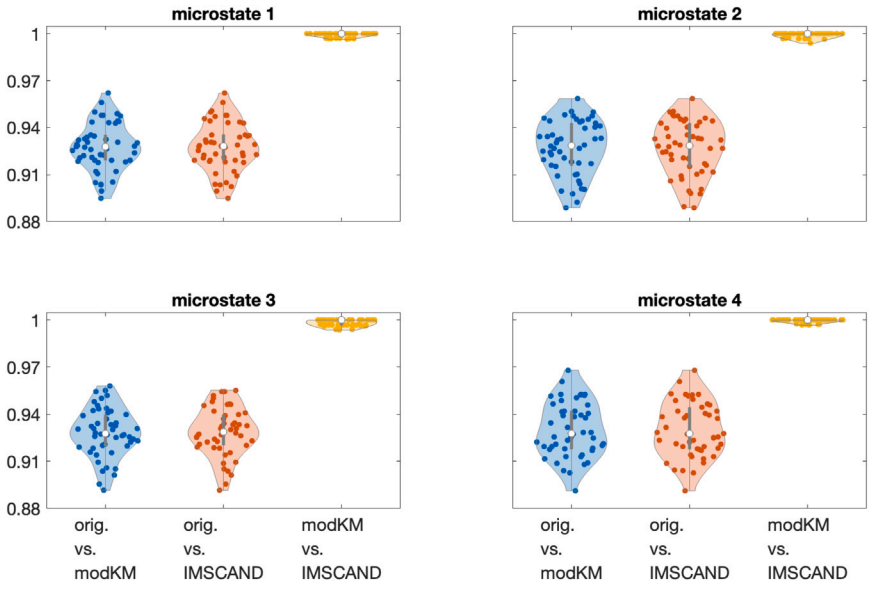
The matrix Z from *IMSCAND* contains values in the range $[0, 1]$, akin to soft clustering. These values can be interpreted as probabilities or weights indicating the degree to which a given time point belongs to a given microstate. At this step, we therefore focused on the proportion of GFP time points where the leading probability falls below the threshold of 0.5, indicating that this point is not dominantly similar to any single microstate.

The results shown in Fig. 5 demonstrate that the overall proportion of points not strongly similar to any microstate spatial map is low (1.2% in average; Fig. 5, blue) in our simulated data. But more importantly, these ambiguous points are absent from the set of points associated with the top 25% of GFP peaks (Fig. 5, orange). Within the subset of time points belonging to the lowest 25% of GFP peaks, however, 4%–6% exhibit a leading probability below 0.5, suggesting they lie between microstates (Fig. 5, yellow). Consequently, even minor variations in the estimated microstate spatial maps in comparison to their true simulated versions can cause such points to be assigned to a different microstate than the one from which they were simulated.

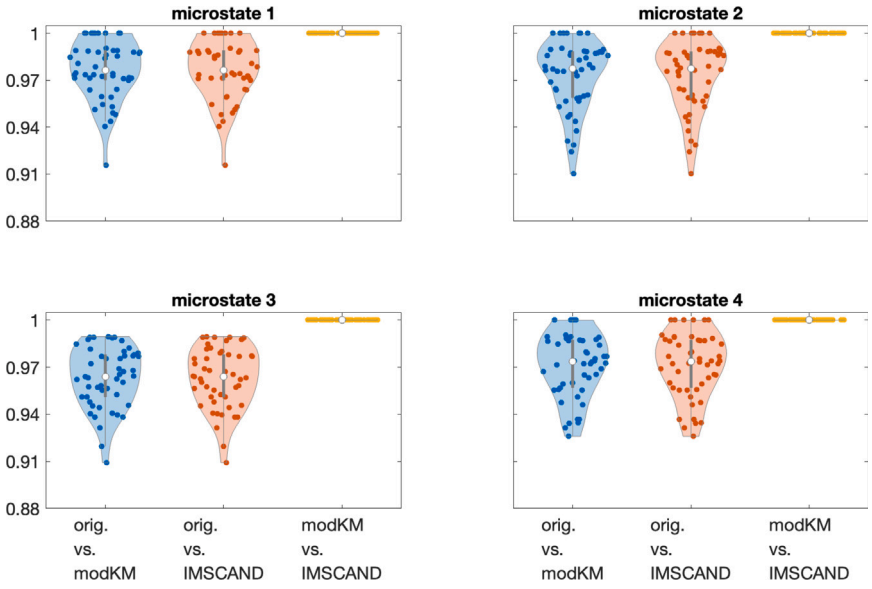
6.2. Real EEG signal

In the initial step, we compared GEV of three microstates identified by *modKM* and *IMSCAND*. Fig. 6 shows that both EEG detection methods produced similar distributions of GEV with median values of 0.525 for *modKM* (IQR = 0.023) and 0.523 for *IMSCAND* (IQR = 0.024). Increasing the number of microstates to five resulted in median GEV values of 0.575 for *modKM* and 0.573 for *IMSCAND*, with IQR values of approximately 0.027 for both methods (Fig. 6).

Subsequent analyses compared the microstate spatial maps and microstate sequences estimated by *modKM* and *IMSCAND*. It is important to note that these results are presented at the subject-specific level rather than the population level. This approach was chosen for two main reasons. First, the dataset size is insufficient to support robust population-level inferences. Second, as illustrated in Fig. 7, the



(a) All GFP peak time points were taken into account in the analysis.



(b) Only points related to the 25% of highest GFP peaks were taken into account in the analysis.

Fig. 4. Proportion of the GFP peak time points from the original microstate $k, k = 1, \ldots, 4$ assigned to the same microstate by *modKM* (blue) or *IMSCAND* (red). The yellow violin plot represents the proportion of GFP peak time points from the kth microstate according to *modKM* which were assigned the same microstate also by *IMSCAND*.

detected microstates differ across subjects, complicating aggregation of results across microstates. Additionally, as highlighted in Section 5, the real EEG data analysis primarily serves as an illustrative example to demonstrate that the correspondence between *modKM* and *IMSCAND* is observable in a complex real-world dataset.

Table 2 presents the median absolute Spearman correlation coefficient ($|\rho_{Spear.}|$) and GMD between estimated microstate spatial maps. There is no consensus on which correlation coefficient values indicate acceptable level of similarity across the literature (Akoglu, 2018). In this study, we therefore adopt approximate correlation intervals from the medical literature (Schober et al., 2018; Akoglu, 2018): correlations of 0.6 - 0.7 typically suggest moderate similarity, 0.7 - 0.9 indicate a strong similarity, and values exceeding 0.9 indicate a very strong similarity or perfect match.

For three microstates, the median $|\rho_{Spear.}|$ exceeded 0.9 for all eight subjects. The minimum $|\rho_{Spear.}|$ observed was 0.81 for Subject 6, which still reflects strong similarity. Regarding median $|\rho_{Spear.}|$ values, comparable results were obtained for the five-microstate condition (Table 2, bottom). Although the lowest $|\rho_{Spear.}| = 0.693$ was observed for Subject 8, this value still indicates a moderate to strong similarity.

However, correlation does not always adequately capture numerical differences in microstate spatial maps, and it is appropriate to evaluate their similarity from other perspectives as well. Fig. 7 illustrates two such examples. The spatial maps for the 5th microstate of Subject 3 (Fig. 7(a)) and the 4th microstate of Subject 8 (Fig. 7(b)) exhibit high correlations, yet visual differences are more effectively reflected by slightly higher GMD values.

Overall, GMD values were higher than those observed in the simulated data (Table 1), indicating greater numerical differences in the

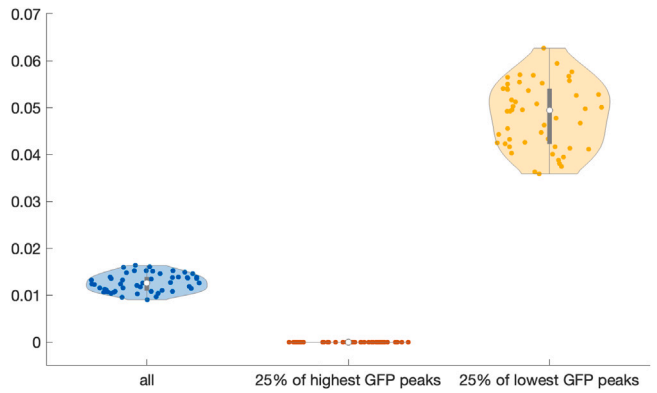


Fig. 5. The proportion of GFP peak time points with a leading Z-matrix element below 0.5. Results are shown for the time points corresponding to all GFP peaks (blue), the highest 25% of GFP peaks (orange), and the lowest 25% of GFP peaks (yellow).

Table 2
The median and interquartile range (IQR) of the global map dissimilarity (GMD) and the absolute Spearman correlation coefficient between the microstate spatial maps estimated by *modKM* and *IMSCAND* when applied to real EEG data of eight subjects. In the analysis, either three (top) or five (bottom) microstates were considered.

		subject							
3 microstates		1	2	3	4	5	6	7	8
GMD	median	0.09	0.14	0.31	0.23	0.03	0.31	0.13	0.21
	IQR	0.04	0.10	0.22	0.18	0.07	0.22	0.09	0.06
$\lvert\rho_{Spear.}\rvert$	median	0.99	0.98	0.91	0.97	1.00	0.93	0.98	0.96
	IQR	0.01	0.01	0.05	0.05	0.00	0.10	0.02	0.02
		subject							
5 microstates		1	2	3	4	5	6	7	8
GMD	median	0.23	0.15	0.15	0.17	0.05	0.44	0.33	0.43
	IQR	0.16	0.17	0.15	0.05	0.04	0.32	0.08	0.29
$\lvert\rho_{Spear.}\rvert$	median	0.97	0.98	0.98	0.98	0.99	0.88	0.95	0.90
	IQR	0.04	0.04	0.02	0.01	0.01	0.14	0.06	0.13

estimated spatial maps. This outcome is expected given the increased complexity of real EEG data and the iterative nature of both *modKM* and *IMSCAND*, which may converge to different local optima.

The final stage of the real-data analysis involved comparing microstate sequences. As in the simulated data analysis, all GFP peak time points were considered, as well as those corresponding to the highest 10% and lowest 10% of GFP peaks.

However, the proportion of GFP peak time points – whether all, or associated to the highest, or the lowest peaks – assigned to the same microstate by both algorithms varied across subjects and was generally lower compared to the simulated data analysis (Fig. 8). The observed better correspondence between *modKM* and *IMSCAND* microstate sequences when only three microstates were considered (Fig. 8(a)), was primarily due to the reduced number of microstates.

Upon closer examination, the results were found to be consistent with Table 2 and can be attributed to numerical differences in the estimated microstate spatial maps. For example, Subjects 1 and 5 exhibited high absolute correlation and low GMD between spatial microstate maps from *modKM* and *IMSCAND* (Table 2), resulting in a high degree of similarity in the obtained microstate sequences for both three and five microstates (Fig. 8). In contrast, for Subject 8, spatial microstate maps for microstates 3 and 5 were numerically different (Fig. 7(b)), as indicated by lower absolute correlations and higher GMD. Consequently, only 38% of GFP peak points assigned to microstate 5 by *modKM* were assigned to the same microstate by *IMSCAND* (Fig. 8(b), last row, fifth column). For microstate 3, this ratio was even lower and equal to 28% (Fig. 8(b), last row, third column). A similar pattern was observed when considering only time points associated to the highest or lowest 10% of GFP peaks.

These findings indicate that even small numerical differences in estimated microstate spatial maps can result in divergent microstate assignments between *modKM* and *IMSCAND*. These differences may naturally arise from a combination of several factors: although the simulated data are designed to approximate real EEG signals as closely as possible, they represent a simplified scenario and may not fully capture certain EEG-specific properties. Moreover, tensors constructed from real EEG signals in practice do not fully conform to the *IMSCAND* model due to noise or other factors disrupting the trilinear structure. Consequently, given the iterative nature of both *modKM* and *IMSCAND*, the algorithms may converge to slightly different local optima when estimating microstate spatial maps.

This conclusion was also confirmed by the following analysis: To avoid inferior results driven only by numerical reasons, we re-ran the *IMSCAND* algorithm with the matrix U fixed to the microstate spatial maps estimated by the *modKM* algorithm. We then compared the assigned microstate sequences from the re-run *IMSCAND* to those from the original *modKM*. In this case, both algorithms produced exactly the same microstate assignment, reflected by a GFP peak time point ratio of 1. This was true when considering both three and five microstates.

Similarly, we took the microstate spatial maps from the original *IMSCAND* algorithm and assigned the GFP peak time points to them, as is done within the *modKM* algorithm. The resulting microstate sequence was identical to those obtained from *IMSCAND* for both three and five microstates and all eight subjects.

7. Discussion

In this study, we derived the relationship between the polarity-invariant modified k-means algorithm (*modKM*) and the Implicit Slice Canonical Decomposition (*IMSCAND*). Although *modKM* operates on two-dimensional data and *IMSCAND* is designed for three-way tensors with symmetric slices, the optimization criteria used in both methods differ only by a constant, or are equal when the data are normalized.

We further demonstrated the equivalence of both approaches using simulated and real-world EEG data for EEG microstate detection. Due to its polarity-invariance property, *modKM* is a widely used tool for this application. In the simulated scenario, both algorithms accurately detected four microstates with spatial maps similar to the true prototype maps, as confirmed by a high Spearman correlation coefficient and low global map dissimilarity values. The assigned microstate sequences also exhibited high similarity with the original generated sequence. Moreover, *modKM* and *IMSCAND* produced identical or highly similar microstate sequences, as indicated by the proportion of GFP peak time points assigned to the same microstate being close to one.

For real-world EEG data, the true number of microstates is unknown and must be determined using appropriate methods. In this study, the *kneedle* algorithm was applied to the global explained variance (GEV) of both *modKM* and *IMSCAND*, while tensor-based core consistency diagnostics (CCD) was available only for *IMSCAND*. The *kneedle* approach identified five microstates for both methods, whereas CCD identified only three. Nevertheless, we would like to emphasize that this potential discrepancy arose solely from the distinct selection criteria employed by the two approaches. Namely, *kneedle* detects the most significant change in GEV which is a measure of variance. On other hand, CCD identifies the largest model that preserves the trilinear structure (Bro and Kiers, 2003). Therefore, it can be considered as a measure of structure. Consequently, these findings should be interpreted as complementary rather than competitive, indicating that the optimal number of microstates, depending on the evaluation criterion, ranges from three to five. To ensure a fair comparison, both algorithms were run for $K = 3$ and $K = 5$.

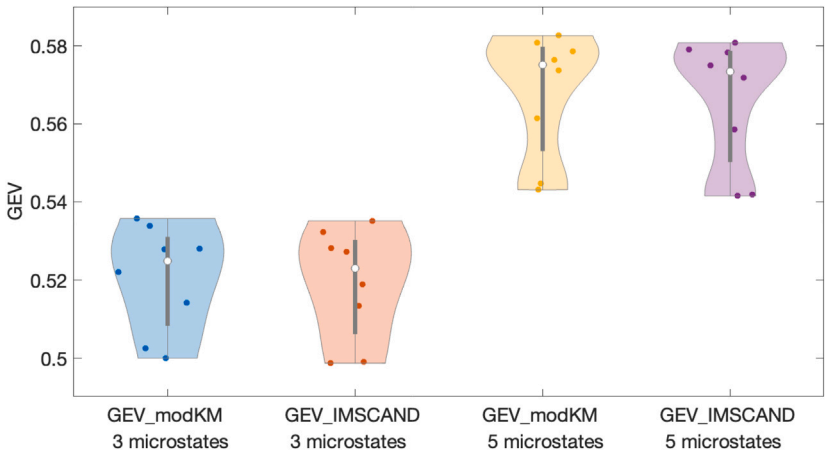


Fig. 6. Violin plot representing the global explained variance (GEV) of three (two violin plots on the left) or five microstates (two violin plots on the right) detected by *modKM* and *IMSCAND*.

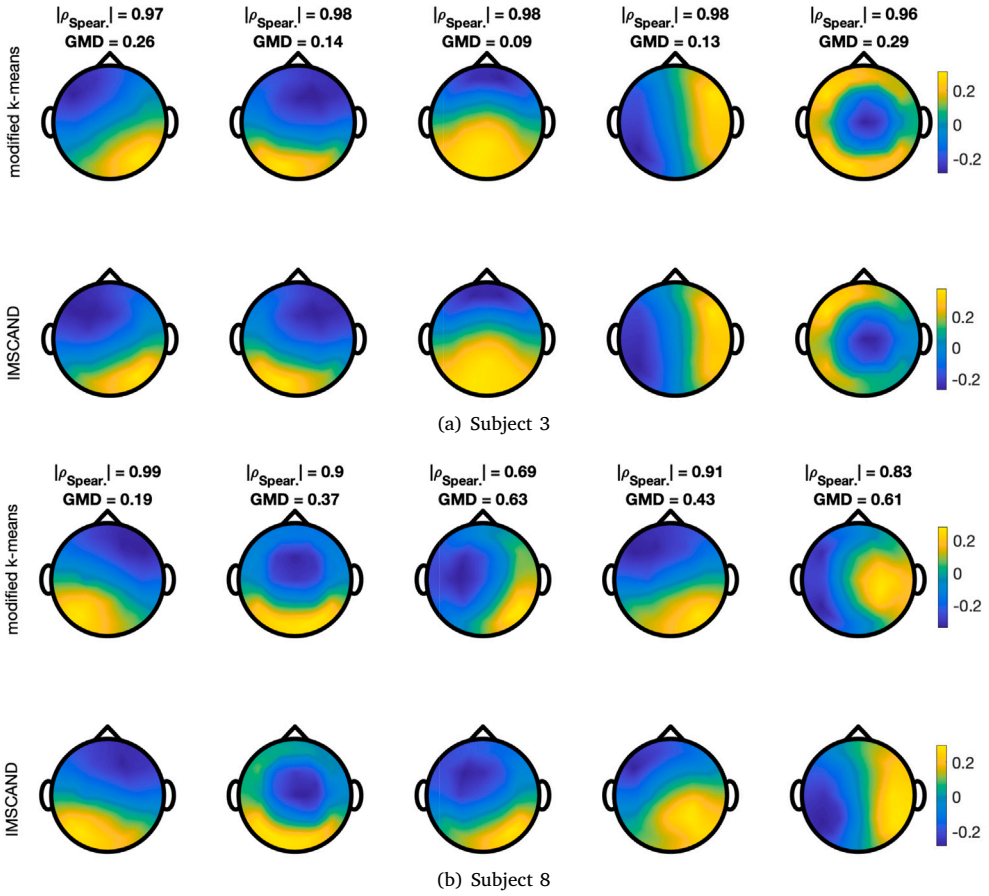


Fig. 7. EEG microstate spatial maps estimated by the *modKM* algorithm (top row) and *IMSCAND* (bottom row) with $K = 5$ for Subjects 3 and 8.

Spatial maps estimated by *modKM* and *IMSCAND* demonstrated fair to excellent similarity, with absolute Spearman correlation coefficients exceeding 0.69. However, visual inspection revealed small numerical differences that were not adequately captured by absolute $\rho_{Spear.}$, but were reflected in higher GMD values.

When comparing the assigned microstate sequences, the ratio of GFP peak time points assigned to the same microstate by both methods varied across subjects. For some subjects (e.g., Subjects 1 and 5), the estimated microstate sequences achieved a near-ideal match between methods. In contrast, for Subject 8, the ratio of GFP peak time points assigned to microstates 3 or 5 by both methods was below 40%. This discrepancy was consistent with the observed correlation and GMD between *modKM* and *IMSCAND* spatial maps for the respective microstates. These observations support the hypothesis that even small numerical differences in microstate spatial maps may result in a greater

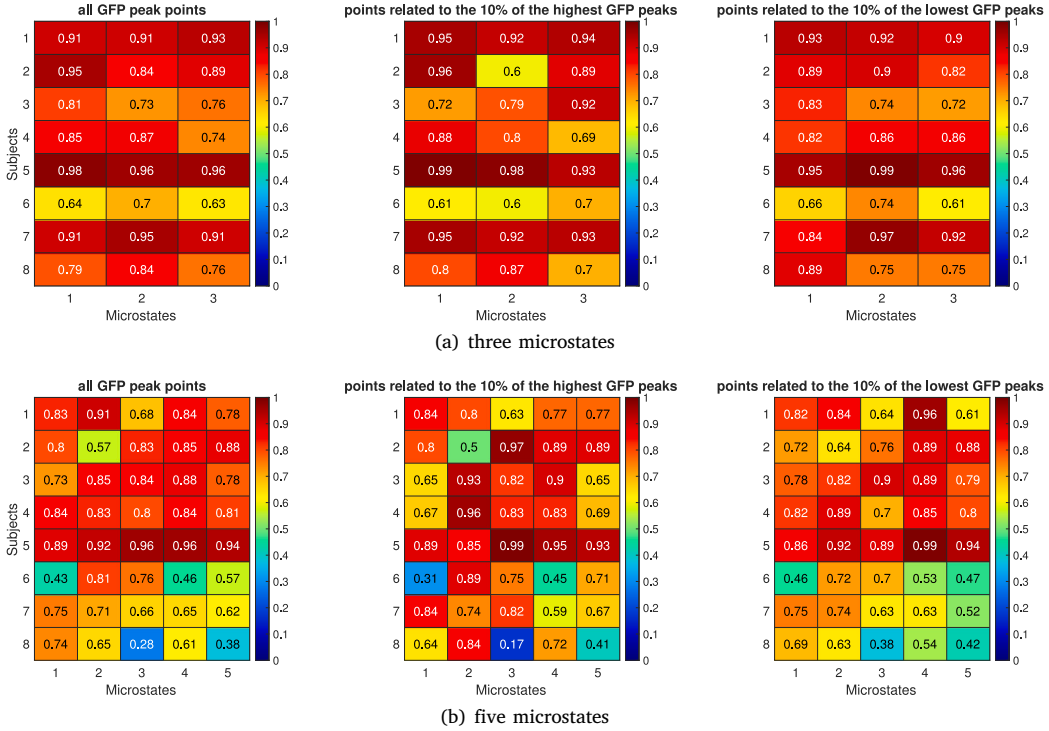


Fig. 8. Comparison between microstate sequences estimated by the *modKM* algorithm and *IMSCAND* with the number of microstates $K = 3$ (top) and $K = 5$ (bottom). Either all GFP peak time points (left), time points corresponding to the highest 10% (middle) of GFP peaks, or time points corresponding to the lowest 10% (right) of GFP peaks were considered. We would like to highlight that the kth microstate for two different subjects does not necessarily follow the same spatial map.

number of GFP peak time points being assigned to different microstates by the *IMSCAND* and *modKM* algorithms.

To verify this hypothesis, the *IMSCAND* algorithm was re-run with the matrix U fixed, setting its columns to the spatial maps derived from *modKM*. In this configuration, only the matrix Z, representing the microstate assignment, was estimated. Comparison of the resulting *IMSCAND* microstate sequence with the original *modKM* sequence showed an exact match. This analysis was also repeated in reverse order, using the spatial maps from *IMSCAND* for microstate assignment after applying the *modKM* algorithm. The newly obtained microstate sequences were again identical to those originally provided by *IMSCAND*.

The equivalence between *IMSCAND* and *modKM* may yield new perspectives for EEG microstate analysis. In cluster analysis, and consequently in EEG microstate detection, there is no generally established guideline for selecting the optimal number of clusters. Usually, multiple approaches should be employed, and the optimal value of K should be determined based on their results. Core consistency diagnostics (Bro and Kiers, 2003), automatic relevance determination (Mørup and Hansen, 2009), and other methods developed for tensor decomposition can provide supplementary information about the optimal number of microstates in real EEG data analysis, complementing traditional approaches for *modKM*.

Theoretical properties of the CP model represent a second set of potential benefits for EEG microstate analysis, specifically for *modKM*. The CP model (Carroll and Chang, 1970) is widely used for EEG signal analysis due to its interpretability and established criteria for assessing the essential uniqueness of the decomposition (Kruskal, 1977, 1989). As *IMSCAND* represents a specific case of the CP model, analogous results can be derived for this approach. However, theoretical results regarding the uniqueness of *modKM* or k-means in general are not yet available. In future research, the equivalence between *IMSCAND* and *modKM* could help derive similar uniqueness criteria for *modKM* as well.

A notable limitation of *IMSCAND* is its computational complexity. Although the algorithm converged within a reasonable time – approximately five minutes for five repetitions of *IMSCAND* with the number of microstates ranging from two to ten when applied to real EEG data[3] – potential challenges may arise when processing longer EEG recordings. This limitation highlights opportunities for future research and methodological improvements. Nevertheless, Bauckhage (2016) explores a similar relationship between k-means and matrix decomposition or archetypal analysis, suggesting that these latter methods can be addressed using simpler algorithms than currently used. Therefore, we believe that the equivalence derived in this study will help streamline the algorithm for *IMSCAND* in the future, as efficient algorithms for *modKM* and k-means are already available.

Our previous work (Rosipal et al., 2022) demonstrated that representing the multi-electrode EEG amplitude spectrum as a tensor and decomposing it into latent components provides more comprehensive information about signal structure and oscillatory rhythms than traditional time-space domain approaches, such as spatio-temporal decomposition, ICA, or PCA. This advantage may also extend to EEG microstates. Although the tensor input for *IMSCAND* is currently derived from EEG signals in the time-space domain, this approach represents an initial step toward broader applications of tensor decomposition methods in EEG microstate analysis, potentially enabling the integration of frequency information or other modalities, which is not feasible with the *modKM* algorithm.

[3] The experiments were conducted in MATLAB 2024b on a standard computer running OS X, with an M1 chip and 16 GB of memory.

Beyond the practical and methodological benefits outlined above, the equivalence between *modKM* and *IMSCAND* enables a reinterpretation of EEG microstates. Traditionally, since Lehmann's pioneering work (Lehmann, 1971), EEG microstates have been characterized primarily as spatial map clusters under a winner-takes-all assumption, where only one microstate is active at any given time point. However, recent studies (Mishra et al., 2020; Jajcay and Hlinka, 2023; Shaw et al., 2019; Haydock et al., 2025) have analyzed microstates in a continuous space, proposing mixed-state representations. The equivalence between *modKM* and *IMSCAND* facilitates this exploration by framing microstates as rank-1 latent components that describe the evolution of EEG spatial covariance, thereby integrating them into a broader tensor decomposition or blind-source separation framework.

The process of constructing the EEG tensor for the *IMSCAND* algorithm and its subsequent interpretation may not be immediately intuitive. Nevertheless, each slice of the resulting tensor corresponds to the autocovariance matrix of a spatial map, representing the EEG signal from M electrodes at a time point associated with a specific GFP peak and centered across electrodes. Collectively, the tensor captures the temporal dynamics of EEG spatial covariance, reflecting changes in the similarity between electrodes over time. Microstates identified from this tensor correspond to short intervals during which the EEG spatial covariance structure remains quasi-stable, consistent with the established EEG microstate theory.

Finally, we would like to highlight that the application of the *IMSCAND* algorithm within the EEG microstate detection pipeline affects only the clustering step. Standard preprocessing procedures used prior to *modKM* are equally applicable before *IMSCAND*, and no additional assumptions regarding data structure are required. Similarly, post-clustering analyses, such as the *kneedle* algorithm for selecting the appropriate number of microstates, are suitable for both *modKM* and *IMSCAND*. Standard microstate parameters – including global explained variance, microstate duration, and coverage – can be computed identically to conventional microstate analysis, regardless of whether *modKM* or *IMSCAND* is employed.

8. Conclusion

The modified k-means algorithm (*modKM*) is widely used for detecting EEG microstates in resting-state EEG data. In contrast, the Implicit Slice Canonical Decomposition (*IMSCAND*) was introduced to detect similarity between data described by different sets of variables (Selee et al., 2007). Although both algorithms were developed for different purposes and operate on slightly different versions of the same data, we theoretically derived equivalence between them. These theoretical findings were further validated using simulated EEG data with known and well-controlled properties. Both *modKM* and *IMSCAND* successfully detected four microstates with true prototype spatial maps, and the correspondence between the estimated microstate sequences was high both between methods and when compared to the original simulated sequence.

Both methods were also applied to real EEG data to detect subject-specific EEG microstates. In an optimal noiseless scenario, both approaches would be expected to yield identical results. However, due to noise, the complex structure of the EEG signal, numerical factors, and the iterative nature of both algorithms, small differences in the solutions were expected.

Applying the *IMSCAND* algorithm to EEG microstate detection may provide new insights by leveraging the advantages of tensor decomposition. These include the essential uniqueness of CP decomposition, the availability of diverse criteria for selecting the optimal number of latent components and consequently microstates, and established methods for tensor post-hoc analysis. Conversely, this relationship may also facilitate the development of more efficient and faster algorithms for *IMSCAND* or other tensor decomposition methods, given the availability of efficient algorithms for *modKM* and k-means clustering. Finally, the derived equivalence enables a shift in the understanding of EEG microstates from cluster representatives of spatial maps to latent tensor components, potentially opening new avenues to study EEG microstates from a continuous perspective.

CRediT authorship contribution statement

Zuzana Rošťáková: Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Roman Rosipal:** Writing – review & editing, Supervision, Data curation.

Ethics statement

In this study, we analyzed multichannel EEG signals from eight healthy volunteers. Ethical approval for this study was granted by the Ethics Committee of the Institute of Measurement Science, Slovak Academy of Sciences on August 8, 2025. The subjects received an explanation of the purpose and procedures of the study, which they fully understood, and gave written informed consent to participate. The obtained data were handled in compliance with GDPR and institutional data protection policies.

Declaration of competing interest

The authors declare no competing financial interests.

Acknowledgments

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 code and availability

The simulated and real EEG data supporting the findings of this study are available from the corresponding author upon request.
The custom code for the *IMSCAND* algorithm is available from the MATLAB Central File Exchange repository: https://www.mathworks.com/matlabcentral/fileexchange/183118-imscand-algorithm-for-eeg-microstate-analysis.

References

Akoglu, H., 2018. User's guide to correlation coefficients. Turk. J. Emerg. Med. 18 (3), 91–93. http://dx.doi.org/10.1016/j.tjem.2018.08.001.

Bauckhage, C., 2015. K-means clustering is matrix factorization. arXiv: Machine Learning. URL https://api.semanticscholar.org/CorpusID:88514257.

Bauckhage, C., 2016. K-means clustering via the Frank-Wolfe algorithm. In: Proc. of KDML-LWDA. pp. 311–322.

Brain Products, GmbH, 2013. *BrainVision Analyser 2*.

Bro, R., 1997. PARAFAC. Tutorial and applications. Chemometr. Intell. Lab. Syst. 38 (2), 149–171. http://dx.doi.org/10.1016/S0169-7439(97)00032-4, URL https://www.sciencedirect.com/science/article/pii/S0169743997000324.

Bro, R., Kiers, H.A.L., 2003. A new efficient method for determining the number of components in PARAFAC models. J. Chemom. 17 (5), 274–286. http://dx.doi.org/10.1002/cem.801.

Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. Comput. Intell. Neurosci. 2011 (1), 813870. http://dx.doi.org/10.1155/2011/813870.

Carroll, J.D., 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. http://dx.doi.org/10.1007/BF02310791.

Cichocki, A., Zdunek, R., Phan, A.H., Amari, S., 2009. Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation. John Wiley & Sons, United Kingdom.

Evetović, N., Rosipal, R., Polyanskaya, A., Rošťáková, Z., Trejo, L., 2025. Assessment of mental fatigue in healthy participants during extended BCI-HMD sessions. In: 2025 IEEE International Conference on Systems, Man, and Cybernetics. SMC, pp. 4366–4369. http://dx.doi.org/10.1109/SMC58881.2025.11343462.

Harshman, R.A., 1970. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis. UCLA Work. Pap. Phon. 16, 1–84.

Haydock, D., Kadir, S., Leech, R., Nehaniv, C.L., Antonova, E., 2025. EEG microstate syntax analysis: A review of methodological challenges and advances. NeuroImage 309, 121090. http://dx.doi.org/10.1016/j.neuroimage.2025.121090.

Hu, W., Zhang, Z., Zhang, L., Huang, G., Li, L., Liang, Z., 2022. Microstate detection in naturalistic electroencephalography data: A systematic comparison of topographical clustering strategies on an emotional database. Front. Neurosci. 16, http://dx.doi.org/10.3389/fnins.2022.812624.

Jajcay, N., Hlinka, J., 2023. Towards a dynamical understanding of microstate analysis of M/EEG data. NeuroImage 281, 120371. http://dx.doi.org/10.1016/j.neuroimage.2023.120371.

Khanna, A., Pascual-Leone, A., Michel, C.M., Farzan, F., 2015. Microstates in resting-state EEG: Current status and future directions. Neurosci. Biobehav. Rev. 49, 105–113. http://dx.doi.org/10.1016/j.neubiorev.2014.12.010.

Kruskal, J.B., 1977. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algebra Appl. 18 (2), 95–138. http://dx.doi.org/10.1016/0024-3795(77)90069-6.

Kruskal, J.B., 1989. Rank, decomposition, and uniqueness for 3-way and N-way arrays. In: Multiway Data Analysis. North-Holland Publishing Co. Netherlands, pp. 7–18.

Lehmann, D., 1971. Multichannel topography of human alpha EEG fields. Electroencephalogr. Clin. Neurophysiol. 31 (5), 439–449. http://dx.doi.org/10.1016/0013-4694(71)90165-9.

Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48 (6), 609–621. http://dx.doi.org/10.1016/0013-4694(80)90419-8.

Lloyd, S., 1982. Least squares quantization in PCM. IEEE Trans. Inform. Theory 28 (2), 129–137. http://dx.doi.org/10.1109/TIT.1982.1056489.

Michel, C.M., Koenig, T., 2017. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. Neuroimage 180, 577–593. http://dx.doi.org/10.1016/j.neuroimage.2017.11.062.

Mishra, A., Englitz, B., Cohen, M.X., 2020. EEG microstates as a continuous phenomenon. NeuroImage 208, 116454. http://dx.doi.org/10.1016/j.neuroimage.2019.116454.

Miwakeichi, F., Martínez-Montes, E., Valdés-Sosa, P.A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y., 2004. Decomposing EEG data into space–time–frequency components using Parallel Factor Analysis. NeuroImage 22 (3), 1035–1045. http://dx.doi.org/10.1016/j.neuroimage.2004.03.039.

Mørup, M., Hansen, L.K., 2009. Automatic relevance determination for multi-way models. J. Chemom. 23 (7–8), 352–363. http://dx.doi.org/10.1002/cem.1223.

Murray, M., Brunet, D., Michel, C., 2008. Topographic ERP analyses: A step-by-step tutorial review. Brain Topogr. 20, 249–264. http://dx.doi.org/10.1007/s10548-008-0054-5.

Pascual-Marqui, R., Michel, C., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42 (7), 658–665. http://dx.doi.org/10.1109/10.391164.

Poulsen, A.T., Pedroni, A., Langer, N., Hansen, L.K., 2018. Microstate EEGlab toolbox: An introductory guide. http://dx.doi.org/10.1101/289850, BioRxiv. URL https://www.biorxiv.org/content/early/2018/03/27/289850.

Rosipal, R., Rošťáková, Z., Trejo, L.J., 2022. Tensor decomposition of human narrow-band oscillatory brain activity in frequency, space and time. Biol. Psychol. 169, 108287. http://dx.doi.org/10.1016/j.biopsycho.2022.108287.

Schober, P., Boer, C., Schwarte, L., 2018. Correlation coefficients: Appropriate use and interpretation. Anesth. Analg. 126, 1. http://dx.doi.org/10.1213/ANE.0000000000002864.

Selee, T.M., Kolda, T.G., Kegelmeyer, W.P., Griffin, J.D., 2007. Extracting clusters from large datasets with multiple similarity measures using IMSCAND. In: Parks, M.L., Collis, S.S. (Eds.), CSRI Summer Proceedings 2007. Tech. Rep. SAND2007-7977, Sandia National Laboratories, pp. 87–103, URL http://www.cs.sandia.gov/CSRI/Proceedings/CSRI2007.pdf.

Shaw, S.B., Dhindsa, K., Reilly, J.P., Becker, S., 2019. Capturing the forest but missing the trees: Microstates inadequate for characterizing shorter-scale EEG dynamics. Neural Comput. 31 (11), 2177–2211. http://dx.doi.org/10.1162/neco_a_01229.

Skrandies, W., 1989. Data reduction of multichannel fields: global field power and principal component analysis. Brain Topogr. 2 (1), 73–80.

Tait, L., Zhang, J., 2022a. MEG cortical microstates: Spatiotemporal characteristics, dynamic functional connectivity and stimulus-evoked responses. NeuroImage 251, 119006. http://dx.doi.org/10.1016/j.neuroimage.2022.119006.

Tait, L., Zhang, J., 2022b. +microstate: A MATLAB toolbox for brain microstate analysis in sensor and cortical EEG/MEG. NeuroImage 258, 119346. http://dx.doi.org/10.1016/j.neuroimage.2022.119346.

Yuan, H., Zotev, V., Phillips, R., Drevets, W.C., Bodurka, J., 2012. Spatiotemporal dynamics of the brain at rest - Exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. NeuroImage 60 (4), 2062–2072. http://dx.doi.org/10.1016/j.neuroimage.2012.02.031.