

# An improved multi-objective optimization-based CICA method with data-driver temporal reference for group fMRI data analysis

Yuhu Shi<sup>1</sup> · Weiming Zeng<sup>1,2</sup> · Xiaoyan Tang<sup>1</sup> · Wei Kong<sup>1</sup> · Jun Yin<sup>1</sup>

Received: 11 March 2017 / Accepted: 17 August 2017 / Published online: 2 September 2017  
© International Federation for Medical and Biological Engineering 2017

**Abstract** Group independent component analysis (GICA) has been successfully applied to study multi-subject functional magnetic resonance imaging (fMRI) data, and the group independent component (GIC) represents the commonality of all subjects in the group. However, some studies show that the performance of GICA can be improved by incorporating a priori information, which is not always considered when looking for GICs in existing GICA methods. In this paper, we propose an improved multi-objective optimization-based constrained independent component analysis (CICA) method to take advantage of the temporal a priori information extracted from all subjects in the group by incorporating it into the computational process of GICA for group fMRI data analysis. The experimental results of simulated and real data show that the activated regions and the time course detected by the improved CICA method are more accurate in some sense. Moreover, the GIC computed by the improved CICA method has a higher correlation with the corresponding independent component of each subject in the group, which means that the improved CICA method with the temporal a priori information extracted from the group can better reflect the commonality of the subjects. These results demonstrate that the improved CICA method has its own advantages in fMRI data analysis.

**Keywords** fMRI · GICA · CICA · Temporal a priori information · Multi-objective optimization

## 1 Introduction

In the past two decades, functional neuroimaging has become an important tool for studying various neural mechanisms in the brain. In particular, functional magnetic resonance imaging (fMRI) has drawn considerable attention due to being noninvasive and having high spatio-temporal resolution [1–3]. Many methods have been used for fMRI data analysis, and these methods are generally divided into model-based methods and data-driven methods [4–6]. The model-based methods require a priori information about the experimental paradigm, and usually only local brain area data are considered in these methods rather than data from the whole brain. On the contrary, data-driven methods do not depend on any a priori information, and several of these methods have been fruitfully applied to the field of fMRI data analysis, such as principle component analysis (PCA), independent component analysis (ICA), and clustering analysis (CA).

Among them, the purpose of ICA is to decompose the observed multivariate data into the source signals, which are assumed statistically independent and non-Gaussian. Since it was first introduced into this field for single-subject fMRI data analysis [7], it has become one of the most popular methods to analyze fMRI data [8–10]. Compared with univariate methods such as general linear model (GLM) methods based on the single voxel level [11], ICA is a multivariate method that considers the interactions between the voxels and is increasingly being applied to extract functional neural networks from the fMRI data of various cognitive activities without relying

✉ Weiming Zeng  
zengwm86@163.com

<sup>1</sup> Laboratory of Digital Image and Intelligent Computation, Shanghai Maritime University, 1550 Harbor Avenue, Pudong, Shanghai 201306, China

<sup>2</sup> Information Engineering College, Shanghai Maritime University, 1550 Harbor Avenue, Pudong, Shanghai 201306, China

on any a priori information. Currently, ICA has been widely used for fMRI data analysis in a resting state [12, 13] or a specific cognitive task-related state [14, 15]. In many circumstances, ICA needs to be used for multiple-subject fMRI data analysis, which is usually referred to as group ICA (GICA) [16].

As a kind of purely data-driven blind source separation technique, ICA does not require any a priori information. However, many studies have shown that the capabilities of ICA can be greatly improved if some a priori information is incorporated into the estimation process when it is available [17–19]. This kind of method is usually called constrained ICA (CICA) or ICA with reference (ICA-R) [20, 21]. Compared with the classical ICA, CICA or ICA-R only extracts the sources of interest without extracting all the sources by introducing the a priori information into the calculation process, and this method avoids computing uninteresting sources, facilitates subsequent applications, and reduces the computation time and storage requirements [22, 23]. In addition, the separation quality and accuracy of interesting sources can also be improved through the incorporation of a priori information [24, 25]. On this basis, a great number of other extension methods have been proposed [26–29].

Currently, the existing methods with a priori information generally consider specific knowledge associated with the sources. For example, the spatial template of some mature networks, such as the visual network or the default mode network, and the specific experimental paradigms of some cognitive task experiments, such as block stimulation mode in a visual cognitive experiment, are considered [30–32]. However, this knowledge about the sources is not always known, especially in the case of complex cognitive activity. It is important to get available a priori information from the existing data itself. Recently, we proposed a method called GICA-IR to extract some intrinsic, spatial a priori information from data from groups of subjects. The results demonstrated that the group independent component (GIC) computed by GICA-IR is more representative of the commonality of the subjects in the group through incorporating this information into the GICA extraction procedure [33].

However, as far as we know, there are very few papers that study how to obtain temporal a priori information from the subjects in a group for multi-subject fMRI data analysis. In this paper, we propose a novel method to extract the temporal a priori information from the data from groups of subjects and then incorporate it into a GICA computational process for group fMRI data analysis using the improved multi-objective optimization-based CICA method. The experimental results showed that the GIC computed by the improved CICA method is more representative of the commonality of subjects in the group.

## 2 Methods

In this section, we first briefly introduce the relevant knowledge about ICA and GICA. Then, a detailed description of the proposed improved CICA method with temporal reference signal is presented. Finally, we provide a description of the experimental data and data processing.

### 2.1 Independent component analysis

Assume  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)'$  is a  $T \times V$  matrix of observed fMRI data from a single subject, where  $T$  and  $V$  represent the number of time points and the number of voxels within the brain, respectively. Then, the classical spatial ICA can be formulated as the following linear generative model:

$$\mathbf{X} = \mathbf{M}\mathbf{S} \quad (1)$$

where  $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N)'$  is an  $N \times V$  matrix in which each row represents a spatial source and  $N$  denotes the number of sources. These sources are assumed to be unobservable, independent, and non-Gaussian.  $\mathbf{M}$  is a  $T \times N$  unknown mixing matrix that mixes the  $N$  sources to generate the observed fMRI data, whose columns contain the associated time courses of the  $N$  source signals. Solving the ICA is estimating an  $N \times T$  unmixing matrix  $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N)'$  such that  $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)'$  is a good approximation of the sources  $\mathbf{S}$  according to the following equation:

$$\mathbf{Y} = \mathbf{W}\mathbf{X} \quad (2)$$

Many algorithms can be used to solve the ICA model in (1), and currently the most widely used ICA algorithms include InfoMax [34] and FastICA [35].

### 2.2 Group independent component analysis

ICA was implemented on the multi-subject fMRI data, often referred to as GICA. Temporal concatenation GICA (TCGICA) is the most widely used method in the existing GICA approaches, which assumes that all subjects have common, spatially independent components (ICs). Specifically, assuming that there are  $K$  subjects in total, and  $\mathbf{X}_i$  is a  $T \times V$  matrix that represents the fMRI data of each subject  $i$  ( $i = 1, \dots, K$ ), the TCGICA concatenates the fMRI data of  $K$  subjects along the temporal dimension and then decomposes the  $KT \times V$  group data as follows:

$$\hat{\mathbf{X}} = \left( \mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_K \right)' = \hat{\mathbf{M}}\hat{\mathbf{S}} \quad (3)$$

where  $\hat{\mathbf{M}}$  is a  $KT \times L$  group mixing matrix,  $\hat{\mathbf{S}}$  is a  $L \times V$  matrix in which each row represents a GIC, and  $L$  denotes the number of GICs.

### 2.3 The improved CICA with temporal reference signal

In this subsection, we first give the specific steps of the method for extracting the temporal reference signal from the group of subjects, and then we present a detailed description of the improved multi-objective optimization-based CICA method.

#### 2.3.1 The extraction of the temporal reference signal

We assume there are a total of  $K$  subjects in the group, and all subjects have  $T$  time points and  $V$  voxels after normalization. First, we implemented ICA on each of the subjects in the group. For each subject  $i$ , ICA was defined as

$$X_i = M_i S_i, \quad (i = 1, 2, \dots, K) \tag{4}$$

where  $X_i$  is a  $T \times V$  fMRI observed data,  $S_i = (s_{i1}, s_{i2}, \dots, s_{iN_i})'$  is an  $N_i \times V$  matrix, and each row represents an IC of subject  $i$ .  $M_i = (m_{i1}, m_{i2}, \dots, m_{iN_i})$  is a  $T \times N_i$  mixing matrix. For the sake of simplicity, we only considered the case where each subject has just one IC of interest, and the correspondence of the ICs of different subjects could be obtained by the absolute value of the spatial correlation [36].

Now we denoted  $s_{in_i}$  ( $i = 1, 2, \dots, K$ ) as the  $n_i$ st IC, which is the source of interest for the subject  $i$ , and  $m_{in_i}$  ( $i = 1, 2, \dots, K$ ) is the corresponding time course. Finally, these time courses were concatenated into a time course with greater length:

$$m = \left( m_{1n_1}', m_{2n_2}', \dots, m_{Kn_K}' \right)' \tag{5}$$

where  $m$  represents the temporal reference signal, which is a column vector of size  $KT \times 1$ , and it is used as  $rt_i$  in the following improved multi-objective optimization-based CICA method for group fMRI data analysis.

#### 2.3.2 The improved multi-objective optimization-based CICA method

In this paper, the proposed improved CICA with temporal reference signal method was established with the multi-objective optimization framework as follows:

$$\begin{aligned} & \text{Maximize } \begin{cases} J(\mathbf{w}_i) \approx \{E[G(\hat{\mathbf{s}}_i)] - E[G(\mathbf{v})]\}^2 \\ \varepsilon_1(\mathbf{w}_i) = \text{abs}(E[\hat{\mathbf{m}}_i \cdot \mathbf{rt}_i]) \end{cases} \\ & \text{Subject to } \|\mathbf{w}_i\|^2 = 1 \end{aligned} \tag{6}$$

where  $J(\mathbf{w}_i)$  is the negentropy of the estimated IC,  $\hat{\mathbf{s}}_i = \mathbf{w}_i^T \hat{\mathbf{X}}$ .  $\hat{\mathbf{m}}_i = \mathbf{Z}^{-1} \mathbf{w}_i$  denotes the time course corresponding to  $\hat{\mathbf{s}}_i$ , and it is a column vector of size  $KT \times 1$ .  $\mathbf{Z}$  denotes the  $L \times KT$  whitening matrix, which is obtained by eigenvalue decomposition, and  $L$  denotes the number of ICs.  $\mathbf{v}$  is a Gaussian random variable with a zero mean and a unit variance.  $G(\cdot)$  is a non-quadratic function, and  $G(\mathbf{v}) = \log(\cosh(\mathbf{v}))$  is used in this paper.  $\mathbf{rt}_i$  denotes a temporal reference signal, which is a column vector of size  $KT \times 1$ , and  $\varepsilon_1(\mathbf{w}_i) = \text{abs}(E[\hat{\mathbf{m}}_i \cdot \mathbf{rt}_i])$  is specifically defined as the Pearson correlation coefficient to measure the closeness between  $\hat{\mathbf{m}}_i$  and  $\mathbf{rt}_i$  where both  $\hat{\mathbf{m}}_i$  and  $\mathbf{rt}_i$  have a zero mean and unit variance. Each solution of the multi-objective optimization problem in (6) corresponds to an optimal unmixing column vector  $\mathbf{w}_i$  that is constrained to  $\|\mathbf{w}_i\|^2 = 1$ .

For multi-objective optimization problems, there is no global solution that makes all of the cost functions achieve the optimum simultaneously. Therefore, a trade-off solution is needed to balance the optimality of all cost functions. Among the methods of solving the multi-objective optimization problem, the weighted summing method is simple and efficient, and it is achieved by optimizing the weighted sum function of the objective functions on the condition that the weight value of each objective function is positive and the sum of all weights is 1 [37]. Therefore, this method was adopted to solve the multi-objective optimization problem of (6) in our study. To avoid the calculation process being controlled by the objective function with a larger value, the arc-tangent function was used to normalize the objective function  $J(\mathbf{w}_i)$  in (6):

$$f_1(\mathbf{w}_i) = (2/\pi) \cdot \arctan[c_i \cdot J(\mathbf{w}_i)] \tag{7}$$

where  $c_i$  in (7) is automatically determined so that the possible values of  $f_1(\mathbf{w}_i)$  and  $\varepsilon_1(\mathbf{w}_i)$  range from 0 to 1 [36]. Then, the reformulated linear weighted objective function is

$$f(\mathbf{w}_i) = a_1 \cdot f_1(\mathbf{w}_i) + a_2 \cdot \varepsilon_1(\mathbf{w}_i) \tag{8}$$

where  $a_i$  ( $i = 1, 2$ ) is the weight parameters and  $a_1 + a_2 = 1$ . Then, the iteration algorithm for optimizing  $f(\mathbf{w}_i)$  can be derived as follows:

$$\begin{aligned} \nabla f(\mathbf{w}_i) = & a_1 \cdot \nabla f_1(\mathbf{w}_i) + a_2 \cdot \nabla \varepsilon_1(\mathbf{w}_i) = 2a_1 \cdot (2/\pi) \cdot c_i / \left(1 + [c_i \cdot J(\mathbf{w}_i)]^2\right) \\ & \cdot \left( E[G(\mathbf{w}_i^T \hat{\mathbf{X}})] - E[G(\mathbf{v})] \right) \cdot E[\hat{\mathbf{X}} \cdot g(\mathbf{w}_i^T \hat{\mathbf{X}})] + a_2 \cdot E[\mathbf{Z}^{-1} \mathbf{rt}_i] \end{aligned} \tag{9}$$

where  $g(\cdot)$  is the derivative of  $G(\cdot)$ ,  $\text{sog}(\mathbf{v}) = \tanh(\mathbf{v})$ .  $E(\cdot)$  can be estimated as the mean of all samples. Once the

gradient of the weighted sum function in (8) is calculated, the steepest ascent iteration formula can be set up as follows:

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu(k) \cdot d_i(k) \quad (10)$$

where  $\mathbf{w}_i(k)$  denotes the value of  $\mathbf{w}_i$  after the  $k$ th iteration.  $d_i(k) = \nabla f(\mathbf{w}_i(k)) / \|\nabla f(\mathbf{w}_i(k))\|$ , and  $\mu(k)$  denotes the step-length. Finally, the corresponding time course can be calculated by using the following formula when the IC  $s_i$  is obtained:

$$\hat{\mathbf{m}}_i = \mathbf{Z}^{-1} \mathbf{w}_i \quad (11)$$

where  $\mathbf{w}_i$  is the unmixing column vector corresponding to  $s_i$ .

In the last subsection, the whole process for the proposed method is summarized as a flowchart (see Fig. 1).

## 2.4 Experimental data

In this subsection, the simulated and real fMRI data were used to evaluate the performance of the improved CICA method at the group level for fMRI data analysis.

### 2.4.1 Simulated data

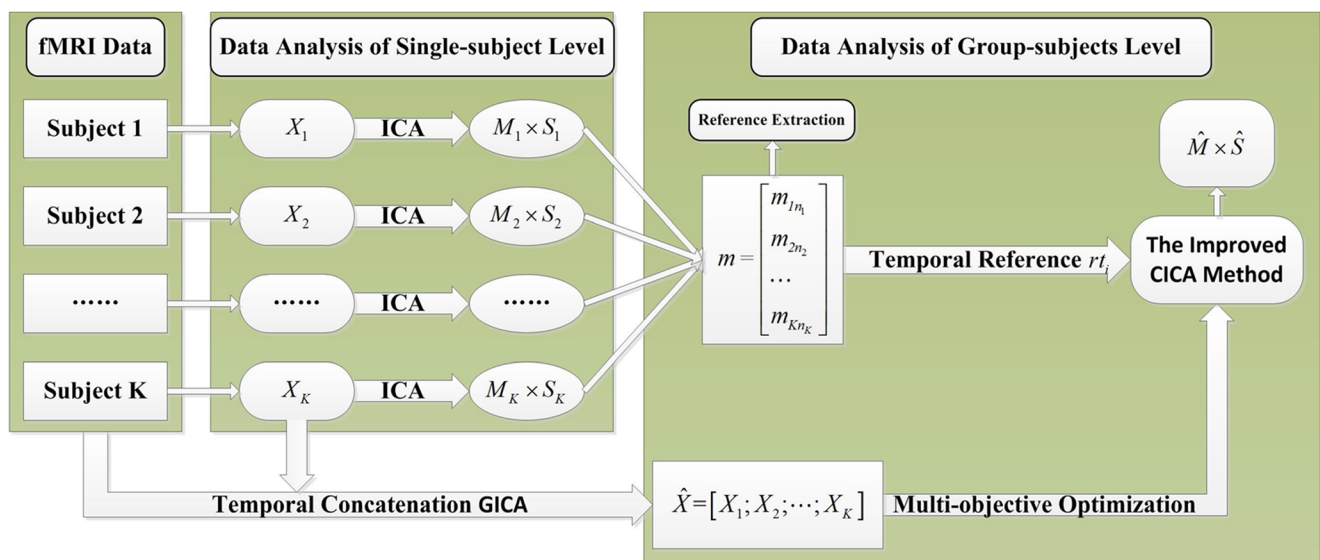
The simulated fMRI data were obtained using the code downloaded from [http://mlsp.umbc.edu/simulated\\_fmri\\_data.html](http://mlsp.umbc.edu/simulated_fmri_data.html) [38], where a set of spatial sources with different simulated hemodynamic time courses

were designed to generate the data through linear superposition. Specifically, the data for each simulated subject were produced by mixing the eight original sources with their corresponding time courses, and a total of 100 sample images were included in the data, where each source image had  $60 \times 60$  pixels with 100 time points (see Fig. 2). The eight original sources were designed such that source 1 was task related, source 2 and source 6 were transiently task related, source 5 was function related, and source 3–source 4 and source 7–source 8 were artifact related. In particular, the task-related source 1 had a time course similar to the block-like shape that is often used to imitate an experimental paradigm.

In this experiment, 20 groups of simulated datasets were produced from the same original sources/TCs by adding specific variability to each subject, and each simulated dataset included five subjects. The spatial variation in the sources of each subject was portrayed by adding Gaussian noise with a different signal-to-noise ratio (SNR) to the source images, and the signal-to-noise ratios ranged from 0.3 to 0.4 and were randomly determined for different subjects. The temporal variation in time courses was simulated by applying time delay and amplitude modulation, which were also randomly determined with the time courses.

### 2.4.2 Real fMRI data

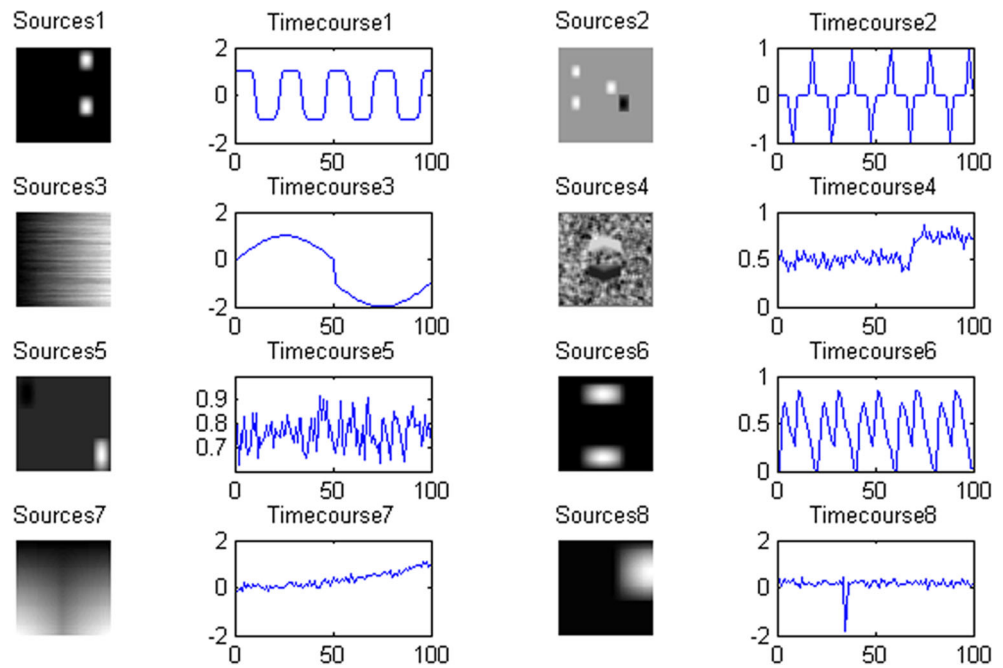
A real fMRI dataset from five subjects who completed a visual task was included in this study. All five subjects



**Fig. 1** The flowchart of the proposed method.  $X_i$  ( $i = 1, 2, \dots, K$ ) represents the fMRI data of subject  $i$ .  $M_i$  ( $i = 1, 2, \dots, K$ ) and  $S_i$  ( $i = 1, 2, \dots, K$ ) represent the temporal and spatial components of subject  $i$ , respectively, which are obtained using ICA.  $\mathbf{m}$  represents the

temporal reference  $rt_i$ .  $\mathbf{M}$  and  $\mathbf{S}$  represent the group temporal and spatial components, respectively, which are obtained using the improved CICA method

**Fig. 2** The simulated original sources and their corresponding time courses, which are all normalized to have a mean of zero and a unit variance. Specifically, source 1 represents task-related information, source 2 and source 6 represent transiently task-related information, source 5 represents function-related information, and the others four sources (source 3–source 4 and source 7–source 8) represent artifact-related information



were notified about the aim of this study, and they signed a written consent letter. The block pattern of OFF-ON-OFF-ON-OFF-ON-OFF-ON was used as the experimental paradigm, and each block lasted 20 s. In the “ON” state, the visual stimulus corresponded to a radial blue/yellow checkerboard that reversed at 7 Hz. In the “OFF” state, the participants were required to focus on a cross at the center of the screen. The BOLD fMRI data from two subjects were acquired using single-shot SENSE gradient echo EPI with 37 slices, providing whole-brain coverage and 70 volumes, a TR of 2.0 s, and a scan resolution of  $64 \times 64$ . The in-plane resolution was  $4 \text{ mm} \times 4 \text{ mm}$ , and the slice thickness was 4 mm. The other three subjects were acquired using single-shot SENSE gradient echo EPI with 40 slices, providing whole-brain coverage and 70 volumes, a TR of 2.0 s, and a scan resolution of  $80 \times 80$ . The in-plane resolution was  $3 \text{ mm} \times 3 \text{ mm}$ , and the slice thickness was 3 mm.

### 2.5 Data processing

All of the calculations in this study were implemented on a workstation whose operation system platform was Windows 7 Unlimited Service Pack 1, with an Intel(R) Xeon(R) E5-1620 3.60 GHz processor and 40 GB RAM. The preprocessing and calculation steps from FastICA and the improved CICA methods were run using Matlab (Matlab, 2012b, MathWorks Inc., Sherborn, MA, USA) [39].

The preprocessing steps of the real-data experiment were implemented using SPM8 software (<http://www.fil.ion.ucl.ac.uk/spm/>), which included slice timing, motion

correction, spatial normalization, and smoothing with a Gaussian kernel of 8 mm. In all experiments, FastICA, the newly published method [39] (which is denoted as CICA in this paper), and the improved CICA method were used. Specifically, FastICA was implemented using GIFT software (v2.0e) (<http://mialab.mrn.org/software/>) for the purpose of comparison. Moreover, ICASSO [40] with 20 runs of ICA was used to obtain reliable ICs, and MDL [41] was used to estimate the number of ICs. Furthermore, the positioning and display of the spatial networks were implemented using MRICro software (<http://www.mricro.com>).

### 3 Results

In this section, the advantages of the improved CICA method are demonstrated by comparing the experimental results obtained with the improved CICA method with those obtained with FastICA and CICA for the group fMRI data analysis. Specifically, the spatial a priori information in CICA was obtained using a previously described method [33], and a detailed description is presented in Appendix 1.

First, a power analysis of the receiver operating characteristic (ROC) curve was adopted to evaluate the spatial detection ability of these methods in the simulated experiment, which is denoted by the area surrounded by the ROC curve, and a larger area under the curve (AUC) is usually better [42]. Second, the correlations among the time courses computed by FastICA, CICA, and the

improved CICA methods with the true time course were used to measure the temporal performance, which can be calculated using the following formula:

$$\text{corrcoef\_TC} = \text{abs}\left(\frac{\text{cov}(\mathbf{TTC}, \mathbf{TC})}{\sqrt{\text{cov}(\mathbf{TTC})\text{cov}(\mathbf{TC})}}\right) \quad (12)$$

where  $\mathbf{TTC}$  represents the true time course and  $\mathbf{TC}$  represents the time courses computed by each method. Finally, the correlations among the GICs computed by FastICA, CICA, and the improved CICA methods with the corresponding IC of each subject were used to evaluate the group-level analysis, which can be calculated as follows:

$$\text{corrcoef\_IC}_i = \text{abs}\left(\frac{\text{cov}(\mathbf{GIC}, \mathbf{IC}_i)}{\sqrt{\text{cov}(\mathbf{GIC})\text{cov}(\mathbf{IC}_i)}}\right) \quad (13)$$

where  $\mathbf{GIC}$  represents the group independent component and  $\mathbf{IC}_i$  represents the corresponding independent component of subject  $i$ .

In the experiments in this paper, the weighting parameter “a” of the improved CICA method was a value from 0.1 to 0.9 with a step length of 0.1, and then we decided which “a” to use according to the evaluation of the experimental results of each “a.” The corresponding results were the final experimental results. Specifically, in order to guarantee the consistency of the selection of the optimal weighting parameters for the simulated data and real data experiments, the index obtained by the following formula (14) was used to select the best situation. We first calculated the average of the correlation coefficients between the GIC and the corresponding IC of each subject across all subjects and then combined this average with the correlation coefficient between the true time course and the time course calculated with a new average, which was used as the quantitative indicator to select the weighting parameter of the improved CICA method:

$$\text{index} = \text{corrcoef\_TC} + \sum_{i=1}^K \text{corrcoef\_IC}_i / K \quad (14)$$

where  $K$  denotes the number of subjects in the group and  $K = 5$  for both simulated data and real data experiments in this paper.

### 3.1 Simulated data results

In this experiment, we focused on the task-related source 1, and its corresponding time course had a block-like shape that closely matched the experimental paradigm. The results of using weighting parameter  $a = 0.6$  for dataset 8,  $a = 0.1$  for dataset 14, and  $a = 0.9$  for other datasets are presented according to formula (14).

Figure 3 shows the AUCs of the ROC curves of the GICs that were computed by FastICA, CICA, and the improved CICA methods on the 20 simulated datasets. It can be seen clearly from the figure that the AUCs of CICA were significantly higher than those of the improved CICA, except dataset 14, and those of FastICA across all datasets, and the AUCs of the improved CICA are significantly higher than those of FastICA, except dataset 8. These significant differences were verified by  $T$  test with a confidence level of 95%, which demonstrated that CICA has the best source recovery ability, and the improved CICA method with the temporal reference signal extracted from the group of subjects had better source recovery ability compared with FastICA.

Figure 4 shows the correlation coefficients (CCs) among the true time course and the group time courses computed by FastICA, CICA, and the improved CICA methods on the 20 simulated datasets. It can be seen from the figure that the CCs of the improved CICA are significantly higher than those of CICA and FastICA across all simulated datasets using  $T$  tests with a confidence level of 95%, while the CCs between CICA and FastICA were not significantly different. These results demonstrate that the improved CICA method had better temporal detection performance compared with CICA and FastICA methods.

Figure 5 shows the average CCs of the GIC with the corresponding IC of each subject in the group across the 20 simulated datasets and their standard deviations, which are obtained using GICs computed by FastICA, CICA, and the improved CICA methods. We can see from the figure that the CCs calculated by CICA and the improved CICA methods were significantly higher than those of FastICA, which was verified by  $T$  tests at a confidence level of 95%. At the same time, the CCs of the improved CICA were slightly higher than those of CICA, but they were not significantly different. These results demonstrate that the GIC computed by the improved CICA method was more representative of the commonality of subjects

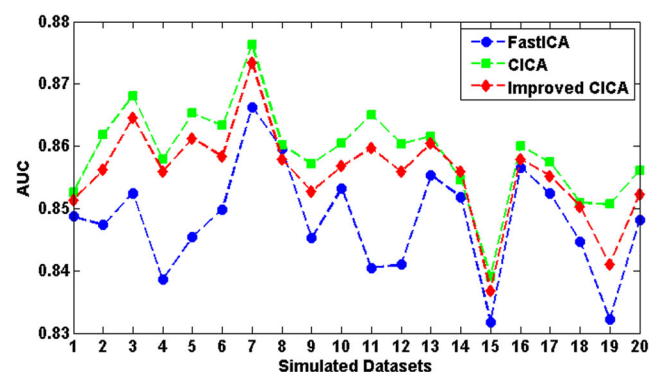
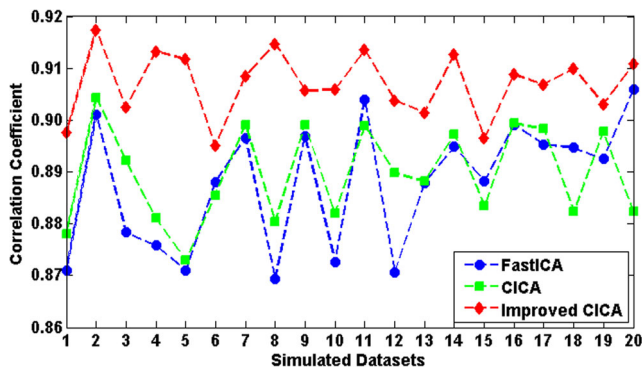


Fig. 3 The AUCs of FastICA, CICA, and the improved CICA methods on the 20 simulated datasets



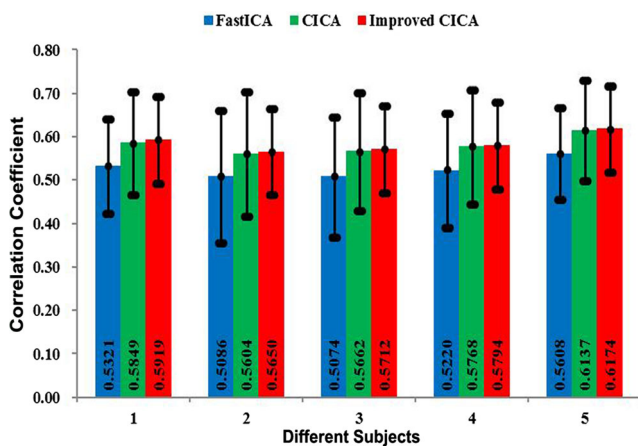
**Fig. 4** The CCs among the true time course and the time courses computed by FastICA, CICA, and the improved CICA methods on the 20 simulated datasets

in the group. That is, the temporal reference signal extracted from the group of subjects improved the analysis of the group data.

### 3.2 Real data results

In this section, only the task-related independent component is considered. The performance of the improved CICA method was compared with FastICA and CICA for fMRI data analysis at the group level. The results using weighting parameter  $\alpha = 0.9$  with the improved CICA method are presented according to formula (14).

Figure 6 shows the visual regions detected by FastICA, CICA, and the improved CICA methods. We can see from the figure that the regions of the improved CICA method are better than those of FastICA and similar to those of CICA, which means that the improved CICA method is superior to FastICA in source recovery using the temporal



**Fig. 5** The average CCs of the 20 simulated datasets among the GIC and the corresponding IC of each of the five subjects and their standard deviations, which are obtained using GICs computed by FastICA, CICA, and the improved CICA methods

reference signal from the group of subjects and has the same performance as the CICA method.

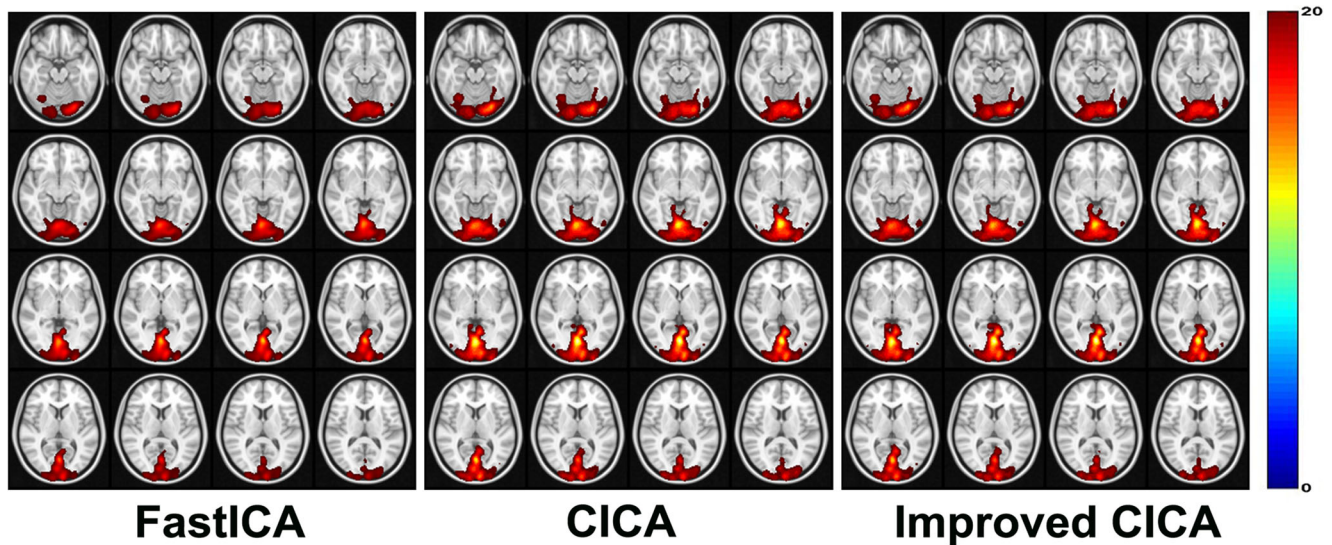
Figure 7 shows the prior block and the time courses computed by FastICA, CICA, and the improved CICA methods as well as the CCs between the prior block and the time course of each method. It can be clearly seen from the figure that the CC of the improved CICA method is higher than those of CICA and FastICA, which means that the time course computed by the improved CICA method is more accurate than those of FastICA and CICA and further demonstrates its better temporal performance.

Figure 8 shows the CCs between the GIC and the corresponding IC of each subject in the group. The results were obtained using GICs computed by FastICA, CICA, and the improved CICA methods. We can see from the figure that the GICs computed by CICA and the improved CICA methods have a higher correlation with the corresponding IC of each subject than with that of FastICA. These significant differences were found using *T* tests at a confidence level of 95%. Although the CCs of the improved CICA were slightly lower than those of CICA, they were not significantly different. This result indicates that the GIC calculated by the improved CICA method can better reflect the commonality of subjects in the group.

### 4 Discussion

In this study, the improved CICA method with the temporal a priori information extracted from the group data had better performance in detecting brain functional connectivity through the experimental results with the simulated and real fMRI data. First, the results in Figs. 3 and 6 show that the spatial source recovery of the improved CICA method was better than that of FastICA, but it was not as good as CICA. Second, the results in Figs. 4 and 7 show that the time courses of the corresponding sources computed by the improved CICA method were more accurate than those of the FastICA and CICA methods. Finally, the results in Figs. 5 and 8 demonstrate that the correlation between the GIC computed by the improved CICA method and the corresponding IC of each subject in the group was improved in comparison with that of FastICA, but there was no significant difference with CICA, which means that the GIC computed by the improved CICA method was more representative of the commonality of the subjects in the group.

In this paper, in order to imitate the situation of noise contained in the real fMRI data, noises with different SNRs that ranged from 0.3 to 0.4 were added to the simulated data, which were used for the evaluation of the

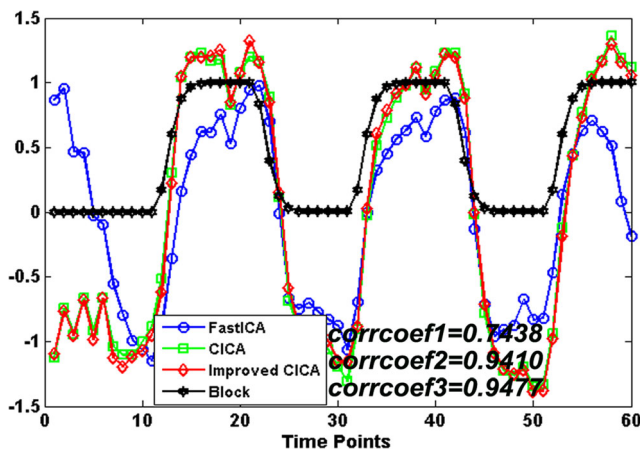


**Fig. 6** The visual areas from slices 28 to 43 are detected by FastICA, CICA, and the improved CICA methods. All the spatial maps are z-scored with the same threshold of 2

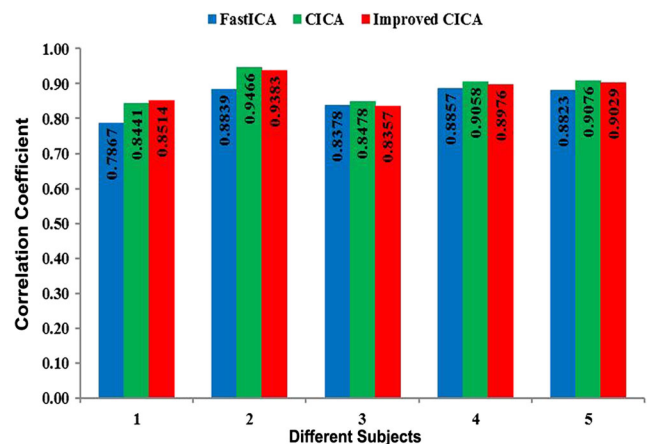
performance of the different methods. However, if the SNR of the added noise was too small, all methods showed poor performance in signal detection and thus lost the significance of evaluation. On the contrary, if the SNR of the added noise was too large, all methods produced better detection results and that there was no obvious comparability between the different methods. The SNR range adopted in this paper was a trade-off between these two situations, and it was much closer to the noise found in real data.

In the classical CICA method, the a priori reference signal was incorporated using a constraint condition  $g(y) = \varepsilon(y, r) - \xi \leq 0$ , where  $y$  denotes the output signal,  $r$  denotes a

reference signal,  $\varepsilon(y, r)$  is a distance criterion, and  $\xi$  is a threshold parameter that needs to limit the distance such that the desired output signal should be the only one satisfying the inequality constraint. However, it is difficult to predetermine the threshold parameter  $\xi$  in practical application because the ICs are blind, so the choice of a suitable  $\xi$  is quite dependent on the applied CICA. Improper  $\xi$  often leads to two possible consequences. When  $\xi$  is beyond the upper bound of the feasible range, the output may produce an undesired IC. On the other hand, when  $\xi$  is smaller than the lower bound of the range, the output cannot produce any IC. Therefore, special effort has to be made to determine a proper parameter. In this paper, the multi-objective optimization strategy was applied to estimate ICs with the CICA method, which circumvents the



**Fig. 7** The prior block (black) and time courses computed by FastICA (blue), CICA (green), and the improved CICA method (red), and the CCs between the prior block and the time course of each method, including *corrcoef1* for FastICA, *corrcoef2* for CICA, and *corrcoef3* for the improved CICA method (color figure online)



**Fig. 8** The CCs between the GIC and the corresponding IC of each subject in the group. The results were obtained using GICs computed by FastICA, CICA, and the improved CICA methods



selection of threshold parameter  $\xi$ , and the results demonstrated its improved performance.

In this paper, the weighted summation method was used to solve the multi-objective optimization problem of Eq. (6). The weight parameters “ $a_1$ ” and “ $a_2$ ” in formula (8) reflect the importance of the corresponding objective function  $f_1(\mathbf{w}_i)$  and  $\varepsilon_1(\mathbf{w}_i)$  in the summation function  $f(\mathbf{w}_i)$ . The goal of the weighted summation method is to seek a balance between the independence of the output signal and the similarity with the reference signal and then to obtain a source signal that is the closest to the reference signal with the largest independence. According to theory, applying the linear weighted summation method to solve the multi-objective optimization problem in [37] was proposed by Klamroth et al. As long as the weight parameters satisfied the conditions that they were strictly positive and added to 1, then one point of the Pareto optimal set can be found with one choice of such weights [43].

When using the linear weighted summation method to solve the multi-objective optimization problem, the summation function will contain the corresponding weight parameters that are usually determined manually according to artificial experience, and this process means that the experimental results obtained by this approach will contain many kinds of situations. Therefore, it is necessary to choose the best situation according to certain evaluation indicators from these results by additional post-processing steps. In the experiments in this paper, the weighting parameter “ $a$ ” of the improved CICA method was a value from 0.1 to 0.9 with a step length of 0.1, thus making the results include nine kinds of situations. To guarantee the consistency of the selection of the optimal weighting parameters for the simulated data and real data experiments, the index obtained by formula (14) was used to select the best situation, and the evaluation results of all situations obtained by formula (14) are shown in Appendix Tables 1 and 2, which correspond to simulated-data and real-data experiments, respectively, and are presented in Appendix 2. However, sometimes the final evaluation results may be different when adopting a different evaluation index. For example, in the simulated data experiment in this paper, if we use the average of AUCs (see Fig. 3) and CCs (see Fig. 4) to choose the optimal weighting parameter, the best situation is  $a = 0.9$  for dataset 18, which is different from the results obtained by formula (14). The evaluation results of all situations obtained by this approach are shown in Appendix Table 3, which is also presented in Appendix 2.

In addition, tensor decomposition (TD) has also shown better performance with multi-subject fMRI data analysis in recent years due to its ability to retain multi-way linkages and interactions presented in the data [44], and it can be used to obtain common spatial maps (SMs), common time courses (TCs), and subject-specific intensities [45, 46]. However, the TD method sometimes may converge to a local optimal solution because of the noise in the fMRI data, such as canonical

polyadic decomposition (CPD), which is a popular TD method. To improve the robustness of CPD with respect to noise, some additional properties have been efficiently incorporated into CPD as modality constraints [47]. For example, Beckmann and Smith proposed a solution using the statistical independence as a spatial modality constraint in TD by combining ICA with CPD [45]. Recently, Kuang et al. propose a new combined ICA and CPD method by incorporating TC delays into a CP model as the temporal constraint to obtain the shared TC, and then estimated the shared SM using a least-square fit post shift-invariant CPD [48]. Therefore, how to extract a priori information from the data itself and how to introduce it into the TD method to improve fMRI data analysis will be questions worth studying in the future.

## 5 Conclusions

In this paper, we proposed a multi-objective optimization-based improved CICA method with temporal a priori information extracted from group subject data and then used it for group fMRI data analysis. The experimental results of simulated and real fMRI data showed that the group data analysis using the improved CICA method was better than that of FastICA in both spatial and temporal domains, and it not only increased the accuracy of spatial sources and time courses but also improved the correlation of the GIC with the corresponding IC of each subject in the group. Compared to the CICA method, it only performed better temporally, and there was a slight deficiency in spatial source signal recovery. On the whole, it has its own advantages in fMRI data analysis as a blind source separation method.

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (Grants No. 31470954, No. 61271446), the Research Foundation from Shanghai Science and Technology Project (Grant No. 14590501700), the Innovation Program of Shanghai Municipal Education Commission (Grant No. 15ZZ079), the Programs for Graduate Special Endowment Fund for Innovative Developing (Grant No. 2015ycx081), and Excellent Doctoral Dissertation Cultivation (Grant No. 2015bxl005) of Shanghai Maritime University.

## Appendix 1

In this section, we provide a detailed description of extracting spatial a priori information from group data [33]. Similar to extracting temporal a priori information from group data as described in this paper, we first need to implement ICA at the single-subject level. Now assuming that we have obtained the ICs,  $S_i (i = 1, 2, \dots, K)$ , of each subject in the group using formula (6), then these ICs will be used to extract the spatial a priori information by principal component analysis (PCA). For

simplicity, we consider the situation in which each subject in the group has one IC of interest corresponding to the group IC (GIC). The correspondence of the ICs across different subjects corresponding to the GIC can be measured using the absolute value of the spatial correlation [36].

We denote the location set of voxels in the mask of subject  $i$  as  $VLS_i$  ( $i = 1, 2, \dots, K$ ), and  $s_{in_i}$  ( $i = 1, 2, \dots, K$ ) denotes the  $n_i$ th IC of subject  $i$  corresponding to the GIC of interest. Then, we can calculate the location set of common activated voxels in all  $s_{in_i}$  ( $i = 1, 2, \dots, K$ ) at the same threshold  $\theta$  and denote it as  $CAVLS$ :

$$CAVLS = \{j | abs(s_{in_i}(j)) \geq \theta, j \in VLS_i, i = 1, 2, \dots, K\} \quad (s1)$$

Let  $s_{in_i}^c$  ( $i = 1, 2, \dots, K$ ) denote the common voxels from  $s_{in_i}$  with regard to the index  $CAVLS$  where  $s_{in_i}^c$  is a column vector of size  $v \times 1$  and can be retrieved as

$$s_{in_i}^c = \{abs(s_{in_i}(j)) | j \in CAVLS\}, (i = 1, 2, \dots, K) \quad (s2)$$

Here, the absolute value in  $s_{in_i}$  is used in formula (s1) and formula (s2) due to the network of interest possibly having negative activation in the IC. Although it may mean that  $s_{in_i}^c$  ( $i = 1, 2, \dots, K$ ) contains some noise, the

spatial reference is extracted from all  $s_{in_i}^c$  ( $i = 1, 2, \dots, K$ ) by PCA, which has the ability to reduce the noise.

Now we use PCA to calculate the spatial reference signal from the  $K \times v$  matrix  $R$  which consists of all  $s_{in_i}^c$  ( $i = 1, 2, \dots, K$ ):

$$R = [s_{1n_1}^c, s_{2n_2}^c, \dots, s_{Kn_K}^c]' \quad (s3)$$

Then, the eigenvalue  $\lambda_k$  ( $k = 1, 2, \dots, K$ ) such that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K \geq 0$ , and the corresponding eigenvectors  $e_k$  ( $k = 1, 2, \dots, K$ ) of the covariance matrix  $C = E[RR']$  can be calculated, where  $e_k$  is a column vector of size  $K \times 1$ . Finally, we selected the first principal component as the spatial reference  $r$ :

$$r = e_1' R \quad (s4)$$

where  $r$  is a row vector of size  $1 \times v$  and the corresponding contribution of  $r$  can be calculated by  $c_r = \lambda_1 / \sum_{k=1}^K \lambda_k$ . In particular, if all subjects in the group have the same mask, then the spatial a priori information can be obtained directly through formulas (s3) and (s4).

### Appendix 2

**Table 1** The evaluation results of the nine situations of the improved CICA method obtained by formula (14) on the 20 datasets in the simulated-data experiment

a = 0.1	a = 0.2	a = 0.3	a = 0.4	a = 0.5	a = 0.6	a = 0.7	a = 0.8	a = 0.9	index
0.5703	0.6510	0.6611	0.6666	0.6700	0.6723	0.6739	0.6752	<b>0.6762</b>	9
0.7606	0.7715	0.7752	0.7770	0.7781	0.7788	0.7793	0.7797	<b>0.7800</b>	9
0.5493	0.5923	0.6090	0.6185	0.6248	0.6292	0.6324	0.6349	<b>0.6369</b>	9
0.6185	0.6413	0.6486	0.6521	0.6542	0.6556	0.6566	0.6573	<b>0.6579</b>	9
0.7259	0.7291	0.7301	0.7307	0.7310	0.7312	0.7313	0.7314	<b>0.7315</b>	9
0.6736	0.7375	0.7551	0.7639	0.7691	0.7726	0.7750	0.7769	<b>0.7783</b>	9
0.7352	0.7761	0.7793	0.7809	0.7819	0.7826	0.7830	0.7834	<b>0.7836</b>	9
0.6130	0.6445	0.6992	0.7137	0.7421	<b>0.7424</b>	0.7396	0.7355	0.7218	6
0.4108	0.6665	0.6949	0.7099	0.7190	0.7251	0.7295	0.7327	<b>0.7352</b>	9
0.7146	0.7332	0.7393	0.7422	0.7441	0.7453	0.7461	0.7468	<b>0.7472</b>	9
0.5980	0.6487	0.6638	0.6709	0.6751	0.6778	0.6797	0.6811	<b>0.6822</b>	9
0.6859	0.7460	0.7638	0.7730	0.7785	0.7822	0.7849	0.7868	<b>0.7883</b>	9
0.7357	0.7611	0.7695	0.7737	0.7761	0.7777	0.7789	0.7798	<b>0.7804</b>	9
<b>0.7333</b>	0.7313	0.7302	0.7294	0.7289	0.7286	0.7283	0.7281	0.7280	1
0.5475	0.6185	0.6744	0.6954	0.7084	0.7171	0.7232	0.7278	<b>0.7313</b>	9
0.5467	0.6152	0.6456	0.6574	0.6647	0.6698	0.6735	0.6763	<b>0.6785</b>	9
0.5655	0.6529	0.6684	0.6766	0.6817	0.6851	0.6876	0.6894	<b>0.6908</b>	9
0.6004	0.6835	0.6860	0.6870	<b>0.6886</b>	0.6879	0.6882	0.6883	0.6885	5
0.6429	0.6651	0.6723	0.6760	0.6781	0.6795	0.6805	0.6813	<b>0.6818</b>	9
0.6029	0.6691	0.6789	0.6841	0.6873	0.6895	0.6911	0.6923	<b>0.6932</b>	9

The bold numbers indicated the "index" values of the best situation obtained by formula (14) in each simulated dataset

**Table 2** The evaluation results of the nine situations of the improved CICA method obtained by formula (14) in the real-data experiment

a = 0.1	a = 0.2	a = 0.3	a = 0.4	a = 0.5	a = 0.6	a = 0.7	a = 0.8	a = 0.9	index
0.8976	0.8986	0.8998	0.9012	0.9030	0.9050	0.9077	0.9115	<b>0.9164</b>	9

The bold number indicated the "index" value of the best situation obtained by formula (14) in the real-data experiment

**Table 3** The average of AUCs and CCs of the nine situations of the proposed improved CICA method on the 20 datasets in the simulated-data experiment

a = 0.1	a = 0.2	a = 0.3	a = 0.4	a = 0.5	a = 0.6	a = 0.7	a = 0.8	a = 0.9	index
0.7784	0.8496	0.8591	0.8649	0.8684	0.8709	0.8724	0.8738	<b>0.8745</b>	9
0.8686	0.8801	0.8852	0.8859	0.8853	0.8859	0.8865	0.8862	<b>0.8868</b>	9
0.8060	0.8433	0.8579	0.8673	0.8725	0.8761	0.8792	0.8822	<b>0.8835</b>	9
0.8386	0.8649	0.8730	0.8767	0.8800	0.8818	0.8832	0.8841	<b>0.8846</b>	9
0.8806	0.8841	0.8853	0.8859	0.8863	0.8864	0.8862	0.8864	<b>0.8865</b>	9
0.7841	0.8406	0.8547	0.8627	0.8674	0.8704	0.8730	0.8750	<b>0.8767</b>	9
0.8446	0.8855	0.8870	0.8880	0.8889	0.8900	0.8903	0.8907	<b>0.8910</b>	9
0.7628	0.7924	0.8402	0.8554	0.8853	<b>0.8857</b>	0.8796	0.8735	0.8643	6
0.6001	0.8102	0.8377	0.8525	0.8628	0.8686	0.8729	0.8766	<b>0.8792</b>	9
0.8487	0.8663	0.8736	0.8767	0.8788	0.8794	0.8795	0.8808	<b>0.8813</b>	9
0.8097	0.8565	0.8705	0.8767	0.8808	0.8835	0.8852	0.8862	<b>0.8866</b>	9
0.7904	0.8385	0.8551	0.8636	0.8697	0.8737	0.8757	0.8781	<b>0.8798</b>	9
0.8405	0.8612	0.8695	0.8740	0.8768	0.8786	0.8799	0.8802	<b>0.8809</b>	9
<b>0.8843</b>	0.8830	0.8821	0.8834	0.8841	0.8832	0.8833	0.8833	0.8832	1
0.6949	0.7580	0.8074	0.8290	0.8446	0.8538	0.8600	0.8629	<b>0.8666</b>	9
0.7655	0.8232	0.8497	0.8600	0.8679	0.8730	0.8768	0.8799	<b>0.8833</b>	9
0.7730	0.8450	0.8581	0.8658	0.8696	0.8735	0.8751	0.8770	<b>0.8810</b>	9
0.8063	0.8772	0.8772	0.8788	0.8798	0.8793	0.8800	0.8801	<b>0.8808</b>	9
0.8341	0.8555	0.8605	0.8653	0.8675	0.8690	0.8702	0.8713	<b>0.8720</b>	9
0.8032	0.8569	0.8652	0.8719	0.8747	0.8776	0.8791	0.8797	<b>0.8815</b>	9

The bold numbers indicated the average of AUCs and CCs of the best situation in each simulated dataset

**References**

1. Logothetis NK (2008) What we can do and what we cannot do with fMRI. *Nature* 453:869–878
2. Im CH (2007) Dealing with mismatched fMRI activations in fMRI constrained EEG cortical source imaging: a simulation study assuming various mismatch types. *Med Bio Eng Comput* 45:79–90
3. Vargas ER, Mitchell DGV, Greening SG, Wahl LM (2016) Network analysis of human fMRI data suggests modular restructuring after simulated acquired brain injury. *Med Bio Eng Comput* 54:235–248
4. Li KM, Guo L, Nie JX, Li G, Liu T (2009) Review of methods for functional brain connectivity detection using fMRI. *Comput Med Imaging Graph* 33:131–139
5. Li Z, Zang YF, Ding J, Wang Z (2017) Assessing the mean strength and variations of the time-to-time fluctuations of resting-state brain activity. *Med Bio Eng Comput* 55:631–640
6. Sun F, Morris D, Babyn P (2009) The optimal linear transformation-based fMRI feature space analysis. *Med Bio Eng Comput* 47:1119–1129
7. McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998) Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp* 6:160–188
8. Zhang S, Tsai SJ, Hu S, Xu J, Chao HH, Calhoun VD, Li CR (2015) Independent component analysis of functional networks for response inhibition: Inter-subject variation in stop signal reaction time. *Hum Brain Mapp* 36:3289–3302
9. Long Z, Chen K, Wu X, Reiman E, Peng D, Yao L (2009) Improved application of independent component analysis to functional magnetic resonance imaging study via linear projection techniques. *Hum Brain Mapp* 30:417–431
10. Long Z, Li R, Hui M, Jin Z, Yao L (2013) An improvement of independent component analysis with projection method applied to multi-task fMRI data. *Comput Biol Med* 43:200–210
11. Friston KJ, Frith CD, Turner R, Frackowiak RSJ (1995) Characterizing evoked hemodynamics with fMRI. *NeuroImage* 2:157–165
12. Damoiseaux JS, Rombouts S, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF (2006) Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci U S A* 103:13848–13853
13. Mantini D, Perrucci MG, Del Gratta C, Romani GL, Corbetta M (2007) Electrophysiological signatures of resting state networks in the human brain. *Proc Natl Acad Sci U S A* 104:13170–13175
14. Calhoun VD, Kiehl KA, Pearlson GD (2008) Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. *Hum Brain Mapp* 29:828–838

15. Schmithorst VJ (2005) Separate cortical networks involved in music perception: Preliminary functional MRI evidence for modularity of music processing. *NeuroImage* 25:444–451
16. Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001) A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp* 14:140–151
17. Wang Z, Xia MG, Jin Z, Yao L, Long Z (2014) Temporally and spatially constrained ICA of fMRI data analysis. *PLoS One* 9:e94211
18. Ma X, Zhang H, Zhao X, Yao L, Long Z (2013) Semi-blind independent component analysis of fMRI based on real-time fMRI system. *IEEE Trans Neural Syst Rehabil Eng* 21:416–426
19. Liu H, Xie X, Xu S, Wan F, Hu Y (2013) One-unit second-order blind identification with reference for short transient signals. *Inf Sci* 227:90–101
20. Lu W, Rajapakse JC (2005) Approach and applications of constrained ICA. *IEEE Trans Neural Netw* 16:203–212
21. Lu W, Rajapakse JC (2006) ICA with reference. *Neurocomputing* 69:2244–2257
22. Barros AK, Vigario R, Jousmaki V, Ohnishi N (2000) Extraction of event related signals from multi-channel bioelectrical measurements. *IEEE Trans Biomed Eng* 47:583–588
23. Lin QH, Zheng YR, Yin FL, Liang H, Calhoun VD (2007) A fast algorithm for one unit ICA-R. *Inf Sci* 177:1265–1275
24. Calhoun VD, Adali T, Stevens MC, Kiehl KA, Pekar JJ (2005) Semi-blind ICA of fMRI: A method for utilizing hypothesis-derived time courses in a spatial ICA analysis. *NeuroImage* 25:527–538
25. Lin QH, Liu JY, Zheng YR, Liang H, Calhoun VD (2010) Semiblind spatial ICA of fMRI using spatial constraints. *Hum Brain Mapp* 31:1076–1088
26. Sun ZL, Shang L (2010) An improved constrained ICA with reference based unmixing matrix initialization. *Neurocomputing* 73:1013–1017
27. Li CL, Liao GS, Shen YL (2010) An improved method for independent component analysis with reference. *Digit Signal Process* 20:575–580
28. Mi JX (2014) A novel algorithm for independent component analysis with reference and methods for its applications. *PLoS One* 9:e93984
29. Mi JX, Xu Y (2014) A comparative study and improvement of two ICA using reference signal methods. *Neurocomputing* 137:157–164
30. Valente G, De Martino F, Filosa G, Balsi M, Formisano E (2009) Optimizing ICA in fMRI using information on spatial regularities of the sources. *Magn Reson Imaging* 27:1110–1119
31. Zhang ZL (2008) Morphologically constrained ICA for extracting weak temporally correlated signals. *Neurocomputing* 71:1669–1679
32. James CJ, Gibson OJ (2003) Temporally constrained ICA: an application to artifact rejection in electromagnetic brain signal analysis. *IEEE Trans Biomed Eng* 50:1108–1116
33. Shi YH, Zeng WM, Wang NZ, Chen DTL (2015) A novel fMRI group data analysis method based on data-driven reference extracting from group subjects. *Comput Methods Prog Biomed* 122:362–371
34. Bell AJ, Sejnowski TJ (1995) An information maximization approach to blind separation and blind deconvolution. *Neural Comput* 7:1129–1159
35. Hyvarinen A, Oja E (1997) A fast fixed-point algorithm for independent component analysis. *Neural Comput* 9:1483–1492
36. Du YH, Fan Y (2013) Group information guided ICA for fMRI data analysis. *NeuroImage* 6:157–197
37. Klamroth K, Tind J (2007) Constrained optimization using multiple objective programming. *J Glob Optim* 37:325–355
38. Correa N, Adali T, Li YO, Calhoun VD (2005) Comparison of blind source separation algorithms for FMRI using a new Matlab toolbox: Gift. *IEEE Int Conf Acoust Speech Signal Process* 5:401–404
39. Shi YH, Zeng WM, Wang NZ, Zhao L (2017) A new method for independent component analysis with priori information based on multi-objective optimization. *J Neurosci Methods* 283:72–82
40. Himberg J, Hyvarinen A, Esposito F (2004) Validating the independent components of neuro- imaging time series via clustering and visualization. *NeuroImage* 22:1214–1222
41. Li YO, Adali T, Calhoun VD (2007) Estimating the number of independent components for functional magnetic resonance imaging data. *Hum Brain Mapp* 28:1251–1266
42. Wang NZ, Zeng WM, Chen L (2013) SACICA: a sparse approximation coefficient-based ICA model for functional magnetic resonance imaging data analysis. *J Neurosci Methods* 216:49–61
43. Marler RT, Arora JS (2004) Survey of multi-objective optimization methods for engineering. *Struct Multidiscip Optim* 26:369–395
44. Andersen AH, Rayens WS (2004) Structure-seeking multilinear methods for the analysis of fMRI data. *NeuroImage* 22:728–739
45. Beckmann CF, Smith SM (2005) Tensorial extensions of independent component analysis for multi-subject fMRI analysis. *NeuroImage* 25:294–311
46. Kuang LD, Lin QH, Gong XF, Cong FY, Calhoun VD (2013) Multi-subject fMRI data analysis: shift-invariant tensor factorization vs. group independent component analysis. In: 2013 I.E. China summit and international conference on signal and information processing, 269–272
47. Cichocki A, Mandic D, Phan AH, Caiafa C, Zhou G, Zhao Q, Lathauwer L (2015) Tensor decompositions for signal processing applications from two-way to multiway component analysis. *IEEE Signal Process Mag* 32:145–163
48. Kuang LD, Lin QH, Gong XF, Cong F, Sui J, Calhoun VD (2015) Multi-subject fMRI analysis via combined independent component analysis and shift-invariant canonical polyadic decomposition. *J Neurosci Methods* 256:127–140

**Yuhu Shi** is working toward the Ph.D. degree in Shanghai Maritime University, Shanghai, China. His research interests include functional magnetic resonance imaging data analysis, image processing, and neuroscience computation.

**Weiming Zeng** received the Ph.D. degree from Southeast University. Now he is a full professor in Shanghai Maritime University. His research interests include image processing and analyzing, functional connectivity, and neuroscience computation.

**Xiaoyan Tang** served as a lecturer at Changshu Institute of Technology, China. Now she is working toward the Ph.D. degree in Shanghai Maritime University. Her research interests concentrate on the processing of fMRI data.

**Wei Kong** received the Ph.D. degree from Shanghai Jiaotong University. Now she works in Shanghai Maritime University. Her current research interests include the pathway and regulatory network reconstruction of dementia and cancer.

**Jun Yin** received the Ph.D. degree from Nanjing University of Science and Technology. Now he is a lecturer in Shanghai Maritime University. His current research interests include pattern recognition, face recognition, and machine learning.