ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-6654-0540-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICASSP43922.2022.9746404

AN ACCELERATED RANK-(L,L,1,1) BLOCK TERM DECOMPOSITION OF MULTI-SUBJECT FMRI DATA UNDER SPATIAL ORTHONORMALITY CONSTRAINT

Li-Dan Kuang[1], Biao Wang[1], Qiu-Hua Lin[2], Hao-Peng Zhang[1], Jianming Zhang[1], Wenjun Li[1], Feng Li[1], and Vince D. Calhoun[3]

[1] School of Computer and Communication Engineering,
Changsha University of Science and Technology, Changsha 410114, China
kuangld@csust.edu.cn
[2] School of Information and Communication Engineering,
Dalian University of Technology, Dalian 116024, China
[3] Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS),
Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA USA

ABSTRACT

The decomposition of multi-subject fMRI data using rank-(L,L,1,1) block term decomposition (BTD) can preserve higher-way data structure and is more robust to noise effects by decomposing shared spatial maps (SMs) into a product of two rank-L loading matrices. However, since the number of whole-brain voxels is very large and rank L is larger than 1, the rank-(L,L,1,1) BTD requires high computation and memory. Therefore, we propose an accelerated rank-(L,L,1,1) BTD algorithm based upon the method of alternating least squares (ALS). We speed up updates of loading matrices by reducing fMRI data into subspaces, and add an orthonormality constraint on shared SMs to improve the performance. Moreover, we evaluate the rank-L effect on the proposed method for actual task-related fMRI data. The proposed method shows better performance when L=35. Meanwhile, experimental comparison results verify that the proposed method largely reduced (17.36 times) computation time compared to ALS while also providing satisfying separation performance.

Index Terms— rank-(L,L,1,1) BTD; ALS; acceleration; fMRI; tensor decomposition.

1. INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a noninvasive technique for studying brain activity and by measuring blood-oxygenation-level dependent (BOLD) signal to get fMRI data [1]. The multi-subject fMRI data inherently possess multi-way nature including 3-dimensional space, time, and subject modes. The majority of studies, such as group independence component analysis (ICA) [2], canonical polyadic decomposition (CPD) [3], and independence vector analysis (IVA) [4], unfold the 3D spatial brain images into a vector and select the in-brain voxels. More specifically, group ICA temporally concatenates fMRI data of each subject to compose a matrix with space and time modes, and CPD and IVA treat multi-subject fMRI data as a three-way tensor with space, time and subject modes.

However, such approach does not leverage the 3D spatial structure of fMRI. Block term decomposition (BTD) has been efficiently applied to four-way multi-subject fMRI data [5, 6]. Instead of vectorization of spatial mode, BTD unfolds the 3D spatial brain image into a matrix by concatenating different frontal slices [6]. More specifically, the rank-(L_r,L_r,1,1) model, one of the simplified and typical BTD model, decomposes the 4-way multi-subject fMRI data into N rank-L_r tensors (N is the number of components). Moreover, each tensor is the out-product of a shared spatial map (SM) matrix which is the product of two rank-L_r loading matrices, a time course (TC) vector and a subject intensity vector. The uniqueness of rank-(L_r,L_r,1,1) BTD has been verified in [5-7]. Moreover, rank-(L_r,L_r,1,1) BTD is more robust than CPD especially in cases of processing strong noise [5, 6]. However, these following two problems have been aroused when applying rank-(L_r,L_r,1,1) BTD to multi-subject fMRI data.

On one hand, since brain images for rank-(L_r,L_r,1,1) BTD include both in- and out-brain voxels, the number of voxels of brain images are much larger than the product of the number of time points and subjects (e.g., 120,060 vs. 2,640 in this paper). Moreover, the rank of each shared SM component L_r is generally larger than 1. Therefore, rank-(L_r,L_r,1,1) BTD is more computationally demanding than CPD. The classical alternating least squares (ALS) has been widely used for CPD and BTD. Some accelerated methods based on ALS have been proposed to speed up CPD, such as ALS with compression [8, 9], with line search methods [10, 11], using mode reduction and Khatri-Rao product projection procedure [12], through tensor reshaping [13], by partitioning loading matrices into blocks [14], via a matrix

polynomial predictive model [15], and so on. Recently, a rank-(L,L,1) BTD based on an alternating hierarchical iteratively reweighted least squares exhibiting fast convergence was proposed [16].

Once challenge is the selection of rank is a NP-hard problem. For ICA, IVA, and CPD of multi-subject fMRI data, we can use model order selection such as minimum description length (MDL) [17], rank estimation tensor nuclear norm [18] and tensor rank network [19], to estimate the number of components. Besides the number of components, rank-(L_r,L_r,1,1) BTD also needs to estimate the L_r values. Chatzichristos et al. proposed the maximum L_r estimation method determined by the number of columns containing at least one effective voxel after performing Z-test and Bonferroni correction. Han et al. treated rank-(L,L,1) BTD as CPD and estimated the rank L by group sparsity [20]. For simplicity, we also assume L_r same for all components, *i.e.*, rank-(L,L,1,1) BTD.

In consider of these above two problems, we here firstly propose a novel accelerated ALS method for rank-(L,L,1,1) BTD applied to multi-subject fMRI data. The proposed method accelerates the updates of two loading matrices of shared SMs by reducing the fMRI data in the subspace of Khatri-Rao product of shared TCs and subject intensities, and accelerates the updates of shared TCs and subject intensities by reducing the fMRI data in the subspace of shared SMs. Besides, since the orthogonality between SMs has been verified that can improve the separation performance the CPD when applied to multi-subject fMRI data [3, 21-22], we incorporate the spatial orthonormality constraint on the shared SMs. We secondly comprehensively test the separation performance effect on the rank L_r by varying from 5 to 45 based upon the actual task-related fMRI data experiments. We also compare the proposed method with ALS and nonlinear least squares (NLS) methods of rank-(L_r,L_r,1,1) BTD. The rest of paper is organized as follows. Section 2 descripts the proposed accelerated ALS method for rank-(L_r,L_r,1,1) BTD with spatial orthonormality constraint. Section 3 narrates the experimental methods. Section 4 concludes the results of actual task-related fMRI data. Section 5 has the conclusions.

2. THE PROPOSED METHOD

Notations: Scalar, vector, matrix and tensor are expressed as italic lowercase letters (e.g., x), bold lowercase letters (e.g., $\mathbf{x}$), bold uppercase letters (e.g., $\mathbf{X}$), and bold uppercase calligraphic letters (e.g., $\boldsymbol{\mathcal{X}}$), respectively. The transpose, conjugate and pseudo-inverse are denoted by "T", "*" and "†" respectively. "$\|\cdot\|$" is a ℓ_2 norm function. "$\circ$" and "$\otimes$" are expressed as outer product and Kronecker product. We here use two kinds of Khatri-Rao product. First, let $\mathbf{A}=[\mathbf{A}_1,\cdots,\mathbf{A}_N]$ and $\mathbf{B}=[\mathbf{B}_1,\cdots,\mathbf{B}_N]$ denote two partitioned matrices, and the Khatri-Rao product of these two matrices is $\mathbf{A}\odot\mathbf{B}=[\mathbf{A}_1\otimes\mathbf{B}_1,\cdots,\mathbf{A}_N\otimes\mathbf{B}_N]$ [23]. Second, when $\mathbf{A}=[\mathbf{a}_1,\cdots,\mathbf{a}_N]$ and $\mathbf{B}=[\mathbf{b}_1,\cdots,\mathbf{b}_N]$ are partitioned into vectors, we calculate the Khatri-Rao product of these two matrices as $\mathbf{A}\odot_c\mathbf{B}=[\mathbf{a}_1\otimes\mathbf{b}_1,\cdots,\ \mathbf{a}_N\otimes\mathbf{b}_N]$ [23].

2.1. The model and cost function

Assume the 3D spatial brain image with size $I_x\times I_y\times I_z$ can be reshaped as 2D spatial brain image with size $I_x\times I_{yz}$ ($I_{yz}=I_y\times I_z$), J and M denote the number of scans and the number of subjects. As such, the four-way multi-subject whole-brain fMRI data $\boldsymbol{\mathcal{X}}\in\mathbb{R}^{I_x\times I_{yz}\times J\times M}$ can be treated a sum of N rank-(L,L,1,1) terms which satisfy:

$$\boldsymbol{\mathcal{X}}=\sum_{n=1}^{N}\mathbf{S}_n\circ\mathbf{c}_n\circ\mathbf{d}_n=\sum_{n=1}^{N}\left(\mathbf{A}_n\mathbf{B}_n^T\right)\circ\mathbf{c}_n\circ\mathbf{d}_n\ , \qquad (1)$$

where $\mathbf{S}_n=\mathbf{A}_n\mathbf{B}_n^T\in\mathbb{R}^{I_x\times I_{yz}}$ denotes shared SM component which is the product of rank-L matrix $\mathbf{A}_n\in\mathbb{R}^{I_x\times L}$ and rank-L matrix $\mathbf{B}_n\in\mathbb{R}^{I_{yz}\times L}$ ($n=1,\ldots,N$). Notice that the rank-(L,L,1,1) model assumes all shared SM components have same ranks L for simplicity. $\mathbf{C}=[\mathbf{c}_1,\cdots,\mathbf{c}_N]\in\mathbb{R}^{J\times N}$ and $\mathbf{D}=[\mathbf{d}_1,\cdots,\mathbf{d}_N]\in\mathbb{R}^{M\times N}$ correspond to shared TCs and subject-specific intensities, respectively.

Since the orthogonality between SMs has been shown improved performance of CPD of multi-subject fMRI data [3, 21-22], we aim to minimize the error of rank-(L,L,1,1) BTD model under spatial orthonormality constraint:

$$\begin{aligned}&\min_{\mathbf{A}_n,\mathbf{B}_n,\mathbf{c}_n,\mathbf{d}_n}\left\|\boldsymbol{\mathcal{X}}-\sum_{n=1}^{N}\left(\mathbf{A}_n\mathbf{B}_n^T\right)\circ\mathbf{c}_n\circ\mathbf{d}_n\right\|,\\ &s.t.\ \ \mathbf{S}^H\mathbf{S}\ =\mathbf{I},\end{aligned} \qquad (2)$$

where $\mathbf{S}=[\mathrm{vec}(\mathbf{A}_1\mathbf{B}_1^T),\cdots,\mathrm{vec}(\mathbf{A}_N\mathbf{B}_N^T)]\in\mathbb{R}^{I_{xyz}\times N}$ are shared SMs, function vec(·) denotes vectorization of a matrix, $I_{xyz}=I_xI_{yz}$, and $\mathbf{I}$ is the identity matrix.

2.2. The accelerated ALS updating rules

Since the size of whole-brain fMRI image are much larger than in-brain fMRI image (e.g., 120,060 vs. 59,610 in this paper), it requires more computation time and memory to calculate the loading matrices for ALS of rank-(L,L,1,1) BTD. Inspired by the accelerated ALS of CPD [9, 24], we here propose an accelerated ALS for rank-(L,L,1,1) BTD.

Let $\mathbf{A}=[\mathbf{A}_1,\cdots,\mathbf{A}_N]\in\mathbb{R}^{I_x\times LN}$ and $\mathbf{B}=[\mathbf{B}_1,\cdots,\mathbf{B}_N]\in\mathbb{R}^{I_{yz}\times LN}$, we can unfold the fMRI data $\boldsymbol{\mathcal{X}}$ into the following matrix representation of CPD model:

$$\mathbf{X}_{I_{xyz}\times JM}=\mathbf{S}(\mathbf{C}\odot_c\mathbf{D})^T\ . \qquad (3)$$

When updating $\mathbf{A}$, ALS requires large computation time to compute the Khatri-Rao product of $\mathbf{B}$ (e.g., 2,610×NL) and $\mathbf{C}\odot_c\mathbf{D}$ (e.g., 2,640×N), with dimension $I_{yz}JM\times NL$ (e.g., 6,890,400×NL). As such, we project (3) into $\mathbf{C}\odot_c\mathbf{D}$ subspace to reduce the dimensionality of the fMRI tensor:

$$\begin{aligned}\mathbf{X}_{I_{xyz}\times JM}(\mathbf{C}\odot_c\mathbf{D})&=\mathbf{S}(\mathbf{C}\odot_c\mathbf{D})^T(\mathbf{C}\odot_c\mathbf{D})\\ &=\mathbf{S}\mathbf{M}=\hat{\mathbf{X}}_{I_{xyz}\times N}\in\mathbb{R}^{I_{xyz}\times N},\end{aligned} \qquad (4)$$

where $\mathbf{M}=\{\mathbf{m}_n\}=(\mathbf{C}\odot_c\mathbf{D})^T(\mathbf{C}\odot_c\mathbf{D})\in\mathbb{R}^{N\times N}$. The reduced data $\hat{\mathbf{X}}_{I_{xyz}\times N}$ can be transformed as a three-way tensor

$\hat{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{I_x \times I_{yz} \times N}$. As the number of components N (generally smaller than J) is obviously smaller than JM. The size of $\hat{\boldsymbol{\mathcal{X}}}$ is significantly smaller than that of $\boldsymbol{\mathcal{X}}$. The $\hat{\boldsymbol{\mathcal{X}}}$ satisfies the following rank-(L,L,1) BTD model:

$$\hat{\boldsymbol{\mathcal{X}}} = \sum_{n=1}^{N} \left(\mathbf{A}_n \mathbf{B}_n^T \right) \circ \mathbf{m}_n , \tag{5}$$

and has the following matrix representations:

$$\hat{\mathbf{X}}_{I_x \times I_{yz} N} = \mathbf{A}(\mathbf{B} \odot \mathbf{M})^T , \tag{6}$$

$$\hat{\mathbf{X}}_{I_{yz} \times I_x N} = \mathbf{B}(\mathbf{M} \odot \mathbf{A})^T . \tag{7}$$

Therefore, we can derive the following ALS updating rules of $\mathbf{A}$ and $\mathbf{B}$ respectively based on (6) and (7):

$$\mathbf{A} \leftarrow \hat{\mathbf{X}}_{I_x \times I_{yz} N} (\mathbf{B} \odot \mathbf{M})^{T\dagger} , \tag{8}$$

$$\mathbf{B} \leftarrow \hat{\mathbf{X}}_{I_{yz} \times I_x N} (\mathbf{M} \odot \mathbf{A})^{T\dagger} . \tag{9}$$

When updating $\mathbf{C}$ and $\mathbf{D}$, as the number of whole-brain voxels of share SMs $\mathbf{S}$ is large (e.g., 120,060), we further project (7) into the $\mathbf{S}$ subspace to reduce the dimensionality of fMRI tensor:

$$\begin{aligned} \mathbf{S}^T \mathbf{X}_{I_{xyz} \times JM} &= \mathbf{S}^T \mathbf{S} (\mathbf{C} \odot_c \mathbf{D})^T \\ &= \hat{\mathbf{S}} (\mathbf{C} \odot_c \mathbf{D})^T = \hat{\mathbf{X}}_{N \times JM} \in \mathbb{R}^{N \times JM} , \end{aligned} \tag{10}$$

where $\hat{\mathbf{S}} = \{\hat{\mathbf{s}}_n\} = \mathbf{S}^T \mathbf{S} \in \mathbb{R}^{N \times N}$. We transform the reduced data $\hat{\mathbf{X}}_{N \times JM}$ into a tensor $\hat{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{N \times J \times M}$ which is evidently smaller than $\boldsymbol{\mathcal{X}}$ due to $N \ll I_{xyz}$. The $\hat{\boldsymbol{\mathcal{X}}}$ conforms the following CPD model:

$$\hat{\boldsymbol{\mathcal{X}}} = \sum_{n=1}^{N} \hat{\mathbf{s}}_n \circ \mathbf{c}_n \circ \mathbf{d}_n , \tag{11}$$

and has the following matrix representations:

$$\hat{\mathbf{X}}_{J \times MN} = \mathbf{C}(\mathbf{D} \odot_c \hat{\mathbf{S}})^T , \tag{12}$$

$$\hat{\mathbf{X}}_{M \times JN} = \mathbf{D}(\mathbf{C} \odot_c \hat{\mathbf{S}})^T . \tag{13}$$

Along this line, we can derive the following ALS updating rules of $\mathbf{C}$ and $\mathbf{D}$ based on (12) and (13):

$$\mathbf{C} \leftarrow \hat{\mathbf{X}}_{J \times MN} (\mathbf{D} \odot_c \hat{\mathbf{S}})^{T\dagger} , \tag{14}$$

$$\mathbf{D} \leftarrow \hat{\mathbf{X}}_{M \times JN} (\mathbf{C} \odot_c \hat{\mathbf{S}})^{T\dagger} . \tag{15}$$

2.3. The orthonormality constraint on shared SMs

We here add the orthonormality constraint on shared SMs to reduce the crosstalk between shared SM components [21, 22]. More precisely, we perform economical singular value decomposition (SVD) on $\mathbf{S}$, and calculate an orthonormal $\mathbf{S}$ by using the left singular matrix $\mathbf{U} \in \mathbb{R}^{I_{xyz} \times N}$ and right singular matrix $\mathbf{V} \in \mathbb{R}^{N \times N}$:

$$\begin{cases} [\mathbf{U}, \boldsymbol{\Sigma}, \mathbf{V}] = \text{svd}(\mathbf{S}), \\ \mathbf{S} = \mathbf{U}\mathbf{V}^H , \end{cases} \tag{16}$$

In reality, the complexity of BTD with ALS amounts to $\mathcal{O}\left(I_x I_{yz} MJNL\right)$. Since N is generally smaller than J, the complexity of the proposed method is obviously reduced to $\mathcal{O}\left(I_x I_{yz} LN^2\right)$ (e.g., 75.429 times lower complexity when N = 35, J = 165 and M = 16 in our experiment). As a whole, we conclude the detailed procedure of the proposed method in Table I.

Table I. The detailed implementation of the proposed algorithm.

Input: multi-subject fMRI data $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_x \times I_{yz} \times J \times M}$, the number of components N, rank L, and the maximum number of iterations $\text{iter}_{\max}$, the termination threshold of error $\varepsilon_{\min}$.

Randomly initialize $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, and let iter = 0 and $\varepsilon_0 = 1$;

while $\varepsilon_{\text{iter}} > \varepsilon_{\min}$ or iter < $\text{iter}_{\max}$ **do**

- iter = iter + 1;
- calculate $\mathbf{C} \odot_c \mathbf{D}$ and obtain $\hat{\mathbf{X}}_{I_{xyz} \times N}$ using (4);
- update $\mathbf{A}$ based on (8);
- update $\mathbf{B}$ based on (9);
- calculate $\mathbf{S} = [\text{vec}(\mathbf{A}_1 \mathbf{B}_1^T), \cdots, \text{vec}(\mathbf{A}_N \mathbf{B}_N^T)]$;
- add spatial orthonormality constraint on $\mathbf{S}$ based on (16);
- obtain $\hat{\mathbf{X}}_{N \times JM}$ based on (10);
- update $\mathbf{C}$ based on (14);
- update $\mathbf{D}$ based on (15);
- calculate $\varepsilon_{\text{iter}}$ using (2);

end while

Output: $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$

3. EXPERIMENTS AND RESULTS

In order to evaluate the efficacy of the proposed accelerated rank-(L,L,1,1) BTD under spatial orthonormal constraint (shorted as accBTD-O), we carry out 16-subject finger-tapping task-related fMRI data experiments (see [3, 24] for detail). As the size of 3D whole-brain images is 53×63×46, we reduce the size of 3D whole-brain images into 46×58×45 to remove some out-brain voxels. We unfold 3D brain images into matrix 46×2610 by concatenating frontal slices. The number of scans J is 165 for all M = 16 subjects.

Meanwhile, we compare the proposed accBTD-O and accBTD, BTD with ALS (BTD-ALS), BTD-ALS under spatial orthonormality constraint (BTD-ALS-O), BTD with NLS (BTD-NLS) [25]. We run each algorithm 10 times on MATLAB2018a under Win10 operating system with Inter(R) Xeon(R) Gold 6226R and 500G memory. We set $\text{iter}_{\max}$ = 200 and $\varepsilon_{\min} = 10^{-6}$ for each algorithm. We select the number of components R = 35 which suggested in [3]. As rank L is unknown, we change L from 5 to 45 with interval 5 to evaluate the L effect on the proposed method.

We here examine the task-related shared SM and TC. The task-related SM and TC references in [3] are used and the absolute Pearson correlation coefficients ρ between estimates and references are calculated to evaluate the performance. The higher ρ is, the better performance is. We also calculate the computation time of each algorithm without other programs running for the sake of fairness.

3.1. Results of different rank L values

We first compare the proposed accBTD-O and accBTD at different rank L values from 5 to 45 with interval 5, as shown in Fig. 1. Due to the incorporation of spatial orthonormality constraint, accBTD-O achieves obviously higher average ρ values of shared task-related SMs (see Fig.

1A) and TCs (see Fig. 1B) than accBTD. Furthermore, accBTD-O shows relatively higher average ρ values of both shared task-related SMs and TCs when L = 30 and 35 than other L values in Fig. 1. Meanwhile, accBTD exhibits higher average ρ values of SMs and TCs when L = 35 and 40 as shown in Fig. 1. On the whole, accBTD-O and accBTD at L = 35 show relatively better shared SM and TC estimates than other L values. Therefore, we choose L = 35 for the following experiments, which is also close to the rank L = 40 estimated in [6].

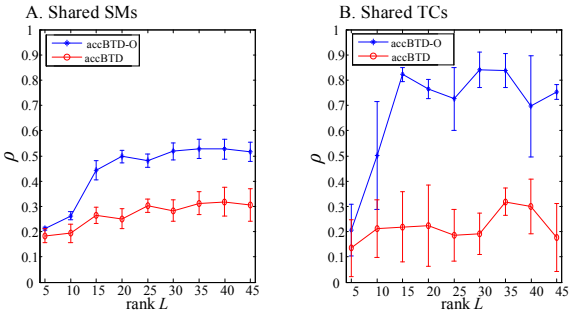

Fig. 1. Effects of different rank L values (from 5 to 45 with interval 5) on accBTD-O and accBTD in terms of means and standard deviations of ρ values of task-related SMs (A) and TCs (B) for multi-subject task-related fMRI data.

3.2. Comparison results

We further compare the separation performance of proposed accBTD-O and accBTD with BTD-ALS-O, BTD-ALS, and BTD-NLS at L = 35, as shown in Table II. First, accBTD and BTD-ALS show higher average ρ values of shared task-related SMs and TCs than BTD-NLS (see Table II), indicating that ALS is more suitable for multi-subject fMRI data than NLS. Second, owing to the spatial orthonormality constraint, the proposed accBTD-O and BTD-ALS-O extract apparently higher average ρ values of shared task-related SMs and TCs than accBTD and BTD-ALS, respectively. Third, due to accelerating the update of loading matrices, accBTD-O costs 17.36 times less computation time than BTD-ALS-O and slightly higher average ρ values for the shared TCs, although slightly lower average ρ values for the shared SMs.

Some example shared task-related SMs and TCs estimated by the proposed accBTD-O and accBTD, BTD-ALS-O, BTD-ALS, and BTD-NLS at L = 35 are shown in Fig. 2. We can conclude from Fig. 2A that accBTD-O and BTD-ALS-O show obviously larger activations in right primary motor areas (RPMA) and supplementary motor areas (SMA) than other three methods. Besides, accBTD-O and BTD-ALS-O additionally exhibits activated voxels in left primary motor areas (LPMA). In addition, the shared TC waveforms estimated by accBTD-O and BTD-ALS-O are more closed to TC reference than other three methods as displayed in Fig. 2B. These results further show the advantage of spatial orthonormality constraint.

Table II. Comparison of accBTD-O, accBTD, BTD-ALS-O, BTD-ALS and BTD-NLS for analyzing the multi-subject task-related fMRI data over 10 runs in terms of the mean and standard deviation of ρ values of the shared task-related SM and TC estimates and average computation time (s).

	accBTD-O	accBTD	BTD-ALS-O	BTD-ALS	BTD-NLS
SM	0.53±0.03	0.31±0.05	0.57±0.05	0.48±0.03	0.12±0.03
TC	0.84±0.07	0.32±0.05	0.82±0.14	0.32±0.21	0.21±0.07
Time	13097.42	13901.08	240408.89	239280.62	170301.12

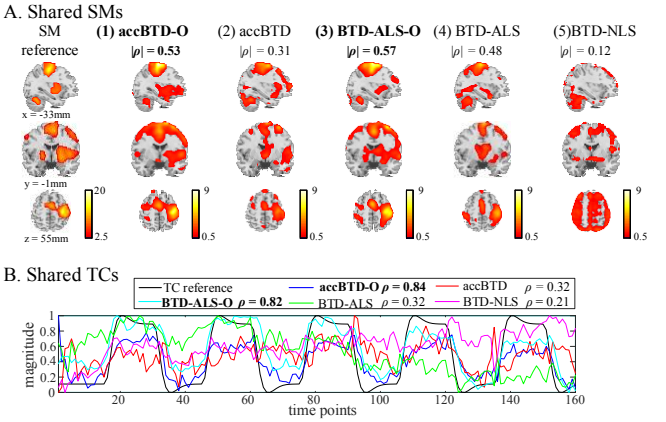

Fig. 2. Example shared task-related SMs ($Z \geq 0.5$) and TCs estimated by accBTD-O, accBTD, BTD-ALS-O, BTD-ALS, and BTD-NLS. The ρ values with corresponding references are also shown.

4. CONCLUSIONS

This study proposes a novel accelerated rank-(L,L,1,1) BTD under spatial orthonormality constraint, to conquer the problems of high computation time of BTD applied to multi-subject fMRI data and to reduce the crosstalk between components. Experiment results show that the proposed method gets better separation performance when rank L = 35, is obviously 17.36 times faster than BTD-ALS-O and achieves similar separation performance with BTD-ALS-O. Multi-subject fMRI data are originally complex-valued and possess high spatiotemporal variability, and given the performance improvement, we can also apply the proposed method to complex-valued fMRI data with some spatial and temporal constraints.

5. ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China under Grants 61901061, 61871067, 61379012, 61972056, Natural Science Foundation of Hunan Province under Grant 2020JJ5603, the Research Foundation of Education Bureau of Hunan Province under Grant 19C0031, Postgraduate Research Innovation Project of Changsha University of Science and Technology under Grant CX20201SS75, NSF grants 2112455, 0840895 and 0715022, NIH grants R01MH104680, R01MH107354, and R01EB005846, and in part by the Fundamental Research Funds for the Central Universities, China, under Grant DUT20ZD220.

6. REFERENCES

[1] M. A. Lindquist, "The statistical analysis of fMRI data," *Statistical Science*, vol. 23, no. 4, pp. 439-464, June 2009.

[2] V.D. Calhoun, J. Liu, and T. Adali, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," Neuroimage, vol. 45, no. 1, Suppl, pp. S163-72, 2009.

[3] L.D. Kuang, Q.H. Lin, X.F. Gong, J. Sui and V.D. Calhoun "Multi-subject fMRI analysis via combined independent component analysis and shift-invariant canonical polyadic decomposition," *Journal of Neuroscience Methods*, vol. 256, pp. 127-140, December 2015.

[4] .H. Lee, T.W. Lee, F.A. Jolesz, and S.S. Yoo, "Independent vector analysis (IVA): multivariate approach for fMRI group study," NeuroImage, vol. 40, no. 1, pp. 86-109, March 2008.

[5] C. Chatzichristos, E. Kofidis, Y. Kopsinis, M.M. Moreno and S. Theodoridis, "Higher-Order block term decomposition for spatially folded fMRI data," *International Conference on Latent Variable Analysis and Signal Separation*, vol. 10169, pp. 3-15, February 2017.

[6] C. Chatzichristos, E. Kofidis, M. Morante and S. Theodoridis, "Blind fMRI source unmixing via higher-order tensor decompositions," *Journal of Neuroscience Methods*, vol. 315, pp. 17-47, March 2019.

[7] L.D. Lathauwer, "Decompositions of a higher-order tensor in block terms—Part II: definitions and uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1033-1066, September 2008.

[8] N.D. Sidiropoulos, G.B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 810−823, 2000.

[9] G. Tomasi, "Practical and computational aspects in chemometric data analysis," *Ph.D. thesis*, The Royal Veterinary and Agricultural University, Frederiksberg, Denmark, May 2006.

[10] M. Rajih, P. Comon and R.A. Harshman, "Enhanced line search: a novel method to accelerate PARAFAC," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1128-1147, September 2008.

[11] L. Sober, I. Domanov, M.V. Barel and L.D. Lathauwer, "Exact line and plane search for tensor optimization," *Computational Optimization and Applications*, vol. 63, pp. 121-142, May 2015.

[12] G. Zhou, A. Cichocki, and S. Xie, "Accelerated canonical polyadic decomposition by using mode reduction," *IEEE Transactions on Neural Networks and Learning System*, vol. 24, no. 12, pp. 2051-2062, December 2012.

[13] A. Phan, P. Tichavsky and A. Cichocki, "CANDECOMP/ PARAFAC decomposition of high-order tensors through tensor reshaping," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4847-4860, October 2013.

[14] P. Tichavsky, A.H. Phan and Z. Koldovsky, "Cramér-Rao-Induced bounds for CANDECOMP/PARAFAC tensor decomposition," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1986-1997, April 2013.

[15] M. Shi, D. Li and J.Q. Zhang, "Matrix polynomial predictive model: a new approach to accelerating the PARAFAC decomposition," *IEEE Access*, vol. 7, pp. 91872-91884, July 2019.

[16] A.A. Rontogiannis, P.V. Giampouras and E. Kofidis, "Rank-Revealing block-term decomposition for tensor completion," In: *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, pp. 2915-2919, June 2021.

[17] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[18] Q. Shi, Y.M. Cheung and J. Lou, "Robust tensor SVD and recovery with rank estimation," *IEEE Transactions on Cybernetics*, pp. 1-16, April 2021.

[19] M.Y. Zhou, Y.P Liu, Z. Long, L.X. Chen and C. Zhu, "Tensor rank learning in CP decomposition via convolutional neural network," *Signal Processing: Image Communication*, vol. 73, pp. 12-21, April 2019.

[20] X. Han, L. Albera, A. Kachenoura, H. Shu and L. Senhadji, "Block term decomposition with rank estimation using group sparsity," In: *2017 IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Curacao, pp. 1-5, March 2018.

[21] B. Sen and K.K. Parhi, "Constrained tensor decomposition optimization with applications to fMRI data analysis," In: *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, pp. 1923-1928, February 2019.

[22] B. Sen and K.K. Parhi, "Extraction of common task signals and spatial maps from group fMRI using a PARAFAC-based tensor decomposition technique," In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, pp. 1113-1117, March 2017.

[23] L.D. Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms—Part III: alternating least squares algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067-1083, September 2008.

[24] L.D. Kuang, Q.H. Lin, X.F. Gong, F. Cong, Y.P. Wang, and V.D. Calhoun, "Shift-invariant canonical polyadic decomposition of complex-valued multi-subject fMRI data with a phase sparsity constraint," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 844–853, 2020.

[25] N. Vervliet, O. Debals, L. Sorber, M. Van Barel and L.D. Lathauwer, *Tensorlab 3.0*, Available online, Mar. 2016. URL: http://www.tensorlab.net/