SCIENCE CHINA
Technological Sciences

Special Topic: Tensor Methods in Machine Learning
• Article •

https://doi.org/10.1007/s11431-020-1839-5

On the robustness of EEG tensor completion methods

DUAN Feng[1*], JIA Hao[1], ZHANG ZhiWen[1], FENG Fan[1], TAN Ying[1], DAI YangYang[1], CICHOCKI Andrzej[2,3,4,5], YANG ZhengLu[6], CAIAFA Cesar F.[1,7*], SUN Zhe[1,8*] & SOLÉ-CASALS Jordi[1,9,10*]

[1] *College of Artificial Intelligence, Nankai University, Tianjin 300350, China;*
[2] *Skolkowo Institute of Science and Technology, Moscow 121205, Russia;*
[3] *College of Computer Science, Hangzhou Dianzi University, Hangzhou 310018, China;*
[4] *Department of Informatics, Nicolaus Copernicus University, Torun 87-100, Poland;*
[5] *Systems Research Institute, Polish Academy of Sciences, Warsaw 01-447, Poland;*
[6] *College of Computer Science, Nankai University, Tianjin 300350, China;*
[7] *Instituto Argentino de Radioastronomía IAR-CCT La Plata, CONICET/CIC-PBA/UNLP, Villa Elisa 1894, Argentina;*
[8] *Computational Engineering Applications Unit, Head Office for Information Systems and Cybersecurity, RIKEN, Wako-Shi 351-0106, Japan;*
[9] *Department of Psychiatry, University of Cambridge, Cambridge CB2 8AH, UK;*
[10] *Data and Signal Processing Research Group, University of Vic-Central University of Catalonia, Catalonia 08500, Spain*

Received November 16, 2020; accepted April 21, 2021; published online July 27, 2021

During the acquisition of electroencephalographic (EEG) signals, data may be missing or corrupted by noise and artifacts. To reconstruct the incomplete data, EEG signals are firstly converted into a three-order tensor (multi-dimensional data) of shape time × channel × trial. Then, the missing data can be efficiently recovered by applying a tensor completion method (TCM). However, there is not a unique way to organize channels and trials in a tensor, and different numbers of channels are available depending on the EEG setting used, which may affect the quality of the tensor completion results. The main goal of this paper is to evaluate the robustness of EEG completion methods with several designed parameters such as the ordering of channels and trials, the number of channels, and the amount of missing data. In this work, the results of completing missing data by several TCMs were compared. To emulate different scenarios of missing data, three different patterns of missing data were designed. Firstly, the amount of missing data on completion effects was analyzed, including the time lengths of missing data and the number of channels or trials affected by missing data. Secondly, the numerical stability of the completion methods was analyzed by shuffling the indices along channels or trials in the EEG data tensor. Finally, the way that the number of electrodes of EEG tensors influences completion effects was assessed by changing the number of channels. Among all the applied TCMs, the simultaneous tensor decomposition and completion (STDC) method achieves the best performance in providing stable results when the amount of missing data or the electrode number of EEG tensors is changed. In other words, STDC proves to be an excellent choice of TCM, since permutations of trials or channels have almost no influence on the complete results. The STDC method can efficiently complete the missing EEG signals. The designed simulations can be regarded as a procedure to validate whether or not a completion method is useful enough to complete EEG signals.

electroencephalogram, corrupted data, missing data, tensor completion, tensor decomposition

Citation: Duan F, Jia H, Zhang Z W, et al. On the robustness of EEG tensor completion methods. Sci China Tech Sci, 2021, 64, https://doi.org/10.1007/s11431-020-1839-5

*Corresponding authors (email: duanf@nankai.edu.cn; ccaiafa@fi.uba.ar; zhe.sun.vk@riken.jp; jordi.sole@uvic.cat)

1 Introduction

An electroencephalogram (EEG) can be used to record the electrical activity from a person's brain [1,2]. With the analysis of EEG signals, certain brain activities such as motor imagery and visual evoked potentials can be decoded from the brain [3–6]. However, EEG signals may be missing or corrupted due to a temporary disconnection of electrodes during the acquisition process, resulting in the loss of information and seriously affecting the subsequent analysis of EEG data [7]. For example, event-related desynchronization/synchronization (ERD/ERS) is one of the brain activities concerning the decrease or increase of power in the motor cortex [8,9]. The ERD can be detected as the changes of variances of EEG signals in alpha and beta rhythms [10]. The ERD may exactly occur in the acquisition process of missing data, which has an impact on the accuracy of ERD detection. Therefore, completing missing data and minimizing errors between original EEG signals and completed EEG signals constitute a crucial step in EEG analysis.

Interpolation is a basic method used to complete one-dimensional or multi-dimensional missing signals in a form of time series [11]. However, because the information used for interpolation is local, this method may fall into a local trap of complex signals. Gunnarsdottir et al. [12] proposed a method to complete the missing EEG signals under the assumption that some electrodes are missing. The EEG data were represented as sequences of linear time-variant models. Since EEG signals in some electrodes are entirely unknown, they proposed an estimated network model to complete missing data with the simultaneously acquired two-dimensional EEG signals. Tensor completion methods (TCMs), on the other hand, complete the missing data by using global and multi-dimensional information relying on tensor factorization models [13]. Instead of copying from adjacent signals, TCMs impose low-rankness of the tensors involved in the factorization. TCMs utilize the information from not only the one-dimensional EEG signals or simultaneously acquired two-dimensional EEG signals, but also EEG signals from other trials which are multi-dimensional. In the analysis of EEG data, TCMs can also incorporate the multiway nature of data, similar to the applications of image processing and social network data [14–19]. Solé-Casals et al. [7] proposed to reconstruct EEG signals with missing data using TCMs in a brain computer interface (BCI) context. Compared to the simple interpolation methods, the TCMs can minimize the errors between the original and completed EEG signals, which can also improve the performance of EEG signal analysis.

Despite the successful completion of missing data with TCMs on EEG signals, there are still some problems with the procedure. It is important to note that two main differences exist between the completion on EEG signals and images. The differences are mainly reflected during the transformation from EEG signals or images to tensors. One is related to the permutations of channels in EEG signals. The height and width of images correspond to the channel and time of EEG signals in a tensor. In images, the sequences of pixels along the height or width are fixed. However, in EEG signals, the sequences of channels can be randomly permuted without influencing the analysis results. The permutation of channel sequences may have an influence on the numerical stability when applying the completion methods. The other difference concerns the number of channels in EEG signals. In images, the number of pixels along the height or width is greater than ten, depending on the camera settings. In EEG signals, the number of EEG channels ranges from one to one hundred without a lower limit. Multiple pixels ensure that TCMs can find a valid relationship within image tensors. In other words, TCMs may fail on EEG signals when the number of EEG channels is smaller than a certain threshold.

These two differences lead to the completion problems, that is, not all TCMs can efficiently complete the missing EEG signals. Consequently, it is necessary to find out which TCM is the most appropriate for the completion of missing EEG signals. In order to minimize influences induced by the permutations of EEG channels/trials or the number of EEG channels, different available TCMs are compared when used to complete the missing EEG signals. We design simulations to emulate different practical scenarios of missing data in EEG signals and compare the ability of each TCM to recover the missing data. The simulations of missing data focus on the following three problems: (1) how the amount of missing data in EEG signals influences the completion effects; (2) the influence of EEG tensor ordering on the numerical stability; (3) the minimal number of EEG channels necessary to achieve acceptable completion results.

Thus, to solve these three problems, three types of experiments are proposed to measure and evaluate the completion effects of TCMs on EEG signals. The three experiments are (1) the amount of missing data; (2) the initial condition of the EEG signals (with different tensor permutations); (3) the number of channels of the EEG signals, separately. The completion effects of the prevalent TCMs and those of a basic interpolation method are deeply analyzed through simulations.

In the following sections, we first introduce the used EEG dataset, how we design the experiments, basic concepts on TCMs, and how to evaluate the completion effects in Sect. 2. Second, we analyze the results of the simulations in Sect. 3. We then discuss the assumptions of the proposed simulations in Sect. 4. Finally, Sect. 5 is a summary of this work.

2 Materials and methods

EEG completion is the task related to estimating missing values in EEG signals through completion methods. In EEG completion, the first step is to tensorize the original EEG signals into a 3-dimensional EEG tensor. The initial EEG signals are given as multi-channel time series (a two-dimensional dataset). In most EEG experiments, subjects are expected to execute several repeated actions, such as imagining the movement of the hand or the extension/flexion of the elbow. EEG signals are used to identify the pattern distribution of certain brain activities. In EEG experiments, a trial comprises of the signals recorded from the start to the end of an action. EEG signals can be arranged as a 3-dimensional tensor of shape time × channel × trial with more than one trial. We term the missing data in a tensor as missing entries. To facilitate the description of the simulations of missing entries, the notations that are used the most often throughout this work are described in Table 1. In the EEG analysis, we can take advantage of the 3-D EEG tensor to complete missing or corrupted entries with TCMs.

To evaluate the TCMs, Solé-Casals et al. [7] applied a simulation of missing entries in the EEG data. Specifically, we have a 3-dimensional tensor $\mathcal{X}$ containing intact signals (full EEG data tensor). Then, missing entries are randomly selected in $\mathcal{X}$. This random selection is defined by a mask $\mathcal{O}$. Each entry in the mask $\mathcal{O}$ is either 0 or 1. Information in entries with 1 is known or intact; on the other hand, information in entries with 0 is deleted or contaminated. The incomplete tensor is $\mathcal{Y} = \mathcal{X} * \mathcal{O}$, where $*$ denotes the elementwise product. TCMs are then applied on $\mathcal{Y}$ to recover the information of the missing entries. Through the process, the recovered/completed tensor $\hat{\mathcal{X}}$ can be obtained.

In the following subsections, the used dataset is first introduced. Then, the method to simulate incomplete EEG data tensors is described. Finally, the related TCMs are introduced and a measurement of the reconstruction quality is properly defined.

2.1 Dataset description

In most cases, EEG signals are acquired to analyze the activities of the brain. These are primarily associated with changes of power in the electrical activity in certain areas of the brain. In this paper, the associated brain area is the motor cortex. A dataset with 15 subjects was used, obtained from ref. [20][1)]. This dataset contains both EEG signals and simultaneously acquired hand positions. The channels of the used data include F_3, F_z, F_4, FC_z, C_1, C_z, C_2, CP_z, P_3, P_z, and P_4 (Figure 1). EEG signals are first downsampled to F_s=256 Hz.

Table 1 Description of notations used in this work

Notation	Description
Fs	Fs= 256 Hz is the sample rate/frequency of EEG signals
$\mathcal{X}$	Tensors containing full EEG signals
$\mathcal{O}$	Masks indicating positions of the missing entries
$\mathcal{Y}$	Tensors containing missing entries, equaling to $\mathcal{X} * \mathcal{O}$
$\hat{\mathcal{X}}$	Completed tensors
N_{channel}	Number of channels of EEG signals
N_{time}	Number of sample points of EEG signals
N_{trial}	Number of trials of EEG signals
N_{mTL}	Time length of missing entries with unit of second (s)
N_{mCT}	Number of time series containing missing entries

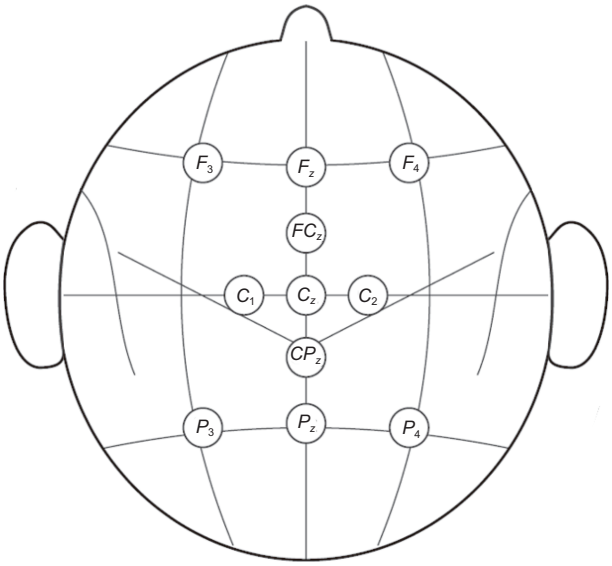

Figure 1 Channel locations of EEG signals. The distribution of electrodes is in accordance with the 10–20 international system of electrode placement, which is an internationally recognized method to describe the location of scalp electrodes.

Here F_s denotes the sample frequency. The hand positions were recorded using an exoskeleton sensor. The subjects were asked to implement seven different actions, which included six movement states and the rest state. In our experiments, two movement states were adopted (elbow flexion, elbow extension) along with the rest state (resting). For the movement states, the movement onset is the time when the subjects began to implement their movements, which can be located from the hand positions. The brain remains in a rest state prior to the movement onset. The motor cortex becomes active around the movement onset. During the data acquisition, the trial started after a beep sound. Two seconds (2 s) later, a cue was shown indicating that the action should be executed. Subjects were then supposed to execute the corresponding action.

1) Upper limb movement decoding from EEG (001-2017), 15 subjects. http://bnci-horizon-2020.eu/database/data-sets

The EEG signals in EEG tensors were extracted from signals around the movement onset. Specifically, a 1-order Savitzky-Golay finite impulse response filter was first used to smooth the hand position acquired with an exoskeleton sensor. The frame length of the filter was 31. Note that a higher or lower frame length will lead to inaccurate detection of the movement onset. A higher frame will induce a latency on the located movement onsets and a lower frame cannot ensure the smoothness of the filtered hand positions. The movement onsets were located from the filtered smooth hand positions. EEG signals 2 s before and 1 s after the movement onset were cut off and were therefore selected to generate the EEG tensors. Figure 2 illustrates how the EEG tensors are generated from the original EEG signals. If subjects are supposed to stay at rest, the obtained signals start at 0.5 s post-cue. By concatenating trials from the same subject, EEG tensors for each subject can be acquired. Each EEG tensor corresponds to a subject's certain executed action. 15 subjects repeat each action 60 times, so the number of trials was 60 for each subject. Due to disruptions occurring in some trials when locating the movement onset, 50 trials were randomly selected without disruptions for each action. Therefore, the size of the EEG tensor $\mathcal{X}$ is $N_{\text{channel}} = 11$, $N_{\text{time}} = 3\text{ s} \times 256\text{ Hz} = 768$, $N_{\text{trial}} = 50$.

In the processing of the EEG dataset above, the data belonging to the movement execution was adopted. In movement execution, the movement onset can be captured and located with the exoskeleton sensor. The localization of the movement onset has two advantages for the implementation and evaluation of the simulations:

(1) The EEG tensors can cover active brain activities because the motor cortex becomes active around the movement onset;

(2) The premovement encoding method can be directly applied to the EEG tensors, which can classify the EEG data in the movement state and the rest state.

In Sect. 2.4.2, further details about the premovement encoding method are given.

2.2 Simulation of missing entries in EEG data tensors

In most TCM-related studies, a simulation of missing entries is frequently applied because it is necessary to compare the information present in the completed tensors against the information in the original tensors [7]. In Figure 3, a procedure of the simulation of an incomplete EEG data tensor and the procedure for completing its missing entries is provided.

The first step in a simulation of an incomplete EEG data tensor is to select the missing entries (Figure 3). By default, it is assumed that the missing entries in the time series are continuous. The selection of missing entries comprises of two steps (1) select the time series that contains the missing entries; and (2) decide the time length of the missing entries as well as randomly choose their start time.

Selection of time series. The missing signals are characteristic in which they may occur at any channel or trial during the acquisition of EEG signals. In order to simulate this characteristic, we define a two-dimensional matrix

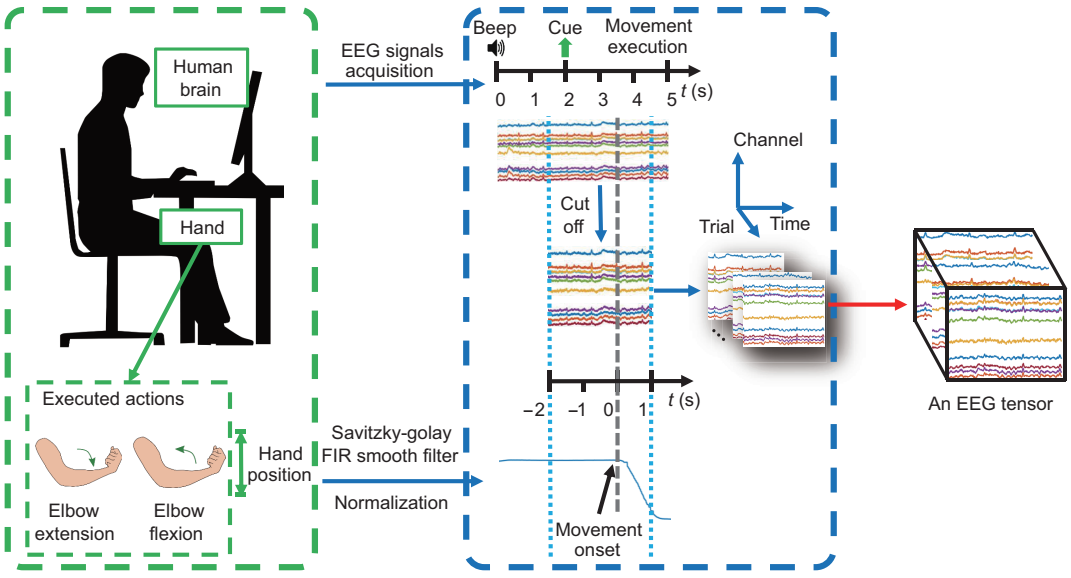

Figure 2 (Color online) Procedure of generating an EEG tensor from the original dataset. After a beep, subjects were asked to prepare the movement. Two seconds (2 s) later, a cue indicating the action to be performed was given. Subjects immediately executed the action after the cue. EEG signals and hand position were simultaneously acquired. Movement onsets were decided using the hand positions. EEG signals 2 s before and 1 s after movement onset were cut-off and were thus selected to generate the EEG tensors.

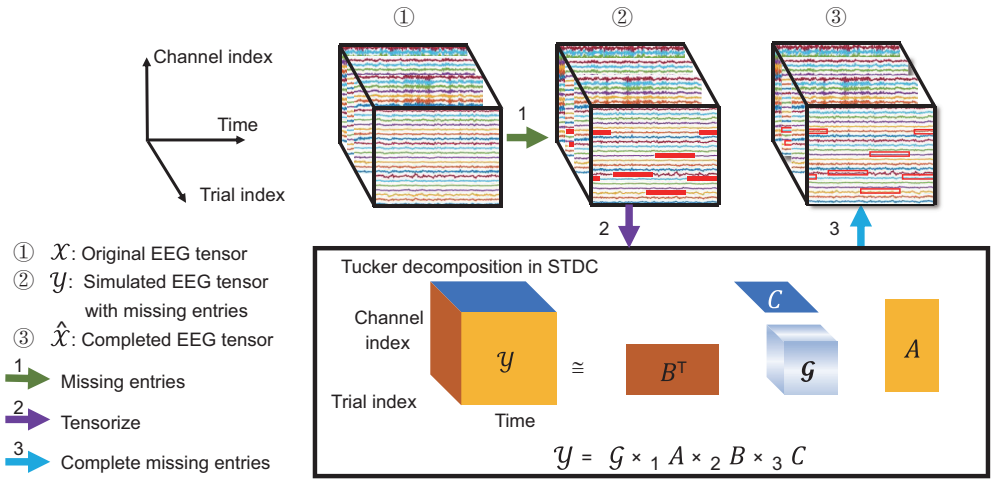

Figure 3 (Color online) The simulation and completion procedures for missing entries in EEG tensors. The first step of the simulation is to select the entries with missing information. Then, TCMs are applied to complete the missing entries in the EEG tensor. The Tucker model is one of the basic models used in STDC. Details about the Tucker decomposition are given in eq. (6).

$M \in \mathbb{R}^{N_{\text{channel}} \times N_{\text{trial}}}$ and initialize M with $M_{j,k} = 1$, where $j = 1, \ldots, N_{\text{channel}}$, $k = 1, \ldots, N_{\text{trial}}$. $M_{j,k} = 1$ means that the time series at the jth channel and kth trial does not contain any missing entries. In contrast, $M_{j,k} = 0$ means that the time series at the jth channel and kth trial contains missing entries. The values in M have the following definition:

$$M_{j,k} = \begin{cases} 1, & \text{The time series does not have missing entries,} \\ 0, & \text{The time series has missing entries.} \end{cases} \tag{1}$$

We denote the number of '0's in M as N_{mCT}, indicating how many time series contain missing entries.

Decision of missing entries' positions. We suppose that a time series at the jth channel and kth trial has missing entries ($M_{j,k} = 0$). The time length of the missing entries is denoted as N_{mTL}. A random start time t_{random} is decided. The information between t_{random} and $t_{\text{random}} + N_{\text{mTL}}$ is set to be discarded during this time series. For other time series that contain missing entries, t_{random} are totally different because the values of t_{random} are all randomly selected.

With the two steps, the mask O in the simulation can be expressed as

$$O = \begin{cases} 1, & \text{Information in this entry is not missing,} \\ 0, & \text{Information in this entry is missing.} \end{cases} \tag{2}$$

In comparison to the original EEG tensor $\mathcal{X}$, information in the missing entries of $\mathcal{Y}$ is lost and set to zero. The elementwise relationship between $\mathcal{X}$, $\mathcal{Y}$, and O can be written as follows:

$$y_{ijk} = x_{ijk} o_{ijk}. \tag{3}$$

Previous studies have proven that TCMs are able to recover missing entries in EEG tensor $\mathcal{Y}$ [7]. In this work, in order to measure whether or not a TCM is suitable for EEG completion, three different simulations are designed focusing on (1) the number of missing entries; (2) the permutation of channels/trials of EEG tensors; and (3) the number of channels of EEG tensors.

2.2.1 *Simulation I: Number of missing entries*

The number of missing entries can be described using two values: N_{mCT} and N_{mTL}. In order to observe how the completion effects change as the number of missing entries increases, two methods were carried out to increase the number of missing entries:

(1) Set N_{mTL} to a constant and increase N_{mCT};

(2) Set N_{mCT} to a constant and increase N_{mTL}.

Therefore, there are two cases in simulation I: increasing N_{mCT} and increasing N_{mTL}.

(1) Increasing N_{mCT}: Set $N_{\text{mTL}} = 0.25$ and 0.5 s separately; increase N_{mCT} from 5 to 100 with step size 5.

(2) Increasing N_{mTL}: Set $N_{\text{mCT}} = 10$ and 20 separately; increase N_{mTL} from 0.125 to 2 s with step size 0.125 s.

2.2.2 *Simulation II: Numerical stability of completion methods*

In this simulation, the stability of the completion methods is explored when the initial condition changes. The permutation of the EEG tensors is changed while the rank and the content of the tensors remain unchanged. This simulation is applied to the EEG tensor $\mathcal{Y}$ with $N_{\text{mCT}} = 5$ and $N_{\text{mTL}} = 0.25$ s. The permutation of channels of $\mathcal{Y}$ is randomly shuffled, thus ob-

taining the shuffled EEG tensor $\mathcal{Y}'_{\text{channel}}$. $\hat{\mathcal{X}}$, $\hat{\mathcal{X}}'_{\text{channel}}$ are the completed EEG tensor for $\mathcal{Y}$ and $\mathcal{Y}'_{\text{channel}}$, separately. Subsequently, the permutation of $\hat{\mathcal{X}}'_{\text{channel}}$ to $\hat{\mathcal{X}}_{\text{channel}}$ is recovered, i.e., $\hat{\mathcal{X}}_{\text{channel}}$ has the same channel sequence and trial sequence with $\hat{\mathcal{X}}$.

In addition to the permutation of channels, we note that the permutation of trials in EEG tensors also shows a random sequence. Consequently, the permutation of trials in EEG tensors is also compared. Similar to the case with channels, the related variables are $\mathcal{Y}'_{\text{trial}}$, $\hat{\mathcal{X}}'_{\text{trial}}$, and $\hat{\mathcal{X}}_{\text{trial}}$.

2.2.3 *Simulation III: Number of channels of EEG tensors*

This simulation aims to display how N_{channel} influences the completion effects of TCMs. We need to change N_{channel} from 1 to 11 (N_{channel} is fixed to 11 in simulations I and II). Since the EEG tensor must contain missing entries when $N_{\text{channel}} = 1$, there is only one channel containing missing entries. We set $N_{\text{mCT}} = 1$ and $N_{\text{mTL}} = 0.25$ s. For convenience, we denote the channel which contains missing entries as Channel 1. When N_{channel} increases from 1 to 11, other channels are appended after Channel 1 referred to as Channel 2, …, Channel 11.

2.3 Tensor completion methods

Tensor completion methods play a critical role in recovering information from missing entries. Specifically, TCMs utilize the inner relationship within the tensors to complete the missing entries instead of directly copying information from adjacent entries. The inner relationship of the tensors used in most TCMs is the rank of the tensors. Tucker decomposition and canonical polyadic (CP) decomposition are two basic tensor decomposition methods [21, 22]. In the following statement on TCM-related concepts, $\mathcal{X}$ is the original tensor without missing entries; $\mathcal{Y}$ is the tensor with missing entries; $\mathcal{Y} = \mathcal{X} * \mathcal{O}$; finally, $\hat{\mathcal{X}}$ is the completed tensor ($\mathcal{X}, \mathcal{Y}, \hat{\mathcal{X}} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$, where N_n is the length of the ith dimension, $i = 1, 2, 3$).

CP decomposition. In CP decomposition [13], the tensor $\mathcal{Y}$ can be factorized as

$$\mathcal{Y} \approx [\![A, B, C]\!] = \sum_{r=1}^{R} a_r \circ b_r \circ c_r, \tag{4}$$

where R is a positive integer; $a_r \in \mathbb{R}^{N_1}$, $b_r \in \mathbb{R}^{N_2}$, and $c_r \in \mathbb{R}^{N_3}$ where $r = 1, 2, \ldots, R$; finally, $\circ$ denotes the outer product of vectors. In an elementwise form, CP decomposition can be written as

$$y_{i_1 i_2 i_3} \approx \sum_{r=1}^{R} a_{i_1 r} b_{i_2 r} c_{i_3 r}. \tag{5}$$

Tucker decomposition. In Tucker decomposition [13], the tensor $\mathcal{Y}$ can be factorized as

$$\mathcal{Y} \approx \mathcal{G} \times_1 A \times_2 B \times_3 C = [\![\mathcal{G}; A, B, C]\!], \tag{6}$$

where $\mathcal{G}$ is the core tensor; $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$; A, B, and C are the $A \in \mathbb{R}^{N_1 \times R_1}$, $B \in \mathbb{R}^{N_2 \times R_2}$, and $C \in \mathbb{R}^{N_3 \times R_3}$, respectively. R_1, R_2, and R_3 are the ranks of matrices A, B, and C with limitation $\text{rank}(\mathcal{Y}) \le R_1 \le N_1$, $\text{rank}(\mathcal{Y}) \le R_2 \le N_2$, and $\text{rank}(\mathcal{Y}) \le R_3 \le N_3$, respectively; and $\times_n$ is the n-mode product. For example, the 1-mode product $\mathcal{Y} = \mathcal{G} \times_1 A$ is defined as

$$y_{i_1, r_2, r_3} = \sum_{r_1=1}^{R_1} g_{r_1 r_2 r_3} a_{i_1 r_1}, \quad i_1 = 1, 2, \ldots, N_1. \tag{7}$$

In this work, four TCMs are compared: CP-weighted optimization (WOPT), Bayesian CP factorization (BCPF), simultaneous tensor decomposition, and completion (STDC), and high accuracy low-rank tensor completion (HaLR) [21, 23–25]. These four TCMs are based on rank optimization. WOPT and BCPF are two TCMs based on CP decomposition while STDC is based on Tucker decomposition. The interpolation method is also mentioned as a baseline. The main concepts of the four TCMs and the interpolation method are now briefly introduced.

2.3.1 *WOPT*

This method solves the weighted minimum problem by using a first-order optimization approach [21]. In order to minimize the approximation error in eqs. (4) and (5), the objective function of weighted CP decomposition is the following error function:

$$\begin{aligned} f_{\mathcal{O}}(A, B, C) &= \frac{1}{2} \left\| (\mathcal{Y} - [\![A, B, C]\!]) \right\|_{\mathcal{O}}^2 \\ &= \frac{1}{2} \left\| (\mathcal{Y} * \mathcal{O} - [\![A, B, C]\!] * \mathcal{O}) \right\|^2 \\ &= \frac{1}{2} \left\| (\mathcal{Y} - \mathcal{Z}) \right\|^2, \end{aligned} \tag{8}$$

where $\mathcal{Y}$ can be inferred from eq. (3): $\mathcal{Y} = \mathcal{X} * \mathcal{O} = \mathcal{X} * \mathcal{O} * \mathcal{O} = \mathcal{Y} * \mathcal{O}$. We define $\mathcal{Z}$ as $\mathcal{Z} = [\![A, B, C]\!] * \mathcal{O}$. By minimizing the objective function, WOPT finds the completed tensor $\hat{\mathcal{X}} = [\![A, B, C]\!] * (1 - \mathcal{O}) + \mathcal{Y} * \mathcal{O}$. Tensor gradient deviation with weight is used to minimize the objective function. For example, the gradient deviation of A is

$$\frac{\partial f_{\mathcal{O}}}{\partial A} = (\mathcal{Z}_{(1)} - \mathcal{Y}_{(1)}) A^{-1}, \tag{9}$$

where $A^{-1} = B \odot C$, $\odot$ denotes the Khatri-Rao product and $\mathcal{Z}_{(1)}$ is the 1-mode unfolding of $\mathcal{Z}$.

2.3.2 *BCPF*

BCPF is also an extension of the CP method [23]. In CP decomposition, the tensor rank R is a hyper-parameter that needs to be specified manually. The Bayesian method is applied to CP to resolve the rank determination problem. This method can deal with incomplete and noisy data with a fully Bayesian inference framework and a hierarchical sparsity inducing prior.

2.3.3 *STDC*

STDC is a TCM based on Tucker decomposition [24, 25]. The optimized tensor $\hat{\mathcal{X}}$ can be decomposed into four components through eq. (6):

$$\hat{\mathcal{X}} = [\![\mathcal{G}; A, B, C]\!] \quad \text{s.t. } \mathcal{Y} = \hat{\mathcal{X}} * \mathcal{O}. \tag{10}$$

In order to find $\hat{\mathcal{X}}$, there are two middle steps present in STDC. The two steps alternately optimize matrices A, B, C, and the core tensor $\mathcal{G}$. The maximum a posteriori (MAP) strategy is used to find $\hat{\mathcal{X}}$, $\mathcal{G}$, A, B, and C. The strategy minimizes the objective function:

$$\begin{aligned} f_{\mathcal{O}}(\mathcal{G}, A, B, C) =& \alpha_1 \|A\|_* + \alpha_2 \|B\|_* + \alpha_3 \|C\|_* \\ &+ \beta \mathrm{tr}((A \otimes B \otimes C) L (A \otimes B \otimes C)^{\mathrm{T}}) \\ &+ \gamma \|\mathcal{G}\|_{\mathrm{F}}^2, \end{aligned} \tag{11}$$

where $\|A\|_*$ denotes the nuclear norm and is the sum of all of the singular values; $\otimes$ denotes the Kronecker product; and $L \in \mathbb{R}^{N_1N_2N_3 \times N_1N_2N_3}$ is the so-called Laplacian matrix. The augmented Lagrange multiplier method is used to update the parameters for several iterations. Finally, the augmented Lagrange multiplier method converges and the sub-manifold A, B, C, and the core tensor $\mathcal{G}$ are obtained. The optimized tensor $\hat{\mathcal{X}}$ can be constructed using eq. (10).

2.3.4 *HaLR*

HaLR is developed to estimate the missing values in the low-rank tensor completion problems [25]. For low-rank matrix completions, it can optimize the tensor with missing entries by minimizing the rank. Because the rank of a tensor is non-convex, and computing the rank of tensors with more than two modes constitutes an NP-hard problem, the nuclear norm can be used to approximate the rank of a matrix, which is the tightest convex envelope of the rank. HaLR aims at minimizing the objective function:

$$\begin{aligned} &f_{\mathcal{O}}(\hat{\mathcal{X}}, \mathcal{A}, \mathcal{B}, C) = \alpha_1 \left\|\mathcal{A}_{(1)}\right\|_* + \alpha_2 \left\|\mathcal{B}_{(2)}\right\|_* + \alpha_3 \left\|C_{(3)}\right\|_*, \\ &\text{s.t. } \hat{\mathcal{X}} * \mathcal{O} = \mathcal{X},\ \hat{\mathcal{X}} = \mathcal{A},\ \hat{\mathcal{X}} = \mathcal{B},\ \hat{\mathcal{X}} = C,\ \alpha_1 + \alpha_2 + \alpha_3 = 1. \end{aligned} \tag{12}$$

This function has multiple dependent terms so that it cannot converge efficiently. HaLR overcomes this issue with the alternating direction method of the multipliers algorithm, which solves optimization problems with non-smooth terms.

2.3.5 *Interpolation*

Interpolation is the basic completion method used for completing missing EEG signals. In this paper, this method is used as a baseline. For simplicity, we adopt a simple one-dimension linear interpolation method in the one-dimensional time series. The linear interpolation completes the missing entry with a line connecting two adjacent known points. If the missing entry starts from the beginning of the time series or ends at the ending of the time series (which only have one known adjacent point), the missing entries are completed with the value of the adjacent point.

2.4 Performance measurement

2.4.1 *Negative logarithm normalized root mean square error*

The normalized root mean square error (NRMSE) can be regarded as the reciprocal of the simplified signal to noise ratio (SNR). It is used as the criterion to assess the difference between the completed EEG tensor $\hat{\mathcal{X}}$ and the original EEG tensor $\mathcal{X}$.

The general NRMSE can be obtained by dividing the root mean square error by the difference between the maximum and the minimum value of the signals. The maximum and minimum of these time series in EEG tensors may differ significantly. Therefore, the NRMSE is computed as follows: firstly, the NRMSE of each channel is computed and then the mean of the NRMSE of all channels is calculated. These steps can be expressed as follows:

$$\mathrm{NRMSE}_{jk} = \frac{\sqrt{\sum_{i=1}^{N_{\mathrm{time}}} (\hat{\mathcal{X}}_{ijk} - \mathcal{X}_{ijk})^2 / (N_{\mathrm{mTL}} \times \mathrm{Fs})}}{\max(\mathcal{X}_{:jk}) - \min(\mathcal{X}_{:jk})}, \tag{13}$$

$$\mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle} = \frac{\sum_{j=1}^{N_{\mathrm{channel}}} \sum_{k=1}^{N_{\mathrm{trial}}} \mathrm{NRMSE}_{jk}}{N_{\mathrm{mCT}}}. \tag{14}$$

The $-\log_{10}$ function is applied on $\mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle}$, thereby obtaining the negative logarithm normalized root mean square error (LNRMSE):

$$\mathrm{LNRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle} = -\log_{10}(\mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle}). \tag{15}$$

In simulation I, the differences can be easily measured using eqs. (13) and (14). The factors N_{mTL} and N_{mCT} in these equations can avoid the influences induced by the change of N_{mCT} and N_{mTL}, even though these vary as the number of missing entries changes. In simulation II, we first need to calculate $\mathrm{LNRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle}$, $\mathrm{LNRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}}_{\mathrm{channel}} \rangle}$, and

$\mathrm{LNRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}}_{\mathrm{trial}} \rangle}$. The differences between two permutations can be measured using

$$\delta\mathrm{LNRMSE}_{\mathrm{channel}} = -\log_{10}\left(\mathrm{abs}\left(\mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}}_{\mathrm{channel}} \rangle} - \mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle}\right)\right), \tag{16}$$

or

$$\delta\mathrm{LNRMSE}_{\mathrm{trial}} = -\log_{10}\left(\mathrm{abs}\left(\mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}}_{\mathrm{trial}} \rangle} - \mathrm{NRMSE}_{\langle \mathcal{X}, \hat{\mathcal{X}} \rangle}\right)\right). \tag{17}$$

In simulation III, NRMSE is based on the differences in the missing entries, so the influences induced by the change of N_{channel} can also be avoided. Therefore, LNRMSE can be used as the criterion to measure the differences between the original EEG tensor $\mathcal{X}$ and the completed EEG tensor $\hat{\mathcal{X}}$.

Since there are 15 subjects in the dataset used, the average LNRMSE is calculated across subjects when the comparison results are measured. The calculation of the average LNRMSE can avoid the uniqueness for a single subject.

2.4.2 *Classification on movement-related cortical potential*

The EEG data is usually used for classification tasks thus the classification performance of different tensor completion methods is also investigated in ref. [15]. From the original dataset, three EEG tensors are extracted for each subject (Sect. 2.1). The three tensors refer to three actions: elbow flexion, elbow extension, and resting. In tensors related to movement states (elbow flexion or elbow extension), the movement onsets are exactly located at the 2 s. An application based on the movement-related cortical potential (MRCP) can be implemented on these EEG tensors. This application aims to decode and classify premovement patterns from low frequency bands of EEG signals. The method of task-related component analysis (TRCA) + canonical correlation patterns (CCPs) can be applied to classify the movement states (elbow flexion or elbow extension) and rest state (resting). More specifically, the classification has four steps:

(1) Filter EEG signals with a bandpass filter between 0.05 and 4 Hz;

(2) Cast the bandpassed signals into the spatial filter, TRCA;

(3) Extract CCPs from the EEG signals filtered with TRCA;

(4) Classify the CCPs with a linear discriminated analysis (LDA) classifier.

TRCA designs a spatial filter that extracts task-related components by maximizing the reproducibility during the task [26]. To simplify the description of TRCA, we denote the multichannel EEG tensors $\mathcal{X}$ as $X^k(t) \in \mathbb{R}^{N_{\mathrm{channel}} \times N_{\mathrm{trial}}}$, $t = 1, 2, \ldots, N_{\mathrm{time}}$. The symbol k represents the trials belonging to the movement states ($k = 1$) or the rest state ($k = 2$). $X^k(t)$ consists of two components: (1) task-related signal $s(t) \in \mathbb{R}$ and (2) task-unrelated noise $n(t) \in \mathbb{R}$. The relationship between $X(t)$, $s(t)$, and $n(t)$ can be described with a linear model:

$$X^k_{ij}(t) = a^k_{1,i,j} s(t) + a^k_{2,i,j} n(t), \quad i = 1, \ldots, N_{\mathrm{channel}},\ j = 1, \ldots, N_{\mathrm{trial}}. \tag{18}$$

If we define the linear sum of EEG signals $X(t)$ as $y(t) \in \mathbb{R}^{N_{\mathrm{trial}}}$, then

$$y^k_j(t) = \sum_{i=1}^{N_{\mathrm{channel}}} w^k_i X^k_{ij}(t) = \sum_{i=1}^{N_{\mathrm{channel}}} \left(w^k_i a^k_{1,i,j} s(t) + w^k_i a^k_{2,i,j} n(t)\right), \quad j = 1, \ldots, N_{\mathrm{trial}}. \tag{19}$$

TRCA aims to recover the task-related signal $s(t)$ from $y(t)$. The ideal solution to the equation is $\sum_{i=1}^{N_{\mathrm{channel}}} w^k_i a^k_{1,i,j} = 1$ and $\sum_{i=1}^{N_{\mathrm{channel}}} w^k_i a^k_{2,i,j} = 0$. The solution can be approached by maximizing the intertrial covariance. The covariance $C^k_{j_1,j_2}$ between the j_1th trial and the j_2th trial can be computed by

$$C^k_{j_1,j_2} = \mathrm{Cov}\left(y^k_{j_1}(t), y^k_{j_2}(t)\right) = \sum_{i_1,i_2}^{N_{\mathrm{channel}}} w^k_{i_1} w^k_{i_2} \mathrm{Cov}\left(X^k_{i_1 j_1}(t), X^k_{i_2 j_2}(t)\right). \tag{20}$$

For all combinations of the trials, the sums of covariances are calculated:

$$\begin{aligned} \sum_{\substack{j_1,j_2=1 \\ j_1 \neq j_2}}^{N_{\mathrm{trial}}} C^k_{j_1,j_2} &= \sum_{\substack{j_1,j_2=1 \\ j_1 \neq j_2}}^{N_{\mathrm{trial}}} \mathrm{Cov}\left(y^k_{j_1}(t), y^k_{j_2}(t)\right) \\ &= \sum_{\substack{j_1,j_2=1 \\ j_1 \neq j_2}}^{N_{\mathrm{trial}}} \sum_{i_1,i_2=1}^{N_{\mathrm{channel}}} w^k_{i_1} w^k_{i_2} \mathrm{Cov}\left(X^k_{i_1 j_1}(t), X^k_{i_2 j_2}(t)\right) \\ &= w^{\mathrm{T}} S^k w. \end{aligned} \tag{21}$$

To obtain a finite solution, the variance of $y^k_j(t)$ is constrained to 1:

$$\begin{aligned} \sum_{j_1,j_2=1}^{N_{\mathrm{trial}}} C^k_{j_1,j_2} &= \sum_{j_1,j_2=1}^{N_{\mathrm{trial}}} \mathrm{Cov}\left(y^k_{j_1}(t), y^k_{j_2}(t)\right) \\ &= \sum_{j_1,j_2=1}^{N_{\mathrm{trial}}} \sum_{i_1,i_2=1}^{N_{\mathrm{channel}}} w^k_{i_1} w^k_{i_2} \mathrm{Cov}\left(X^k_{i_1 j_1}(t), X^k_{i_2 j_2}(t)\right) \\ &= w^{\mathrm{T}} Q^k w = 1. \end{aligned} \tag{22}$$

The constrained spatial filter can be obtained by calculating the solution of maximizing $J = (w^{\mathrm{T}} S^k w)(w^{\mathrm{T}} Q^k w)^{-1}$. After solving the generalized eigenvalue problem, the eigenvectors are sorted and arranged in the descending order of eigenvalues. Half of the eigenvectors whose corresponding eigenvalues are the maximum ones are selected as the eigenvectors

used in the spatial filter. The eigenvectors from two classes are then concatenated into the TRCA spatial filter W.

After being filtered with TRCA, three CCPs were calculated for each class from the filtered EEG signals. Given the filtered training data $\mathcal{X}^k \in \mathbb{R}^{N_{\text{time}} \times N_{\text{channel}} \times N_{\text{trial}}}, k = 1, 2$, we can obtain the templates $X^k = \sum_{j=1}^{N_{\text{trial}}} \mathcal{X}^k_{::j} / N_{\text{trial}} \in \mathbb{R}^{N_{\text{time}} \times N_{\text{channel}}}$ for each class. The EEG signal of the trial from which we aim to extract features is $X \in \mathbb{R}^{N_{\text{time}} \times N_{\text{channel}}}$.

(1) Correlation coefficients between filtered signals:

$$X_k = X^k,\ X_* = X, \tag{23}$$

$$\rho_1^k = \text{corr}(X_* W, X_k W),\ k = 1, 2; \tag{24}$$

(2) Correlation coefficients between filtered signals with a canonical correlation analysis projection:

$$X_k = \hat{X}^k,\ X_* = X,\ [A_k, B_k] = \text{cca}(X_* W, X_k W), \tag{25}$$

$$\rho_2^k = \text{corr}(X_* W B_k, X_k W B_k),\ k = 1, 2; \tag{26}$$

(3) Correlation coefficients between the distances of filtered signals:

$$X_k = \hat{X}^k - \hat{X}^{3-k},\ X_* = X - \hat{X}^{3-k}, \tag{27}$$

$$[A_k, B_k] = \text{cca}(X_* W, X_k W), \tag{28}$$

$$\rho_3^k = \text{corr}(X_* W B_k, X_k W B_k),\ k = 1, 2. \tag{29}$$

For each trial, six features can be obtained. In the above three equations, cca denotes the function to calculate the projection of the canonical correlation analysis and returns the canonical coefficients of the two input data matrices. The function corr returns the 2-dimensional correlation coefficient between the two input data matrices.

The TRCA+CCP method is applied to evaluate the influence of the number of missing entries in simulation I. In the two-class classification task, the classification is implemented between elbow flexion and resting or elbow extension and resting. In the classification between movement states (elbow flexion or elbow extension) and rest state (resting), the EEG signals in the resting state are the original ones (no completion procedure) while there are three types of EEG signals in the movement states: completed with simple interpolation, completed with TCMs and original. A 10-fold cross validation is applied.

3 Results

In the following section, we compare the LNRMSE or δLNRMSE of four TCMs (STDC, HaLR, WOPT, and BCPF) with those of the interpolation method (Interp) in three simulations. The LNRMSE and δLNRMSE results are averaged across three actions (elbow flexion, elbow extension, and rest) because of their similar conclusions. In the classification task in simulation I, we classify two states using the TRCA+CCP method: elbow flexion versus rest and elbow extension versus rest. The classification results are also averaged across the two cases.

Before the simulations, the singular values of the unfolded matrices of the EEG tensors are tested. The singular value decomposition (SVD) is applied to the matrices, which are unfolded from the EEG tensors along three axes [27]. Figure 4 presents the singular value curves along sample time, channel, and trial, individually. These singular value curves are averaged across the original tensors of the 15 subjects. The curves decrease sharply, which means that these EEG tensors can be compressed into a low-rank tensor.

3.1 Quality of reconstructions versus the number of missing entries

In simulation I, the aim is to determine how the number of missing entries influences the completion effects of TCMs. We have two cases for the change of the number of missing entries: (1) changing N_{mCT}; and (2) changing N_{mTL}.

(1) Changing N_{mCT}. Figures 5 and 6 show the comparison between TCMs as N_{mCT} changes for LNRMSE and classification accuracy, individually. In Figure 5, it can be seen that the LNRMSE remains stable as N_{mCT} increases for all four TCMs. When N_{mCT} is very small, the LNRMSE is slightly higher than these constants for the four TCMs. STDC has the

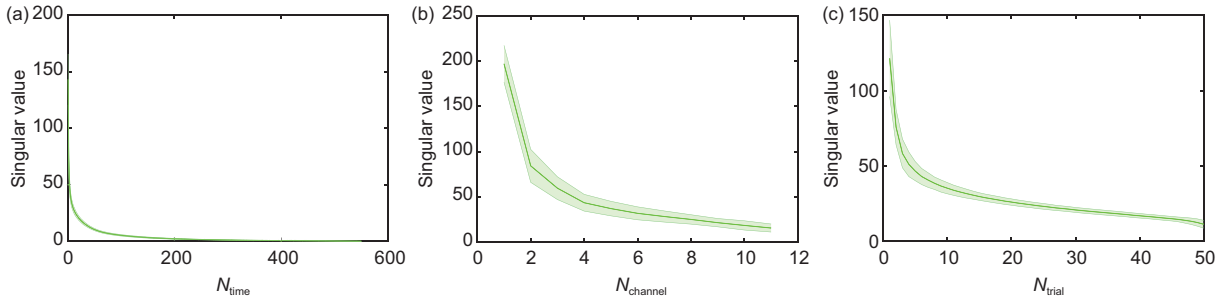

Figure 4 (Color online) SVD along three axes in EEG tensors. (a) SVD along sample; (b) SVD along channel; (c) SVD along trial.

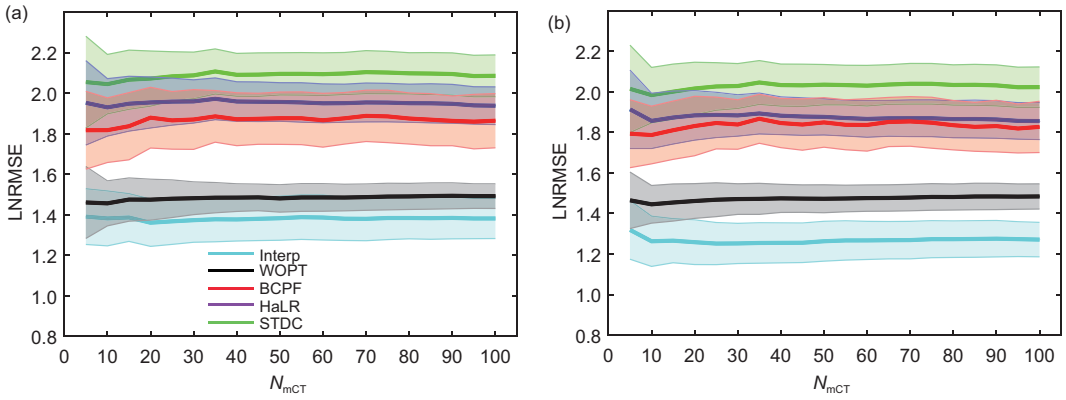

Figure 5 Simulation I: Evaluation of the LNRMSE as N_{mCT} is increased. N_{mTL} is set to 0.25 s (a) and 0.5 s (b) separately. The change of N_{mCT} leads to very small changes on the completion effects of TCMs. STDC shows better LNRMSE performance compared to the other compared TCMs.

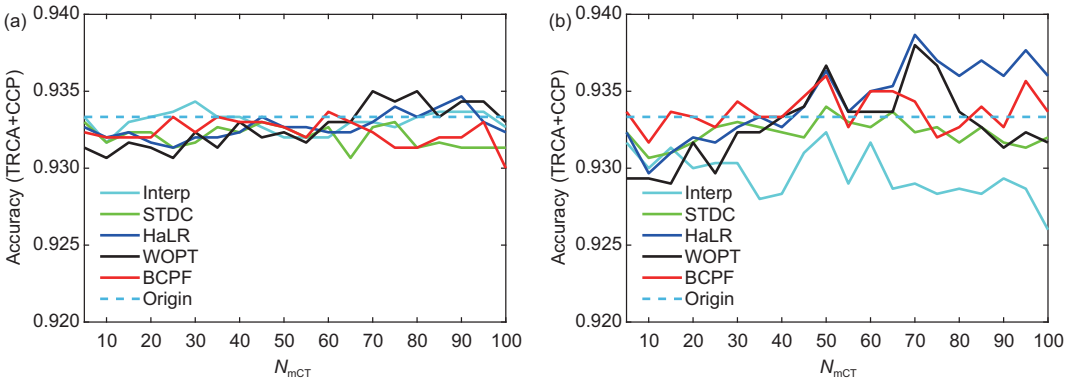

Figure 6 Simulation I: Evaluation of the classification against N_{mCT} on MRCP encoding with TRCA+CCP. (a) N_{mTL} = 0.25; (b) N_{mTL} = 0.5.

best LNRMSE performance among the compared completion methods. In Figure 6, we can analyze how the number of missing entries and completion influences the classification result between the movement and rest states. The accuracy shows no major differences between the completed data and original data.

(2) Changing N_{mTL}. Figures 7 and 8 show the comparison between TCMs as N_{mTL} changes for LNRMSE and classification accuracy, individually. In Figure 7, the LNRMSE decreases gradually as N_{mTL} increases. However, in real EEG acquisition, the value of N_{mTL} will be minuscule. Therefore, missing information in EEG signals can be easily recovered.

In the analysis of the influence by the number of missing entries, we evaluate the influence from both the LNRMSE and the classification with TRCA+CCP. From the perspective of the LNRMSE, increasing the number of missing entries will lead to unrecoverable information loss, even with TCMs involved. Nevertheless, compared to the simple interpolation method and the other TCMs, STDC has the highest LNRMSE, which means that data completed using STDC has the smallest error among the compared completion methods. From the perspective of classification accuracy with TRCA+CCP, accuracy has no obvious decrease until the time length of missing entries is large enough. In this case, data completed using TCMs have better performance than that completed using basic interpolation techniques.

3.2 Analysis of robustness to initial conditions

In simulation II, the effect of changing the initial conditions of the completion methods is investigated. A way to do that is to randomly permute channels or trials. Since this permutation will not affect the rank and content of the tensor, the result should remain stable (theoretically, it should remain the same). Any changes will then be due to the re-running of the method with different initial conditions, leading to a different convergence point. In Figure 9, the differences of four TCMs are calculated using eqs. (16) and (17). The results displayed in Figure 9 show that the random permutation in EEG signals has less influence on the completion effects of STDC and HaLR. In comparison to WOPT, the other three TCMs have more stable variances. Therefore, when completing missing entries in EEG signals with WOPT or BCPF, the completed signals suffer the risk of being more dependent on the initial conditions.

The methods that were used assume the low-rank of the original tensor prior to completing the missing entries, and

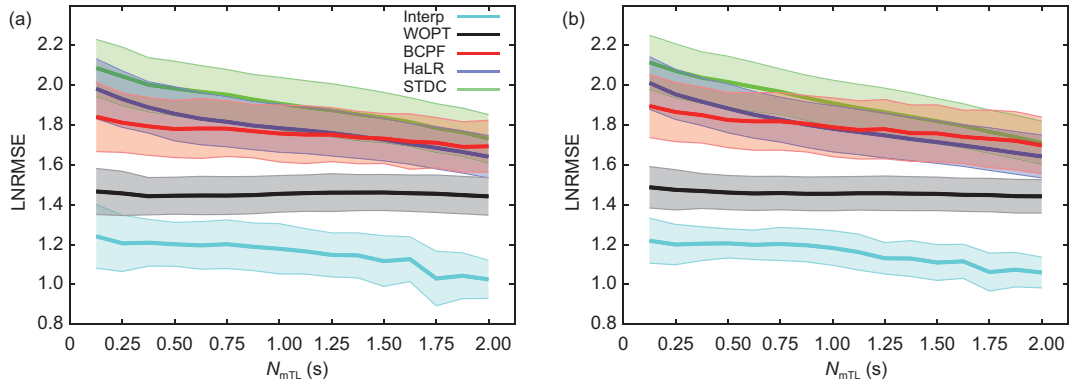

Figure 7 Simulation I: Evaluation of the LNRMSE as N_{mTL} is increased. N_{mCT} is set to 10 (a) and 20 (b) separately. The decrease of LNRMSE indicates that the completion effects become worse as N_{mTL} increases. STDC achieves the best LNRMSE performance among the compared TCMs.

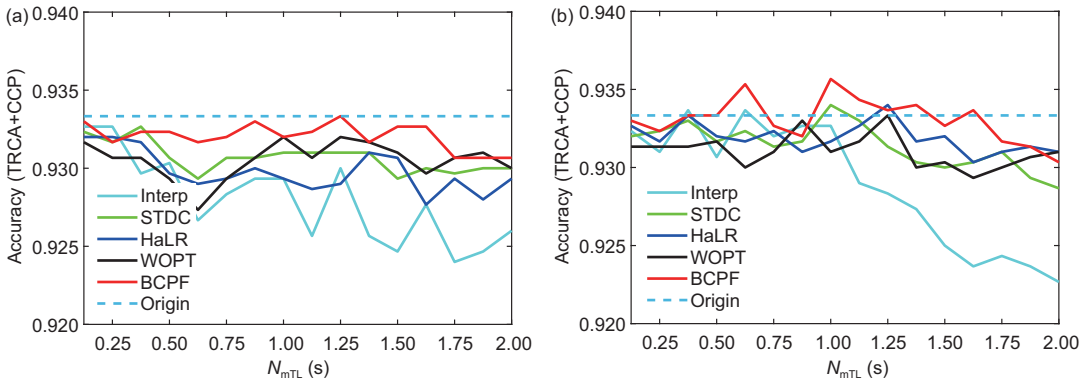

Figure 8 Simulation I: Evaluation of the classification against N_{mTL} on MRCP encoding with TRCA+CCP. (a) $N_{\mathrm{mCT}} = 10$; (b) $N_{\mathrm{mCT}} = 20$.

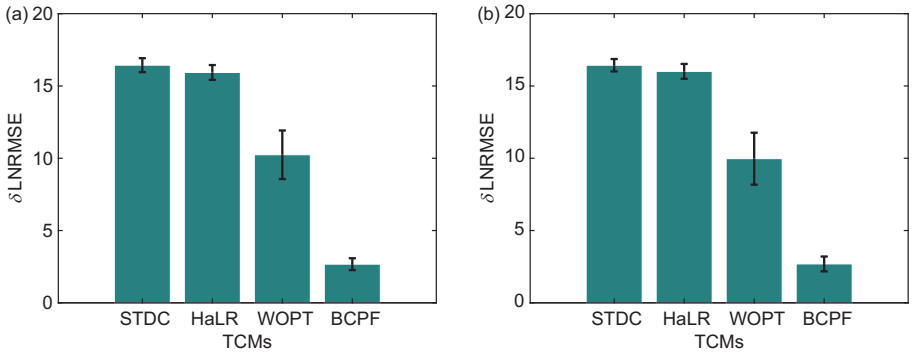

Figure 9 Simulation II: Influences of channel/trial permutations. The shuffling of channel ordering (a) and trial ordering (b) are implemented in this simulation. It can be seen that the permutations have less influence on the completion effect of STDC.

we have checked the low-rank of the EEG tensors in Figure 4. Any permutation in a particular mode is supposed not to change the results because the rank of the tensor is invariable under any permutation. In theory, all the methods used here are supposed to give the same results independently of the permutation order in any of the dimensions of the tensor. The differences between two permutations may be induced by two numerical convergence problems for WOPT and BCPF.

(1) In WOPT, the initial rank cannot be automatically located. When applying the WOPT method, the errors between the completed tensor and original tensor cannot be reduced to smaller values and the errors change within a small range before the end of the iteration loops.

(2) In BCPF, the rank of the tensor can be determined automatically by the Bayesian reference. The rank of the EEG tensor changes as the number of missing entries changes. Furthermore, the Bayesian reference enables BCPF to locate the rank of the EEG tensors. In the completion experiment, the maximum rank is limited by the computer memory. The most suitable rank is unlikely to be reached when the rank of the the EEG tensor is out of the memory limit. The other possible reason is the number of iteration loops. Due to the

time-consuming computation carried out in BCPF, loops are stopped early when the errors are stable and hardly change.

To check the completion effects of different permutations, the time series with missing entries are depicted in Figure 10. The terms contained in the legend 'original', 'completed', 'completed (channel shuffled)', 'completed (trial shuffled)' refer to $\mathcal{X}$, $\hat{\mathcal{X}}$, $\hat{\mathcal{X}}_{\text{channel}}$, $\hat{\mathcal{X}}_{\text{trial}}$ in eqs. (16) and (17). To compare the differences between these permutations in the missing entries more easily, the amplitudes of 'completed', 'completed (channel shuffled)', 'completed (trial shuffled)' are added to 0.5, 1, 1.5, individually.

3.3 Analysis of quality of reconstructions versus the total number of channels

In simulation III, the number of N_{channel} of the EEG tensors is changed. With it, we aim to simulate the different numbers of electrodes for EEG experiments. Figure 11(a) presents the results of simulation III. There are two points that need to be explained in this figure: (1) STDC still achieves the best performance among the other compared TCMs, but note that the LNRMSE becomes considerably reduced when N_{channel} is less than 3, so it is better to adopt at least three electrodes in EEG acquisition; (2) BCPF experiences an unexpected decrease when $N_{\text{channel}} = 7$. In Figure 11(b), the original EEG signals and the completed EEG signals of Subject 1 in the missing entries are given. Compared to the original EEG signals and BCPF-completed signals with $N_{\text{channel}} = 4$, BCPF-completed signals with $N_{\text{channel}} = 7$ have a higher amplitude. This phenomenon may be induced by insufficient convergence in the implementation of BCPF. Considering the time consumption and computer memory limitation, loops are stopped early when the errors are stable and hardly change.

In Figure 12, the spectrum differences on the missing entries are analyzed for the EEG tensor. In Figure 12(a), the time series with missing entries are depicted, with the missing entries being completed. The section of the missing data is comprised within the two red lines in the figure. Figure 12(b) shows the spectrograms of these time series. In Figure 12(c), the part of spectrogram which shows the differences can be more clearly discerned, which allows for easier and better comparisons.

In Supporting Information, the differences in the spectrum are compared to a numeric measurement, the average spectrum error (ASE). The ASE has similar conclusions as the LNRMSE for simulations I and II. The ASE results also

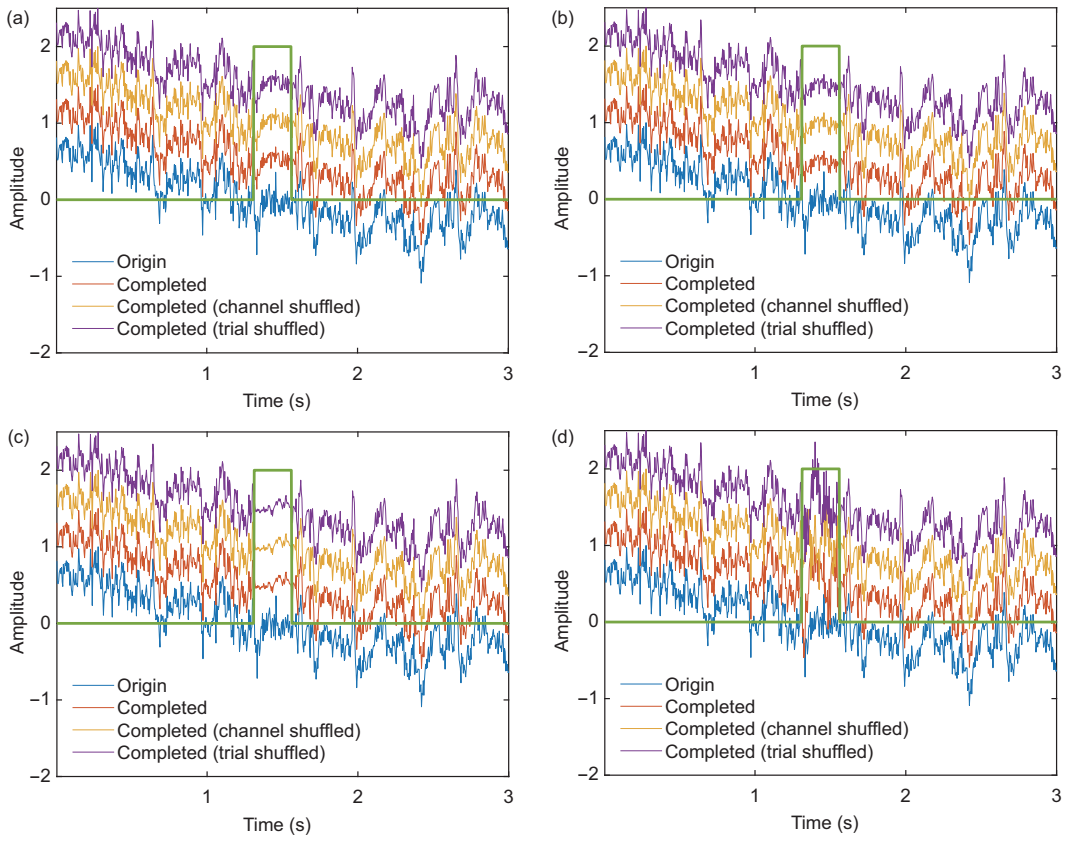

Figure 10 The time series with missing entries of different permutations. To facilitate the comparison of amplitudes in the missing entries, the amplitudes of 'completed', 'completed (channel shuffled)', and 'completed (channel trial shuffled)' are added to 0.5, 1, 1.5, individually. (a) STDC; (b) HaLR; (c) WOPT; (d) BCPF.

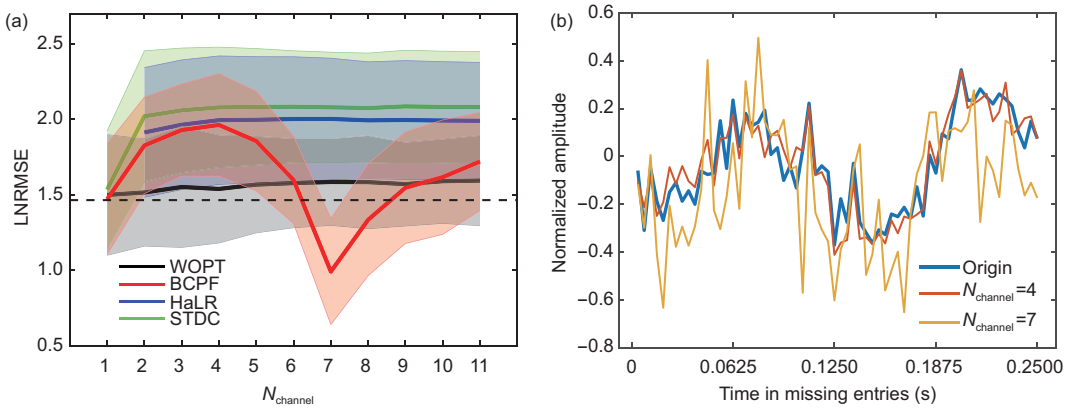

Figure 11 Simulation III. (a) Completion effects comparison as N_{channel} changes. Because the missing entries remain unchanged in simulation III, LNRMSE is a constant for the interpolation method, as represented with the black dotted line. Thus, here it serves as a baseline. HaLR fails to complete the missing entries and the LNRMSE remains empty with $N_{\text{channel}} = 1$ for this method. (b) Signals comparison in BCPF-completed missing entries.

demonstrate that STDC has a better performance compared to the other completion methods.

From the analysis of the results above, it can be seen that STDC achieves a superior completion effect in comparison to the other completion methods in all three simulations.

4 Discussion

In this work, we have proven through three experiments that STDC is the most suitable TCM for completing missing EEG signals among the compared completion methods. The three experiments are based on the following assumptions:

(1) The LNRMSE measurement is a suitable way to evaluate completion effects;

(2) The used EEG dataset is representative of most EEG signals;

(3) In simulation III, it is acceptable that missing entries are set to the first channel.

The first assumption is that the LNRMSE measurement can assess the completion effects of TCMs in three simulations. The independent variables in these three are the number of missing entries, the permutation of channels or trials, and the number of channels in EEG tensors, respectively. The measurement should avoid influences from the potential changes of the independent variables. In eqs. (13) and (14), the factor N_{mTL} and N_{mCT} can balance the number of missing entries. In other words, the LNRMSE used here measures the averaged completion effects on each missing entry. Therefore, the LNRMSE still works here even when the size of the EEG tensor changes.

The second assumption concerns whether the used EEG signals are representative. The active area of the brain changes for different brain activities, e.g., brain activities in the visual cortex and the motor cortex. The main difference is the location of the active areas. Regardless of the location of these areas, however, the different brain activities are mainly characterized by the change of power, e.g., ERD phenomena. In fact, we do not know the exact channel location in EEG completion. As a result, the difference can be ignored in EEG completion. The used EEG signals consist of signals in both the rest and the movement states. For the EEG signals in the movement state, the movement onset is located using an exoskeleton sensor. The position of the movement onset is fixed in the EEG tensors. In most offline EEG analysis, it is common for EEG signals during certain brain activities to be extracted and then analyzed. For EEG signals in the rest state, the completion is also successful even without localization. Therefore, the used EEG signals can be considered to be representative of most EEG signals.

The third assumption focuses on the design of simulation III. In this simulation, we set the only channel with missing entries in the first channel of the EEG tensor. Although this hardly occurs in real EEG completion, this operation is reasonable because we have explored the permutation of channels. In simulation II, we calculated the differences between two permutations of channel sequence $\delta\text{LNRMSE}_{\text{channel}}$ with eq. (16). The results in simulation II showed that changing the permutations of channels has only very slight influences on the completion effects for STDC, HaLR, and even WOPT. Therefore, the third assumption is also reasonable.

5 Conclusions

Missing EEG signals deteriorate the performance of EEG signal analysis. In this work, we propose to use STDC to complete missing EEG signals. We analyze the completion effects with LNRSME with four kinds of TCMs (STDC, BCPF, WOPT, and HaLR) and with the interpolation method. Three simulations are designed to demonstrate the outstanding features of STDC.

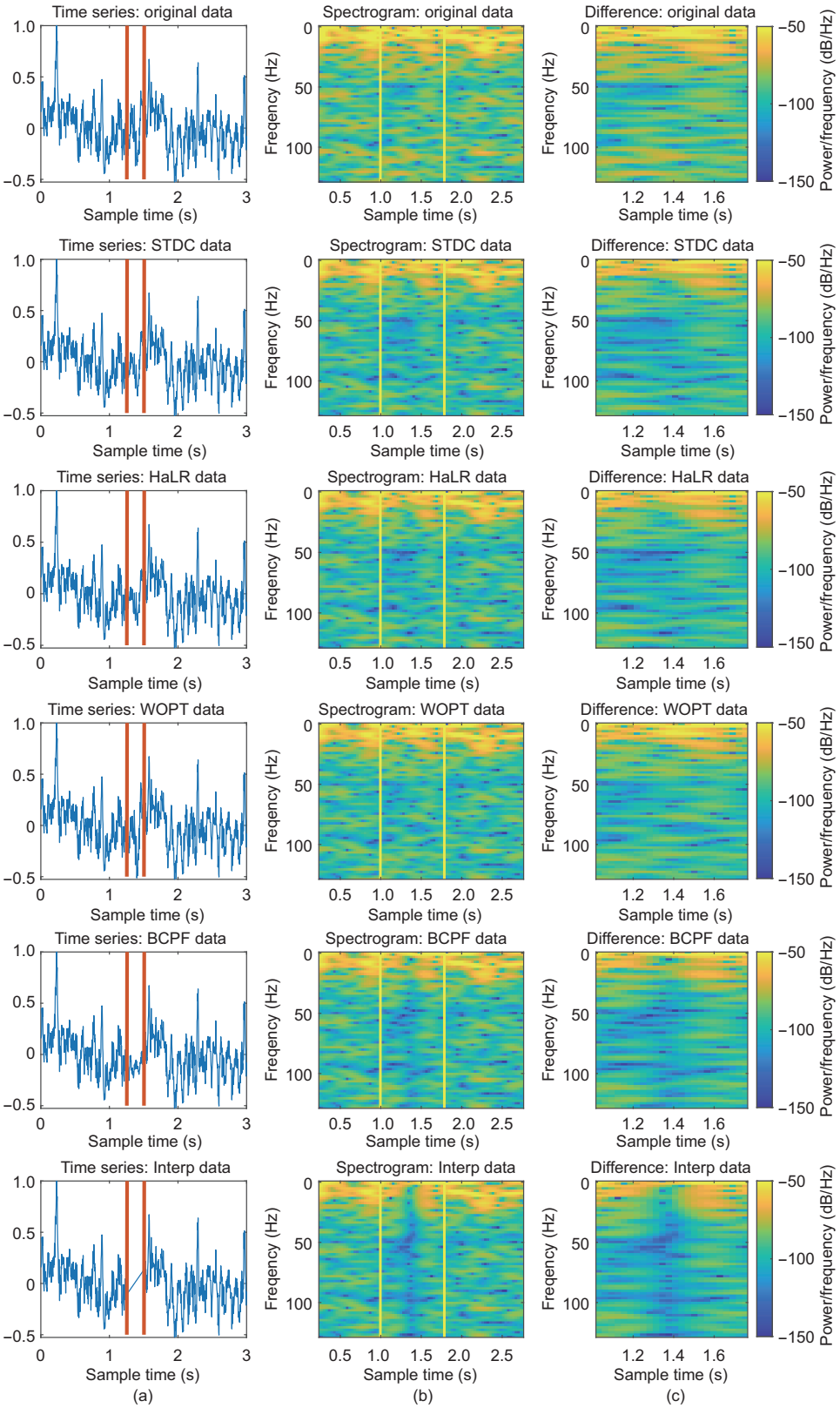

Figure 12 Spectrogram comparison on the missing entries in simulation III with the EEG tensor whose N_{channel} is equal to 11. (a) Completion effects in the time domain; (b) completion effects in frequency domain; (c) difference of completion effects in the frequency domain.

(1) Simulation I emulates how the completion effects change as the amount of the missing entries changes. The average LNRMSE of STDC is approximately 4.75, which is greater than the other completion methods. This means

that the completion effects of STDC are better than those of the other completion methods. Classification accuracy with TRCA+CCP is also applied on simulation I to analyze the influence from the number of missing entries.

(2) Simulation II evaluates the influences of changes in the order of the EEG channels (shuffling), which are supposed not to affect anything, but in a practical scenario, we demonstrate that this may change due to numerical convergence and the initial conditions of the method. The average δLNRMSE of STDC is approximately 15. This means that STDC experiences minimal influences after changing the permutation of the EEG channels. Due to the possibly insufficient iteration loops induced by memory limitations and time consumption, BCPF fails to meet the expectations.

(3) Simulation III studies the completion effects as the number of EEG channels changes. The result in this simulation shows that the EEG signals must consist of at least three channels; otherwise, the TCMs will have an inferior completion effect. In addition, note that here STDC also has the best LNRMSE performance among the compared TCMs.

When comparing the different completion methods, it can be seen that overall STDC best fits the characteristics of EEG signals. In future work, we will explore the influences with a higher number of missing entries and will develop real-time applications to complete missing EEG signals.

This work was supported by the National Key R&D Program of China (Grant No. 2017YFE0129700), the National Natural Science Foundation of China (Key Program) (Grant No. 11932013), the National Natural Science Foundation of China (Grant No. 61673224), the Tianjin Natural Science Foundation for Distinguished Young Scholars (Grant No. 18JCJQJC46100), the Tianjin Science and Technology Plan Project (Grant No. 18ZXJMTG00260), and in part by the Ministry of Education and Science of the Russian Federation (Grant No. 14.756.31.0001). SOLÉ-CASALS Jordi work was supported by COST (European Cooperation in Science and Technology) Action (Grant No. CA18106). CAIAFA Cesar F. Work was supported by Proyectos de Investigación Científica y Tecnológica (PICT) (Grant No. 2017-3208) and Proyectos Universidad de Buenos Aires Ciencia y Técnica (UBACyT) (Grant No. 20020170100192BA).

Supporting Information

The supporting information is available online at tech.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

1 Lotte F, Congedo M, Lécuyer A, et al. A review of classification algorithms for EEG-based brain computer interfaces. J Neur Eng, 2007, 4: 1–13
2 Sakhavi S. Application of deep learning methods in brain-computer interface systems. Dissertation for Doctoral Degree. Singapore: National University of Singapore, 2017
3 Lin Z, Zhang C, Wu W, et al. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. IEEE Trans Biomed Eng, 2006, 53: 2610–2614
4 Donchin E, Spencer K M, Wijesinghe R. The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. IEEE Trans Rehab Eng, 2000, 8: 174–179
5 Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication. Proc IEEE, 2001, 89: 1123–1134
6 Jin J, Li S, Daly I, et al. The study of generic model set for reducing calibration time in P300-based brain-computer interface. IEEE Trans Neur Syst Rehabil Eng, 2020, 28: 3–12
7 Solé-Casals J, Caiafa C F, Zhao Q, et al. Brain-computer interface with corrupted EEG data: A tensor completion approach. Cogn Comput, 2018, 10: 1062–1074
8 Neuper C, Wrtz M, Pfurtscheller G. ERD/ERS patterns reflecting sensorimotor activation and deactivation. Progress Brain Res, 2006, 159: 211–222
9 Jin J, Allison B Z, Zhang Y, et al. An ERP-based BCI using an oddball paradigm with different faces and reduced errors in critical functions. Int J Neur Syst, 2014, 24: 1450027
10 Lisi G, Rivela D, Takai A, et al. Markov switching model for quick detection of event related desynchronization in EEG. Front Neurosci, 2018, 12: 24
11 Zhang L, Song L, Du B, et al. Nonlocal low-rank tensor completion for visual data. IEEE Trans Cybern, 2021, 51: 673–685
12 Gunnarsdottir K M, Bulacio J, Gonzalez-Martinez J, et al. Estimating intracranial EEG signals at missing electrodes in epileptic networks. In: Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Berlin, 2019
13 Cichocki A, Lee N, Oseledets I, et al. Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions. FNT Mach Learn, 2016, 9: 249–429
14 Zhou G, Zhao Q, Zhang Y, et al. Linked component analysis from matrices to high-order tensors: Applications to biomedical data. Proc IEEE, 2016, 104: 310–331
15 Shi Q, Cheung Y M, Zhao Q, et al. Feature extraction for incomplete data via low-rank tensor decomposition with feature regularization. IEEE Trans Neur Netw Learn Syst, 2019, 30: 1803–1817
16 Yu J, Li C, Zhao Q, et al. Tensor-ring nuclear norm minimization and application for visual: Data completion. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton, 2019. 3142–3146
17 Sorensen M, de Lathauwer L. Fiber sampling approach to canonical polyadic decomposition and application to tensor completion. SIAM J Matrix Anal Appl, 2019, 40: 888–917
18 Gu Y, Liu T, Li J. Superpixel tensor model for spatial-spectral classification of remote sensing images. IEEE Trans Geosci Remote Sens, 2019, 57: 4705–4719
19 Cui G, Gui L, Zhao Q, et al. Bayesian CP factorization of incomplete tensor for EEG signal application. In: Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Vancouver, 2016. 2170–2173
20 Ofner P, Schwarz A, Pereira J, et al. Upper limb movements can be decoded from the time-domain of low-frequency EEG. PLoS ONE, 2017, 12: 0182578
21 Acar E, Dunlavy D M, Kolda T G, et al. Scalable tensor factorizations for incomplete data. Chemometr Intell Lab Syst, 2011, 106: 41–56
22 Cichocki A, Mandic D, de Lathauwer L, et al. Tensor decompositions for signal processing applications: From two-way to multiway component analysis. IEEE Signal Process Mag, 2015, 32: 145–163
23 Zhao Q, Zhang L, Cichocki A. Bayesian CP factorization of incomplete tensors with automatic rank determination. IEEE Trans Pattern Anal Mach Intell, 2015, 37: 1751–1763
24 Chen Y L, Hsu C T, Liao H Y M. Simultaneous tensor decomposition and completion using factor priors. IEEE Trans Pattern Anal Mach Intell, 2014, 36: 577–591
25 Liu J, Musialski P, Wonka P, et al. Tensor completion for estimating missing values in visual data. IEEE Trans Pattern Anal Mach Intell, 2013, 35: 208–220
26 Nakanishi M, Wang Y, Chen X, et al. Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. IEEE Trans Biomed Eng, 2018, 65: 104–112
27 Xie Q, Zhao Q, Meng D, et al. Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery. IEEE Trans Pattern Anal Mach Intell, 2018, 40: 1888–1902