JYX
JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Wang, Deqing; Zhu, Yongjie; Ristaniemi, Tapani; Cong, Fengyu

Title: Extracting multi-mode ERP features using fifth-order nonnegative tensor decomposition

Year: 2018

Version: Accepted version (Final draft)

Copyright: © 2018 Elsevier B.V. All rights reserved.

Rights: CC BY-NC-ND 4.0

Rights url: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the original version:

Wang, D., Zhu, Y., Ristaniemi, T., & Cong, F. (2018). Extracting multi-mode ERP features using fifth-order nonnegative tensor decomposition. Journal of Neuroscience Methods, 308, 240-247. https://doi.org/10.1016/j.jneumeth.2018.07.020

Accepted Manuscript

Title: Extracting multi-mode ERP features using fifth-order nonnegative tensor decomposition

Author: Deqing Wang Yongjie Zhu Tapani Ristaniemi Fengyu Cong

PII: S0165-0270(18)30235-8
DOI: https://doi.org/doi:10.1016/j.jneumeth.2018.07.020
Reference: NSM 8071

To appear in: *Journal of Neuroscience Methods*

Received date: 23-4-2018
Revised date: 8-7-2018
Accepted date: 31-7-2018

Please cite this article as: Deqing Wang, Yongjie Zhu, Tapani Ristaniemi, Fengyu Cong, Extracting multi-mode ERP features using fifth-order nonnegative tensor decomposition, *<![CDATA[Journal of Neuroscience Methods]]>* (2018), https://doi.org/10.1016/j.jneumeth.2018.07.020

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Extracting Multi-mode ERP Features Using Fifth-order Nonnegative Tensor Decomposition

Deqing Wang[a,b], Yongjie Zhu[a,b], Tapani Ristaniemi[b], Fengyu Cong[a,b,*]

[a] *School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China*
[b] *Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland*

Abstract

Background: Preprocessed Event-related potential (ERP) data are usually organized in multi-way tensor, in which tensor decomposition serves as a powerful tool for data processing. Due to the limitation of computation burden for multi-way data and the low algorithm performance of stability and efficiency, multi-way ERP data are conventionally reorganized into low-order tensor or matrix before further analysis. However, the reorganization may hamper mode specification and spoil the interaction information among different modes.

New Method: In this study, we applied a fifth-order tensor decomposition to a set of fifth-order ERP data collected by exerting proprioceptive stimulus on left and right hand. One of the latest nonnegative CANDECOMP/PARAFAC (NCP) decomposition methods implemented by alternating proximal gradient (APG) was employed. We also proposed an improved DIFFIT method to select the optimal component number for the fifth-order tensor decomposition.

Results: By the fifth-order NCP model with a proper component number, the ERP data were fully decomposed into spatial, spectral, temporal, subject and condition factors in each component. The results showed more pairs of components with symmetric activation region in left and right hemisphere elicited by

*Corresponding author
Email addresses: deqing.wang@foxmail.com (Deqing Wang), yongjie.zhu@foxmail.com (Yongjie Zhu), tapani.e.ristaniemi@jyu.fi (Tapani Ristaniemi), cong@dlut.edu.cn (Fengyu Cong)
URL: www.deqing.me (Deqing Wang), users.jyu.fi/~riesta (Tapani Ristaniemi), www.escience.cn/people/cong (Fengyu Cong)

Accepted Manuscript

ACCEPTED MANUSCRIPT

contralateral stimuli on hand.

Comparison with Existing Method(s): In our experiment, more interesting components and coherent brain activities were extracted, compared with previous studies.

Conclusions: The discovered activities elicited by proprioceptive stimulus are in line with those in relevant cognitive neuroscience studies. Our proposed method has proved to be appropriate and viable for processing high-order EEG data with well-preserved interaction information among all modes.

Keywords: nonnegative tensor decomposition, CANDECOMP/PARAFAC, event-related potential, multi-mode features, component number selection

1. Introduction

Tensor decomposition, as a versatile tool for signal processing and machine learning (Cichocki et al., 2015; Sidiropoulos et al., 2017), has become more and more popular for EEG data processing and cognitive neuroscience in recent years (Cong et al., 2015; Zhou et al., 2016; Mahyari et al., 2017; Idaji et al., 2017; Wang et al., 2018). Event-related potential (ERP) is a time-locked EEG activity measuring brain response elicited by perceptual, cognitive or motor events (Handy, 2005). ERP can be represented naturally in tensor form. In these studies (Zhang et al., 2013; Idaji et al., 2017; Vanderperren et al., 2012; Niknazar et al., 2014), ERP data are arranged in channel × time × trial tensor. If ERP data are transformed into time-frequency representation (by Continuous Wavelet Transform), a new frequency mode will be introduced. Consequently, the ERP data of single-trial or the average of trials can be represented by channel × frequency × time tensor (Mørup et al., 2006a; Weis et al., 2009; Zhao et al., 2011; Cong et al., 2012, 2013). In an experiment, ERP data can be represented by high-order tensor including modes such as space, frequency, time, trial, subject, condition and group (Cong et al., 2015). Afterwards, tensor decomposition can be performed for the multi-way ERP data.

CANDECOMP/PARAFAC (CP), as a basic tensor decomposition method

ACCEPTED MANUSCRIPT

(Sidiropoulos et al., 2017), has remarkable advantage in processing high-order EEG data, in which a group of related feature factors can be extracted from each mode (a feature factor refers to a column of the matrix in some mode after tensor decomposition). When time-frequency representation is applied on EEG time series, the transformed data are nonnegative, and nonnegative CANDECOMP/PARAFAC (NCP) is preferred with nonnegative constraints in each mode (Cichocki et al., 2009). The rationale for using NCP in time-frequency transformed EEG tensor is that: after decomposition, (1) the temporal factor representing temporal envelope is nonnegative; (2) the spectral factor representing spectrum is nonnegative; and (3) the spatial factor representing topography can also be nonnegative. Specifically, for a brain activity, the temporal envelope exhibits the temporal evolution; the spectrum reveals the most prominent frequency band; and the topography indicates the location on the scalp. The meaning of NCP for third-order EEG tensor (channel $\times$ frequency $\times$ time) is illustrated in Fig. 1. When higher-order (> 3) EEG tensor is decomposed by NCP, more feature factors from different modes (e.g. subject, condition) will be extracted simultaneously, with the degree of strength given by score vectors being nonnegative.

In general, the processing of high-order EEG tensor data is time-consuming, in which the stability and convergence of tensor decomposition algorithms cannot be guaranteed. Hence high-order EEG data are often reshaped into low-order tensor by merging several modes together, or unfolded into two-order matrix. In a study of wavelet transformed ERP (Mørup et al., 2006b), the fifth-order ERP data (channel $\times$ frequency $\times$ time $\times$ subject $\times$ condition) were reorganized into a third-order tensor (channel $\times$ time-frequency $\times$ subject-condition). In (Cong et al., 2012), a fourth-order ERP tensor (channel $\times$ frequency $\times$ time $\times$ subject-condition) was generated with subject and condition modes being merged. However, the merging or unfolding of modes can potentially hamper data interpretation, dismiss mode specification, and spoil the interaction information among these modes (Cong et al., 2015; Mørup et al., 2006b).

Accepted Manuscript

ACCEPTED MANUSCRIPT

Mørup et al. applied fifth-order NCP to decompose another fifth-order ERP data, the condition mode size of which is two (Mørup et al., 2006a). However, there existed an unreasonable assumption that the data could only entail a two-component CP model (Mørup et al., 2006a). The rationale behind this assumption might be that the component number was no more than the minimal mode size of two. In fact, the selection of component number for CP model is related to the rank of tensor. Since the rank of tensor can be larger than the maximal mode size (Sidiropoulos et al., 2017), a larger component number can be selected.

In this study, we analyze a set of fifth-order ERP data (channel × frequency × time × subject × condition) elicited by proprioceptive stimulus. Fifth-order NCP is applied with a large component number. The data are fully decomposed in each mode, and the interaction information among these five modes is retained. One of the latest NCP models implemented by alternating proximal gradient (APG) method is employed, which has been proved to be mathematically convergent and stable (Xu & Yin, 2013).

It is nontrivial to determine the component number by the tensor rank directly. Some studies have focused on the selection of component number for multi-way models, such as core consistency diagnostic (CORCONDIA) (Bro & Kiers, 2003) and automatic relevance determination (ARD) (Mørup & Hansen, 2009). Nevertheless, these methods often indicate very few components, which are not adequate to reveal the physiological properties of EEG signals (Cong et al., 2015). Conventional DIFFIT (Timmerman & Kiers, 2000) is a preferred method for component number selection in EEG data processing (Cong et al., 2015), but it usually fails to provide useful information due to fluctuations on DIFFIT curve. Therefore, we propose an improved smooth DIFFIT method that can select a proper number of components to make sure that the most important components are included.

The assumption of the fifth-order analysis is that the underlying spatial, spectral, and temporal factors are the same among all subjects, however, with a subject-dependent strength given by the subject score vector and with a variable

ACCEPTED MANUSCRIPT

strength in all conditions (Mørup et al., 2006a). Our method satisfies the above assumption and meanwhile reveals more interesting components and coherent activities compared with previous study (Mørup et al., 2006b). Our findings are also in line with related cognitive neuroscience explanations of proprioceptive stimulus (Herrmann et al., 2004b; Arnfred et al., 2011).

2. Nonnegative CANDECOMP/PARAFAC Decomposition

2.1. Notation

In this paper, we denote a vector by boldface lowercase letter, such as $\boldsymbol{x}$; a matrix by boldface uppercase letter, such as $\boldsymbol{X}$; and a high order tensor by boldface Euler script letter, such as $\boldsymbol{\mathcal{X}}$. Operator $\circ$ represents outer product of vectors, $\langle\ \rangle$ represents inner product, $[\![\ ]\!]$ represents Kruskal operator, and $\|\ \|_F$ means Frobenius norm. Nonnegative CANDECOMP/PARAFAC decomposition is abbreviated as NCP for convenience in following contents.

2.2. Mathematical Model

Given a nonnegative Nth-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, NCP is to solve the following minimization problem:

$$\begin{aligned} \min_{\boldsymbol{A}^{(1)},\ldots,\boldsymbol{A}^{(N)}} & \frac{1}{2}\left\|\boldsymbol{\mathcal{X}} - [\![\boldsymbol{A}^{(1)},\ldots,\boldsymbol{A}^{(N)}]\!]\right\|_F^2 \\ \text{s.t. } & \boldsymbol{A}^{(n)} \geqslant 0 \text{ for } n = 1,\ldots,N, \end{aligned} \tag{1}$$

where $\boldsymbol{A}^{(n)} \in \mathbb{R}^{I_n \times R}$ for $n = 1,\ldots,N$ are the estimated factor matrices in different modes, I_n is the size in mode-n, and R is the selected rank-1 tensor number (component number). The estimated factor matrices in Kruskal operator can be represented by the sum of R rank-1 tensors in outer product form:

$$[\![\boldsymbol{A}^{(1)},\ldots,\boldsymbol{A}^{(N)}]\!] = \sum_{r=1}^{R} \widetilde{\boldsymbol{\mathcal{X}}}_r = \sum_{r=1}^{R} \boldsymbol{a}_r^{(1)} \circ \cdots \circ \boldsymbol{a}_r^{(N)}, \tag{2}$$

where $\boldsymbol{a}_r^{(n)}$ represents the rth column of $\boldsymbol{A}^{(n)}$.

ACCEPTED MANUSCRIPT

2.3. Optimization Scheme

Conventionally, there are many optimization methods that can be applied to solve NCP problem, such as multiplicative updating (MU), alternating least squares (ALS), hierarchical alternating least squares (HALS) (Cichocki et al., 2009). Recently, Xu et al. (Xu & Yin, 2013) have proposed the alternating proximal gradient (APG) method to solve nonnegative matrix and tensor decomposition problems in block coordinate descend (BCD) framework. APG outperforms many conventional methods both in accuracy and efficiency. More importantly, it has better convergence properties. Therefore, APG method is employed to solve NCP in our study. We don't make a comparison among different optimization methods in this study, which might be done in future.

Supposing that $\widehat{\boldsymbol{A}}^{(n)}$ is an extrapolated point, $\widehat{\boldsymbol{G}}^{(n)}$ is the block-partial gradient at $\widehat{\boldsymbol{A}}^{(n)}$ and $L^{(n)}$ is a Lipschitz constant, factor matrix $\boldsymbol{A}^{(n)}$ can be updated alternatively by

$$\boldsymbol{A}^{(n)} \leftarrow \underset{\boldsymbol{A}^{(n)} \geqslant 0}{\arg\min} \left\langle \widehat{\boldsymbol{G}}^{(n)}, \boldsymbol{A}^{(n)} - \widehat{\boldsymbol{A}}^{(n)} \right\rangle + \frac{L^{(n)}}{2} \left\| \boldsymbol{A}^{(n)} - \widehat{\boldsymbol{A}}^{(n)} \right\|_F^2 . \tag{3}$$

Furthermore, (3) can be rewritten in the closed form

$$\boldsymbol{A}^{(n)} \leftarrow \max\left(0, \widehat{\boldsymbol{A}}^{(n)} - \widehat{\boldsymbol{G}}^{(n)}/L^{(n)}\right). \tag{4}$$

The detailed solution of NCP model (1) using APG and its convergence analysis can be found in (Xu & Yin, 2013).

3. Component Number Estimation

Before applying tensor decomposition to EEG data, it is required to determine a proper component number, which is the rank-1 tensor number R in (2). The component number is closely related to the rank of tensor. For a two-way matrix $\boldsymbol{X} \in \mathbb{R}^{I \times J}$, it follows $\text{rank}(\boldsymbol{X}) \leqslant \min(I, J)$, but for a multi-way tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, $\text{rank}(\boldsymbol{\mathcal{X}})$ can be even larger than $\max(I_1, I_2, \ldots, I_N)$ (Sidiropoulos et al., 2017). Therefore, the component number for tensor decomposition can be selected with a large number, which is not restricted by the size

ACCEPTED MANUSCRIPT

in each mode. However, it is nontrivial to determine R by tensor rank. DIFFIT is a conventional method to determine component number, and has been used in ERP tensor data decomposition (Cong et al., 2012).

3.1. DIFFIT Method

DIFFIT refers to the difference in data fitting, and is calculated based on relative error/residual and the explained variance (or the explained sum of squares) (Timmerman & Kiers, 2000; Mørup & Hansen, 2009). The relative error/residual of NCP is defined as

$$\text{RelErr} = \frac{\left\| \boldsymbol{\mathcal{X}} - [\![\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(N)}]\!] \right\|_F}{\|\boldsymbol{\mathcal{X}}\|_F}. \tag{5}$$

Let component number $R \in [1, \mathcal{R}]$, where $\mathcal{R}$ is the empirically maximal component number. When R is selected, the explained variance of the estimated R components from NCP is

$$\text{ExpVar}(R) = 1 - \frac{\left\| \boldsymbol{\mathcal{X}} - [\![\boldsymbol{A}^{(1)}, \ldots, \boldsymbol{A}^{(N)}]\!] \right\|_F^2}{\|\boldsymbol{\mathcal{X}}\|_F^2}. \tag{6}$$

In order to obtain a more precise value, NCP model is usually run many times (e.g. T times) for each R and the average explained variance is calculated yielding an averaged sequence

$$\boldsymbol{e} = \left[\overline{\text{ExpVar}}(1), \ldots, \overline{\text{ExpVar}}(R), \ldots, \overline{\text{ExpVar}}(\mathcal{R}) \right]. \tag{7}$$

Then, calculate the DIF (difference of explained variance) and DIFFIT as follows:

$$\text{DIF}(R) = \overline{\text{ExpVar}}(R) - \overline{\text{ExpVar}}(R-1), \tag{8}$$

$$\text{DIFFIT}(R) = \frac{\text{DIF}(R)}{\text{DIF}(R+1)}. \tag{9}$$

Sometimes, NCP model is run tens or hundreds of times for each R, which is very time-consuming and even worse for big data or higher-order tensor. Meanwhile, since the discrete sequence of averaged explained variances has only

Accepted Manuscript

ACCEPTED MANUSCRIPT

$\mathcal{R}$ values (usually several tens), there probably exist serious fluctuations and fake spikes on the final DIFFIT curve. Then, we propose an improved method named "smoothed DIFFIT".

3.2. Smoothed DIFFIT

Run NCP fewer times ($T = 10, 20$, or other values according to the computer's computation ability) for each R, and obtain vector $\boldsymbol{e}$ based on (7). From vector $\boldsymbol{e}$, a smooth curve is generated by polynomial curve fitting (PCF) using a mth-order equation

$$p(x) = p_1 x^m + p_2 x^{m-1} +, \ldots, + p_m x^1 + p_{m+1}. \tag{10}$$

After polynomial curve fitting, the new sequence is

$$\boldsymbol{e}_{\mathrm{PCF}} = \left[\mathrm{ExpVar}_{\mathrm{PCF}}(1), \ldots, \mathrm{ExpVar}_{\mathrm{PCF}}(\mathcal{R})\right]. \tag{11}$$

We recompute the DIF and DIFFIT as follows:

$$\mathrm{DIF}_{\mathrm{PCF}}(R) = \mathrm{ExpVar}_{\mathrm{PCF}}(R) - \mathrm{ExpVar}_{\mathrm{PCF}}(R-1), \tag{12}$$

$$\mathrm{DIFFIT}_{\mathrm{PCF}}(R) = \frac{\mathrm{DIF}_{\mathrm{PCF}}(R)}{\mathrm{DIF}_{\mathrm{PCF}}(R+1)}. \tag{13}$$

4. Experiments and Results

4.1. Data Description

The ERP data in our experiment come from an open preprocessed dataset associated with ERPWAVELAB toolbox (Mørup et al., 2007), which can be downloaded from www.erpwavelab.org. The data were collected from a proprioceptive experiment, in which two conditions (left and right hand) were manipulated with the increment of handhold load. An important part of the stimuli is the change of applied force on a static muscle contraction, which is conceived as proprioceptive stimulus (Mørup et al., 2006b). Fourteen subjects participated in the experiment and 64 scalp electrodes were used to record EEG data. A

ACCEPTED MANUSCRIPT

total of 360 trials (epochs) were obtained from each subject under each condition. All epochs were transformed into time-frequency representation (TFR) by complex Morlet wavelet. In the wavelet transform, only the frequency band from 15 Hz to 75 Hz were analyzed with linear interval of 1Hz. Then, inter-trial phase coherence (ITPC) (Delorme & Makeig, 2004) was computed as an average spectral estimate across all trials. Since the TFR was first applied to each trial and then the average was calculated across trials, ITPC can be seen as induced oscillations of brain (David et al., 2006). Meanwhile, ITPC only takes values between 0 and 1 (Delorme & Makeig, 2004; Cohen, 2014). Finally, a fifth-order nonnegative tensor (channel $\times$ frequency $\times$ time $\times$ subject $\times$ condition $= 64 \times 61 \times 72 \times 14 \times 2$) was generated. The structure of the fifth-order tensor is shown in Fig. 2. The 61 frequency points represent $15 - 75$ Hz, and the 72 time points represent $0 - 346.68$ ms.

The detailed EEG data collection and preprocessing procedures are described in (Mørup et al., 2006b).

4.2. Component Number Selection

Both the original DIFFIT and our proposed smoothed DIFFIT methods were tested. The results are shown in Fig. 3.

We tested NCP on the fifth-order tensor data by increasing R from 1 to 50 ($\mathcal{R} = 50$). With each selected component number R, NCP was run 20 times. The averaged explained variance vector $\boldsymbol{e}$, DIF and DIFFIT curves are shown in Fig. 3 (a)-(c). Obviously, many fluctuations appear on the DIF and DIFFIT curves. From Fig. 3 (b) we can find that the DIF values at $R = 31$ and $R = 44$ are very close to 0. According to (9), this will cause large DIFFIT values with two long fake spikes in Fig. 3 (c) at $R = 30$ and $R = 43$. The DIFFIT curve in Fig. 3 (c) entails very limited useful information.

We fit $\boldsymbol{e}$ by a 7th-order polynomial, and the obtained polynomial is

$$\begin{aligned} p(x) =& 2.0486^{-11}x^7 - 3.9518^{-9}x^6 + 3.0904^{-7}x^5 \\ & - 1.2586^{-5}x^4 + 2.8604^{-4}x^3 - 3.6465^{-3}x^2 \\ & + 2.6685^{-2}x + 7.3496^{-1}. \end{aligned}$$

ACCEPTED MANUSCRIPT

The fitted curve $\boldsymbol{e}_{\mathrm{PCF}}$ is shown in Fig. 3 (d). By visual inspection, the results of 5th and 6th-order polynomials were under-fitting, while 8th and 9th-order were over-fitting. 7th-order polynomial has the best fitting for the data in our experiment.

Subsequently, $\mathrm{DIF}_{\mathrm{PCF}}$ and $\mathrm{DIFFIT}_{\mathrm{PCF}}$ were calculated, as shown in Fig. 3 (e) and (f). In Fig. 3 (f), when the DIFFIT value is larger than 1, DIF curve decreases; when it is smaller than 1, DIF curve increases; and when it equals to 1, DIF curve doesn't change. The two local maximums on DIFFIT curve at $R = 4$ and $R = 24$ indicate two positions on DIF curve that have fast dropping rate. The values after $R = 45$ can be ignored due to the poor fitting at the end of the curve. From Fig. 3 (e) and (f) we find that at $R = 20$ the DIF value decreases rapidly again approaching to a rather low value, which indicates that, after $R = 20$, the explained variance doesn't increase that much. Based on above analysis, we assume that, at $R = 20$, the estimated factors of tensor decomposition contain the most important EEG components. Therefore, $R = 20$ is selected as the optimal component number in NCP model for this fifth-order tensor data. In fact, our smoothed DIFFIT method can estimate a narrow range of proper component numbers. For example, in case of $R = 19$ or 21, similar results can be obtained as $R = 20$. However, if $R = 30$, more unimportant components, such as weak signals or noise, might be included in the results.

4.3. Tensor Decomposition Implementation

4.3.1. Factor Initialization

All factors of the five modes were initialized using normally distributed random numbers as in (Xu & Yin, 2013).

4.3.2. Stop Criteria

Based on (5), the relative error/residual of NCP at the kth iteration was defined as $\mathrm{RelErr}_k = \frac{\left\|\boldsymbol{\mathcal{X}} - [\![\boldsymbol{A}_k^{(1)}, \ldots, \boldsymbol{A}_k^{(N)}]\!]\right\|_F}{\|\boldsymbol{\mathcal{X}}\|_F}$. We terminated the NCP decomposition process when the following condition between two iterations k and $k+1$ was

Accepted Manuscript

ACCEPTED MANUSCRIPT

satisfied:

$$T_{\text{stop}} = |\text{RelErr}_k - \text{RelErr}_{k+1}| \leq \epsilon,$$

where $\epsilon = 10^{-6}$ in this study.

4.4. Components Analysis

In this section, we name the rth rank-1 tensor $\widetilde{\boldsymbol{\mathcal{X}}}_r$ after decomposition as the rth component. A component further contains spatial, temporal, temporal, subject, and condition factor, which can be represented by $\widetilde{\boldsymbol{\mathcal{X}}}_r = \boldsymbol{a}_r^{(spatial)} \circ \boldsymbol{a}_r^{(spectral)} \circ \boldsymbol{a}_r^{(temporal)} \circ \boldsymbol{a}_r^{(subject)} \circ \boldsymbol{a}_r^{(condition)}$ according to (2).

After NCP decomposition, a total of 20 components were obtained. The assumption of the fifth-order analysis is that the underlying spatial, spectral, and temporal factors are the same among all subjects only with differences in subject-dependent and condition-dependent strength. In the results, we discovered five pairs of components (10 components) with symmetric responses on topographies, which are shown in Fig. 4. and Fig. 5. It is clear that left-hand stimuli elicit activities in the right hemisphere, and right-hand stimuli elicit activities in the left hemisphere. We also calculated time-frequency representation (TFR) for each component, which is the outer product of the spectral and temporal factor. The frequency of TFR is presented in logarithm scale.

Fig. 4 shows the components with frequency oscillations in beta-band (15-30Hz). Fig. 4 (a) displays a pair of components with symmetric activities in right and left frontal lobe elicited by contralateral hand stimuli. Both components appear within 15-20Hz at 75ms. This pair of components was also found in (Mørup et al., 2006b). Fig. 4 (b) shows symmetric activities in right and left temporal lobe within 15-20Hz at 75ms. Activities in Fig. 4 (c) appear in right and left frontal lobe within 25-30Hz at 60ms. These beta-band activities emerge before 100ms after stimulus onset, which is consistent with previous study about beta-frequency oscillations of proprioceptive information processing (Arnfred et al., 2011).

Fig. 5 demonstrates the components with frequency oscillations in gamma-band (30-75Hz). Fig. 5 (a) shows symmetric activities in right and left temporal

ACCEPTED MANUSCRIPT

lobe, and they occur within 30-40Hz at 60ms. This pair of components was also found in (Mørup et al., 2006b). In Fig. 5 (b), the other two components with the frequency of 40-75Hz in the region between parietal and temporal lobe appear respectively at 75ms and 45ms. It has been reported in (Mørup et al., 2006b; Arnfred et al., 2011) that proprioceptive stimulus could elicit gamma-band activities (GBA, 30-80Hz). Our fifth-order tensor decomposition can extract more gamma-band components compared with previous study (Mørup et al., 2006b). These gamma-band activities, occurring before 100ms after stimulus onset in parietal and temporal region, are in accordance with the match step of match-and-utilization model (MUM) in cognitive neuroscience (Herrmann et al., 2004a,b). It is assumed in MUM that the 'early' gamma-band response reflects the matching of stimulus-related information with memory contents that occur rapidly (before 150 ms) after stimulus onset (Herrmann et al., 2004b). We infer that, when a weight increment is exerted on subjects' hand repeatedly during the experiment, there will be a matching and comparison process between the new proprioceptive-stimulus-related signals and existing memory contents. These gamma-band activities in Fig. 5 can be regarded as the results of the matching and comparison process.

In addition, we also discovered two extra components of high strength in both left-hand and right-hand conditions. One of the components occurs in occipital lobe with the frequency of 22Hz at 80ms, while another occurs at the center of parietal lobe with the frequency of 28Hz at 130ms. Perhaps, these two components were elicited by non-proprioceptive stimulus.

5. Conclusion

In this study, we applied fifth-order NCP model to decompose a set of fifth-order ERP tensor data collected by exerting proprioceptive stimulus on left and right hand. The data were fully decomposed in all modes (space, frequency, time, subject, condition), and the interaction information among these modes was well retained. We also proposed an improved smoothed DIFFIT method to

ACCEPTED MANUSCRIPT

select a proper component number for NCP with the most important components being well-preserved after decomposition. Compared with previous study, we discovered more interesting components with symmetric and coherent activities elicited by left-hand and right-hand stimuli. Our findings are consistent with the explanations in related cognitive neuroscience studies. In future studies, some aspects still need further consideration. From the perspective of signal processing algorithm, it would be interesting to compare the performance and stability of NCP tensor decomposition in fifth-order ERP data implemented by different optimization methods, such as APG, MU, ALS, and HALS. From the perspective of cognitive neuroscience methodology, it is worth analyzing experiments results using diverse data forms, such as fifth-order tensor analysis, reshaped third-order tensor analysis, subject-based analysis and group-based analysis.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. 81471742), the Fundamental Research Funds for the Central Universities [DUT16JJ(G)03] in Dalian University of Technology in China, and the scholarships from China Scholarship Council (Nos. 201600090043 and 201600090042). The authors would like to express their thanks to Lili Tian for refining the written English.

References

Arnfred, S. M. H., Mørup, M., Thalbitzer, J., Jansson, L., & Parnas, J. (2011). Attenuation of beta and gamma oscillations in schizophrenia spectrum patients following hand posture perturbation. *Psychiatry Research*, *185*, 215–224. doi:10.1016/j.psychres.2009.10.005.

Bro, R., & Kiers, H. A. L. (2003). A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics*, *17*, 274–286. doi:10.1002/cem.801.

ACCEPTED MANUSCRIPT

Cichocki, A., Mandic, D., Lathauwer, L. D., Zhou, G., Zhao, Q., Caiafa, C., & PHAN, H. A. (2015). Tensor decompositions for signal processing applications: From two-way to multiway component analysis. *IEEE Signal Processing Magazine*, *32*, 145–163. doi:10.1109/msp.2013.2297439.

Cichocki, A., Zdunek, R., Phan, A. H., & Amari, S.-i. (2009). *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. Issues in Clinical and Cognitive Neuropsychology. MIT press.

Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., & Ristaniemi, T. (2015). Tensor decomposition of EEG signals: A brief review. *Journal of Neuroscience Methods*, *248*, 59–69. doi:10.1016/j.jneumeth.2015.03.018.

Cong, F., Phan, A.-H., Astikainen, P., Zhao, Q., Wu, Q., Hietanen, J. K., Ristaniemi, T., & Cichocki, A. (2013). Multi-domain feature extraction for small event-related potentials through nonnegative multi-way array decomposition from low dense array EEG. *International Journal of Neural Systems*, *23*, 1350006. doi:10.1142/s0129065713500068.

Cong, F., Phan, A. H., Zhao, Q., Huttunen-Scott, T., Kaartinen, J., Ristaniemi, T., Lyytinen, H., & Cichocki, A. (2012). Benefits of multi-domain feature of mismatch negativity extracted by non-negative tensor factorization from EEG collected by low-density array. *International Journal of Neural Systems*, *22*, 1250025. doi:10.1142/s0129065712500256.

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, *31*, 1580–1591. doi:10.1016/j.neuroimage.2006.02.034.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analy-

ACCEPTED MANUSCRIPT

sis. *Journal of Neuroscience Methods*, *134*, 9–21. doi:10.1016/j.jneumeth.2003.10.009.

Handy, T. C. (2005). *Event-related potentials: A methods handbook*. MIT press.

Herrmann, C. S., Lenz, D., Junge, S., Busch, N. A., & Maess, B. (2004a). Memory-matches evoke human gamma-responses. *BMC Neuroscience*, *5*, 13. doi:10.1186/1471-2202-5-13.

Herrmann, C. S., Munk, M. H., & Engel, A. K. (2004b). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences*, *8*, 347–355. doi:10.1016/j.tics.2004.06.006.

Idaji, M. J., Shamsollahi, M. B., & Sardouie, S. H. (2017). Higher order spectral regression discriminant analysis (HOSRDA): A tensor feature reduction method for ERP detection. *Pattern Recognition*, *70*, 152–162. doi:10.1016/j.patcog.2017.05.004.

Mahyari, A. G., Zoltowski, D. M., Bernat, E. M., & Aviyente, S. (2017). A tensor decomposition-based approach for detecting dynamic network states from EEG. *IEEE Transactions on Biomedical Engineering*, *64*, 225–237. doi:10.1109/tbme.2016.2553960.

Mørup, M., & Hansen, L. K. (2009). Automatic relevance determination for multi-way models. *Journal of Chemometrics*, *23*, 352–363. doi:10.1002/cem.1223.

Mørup, M., Hansen, L. K., & Arnfred, S. M. (2007). ERPWAVELAB: A toolbox for multi-channel analysis of timefrequency transformed event related potentials. *Journal of Neuroscience Methods*, *161*, 361–368. doi:10.1016/j.jneumeth.2006.11.008.

Mørup, M., Hansen, L. K., Herrmann, C. S., Parnas, J., & Arnfred, S. M. (2006a). Parallel factor analysis as an exploratory tool for wavelet transformed event-related EEG. *NeuroImage*, *29*, 938–947. doi:10.1016/j.neuroimage.2005.08.005.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Mørup, M., Hansen, L. K., Parnas, J., & Arnfred, S. M. (2006b). Decomposing the time-frequency representation of EEG using non-negative matrix and multi-way factorization. *Technical University of Denmark Technical Report*, (pp. 1–28).

Niknazar, M., Becker, H., Rivet, B., Jutten, C., & Comon, P. (2014). Blind source separation of underdetermined mixtures of event-related sources. *Signal Processing*, *101*, 52–64. doi:10.1016/j.sigpro.2014.01.031.

Sidiropoulos, N. D., Lathauwer, L. D., Fu, X., Huang, K., Papalexakis, E. E., & Faloutsos, C. (2017). Tensor decomposition for signal processing and machine learning. *IEEE Transactions on Signal Processing*, *65*, 3551–3582. doi:10.1109/tsp.2017.2690524.

Timmerman, M. E., & Kiers, H. A. L. (2000). Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology*, *53*, 1–16. doi:10.1348/000711000159132.

Vanderperren, K., Mijović, B., Novitskiy, N., Vanrumste, B., Stiers, P., den Bergh, B. R. H. V., Lagae, L., Sunaert, S., Wagemans, J., Huffel, S. V., & Vos, M. D. (2012). Single trial ERP reading based on parallel factor analysis. *Psychophysiology*, *50*, 97–110. doi:10.1111/j.1469-8986.2012.01405.x.

Wang, D., Wang, X., Zhu, Y., Toiviainen, P., Huotilainen, M., Ristaniemi, T., & Cong, F. (2018). Increasing stability of EEG components extraction using sparsity regularized tensor decomposition. In *Advances in Neural Networks – ISNN 2018* (pp. 789–799). Springer International Publishing. doi:10.1007/978-3-319-92537-0_89.

Weis, M., Romer, F., Haardt, M., Jannek, D., & Husar, P. (2009). Multi-dimensional space-time-frequency component analysis of event related EEG data using closed-form PARAFAC. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. doi:10.1109/icassp.2009.4959592.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Xu, Y., & Yin, W. (2013). A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM Journal on Imaging Sciences*, *6*, 1758–1789. doi:10.1137/120887795.

Zhang, Y., Zhou, G., Jin, J., Wang, M., Wang, X., & Cichocki, A. (2013). L1-regularized multiway canonical correlation analysis for SSVEP-based BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*, 887–896. doi:10.1109/tnsre.2013.2279680.

Zhao, Q., Caiafa, C. F., Mandic, D. P., Zhang, L., Ball, T., Schulze-bonhage, A., & Cichocki, A. S. (2011). Multilinear subspace regression: An orthogonal tensor decomposition approach. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 24* (pp. 1269–1277). Curran Associates, Inc. URL: http://papers.nips.cc/paper/4328-multilinear-subspace-regression-an-orthogonal-tensor-decomposition-approach.pdf.

Zhou, G., Zhao, Q., Zhang, Y., Adali, T., Xie, S., & Cichocki, A. (2016). Linked component analysis from matrices to high-order tensors: Applications to biomedical data. *Proceedings of the IEEE*, *104*, 310–331. doi:10.1109/jproc.2015.2474704.

Accepted Manuscript

ACCEPTED MANUSCRIPT

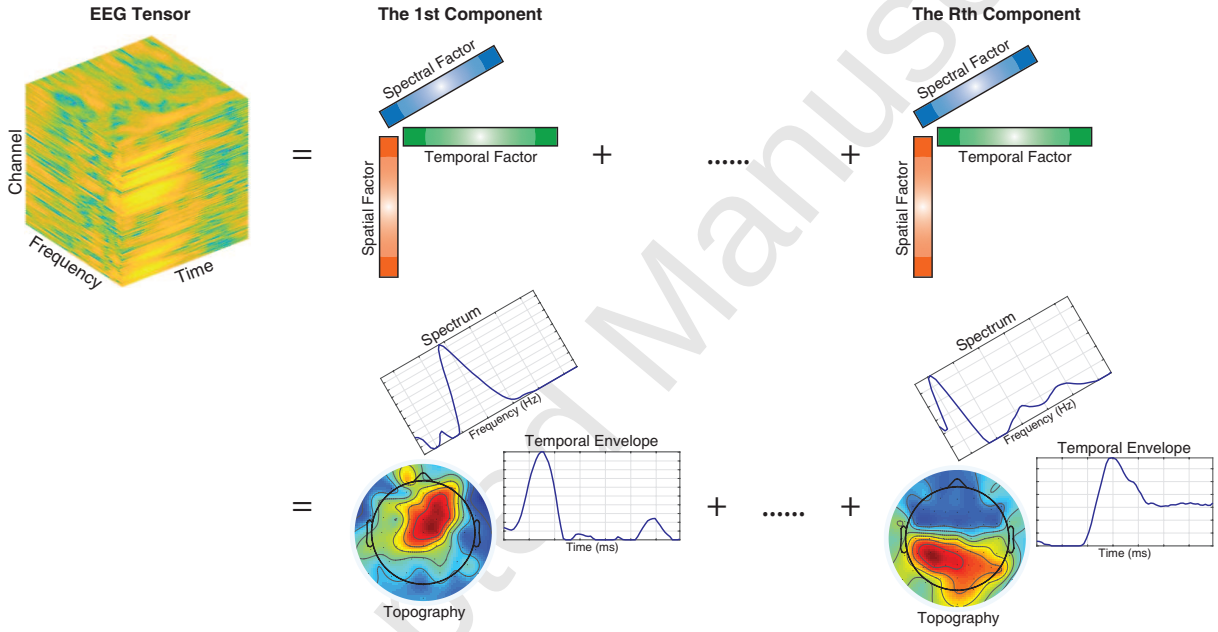


Figure 1: Tensor decomposition of third-order EEG tensor (channel × frequency × time) by NCP. The time-frequency transformed EEG tensor is nonnegative. After decomposition, the temporal factor representing temporal envelop is nonnegative, the spectral factor representing spectrum is nonnegative, and the spatial factor representing topography can also be nonnegative.

Accepted Manuscript

ACCEPTED MANUSCRIPT

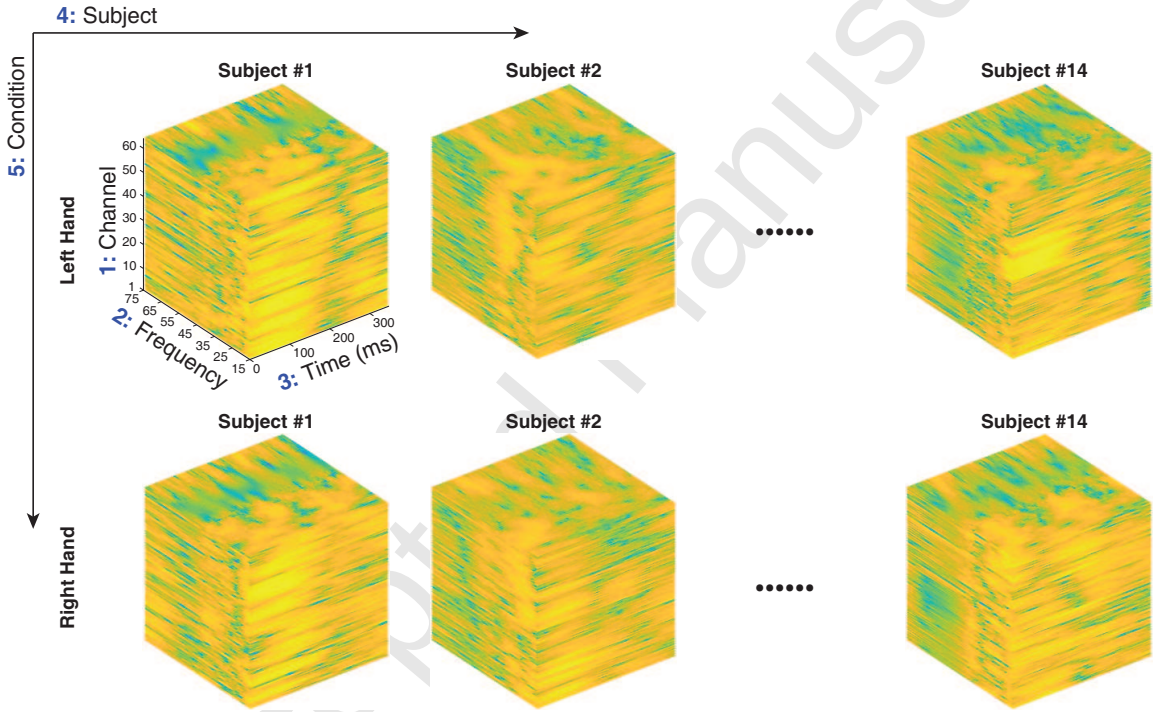


Figure 2: Fifth-order ERP tensor (channel × frequency × time × subject × condition). This tensor has 64 channels, 61 frequency points representing $15-75$ Hz, 72 time points representing $0-346.68$ ms, 14 subjects and 2 conditions of stimuli on left and right hand.

ACCEPTED MANUSCRIPT

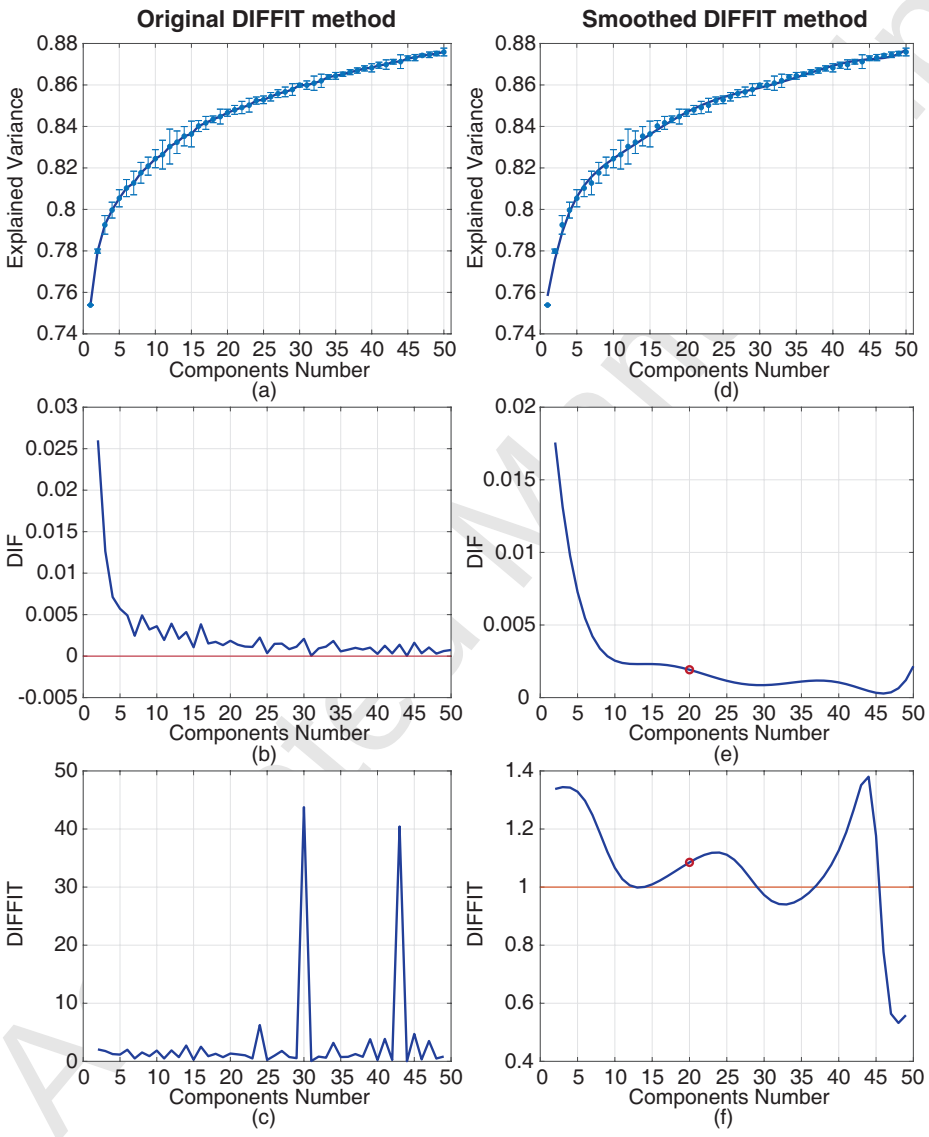


Figure 3: DIFFIT for component number selection. (a)-(c) represent original DIFFIT method. (d)-(e) represent the proposed smoothed DIFFIT method.

ACCEPTED MANUSCRIPT

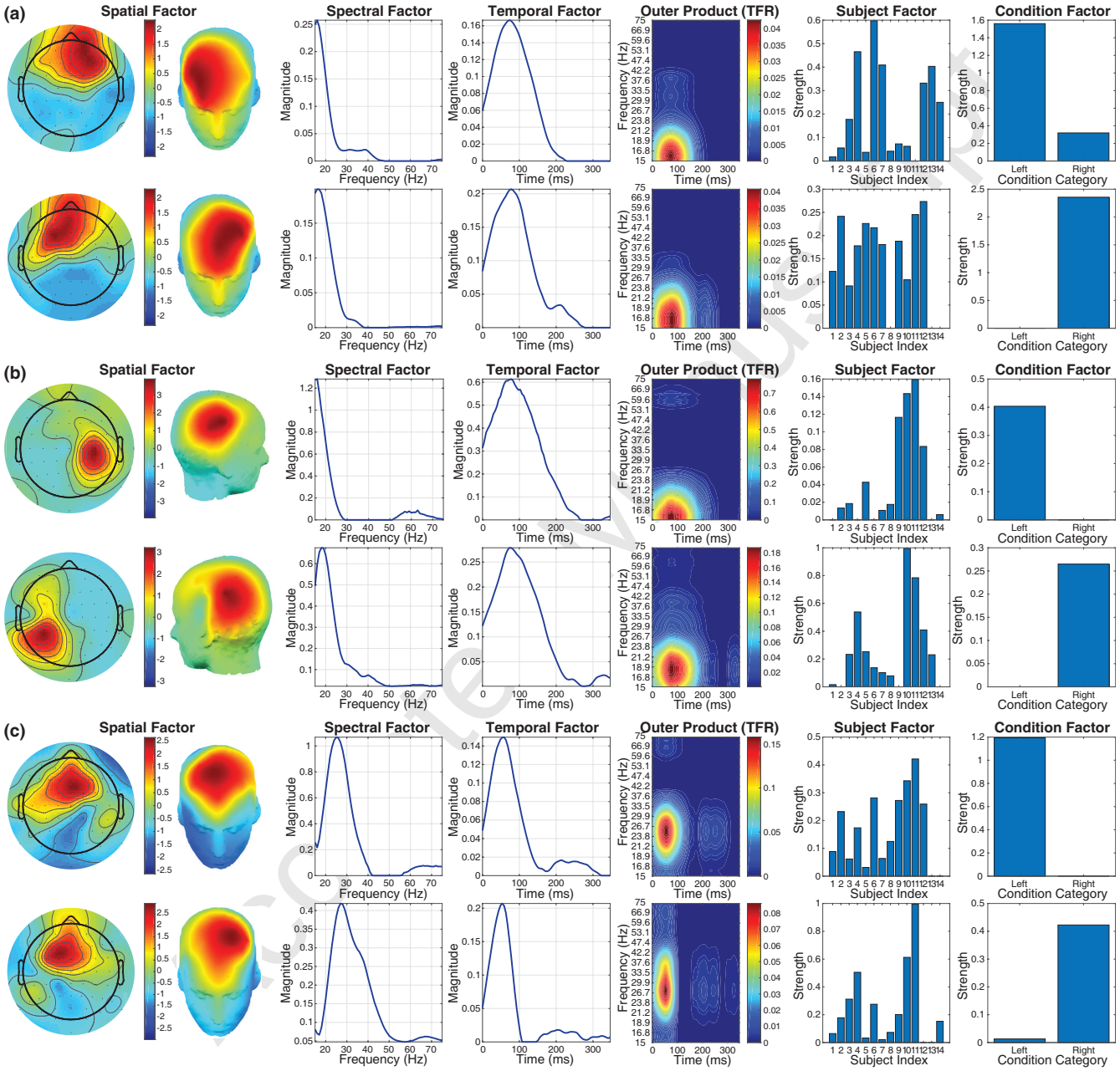


Figure 4: Components of beta-band oscillations. These components have contralateral brain activities on topographies elicited by left-hand and right-hand stimuli respectively. Activities in (a) appear in frontal lobe within 15-20Hz at 75ms. Activities in (b) occur in temporal lobe within 15-20Hz at 75ms. Activities in (c) emerge in frontal lobe within 25-30Hz at 60ms.

ACCEPTED MANUSCRIPT

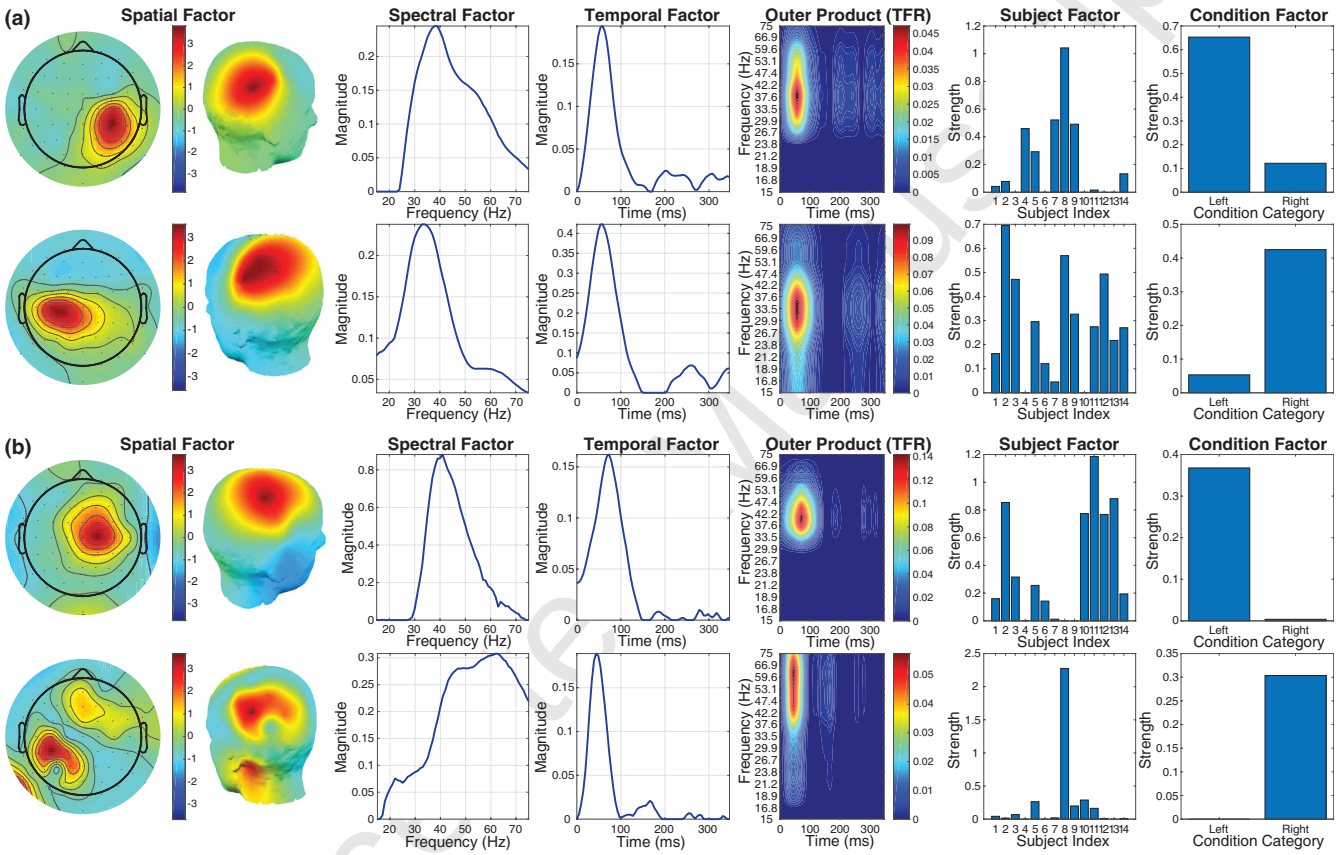


Figure 5: Components of gamma-band oscillations. These components have contralateral brain activities on topographies elicited by left-hand and right-hand stimuli respectively. Activities in (a) appear in frontal lobe within 30-40Hz at 60ms. Activity of the first component in (b) occurs between right parietal and temporal lobes at 40Hz and 75ms. Activity of the second component in (b) occurs between left parietal and temporal lobes within 40-75Hz at 45ms.

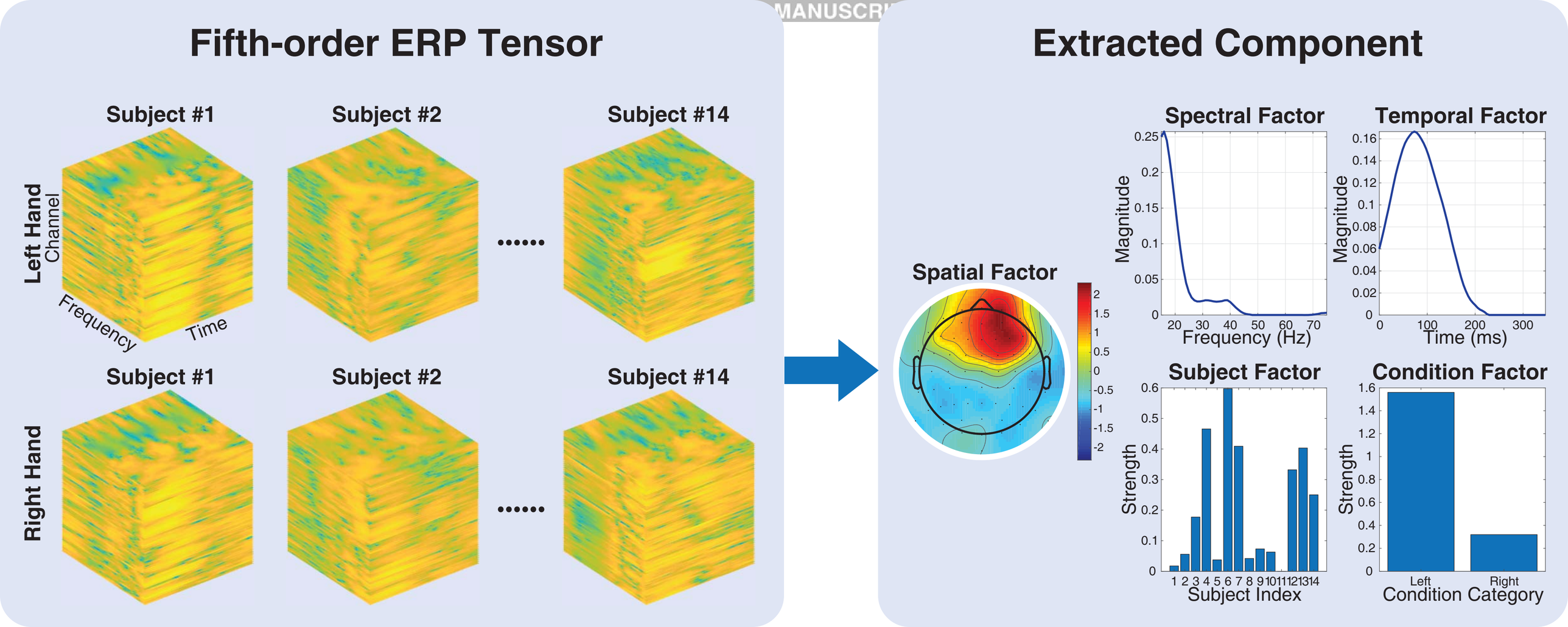
ACCEPTED MANUSCRIPT

Highlights:

- Fifth-order nonnegative CANDECOMP/PARAFAC model was applied to fifth-order ERP tensor with full decomposition in all modes.
- A smoothed DIFFIT method was proposed to select optimal component number.
- More interesting components and coherent activities were extracted than previous study.

Accepted Manuscript

MANUSCRI


Fifth-order ERP Tensor
Subject #1
Subject #2
Subject #14
Left Hand
Channel
Frequency
Time
Subject #1
Subject #2
Subject #14
Right Hand
Extracted Component
Spatial Factor
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
Spectral Factor
Magnitude
0.25
0.2
0.15
0.1
0.05
0
20 30 40 50 60 70
Frequency (Hz)
Temporal Factor
Magnitude
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0 100 200 300
Time (ms)
Subject Factor
Strength
0.6
0.5
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Subject Index
Condition Factor
Strength
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Left
Right
Condition Category