Experimental Comparisons of Sparse Dictionary Learning and Independent Component Analysis for Brain Network Inference From fMRI Data

Wei Zhang, Jinglei Lv, Xiang Li [ID], Dajiang Zhu, Xi Jiang, Shu Zhang, Yu Zhao, Lei Guo, Jieping Ye, Dewen Hu [ID], and Tianming Liu [ID], *Senior Member, IEEE*

***Abstract*—In this work, we conduct comprehensive comparisons between four variants of independent component analysis (ICA) methods and three variants of sparse dictionary learning (SDL) methods, both at the subject-level, by using synthesized fMRI data with ground-truth. Our results showed that ICA methods perform very well and slightly better than SDL methods when functional networks' spatial overlaps are minor, but ICA methods have difficulty in differentiating functional networks with moderate or significant spatial overlaps. In contrast, the SDL algorithms perform consistently well no matter how functional networks spatially overlap, and importantly, SDL methods are significantly better than ICA methods when spatial overlaps between networks are moderate or severe. This work offers empirical better understanding of ICA and SDL algorithms in inferring functional networks from fMRI data and provides new guidelines and caveats when constructing and interpreting functional networks in the era of fMRI-based connectomics.**

***Index Terms*—Resting state fMRI, functional network, independent component analysis, sparse dictionary learning.**

W. Zhang, X. Li, S. Zhang, and Y. Zhao are with the Cortical Architecture Imaging and Discovery Lab, Department of Computer Science and Bioimaging Research Center, The University of Georgia.

J. Lv is with the Translational Neuroscience, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia.

D. Zhu is with the Department of Computer Science and Engineering, University of Texas at Arlington.

X. Jiang is with the Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, School of Life Science and Technology, University of Electronic Science and Technology of China.

L. Guo is with the School of Automation, Northwestern Polytechnical University.

J. Ye is with the Department of Electrical Engineering and Computer Science, University of Michigan.

D. Hu is with the College of Mechatronics and Automation, National University of Defense Technology.

T. Liu is with the Cortical Architecture Imaging and Discovery Lab, Department of Computer Science and Bioimaging Research Center, The University of Georgia, Athens, GA 30602 USA (e-mail: tianming.liu@gmail.com).

I. INTRODUCTION

SINCE the introduction of Independent Component Analysis [1]–[6]. During the past decade, ICA has become one of the (ICA) a few decades ago, ICA has achieved great successes in separating independent blind sources from observed mixtures dominant methods in functional magnetic resonance imaging (fMRI) data modeling and analysis, particularly for inferring brain networks. Though comprehensive reviews of ICA methods for fMRI data modeling are beyond the scope of this paper, here we briefly discuss a few exemplar ICA works that are most relevant to this paper. McKeown, in 2000, provided a hybrid ICA (HYBICA) that combined the advantage of powerful data-driven techniques and a priori hypotheses to guide the fMRI data analysis [7]. In 2001, Calhoun *et al.* introduced an effective group ICA (GICA) approach for fMRI data modeling [8]–[10]. An interesting observation of these early ICA methods is that they attempted to explore the efficiency of ICA on fMRI data analysis based on simulated experiments. These works used two different simulated experiments that employed independent sources to generate simulated data and then examined the reconstruction performance of ICA [8]–[10]. A few years later, a probabilistic ICA (PICA) approach and other variant ICAs were introduced and optimized for fMRI data analysis [11], [76]–[79]. The PICA was applied on fMRI data acquired at resting state, and it was demonstrated that PICA is an effective and robust tool for the identification of low-frequency resting-state oscillation patterns from fMRI data acquired at various spatial and temporal resolutions [11], [12]. Additional mathematic introductions of these ICA method are referred to Supplemental materials.

However, there is growing concern that ICA methods might have potential limitations in inferring brain networks when functional networks have significant spatial overlaps. From a neuroscience perspective, it has been demonstrated that cortical microcircuits are not independent and segregated in space, but they rather overlap and interdigitate with each other, given the convergent and divergent axonal projections in the brain and heterogeneous activities of intermixed neurons in the same brain region [13]–[15]. Multiple investigators have discussed and explicitly demonstrated extensive overlaps of large-scale functional networks (FNs) in the brain [14], [16]–[18]. Furthermore, many studies have already described and modeled spa-

tial overlaps of two or more FNs, indicating that fMRI signal mixes individual voxels into two or more FNs in the overlapping regions [19]–[29]. Several other research groups reported that task-evoked brain networks, such as in emotion, gambling, language and motor tasks, have large overlaps with each other [14], [16]–[18], [30]. More recently, several research groups systematically described and examined the important issue of FN overlaps [31]–[38]. These FN overlap studies raise an important question to ICA methods: can the ICA and its variant methods effectively and sufficiently separate blind source components when/if these components have strong temporal correlations or large spatial overlaps, given the basic assumption in ICA that those decomposed components should be as independent as possible [4], [5]?

Alternatively, sparse dictionary learning (SDL), such as the online dictionary learning (ODL) algorithm [39], stochastic coordinate coding (SCC) algorithm [40], [41] and KSVD [57], has been successfully applied to reconstruct concurrent brain networks from both task-based and resting-state fMRI datasets [41]–[44]. In this category of SDL-based methods, briefly, fMRI signals from all voxels within the whole brain are extracted and are then reorganized as a big 2-dimensional matrix, where the number of columns represents the total brain voxels and the number of rows stands for the time points. The 2D data matrix is then decomposed into the production of an over-complete dictionary basis matrix (each atom representing a functional network) and a reference weight matrix (representing this network's spatial volumetric distribution). A particularly important characteristic of this SDL framework for fMRI data modeling is that the reference weight matrix naturally reveals the spatial overlap and interaction patterns among those reconstructed brain networks [42]–[44]. In general, the SDL framework effectively provides compact high-fidelity representation of the whole brain fMRI signals and reveals meaningful spatial patterns such as brain FNs [41]–[50], [80]. For example, extensive experiments on the Human Connectome Project (HCP) Q1 fMRI datasets have demonstrated that the SDL methodology can effectively and robustly uncover multiple functional networks, including both task-evoked networks (TENs) and resting-state networks (RSNs) from task fMRI data that can be well-characterized and interpreted in spatial and temporal domains. In particular, these well-characterized TENs and RSNs networks are quite reproducible across different tasks and individuals (results publicly available on our website [42], [43] and exhibit substantial spatial overlap with each other, thus forming the Holistic Atlases of Functional Networks and Interactions [42], [43].

From a mathematics perspective, SDL is different from ICA, despite that both algorithms use linear matrix decomposition. The objective function for SDL, generally, is considered as a convex optimization problem based on ℓ_1 or ℓ_2 norm regression by using alternatively optimization. However, the objective for ICA is usually to maximize a negentropy [4], [5], which indicates that ICA needs to maintain the independence of components from each other. From a human neuroscience perspective, a variety of cortical regions and networks exhibit strong functional diversity and heterogeneity; that is, a cortical region could participate in multiple functional domains/processes and a functional network might recruit various heterogeneous neuroanatomical areas [13]–[17], [30]. Meanwhile, the number of functional networks that a cortical region is involved in at a specific moment is sparse, typically from several to one or two dozen [13], [30], [51]. Therefore, a computational modeling framework that can simultaneously investigate both functional network heterogeneity and sparsity, instead of independence, seems more desirable, such as the SDL methods [41]–[44]. Therefore, the next natural question is: is SDL methodology truly better than ICA methodology? Or, do SDL and ICA have their own superiority in different contexts?

Unfortunately, there have been very few experimental comparisons between ICA and SDL in the literature so far. Recently, a series of our research works have focused on the application of SDL on resting-state fMRI and/or task-based fMRI [41]–[44]. In this paper, we design a series of simulated experiments to qualitatively and quantitatively compare three variants of SDL (ODL, SCC and KSVD algorithms) with four variants of ICAs (basic ICA, fast ICA, fast radical ICA and efficient ICA), by examining the spatial overlap level, cross-validation and statistics analysis. Specifically, five sets of simulated experiments with ground-truth are designed to compare the network reconstruction performances by three SDL and four ICA methods. Our results showed that ICA methods perform very well and slightly better than SDL methods when functional networks' spatial overlaps are minor, but ICA methods have substantial difficulty in separating functional networks with moderate or significant spatial overlaps. In comparison, the SDL algorithms perform consistently well no matter how functional networks spatially overlap, and importantly, SDL methods are significantly better than ICA methods when spatial overlaps between networks are moderate or severe. In general, our works offer better understanding of ICA and SDL algorithms in inferring functional networks from fMRI data, and provides new guidelines and caveats when constructing and interpreting functional networks in the era of fMRI-based connectomics.

II. MATERIALS AND METHODS

In this section, we will firstly introduce the algorithms of two variants of Sparse Dictionary Learning including Online Dictionary Learning and Stochastic Coordinate Coding, as well as the four variants of Independent Component Analysis methods (see Section II-A). Secondly, we will introduce our methods for fMRI data synthesis (see Section II-B), based on which the quantitative comparisons among different algorithms will be performed. In general, 4 ICAs and 3 SDLs employed in our following simulated experiments are all at subject-level. Since the vital idea in this paper is to examine the spatial reconstruction performance of ICAs and SDLs based on human brain networks at different overlap levels, we do not concentrate on the comparison of time course reconstruction. We should emphasize that, if more factors (e.g., overlap level or random noise) are introduced simultaneously in all simulated experiments, it is difficult to clarify which factor is the dominant one that influences the reconstruction performance in the simulated experiments. This is why only overlap level is included in the series of simulated ex-
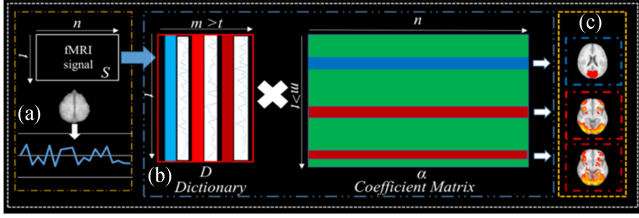
Fig. 1. The computational framework of SDL at subject-level applying on whole-brain fMRI signals for identification of FNs. (a) The aggregated whole-brain fMRI signals matrix S of a single subject. (b) The result dictionary matrix D. For the dictionary matrix, each column is a dictionary atom representing the temporal pattern identified by a single subject's fMRI data. (c) The result sparse coefficient matrix α. Each row in α represents the involvement of the corresponding dictionary atom in the whole brain functional activity, which can be further mapped back to 3-D volume space for representing the spatial distribution pattern of the FN. The whole coefficient matrix is decomposed after obtained dictionary, based on a single subject's fMRI signal.
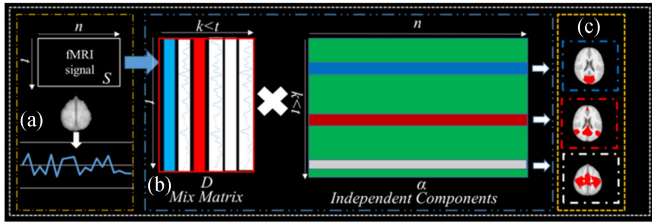

Fig. 2. The computational framework of basic spatial ICA for inferring FNs from whole-brain resting-state fMRI signals. (a) The aggregated whole-brain fMRI signals matrix S from an individual. (b) ICA decomposes the signal matrix S into mix matrix D and spatially independent components. (c) The obtained matrix α of spatially independent components. Furthermore, each row of α can be mapped back to brain volume to represent the spatial pattern of an FN.

periments. To examine ICAs and SDLs more comprehensively, additional experiments including random noise and comparison of time course reconstruction are provided in Supplementary materials.

A. ODL and SCC Algorithms for Brain Network Inference

The computational framework of identifying functional networks (FNs) from whole-brain fMRI signal using SDL is illustrated in Fig. 1. At the first step, the whole-brain fMRI signal is aggregated into a 2D data matrix $S \in R^{t \times n}$ containing t time points on n voxels [see Fig. 1(a)]. Then, **S** is decomposed into an over-complete dictionary matrix $D \in R^{t \times m}$ and a sparse coefficient matrix $\alpha \in R^{m \times n}$ [see Fig. 1(c)], i.e., $S = D \times \alpha$ by applying the ODL or SCC methods [39]–[44]. Specifically, each column of $\boldsymbol{D}$ is a dictionary atom representing the temporal pattern of a FN and the corresponding row of $\boldsymbol{\alpha}$ can be mapped back to brain volume in order to identify the spatial distribution pattern of the FNs [see Fig. 2(c)].

Mathematically, the learning process is performed as follows: defining the empirical cost function of **S** considering the averaged loss of regression. Indeed, if we assume that the parameter s_i represents the simple vector/column of a single matrix, the following (1) is considered as a traditional sparsity model; meanwhile, if we assume that the parameter s_i represents a combination of multiple fMRI signal matrices, the following (1) is

considered as a group-wise sparsity model:

$$f_n(D) \triangleq \frac{1}{n} \sum_{i=1}^{n} \ell(s_i, D) \tag{1}$$

where is i-th simulation data. Defining the loss function $\ell(s_i, D)$ or the sparse representation:

$$\ell(s_i, D) = \frac{1}{2} \min \|s_i - D\alpha\|_2^2 + \lambda \|\alpha_i\|_1 \tag{2}$$

where α_i is i-th column of α. The ℓ_1-regularization will shrinkage the solution which leads to a sparse solution of α. The parameter λ is selected to achieve tradeoff between regression fidelity versus sparsity. The computational framework of ODL [39] targets to solve (2) to learn D by alternatively optimizing D and α while fixing the other., It should be noted that while (2) is non-convex (i.e., difficult to find global optimum) [39]. Once D is successfully learned, α can be obtained by solving the ℓ_1-regularized linear least-squares problem based on D [39]. The ODL applies least angle regression (LARS) to train the dictionary. However, when the dimensionality of the data is high, the LARS can become very time-consuming. Therefore, ODL adopts the parallel computation to solve the problem. Based on our extensive experiments [42]–[44], [49], ODL has been shown as a robust and effective method to identify the FNs from both resting-state and task-based fMRI data.

While the underlying mathematical model of SCC is similar to ODL [as in (1) and (2)], SCC utilizes a Fast-Iterative Shrinkage-Thresholding Algorithm (FISTA) [52] for reducing the computational cost and improving its speed. It uses stochastic gradient descend (rather than computing the full gradient) to update the dictionary and coefficient matrix. More importantly, SCC algorithm utilizes a screening operator to skip atom with 0 coefficient during the optimization, which greatly accelerate the training process [40], [41]. Moreover, SCC solves the $l - 1$ norm problem in (2) by using shrinkage-thresholding function to simplify the training process. Our experiments have shown that SCC can achieve reasonably good performance for FN identification [41].

B. Four Variants of ICA Methods for Brain Network Inference

ICA has been widely used for fMRI data modeling and analysis [8]–[10], [19], [23], [24], [30], [53]–[55]. In general, ICA attempts to decompose the presumably independent signal sources that are mixed together. The primary assumption in ICA is the decomposed components' independence in space (non-systematically overlapping networks) or time (non-systematically varying networks) [8]. Mathematically, the separation of blind sources from the mixed observation is done by performing a matrix decomposition (5). The concept for applying ICA for the FN identification is shown in Fig. 2.

Comparing with the computational framework for SDL, the matrix D in (5) is fundamentally different. Specifically, for the application on the fMRI field, the dictionary learned in SDL is over-complete, which indicates that the dictionary size should be larger than the temporal length (i.e., total number of 3-D im-
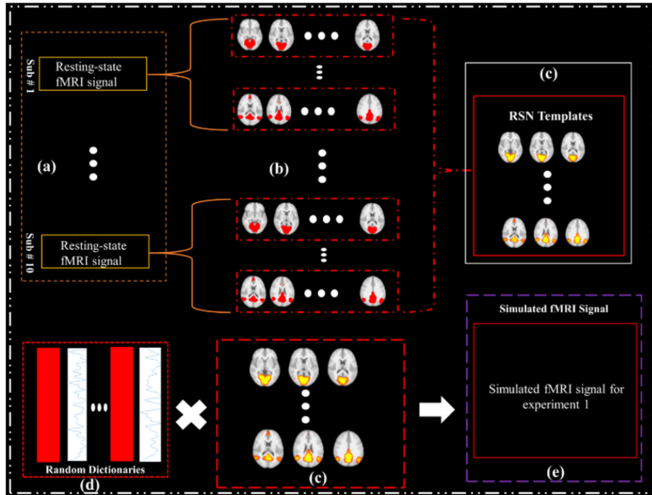
Fig. 3. Computational pipeline for fMRI data synthesis from resting-state fMRI data in experiment 1.

Fig. 4. Computational pipeline for fMRI data synthesis from task-based fMRI data in experiment 2.

ages). For example, in our previous work, we set the dictionary size as 400 to identify the meaningful spatial maps from the task-based fMRI data with temporal length of 284 [42], [43]. However, for spatial ICAs, the number of independent components (ICs) should be smaller than the temporal length. In practice, spatial ICA methods adopt the automatic estimation principle for ICs and the estimated number is usually substantially smaller than the temporal length [54]. More importantly, D in SDL [see Fig. 1(b)] can represent the temporal activities of the FNs. For example, in our prior work, it was reported that specific dictionary atoms have very high similarity with the task-design paradigm curves [42], [43]. Similarly, ICA can also reflect the temporal task design via task-based fMRI [63], [72], [73].

C. Experiment Design for Simulated fMRI Data

In this section, we describe the details of three fMRI data synthesis experiments to generate the ground-truth data for quantitative comparisons. The spatial overlap of all ground-truth data is measured by the Jaccard similarity coefficient to construct the simulated data with different overlap level. The details of the parameters tuning for SDLs are discussed in the Supplemental materials. Briefly, we employ the estimation of maximum value of sparse tradeoff [69] and rank estimator propose in LAMAFIT [70] to avoid capriciousness. For all ICAs in the simulated experiments, only the number of potential independent components is set the same as the number of ground truth, and other parameters are set by default.

Specifically, experiment 1 aims to generate the synthesized fMRI data from resting-state fMRI data, as illustrated in Fig. 3. Firstly, we randomly select the resting-state fMRI signals from 10 subjects [see Fig. 3(a)]. For each subject, we then identify the 10 well-characterized resting state networks (RSNs) by FSL Melodic ICA tools [see Fig. 3(b)]. The templates [see Fig. 3(c)] are derived using (3), based on corresponding individual networks [see Fig. 3(b)]. The detailed visualization and explanation of these 10 RSNs can be referred to Supplemental Fig. 1.
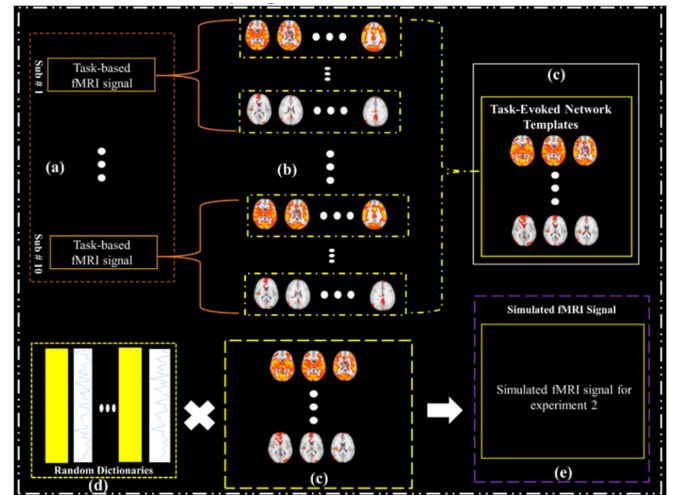
RSN templates are derived by the group-wise statistically determined RSNs using FSL Melodic [see Fig. 3(c)]. After that, 10 time series from the dictionaries derived from resting-state fMRI data in our prior studies [42], [43] are randomly selected [see Fig. 3(d)]. We also provide the simulated experiments using time course identified by Melodic ICA, which can be viewed in our Supplementary documents. The simulated results demonstrate that we obtain the similar results, compared with following 6 simulated experiments. The temporal length of the time series is 1, 200. Given the influence of dictionary correlation [63], we maintain the correlation of random selected temporal dictionary atoms below 10%. Finally, synthetic fMRI signals are computed by multiplying the 10 time series with the 10 templates [see Fig. 3(e)]. In experiment 1, the range of spatial overlap between the 10 RSN networks is 0.00%–14.74%, using (4).

$$\text{Template}_j = \frac{1}{10} \sum_{i=1}^{10} \text{individual\_network}_{k_{i,j}} \quad (3)$$

Experiment 2 aims to generate synthesized fMRI data from task-based fMRI data, as illustrated in Fig. 4. Firstly, 10 subjects' task-based fMRI datasets are randomly selected [see Fig. 4(a)]. Then 10 task-evoked networks are identified by FSL FEAT using HCP motor task fMRI data [see Fig. 4(b)]. The detailed visualization and explanation of these 10 task-evoked networks can be referred to Supplemental Fig. 2. The templates are then derived by the group-wise statistically determined task-evoked networks using FSL FEAT [see Fig. 4(c)]. Similar to experiment 1, 10 time series are randomly selected from the dictionaries obtained in our previous study [see Fig. 4(d)]. The temporal length of the time series is 284. Considering the correlation of dictionary atoms [63], we examine the correlation of random selected temporal dictionary, and select these atoms whose correlations below 10%. The synthetic fMRI signals are then computed as shown in Fig. 4(e). In experiment 2, the range of spatial overlap between the 10 task-evoked networks is 0.07%–70.00%, using (4).
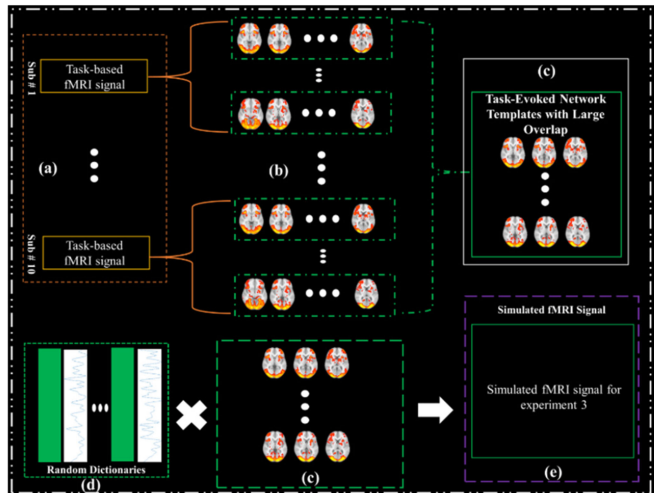
Fig. 5. Computational pipeline for fMRI data synthesis from 3 task-based fMRI data in experiment 3.
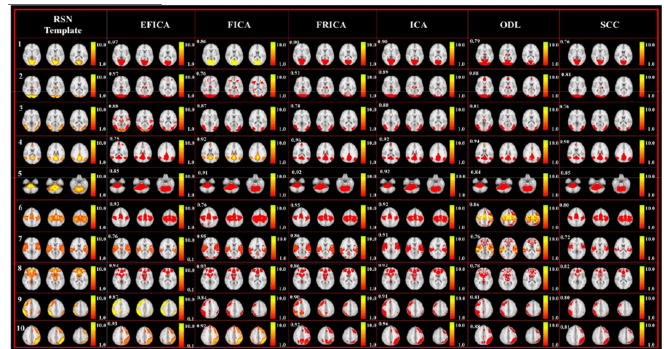

Fig. 6. Reconstructed RSNs by 2 SDL (ODL and SCC) algorithms and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms based on experiment 1. The spatial overlap between the reconstructed RSN and the template is also provided as the upper left corner of each sub-figure.

Experiment 3 aims to generate synthesized fMRI data with significantly overlapping functional networks, as illustrated in Fig. 5. Firstly, three task-based fMRI datasets (Emotion, Gambling and Working Memory) are selected from 10 random subjects. The 10 task-evoked networks are identified by FSL FEAT from the 3 task-based fMRI data [see Fig. 5(b)]. The detailed visualization and explanation of these 10 task-evoked networks can be referred to Supplemental Fig. 3. After that, 10 template networks are computed by averaging the task-evoked networks across 10 subjects [see Fig. 5(c)]. As shown in Supplemental Fig. 3, five pairs of these 10 task-evoked networks show significant spatially overlap with each other. 10 time series are randomly selected from the dictionaries obtained in our previous study (Lv *et al.*, 2015a, b) [see Fig. 5(d)], with temporal length of 284. To avoid the high correlation of dictionaries atoms [63], at first, we calculate the correlation of random selected temporal dictionary atoms, and only the dictionary atoms' correlation below 10% are selected in our simulated experiment. Synthetic fMRI datasets for experiment 3 are then computed as shown in Fig. 5(e). In this experiment, the range of spatial overlap between 10 networks is 30.00%–90.00%, using (4).

III. RESULTS

Based on the synthesized fMRI data at subject-level with ground truth discussed above, three SDL algorithms and four ICA variants are applied to reconstruct functional networks from such synthesized fMRI data. Additional comparison, including all ICAs and ODL, SCC, KSVD is discussed in our Supplementary document. The matching corresponding reconstructed networks is to calculate spatial overlaps with the released standard templates [42], [43], [71]. Through all simulated experiments, the spatial overlaps between reconstructed networks by SDL/ICA algorithms with ground truth networks in the above-mentioned simulations are used to measure the performances of these compared algorithms. A Jaccard similarity coefficient is

applied to calculate the overlap as below [64]:

$$R = \frac{\text{IndividualMap} \cap \text{Tempalte}}{\text{IndividualMap} \cup \text{Template}} \qquad (4)$$

To be specific, a large R indicates the individual reconstructed network is more similar to the template; if R is equal to 1, the networks should be considered as the same; otherwise, a small R means that the individual map is more different from template. We can find the maximum value of R, which demonstrates that we can match the components across multiple subjects. More details of matching procedure can be viewed in our previous research work [42], [43]. In these synthesis experiments, we set the IC number as the same as the number of ground truth and adopt the other default parameters, according to GIFT ICA tools [19], [55]. For ODL, the sparsity level is set as 0.01; the dictionary size is set as 10, and the iteration number is set as 500 [42]–[44]. The details of SDL parameter tuning are included in the Section II of the Supplementary document.

A. Comparison Results on Synthesized fMRI Data in Experiment 1

As mentioned before, experiment 1 synthesized RSNs with spatial overlaps ranging from 0% to 14.74%, which is considered as minor level of overlap. The network reconstruction results by EFICA/FICA/FRICA/ICA and ODL/SCC are shown in Fig. 6. Specifically, the first column shows three representative slices from RSN templates in Fig. 1, in the Supplementary document. Then, in the second to fifth column, the reconstructed RSNs in the same slices by 4 variant ICA (EFICA/FICA/FRICA/ICA) are presented, respectively. The last two columns present the reconstructed RSNs by ODL and SCC. In Fig. 6. Each row represents one RSN network, and the detailed visualization of these 10 RSNs are provided in Supplemental Fig. 1. In addition, all slices of the reconstructed RSNs by ICA algorithms and SDL algorithms are available on our webpage: http://hafni.cs.uga.edu/SDL_ICA_Comparison/ RSN/SDL_ICAs_V1_ReconstructionCompare_RestingSpatial Map_presentation.html

In Fig. 7, we provide the quantitative comparisons for SDL and ICA algorithms. In general, the overlaps between recon-
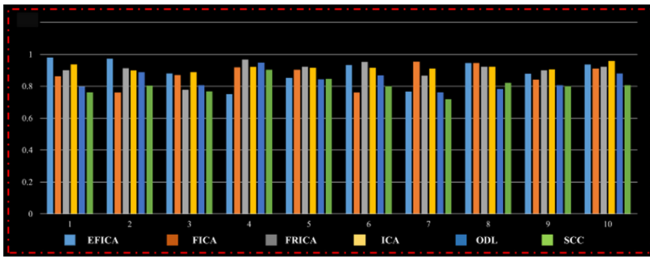
Fig. 7. Quantitative comparisons of 10 reconstruction RSNs obtained by ICA algorithms and SDL algorithms. The 10 light blue, orange, gray and yellow bars represent the overlap values obtained by EFICA, FICA, FRICA and ICA, respectively, compared with the 10 ground truth RSNs. The 10 dark blue and green bars represent the overlap values obtained by ODL and SCC, respectively, compared with the 10 ground truth RSNs.
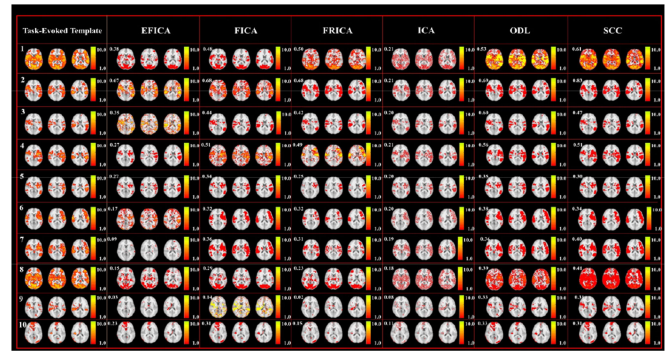

Fig. 8. Reconstructed task-evoked networks by 2 SDL (ODL and SCC) algorithms and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms based on experiment 2. The spatial overlaps between the reconstructed task-evoked networks and the template are also provided as the upper left corner of each sub-figure.
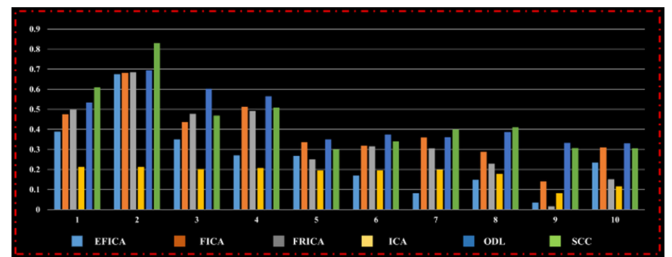

Fig. 9. Quantitative comparison for 10 reconstruction task-evoked networks obtained by ICAs and SDL with 10 task-evoked networks. The 10 light, orange, gray and yellow bars represent the overlap values obtained by EFICA, FICA, FRICA and ICA, compared with the ground truth task-evoked networks. The 10 dark blue and green bars represent the overlap values obtained by ODL and SCC, compared with the ground truth task-evoked networks.

structed spatial patterns by ICAs/SDL and those 10 RSNs templates are all larger than 80%. Specifically, the average overlap values between reconstructed spatial maps and RSNs templates for EFICA, FICA, FRICA, ICA, ODL and SCC are 88.98%, 87.30%, 90.46%, 91.76%, 83.87%, and 80.36%, respectively. It is evident that ICA algorithms perform slightly better than SDL algorithms. Also, it is clear that ODL and SCC have similar reconstruction performances, and all reconstructed spatial patterns by ODL and SCC have large overlaps compared with the ground truth RSNs, as shown in Fig. 6. Moreover, in experiment 4, we examine the performance of reconstruction for another vital SDL algorithm named KSVD. Given the ground truth shown in Supplementary material Fig. 1, we compare the performance between ODL, SCC and KSVD, qualitatively and quantitatively.

B. Comparison Results on Synthesized fMRI Data in Experiment 2

The synthesized task-evoked networks in experiment 2 have spatial overlaps ranging from 0% to 70%, which is considered as the moderate level of overlap. The network reconstruction results by EFICA/FICA/FRICA/ICA and ODL/SCC are shown in Fig. 8. In details, the first column shows three representative slices from task-evoked templates in Supplementary Fig. 2. Then, in the second to fifth column, the reconstructed task-evoked networks in the same slices by 4 variants of ICA (EFICA/FICA/FRICA/ICA) are presented, respectively. The last two columns present the reconstructed task-evoked networks by ODL and SCC. In Fig. 8. each row represents one task-evoked network, and the detailed visualization of these 10 task-evoked templates are provided in Supplemental Fig. 2. In addition, all slices of the reconstructed task-evoked networks by ICA and SDL algorithms are available on our webpage: http://hafni.cs.uga.edu/SDL_ICA_Comparison/TaskEvoked/ SDL_ICAs_V1_SimulationReconstruction_MOTORTaskMap _presentation.html.

In Fig. 9, we provide the quantitative comparisons for SDL and ICA algorithms. In general, the average overlap values between reconstructed task-evoked network spatial maps and template networks for EFICA, FICA, FRICA, ICA, ODL and SCC are 26.17%, 38.56%, 34.18%, 17.98%, 45.29%, and 44.80%,

respectively. It can be seen that the results by SDL algorithms are significantly better than ICA methods. Meanwhile, ODL and SCC have similar reconstruction performances, and all reconstructed network spatial patterns by ODL and SCC have larger overlaps compared with the ground truth task-evoked networks, as shown in Fig. 8. Furthermore, we do the further analysis on the reconstruction performance of 4 variant ICAs (EFICA, FICA, FRICA, ICA) and 3 variant SDLs (ODL, SCC, KSVD) by another 6 group-wised task-evoked networks from Language and Relational task as the ground truth, in the experiment 5. The simulated data is produced by the same framework discussed in Fig. 4 in this paper. Then all algorithms are adopted to decompose the simulated data. The visualization of ground truth can be viewed in Supplementary Fig. 2.

C. Comparison Results on Synthesized fMRI Data in Experiment 3

It was mentioned before that the synthesized task-evoked networks in experiment 3 have spatial overlaps ranging from 30% to 90%, which is considered as the significant or severe level of overlap. The network reconstruction results by EFICA/FICA/FRICA/ICA and ODL/SCC are shown in Fig. 10. Specifically, the first column shows three representative slices from task-evoked templates in Supplementary Fig. 3.
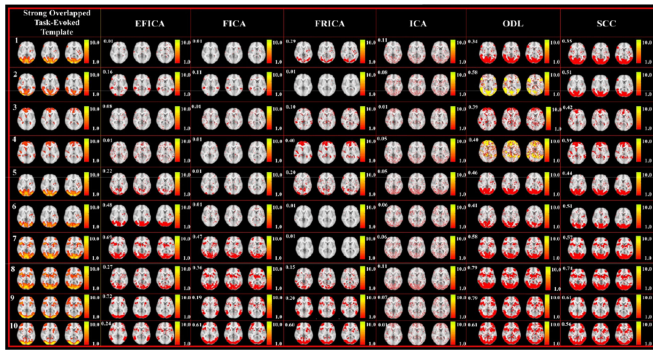
Fig. 10. Reconstructed task-evoked networks which have severe overlaps by 2 SDL (ODL and SCC) and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms based on experiment 3. The spatial overlaps between the reconstructed task-evoked networks and the template networks are also provided as the upper left corner of each sub-figure.
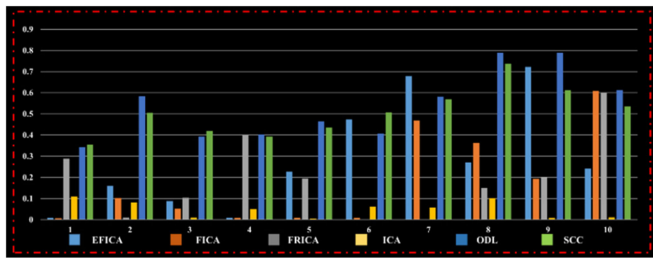

Fig. 11. Quantitative comparison for 10 reconstruction networks obtained by ICAs and ODL with 10 task-evoked networks with strong overlap. The coloring schemes are similar as those in Figs. 7 and 9.

Then in the second to fifth column, the reconstructed task-evoked networks in the same slices by 4 variant ICA (EFICA/FICA/FRICA/ICA) are presented, respectively. The last two columns present the reconstructed task-evoked networks by ODL and SCC. In Fig. 10, each row represents one task-evoked network, and the detailed visualization of these 10 task-evoked templates are provided in Supplemental Fig. 3. In addition, all slices of the reconstructed task-evoked networks by ICA and SDL algorithms are available on our webpage: http://hafni.cs.uga.edu/SDL_ICA_Comparison/TaskEvoked_ StrongOverlap/SDL_ICAs_V1_SimulationReconstruction_ OverlapRateMap_presentation.html

In Fig. 11, we show the quantitative comparisons for SDL and ICA algorithms. On average, the overlap values between reconstructed network spatial maps and network templates for EFICA, FICA, FRICA, ICA, ODL and SCC are 28.73%, 18.18%, 19.47%, 4.92%, 53.65%, and 50.66%, respectively. Again, it is clear that the SDL algorithms perform significantly better than ICA algorithms. Meanwhile, ODL and SCC have similar network reconstruction performances, and all reconstructed network spatial patterns by ODL and SCC have much larger overlap compared with the ground truth task-evoked networks, as shown in Fig. 10. What is more, in our Supplementary material, we compare the reconstructive performance of 7 algorithms (4 ICAs and 3 SDLs) by using the extreme strong group-wised task-evoked networks from WM task. We generate the simulated data by the same framework shown in Fig. 5 in
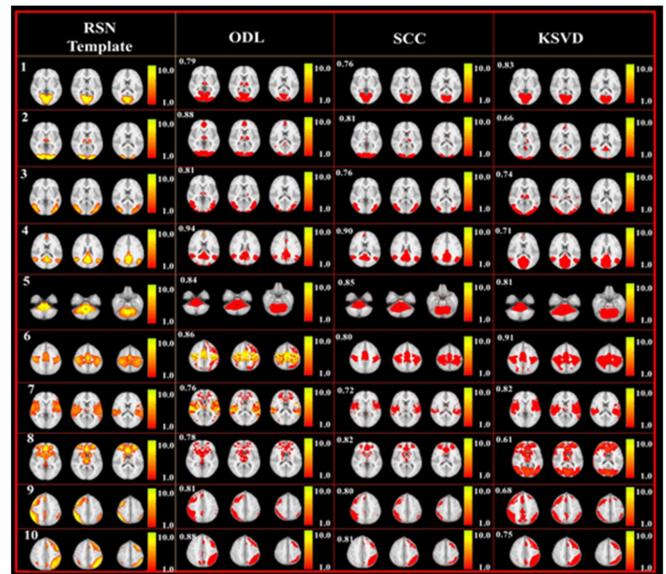

Fig. 12. Reconstructed RSNs which have small overlaps by using 3 SDL (ODL, SCC and KSVD) algorithms, based on the 10 RSN templates of experiment 1. The spatial overlaps between the reconstructed RSNs and the template networks are also provided as the right bottom of each sub-figure.

this paper. And then all 7 algorithms decompose the simulated data and reconstruct the group-wise task-evoked networks. Besides, these constructed networks have the spatial overlap range 70%–85%, and will be adopted to generate the ground truth which is a very difficult task for all 7 algorithms.

In the following sections, we present additional experiments 4, 5 and 6, which concentrates on the qualitative and quantitative comparison with ODL, SCC and KSVD with other 4 variant ICAs. All synthesized fMRI data in experiments 4, 5 and 6 is generated via the same framework discussed in Section II, part C.

D. Comparison Results on Synthesized fMRI Data as Experiment 4

As mentioned before, in this section, since KSVD is a vital method using singular value decomposition (SVD) to train the dictionary at the early stage of SDL method, we provide the qualitative and quantitative analysis of three variant SDLs such ODL, SCC and KSVD using the same synthesized fMRI signal in experiment 1.

In Fig. 12, based on selected 10 RSN templates, shown as Supplementary Fig. 1 as our ground truth in experiment 1, we compare the performance of reconstruction by 3 SDLs including ODL, SCC and KSVD. The qualitative and quantitative comparison is presented in the following Figs. 12 and 13, in an effort to comprehensively analyze the performance of functional networks (FNs) reconstruction by SDLs and ICAs, by comparing with the results shown in Figs. 6 and 7 in this paper.

It can be learned that the results of KSVD algorithm is very similar to ODL and SCC. For RSN templates such as #6 and #7, the reconstructed networks by KSVD have the larger spatial overlap, compared with ODL and SCC. Meanwhile, for RSN
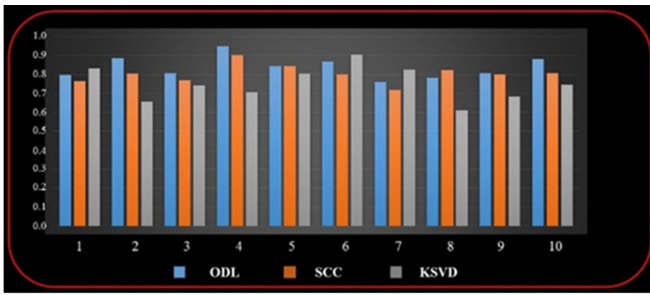
Fig. 13.   Quantitative comparison for 10 reconstruction networks obtained by ODL, SCC and KSVD with 10 RSN templates with small overlap. The 10 blue, orange and gray bars represent the overlap values obtained by ODL, SCC, and KSVD, respectively, compared with the 10 ground truth RSNs.
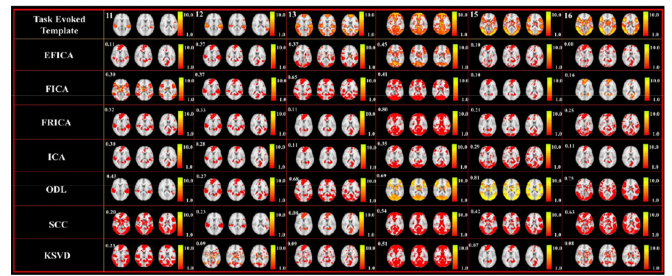

Fig. 14.   Reconstructed task-evoked networks by 3 SDL algorithms and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms based on experiment 2. The spatial overlaps between the reconstructed task-evoked networks and the template are also provided as the right bottom of each sub-figure.

template #9, the ODL and SCC have better reconstruction performances than KSVD. In a conceptual level, the method of training dictionary of KSVD is fundamentally based on SVD, which is also an important method applied widely for ICA. So, KSVD can provide a qualified reconstruction performance when the RSN templates have a small overlap. By increasing the spatial overlap of RSN templates in our experiments, the performance of KSVD should be degrading, compared with ODL and SCC. We provide the quantitative analysis for these reconstruction results shown in Fig. 13.

E. Comparison Results on Synthesized fMRI Data as Experiment 5

In the experiment 5, there are totally 6 task-evoked networks identified by the FSL FEAT from the HCP language and relational task in our previous HAFNI project. Since we aim to examine the reconstruction performance of 3 SDLs and 4 ICAs as the experiment 2, we select 6 task-evoked networks based on 6 averaged networks (cope #1-#6) as the reference networks. Meanwhile, based on our previous HAFNI project, we adopt 6 randomly selected time series dictionaries from task-based fMRI as the temporal signals as the experiment 2. The visualization of the 6-selected task-evoked networks is provided in Fig. 4 in the Supplementary material.

In Fig. 15, we provide the qualitative comparisons for SDL and ICA algorithms. In general, the average overlap values between reconstructed task-evoked network spatial maps and template networks for EFICA, FICA, FRICA, ICA, ODL, SCC and KSVD are 24.82%, 33.24%, 33.61%, 23.90%, 60.67%, 35.13%, and 17.86%, respectively. In general, it can be clearly seen that the results by ODL and SCC algorithms are better than ICA methods. The reconstruction performance of KSVD is not as good as ODL and SCC. And, ICA and KSVD have the similar reconstruction performance, since both ICA and ODL adopts SVD to train the weight matrix or dictionary; SVD is very sensitive to strong correlated components such as template #15 and #16. Meanwhile, ODL and SCC have similar reconstruction performances in template #14, #15 and #16, and all reconstructed network spatial patterns by ODL and SCC have larger overlaps compared with the ground truth task-evoked networks, as shown in Fig. 14.
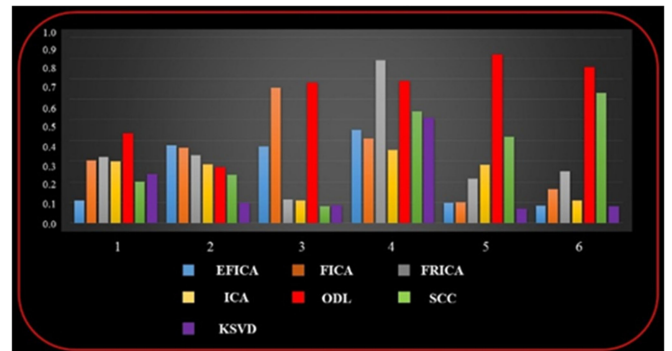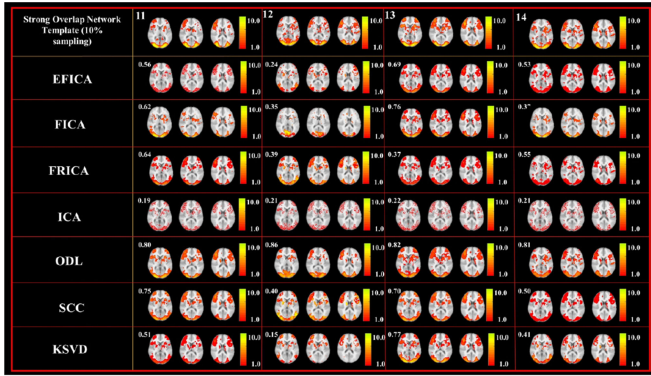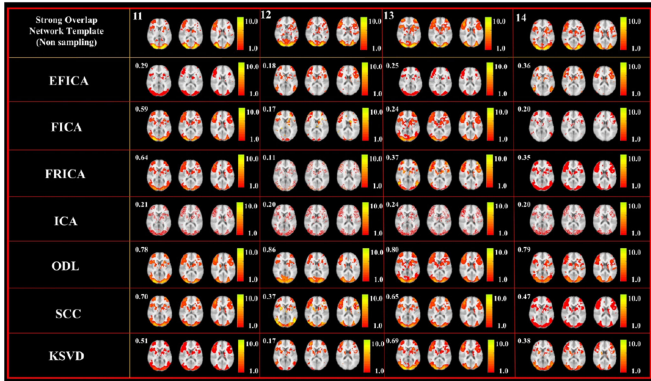

Fig. 15.   Quantitative comparison for 10 reconstruction networks obtained by 3 SDL (ODL, SCC and KSVD) algorithms and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms with 6 task-evoked templates with moderate overlap. The 6 blue, orange, gray and yellow bars represent the overlap values obtained by EFICA, FICA, FRICA and ICA, respectively, compared with the 6 ground truth task-evoked networks in Supplementary Fig. 4. The 6 red, green and purple bars represent the overlap values obtained by ODL and SCC, respectively, compared with the 6-ground truth task-evoked networks.

F. Comparison Results on Synthesized fMRI Data as Experiment 6

In the following experiment, we focus on examination of reconstruction performance via very strongly overlapped synthesized fMRI signal and cross-validation. We select 4 very strongly overlap FNs from working memory task (cope #1-cope #4). The spatial overlap between each other is approximately 90% so that it is easy to differentiate the reconstruction performance of 3 SDLs and 4 ICAs in the extreme case. And we employ a repeated random sub-sampling validation, also known as the Monte Carlo cross-validation, to examine the reconstruction performance for all ICAs and SDLs [68]. This method randomly splits the dataset into training and validation data. For each such split, the model is fit to the training data, and predictive accuracy is assessed using the validation data. The results are then averaged over the splits [68]. ICA and SDL are different from classification algorithms, both algorithms are data-driven, since they do not require the training process to provide the model parameters such as the deep neural network adopted for the classification. However, the random cross-validation can provide an opportunity to examine the influence for ICAs and SDLs. In this experiment, we employ the very strongly

(a)

(b)

Fig. 16.   (a) Averaged reconstructed task-evoked networks by 3 SDL (ODL, SCC and KSVD) algorithms and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms using Monte Carlo cross-validation, based on the extreme strong overlap templates shown in Supplementary Fig. 5. The spatial overlaps between the reconstructed task-evoked networks and the template are also provided as the upper left corner of each sub-figure. (b) Averaged reconstructed task-evoked networks by 3 SDL (ODL, SCC and KSVD) algorithms and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms without using Monte Carlo cross-validation, based on the extreme strong overlap templates shown in Supplementary Fig. 5. The spatial overlaps between the reconstructed task-evoked networks and the template are also provided as the upper left corner of each sub-figure.

overlap components; therefore, we can be better understanding the difference between 4 ICAs and 3 SDLs by compared the reconstruction results. To be detailed, we repeated the Monte Carlo cross-validation for 100 times in total, and, in each split, we only randomly take 10% data as the sub-sampling (temporal sub-sampling) from 10 subjects' simulated fMRI signal as the input simulated task-based fMRI data. In the Fig. 16(a), we present the averaged identified reconstructed brain networks using 4 ICAs and 3 SDLs at subject-level. And in the Fig. 16(b), we provide the averaged reconstructed results using the same repeated random cross-validation method and parameters but no temporal sub-sampling. Also, we have provided all reconstruction results for each simulated subject from #1 to #10, using Monte Carlo cross-validation. All comparisons can be viewed in Supplementary Figs. 8–17.

In Fig. 16(a) and (b), we show the qualitative comparisons for SDL and ICA algorithms. Using Monte Carlo cross-validation, ICAs can provide much better reconstruction results. Some
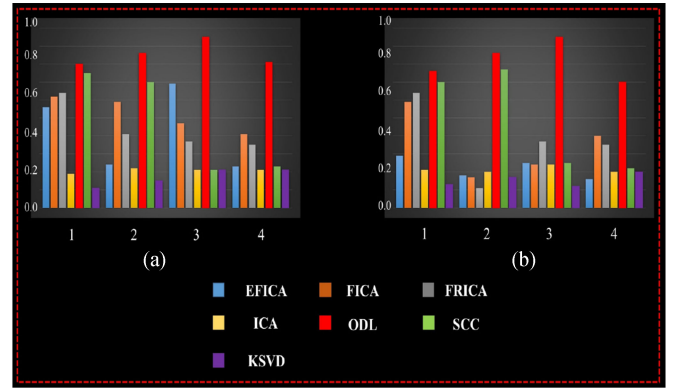

Fig. 17.   Quantitative comparison for 4 reconstruction networks obtained by 3 SDL (ODL, SCC and KSVD) and 4 ICA (EFICA, FICA, FRICA, and ICA) algorithms with 4 ground truth task-evoked networks with extreme strong overlap, using 10% repeated random cross-validation (part (a)) and without repeated random cross-validation (part (b)). The 4 blue, orange, gray and yellow bars represent the overlap values obtained by EFICA, FICA, FRICA and ICA, respectively, compared with the 6-ground truth task-evoked networks. The 4 red, green and purple bars represent the overlap values obtained by ODL, SCC and KSVD, respectively, compared with the 4 ground truth task-evoked networks.

reconstruction overlap (ground truth #1 and #2 of FICA) is even comparable with SDLs. But, without using Monte Carlo cross-validation, it is clear that the SDL algorithms ODL and SCC perform significantly better than ICA algorithms. Meanwhile, ODL has the best reconstruction performances in these very strongly overlapped templates in both two parts of this experiment, and all reconstructed network spatial patterns by ODL have much larger overlap compared with the ground truth task-evoked networks, as quantitatively shown in Fig. 17.

As a result, our conclusion is that ICA can perform efficiently only using massive repeated cross-validation, which indicates the ICA can achieve the comparable reconstruction with SDLs only using repeated random cross-validation that is very time consuming. It is also worth noticing that ODL is very robust, whether employed or unemployed the repeated cross-validation method. Obviously, to explore the current massive human brain fMRI data, SDLs, especially for ODL, have the obvious superiority.

G. Statistics Analysis on Experiment 1, 2 and 3 via Multiple Subjects

As discussed before, we generated 10 different subjects' simulated resting-state and tsk-based fMRI data to compare the reconstruction performance of ICAs and SDLs. Hereby we provide the further statistics analysis by two-sample *t-test*. In Fig. 18, we provide the t-test between each two algorithms from ICAs and SDLs in the experiment 1, 2 and 3, respectively. By review all three matrices, the p-value pair of ICAs and SDLs is continuously decreased. In the last matrix, most p-value pairs of ICAs and SDL are decreased to small values (<0.05), which further indicates that the significant differences between ICAs and SDLs, when these two types of methods are applied on the very strongly overlapped data.
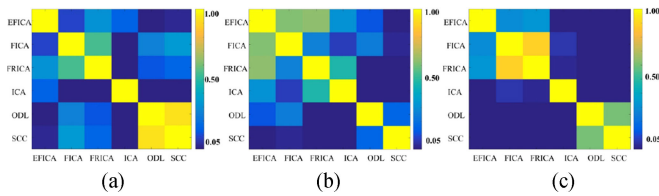
Fig. 18.   Statistics analysis of reconstruction accuracy (spatial overlap) on all 10 subjects' simulated fMRI data in experiments 1 (a), 2 (b) and 3 (c) discussed before in our paper. We employ the two-sample *t-test* (*alpha* = 0.05) to examine the difference between each subject's spatial overlap obtained by two algorithms. And we visualize the averaged *p-value* of each pair of two algorithms using two-sample *t-test*. For example, a single element in first row and fifth column in matrix (a) represents the averaged *p-value* based on the *t-test* applied on 10 subjects' reconstruction accuracy obtained by ODL and EFICA.

IV. CONCLUSION

In this study, we proposed three simulation experiments to synthesize fMRI data with ground-truth in order to compare the network reconstruction performance by three SDL and four ICA algorithms. Overall, SDL algorithms can provide consistently high reconstruction accuracy for all kinds of fMRI data including both resting state fMRI and task-based fMRI data, no matter how strong spatial overlaps brain networks have. Our recommendation is that SDL algorithms are more universally efficient in reconstructing brain networks in various combinations of scenarios in the era of fMRI-based connectomics, given the known extensive overlaps of large-scale functional networks in the brain [14], [16]–[18], [30].

However, it should be pointed out that SDL algorithms have their own caveats. First of all, there are lacking solid theoretic guidelines on how to select and optimize the parameters such as the sparsity level and dictionary size. Future efforts should be devoted to investigating the optimal ways to estimate and select such key parameters [41]–[44]. In addition, in the future, we will employ group-wise and structured sparsity to perform further comparison with corresponding ICA methods [74], [75]. Second, the regularity and variability of those reconstructed functional brain networks by SDL algorithms across different fMRI scans and individual brains should be assessed and modeled [44], [46], [49], [65]. Finally, it is worth noting that SDL methods are still the so-called "shallow" models [66], [67]. In contrast, deep neural network structures, such as the recently developed Deep Convolutional Auto-Encoder (DCAE) [67], can take the advantages of both data-driven approach and hierarchical feature abstraction ability. We envision that deep learning models would potentially further improve SDL methods in reconstructing functional brain networks in the future.

REFERENCES

[1]   E. Bingham, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *Int. J. Neural Syst.*, vol. 10, no. 1, pp. 1–8, Feb. 2000.
[2]   P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
[3]   A. J. Bell, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, no. 6, pp. 1129–1159, Nov. 1995.

[4]   A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.
[5]   A. Hyvärinen, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, no. 4, pp. 411–430, Jun. 2000.
[6]   T. Adali, "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, Apr. 2014.
[7]   M. J. McKeown, "Detection of consistently task-related activations in fMRI data with hybrid independent component analysis," *Neuroimage*, vol. 11, no. 1, pp. 24–35, Jan. 2000.
[8]   V. D. Calhoun, "A method for making group inferences from functional MRI data using independent component analysis," *Hum. Brain Mapping*, vol. 14, no. 3, pp. 140–151, Aug. 2001.
[9]   V. D. Calhoun, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *Neuroimage*, vol. 45, no. 1, pp. 163–172, Mar. 2009.
[10] V. D. Calhoun, "Multisubject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery," *IEEE Rev. Biomed. Eng.*, vol. 5, pp. 60–73, 2012.
[11] C. F. Beckmann, "Investigations into resting-state connectivity using independent component analysis," *Philos. Trans. Roy. Soc. Lond. B, Biol. Sci.*, vol. 360, no. 1457, pp. 1001–1013, May 2005.
[12] C. F. Beckmann, "Modelling with independent components," *Neuroimage*, vol. 62, no. 2, pp. 891–901, Feb. 2012.
[13] K. D. Harris, "Cortical connectivity and sensory coding," *Nature*, vol. 503, no. 7474, pp. 51–58, Nov. 2013.
[14] T. D. Hermansen, "Human development XI: The structure of the cerebral cortex. Are there really modules in the brain?" *Sci. World J.*, vol. 7, pp. 1922–1929, Jan. 2007.
[15] R. Perin, "A synaptic organizing principle for cortical neuronal groups," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 13, pp. 5419–5424, Oct 2011.
[16] J. M. Fuster, "Cortex and memory: Emergence of a new paradigm," *J. Cogn. Neurosci.*, vol. 21, no. 11, pp. 2047–2072, Sep. 2009.
[17] J. M. Fuster, "Past makes future: Role of pFC in prediction," *J. Cogn. Neurosci.*, vol. 4, pp. 639–654, Feb. 2015.
[18] N. V. Swindale, "Orientation tuning curves: Empirical description and estimation of parameters," *Biol. Cybern.*, vol. 78, no. 1, pp. 45–56, Jan. 1998.
[19] V. D. Calhoun, "A method for comparing group fMRI data using independent component analysis: Application to visual, motor and visuomotor tasks," *Magn. Reson. Imag.*, vol. 22, no. 9, pp. 1181–1191, Nov. 2004.
[20] A. Domagalik, "Neural networks related to pro-saccades and anti-saccades revealed by independent component analysis," *Neuroimage*, vol. 62, no. 3, pp. 1325–1333, Sep. 2012.
[21] D. I. Kim, "Dysregulation of working memory and default-mode networks in schizophrenia using independent component analysis, an fBIRN and MCIC study," *Hum. Brain Mapping*, vol. 30, no. 11, pp. 3795–3811, May 2009.
[22] D. I. Kim, "Auditory oddball deficits in schizophrenia: An independent component analysis of the fMRI multisite function BIRN study," *Schizophrenia Bull.*, vol. 35, no. 1, pp. 67–81, Jan. 2009.
[23] M. Li, "OI and fMRI signal separation using both temporal and spatial autocorrelations," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1917–1926, May 2010.
[24] M. Li, "Including signal intensity increases the performance of blind source separation on brain imaging data," *IEEE Trans. Med. Imag.*, vol. 34, no. 2, pp. 551–563, Oct. 2015.
[25] M. Menz, "Dissociating networks of imitation," *Hum. Brain Mapping*, vol. 30, no. 10, pp. 3339–3350, Apr. 2009.
[26] P. L. S. Jacques, "Dynamic neural networks supporting memory retrieval," *Neuroimage*, vol. 57, no. 2, pp. 608–616, Jul. 2011.
[27] H. van Wageningen, "The effects of the glutamate antagonist memantine on brain activation to an auditory perception task," *Hum. Brain Mapping*, vol. 30, no. 11, pp. 3616–3624, May 2009.
[28] X. Wu, "Multiple neural networks supporting a semantic task: An fMRI study using independent component analysis," *Neuroimage*, vol. 45, no. 4, pp. 1347–1358, May 2009.
[29] S. Zhang, "Functional networks for cognitive control in a stop signal task: Independent component analysis," *Hum. Brain Mapping*, vol. 33, no. 1, pp. 89–104, Mar. 2012.
[30] J. Xu, "Large-scale functional network overlap is a general property of brain functional organization: Reconciling inconsistent fMRI findings from general-linear-model-based analyses," *Neurosci. Biobehav. Rev.*, vol. 71, pp. 83–100, Dec. 2016.

[31] E. Beldzik, "Contributive sources analysis: A measure of neural networks' contribution to brain activations," *Neuroimage*, vol. 76, pp. 304–312, Aug. 2013.

[32] R. M. Braga, "Echoes of the brain within default mode, association, and heteromodal cortices," *J. Neurosci.*, vol. 33, no. 35, pp. 14031–14039, Aug. 2013.

[33] F. Geranmayeh, "Overlapping networks engaged during spoken language production and its cognitive control," *J. Neurosci.*, vol. 34, no. 26, pp. 8728–8740, Jun. 2014.

[34] R. Leech, "Echoes of the brain within the posterior cingulate cortex," *J. Neurosci.*, vol. 32, no. 1, pp. 215–222, Jan. 2012.

[35] J. Xu, "Spatial ICA reveals functional activity hidden from traditional fMRI GLM-based analyses," *Front. Neurosci.*, vol. 7, Aug. 2013, Art. no. 154.

[36] J. Xu, "Task-related concurrent but opposite modulations of overlapping functional networks as revealed by spatial ICA," *Neuroimage*, vol. 79, pp. 62–71, Oct. 2013.

[37] J. Xu, "Opposite modulation of brain functional networks implicated at low vs. high demand of attention and working memory," *PLoS One*, vol. 9, no. 1, Jan. 2014, Art. no. e87078.

[38] R. L. Buckner, "Opportunities and limitations of intrinsic functional connectivity MRI," *Nature Neurosci.*, vol. 16, no. 7, pp. 832–837, May 2013.

[39] J. Mairal, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, Jan. 2010.

[40] J. Lv, "Modeling task FMRI data via supervised stochastic coordinate coding," in *Proc. Int. Conf. Med. Image Comput. Computer-Assisted Interv.*, Springer, Chem, Switzerland, 2015, vol. 9349, pp. 239–246.

[41] J. Lv, "Task fMRI data analysis based on supervised stochastic coordinate coding," *Med. Image Anal.*, vol. 38, pp. 1–16, May 2017.

[42] J. Lv, "Holistic atlases of functional networks and interactions reveal reciprocal organizational architecture of cortical function," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1120–1131, Nov. 2015.

[43] J. Lv, "Sparse representation of whole-brain FMRI signals for identification of functional networks," *Med. Image Anal.*, vol. 20, no. 1, pp. 112–134 Feb. 2015.

[44] J. Lv, "Assessing effects of prenatal alcohol exposure using group-wise sparse representation of fMRI data," *Psychiatry Res., Neuroimag.*, vol. 233, no. 2, pp. 254–268, Aug. 2015.

[45] B. Ge, "Signal sampling for efficient sparse representation of resting state FMRI data," *Brain Imag. Behav.*, vol. 10, no. 4, pp. 1206–1222, Dec. 2016.

[46] X. Hu, "Sparsity-constrained fMRI decoding of visual saliency in naturalistic video streams," *IEEE Trans. Auton. Mental Develop.*, vol. 7, no. 2, pp. 65–75, Mar. 2015.

[47] S. Zhang, "Characterizing and differentiating task-based and resting state fMRI signals via two-stage sparse representations," *Brain Imag. Behav.*, vol. 10, no. 1, pp. 21–32, Mar. 2016.

[48] S. Zhao, "Supervised dictionary learning for inferring concurrent brain networks," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 2036–2045, Apr. 2015.

[49] Y. Zhao, "Connectome-scale group-wise consistent resting-state network analysis in autism spectrum disorder," *Neuroimage, Clin.*, vol. 12, pp. 23–33, Feb. 2016.

[50] M. Makkie, "HAFNI-enabled largescale platform for neuroimaging informatics (HELPNI)," *Brain Informat.*,vol. 2, no. 4, pp. 225–238, Dec. 2015.

[51] L. Pessoa, "Beyond brain regions: Network perspective of cognition–emotion interactions," *Behav. Brain Sci.*, vol. 35, pp. 158–159, Jun. 2012.

[52] A. Beck, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[53] B. B. Biswal, "Blind source separation of multiple signal sources of fMRI data sets using independent component analysis," *J. Comput. Assisted Tomography*, vol. 23, no. 2, pp. 265–271, Mar. 1999.

[54] D. Hu, "Unified SPM–ICA for fMRI analysis," *Neuroimage*, vol. 25, no. 3, pp. 746–755, Apr. 2005.

[55] E. A. Egolf, "Group ICA of fMRI toolbox gift," *Neuroimage*, vol. 22, pp. 14–17, 2004.

[56] M. J. McKeown, "Independent component analysis of fMRI data: Examining the assumptions," *Hum. Brain Mapping*, vol. 6, nos. 5–6, pp. 368–372, 1998.

[57] M. Aharon, "$ rm k $-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Oct. 2006.

[58] G. H. Golub, "Singular value decomposition and least squares solutions," *Numer. Math.*, vol. 14, no. 5, pp. 403–420, Apr. 1970.

[59] S. Kullback, *Information Theory and Statistics*. Chelmsford, MA, USA: Courier Corp., 1959, pp. 329–334.

[60] C. E. Shannon, "The mathematical theory of communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, Jan. 2001.

[61] E. G. Learned-Miller, "ICA using spacings estimates of entropy," *J. Mach. Learn. Res.*, vol. 4, pp. 1271–1295, Dec. 2003.

[62] A. Chen, "Efficient independent component analysis," *Ann. Statist.*, vol. 34, no. 6, pp. 2825–2855, 2006.

[63] V. D. Calhoun, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Hum. Brain Mapping*, vol. 13, no. 1, pp. 43–53, May 2001.

[64] A. B. Waites, "Effect of prior cognitive state on resting state networks measured with functional connectivity," *Hum. Brain Mapping*, vol. 24, no. 1, pp. 59–68, Jan. 2005.

[65] D. Ren, "3D functional brain network classification using convolutional neural networks," in *Proc. IEEE 14th Int. Symp. Biomed. Imag.*, Jun. 2017, pp. 1217–1221.

[66] H. Huang, "Latent source mining in FMRI data via deep neural network," in *Proc. IEEE 13th Int. Symp. IEEE Biomed. Imag.*, Apr. 2016, pp. 638–641.

[67] H. Huang, "Modeling task fMRI data via deep convolutional autoencoder," *Info. Process. Med. Imag.*, pp. 411–424, 2017.

[68] W. R. Gilks *Markov Chain Monte Carlo in Practice*. Boca Raton, FL, USA: CRC Press, 1995, pp. 5–15.

[69] J. Liu, "An efficient algorithm for a class of fused lasso problems," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, Jul. 2010, pp. 323–332.

[70] Z. Wen, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Math. Program. Comput.*, vol. 4, pp. 333–361, Jul. 2012.

[71] J. S. Damoiseaux, "Consistent resting-state networks across healthy subjects," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 37, pp. 13848–13853, Sep. 2006.

[72] E. A. Allen, "Capturing inter-subject variability with group independent component analysis of fMRI data: A simulation study," *Neuroimage*, vol. 59, no. 4, pp. 4141–4159, Feb. 2012.

[73] E. A. Allen, "Tracking whole-brain connectivity dynamics in the resting state," *Cerebral Cortex*, vol. 24, no. 3, pp. 663–676, Mar. 2014.

[74] J. Yao, "An efficient algorithm for dynamic MRI using low-rank and total variation regularizations," *Med. Image Anal.*, vol. 44, pp. 14–27, Feb. 2018.

[75] J. Huang, "Learning with structured sparsity," *J. Mach. Learn. Res.*, vol. 12, pp. 3371–3412, Nov. 2011.

[76] Y. Du, "Artifact removal in the context of group ICA: A comparison of single-subject and group approaches," *Hum. Brain Mapping*, vol. 37, no. 3, pp. 1005–1025, 2016.

[77] Y. Du, "A group ICA based framework for evaluating resting fMRI markers when disease categories are unclear: Application to schizophrenia, bipolar, and schizoaffective disorders," *Neuroimage*, vol. 122, pp. 272–280, 2015.

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

[79] Y. H. Kim *et al.*, "Iterative approach of dual regression with a sparse prior enhances the performance of independent component analysis for group functional magnetic resonance imaging (fMRI) data," *Neuroimage*, vol. 63, no. 4, pp. 1864–1889, 2012.

[80] S. Zhao, "Decoding auditory saliency from brain activity patterns during free listening to naturalistic audio excerpts," *Neuroinformatics*, vol. 4, pp. 1–16, 2018.