Extraction and Automatic Grouping of Joint and Individual Sources in Multisubject fMRI Data Using Higher Order Cumulants

Mansooreh Pakravan and Mohammad Bagher Shamsollahi, *Senior Member, IEEE*

***Abstract*—The joint analysis of multiple data sets to extract their interdependency information has wide applications in biomedical and health informatics. In this paper, we propose an algorithm to extract joint and individual sources of multisubject data sets by using a deflation-based procedure, which is referred to as joint/individual thin independent component analysis (JI-ThICA). The proposed algorithm is based on two cost functions utilizing higher order cumulants to extract joint and individual sources. Joint sources are discriminated by fusing signals of all subjects, whereas individual sources are extracted separately for each subject. Furthermore, JI-ThICA algorithm estimates the number of joint sources by applying a simple and efficient strategy to determine the type of sources (joint or individual). The algorithm also categorizes similar sources automatically across data sets through an optimization process. The proposed algorithm is evaluated by analyzing simulated functional magnetic resonance imaging (fMRI) multisubject data sets, and its performance is compared with existing alternatives. We investigate clean and noisy fMRI signals and consider two source models. Our results reveal that the proposed algorithm outperforms its alternatives in terms of the mean joint signal to interference ratio. We also apply the proposed algorithm on a public-available real fMRI multisubject data set, which was acquired during naturalistic auditory experience. The extracted results are in accordance with the previous studies on naturalistic audio listening and results of a recent study investigated this data set, which demonstrates that the JI-ThICA algorithm can be applied to extract reliable and meaningful information from multisubject fMRI data.**

***Index Terms*—Brain Signals, functional magnetic resonance imaging (fMRI), joint and individual source extraction, multi-subject data analysis, thin independent component analysis (Thin ICA).**

I. Introduction

A. Background and Motivation

IN BIOMEDICAL signal processing, joint analysis of multiple datasets such as multi-subject and multi-modal datasets introduces new degrees of freedom in analyzing datasets, and has wide applications in biomedical and health informatics, e.g., medical diagnosis and eHealth [1]. In traditional techniques, data of each subject is separately analyzed, and then its results are subsequently integrated with results of other subjects. However, recent researches have demonstrated that this approach is suboptimal, because it does not allow the direct interactions among information of multiple subjects. It seems that performance gain of joint analysis is significant, if datasets contain joint information about a specific process. Furthermore, data-driven fusion of multiple datasets provides better understanding of complex processes such as brain signals. Interestingly, diverse multi-subject and multi-modal datasets are nowadays available thanks to advances in brain-signal acquisition systems.

The main challenges in fusion of multiple datasets are the extraction of useful and relevant information and the method of fusing datasets. This is because underlying process depends on large number of unknown variables. It has been shown that various blind source separation (BSS) techniques [2] are useful for data-driven fusion of multiple datasets [1].

In [3], the general BSS sub-problems have been categorized into four classes, based on the number of analyzed datasets and dependency of sources. The first class is single dataset unidimensional (SDU), which analyzes a single dataset whose sources are uncorrelated (independent). For example, the components extracted with principal component analysis (PCA) method, sources of independent component analysis (ICA) and second order blind identification (SOBI) methods [2] have SDU model. The second class is multiple dataset unidimensional (MDU), which analyzes more than one dataset. In this class, sources in each dataset are uncorrelated (independent), and each source has exactly one corresponding correlated (dependent) source in other datasets. The source model in canonical correlation analysis (CCA) [4] and [5], common feature analysis method [6], generalized joint diagonalization of cumulant matrices [7], joint analysis of multiple datasets by cross cumulant tensor block diagonalization [8], group information guided ICA (GIG-ICA) [9] and independent vector analysis (IVA) [10], [11], [12] are categorized in the MDU class.

The third class is single dataset multidimensional (SDM), which analyzes a single dataset with one or more group of sources, where sources of each group are dependent. The source model of multidimensional independent component analysis (MICA) [13] and independent subspace analysis (ISA) [14] are in the SDM class. Finally, the fourth class is multiple dataset

Manuscript received December 4, 2017; revised April 6, 2018; accepted May 9, 2018. Date of publication May 24, 2018; date of current version March 6, 2019. *(Corresponding author: Mansooreh Pakravan.)*

The authors are with the Biomedical Signal and Image Processing Laboratory, Department of Electrical Engineering, Sharif University of Technology, Tehran 11356-9363, Iran (e-mail: mpakravan@ee.sharif.edu; mbshams@sharif.edu).

Digital Object Identifier 10.1109/JBHI.2018.2840085

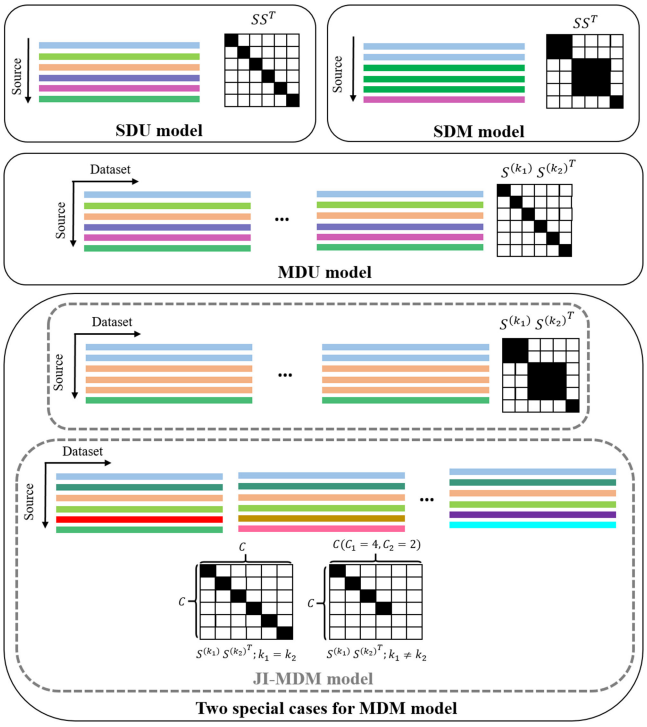

Fig. 1. Four source models with their source correlation pattern. For MDM model only two special cases are presented. Each colorful line indicates a source, and lines with similar colors mean joint (dependent or correlated) sources. Black and white squares in the source correlation matrices mean non-zero and zero correlation, respectively.

multidimensional (MDM), which deals with more than one dataset, where sources in all datasets are organized into a number of groups. Sources in each group are dependent and it is possible that some of groups are joint across all, some or just one of datasets. The source model of multi-dataset independent subspace analysis (MISA) [15] and joint independent subspace analysis (JISA) [16] are examples of the MDM source model.

It is worth mentioning that the SDU model is a special case of MDU, SDM, and MDM, while MDU and SDM models are special cases of MDM [3].

Fig. 1 shows the explained source models with their source correlation pattern for SDU, MDU, SDM and MDM models. Note that only two special cases are shown for MDM source model. In Fig. 1, dependent signals have the same color. In addition, in the shown source correlation matrices, black and white squares mean non-zero and zero, respectively.

Real brain datasets, e.g., brain signal of a number of subjects in the same experiment, are not exactly similar in all regions across subjects. This is partly due to the fact that there are some active processes in brain of each subject that are independent of the experiment effect. On the other hand the anatomical and functional landmarks of individual subjects change within and across subjects even after spatial normalization. Thus, to improve the accuracy of the dataset model, a number of independent (individual) sources should be considered per datasets [17], [18] according to MDM model. Hence, we can have a model in which each dataset is categorized into two parts, namely joint sources and individual sources, where the joint sources in the first part have exactly one equivalent (similar) source across datasets, the same as the MDU model, while the sources in the second part are independent. Intuitively, the joint sources help to discover potential relationships between the members of a dataset, while the individual sources can help to investigate unique information in each member of the dataset. This model is a special case of MDM. In this paper, we refer to this source model as joint/individual MDM (JI-MDM). In Fig. 1, the JI-MDM source model is shown as the second special case for MDM model. It is worth mentioning that the source model of recorded datasets in many biomedical and health informatics applications is JI-MDM, thus, designing efficient algorithms for JI-MDM source model is of utmost importance, which is the main contribution of this paper.

It should be noted that in this paper, each dataset is brain signals of one subject and multi-dataset means data of multiple subjects in the same experiment.

B. Related Works

Recently, in [17] and [18], the JI-MDM source model has been investigated. In [17], the so called joint and individual variation explained (JIVE) method has been introduced to simultaneously extract both joint and individual sources across the members of a multi-dataset. The JIVE method represents an extension of PCA and extracts joint variations between datasets. In [18], two algorithms, namely COBE and COBEC, have been proposed to extract common orthogonal basis in JI-MDM source model. The COBE and COBEC algorithms deal with unknown and known number of common components, respectively, where both algorithms incorporate dimensionality reduction and blind source separation techniques. Moreover, in [18] another method, referred to as common nonnegative features extraction (CNFE), has been introduced, in which low-rank approximation-based (semi) nonnegative matrix factorization model is employed for nonnegative latent sources. It should be noted that the JIVE, COBEC and CNFE are generally referred to as *linked BSS* in which these methods are combined with a BSS method and have a direct connection to group PCA and group ICA [18].

The aforementioned algorithms are also applicable for the MDU source model, if there is no individual sources. Thus, it is worth to compare their performance with those algorithms designed for the MDU problem, such as IVA, and Group ICA [19]. In [20] and [21] some of algorithms which are applicable for MDU source model have been reviewed and their applications for biomedical data analysis have been explained. It is worth mentioning that the novelty of our algorithm compared with the algorithms introduced in [20] and [21] is that our algorithm is applicable for both JI-MDM and MDU source models, whereas algorithms in [20] and [21] can be used only for MDU source model.

In [22], Group ICA algorithm has been applied to analyze simulated fMRI dataset under conditions of spatial, temporal, and amplitude variability. In [23], a graph-theoretical analysis has been applied on IVA algorithm to analyze its ability in capturing group variability of simulated fMRI. In both [22] and [23], the same toolbox, referred to as SimTB [24], has been utilized

to generate fMRI datasets. In this paper, we will use the SimTB toolbox to generate simulated multi-subject fMRI dataset.

In [25], the thin ICA algorithm has been proposed by using a proxy non-Gaussianity measure as cost function. The proposed algorithm in [25] is applicable for SDU source model in which a single dataset with independent sources is analyzed. In this paper, we modify the cost function of [25] and present a new algorithm to extract both joint and individual sources of a multi-subject dataset with the JI-MDM source model. Thus, the novelty of our algorithm compared with [25] is that it extracts both joint and individual sources of multi-subject datasets in a single optimization problem, whereas [25] has been designed to analyze single-subject datasets with SDU model in which sources are independent.

C. Contribution and Paper Organization

The main contribution of this paper is to present a new algorithm to analyze multi-subject datasets with either JI-MDM or MDU source models. The proposed algorithm is based on two cost functions to extract joint and individual sources, where joint sources are obtained by fusing signals of all subjects, whereas individual sources are extracted separately for each subject. This algorithm estimates the number of joint sources by applying a simple and efficient strategy to determine the type of sources (joint or individual). Furthermore, it groups similar sources automatically across datasets through an optimization process in order to solve permutation indeterminacy across datasets. It is worth mentioning that in SDU and SDM problems, the permutation indeterminacy is related to the order of group sources arrangement and is generally unimportant, but sometimes in MDU and MDM problems, the order of extracted sources in each dataset is important. In other words, it is desirable to assign the same source index to joint sources across dataset and solve permutation indeterminacy across datasets.

As mentioned, we compare our algorithm with the existing alternatives, which have been recently introduced for the JI-MDM dataset analysis, [17] and [18]. To do so, we generate clean and noisy fMRI datasets by using SimTB toolbox [24], where in noisy fMRI datasets small variations are applied in the location, rotation and spread of joint sources across subjects, and then a white Gaussian noise is added to the observations. Our results reveal that the proposed algorithm has a better performance in terms of the estimated number of correct join sources (estimated C_1) and mean joint signal to interference ratio (jSIR) for noisy observations.

The novelty of our algorithm compared with its alternatives ([17] and [18]) is that it extracts simultaneously joint and individual sources in a single optimization problem by exploiting higher order cumulants. It is worth mentioning that the algorithm introduced in [17] (JIVE) extracts the sources dominated by variances rather than correlations or higher order cumulants. As a consequence, JIVE may fail to extract the joint sources with high correlations, especially when the joint sources are relatively weak but are consistently present in all the datasets. In addition, our numerical results demonstrate significant advantages of the proposed algorithm in comparison with CNFE [18], where it extracts correctly joint and individual sources of both clean and noisy fMRI observations, whereas CNFE fails to extract joint sources of noisy fMRI observations. This is because in CNFE joint sources are determined by minimizing the error of a factorization problem based on alternating least square (ALS) method, which its performance is degraded with Gaussian noise. However, in our algorithm joint sources are discriminated based on higher order cumulants, which suppress Gaussian noise.

Furthermore, as the proposed algorithm can be used for MDU dataset model, we compare our algorithm with Group ICA [19], GIG-ICA [9], and the IVA algorithm introduced in [26] (as three methods for MDU dataset model). Our results show that the proposed algorithm outperforms Group ICA and GIG-ICA for both clean and noisy observations. In addition, our algorithm outperforms the IVA algorithm for clean observations, while it has approximately the same performance as IVA for noisy observations.

Finally, we apply our algorithm on a public-available multi-subject fMRI dataset obtained from the Study Forrest project [27]. This dataset contains fMRI data of 20 subjects acquired during naturalistic auditory experience. The obtained results of the JI-ThICA algorithm for this dataset are in accordance with the previous studies on naturalistic audio listening [28], [29] as well as results of a recent study investigated this dataset [30], which demonstrates the capability of the proposed JI-ThICA to extract reliable and meaningful information from multi-subject fMRI dataset.

The rest of the paper is organized as follows. In Section II, the signal model and our notations are presented. The proposed algorithm is explained in Section III. In Section IV, numerical results are presented. Finally, Section V is devoted to conclude the paper and present the future works of our study.

II. Signal Model

In our JI-MDM model, C_1 independent components are considered for each dataset, which are joint sources across datasets, (i.e., these sources are dependent across all datasets, the same as the MDU model). In addition, C_2 individual components are considered in each dataset, which are analyzed separately, and C is the total number of sources in each dataset ($C = C_1 + C_2$). Furthermore, K, V, and N are the number of datasets, the number of voxels within a slice of fMRI image (or the number of time points in a EEG signal), and the number of observation time points in a BOLD signal (or the number of channels in a EEG signal), respectively. It should be noted that the vectorized version of the fMRI source images (or volumes) are processed in the algorithm.

Based on these notations, the introduced source models in Fig. 1 can be interpreted as follows. The JI-MDM is reduced to SDU and MDU for the cases of ($K = 1$) and ($K > 1$ and $C_2 = 0$), respectively. Furthermore, if $K > 1$ and $C_1 = 0$, JI-MDM represents K separate and independent SDU source models.

Let $X^{(k)}$ denote the observation signal of kth subject, given by

$$X^{(k)} = A^{(k)} S^{(k)}, \quad k = 1, ..., K. \tag{1}$$

where $A^{(k)} \in \mathbb{R}^{N\times C}$ is the mixing matrix; $S^{(k)} = [\mathbf{s}_1^{(k)^T}, \mathbf{s}_2^{(k)^T}, ..., \mathbf{s}_C^{(k)^T}]^T$ is the source matrix; $\mathbf{s}_c^{(k)} \in \mathbb{R}^{1\times V}$ for $1 \leq c \leq C$ is c'th source component. Here, the superscripts and subscripts denote the indices of datasets and sources, respectively.

In this paper, we assume that the components of $S^{(k)}$ are mutually independent, locally stationary and normalized to zero mean and unit variance, and the mixing matrices $A^{(k)}; k = 1, ..., K$ are full-column rank.

Moreover, cumulants and cross-cumulants are used to extract latent sources, where the second, third and fourth cumulants of real zero-mean stochastic random variables y_i are calculated as follows [2]:

$$\text{cum}\,[y_i] = E\{y_i\} = 0; \tag{2a}$$

$$\text{cum}\,[y_i, y_j] = E\{y_i, y_j\}; \tag{2b}$$

$$\text{cum}\,[y_i, y_j, y_k] = E\{y_i, y_j, y_k\}; \tag{2c}$$

$$\begin{aligned}\text{cum}\,[y_i, y_j, y_k, y_l] &= E\{y_i, y_j, y_k, y_l\} - E\{y_i, y_j\}E\{y_k, y_l\} \\ &\quad - E\{y_i, y_k\}E\{y_j, y_l\} - E\{y_i, y_l\}E\{y_j, y_k\};\end{aligned} \tag{2d}$$

where the operator E{.} is the expected value. It is worth noting that the cross-cumulant of mutually statistically independent (MSI) random variables is zero [2], i.e.,

$$\text{cum}\,[y_1, y_2, ..., y_K] = 0, \text{ if } y_1, y_2, ..., y_K \text{ are MSI.} \tag{3}$$

Furthermore, if two random variables are related by a linear transform, $\mathbf{y} = A\mathbf{x}$, then their cumulants are related multilinearly as follows [2]

$$\text{cum}\,[y_i, y_j, ..., y_k] = \sum_{p,q,...,r} A_{i,p}A_{j,q}...A_{k,r}\,\text{cum}\,[x_i, x_j, ..., x_k]. \tag{4}$$

For MDU and JI-MDM source models, $\text{cum}\left[S^{(k_1)}, S^{(k_2)}\right]$, $\text{cum}\left[S^{(k_1)}, S^{(k_2)}, S^{(k_3)}\right]$ and $\text{cum}\left[S^{(k_1)}, S^{(k_2)}, S^{(k_3)}, S^{(k_4)}\right]$ are all superdiagonal tensors in $\mathbb{R}^{C\times C}$, $\mathbb{R}^{C\times C\times C}$, and $\mathbb{R}^{C\times C\times C\times C}$, respectively, which are given by

$$\left(\text{cum}\left[S^{(k_1)}, S^{(k_2)}\right]\right)_{(c,c)} = \frac{1}{V}\sum_{v=1}^{V}\text{cum}\left[s_c^{(k_1)}(v), s_c^{(k_2)}(v)\right] \tag{5}$$

$$\begin{aligned}&\left(\text{cum}\left[S^{(k_1)}, S^{(k_2)}, S^{(k_3)}\right]\right)_{(c,c,c)} \\ &\quad = \frac{1}{V}\sum_{v=1}^{V}\text{cum}\left[s_c^{(k_1)}(v), s_c^{(k_2)}(v), s_c^{(k_3)}(v)\right]\end{aligned} \tag{6}$$

$$\begin{aligned}&\left(\text{cum}\left[S^{(k_1)}, S^{(k_2)}, S^{(k_3)}, S^{(k_4)}\right]\right)_{(c,c,c,c)} \\ &\quad = \frac{1}{V}\sum_{v=1}^{V}\text{cum}\left[s_c^{(k_1)}(v), s_c^{(k_2)}(v), s_c^{(k_3)}(v), s_c^{(k_4)}(v)\right]\end{aligned} \tag{7}$$

where $s_c^{(k)}(v)$ is the (c, v)th element of $S^{(k)}$. Furthermore, for JI-MDM model, we have

$$\begin{aligned}&\text{cum}\left[s_c^{(k_1)}(v), ..., s_c^{(k_i)}(v)\right] \neq 0 \text{ if } (k_1 = ... = k_i) \text{ or } (c \in \mathcal{J}) \\ &\text{cum}\left[s_c^{(k_1)}(v), ..., s_c^{(k_i)}(v)\right] = 0 \text{ if } (\exists k_{i_1} \neq k_{i_2}) \text{ and } (c \in \mathcal{I})\end{aligned} \tag{8}$$

where $\mathcal{J}$ and $\mathcal{I}$ are the set of joint and individual sources, respectively, $i = 2, 3, 4$ and $i_1, i_2 = 1, ..., 4$.

III. Proposed Algorithm

In the proposed algorithm, joint and individual sources across multiple datasets are extracted by employing a deflation framework [2]. In this algorithm, sources of each dataset are extracted one-by-one, and two cost functions, namely joint thin ICA (J-ThICA) and individual thin ICA (I-ThICA), are employed to extract joint and individual sources, respectively.

The proposed algorithm has two main steps. The first step is preprocessing to prepare the observations for the analysis, and the second one is maximizing a cost function to find the optimum source component in a deflation based framework. In the two following subsections, the aforementioned steps are described, then, in subsection III-C a method is proposed to determine the type of extracted sources, joint or individual.

A. Preprocessing

In order to extract independent components, first we apply PCA to reduce the dimension of the observations to C principle components ($W_{pca}^{(k)}$). The PCA also helps to reduce the computational complexity and artifacts in observations. It should be noted that any other dimensional reduction algorithms can be used here; however, the selected procedure should give the desired information of the signals.

After dimensional reduction, a pre-whitening system ($W_w^{(k)} = {R_{X^{(k)}}}^{-1/2}$) is employed to obtain the preprocessed observations matrices $Z^{(k)}$ [2], where $R_{X^{(k)}}$ is the correlation matrix of $X^{(k)}$.

The number of principle components C (joint and individual sources, $C_1 + C_2$) can be estimated by using the model order selection approaches based on information-theoretic criteria such as Akaike information criterion (AIC) [31] and Bayesian information criterion (BIC) [32]. In this paper, we use BIC to estimate C in each subject. Furthermore, we assume that the number of independent sources are the same across datasets. After estimating C in each dataset, we choose the estimated number with maximum repetition (using a voting approach). It is worth mentioning that there are also other methods for joint estimation of C [33], which can be used alternatively.

B. Optimizing Cost Function

In the following, two cost functions are presented to extract joint and individual sources.

Cost function for individual sources (I-ThICA): We have selected higher order cumulants and cross-cumulants as the

base of our cost functions, because, the sources distributions do not affect the local convergence of higher order cumulants for sources with non-zero cumulant [34]. To this aim, we need non-Gaussianity in independence-based decompositions. Since gaussian sources have zero-valued higher-order cumulants, maximizing the cost functions based on higher-order cumulants leads to non-Gaussianity.

We inspire our cost function from the thin ICA [25], in which the cost function is a proxy non-Gaussianity measure obtained by combining higher order cumulant matrices. The Thin ICA has a multivariate contrast function for the estimation of independent sources from a linear mixture.

The thin ICA processes each dataset individually, and its cost function for kth dataset is given by [25]

$$\Psi_\Phi(y_c^{(k)}(v)) = \sum_{\eta\in\Phi} w_\eta \left|\text{cum}\left[y_c^{(k)}(v-\upsilon_1), y_c^{(k)}(v-\upsilon_2), ..., y_c^{(k)}(v-\upsilon_\eta)\right]\right|^2, \tag{9}$$

where the operator $|.|^2$ is squared norm, Φ is a set containing the order of used cumulants, $w_\eta > 0$ denotes the weight of η'th order cumulant, $v = 1, ..., V$ is the pixel or voxel number, $y_c^{(k)}(v) = \mathbf{u}_c^{(k)} Z^{(k)}(v)$ is the estimation of the source, in which $\mathbf{u}_c^{(k)}$ is the estimation of c'th raw of the inverse of mixing matrix, and $Z^{(k)}(v) \in \mathbb{R}^{C\times 1}$ is the observation in v'th voxel. Note that $y_c^{(k)}(v-\upsilon_m)$ for $m = 1, ..., \eta$ is υ_m delayed version of $y_c^{(k)}(v)$. Hereafter, $\Psi_\Phi\left(y_c^{(k)}(v)\right)$ is referred to as individual thin ICA (I-ThICA).

In [25], in order to maximize I-ThICA, the cost function has been optimized with respect to the first source in $cum[.]$ operator while keeping fixed the other sources (even if some of the other sources is equal to the first source). Furthermore, it has been shown that in a deflation procedure, the solution which extracts one of the independent sources is a local maximum of the thin ICA contrast function (Theorem 1 in [25]).

In this paper, we optimize I-ThICA with respect to $\mathbf{u}_c^{(k)}$ and accordingly estimate the desired source ($y_c^{(k)^*}(v) = \mathbf{u}_c^{(k)^*} Z^{(k)}(v)$, where $\mathbf{u}_c^{(k)^*}$ is the optimum value of $\mathbf{u}_c^{(k)}$).

Cost function for joint sources (J-ThICA): In this case, we modify I-ThICA to utilize the information of similar sources in other datasets. The proposed cost function to extract the cth source of k_1th dataset is formulated as follows

$$\Upsilon_\Phi(y_c^{(k_1)}(v)) = \sum_{\eta\in\Phi} w_\eta \left|\text{cum}\left[y_c^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \tag{10}$$

where $k_j \in \{1, ..., K\}/\{k_1\}$ for $j = 2, ..., \eta$. In the following, this cost function is referred to as joint thin ICA (J-ThICA). In each iteration of maximizing J-ThICA cost function, $\eta - 1$ datasets are selected randomly from $\{1, ..., K\}/\{k_1\}$, which are indexed as $k_2, ..., k_\eta$ datasets.

Joint/individual thin ICA (JI-ThICA): In order to solve the JI-MDM problem, we need to extract both joint and individual sources of the dataset with one algorithm. To this aim, we propose the following algorithm for the cth source in the k_1th dataset ($k_1 = 1, ..., K$ and $c = 1, ..., C$).

$$\mathbf{u}_c^{(k_1)^*} = \begin{cases} \underset{\mathbf{u}_c^{(k_1)}}{\text{argmax}}\, \Upsilon_\Phi(y_c^{(k_1)}(v)), & \text{if } (c \in \mathcal{J}) \\ \quad s.t.: \left|\mathbf{u}_c^{(k_1)}\right|^2 = 1 & \\ \underset{\mathbf{u}_c^{(k_1)}}{\text{argmax}}\, \Psi_\Phi(y_c^{(k_1)}(v)), & \text{if } (c \in \mathcal{I}) \\ \quad s.t.: \left|\mathbf{u}_c^{(k_1)}\right|^2 = 1 & \end{cases} \tag{11}$$

In this algorithm, we need to know the type of source c, joint ($\mathcal{J}$) or individual ($\mathcal{I}$). In subsection III-C, we propose a method to determine the type of sources. Furthermore, the cost functions $\Psi_\Phi(y_c^{(k_1)}(v))$ and $\Upsilon_\Phi(y_c^{(k_1)}(v))$ must have different nonzero values for real sources to correctly extract them [25]. It means that there are two permutation sets α and δ of the indexes $\{1, ..., C\}$ in such a way that the following inequalities are satisfied

$$\Upsilon_\Phi(y_{\delta_1}^{(k_1)^*}(v)) > \Upsilon_\Phi(y_{\delta_2}^{(k_1)^*}(v)) > ... > \Upsilon_\Phi(y_{\delta_C}^{(k_1)^*}(v)), \quad k_1 = 1, ..., K \tag{12}$$

$$\Psi_\Phi(y_{\alpha_1}^{(k_1)^*}(v)) > \Psi_\Phi(y_{\alpha_2}^{(k_1)^*}(v)) > ... > \Psi_\Phi(y_{\alpha_C}^{(k_1)^*}(v)), \quad k_1 = 1, ..., K. \tag{13}$$

In order to maximize these cost functions and extract desired sources, we use [25, Th. 1]. For I-ThICA, the theorem in [25] is directly applied, however, we need to modify it for the case of J-ThICA. In the following, the modified theorem is presented, which shows that at the global maximum of J-ThICA cost function, the output vector is the desired output.

Theorem 1: Considering the extracting vector $\mathbf{u}_c^{(k_1)} \in \mathbb{R}^{1\times C}$ and the output source $y_c^{(k_1)}(v)$ for $c = 1, ..., C$ and $k_1 = 1, ..., K$, at the global maximum of $\Upsilon_\Phi(y_c^{(k_1)}(v))$ subject to $|\mathbf{u}_c^{(k_1)}|^2 = 1$, $y_c^{(k_1)^*}(v)$ represents the desired source $s_c^{(k_1)}(v)$, i.e., $y_c^{(k_1)^*}(v) = s_c^{(k_1)}(v)$.

Proof: If we define $G^{(k)} = (U^{(k)} W_w^{(k)} W_{pca}^{(k)}) A^{(k)}$ and $g_{c_1,c_2}^{(k)} = [G^{(k)}]_{c_1,c_2}$, we have $y_c^{(k_1)}(v) = \sum_{i=1}^C g_{c,i}{}^{(k_1)} s_i^{(k_1)}(v)$. Then, by using (4), we have

$$\begin{aligned} &\left|\text{cum}\left[y_c^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \\ &= \left|\sum_{i_1=1}^C \sum_{i_2=1}^C \cdots \sum_{i_\eta=1}^C g_{c,i_1}^{(k_1)} g_{c,i_2}^{(k_2)} ... g_{c,i_\eta}^{(k_\eta)} \right. \\ &\qquad \left. \cdot\, \text{cum}\left[s_{i_1}^{(k_1)}(v-\upsilon_1), ..., s_{i_\eta}^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \end{aligned} \tag{14}$$

Then, by using Cauchy-Schwarzs inequality and (3), the following relation is obtained

$$\begin{aligned}&\left|\text{cum}\left[y_c^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \\ &\le \sum_{i_1=1}^{C}\left|g_{c,i_1}^{(k_1)}\right|^2 \cdot \left(\sum_{i_1=1}^{C}\prod_{j=2}^{\eta}\left|g_{c,i_1}^{(k_j)}\right|^2\right) \\ &\quad \cdot \left|\text{cum}\left[s_{i_1}^{(k_1)}(v-\upsilon_1), ..., s_{i_1}^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \end{aligned} \tag{15}$$

Since the rows of $G(k)$ are normalized, we have

$$\left(\sum_{i_1=1}^{C}\prod_{j=2}^{\eta}\left|g_{c,i_1}^{(k_j)}\right|^2\right) \le \prod_{j=2}^{\eta}\left(\sum_{i_1=1}^{C}\left|g_{c,i_1}^{(k_j)}\right|^2\right) = 1 \tag{16}$$

Thus, (15) can be rewritten as

$$\begin{aligned}&\left|\text{cum}\left[y_c^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \\ &\le \sum_{i_1=1}^{C}\left|g_{c,i_1}^{(k_1)}\right|^2 \left|\text{cum}\left[s_{i_1}^{(k_1)}(v-\upsilon_1), ..., s_{i_1}^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \end{aligned} \tag{17}$$

Using (17) in the J-ThICA cost function, we have

$$\begin{aligned}&\Upsilon_\Phi\left(y_c^{(k_1)}(v)\right) = \\ &\sum_{\eta\in\Phi} w_\eta \left|\text{cum}\left[y_c^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \\ &\le \sum_{\eta\in\Phi} w_\eta \sum_{i_1=1}^{C}\left|g_{c,i_1}^{(k_1)}\right|^2 \left|\text{cum}\left[s_{i_1}^{(k_1)}(v-\upsilon_1), ..., s_{i_1}^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \\ &\le \sum_{i_1=1}^{C}\left|g_{c,i_1}^{[k_1]}\right|^2 \Upsilon_\Phi(s_{i_1}^{(k_1)}(v-\upsilon_1)). \end{aligned} \tag{18}$$

From the ordering condition in (12), we have $\Upsilon_\Phi(s_{\delta_1}^{(k_1)}(v-\upsilon_1)) > \Upsilon_\Phi(s_{\delta_j}^{(k_1)}(v-\upsilon_1))\ \forall j = 2, ..., C$. Furthermore, for the case of correct extraction, the term $U^{(k)}W_w^{(k)}W_{pca}^{(k)}$ is inverse of the mixing matrix, $A^{(k)}$, which giving rise to $G^{(k)} = I_{C\times C}$. Under these conditions, the J-ThICA cost function is upper bounded by $\Upsilon_\Phi(y_c^{(k_1)}(v-\upsilon_1)) < \Upsilon_\Phi(s_{\delta_1}^{(k_1)}(v-\upsilon_1))$, which means that $\Upsilon_\Phi(y_c^{(k_1)}(v-\upsilon_1))$ extracts the desired independent component for $\upsilon_1 = 0$ in its global maximum, i.e. $y_{\delta_1}^{(k_1)^*}(v) = s_{\delta_1}^{(k_1)}(v)$. ■

In the following, we solve the optimization problem in (11). For simplicity, the J-ThICA cost function is rewritten as

$$\begin{aligned}&\Upsilon_\Phi(y_c^{(k_1)}(v)) = \\ &\sum_{\eta\in\Phi} w_\eta \left|\text{cum}\left[y_c^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]\right|^2 \\ &\qquad = \mathbf{u}_c^{(k_1)}\tilde{M}_{v,c}^{(k_1)}\mathbf{u}_c^{(k_1)^T}, \end{aligned} \tag{19}$$

in which

$$\tilde{M}_{v,c}^{(k_1)} = \sum_{\eta\in\Phi} w_\eta \tilde{C}_{v,c}^{(k_1)}(\eta)\left(\tilde{C}_{v,c}^{(k_1)}(\eta)\right)^T, \tag{20}$$

where

$$\begin{aligned}&\tilde{C}_{v,c}^{(k_1)}(\eta) \\ &= \text{cum}\left[Z^{(k_1)}(v-\upsilon_1), y_c^{(k_2)}(v-\upsilon_2), ..., y_c^{(k_\eta)}(v-\upsilon_\eta)\right]. \end{aligned} \tag{21}$$

In (20), we observe that $\tilde{M}_{v,c}^{(k_1)}$ is independent of $\mathbf{u}_c^{(k_1)}$, thus, the maximum of $\Upsilon_\Phi(y_c^{(k_1)})$ is computed by obtaining the corresponding eigenvector of the dominant eigenvalue of $\tilde{M}_{v,c}^{(k_1)} \in \mathbb{R}^{C\times C}$, where the obtained eigenvector is the optimum value of $\mathbf{u}_c^{(k_1)}$.

In [25], the maximum of the I-ThICA function has been obtained by computing the eigenvector associated to the dominant eigenvalue of $M_{v,c}^{(k)} = \sum_{\eta\in\Phi} w_\eta C_{v,c}^{(k)}(\eta)(C_{v,c}^{(k)}(\eta))^T$ in which $C_{v,c}^{(k)}(\eta) = \text{cum}[Z^{(k)}(v-\upsilon_1), y_c^{(k)}(v-\upsilon_2), ..., y_c^{(k)}(v-\upsilon_\eta)]$.

It is worth mentioning that since sometimes observation process is large term non-stationary, we compute $\mathbf{u}_c^{(k_1)^*}$ in each voxel v. But we can write the I-ThICA and J-ThICA cost functions in batch mode for $\upsilon_m = 0,\ \forall m = 1, .., \eta$. For example for J-ThICA cost function, we have

$$\tilde{C}_c^{(k_1)}(\eta) = \text{cum}\left[Z^{(k_1)}, \mathbf{y}_c^{(k_2)}, ..., \mathbf{y}_c^{(k_\eta)}\right]. \tag{22}$$

The proposed JI-ThICA algorithm is presented in Algorithm 1, which is based on deflation procedure. Note that the algorithm is written in batch model (simultaneously for all voxels not for an individual voxel v).

In Algorithm 1, the operator $PIM(.)$ means power iteration method which computes the eigenvector associated to the dominant eigenvalue of its input matrix. There are some iterative standard eigenvalue and eigenvector finding methods. We use power iteration method [35] to obtain the corresponding eigenvector of the largest eigenvalue. Since $\tilde{M}_c^{(k_1)}$ and $M_c^{(k)}$ are symmetric matrices, the power iteration method generally converges twice as rapidly as an asymmetric matrix.

The maximization of I-ThICA and J-ThICA are performed in an iterative procedure, where the iteration stops if $1 - |\mathbf{u}_{(c)}^{(k_1)}\mathbf{u}_{c_{old}}^{(k_1)^T}|^2 \le \epsilon_0$, (e.g., $\epsilon_0 = 10^{-4}$). We call this iteration as *inner iteration*. We also use an *outer iteration* procedure for JI-ThICA algorithm such that in each iteration all optimum sources of all datasets are extracted. The outer iteration is executed $MaxIter$ times.

C. Determining the Type of Sources

In each step of JI-ThICA algorithm, we can determine whether the extracted independent source is joint across datasets or not. To do so, we compute $\Upsilon_\Phi(\mathbf{y}_c^{(k)^*})$ (which is a measure of cross-cumulant) in each inner iteration of extracting c'th source in k'th dataset. The main reason that cross-cumulants can be used in joint and individual sources discrimination stems from the fundamental property of cumulants, which says that the

Algorithm 1: JI-ThICA.

1 Determine the number of principle components C by using BIC methods;
2 Dimension reduction by using PCA;
3 Initialize source matrices randomly;
4 $i = 1$;
5 **Outer Iteration: while** $i < MaxIter$ **do**
6 **for** $\forall\ c = 1, ..., C$ **do**
7 **for** $\forall\ k = 1, ..., K$ **do**
8 Pre-whiten $Z^{(k)}$;
9 $\mathbf{y}_c^{(k)} = Z_c^{(k)}$;
10 $\varepsilon_0 = 1$;
11 **Inner Iteration: while** $\varepsilon_0 > 10^{-4}$ **do**
12 Compute $\Upsilon_\Phi\left(\mathbf{y}_c^{(k)}\right)$;
13 **if** $\Upsilon_\Phi\left(\mathbf{y}_c^{(k)}\right) > \sigma_{th}$ **then**
14 $\mathbf{u}_c^{(k)} = PIM(\tilde{M}_c{}^{(k)})$;
15 **else**
16 $\mathbf{u}_c^{(k)} = PIM(M_c^{(k)})$;
17 **end**
18 $\varepsilon_0 = 1 - \left|\mathbf{u}_c^{(k)} \mathbf{u}_{c_{old}}^{(k)\,T}\right|^2$;
19 $\mathbf{y}_c^{(k)} = \mathbf{u}_c^{(k)} Z^{(k)}$;
20 $\mathbf{u}_{c_{old}}^{(k)} = \mathbf{u}_c^{(k)}$;
21 **end**
22 $\mathbf{u}_c^{(k)*} = \mathbf{u}_c^{(k)}$;$\mathbf{y}_c^{(k)*} = \mathbf{y}_c^{(k)}$;
23 **Deflation step:** Remove the content of $\mathbf{y}_c^{(k)*}$ from each row of $Z^{(k)}$;
24 **end**
25 **end**
26 $i = i + 1$;
27 **end**

cross-cumulant of MSI random variables is zero, (3). Thus, the value of $\Upsilon_\Phi(\mathbf{y}_c^{(k_1)*})$ is expected to be high and low for joint and individual sources, respectively. Consequently, we can select a threshold (σ_{th}) to decide about the type of source (line 13 in Algorithm 1).

In each outer iteration, the type of extracted source might change, depending on the updated $\mathbf{y}_c^{(k_1)}$. This approach will enable us to resolve the permutation indeterminacy across datasets and to group the dependent sources automatically across subjects. Interestingly, this method does not require to know the number of joint and individual sources, because, it automatically determines their type.

It is worth noting that in deflation based methods, after extracting a desired independent component ($\mathbf{y}_c^{(k)*} \in \mathbb{R}^{1\times V}$), its contribution can be subtracted from the observation matrix ($Z^{(k)}$) by linear regression. By doing so, each row of the new observation matrix ($Z_{new}{}^{(k)}(n,:) = \mathbf{z}_{new,n}^{(k)} \in \mathbb{R}^{1\times V}\ \forall\, n = 1, ..., C$) is obtained as follows [2]:

$$\mathbf{z}_{new,n}^{(k)} = \mathbf{z}_n^{(k)} - \left(\mathbf{z}_n^{(k)}(\mathbf{y}_c^{(k)*})^T\right)\left(\mathbf{y}_c^{(k)*}(\mathbf{y}_c^{(k)*})^T\right)^{-1}\mathbf{y}_c^{(k)}, \tag{23}$$

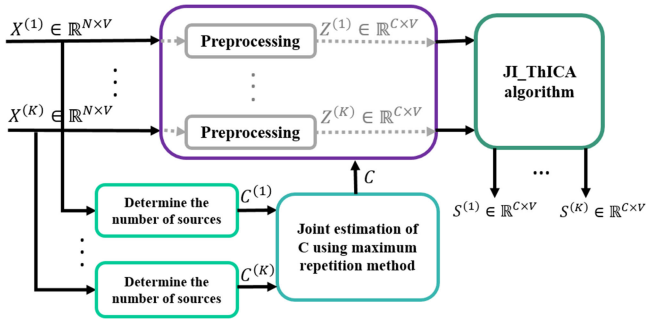

Fig. 2. Flowchart of the proposed algorithm to analyze multi-subject fMRI data, where the nth raw of observation matrix $X^{(k)}; k = 1, ..., K$ is the vectorized version of fMRI image in time n for kth subject.

where $\mathbf{y}_c^{(k)*}$ and $\mathbf{z}_{new,n}^{(k)}\ \forall\, n = 1, ..., C$ are uncorrelated, thus, by processing $Z_{new}{}^{(k)}$ in the next step of the deflation procedure, we extract another independent component.

Fig. 2 shows the flowchart of the proposed algorithm to analyze multi-subject fMRI data.

IV. Numerical Results

A. Performance Measures

We evaluate the performance of the proposed algorithm by computing the Signal-to-Interference Ratio (SIR) [18] as follows

$$SIR(\mathbf{s}, \mathbf{y}^*) = 10\log_{10}\left(\frac{\sum_{v=1}^{V} s^2(v)}{\sum_{v=1}^{V}(s(v) - y^*(v))^2}\right), \tag{24}$$

where $\mathbf{s}$ and $\mathbf{y}^*$ are normalized real and estimated sources with zero mean and unit variance, respectively. The joint SIR (jSIR) measure is obtained by computing the mean value of SIR measure of K datasets. Clearly, based on the above definition, in our performance evaluation higher jSIR is desired.

B. Results on Simulated fMRI Data

To evaluate our algorithm, we generate a synthesized fMRI multi-subject dataset with the SimTB toolbox [24] to determine the capabilities and limitations of JI-ThICA algorithm[1]. This toolbox can simulate 27 spatially independent brain source, as shown in Fig. 3. In our study, the simulated fMRI images with 64×64 pixels are generated. By using SimTb toolbox, we have assigned the same synthesized spatial map across all subjects for joint sources and different spatial maps for individual sources. Furthermore, we have used time courses generated with SimTB toolbox as mixing matrix. For more details about the synthesized fMRI signals interested readers are referred to see [24].

We apply spatial ICA to analyze fMRI data, which is based on the assumption that a weighted linear combination of C spatial independent source images compose each fMRI image over time. Here, we assume that C is equal to the number of principal components in PCA, which is unknown.

[1]This toolbox is available at http://mialab.mrn.org/software/simtb/ [Accessed: 2017-03-20]

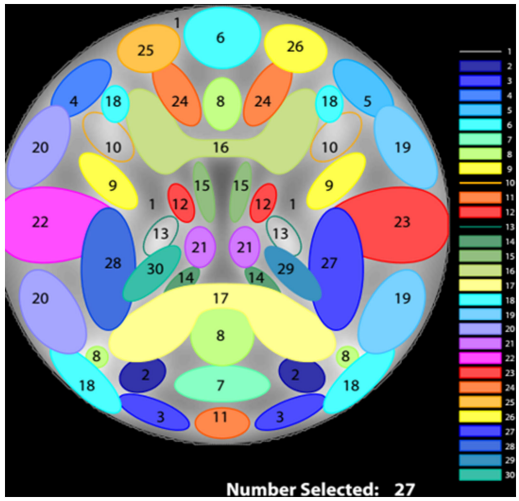

Fig. 3. Configuration of default sources in SimTb toolbox [24].

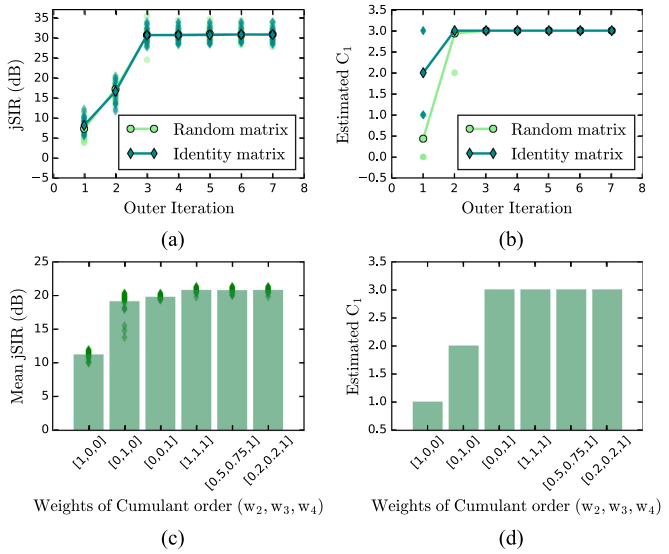

Fig. 4. The convergence rate for the JI-ThICA algorithm versus different outer iterations(a) Mean jSIR and (b) Mean estimated number for joint sources ($K = 8$, $N = 150$, $C_1 = 3$ and $C_2 = 3$). (c) Mean jSIR and (d) Mean estimated number of joint sources for different weights of cumulant orders.

In order to estimate the number of principal components, we use BIC method with the maximum likelihood ICA algorithm [36] for each subject and select the number with maximum repetition as the number of principle components (C).

In the following, K denotes the number of subjects, N is the number of observation time points, and C_1 and C_2 are the number of joint and individual sources, respectively. Note that C_1 and C_2 are unknown and the algorithm should find the correct C_1 and C_2. In the I-ThICA and J-ThICA cost functions, we consider $\Phi = \{2, 3, 4\}$, $w_2 = 0.5$, $w_3 = 0.75$, $w_4 = 1$ and $\upsilon_m = 0$ for all $m = 1, .., \eta$. In addition, for the cases of $K = 2$ and $K = 3$ we set ($k_3 = k_1$, $k_4 = k_2$), and ($k_4 = k_1$), respectively.

It should be noted that the number of simulated sources in SimTB toolbox is limited to 27, thus, K, C_1 and C_2 are limited as $C_1 + C_2 \times K = 27$. Therefore, there is a limitation to increase K, C_1 and C_2 in the JI-MDM source model, however, for the MDU case ($C_2 = 0$), we can increase C_1 up to 27 without any limitation on K.

In Fig. 4(a), the convergence rate of the JI-ThICA algorithm in terms of the mean jSIR is shown, where the results are obtained from 50 Monte Carlo runs under two different initializations (random and identity matrices), and $K = 8$, $N = 150$, $C_1 = 3$ and $C_2 = 3$. Note that in the case of random matrix initialization, a different random initial matrix is used for each Monte Carlo run. It can be observed that the algorithm converges after four iterations and initialization does not affect the convergence rate. Furthermore, Fig. 4(b) shows the estimated number of joint sources (C_1) in each outer iteration. Here, the same setting as Fig. 4(a) is considered, and similarly, after four iterations the mean number of joint sources converges to 3.

In Fig. 4(c) and (d) the performance of the algorithm is analyzed under different weights settings for order cumulants. From the figure, we observe that by setting $w_2 = 1$ and $w_3 = w_4 = 0$ the worst performance is obtained, while if we set $w_2 = w_3 = 0$ and $w_4 = 1$, the performance of the algorithm is significantly improved. This observation reveals that giving higher weight to the fourth order cumulant leads to a better performance. Furthermore, by setting $w_2 = w_3 = w_4 = 1$, the performance of the algorithm is further improved. By changing the values of w_2 and w_3 ($w_2 = 0.5$ or 0.2, $w_3 = 0.75$ or 0.2) for a fixed w_4 ($w_4 = 1$), we get approximately the same results as for $w_2 = w_3 = w_4 = 1$. Thus, it can be deduced that the fourth order cumulant has the most contribution on the performance improvement, and using the second and third order cumulants can slightly improve the performance of the algorithm.

We compare the results of the proposed JI-ThICA algorithm in extracting joint and individual sources with the algorithms proposed in [18][2]. Since the simulated fMRI sources and observations are assumed to be positive, here we can use the CNFE method [18]. Furthermore, as the JIVE algorithm has poor performance with respect to the CNFE alternative [18] and also has a high run time, we do not report the results of JIVE. In the COBEC algorithm [18], the number of joint sources are known, whereas in this section we compare the results of those methods estimating the number of joint sources, thus, in the following, only the results of the JI-ThICA and CNFE methods are compared.

Fig. 5(a) shows the estimated number of joint sources (C_1) versus σ_{th} for the experiment considered in Fig. 4 with $MaxIter = 4$. Results reveals that the proper interval of σ_{th} to estimate C_1 correctly is $[10^{-4}, 10^{-1}]$.

In [18], a parameter called ϵ has been introduced for the CNFE method, which controls the degree of similarity of the extracted sources. In order to find a proper value for ϵ, we have tested different values for ϵ in the same scenario as the experiment considered in Fig. 4 ($K = 8$, $N = 150$, $C_1 = 3$ and $C_2 = 3$). Fig. 5(b) shows the number of estimated joint sources (C_1) by CNFE versus ϵ. We observe that for $10^{-3} < \epsilon \leq 1$ CNFE estimates joint sources correctly.

[2]We have downloaded the MATLAB codes of COBE and CNFE algorithms from http://www.bsp.brain.riken.jp/ zhougx/cifa.html [Accessed: 2017-03-10].

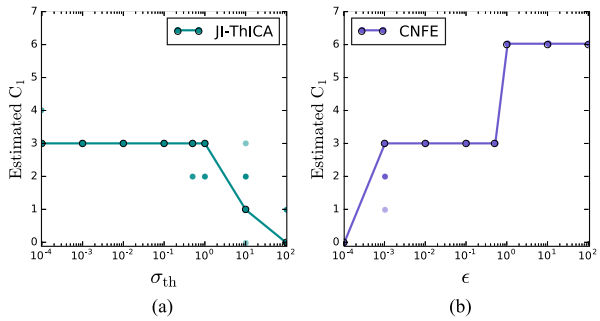

Fig. 5. Proper intervals of σ_{th} and ϵ to estimate the number of joint sources (C_1) correctly (a) in the JI-ThICA algorithm with $MaxIter = 4$, and (b) in the CNFE algorithm, where the correct C_1 is 3.

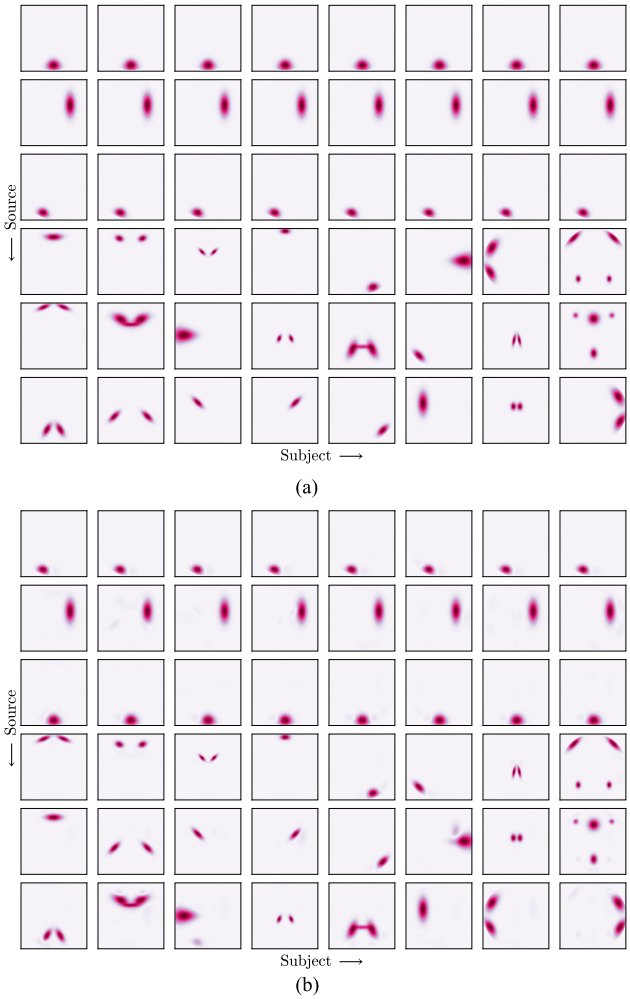

Fig. 6. (a) An example for original simulated fMRI sources ($K = 8$, $C_1 = 3$ and $C_2 = 3$). (b) Estimated sources with JI-ThICA algorithm.

From Fig. 4 we observed that the initialization condition does not affect the results, thus, in the following experiments identity initialization matrix is used. In addition, considering the results in Fig. 5 we set $\sigma_{th} = 10^{-3}$ and $\epsilon = 10^{-2}$ in the rest of this study. Furthermore, in the following the results are obtained by computing the average of 30 Monte Carlo runs, where each run is recorded after 4 outer iterations ($MaxIter = 4$).

In Fig. 6(a), an example of simulated fMRI sources is shown for $K = 8$, $C_1 = 3$ and $C_2 = 3$. The results of applying the JI-ThICA algorithm to extract joint and individual sources of this example are depicted in Fig. 6(b). We see that our algorithm aligned joint sources across subjects, and it extracted individual sources correctly.

In Fig. 7(a), the Mean jSIR of the JI-ThICA and CNFE algorithms are shown versus the number of datasets K ($N = 150$, $C_1 = 3$, $C_2 = 3$) for clear observation. The results of CNFE are depicted for both SOBI and thin ICA. We observe that by using thin ICA, the performance of CNFE is improved with respect to the scenarios employing SOBI. In addition, for $K > 3$, the JI-ThICA algorithm outperforms CNFE-SOBI, whereas it has approximately the same performance as CNFE-ThICA. It is worth mentioning that in the cost function of the JI-ThICA we use second, third and fourth order cumulants and cross-cumulants, and for the case of $K \leq 3$ fourth order cross-cumulants is computed by repeating sources (if $K = 2$ and $K = 3$, we compute fourth order cross-cumulants by setting $\{k_3 = k_1, k_4 = k_2\}$, and $\{k_4 = k_1\}$, respectively), which leads to performance degradation. However, for higher number of sources ($K > 3$) it uses different sources to compute fourth order cross-cumulants and its performance is improved.

The asterisk symbols in Fig. 7 (as well as Figs. 8 and 10) indicate the results of a *one-sided t-test* for the null hypothesis that the average of the jSIR measure of the JI-ThICA algorithm is less or equal to the measure of other algorithms. No asterisk symbol means that the p-value (P_{value}) of the test is higher than 0.01, one asterisk indicates that $10^{-4} < P_{value} \leq 10^{-2}$, two asterisk shows that $10^{-6} < P_{value} \leq 10^{-4}$, and three asterisk indicates that $P_{value} \leq 10^{-6}$. It should be noted that before using the t-test, the normality of the data is checked using Shapiro Wilk test [37][3].

It should be noted that the spatial shape of joint active sources in real fMRI dataset is not exactly the same, and there is a spatial variation. Furthermore, the spatial patterns of joint sources may be affected by aligning brain images of subject on a standardized space [38]. These modifications are likely to obscure the joint sources similarity, and degrade the algorithm performance in terms of joint source discrimination. In Fig. 7(b), the results of a noisy observation experiment are shown. In this experiment, the effect of spatial variability of joint sources between subjects is modeled by applying random small variations in the location, rotation, and spread of active regions of sources. To this aim, two joint sources of subjects are translated down by 10% of the image size, are rotated 10 degrees counter-clockwise and clockwise, and contracted (enlarged) with 0.8 (1.2) factor, to model the location, rotation and spread variations, respectively. In addition, a white Gaussian noise is added to the observations with signal to noise ratio (SNR) of 3 dB, where $SNR = 10(\log_{10} \frac{\text{Signal Power}}{\text{Noise Power}})$.

In Fig. 7(b) the same behavior as Fig. 7(a) is observed, however, for the case of noisy observation our algorithm outperforms both CNFE+SOBE and CNFE+ThICA for $K > 3$, which indicates that our algorithm is a better choice for noisy observations.

[3]We have checked the normality of the data in http://scistatcalc.blogspot.ca/2013/10/shapiro-wilk-test-calculator.html

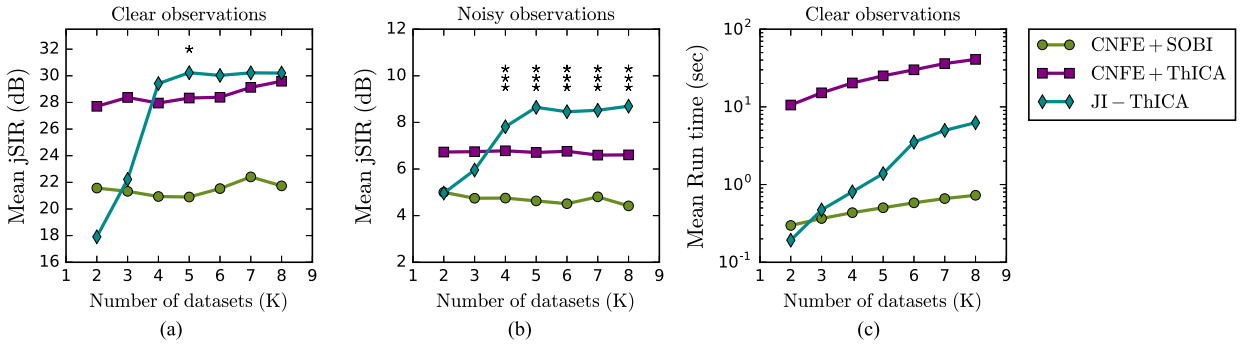

Fig. 7. Performance Comparison among CNFE+SOBE, CNFE+ThICA and JI-ThICA algorithms versus the number of datasets ($N = 150$, $C_1 = 3$, $C_2 = 3$), (a) Mean jSIR (dB) for clean observations, (b) Mean jSIR (dB) for noisy observations, (c) Mean run time for clear observations.

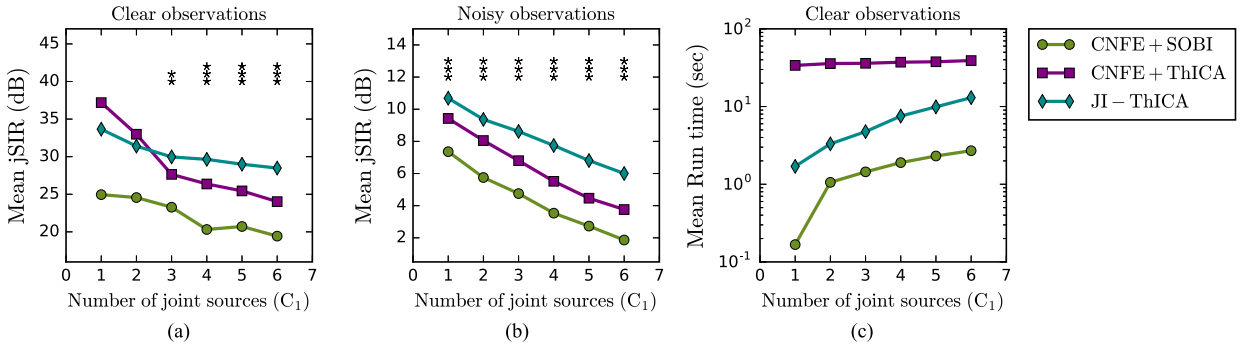

Fig. 8. Performance Comparison among CNFE+SOBE, CNFE+ThICA and JI-ThICA algorithms versus the number of joint sources ($N = 150$, $K = 7$, $C_2 = 3$), (a) Mean jSIR (dB) for clean observations, (b) Mean jSIR (dB) for noisy observations, (c) Mean run time for clear observations.

In Fig. 7(c), the mean run times of CNFE+SOBE, CNFE+ThICA and JI-ThICA are compared versus the number of datasets. Note that all experiments are conducted with an Intel(R) Core(TM) i7- 2.40 GHz computer with 8.0 GB of RAM. We see that CNFE with thin ICA and SOBI have the highest and lowest mean run times, respectively. Furthermore, the run times of algorithms for noisy and clear observations are approximately identical.

In Fig. 8, the mean jSIR of the JI-ThICA and CNFE algorithms are shown versus the number of joint sources C_1 ($N = 150, K = 7, C_2 = 3$) for both clear and noisy observations. Here, we see the same trend as Fig. 7, and the JI-ThICA method outperforms the CNFE+SOBE and CNFE+ThICA alternatives.

In order to evaluate the capability of algorithms in estimating the number of joint sources, in Fig. 9 the mean estimated C_1 versus the exact C_1 of the experiment considered in Fig. 8 is depicted. As it can be seen, for clear observations all methods estimate the joint sources correctly, however, in the case of noisy observations, the JI-ThICA on average estimates joint sources correctly, whereas the CNFE+SOBE and CNFE+ThICA methods can not extract joint sources. Note that in Fig. 9 the results of CNFE+SOBE and CNFE+ThICA are exactly the same, and as the curve of CNFE+SOBE is under the curve of CNFE+ThICA, it is not visible.

It is worth mentioning that in the above experiments changing either the number of fMRI image pixels or the number of observation time points does not affect the performance of the

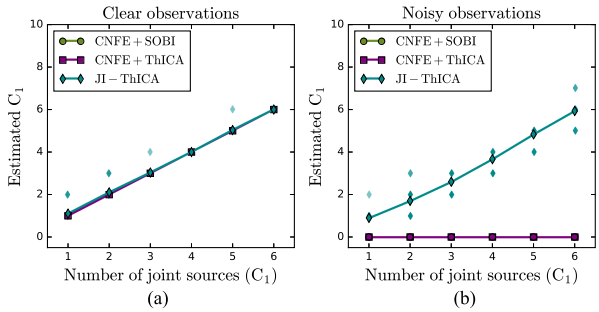

Fig. 9. Estimated number of joint sources (C_1) for CNFE+SOBE, CNFE+ThICA and JI-ThICA algorithms versus the number of joint sources ($N = 150$, $K = 7$, $C_2 = 3$), (a) clean observations, (b) noisy observations.

algorithms. Here, due to the space limitation, we do not report the results of changing these parameters.

As mentioned, if we set the number of individual sources C_2 to zero, the source model in the JI-MDM model will be the same as the MDU one. Thus, we can compare JI-ThICA with algorithms proposed for the MDU model.

In Fig. 10, we compare our algorithm with the IVA algorithm proposed in [26], referred to as IVA-Lap, Group ICA [19] and GIG-ICA [9] algorithms[4] (which are used in GIFT toolbox [39]).

[4]We have downloaded the MATLAB codes of group ICA and GIG-ICA algorithms from http://mialab.mrn.org/software/gift/ [Accessed: 2017-07-20]

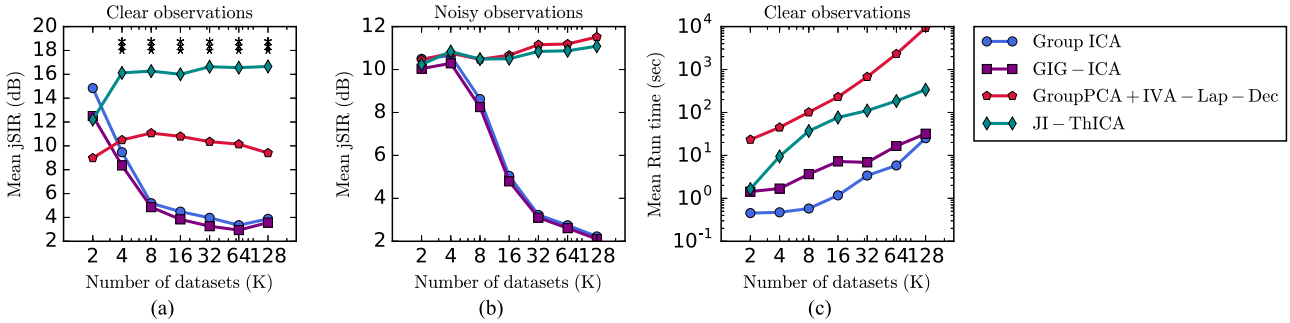

Fig. 10. Performance Comparison among Group ICA, GIG-ICA, GroupPCA+IVA-Lap-Dec and JI-ThICA algorithms versus the number of datasets for clear observations ($N = 150$, $C_1 = 27$, $C_2 = 0$), (a) Mean jSIR (dB), (b) Mean run time in second.

In IVA-Lap, a second-order uncorrelated multivariate Laplace distribution is assumed for the source vectors. We have examined other type of probability density functions for IVA algorithm (e.g., multivariate Gaussian distribution prior and multivariate power exponential distribution prior [40]), however, as the Laplace distribution had the best performance, here we only report the results of Laplace distribution. Furthermore, we use the decoupled version of IVA-Lap (IVA-Lap-Dec), which enables each row of the unmixing matrix to be individually optimized [41][5]. Note that here IVA is initialized with GroupPCA [26].

In Fig. 10, the mean jSIR and run time of four algorithms are depicted versus the number of datasets ($N = 150$, $C_1 = 27$, $C_2 = 0$) for clear and noisy observations. We observe that for clear observations our algorithm outperforms three benchmarking algorithms in terms of mean jSIR for $K \geq 4$. For noisy observations, our algorithm outperforms Group-ICA and GIG-ICA and approximately has the same performance as GroupPCA+IVA-Lap-Dec. In addition, our algorithm has a lower run time than GroupPCA+IVA-Lap-Dec for clear observations. Note that the run times of algorithms for clear and noisy observations are approximately the same. In Fig. 10, we observe that the GroupPCA+IVA-Lap-Dec has a higher mean jSIR for noisy observations with respect to clear observations. This behavior stems from the cost function of IVA which is given by [10]

$$\mathcal{I}_{IVA} = \sum_{c=1}^{C}\sum_{k=1}^{K}\mathcal{H}[y_c^{[k]}] - \sum_{c=1}^{C}\mathcal{MI}[y_c] - \sum_{k=1}^{K}\log|\det(W^{[k]})| \tag{25}$$

where $y_c = [y_c^{[1]}, y_c^{[2]}, y_c^{[K]}]^T$, $\mathcal{H}[y_c^{[k]}]$ and $\mathcal{MI}[y_c]$ are entropy and mutual information of $y_c^{[k]}$ and y_c, respectively. In the case of noisy data, the similarities of sources obscure and the term $\sum_{c=1}^{C}\mathcal{MI}[y_c]$ vanishes for joint sources. Hence, the IVA-Lap-Dec considers joint sources as individual sources. Thus, $\mathcal{I}_{IVA}$ is changed to K individual ICA cost functions, which can extract individual sources correctly. This is because it needs to estimate a multivariate probability density function (pdf) for joint sources with a pre-determined characteristics, but for individual sources it is sufficient to know the pdf of each individual source, which is more straightforward to estimate. In other words, for IVA cost function, the similarity measure of joint sources is modeled by a multivariate pdf and there is no separate measure for detecting the dependence of the joint sources across subjects [26]. Furthermore, we can conclude that the IVA algorithm is very sensitive to multivariate probability density function for joint sources, and choosing a correct and proper multivariate probability density function has a significant effect on the IVA performance to extract joint sources.

C. Results on Real fMRI Data

Dataset: In this subsection, we use a high-quality 7-Tesla fMRI dataset were obtained from the Study Forrest project[6] [27], henceforth refereed to as 7T-fMRI dataset. This dataset contains fMRI data of 20 right-handed participants (aged 2138 years, mean age 26.6 years, 8 females) listening to a two-hour audio-movie version of the film Forrest Gump. All participants were reported to have normal hearing without any history of neurological disorders. The audio-movie was split into 8 segments and the fMRI data was recorded in two separate sessions. Each session consisted 4 runs, and in each run one segment of the audio-movie was presented to the participants.

Data acquisition [27]: Participants were instructed to listen to the audio segments naturally. T2*-weighted echo-planar images (EPI) were obtained by employing a 7-Tesla Siemens MAGNETOM scanner (with scanning parameters of: repetition time = 2000 ms, echo time = 22 ms, 36 axial slices, slice thickness = 1.4 mm, in-plane resolution = 1.4 × 1.4 mm, field of view = 224 mm). The scanner was equipped with a 32-channel brain receive coil (Nova Medical, Inc., Wilmington, MA, USA) and a local circularly polarized head transmit. In order to include the ventral portions of frontal and occipital cortex, slices were oriented. For each participant 3599 fMRI volumes were recorded (the number of voxels in each fMRI volume = 145 × 180 × 71). The number of volumes was different for each run (451, 441, 438, 488, 462, 439, 542, and 338 volumes for runs 18, respectively).

Data processing: In this dataset, fMRI images were skull removed, motion and distortion corrected. A group-specific EPI

[5]We have downloaded the MATLAB codes of IVA-Lap-Dec from http://mlsp.umbc.edu/resources.html [Accessed: 2017-03-20].

[6]Public available at: http://studyforrest.org [Accessed: 2017-03-20]

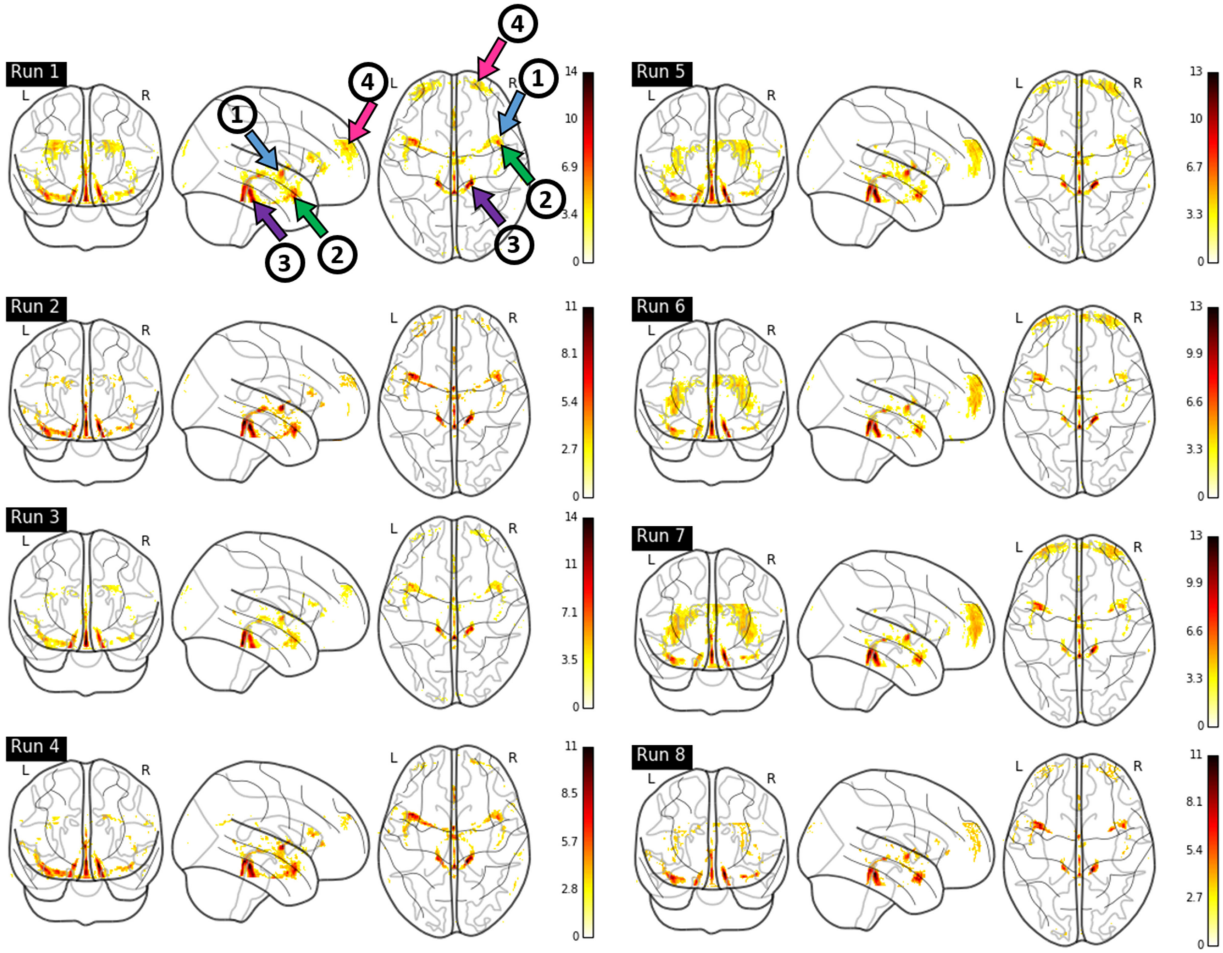

Fig. 11. Spatial distribution of significant joint source ($p < 0.05$ FDR corrected).

template was derived and all fMRI data were aligned to the EPI template [27]. In our analysis, we use data of 18 subjects, because of problems with images of two subjects (4th and 10th subjects). We exclude data of 4th subject due to the motions of participant during the experiment (coughing occasionally), and the data of 10th subject because its distortion corrected fMRI data is not available.

We perform additional fMRI preprocessing steps by using the AFNI (Analysis of Functional Neuro-Images) toolbox [48] for 18 subjects and 8 runs, separately. The following fMRI preprocessing steps are applied:

- Slice-timing correction by using heptic interpolation (using 3dTshift function),
- Band-pass filtering the BOLD signals (using 3dBandpass function with fbot $= 0.001$ and ftop $= 0.2$).

Note that here no spatial smoothing is performed, and the order model (number of principal and independent components) is set to 4 based on the investigated scenarios.

We apply the JI-ThICA method on the preprocessed fMRI data of 18 subjects to extract 4 independent components. As mentioned, one of the main advantages of JI-ThICA is automatic grouping of the sources across subjects by utilizing a measure during optimization process. Thus, the first extracted sources of all subjects can be considered as the first joint source, which have high mutual cumulants across subjects. We choose $\sigma_{th} = 0.1$ and $MaxIter = 5$ according to the investigated scenarios.

Results: By applying the JI-ThICA algorithm on the 7T-fMRI dataset, we observe that it can extract one joint and three individual sources. In Fig. 11, the mean of the first extracted sources across subjects ($p < 0.05$ False Discovery Rate (FDR) corrected [49]) is shown for each run using Nilearn toolbox [50]. We observe that the active regions are approximately the same in all 8 runs. It is expect that the activated regions are related to the experiment task, which will be validated in the following.

We refer to the voxels with maximum activity at each active regions as focus point. Table I (the second column) shows the output of whereami function in AFNI for the focus points of active regions in the extracted joint source. The approximate region of these focus points are shown in the first run of Fig. 11 with numbered and colorful arrows. Note that in Table I only the active regions in right hemisphere are reported.

As summarized in Table I, the extracted brain regions in our analysis is compatible with inter-subject correlation maps during naturalistic audio listening in the other studies such as [28] and [29].

It is worth mentioning that in [30] the 7T-fMRI dataset has been analyzed by applying decoding power-spectral profiles

TABLE I
THE FUNCTION OF SOME OF ACTIVE REGIONS IN EXTRACTED JOINT SOURCE IN THE RIGHT HEMISPHERE

Arrow	Talairach-Tournoux Atlas	Function
1: Blue	Focus point: Right Superior Temporal Gyrus, Within 4 mm: Right Middle Temporal Gyrus -AND- Right Brodmann area 22 -AND- Right Brodmann area 21	Auditory processing, including language and social cognition [42]
2: Green	Focus point: Right Middle Temporal Gyrus, Within 3 mm: Right Brodmann area 21	Auditory processing and language [43] and accessing word meaning while reading [44]
3: Purple	Focus point: Right Brodmann area 47, Within 2 mm: Right Brodmann area 11	Processing of syntax in oral and sign languages, musical syntax, and semantic aspects of language [45]
4: Pink	Focus point: Right Precuneus, Within 5 mm: Right Brodmann area 7	High-level processing tasks, including activation in association with language use [46] and mental imagery [47]

and support vector machine (SVM) classifier. In [30], the Superior Temporal Gyrus (STG) which contains the primary auditory cortex and the Middle Temporal Gyrus (MTG) have been extracted from fMRI brain activities of this database.

It is worth mentioning that the extracted individual sources are uncorrelated and independent across subjects. The corresponding brain voxels of individual sources are not related to the auditory activity. Here, as individual sources are not related to the auditory experiment, we do not further investigate them. Furthermore, as there are no Gold signals for this real fMRI dataset, here we can not present a quantitative analysis between our algorithm and its alternatives.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed a new algorithm to extract both joint and individual sources of multi-subject fMRI datasets. We compared the proposed algorithm, referred to as JI-ThICA, with existing alternatives, namely CNFE, which has been proposed for JI-MDM dataset model. Furthermore, we compared our algorithm with Group ICA, GIG-ICA, and GroupPCA+IVA-Lap-Dec for an MDU source model, as our algorithm is also applicable for this type of dataset model. We have evaluated the introduced algorithms by analyzing simulated fMRI signals. We applied the algorithms on the clean and noisy fMRI signals, where in the noisy signals we considered small variations in the location, rotation, and spread of joint sources across subjects, and added a white Gaussian noise to the observations. Our results reveal that the proposed algorithm outperforms CNFE in terms of mean joint SIR. In the case of MDU source model, the JI-ThICA algorithm outperforms the three benchmarking algorithms for clear observations, and in noisy observations it has a better performance than GroupICA and GIG-ICA, whereas it has approximately the same performance as GroupPCA+IVA-Lap-Dec. Moreover, we have applied our algorithm on a real fMRI dataset with 18 subjects, which was acquired during naturalistic auditory experience. Our algorithm extract one joint sources across subject in the auditory cortex of brain, which is in accordance with the results of previous studies investigated this dataset.

Our experimental results on the synthesized and real fMRI signals reveal that employing higher order cumulants (second, third and especially forth order cumulants) in the cost function of our algorithm improves its capability to extract common information of multi-subject datasets. In addition, as higher order cumulants suppress Gaussian noise, the performance of our algorithm for noisy fMRI observations is significantly better than CNFE. Therefore, we can deduce that higher order cumulants are appropriate metrics to extract interdependency and common information of multi-subjects and multi-modal datasets, which deserve further investigation in future studies to be applied in other source models.

The future works of this paper are summarized as follows:

- The proposed method can be extended to extract joint and individual sources in a general MDM source model. In the current algorithm, we considered a special case of MDM (JI-MDM) in which the joint sources are the same in all subjects. To be more compatible with real conditions, it is worthy to assume that subjects of dataset are categorized in a number of subsets, and each subset has a number of joint sources, which are not necessarily common in other subsets. Furthermore, it is worthy to assume that the numbers of individual sources of subjects are not identical.
- In the proposed method, the number of joint sources of each subject is extracted separately and a voting method is applied over all subjects. In future, the algorithm can be extended to jointly estimate the number of joint sources across subjects in an optimization problem instead of the voting method.
- In the proposed method, the required threshold (σ_{th}) is set manually. It is beneficial to extend the algorithm to set σ_{th} automatically.
- A detailed evaluation of the proposed algorithm to analyze more real datasets and other types of real biomedical signals such as EEG and ECG is left for a future work.

REFERENCES

[1] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Sep. 2015.
[2] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. San Francisco, CA, USA: Academic, 2010.
[3] R. F. Silva *et al.*, "Blind source separation for unimodal and multimodal brain networks: A unifying framework for subspace modeling," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1134–1149, Oct. 2016.
[4] J. Sui *et al.*, "A CCA+ ICA based model for multi-task brain imaging data fusion and its application to schizophrenia," *Neuroimage*, vol. 51, no. 1, pp. 123–134, 2010.
[5] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, 2014, Art. no. 1450013.
[6] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "SSVEP recognition using common feature analysis in brain–computer interface," *J. Neurosci. Methods*, vol. 244, pp. 8–15, 2015.

[7] X.-L. Li, T. Adalı, and M. Anderson, "Joint blind source separation by generalized joint diagonalization of cumulant matrices," *Signal Process.*, vol. 91, no. 10, pp. 2314–2322, 2011.

[8] D. Lahat and C. Jutten, "Joint analysis of multiple datasets by cross-cumulant tensor (block) diagonalization," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, 2016, pp. 1–5.

[9] Y. Du and Y. Fan, "Group information guided ICA for fMRI data analysis," *Neuroimage*, vol. 69, pp. 157–197, 2013.

[10] J.-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, 2008.

[11] T. Adali, M. Anderson, and G.-S. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 18–33, May 2014.

[12] N. P. Wojtalewicz, R. F. Silva, V. D. Calhoun, A. D. Sarwate, and S. M. Plis, "Decentralized independent vector analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2017, pp. 826–830.

[13] D. Lahat, J.-F. Cardoso, and H. Messer, "Second-order multidimensional ICA: Performance analysis," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4598–4610, Sep. 2012.

[14] Z. Szabó, B. Póczos, and A. Lőrincz, "Separation theorem for independent subspace analysis and its consequences," *Pattern Recog.*, vol. 45, no. 4, pp. 1782–1791, 2012.

[15] R. F. Silva, S. M. Plis, T. Adali, and V. D. Calhoun, "Multidataset independent subspace analysis extends independent vector analysis," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 2864–2868.

[16] D. Lahat and C. Jutten, "Joint independent subspace analysis: A quasi-newton algorithm," in *Proc. Int. Conf. Latent Variable Anal. Signal Separation*, 2015, pp. 111–118.

[17] E. F. Lock, K. A. Hoadley, J. S. Marron, and A. B. Nobel, "Joint and individual variation explained (JIVE) for integrated analysis of multiple data types," *Ann. Appl. Statist.*, vol. 7, no. 1, pp. 523–542, 2013.

[18] G. Zhou, A. Cichocki, Y. Zhang, and D. P. Mandic, "Group component analysis for multiblock data: Common and individual feature extraction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 11, pp. 2426–2439, Nov. 2016.

[19] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis," *Human Brain Mapping*, vol. 14, no. 3, pp. 140–151, 2001.

[20] X. Chen, Z. J. Wang, and M. McKeown, "Joint blind source separation for neurophysiological data analysis: Multiset and multimodal methods," *IEEE Signal Process. Mag.*, vol. 33, no. 3, pp. 86–107, Nov. 2016.

[21] G. Zhou, Q. Zhao, Y. Zhang, T. Adalı, S. Xie, and A. Cichocki, "Linked component analysis from matrices to high-order tensors: Applications to biomedical data," *Proc. IEEE*, vol. 104, no. 2, pp. 310–331, Feb. 2016.

[22] E. A. Allen, E. B. Erhardt, Y. Wei, T. Eichele, and V. D. Calhoun, "Capturing inter-subject variability with group independent component analysis of fMRI data: A simulation study," *Neuroimage*, vol. 59, no. 4, pp. 4141–4159, 2012.

[23] S. Ma, R. Phlypo, V. D. Calhoun, and T. Adalı, "Capturing group variability using IVA: A simulation study and graph-theoretical analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2013, pp. 3128–3132.

[24] E. B. Erhardt, E. A. Allen, Y. Wei, T. Eichele, and V. D. Calhoun, "SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability," *Neuroimage*, vol. 59, no. 4, pp. 4160–4167, 2012.

[25] S. Cruces and A. Cichocki, "Combining blind source extraction with joint approximate diagonalization: Thin algorithms for ICA," in *Proc. 4th Int. Symp. Independent Component Anal. Blind Signal Separation*, 2003, pp. 463–468.

[26] A. M. Michael, M. Anderson, R. L. Miller, T. Adalı, and V. D. Calhoun, "Preserving subject variability in group fMRI analysis: Performance evaluation of gica vs. iva," *Frontiers Syst. Neurosci.*, vol. 8, 2014, Art. no. 106.

[27] M. Hanke *et al.*, "A high-resolution 7-tesla fMRI dataset from complex natural stimulation with an audio movie," *Sci. Data*, vol. 1, 2014, Art. no. 140003.

[28] W. Trost, S. Frühholz, T. Cochrane, Y. Cojan, and P. Vuilleumier, "Temporal dynamics of musical emotions examined through intersubject synchrony of brain activity," *Social Cogn. Affective Neurosci.*, vol. 10, no. 12, pp. 1705–1721, 2015.

[29] M. M. Farbood, D. J. Heeger, G. Marcus, U. Hasson, and Y. Lerner, "The neural processing of hierarchical structure in music and speech at different timescales," *Frontiers Neurosci.*, vol. 9, 2015, Art. no. 157.

[30] X. Hu, L. Guo, J. Han, and T. Liu, "Decoding power-spectral profiles from fMRI brain activities during naturalistic auditory experience," *Brain Imag. Behavior*, vol. 11, no. 1, pp. 253–263, 2017.

[31] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, Dec. 1974.

[32] G. Schwarz *et al.*, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, 1978.

[33] T. Hasija, Y. Song, P. J. Schreier, and D. Ramírez, "Detecting the dimension of the subspace correlated across multiple data sets in the sample poor regime," in *Proc. IEEE Statist. Signal Process. Workshop*, 2016, pp. 1–5.

[34] S. Cruces, L. Castedo, and A. Cichocki, "Novel blind source separation algorithms using cumulants," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2000, vol. 5, pp. 3152–3155.

[35] G. H. Golub and H. A. Van der Vorst, "Eigenvalue computation in the 20th century," *J. Comput. Appl. Math.*, vol. 123, no. 1, pp. 35–65, 2000.

[36] L. K. Hansen, J. Larsen, and T. Kolenda, "Blind detection of independent dynamic components," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2001, vol. 5, pp. 3197–3200.

[37] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, nos. 3/4, pp. 591–611, 1965.

[38] J.-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, 2008.

[39] S. Rachakonda, E. Egolf, N. Correa, and V. Calhoun, "Group ICA of fMRI toolbox (gift) manual," *Dostupné z. [Online]. Available: http://www.nitrc. org/docman/view.php/55/295/v1.3d_GIFTManual pdf [cit. 2011-11-5]*, 2007.

[40] M. Anderson, T. Adali, and X.-L. Li, "Joint blind source separation with multivariate gaussian model: Algorithms and performance analysis," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1672–1683, Apr. 2012.

[41] M. Anderson, X.-L. Li, P. Rodriguez, and T. Adali, "An effective decoupling method for matrix optimization and its application to the ICA problem," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2012, pp. 1885–1888.

[42] E. D. Bigler *et al.*, "Superior temporal gyrus, language function, and autism," *Develop. Neuropsychol.*, vol. 31, no. 2, pp. 217–238, 2007.

[43] F. Mirz *et al.*, "Stimulus-dependent central processing of auditory stimuli: A pet study," *Scandinavian Audiology*, vol. 28, no. 3, pp. 161–169, 1999.

[44] D. J. Acheson and P. Hagoort, "Stimulating the brain's language network: Syntactic ambiguity resolution after tms to the inferior frontal gyrus and middle temporal gyrus," *J. Cogn. Neurosci.*, vol. 25, no. 10, pp. 1664–1677, 2013.

[45] D. J. Levitin and V. Menon, "Musical structure is processed in language areas of the brain: A possible role for brodmann area 47 in temporal coherence," *Neuroimage*, vol. 20, no. 4, pp. 2142–2152, 2003.

[46] A. E. Cavanna and M. R. Trimble, "The precuneus: A review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. 3, pp. 564–583, 2006.

[47] P. Fletcher, C. Frith, S. Baker, T. Shallice, R. Frackowiak, and R. Dolan, "The mind's eye precuneus activation in memory-related imagery," *Neuroimage*, vol. 2, no. 3, pp. 195–200, 1995.

[48] R. W. Cox, "Afni: Software for analysis and visualization of functional magnetic resonance neuroimages," *Comput. Biomed. Res.*, vol. 29, no. 3, pp. 162–173, 1996.

[49] T. Nichols and S. Hayasaka, "Controlling the familywise error rate in functional neuroimaging: A comparative review," *Stat. Methods Med. Res.*, vol. 12, no. 5, pp. 419–446, 2003.

[50] A. Abraham *et al.*, "Machine learning for neuroimaging with scikit-learn," *Frontiers Neuroinformatics*, vol. 8, 2014, Art. no. 14.