

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

Author(s): Hu, Guoqiang; Zhou, Tianyi; Luo, Siwen; Mahini, Reza; Xu, Jing; Chang, Yi; Cong, Fengyu

Title: Assessment of nonnegative matrix factorization algorithms for electroencephalography spectral analysis

Year: 2020

Version: Published version

Copyright: © The Author(s) 2020

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Hu, G., Zhou, T., Luo, S., Mahini, R., Xu, J., Chang, Y., & Cong, F. (2020). Assessment of nonnegative matrix factorization algorithms for electroencephalography spectral analysis. *Biomedical Engineering Online*, 19, Article 61. <https://doi.org/10.1186/s12938-020-00796-x>

RESEARCH

Open Access



Assessment of nonnegative matrix factorization algorithms for electroencephalography spectral analysis

Guoqiang Hu^{1†} , Tianyi Zhou^{1,2†}, Siwen Luo³, Reza Mahini¹, Jing Xu^{4*}, Yi Chang^{4*} and Fengyu Cong^{1,5,6,7*}

*Correspondence:

xujing_doc@aliyun.com;
change99@gmail.com;
cong@dlut.edu.cn

[†]Guoqiang Hu, Tianyi Zhou
contributed equally to this
article

¹ School of Biomedical
Engineering, Faculty
of Electronic Information
and Electrical Engineering,
Dalian University
of Technology, Dalian, China⁴
Department of Neurology
and Psychiatry, First Affiliated
Hospital, Dalian Medical
University, Dalian, China
Full list of author information
is available at the end of the
article

Abstract

Background: Nonnegative matrix factorization (NMF) has been successfully used for electroencephalography (EEG) spectral analysis. Since NMF was proposed in the 1990s, many adaptive algorithms have been developed. However, the performance of their use in EEG data analysis has not been fully compared. Here, we provide a comparison of four NMF algorithms in terms of accuracy of estimation, stability (repeatability of the results) and time complexity of algorithms with simulated data. In the practical application of NMF algorithms, stability plays an important role, which was an emphasis in the comparison. A Hierarchical clustering algorithm was implemented to evaluate the stability of NMF algorithms.

Results: In simulation-based comprehensive analysis of fit, stability, accuracy of estimation and time complexity, hierarchical alternating least squares (HALS) low-rank NMF algorithm (IraNMF_HALS) outperformed the other three NMF algorithms. In the application of IraNMF_HALS for real resting-state EEG data analysis, stable and interpretable features were extracted.

Conclusion: Based on the results of assessment, our recommendation is to use IraNMF_HALS, providing the most accurate and robust estimation.

Keywords: Nonnegative matrix factorization, Stability, Clustering, EEG

Background

Nonnegative matrix factorization (NMF) is a low-rank approximation method where both the data and the estimated low-rank factors are constrained to be nonnegative [1]. The method has been widely applied for data analysis and signal processing [2–4]. NMF was first introduced as positive matrix factorization (PMF) by Paatero and Tapper [5] and it was popularized from the application of face recognition by Lee and Seung [6]. Since it was proposed in the 1990s, many adaptive algorithms have been developed [2, 7]. Among numerous algorithms of NMF, multiplicative updating rule-based algorithm [8] is the most popular one. Besides, more efficient algorithms based on hierarchical alternating least squares (HALS) method and low-rank approximation were also proposed.



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

However, the performance of NMF in the implementation of EEG analysis has not been fully compared.

Electroencephalography (EEG) spectral studies deepen the understanding of the neurophysiological processes. It has been shown that the theta frequency band (3–8 Hz) reflects the general motor activity and the delta activity (0.5–3 Hz) has relationship with error-related processing [9, 10]. In working memory tasks, the global field theta synchronization significantly correlated with the blood oxygen level dependent (BOLD) signal [11]. It is also shown that gamma power (60–80 Hz) of EEG positively correlated with BOLD fluctuation and alpha (8–13 Hz) and beta (14–30 Hz) power negatively correlated with BOLD fluctuation [12].

Insomnia is a clinically common and frequent disease. The patient complains of poor sleep and accompanied by impaired day time functions, such as fatigue, lack of energy, cognitive decline, and emotional disorders [13, 14]. Currently, researchers found that the factors that induce insomnia mainly include age, personality, anxious personality, family and personal history of insomnia, genetic factors, etc. [15–17]. The main pathophysiological mechanism of patients with primary insomnia is excessive awakening [18–20]. In order to confirm the theory of excessive awakening, several studies were conducted on autonomic nerve activity, neuroendocrine, neuroimmunology, neuroelectrophysiology, and neuroimaging in patients with insomnia [21–23]. The spectral EEG was used to analyze the sleep electroencephalogram of patients with primary insomnia [24, 25], and it was found that the high-frequency waves of patients with insomnia were more active than normal people, including the period pre-sleep [26, 27], the wake–sleep transition period [28], and the sleep period [27, 29, 30]. Studies have shown that high-frequency brain waves are related to sensory processing, memory formation, and consciousness perception, and the active response of high-frequency brain waves increases cortical arousal levels [31]. Therefore, this study conducted a spectral analysis of the EEG of insomnia patients at rest to further explore the pathophysiological mechanism of primary insomnia.

NMF has been successfully used in the analysis of EEG spectral analysis. The components resulting from the NMF retain the spectral features hence they are interpretable from a physical or physiological point of view. For example, EEG large-scale network was constructed with NMF [32]. NMF was also used to estimate time and frequency components to predict epileptic seizures [33]. Both for static and dynamic functional connectivity, interpretable features also could be extracted with NMF [34, 35]. In the field of brain computer interface (BCI), NMF can obtain components with distinguishing features [36–38]. Several studies used features extracted with NMF to improve the discrimination of different participants [39–42]. Gruve et al. used NMF to extract the weight of the EEG channels to improve the accuracy of motor imagery detection [43] and classification of eye states [44]. With the successful application of NMF algorithm on EEG data, more and more NMF algorithms are constantly emerging [45–48]. In these studies, researchers used different NMF algorithms to analyze EEG data. However, the performance of NMF algorithms has not been fully compared.

Compared with independent component analysis (ICA) [49] and principal component analysis (PCA), NMF identified more meaningful and explainable components in practical application. Compared with PCA, the results of NMF have explicable

physical meaning and they are consistent with the brain intuitive perception [38]. Compared with ICA, some interesting components with higher distinguishing ratio were extracted by NMF [50].

The importance of stability of algorithms is self-evident for scalp EEG data analysis. Stability is a common problem of adaptive algorithms. Same to other blind sources separation (BSS) methods such as ICA, NMF also faces the same issue. Theoretically, just like the stability analysis of ICA [51, 52], the uniqueness of NMF could be verified by geometric interpretation [2, 49]. However, in the practical applications, algorithms may converge to local optima, which has been theoretically analysis with Lyapunov's first and second methods [53]. Thus, the stability of algorithms would be the most important aspect in the assessment of NMF algorithms.

In this study, the performance of four NMF algorithms was compared in terms of accuracy of estimation, stability and time complexity of algorithms with simulated data. The algorithm with excellent performance in simulation data was applied to real resting-state EEG data analysis. Based on stability analysis method, stable and interpretable features were extracted.

Simulation and results

Given a signal $H \in \mathbb{R}_+^{10 \times 1000}$ as component matrices shown in Fig. 1, we generated a mix matrix $W \in \mathbb{R}_+^{10 \times 64}$. Then we constructed $V^* = W^T H \in \mathbb{R}_+^{64 \times 1000}$ and $V = V^* + E$, where E denoted the independent noise and V denoted 64-channel simulated time-series EEG data. Signal noise ratio (SNR) was used to quantitative describe the quality of data. In this simulation, compared the performance of four NMF algorithms (NMF_MU, HALS, lraNMF_MU, lraNMF_HALS) were compared. The number of extracted components r was chosen as 10. The same algorithm was run 50 times. 500 components in total were clustered. All four algorithms were run in Matlab under the Window 10 system (Intel Xeon CPU 3.5 GHz and 32 GB of Random Access Memory (RAM)).

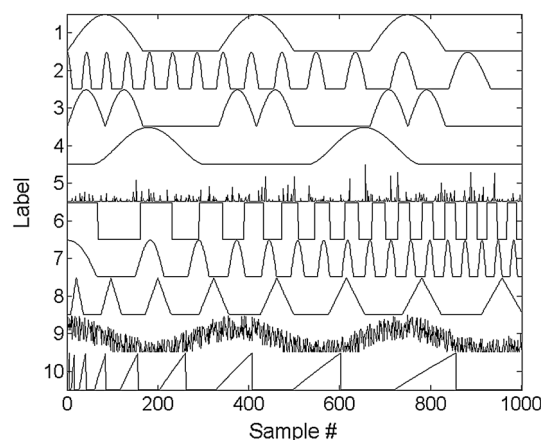


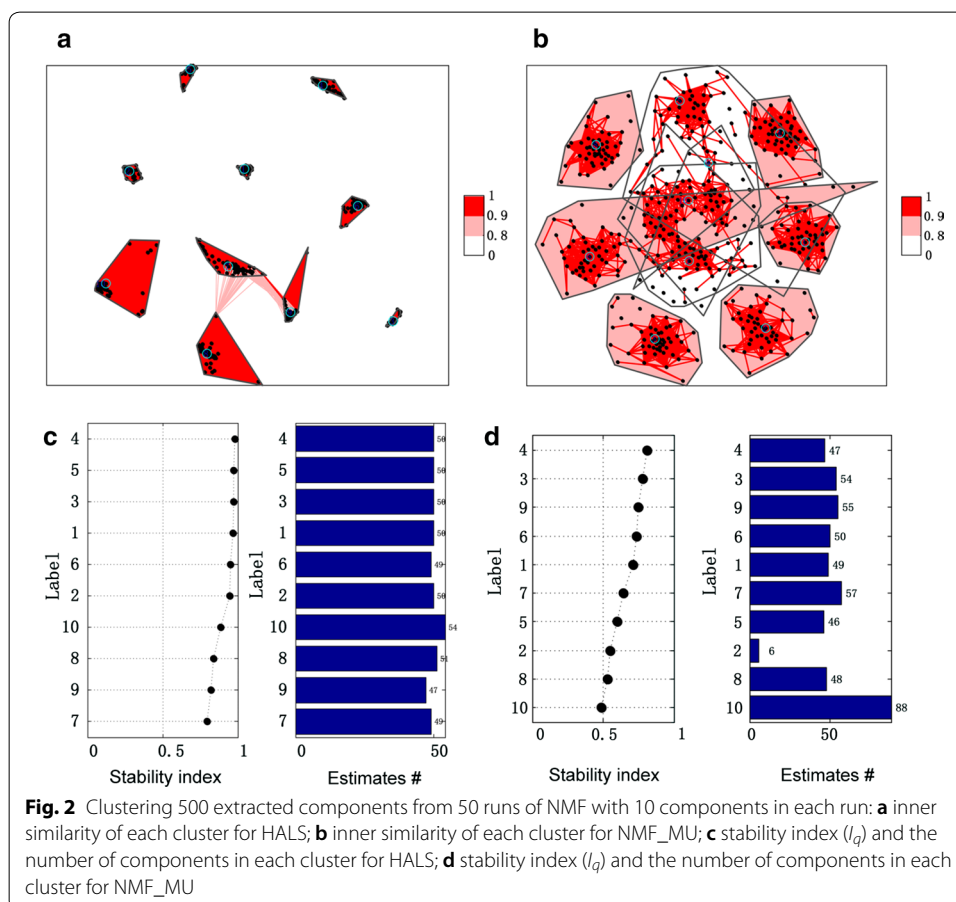
Fig. 1 Waveforms of 10 time-series for H in the NMF model

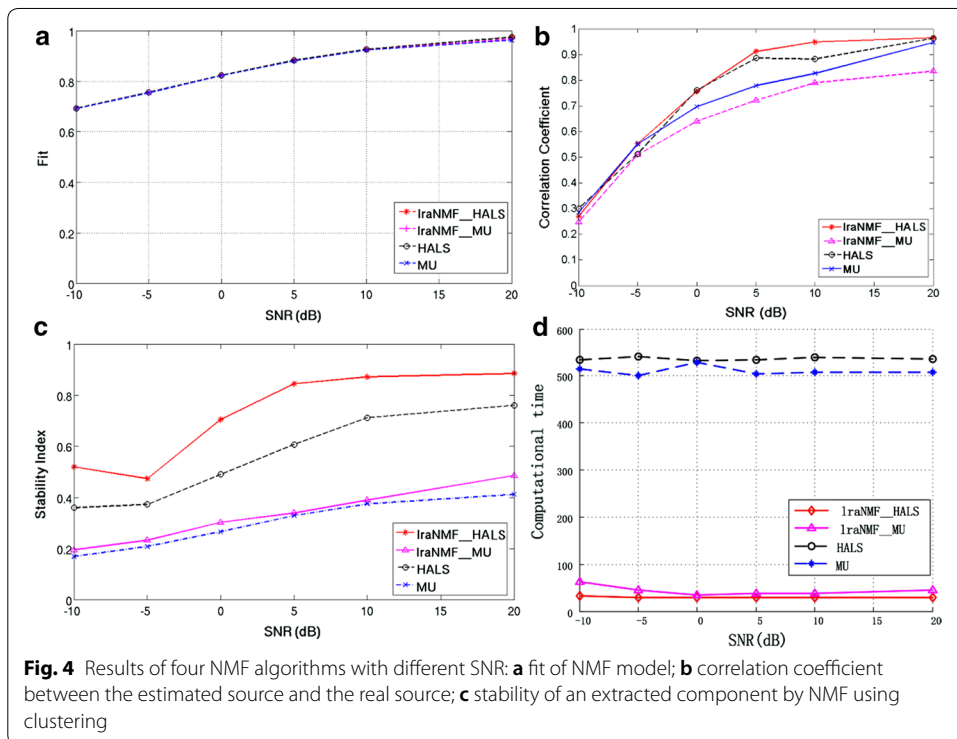
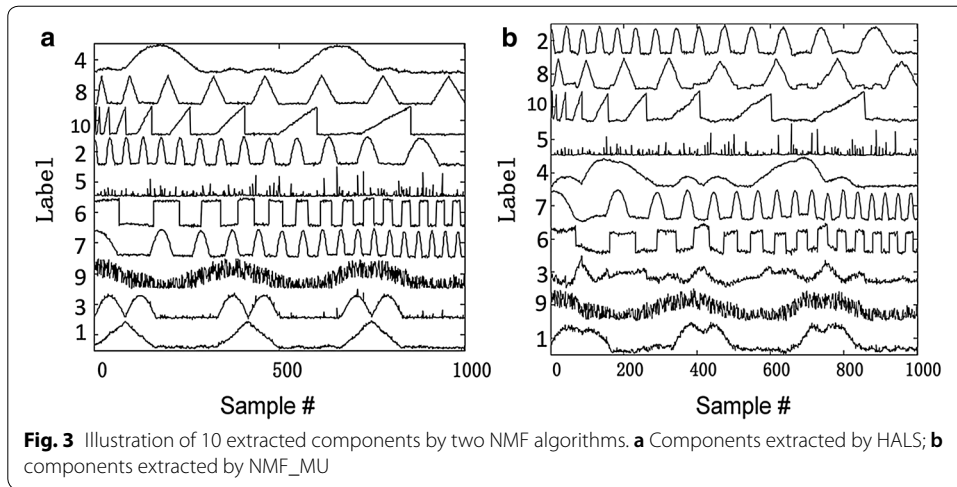
Simulated data with SNR = 20 dB

Simulated data with SNR of 20 dB was used for demonstrating the clustering related results. The 10 estimated sources and the clustering results of all 500 components estimated with HALS and NMF_MU algorithms are shown in Fig. 2. Ideally, the number of components in each cluster equals to the number of times that NMF was run (50 at here), under the condition that the number of clusters is same with the number of extracted components (10 at here). Obviously, HALS outperformed NMF_MU from the view of the stability of extracted components. In Fig. 2, the denser the cluster, the more stable of the components, extracted by an NMF algorithm, are. The purpose of the NMF algorithm is to accurately estimate the potential components in the data. Hence, it is necessary to check the accuracy of estimated sources. Here, readers are guided to have an intuitive feeling of the accuracy of estimation. The quantitative comparison in terms of estimation accuracy would be shown in the next section. The centroid of each cluster was selected as the component extracted by NMF as ICASSO [54] applied. Figure 3 shows the waveforms of 10 extracted components by two NMF algorithms. By visual inspection, HALS outperformed NMF_MU for estimating the sources.

Simulated data with multiple SNR

In order to test the performance of different NMF algorithms under different SNR, Fig. 4 illustrates the results with SNR ranging from $z-10$ dB to 20 dB with the step of



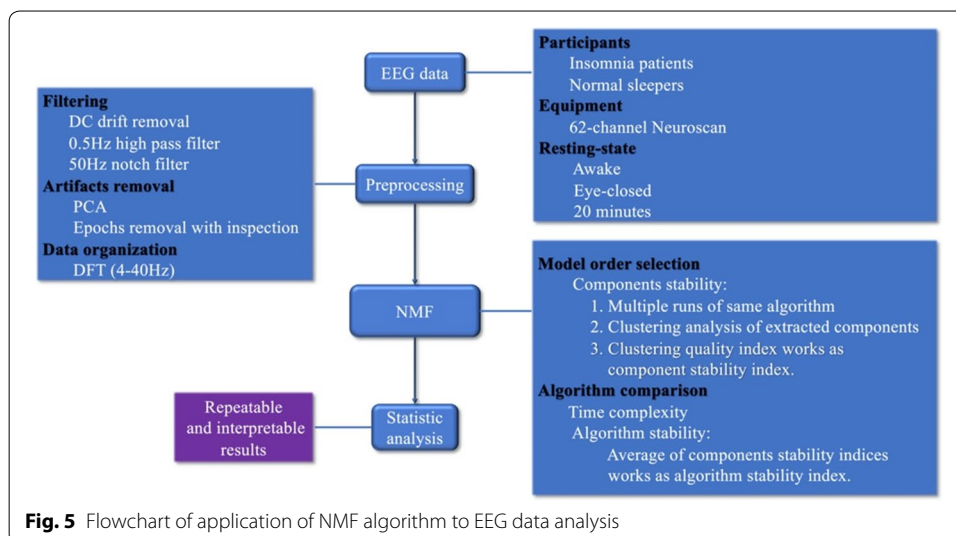


5 dB. Figure 4a shows that fits of four different NMF algorithms were very similar to one another. Figure 4b reveals the estimation accuracy of four algorithms. The estimation accuracy is measured with average correlation coefficient (Pearson correlation was used in this study) cross-ground truth and estimated component. The four algorithms yielded different correlation coefficients, which mean the accuracy of the estimation of four algorithms is different. The situation is a little bit different across SNR. When SNR is lower than -5 dB, the estimation accuracy is mainly dominated with signal quality, the accuracy of all four algorithms is very low. As the SNR increases, the accuracy of the estimated components also increases. When SNR large than 0 dB, the performance of

four algorithm begins to differ in terms of estimation accuracy. By comparison, it can be found that the estimated component of lraNMF_HALS is the most accurate, followed by HALS, followed by lraNMF_MU, and MU is the worst. Figure 4c presents the mean over 10 I_q of 10 components extracted by algorithm for each SNR and the I_q indicates the stability of the components over multiple runs of NMF. The stability of the extracted components of four NMF algorithms was very different. Figure 4d illustrates the computation time of four algorithms for the same data. The difference in computational time mainly is decided by the difference of iteration between algorithms and time complexity of the algorithms, which would be introduced in “Methods” section. The result of computation time shows that lraNMF_HALS needs less time to converge. Evidently, bigger correlation coefficient indicates better estimation of sources in Fig. 4b and the higher I_q implies better stability of extracted components in Fig. 4c. Based on comprehensive analysis of *fit*, stability, accuracy of estimation and time complexity, lraNMF_HALS outperformed the other three NMF algorithms. The algorithm would be used in the application of real EEG data analysis.

Real EEG results

In this section, the application of NMF algorithm to EEG data analysis was introduced. The stability-validating method was extended used not only for algorithm stability evaluation, but also used for determination of number of extracted components. The flowchart of data processing is illustrated in Fig. 5. After EEG data were collected, standard EEG preprocessing procedure was applied. It includes data filtering and artificial removal. Then the EEG data were transformed into frequency domain, which would be fed into NMF algorithm. Performance of several NMF algorithms was then evaluated in terms of algorithm stability and time complexity. Next, the model order of NMF algorithm was selected in terms of reproducibility of estimated components. Based on the above rigorous steps, repeatable and interpretable components were finally obtained. The specific process of each step will be described in the following content.



Data description

To evaluate the performance of four NMF algorithms, we tested it on EEG dataset which was recorded at The First Affiliated Hospital of Dalian Medical University. 16 primary insomnia patients and 17 normal sleepers were included in the study. Gender, age and education were matched between the two groups. All patients fit criteria for primary insomnia in DSM-IV. Patients and normal sleepers are required to complete the Pittsburgh Sleep Quality Index Scale (PSQI), Insomnia Severity Scale (ISI), Hyper-arousal Scale (HAS), Hamilton Anxiety Scale (HAMA) and Hamilton Depression Scale (HRSD-17) for clinical assessment. Continuous EEG recordings were digitally obtained using a 62-channel Neuroscan SynAmps2 brain electrical physiological instrument, EEG were recorded in a resting awake condition with the eyes closed for 20 min.

Data preprocessing

The EEG data were sampled at 1000 Hz for recording. Subsequently, the data were resampled at 500 Hz. DC drifts were removed using 0.5 Hz high-pass filter and 50 Hz notch filters were applied to minimize line noise artifacts. PCA was used to determine linear components associated with eye blinks and saccades. Epochs with strong eye movement or other movement artifacts were manually removed by inspection.

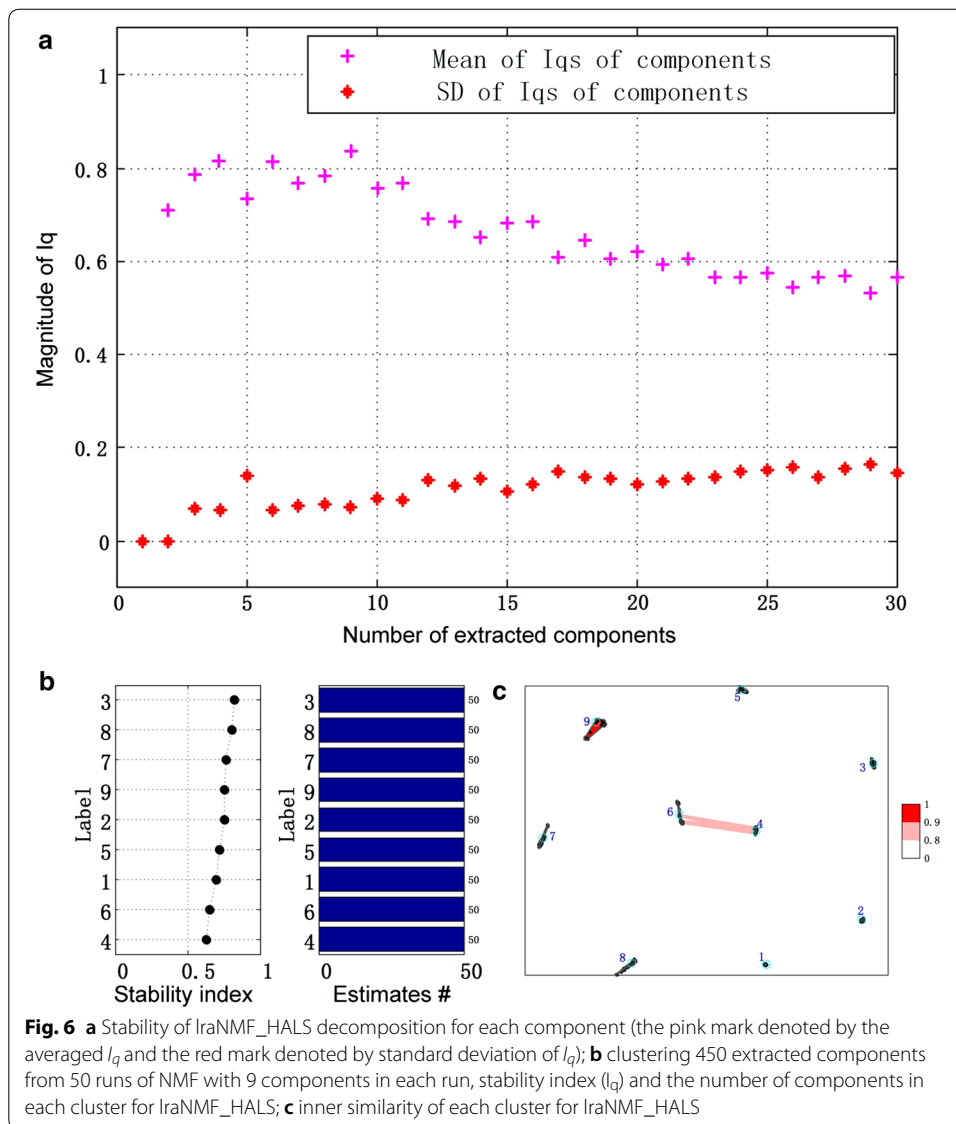
For all further analyses, EEG signals were processed with the method of discrete Fourier transform (DFT) which transform the time series into frequency domain. We cut 4–40 Hz of our interest frequency band. Thus, the sizes of preprocessed data are 73 (frequency bins) by 62 (channels) by 33 (2 groups with 16 primary insomnia patients and 17 health control). Then we transform the three-dimensional matrix from the EEG data of each subject into the matrix (channels \times frequency bins by samples, 62×2409 (73 by 33)) and used it as input to the space-by-frequency decomposition. Based on these processing, the elements in the matrix that would be fed into NMF are nonnegative.

Determination of the model order

Here, lraNMF_HALS was run 50 times with random initialization with the model order (number of extracted components) ranging from 2 to 30. Coefficient matrix was used to evaluate the stability of algorithms.

Figure 6a shows the stability (denoted by I_q) of lraNMF_HALS decomposition for each component. Indeed, each I_q in Fig. 6a is the averaged I_q for different components. It could be found that when the number of extracted components equals 9, the algorithm will be the most stable. As mentioned earlier, except the stability index of NMF decomposition, it is also necessary to check Fig. 6b, c and ensure whether the clustering result is good enough or not.

Different algorithms took different time finishing 50 times runs. NMF_MU took 1157.007s. HALS took 1173.178s. lraNMF_MU took 106.010s. lraNMF_HALS took 72.315s. Obviously, lraNMF_HALS took shortest time to complete the mission of decomposition. The stability indexes of four algorithms are also different. The stability index of NMF_MU is 0.760. The stability index of HALS is 0.717. The stability index of



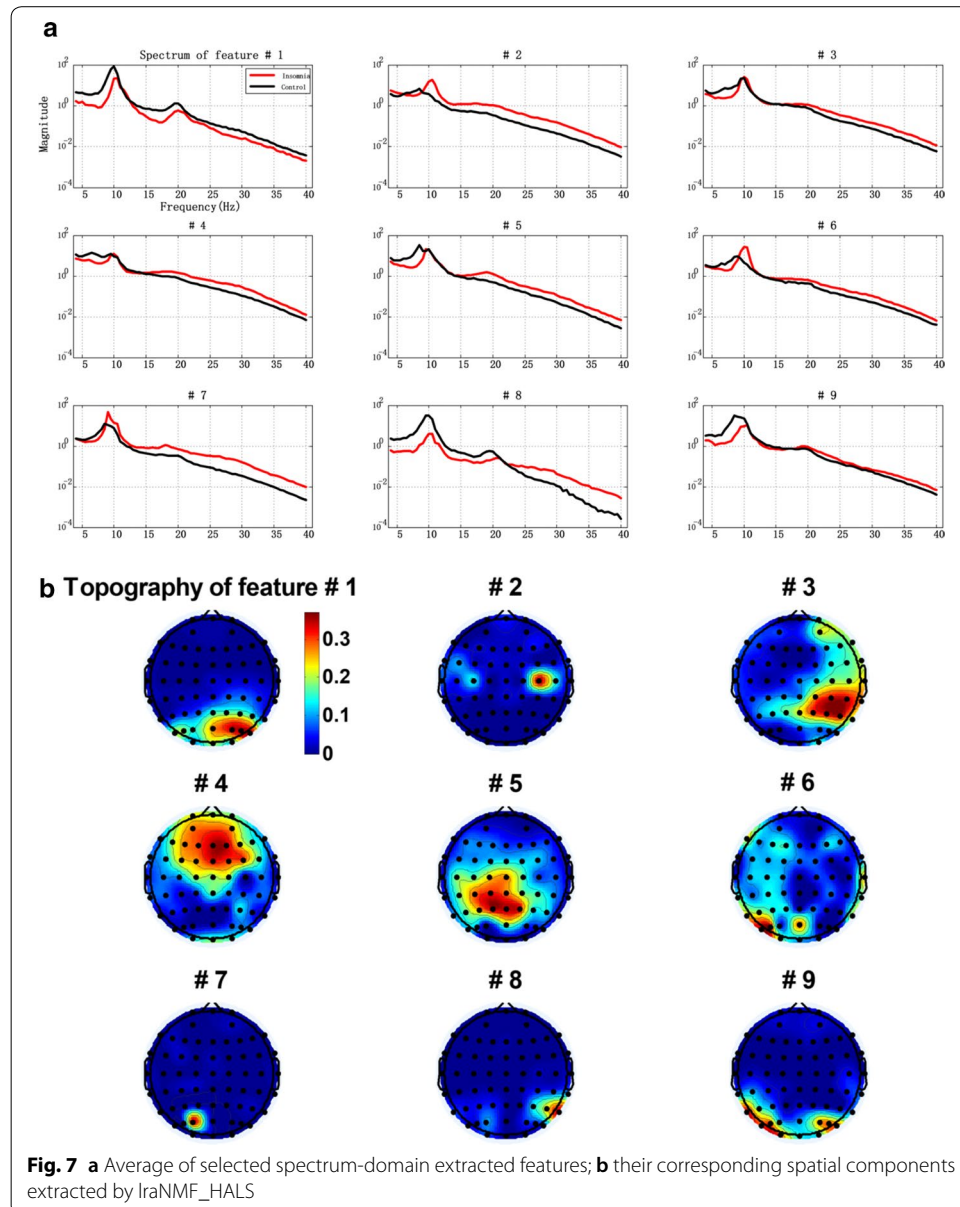
IraNMF_MU is 0.622. The stability index of IraNMF_HALS is 0.833. Obviously, as the result of simulation data, IraNMF_HALS is the most stable algorithm for the data. So, in the following subsection only the results of IraNMF_HALS would be introduced.

Feature selection

After a matrix is decomposed by IraNMF_HALS with 9 components, the features need to be further filtrated. Usually, the criteria of feature selection are determined by the purpose of further feature analysis. As we know that the collected EEG data consist of brain activities of no interest, brain activities of interest, and interference as well as noise. Therefore, even though artifacts are removed in the data preprocessing phase, the features of interest and features of no interest may also be extracted by NMF. Therefore, feature selection step is necessary to choose the features of brain activities of interest for further analysis [55].

The spectrum-domain features and corresponding spatial distribution are illustrated in Fig. 7. In Fig. 7a, the power spectrum values were log-transformed. Group difference was examined between insomnia and normal sleepers in absolute spectrum of 4–40 Hz for the component of interest. Insomnia group is noted by red line and normal sleeper group is noted by black line.

Furthermore, the nine features were statistic by two-way *t* test. We found that the feature#1 and #8 revealed a significant main effect in theta band ($P < 0.05$). From Fig. 7b, the #1 and #8 of brain map, the result indicated that the right occipital region has a main effect of group in theta band. The feature #4 also revealed that a significant main effect in low gamma band (30–40 Hz), $P < 0.05$. On the basis of the #4 of the brain maps in Fig. 7b, there was significant difference in low gamma band at frontal lobe. The feature

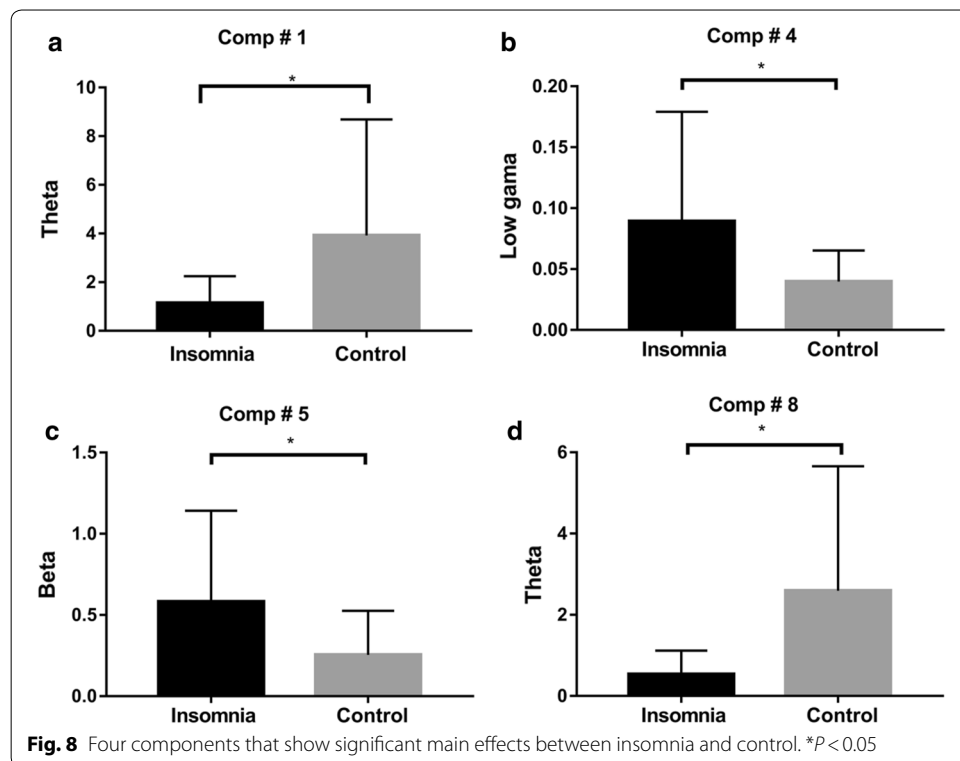


#5 also revealed that a significant main effect in beta band, $P < 0.05$. In accordance with the #5 in the brain map in Fig. 7a, there was the significant difference in beta band at parietal lobe. The significant differences are shown in Fig. 8. For the other five features, no main effect of group was observed.

Discussion

In this study, we provide a comparison of four NMF algorithms in terms of accuracy of estimation, stability (repeatability of the results) and time complexity of algorithms with simulated data. In the practical application of NMF algorithms, stability plays an important role, which was an emphasis in the comparison. A hierarchical clustering algorithm was implemented to evaluate the stability of NMF algorithms. In simulation, based on comprehensive analysis of fit, stability, accuracy of estimation and time complexity, hierarchical alternating least squares (HALS) low-rank NMF algorithm (IraNMF_HALS) outperformed the other three NMF algorithms. In the application of IraNMF_HALS for real resting-state EEG data analysis, stable and interpretable features were extracted.

Based on our assessment, IraNMF_HALS is the most stable algorithm. In order to guarantee the reliability of estimated components, the model order is selected in terms of algorithm stability. With these strategies stable components were reconstructed. These results of IraNMF_HALS for scalp resting-state EEG data analysis are consistent with the previous findings and also provide more reasonable results in support of pathological mechanisms of insomnia. Corsi-Cabrera et al. found that patients with insomnia exhibited significantly higher gamma power at frontoparietal [28]. This is consistent with the result of Comp#4. Szelenberger et al. pointed out that theta band in insomniacs



is lower than normal subjects, but beta band is higher than normal subjects [26]. This is in accordance with the result of the Comp#1 and Comp#5. The consistent results show that the proposed method is effective for spectrum analysis of EEG. In addition, we also found some results that were not found before, which provide support of pathological mechanisms of insomnia. The Comp#4 and Comp#5 with higher power can verify that the wakefulness of the insomniacs is higher than that of the normal subjects, which may lead to their insomnia. Based on the evaluation of algorithm stability, the repeatability of the results is guaranteed. This finding provides theoretical support for clinical treatment of insomnia.

When adaptive algorithm is applied in real-world application, more attention needs to be paid to its stability. Since initialization of adaptive algorithm may be different in different runs, it may converge to local optimum, which makes results different for different runs. It is vital for scientific research. The stability of algorithms could be evaluated with clustering analysis of components extracted from multiple runs of the same algorithm. The study of algorithm stability can not only quantitatively describe the reproducibility of the components, but also provide an effective criterion for the comparison of different algorithms and the selection of model order.

Conclusion

In this study, we proposed a method to compare different NMF algorithms so as to extract stable components. Specifically, we provide a comparison of four NMF algorithms in terms of accuracy of estimation, stability of algorithms and time complexity with simulated data. The performance of NMF algorithms also evaluated scalp resting-state EEG data analysis from aspects of stability and time complexity. In the application of lraNMF_HALS for real resting-state EEG data analysis, stable and interpretable features were extracted. From both simulation and real-world application, lraNMF_HALS outperformed the other three NMF algorithms in terms of algorithm stability and computational time. In simulation, the comparison is illustrated in Fig. 4. From simulation, we also find that lraNMF_HALS can estimate the most accurate potential components. In real-world application, in terms of computational time, lraNMF_HALS (72.315s) is faster than NMF_MU (1157.007s), HALS (1173.178s) and lraNMF_MU (106.010s) and in terms of algorithms stability, lraNMF_HALS (0.833) is more stable than NMF_MU (0.760), HALS (0.717) and lraNMF_MU (0.622). Based on the results of assessment, our recommendation is to use lraNMF_HALS, providing the most accurate and robust estimation as well as offering an intuitive interpretation. In terms of reproducibility of the final estimated components, this work provides a useful pipeline that applied NMF to spectrum analysis of EEG data. The pipeline includes selection of algorithm, selection of model order and stability analysis of estimated components. Based on the novel pipeline, components related to the pathology of insomnia were extracted and provide new ideas for clinical diagnosis and treatment of insomnia. For further study, the method is worthy of being applied on some other EEG spectrum data and identify more reasonable features. When an investigator prepares to use NMF for subsequent analysis, NMF performance needs to be considered in the experimental design to obtain stable and reliable results. Through this research, we found that different algorithms have a great impact on

the decomposition results. In the following research, it is necessary to propose an NMF algorithm specifically applied to EEG spectrum analysis.

Methods

NMF algorithms

For a given nonnegative data matrix $V \in \mathbb{R}^{m \times n}$ and the number of extracted components $r < \min[m, n]$, NMF attempts to find nonnegative matrices $W \in \mathbb{R}^{m \times r}$ and $H \in \mathbb{R}^{r \times n}$ which minimize the cost function as follows:

$$f_{W, H} = \frac{1}{2} \|V - WH\|_F^2, \tag{1}$$

where H and W are coefficient matrix and component matrix, respectively. Their product is a rank- r approximate estimation of V . In practical applications, the selection of r is critical and has great influence on the results. In this study, the model order, r , would be determined by checking the stability of the model with different values:

$$V_{m \times n} \approx W_{m \times r} H_{r \times n}, W \geq 0, H \geq 0. \tag{2}$$

To optimize Eq. (1), Lee and Seung suggested a very popular multiplicative update rule with computational complexity of $\mathcal{O}(mnr)$ [6]. The iterative formulas are as follows:

$$H \leftarrow H \cdot \frac{V^T W}{H W^T W}, \tag{3}$$

$$W \leftarrow W \cdot \frac{V H}{W H^T H}. \tag{4}$$

In 1998, Rasmus Bro proposed another set of updated formula [56] which is a column-wise method. And then it was extended with HALS algorithm and it can be computed in the complexity of $\mathcal{O}(mn)$ [2], where the columns of H and W are updated sequentially:

$$h_i \leftarrow \frac{1}{w_i^T w_i} [V_i^T w_i], \tag{5}$$

$$w_i \leftarrow \frac{1}{h_i^T h_i} [V_i h_i], \tag{6}$$

where $V_i = V - \sum_{j \neq i} w_j h_j^T$. w_j and h_j are the j th column of W and H , respectively. Only one column of W and H are updated in each iteration of HALS algorithm. Since r columns included in W and H , essentially, NMF_MU and HALS have equivalent space complexity and time complexity. However, in practical application, HALS is usually faster than NMF_MU [57].

Although HALS is faster than NMF_MU, the two algorithms still have a lot of room for improvement in computing speed. The common bottleneck of the computing speed is the size of matrix V . In the process of updating W and H , the large original data V will be used for many times. This issue causes not only slow convergence, but also great consuming of computer memory. To address the issue is to replace the large matrix with a

smaller one. Then the efficiency of HALS and NMF_MU could be improved. Motivated by the intuition, low-rank approximation (LRA) based NMF was proposed with computational complexity of $\mathcal{O}(nr^2)$ [57]:

$$\min_{\widehat{W}, \widehat{H}, W, H} F(\widehat{W}, \widehat{H}, W, H) = \min_{\widehat{W}, \widehat{H}, W, H} \left(\|V - \widehat{W} \widehat{H}^T\|_F^2 + \|\widehat{W} \widehat{H}^T - WH^T\|_F^2 \right), \tag{7}$$

where $W \in \mathbb{R}_+^{m \times r}$, $H \in \mathbb{R}_+^{r \times n}$, $\widehat{W} \in \mathbb{R}_+^{m \times l}$, $\widehat{H} \in \mathbb{R}_+^{l \times n}$, $l = pr \ll m$, and p is a small positive constant. In order to solve Eq. (7), the cost function $\min_{\widehat{W}, \widehat{H}, W, H} \|V - \widehat{W} \widehat{H}^T\|_F^2$ was used, where \widehat{W} and \widehat{H}^T are with the low-rank l , $l \ll m$. Then, optimize $\|\widehat{W} \widehat{H}^T - WH^T\|_F^2$ with fixed \widehat{W} and \widehat{H}^T .

The prototypical low-rank NMF algorithms originated by Guoxu Zhou and Andrzej Cichocki [57] are provided as follows:

$$H \leftarrow H * \frac{\widehat{H} \left(\widehat{W}^T W \right)}{H(W^T W)}, \tag{8}$$

$$W \leftarrow W * \frac{\widehat{W} \left(\widehat{H}^T H \right)}{W(H^T H)}. \tag{9}$$

This is lraNMF_MU that low-rank approximation-based multiplicative update (NMF_MU). The initialization of \widehat{W} and \widehat{H} come from PCA singular value decomposition (tSVD). In the first step, suppose that the optimal \widehat{W} and \widehat{H}^T , i.e., $V \approx_T \widehat{W} \widehat{H}^T$. Then above function would be used to optimize the cost function $\min \| \widehat{W} \widehat{H}^T - WH^T \|_F^2$. At first sight, there is no great difference between Eqs. (3) and (8). But note that the dimensions of \widehat{W} and \widehat{H}^T are much smaller than that of V . Under the present circumstances, $l = pr \ll m$, lraNMF_MU has much lower space and time complexity.

Similar to the HALS algorithm, let $V_i = \widehat{W} \widehat{H}^T - \sum_{j \neq i} w_j h_j^T$ and Eqs. (5, 6) become:

$$h_i \leftarrow \frac{1}{w_i^T w_i} \left[\widehat{H} \left(\widehat{W}^T w_i \right) - \overline{H}_1 \left(\overline{W}_1^T w_i \right) \right], \tag{10}$$

$$w_i \leftarrow \frac{1}{h_i^T h_i} \left[\widehat{W} \left(\widehat{H}^T h_i \right) - \overline{W}_1 \left(\overline{H}_1^T h_i \right) \right], \tag{11}$$

where $\overline{W}_1 \in \mathbb{R}^{m \times (r-1)}$ and $\overline{H}_1 \in \mathbb{R}^{(n-r) \times r}$ are the submatrices of W and H by removing their i th column. The low-rank approximation-based HALS is named as lraNMF_HALS, which could be computed in the complexity of $\mathcal{O}(nr)$ [57].

We also provide a summary of the comparison of different NMF algorithms that have been studied by the various published methods, as listed in Table 1, which may serve as a reference of method selection according to the available data types of the analyzers.

Table 1 Comparison of four NMF algorithms

Algorithms	Cost function	Iteration	Advantages
NMF_MU	$f(W, H) = \frac{1}{2} \ V - WH\ _F^2$	$H \leftarrow H * \frac{V^T W}{HW^T W}$ $W \leftarrow W * \frac{VH}{WH^T H}$	The original realization of NMF
HALS		$h_i \leftarrow \frac{1}{w_i^T w_i} [V_i^T w_i]$ $w_i \leftarrow \frac{1}{h_i^T h_i} [V_i h_i]$	Iteration column by column, faster than NMF_MU in practical application
IraNMF_MU	$F(\widehat{W}, \widehat{H}, W, H)$ $= \ V - \widehat{W}\widehat{H}^T\ _F^2$ $+ \ \widehat{W}\widehat{H}^T - WH^T\ _F^2$	$H \leftarrow H * \frac{\widehat{H}(\widehat{W}^T W)}{H(W^T W)}$ $W \leftarrow W * \frac{\widehat{W}(\widehat{H}^T H)}{W(H^T H)}$	Dimension reduction at the start of the algorithm. Fix the problem of the size of matrix to be decomposed. Faster than NMF_MU in theory and in practical application
IraNMF_HALS		$h_i \leftarrow \frac{1}{w_i^T w_i} \left[\widehat{H}(\widehat{W}^T w_i) - \overline{H_i}(\overline{W_i}^T w_i) \right]$ $w_i \leftarrow \frac{1}{h_i^T h_i} \left[\widehat{W}(\widehat{H}^T h_i) - \overline{W_i}(\overline{H_i}^T h_i) \right]$	Dimension reduction at the start of the algorithm. Faster than HALS in theory and practical application

Fitness of the algorithm

The *fit* is defined as follows:

$$fit(V, \widehat{V}) = 1 - \frac{\|V - \widehat{V}_F\|}{\|V_F\|}, \tag{12}$$

where \widehat{V} is a reconstructed version of V by Eq. (2). Apparently, $fit(V, \widehat{V}) = 1$ if and only if $\widehat{V} = V$ [57]. In the previous study, *fit* was used as the performance index of decomposition [2]. However, we find that even when the *fits* of two NMF decomposition are similar, the extracted components could be different. Hence the *fit* is not enough to evaluate the stability of the decomposition.

Hierarchical clustering

Hierarchical clustering is one of most popular clustering methods. In contrast to partitioned clustering, which directly decomposes the features into a set of disjoint clusters, the hierarchical clustering method is the process for transforming a proximity matrix into a nested partition, which can be graphically represented by a tree called dendrogram. Cutting the tree at different selected heights will provide a partitioning cluster at selected precision. So the precision is the one that is “tuned” by the cut [58]. This algorithm was applied to validate the stability of ICA components and it was named ICASSO [59].

Hierarchical clustering algorithms have two different types: agglomerative clustering (bottom-up) and divisive clustering (top-down). In this study, agglomerative clustering was used, and the dendrogram is formed from bottom to up. For this clustering method, at the first iteration, the number of clusters is same as the number of objects N . At the second iteration, the most similarity cluster will merge as a new cluster, so the number of clusters will become $N-1$. At the third iteration, the number of clusters will become $N-2$.

As the clustering goes forward, the number of clusters will become 1. By this way, the relation dendrogram of different objects could be established. In the stability algorithm, the cluster result is the dendrogram at the level that is the number of extracted features of NMF.

A conservative cluster quality index I_q in the previous study [59] was defined to reveal the compactness and isolation of a cluster. It is computed as the difference between the average intra-cluster similarities and average extra-cluster similarities:

$$I_q(C_m) = \frac{1}{|C_m|^2} \sum_{i,j \in C_m} \sigma_{ij} - \frac{1}{|C_m||C_{-m}|} \sum_{i \in C_m} \sum_{j \in C_{-m}} \sigma_{ij}, \quad (13)$$

where C_m means the features of m th cluster. σ_{ij} is the similarity of the i th and j th features. $C_{-m} = C - C_m$ represent the features that not belong to this cluster. Ideally, the same component from different runs will be clustered in the same cluster. The different components are in different clusters. The number of clusters equals to the number of extracted components. The number of features in each cluster is the number of runs of the algorithm. First term $\frac{1}{|C_m|^2} \sum_{i,j \in C_m} \sigma_{ij}$ is the average similarity of m th cluster (average intra-cluster similarities). This part is used to describe the similarity of the components in the same cluster. Second term $\frac{1}{|C_m||C_{-m}|} \sum_{i \in C_m} \sum_{j \in C_{-m}} \sigma_{ij}$ is the average similarity between the features that belong to m th cluster and not belong to m th cluster (average extra-cluster similarities). This part is used to describe the dissimilarity of the components in different clusters. The range of I_q is from 0 to 1. The value closer to 1, the higher stability it is.

Evaluation of stability of NMF algorithms

Given a non-negative data matrix $V \in \mathbb{R}_+^{m \times n}$, we used an NMF algorithm to decompose the data. In order to evaluate the stability and reliability of components decomposed by NMF, three steps are used, which is introduced in our previous study [60]. Three steps are as follows:

Step 1: An NMF algorithm was run K times. Each time with random initialization.

Step 2: All extracted components were clustered. According to their mutual similarities, agglomerative clustering was used.

Step 3: The centroid of each cluster was selected as the component extracted by NMF. The cluster quality index I_q works as stability index for each component.

Abbreviations

NMF: Nonnegative matrix factorization; EEG: Electroencephalography; PMF: Positive matrix factorization; BOLD: Blood oxygen level dependent; ICA: Independent component analysis; PCA: Principal component analysis; BCI: Brain computer interface; BSS: Blind sources separation; NMF_MU: NMF-based multiplicative update; HALS: Hierarchical alternating least squares; lraNMF_MU: Low-rank approximation-based multiplicative update; lraNMF_HALS: Hierarchical alternating least squares (HALS) low-rank NMF algorithm.

Acknowledgements

The authors are grateful to all the study participants.

Authors' contributions

Conceptualization: FC, JX and YC. Methodology: GH. Software: TZ and RM. Validation: GH. Formal analysis: GH. Investigation: JX, YC and SL. Resources: JX, YC and SL. Data curation: SL. Writing—original draft preparation: GH and TZ. Writing—review and editing: GH. Visualization: TZ. Supervision: FC. Project administration: FC, JX and YC. Funding acquisition: FC, JX and YC. All authors read and approved the final manuscript.

Funding

This work was supported by National Natural Science Foundation of China (Grant Nos. 91748105 & 81471742) and the Fundamental Research Funds for the Central Universities [DUT2019] in Dalian University of Technology in China. This work was also supported by China Scholarship Council (No. 201806060038) and Natural Science Foundation of Liaoning Province (2019-MS-099).

Availability of data and materials

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participants

This study was approved by the Ethics Committee of Dalian Medical University in accordance with the Declaration of Helsinki. All participants provided written informed consent following an explanation of the nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interest.

Author details

¹ School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China. ² Center for Cognition and Neuroergonomics, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University at Zhuhai, Zhuhai, China. ³ Affiliated Zhongshan Hospital of Dalian University, Dalian, China. ⁴ Department of Neurology and Psychiatry, First Affiliated Hospital, Dalian Medical University, Dalian, China. ⁵ School of Artificial Intelligence, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China. ⁶ Key Laboratory of Integrated Circuit and Biomedical Electronic System, Dalian University of Technology, Dalian, Liaoning, China. ⁷ Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland.

Received: 28 January 2020 Accepted: 9 June 2020

Published online: 31 July 2020

References

1. Kim J, He Y, Park H. Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework. *J Glob Optim.* 2014;58:285–319.
2. Zhou G, Cichocki A, Zhao Q, Xie S. Nonnegative matrix and tensor factorizations: an algorithmic perspective. *IEEE Signal Process Mag.* 2014;31:54–65.
3. Cichocki A, Zdunek R, Phan AH, Amari S. Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation. Hoboken: Wiley; 2009.
4. Li X. Signal processing in neuroscience. Singapore: Springer; 2016.
5. Paatero P, Tapper U. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics.* 1994;5:111–26.
6. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature.* 1999;401:788–91.
7. Wang YX, Zhang YJ. Nonnegative matrix factorization: a comprehensive review. *IEEE Trans Knowl Data Eng.* IEEE. 2013;25:1336–53.
8. Lee DD, Seung HS. Algorithms for non-negative matrix factorization. *Adv Neural Inf Process Syst.* 2001;62:556–62.
9. Beste C, Saft C, Yordanova J, Andrich J, Gold R, Falkenstein M, et al. Functional compensation or pathology in cortico-subcortical interactions in preclinical Huntington's disease? *Neuropsychologia.* 2007;45:2922–30.
10. Yordanova J, Falkenstein M, Hohnsbein J, Koley V. Parallel systems of error processing in the brain. *Neuroimage.* 2004;22:590–602.
11. Michels L, Lüchinger R, Koenig T, Martin E, Brandeis D. Developmental changes of BOLD signal correlations with global human EEG power and synchronization during working memory. *PLoS ONE.* 2012;7:e39447.
12. Scheeringa R, Fries P, Petersson KM, Oostenveld R, Grothe I, Norris DG, et al. Neuronal dynamics underlying high- and low-frequency EEG oscillations contribute independently to the human BOLD signal. *Neuron.* 2011;69:572–83.
13. Shekleton JA, Rogers NL, Rajaratnam SMW. Searching for the daytime impairments of primary insomnia. *Sleep Med Rev.* Elsevier Ltd. 2010;14:47–60.
14. Riemann D, Baglioni C, Bassetti C, Bjorvatn B, Dolenc Groselj L, Ellis JG, et al. European guideline for the diagnosis and treatment of insomnia. *J Sleep Res.* 2017;26:675–700.
15. Deuschle M, Schredl M, Schilling C, Wüst S, Frank J, Witt SH, et al. Association between a serotonin transporter length polymorphism and primary insomnia. *Sleep.* 2010;33:343–7.
16. de Zambotti M, Goldstone A, Colrain IM, Baker FC. Insomnia disorder in adolescence: diagnosis, impact, and treatment. *Sleep Med Rev.* 2018;39:12–24.
17. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet.* 2019;51:394–403.
18. Bastien CH. Insomnia: neurophysiological and neuropsychological approaches. *Neuropsychol Rev.* 2011;21:22–40.
19. Araújo T, Jarrin DC, Leanza Y, Vallières A, Morin CM. Qualitative studies of insomnia: current state of knowledge in the field. *Sleep Med Rev.* 2017;31:58–69.

20. Jarrin DC, Alvaro PK, Bouchard MA, Jarrin SD, Drake CL, Morin CM. Insomnia and hypertension: a systematic review. *Sleep Med Rev.* 2018;41:3–38.
21. Riemann D, Spiegelhalter K, Feige B, Voderholzer U, Berger M, Perlis M, et al. The hyperarousal model of insomnia: a review of the concept and its evidence. *Sleep Med Rev.* 2010;14:19–31.
22. Bonnet MH, Arand DL. Hyperarousal and insomnia: state of the science. *Sleep Med Rev.* 2010;14:9–15.
23. Colombo MA, Wei Y, Ramautar JR, Linkenkaer-Hansen K, Tagliazucchi E, Van Someren EJW. More severe insomnia complaints in people with stronger long-range temporal correlations in wake resting-state EEG. *Front Physiol.* 2016;7:1–11.
24. Svtnik V, Snyder ES, Ma J, Tao P, Lines C, Herring WJ. EEG spectral analysis of NREM sleep in a large sample of patients with insomnia and good sleepers: effects of age, sex and part of the night. *J Sleep Res.* 2017;26:92–104.
25. Miller CB, Bartlett DJ, Mullins AE, Dodds KL, Gordon CJ, Kyle SD, et al. Clusters of insomnia disorder: an exploratory cluster analysis of objective sleep parameters reveals differences in neurocognitive functioning, quantitative EEG, and heart rate variability. *Sleep.* 2016;39:1993–2004.
26. Wolyńczyk-Gmaj D, Szelenberger W. Waking EEG in primary insomnia. *Acta Neurobiol Exp.* 2011;71:387–92.
27. Corsi-Cabrera M, Rojas-Ramos OA, del Río-Portilla Y. Waking EEG signs of non-restoring sleep in primary insomnia patients. *Clin Neurophysiol.* 2016;127:1813–21.
28. Corsi-Cabrera M, Figueredo-Rodríguez P, del Río-Portilla Y, Sánchez-Romero J, Galán L, Bosch-Bayard J. Enhanced frontoparietal synchronized activation during the wake-sleep transition in patients with primary insomnia. *Sleep.* 2012;35:501–11.
29. Spiegelhalter K, Regen W, Feige B, Holz J, Piosczyk H, Baglioni C, et al. Increased EEG sigma and beta power during NREM sleep in primary insomnia. *Biol Psychol.* 2012;91:329–33.
30. Cervena K, Espa F, Perogamvros L, Perrig S, Merica H, Ibanez V. Spectral analysis of the sleep onset period in primary insomnia. *Clin Neurophysiol.* 2014;125:979–87.
31. Buysse DJ, Germain A, Hall M, Monk TH, Nofzinger EA. A neurobiological model of insomnia. *Drug Discov Today Dis Model.* 2011;8:129–37.
32. Yi C, Chen C, Si Y, Li F, Zhang T, Liao Y, et al. Constructing large-scale cortical brain networks from scalp EEG with Bayesian nonnegative matrix factorization. *Neural Netw.* 2020;125:338–48.
33. Stojanović O, Kuhlmann L, Pipa G. Predicting epileptic seizures using nonnegative matrix factorization. *PLoS ONE.* 2020;15:1–13.
34. Zhou T, Kang J, Cong F, Li X. Stability-driven non-negative matrix factorization-based approach for extracting dynamic network from resting-state EEG. *Neurocomputing.* 2020;134:1–9.
35. Dukic S, McMackin R, Buxo T, Fasano A, Chipika R, Pinto-Grau M, et al. Patterned functional network disruption in amyotrophic lateral sclerosis. *Hum Brain Mapp.* 2019;40:4827–42.
36. Lee H, Cichocki A, Choi S. Kernel nonnegative matrix factorization for spectral EEG feature extraction. *Neurocomputing.* 2009;72:3182–90.
37. Lu N, Li T, Pan J, Ren X, Feng Z, Miao H. Structure constrained semi-nonnegative matrix factorization for EEG-based motor imagery classification. *Comput Biol Med.* 2015;60:32–9.
38. Liu W, Zheng N, You Q. Nonnegative matrix factorization and its applications in pattern recognition. *Chinese Sci Bull.* 2006;51:7–18.
39. Lu N, Yin T. Motor imagery classification via combinatorial decomposition of ERP and ERSP using sparse nonnegative matrix factorization. *J Neurosci Methods.* 2015;249:41–9.
40. Ghoraani B. Class-specific discriminant time-frequency analysis using novel jointly learnt non-negative matrix factorization. *EURASIP J Adv Signal Process.* 2016;2016:95.
41. Lee H, Yoo J, Choi S. Semi-supervised nonnegative matrix factorization. *IEEE Signal Process Lett.* 2010;17:4–7.
42. Cichocki A, Lee H, Kim Y-D, Choi S. Non-negative matrix factorization with α -divergence. *Pattern Recognit Lett.* 2008;29:1433–40.
43. Gurve D, Delisle-Rodríguez D, Romero-Laiseca M, Cardoso V, Loterio F, Bastos T, Krishnan S. Subject-specific EEG channel selection using non-negative matrix factorization for lower-limb motor imagery recognition. *J Neural Eng.* 2020;17:026029.
44. Gurve D, Krishnan S. Deep learning of EEG time-frequency representations for identifying eye states. *Adv Data Sci Adapt Anal.* 2018;10:1840006.
45. Lee H, Cichocki A, Choi S. Nonnegative matrix factorization for motor imagery EEG classification. *Lect Notes Comput Sci.* 2006;4132:250–9.
46. Phan AH, Cichocki A. Seeking an appropriate alternative least squares algorithm for nonnegative tensor factorizations: a novel recursive solution for nonnegative quadratic programming and NTF. *Neural Comput Appl.* 2012;21:623–37.
47. Razavipour F, Boostani R, Kouchaki S, Afrasiabi S. Comparative application of non-negative decomposition methods in classifying fatigue and non-fatigue states. *Arab J Sci Eng.* 2014;39:7049–58.
48. Delis I, Onken A, Schyns PG, Panzeri S, Philiastides MG. Space-by-time decomposition for single-trial decoding of M/EEG activity. *Neuroimage.* 2016;133:504–15.
49. Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. *Neural Netw.* 2000;13:411–30.
50. Cong F, Zhang Z, Kalyakin I, Huttunen-Scott T, Lyytinen H, Ristaniemi T. Non-negative matrix factorization vs. FastICA on mismatch negativity of children. *Int Jt Conf Neural Networks.* 2009; 586–90. .
51. Hu G, Zhang Q, Waters AB, Li H, Zhang C, Wu J, et al. Tensor clustering on outer-product of coefficient and component matrices of independent component analysis for reliable functional magnetic resonance imaging data decomposition. *J Neurosci Methods.* 2019;325:108359.
52. Zhang Q, Hu G, Tian L, Ristaniemi T, Huili W, Chen H, et al. Examining stability of independent component analysis based on coefficient and component matrices for voxel-based morphometry of structural magnetic resonance imaging. *Cogn Neurodyn.* 2018;12:461–70.
53. Bateau R, Bertin N, Vincent E. Stability analysis of multiplicative update algorithms and application to nonnegative matrix factorization. *IEEE Trans Neural Networks.* 2010;21:1869–81.
54. Himberg J, Hyvärinen A. ICASSO: Software for investigating the reliability of ICA estimates by clustering and visualization. *Neural Netw Signal Process Proc IEEE Work.* 2003;2003:259–68.

55. Cong F, Lin Q-H, Kuang L-D, Gong X-F, Astikainen P, Ristaniemi T. Tensor decomposition of EEG signals: a brief review. *J Neurosci Methods*. 2015;248:59–69.
56. Bro R. Multi-way Analysis in the Food Industry. Algorithms, and Applications. Academish proefschrift. Models: Dinamarca; 1998.
57. Zhou G, Cichocki A, Xie S. Fast nonnegative matrix/tensor factorization based on low-rank approximation. *IEEE Trans Signal Process*. 2012;60:2928–40.
58. Abu-Jamous B, Fa R, Nandi AK. Integrative cluster analysis in bioinformatics. Hoboken: Wiley; 2015.
59. Himberg J, Hyvärinen A, Esposito F. Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage*. 2004;22:1214–22.
60. Zhou T, Hu G, Mahini R, Gong X, Lin Q, Cong F. Validating stability of components extracted by nonnegative matrix factorization via clustering. 2016; 8–11.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

