

Sparse Representation of Whole-brain fMRI Signals for Identification of Functional Networks

Jinglei Lv^{1,2*}, Xi Jiang^{2*}, Xiang Li^{2*}, Dajiang Zhu², Hanbo Chen², Tuo Zhang^{1,2}, Shu Zhang²,
Xintao Hu¹, Junwei Han¹, Heng Huang³, Jing Zhang⁴, Lei Guo¹, Tianming Liu²

¹School of Automation, Northwestern Polytechnical University, China; ²Cortical Architecture Imaging and Discovery Lab, Department of Computer Science and Bioimaging Research Center, The University of Georgia, Athens, GA; ³Department of Computer Science and Engineering, University of Texas at Arlington, TX; ⁴Department of Statistics, Yale University, CT. *These authors contributed equally to this work.

ABSTRACT

There have been several recent studies that used sparse representation for fMRI signal analysis and activation detection based on the assumption that each voxel's fMRI signal is linearly composed of sparse components. Previous studies have employed sparse coding to model functional networks in various modalities and scales. These prior contributions inspired the exploration of whether/how sparse representation can be used to identify functional networks in a voxel-wise way and on the whole brain scale. This paper presents a novel, alternative methodology of identifying multiple functional networks via sparse representation of whole-brain task-based fMRI signals. Our basic idea is that all fMRI signals within the whole brain of one subject are aggregated into a big data matrix, which is then factorized into an over-complete dictionary basis matrix and a reference weight matrix via an effective online dictionary learning algorithm. Our extensive experimental results have shown that this novel methodology can uncover multiple functional networks that can be well characterized and interpreted in spatial, temporal

and frequency domains based on current brain science knowledge. Importantly, these well-characterized functional network components are quite reproducible in different brains. In general, our methods offer a novel, effective and unified solution to multiple fMRI data analysis tasks including activation detection, de-activation detection, and functional network identification.

Keywords: Task-based fMRI, activation, intrinsic networks, connectivity.

1. INTRODUCTION

Task-based fMRI has been widely used to identify brain regions that are functionally involved in specific task performance, and has significantly advanced our understanding of functional localizations within the brain (Logothetis, 2008; Friston, 2009). In the human brain mapping community, a variety of fMRI time series analysis methods have been developed for activation modeling and detection, such as correlation analysis (Bandettini et al., 1993), general linear model (GLM) (Friston et al., 1994; Worsley et al., 1997), principal component analysis (PCA) (Andersen et al., 1999), Markov random field (MRF) models (Descombes et al., 1998), mixture models (Hartvig and Jensen, 2000), independent component analysis (ICA) (McKeown et al., 1998), wavelet algorithms (Bullmore et al., 2003; Shimizu et al., 2004), autoregressive spatial models (Woolrich et al., 2001), Bayesian approaches (Bowman et al., 2008), and empirical mean curve decomposition (Deng et al., 2012). Among all of these computational methods, the GLM (Friston et al., 1994; Worsley et al., 1997) is one of the most widely used methods due to its effectiveness, simplicity, robustness and wide availability.

Recently, inspired by the successes of using sparse representation for signal and pattern analysis in the machine learning and pattern recognition fields (Wright et al., 2010), there have been several studies that used sparse representation for fMRI signal analysis and activation detection (e.g., Li et al., 2009; Lee et al., 2011; Li et al., 2012; Oikonomou et al., 2012; Lee et al., 2013; Abolghasemi et al., 2013; Lv et al.,

2013) based on the assumption that the components of each voxel’s fMRI signal are sparse and the neural integration of those components is linear. Actually, the human brain function intrinsically involves multiple complex processes with population codes of neuronal activities (Olshausen 1996; Olshausen and Field, 2004; Quiroga et al., 2008). In the brain science field, a variety of research studies have supported that when determining neuronal activity, sparse population coding of a set of neurons seems more effective than independent exploration (Daubechies et al., 2009). That is, a sparse set of neurons encode specific concepts rather than responding to the input independently (Daubechies et al., 2009). Therefore, it is natural and well-justified to explore sparse representations to describe fMRI signals of the brain. In parallel, significant amount of research efforts from the machine learning and pattern recognition fields has been recently devoted to sparse representations of signals and patterns (Donoho 2006; Huang and Aviyente, 2006; Wright et al., 2008; Wright et al., 2010; Mairal et al., 2010; Yang et al., 2011), and remarkable achievements have been made for both compact high-fidelity representation of the signals and effective extraction of meaningful patterns (Wright et al., 2010). However, despite recent successes of using sparse representation for fMRI signal analysis and activation detection in the human brain mapping field (e.g., Li et al., 2009; Lee et al., 2011; Li et al., 2012; Oikonomou et al., 2012; Lee et al., 2013; Abolghasemi et al., 2013; Lv et al., 2013), it has been rarely explored whether/how sparse representation of fMRI signals can be utilized to infer functional networks within the whole brain at the voxel scale.

To bridge the abovementioned gap, in this paper, we present a novel, alternative methodology which employs sparse representation of whole-brain fMRI signals for functional networks identification in task-based fMRI data. The basic idea here is that we aggregate all of the dozens of thousands of task-based fMRI signals within the whole brain from one subject into a big data matrix, and factorize it by an over-complete dictionary basis matrix and a reference weight matrix via an effective online dictionary learning algorithm (Mairal et al., 2010). Our rationale is that during task performance, there could be multiple, e.g., dozens or even hundreds of, functionally active networks that contribute to the fMRI blood oxygen level dependent (BOLD) signals of the whole brain. The main objectives of this work are to explore the

following three questions: 1) what could these atomic functional networks be; 2) what spatial, temporal and frequency characteristics could those functional network components exhibit; and 3) how do they contribute to the compositions of dozens of thousands of fMRI signals within the whole brain. Given the proven remarkable capability of sparse representation in uncovering meaningful patterns from large amount of data (Wright et al., 2010), we hypothesize that sparse representation of whole-brain fMRI signals via dictionary learning can simultaneously address the abovementioned three questions. In particular, we hypothesize that the identified functional network components can be further characterized and interpreted by existing brain science knowledge, as well as by existing structural and functional brain atlases. To test the above hypotheses, as an example, Fig. 1 illustrates our rationale and the computational methodology. In Fig. 1, three exemplar identified network components including the task related one (Faraco et al., 2011) (yellow), the anti-task related one (or de-activation, Archer et al., 2003; Tomasi et al., 2006) (blue), and the default mode network (DMN) (Raichle and Snyder, 2007) (red), as well as their overlapped areas including task + anti-task (pink), task + DMN (green), anti-task + DMN (cyan), and task + anti-task + DMN (brown), are shown on the inflated cortical surface. It is noted that the visualization on original surface of Fig. 1 is shown in Supplemental Fig. 13(I). It is shown that these three network components exhibit spatially distinct but overlapping distribution patterns, illustrating that multiple functionally active networks simultaneously contribute to the fMRI BOLD signals of the whole brain and that the online dictionary learning method has the great promise to concurrently address the abovementioned three questions.

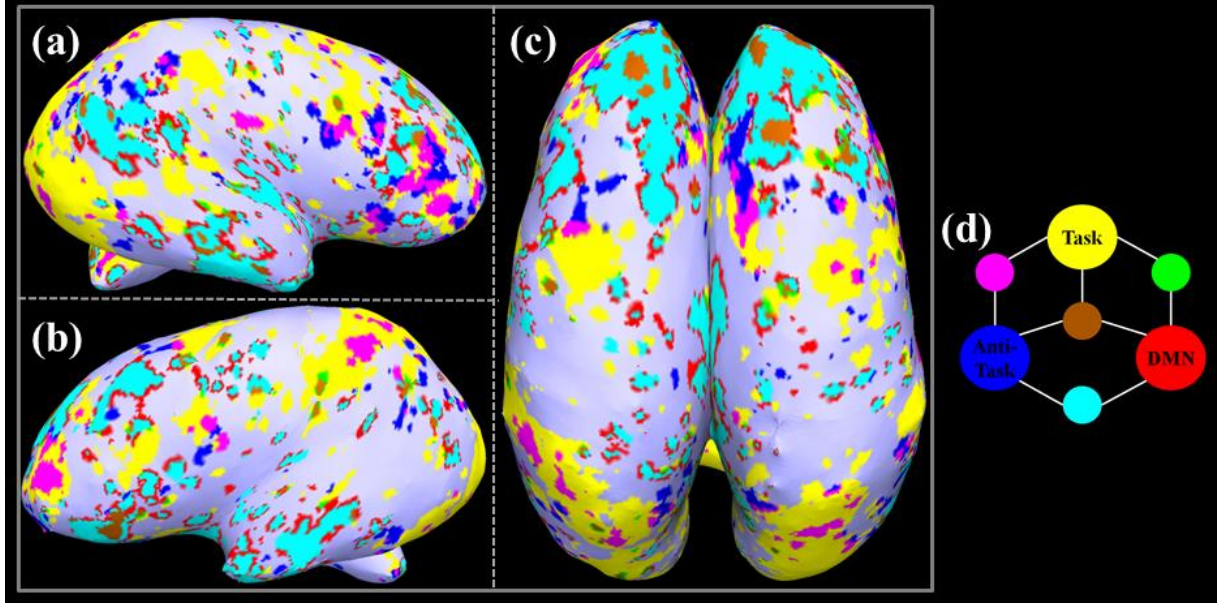


Figure 1. Illustration of spatial distributions of three dictionary components of interest (COI) onto the inflated cortical surface. This illustration is based on a working memory task-based fMRI dataset (Faraco et al., 2011; Zhu et al., 2012). There are task-related component (yellow), anti-task related component (blue) and default mode network (DMN) component (red). (a-c) show different views of representing the spatial distribution patterns of these three network components. (d) demonstrates the color scheme of representing different components and their overlaps. For examples, the regions belonging to both the anti-task and DMN components are represented by cyan, and the green color represents the overlapped areas of the task and DMN components.

In general, the major novelties and contributions of this paper are summarized in three aspects. First, in comparison with previous works of sparse representation of fMRI signals (Li et al., 2009; Lee et al., 2011; Li et al., 2012; Oikonomou et al., 2012; Lee et al., 2013; Abolghasemi et al., 2013; Lv et al., 2013), our methodology systematically considers the whole-brain task-based fMRI signals with each subject, and aims to infer a comprehensive collection of functional networks. In other words, we employ a big-data strategy (Manyika et al., 2011) that include a large number of fMRI signals to uncover multiple functioning brain networks concurrently. Importantly, each fMRI signal is sparsely represented by a

linear combination of those functioning network components’ signals, which offers a novel, alternative window to examine the spatial compositions of meaningful functional brain networks. Second, we have developed an effective computational pipeline to quantitatively characterize those uncovered functional networks in spatial, temporal and frequency domains, which can be potentially used as functional network atlases for specific task performance or functional scenario in the future. This computational pipeline and its results will not only demonstrate the effectiveness of sparse representation of whole-brain fMRI signals and its neuroscience meaning, but also offer a novel approach to identifying and describing functions of the brain. Third, our methodology provides a novel, effective and unified framework for multiple tasks in traditional fMRI data analysis including activation detection, de-activation detection, and functional network identification. Essentially, the data-driven discovered functional network components via online dictionary learning algorithms correspond, to some extent, to different determining factors that have generated the fMRI BOLD signals. Although this paper focuses on the characterization and interpretation of activation, de-activation and default mode network components, quantitative characterization of many other network components in the dataset used in this paper and in other additional task-based fMRI datasets will likely contribute to deeper understanding of the brain’s structure and function in the future.

2. MATERIALS AND METHODS

2.1 Overview

Fig. 2 summarizes the computational pipeline of identifying functional network components via sparse representation of whole-brain fMRI signals. First, the whole-brain fMRI signals are sparsely represented by using online dictionary learning and sparse coding methods, as illustrated by the 400 learned atomic dictionary components in Fig. 2a. That is, dozens of thousands of whole-brain fMRI signals can all be effectively and sparsely represented by linear combinations of these atomic dictionary components. Second, we propose a novel framework for temporal-frequency characteristics analysis of network

components to identify and select network components of interest (COI) within the learned dictionary. For instance, these COIs could be either correlated or anti-correlated with the task paradigm, and exhibit similar frequency domain patterns as the time series of task paradigm. Figs. 2b-2d show the temporal time series shapes and spatial distribution patterns of three selected COIs that correspond to task (Faraco et al., 2011), anti-task (Archer et al., 2003; Tomasi et al., 2006) and DMN (Raichle and Snyder, 2007) network components, respectively, and their dictionary component indices are highlighted by the color circles in Fig. 2a. As mentioned in Section 1, this paper focuses on exploring these atomic COIs (considered as functional networks here) (Section 2.3), characterizing the spatial, temporal and frequency characteristics of these COIs (Section 2.4), and examining how these COIs contribute to the compositions of all of the fMRI signals within a whole brain (Section 3).

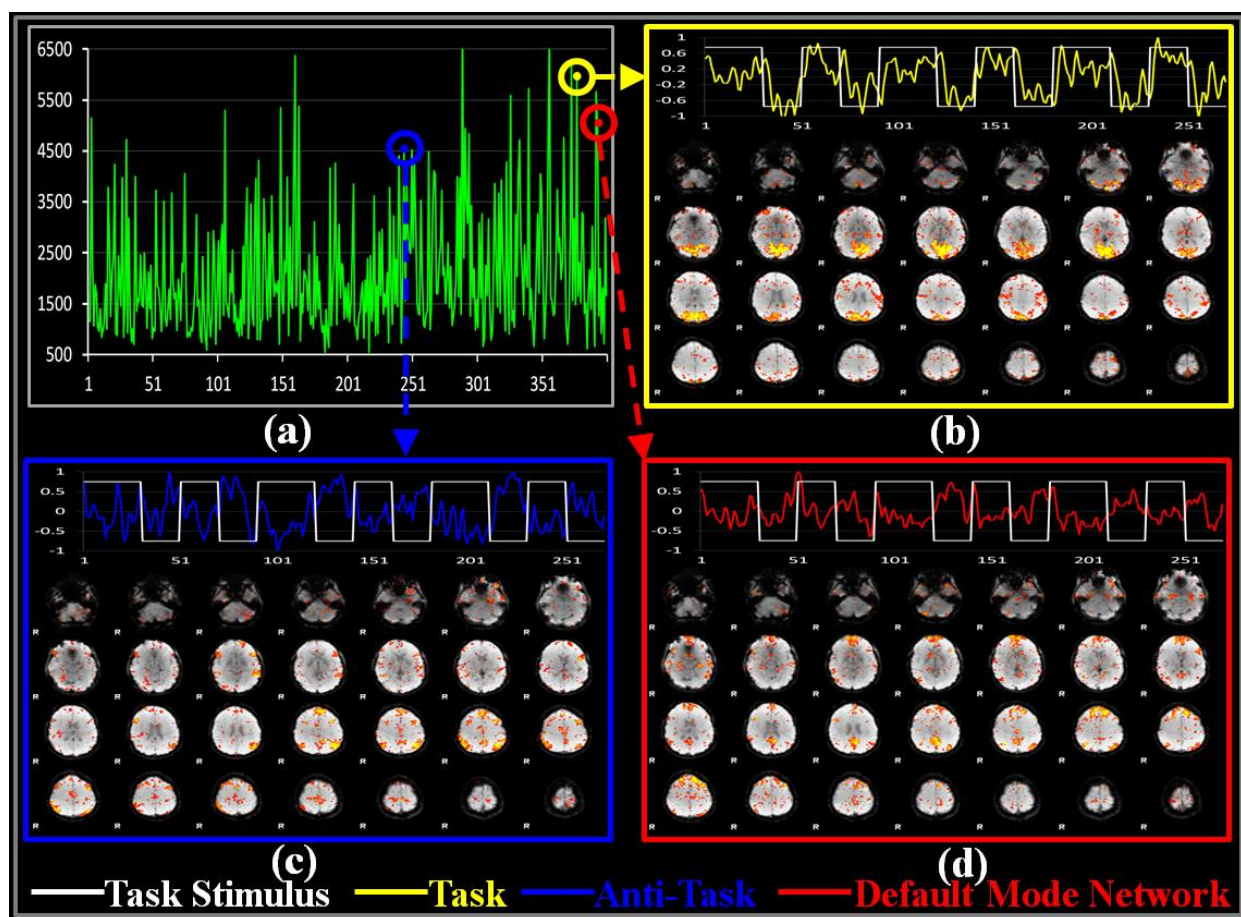


Figure 2. Overview of the computational pipeline of identifying functional brain networks via sparse representation of whole-brain fMRI signals. (a) An example of the learned sparse dictionary of 400 functional components (indexed by the horizontal axis). The vertical axis stands for the occurrence frequency of each component in over 40,000 fMRI BOLD signals in a whole brain. The three dictionary components highlighted by yellow, blue and red circles correspond to different functional networks. They are: (b) task related component in which the response well follows the external block-based task paradigm, (c) anti-task related component in which the response well follows the inverse of external block-based task paradigm, and (d) DMN component. In each component (b-d), the corresponding signals (colored curves) accompanied with the task stimulus (white curve) are shown in the top panels. Their spatial distributions are also back-projected onto the volumetric images in the lower panel. Each voxel is color-coded by the reference weight used in the sparse representation.

2.2 Dataset and Preprocessing

Two different task-based fMRI datasets (block design) and one event-related fMRI data were used in this paper. The first dataset was used as the test bed data to develop and evaluate our sparse representation approaches in Sections 2 and 3. The second dataset was used in Section 3.5 for an independent reproducibility study. For extensive evaluation, the third event-related fMRI data was employed.

Dataset 1: In a working memory task-based fMRI experiment under IRB approval (Faraco et al., 2011; Zhu et al., 2012), fMRI images of 15 subjects were scanned on a 3T GE Signa scanner at the Bioimaging Research Center (BIRC) of The University of Georgia (UGA). Briefly, acquisition parameters are as follows: 64×64 matrix, 4mm slice thickness, 220mm FOV, 30 slices, TR=1.5s, TE=25ms, ASSET=2. Each participant performed a modified version of the operational span (OSPAN) task (3 block types: OSPAN, Arithmetic, and Baseline) (Faraco et al., 2011) while fMRI data was acquired. Preprocessing steps for the fMRI data are referred to Faraco et al., 2011 and Zhu et al., 2012.

Dataset 2: In the semantic decision making task (Zhu et al., 2013), the fMRI scan included 8 on (task) blocks (30 seconds) and 8 off (rest) blocks (15 seconds). During each on-block, ten participants were serially presented with ten pictures (each for 3 seconds), and they made an animacy decision regarding the image (i.e., living/nonliving). Button responses and response times were recorded using a magnetically shielded four-button box in the participant's hand. The task-baseline contrast was used to generate the semantic decision making activation map. FMRI scans were acquired on the 3T GE Signa scanner at UGA BIRC using a T2*-weighted single shot echo planar imaging (EPI) sequence aligned to the AC-PC line, with TE = 25 ms, TR = 1500 ms, 90° RF pulse, 30 interleaved slices, acquisition matrix = 64x64, spacing = 0 mm, slice thickness = 4 mm, FOV = 240 x 240 mm, and ASSET factor = 2. Preprocessing steps of the fMRI data are referred to Zhu et al., 2013.

Dataset 3: Twenty-six right-handed adults (mean age: 28.1±8.5 years) participated in the flanker event-related task fMRI study in New York University (NYU). During the fMRI scan, participants were requested to response to a series of slow-paced Eriksen flanker trials (inter-trial interval (ITI) varied from 8s to 14s, 12s on average). In each trial, the direction the central arrow of five (e.g. < < > > >) was responded by pushing buttons. FMRI images were acquired on a research-dedicated Siemens Allegra 3.0 T scanner in NYU Center for Brain Imaging. The acquisition parameters are as follow: TR=2000 ms; TE=30 ms; flip angle=80, 40 slices, matrix=64×64; FOV=192 mm; acquisition voxel size=3×3×4 mm. Preprocessing includes slice timing correction, motion correction, and spatial smoothing. More details about task design, data acquisition and preprocessing of this open fMRI data are referred to Kelly et al., 2008, Mennes et al., 2010 and Mennes et al., 2011.

2.3 Sparse Representation of Whole-Brain FMRI Signals

Our computational framework of sparse representation of whole-brain fMRI signals is summarized in Fig. 3. Specifically, first, for each single subject's brain, we extract task-based fMRI signals on all voxels within the whole brain. Then, after normalization to zero mean and standard deviation of 1, the fMRI signals are arranged into a big signal data matrix $S \in \mathbb{R}^{t \times n}$ (Fig. 3a), where n columns are fMRI signals from n voxels and t is the fMRI volume number (or time points). By using a publicly available effective online dictionary learning and sparse coding method (Mairal et al., 2010), each fMRI signal vector in S is modeled as a linear combination of atoms of a learned basis dictionary D (Figs. 3b-3c), i.e., $s_i = D \times \alpha_i$ and $S = D \times \alpha$, where α is the coefficient weight matrix for sparse representation and each column α_i is the corresponding reference weight vector for s_i . Finally, we identify components of interests (COIs), namely functional network components in this work, by performing temporal and frequency analysis of atomic signal components (Fig. 3b) in the learned dictionary D . At the same time, we map each row in the α matrix back to the brain volumes and examine their spatial distribution patterns, through which functional network components are characterized and modeled on brain volumes, as shown by the red and yellow areas in Fig. 3c. At the conceptual level, the sparse representation framework in Fig. 3 can effectively achieve both compact high-fidelity representation of the whole-brain fMRI signals (Fig. 3b) and effective extraction of meaningful patterns (Fig. 3c) (Donoho 2006; Huang and Aviyente, 2006; Wright et al., 2008; Wright et al., 2010; Mairal et al., 2010; Yang et al., 2011). In comparison with previous works of sparse representation of fMRI signals (e.g., Li et al., 2009; Lee et al., 2011; Li et al., 2012; Oikonomou et al., 2012; Lee et al., 2013; Abolghasemi et al., 2013), the major novelty here is that our framework holistically considers the whole-brain task-based fMRI signals by using a big-data strategy (Manyika et al., 2011) and aims to infer a comprehensive collection of functional networks concurrently, based on which their spatial, temporal and frequency characteristics are further quantitatively described and modeled.

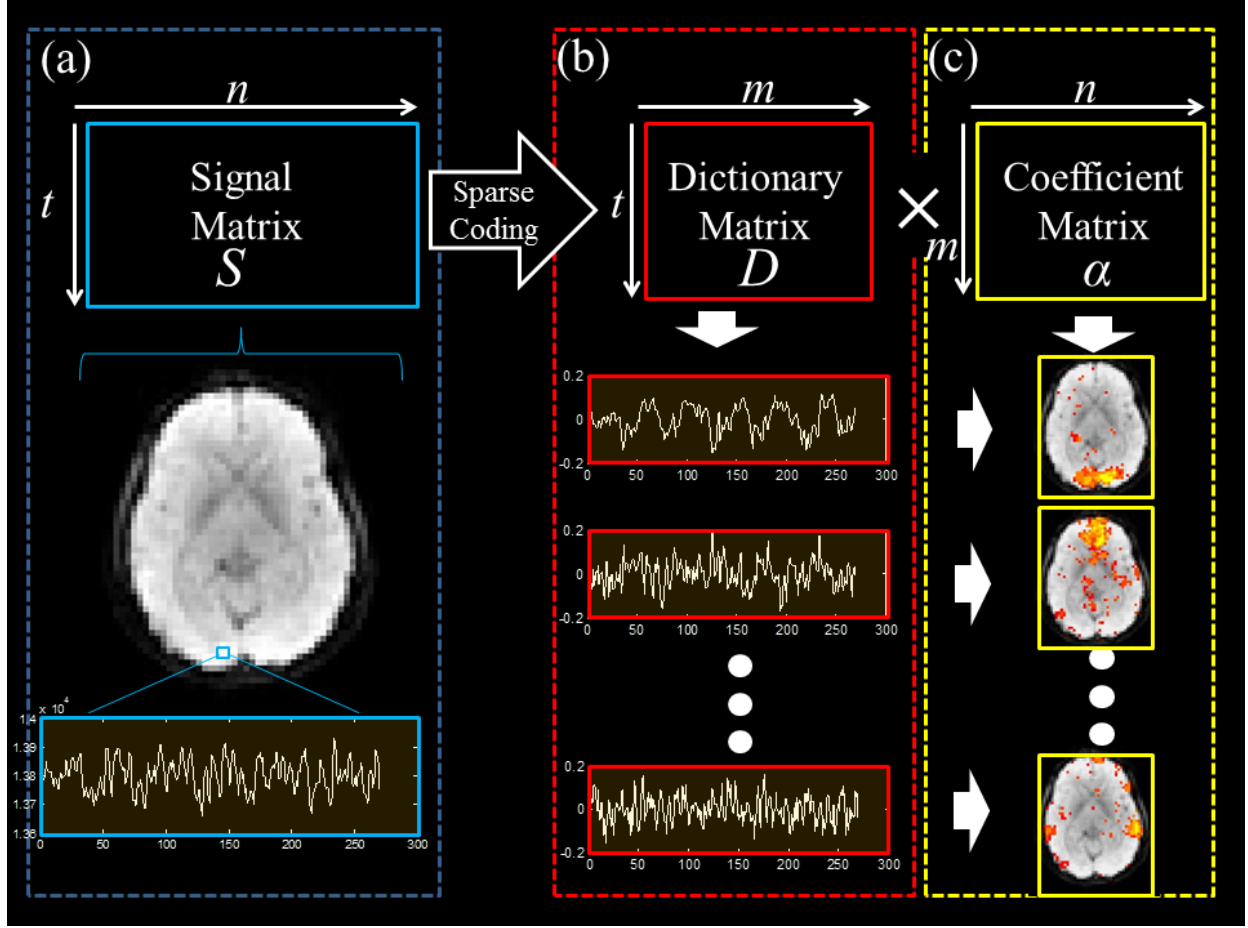


Figure 3. The computational pipeline of sparse representation of whole-brain fMRI signals using an online dictionary learning approach. (a) The whole-brain fMRI signals are aggregated into a big data matrix, in which each row represents the whole-brain fMRI BOLD data in one time point and each column stands for the time series of one single voxel. (b) Illustration of the learned atomic dictionary, each of which represents one functional network component. Three exemplar components of time series are shown in the bottom panels. (c) The decomposed reference weight matrices, each row of which measures the weight parameter of each component in the whole brain. That is, each row defines the contribution of one component to the composition of the fMRI signals.

In our framework, we aim to learn a meaningful and over-complete dictionary $D \in \mathbb{R}^{t \times m}$ ($m > t$, $m \ll n$) (Mairal et al., 2010) for the sparse representation of S . For the task-based fMRI signal set $S =$

$[s_1, s_2, \dots, s_n] \in \mathbb{R}^{t \times n}$, the empirical cost function is summarized in Eq. (1) by considering the average loss of regression of n signals.

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

With the aim of sparse representation using D , the loss function is defined in Eq. (2) with a ℓ_1 regularization that yields to a sparse resolution of α_i , and here λ is a regularization parameter to trade-off the regression residual and sparsity level.

$$\ell(s_i, D) \triangleq \min_{\alpha_i \in \mathbb{R}^m} \frac{1}{2} \|s_i - D\alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \quad (2)$$

As we mainly focus on the fluctuation shapes of basis fMRI BOLD activities and aim to prevent D from arbitrarily large values, the columns d_1, d_2, \dots, d_m are constrained by Eq. (3).

$$C \triangleq \{D \in \mathbb{R}^{t \times m} \text{ s.t. } \forall j = 1, \dots, m, \quad d_j^T d_j \leq 1\} \quad (3)$$

$$\min_{D \in C, \alpha \in \mathbb{R}^{m \times n}} \frac{1}{2} \|S - D\alpha\|_F^2 + \lambda \|\alpha\|_{1,1} \quad (4)$$

In brief, the whole problem of dictionary learning can be rewritten as a matrix factorization problem in Eq. (4) (Lee et al., 2007), and we use the effective online dictionary learning methods in (Mairal et al., 2010) to derive the atomic basis dictionary for sparse representation of whole-brain fMRI signals. Here, we employ the same assumption as previous studies (Li et al., 2009; Lee et al., 2011; Li et al., 2012; Oikonomou et al., 2012; Lee et al., 2013; Abolghasemi et al., 2013) that the components of each voxel's fMRI signal are sparse and the neural integration of those components is linear.

One common use of sparse representation of signals with limited quantity of atoms from a learned dictionary is to de-noise. For our fMRI data analysis application, with the sparse representation, the most relevant basis components of fMRI activities will be selected and linearly combined to represent the original fMRI signals. With the same regularization in Eq. (4), we perform sparse coding of the signal matrix using the fixed dictionary matrix D in order to learn an optimized α matrix for sparse representation

as shown in Eq. (5).

$$\min_{\alpha_t \in \mathbb{R}^m} \frac{1}{2} \|s_t - D\alpha_t\|_2^2 + \lambda \|\alpha_t\|_1 \quad (5)$$

Eventually, the fMRI signal matrix from a subject's whole brain will be represented by a learned dictionary matrix and a sparse coefficient matrix (Fig. 3). Here, each column of the α matrix contains the sparse weights when interpreting each fMRI signal with the atomic basis signals in the dictionary. Meanwhile, each row of the α matrix stores the information of the voxel spatial distributions that have references to certain dictionary atoms. Note that in order to learn task-related and anti-task networks into separate networks and avoid anti-task networks from merging into task-related networks as negative coefficients, we constrained the α matrix positive in both dictionary learning and sparse representation. With these decomposed dictionary components and their reference weight parameters across the whole brain for each subject, our next major task is to characterize and interpret them within a neuroscience context. In particular, the sparse representation and dictionary learning of whole-brain fMRI signals (Fig. 3) are performed for each individual brain separately and thus the spatial, temporal and frequency correspondences of those characterized dictionary components, or components of interests (COIs), across a group of subjects will be another major issue to investigate, as detailed in the next section.

In our approach, the parameter λ not only regularizes the feature selection when reconstructing fMRI signals, but also determines the sparsity and scale of network regions. In other word, if the λ is too small, the network will be too coarse and involve much noise, while if λ is too large, the network will be too sparse. Currently, there is no golden criterion for selection of λ . In our results, the parameter λ was experimentally determined to ensure that the reconstructed networks exhibit meaningful level of sparsity in terms of spatial distributions.

2.4 Temporal-Frequency Analysis of Network Components

In section 2.3, we have obtained the network components by learning a dictionary from the whole-brain fMRI signals for each subject. As each network component has its own time series signal that serves as the basis for sparsely representing the whole-brain fMRI signals, a natural question arises: what are the neuroscience meanings of those hundreds of network components (Figs. 3b-3c)? That is, we need to characterize the structural and functional profiles of those atomic component signals to elucidate the neuroscience meanings of these network components, and potentially establish their correspondences across a group of subjects' brains. It is clear that full understanding and quantitative characterization of all of such hundreds of dictionary network components are beyond our current scope and capability, thus in this paper, our research focus is on the several network components within the learned dictionary that are either correlated or anti-correlated with the task paradigm and exhibit similar frequency domain patterns as the frequency of task performance paradigm. Accordingly, we designed a temporal-frequency analysis framework to identify and select such basic components with more easily interpretable meanings, as shown in the pipeline in Fig. 4a.

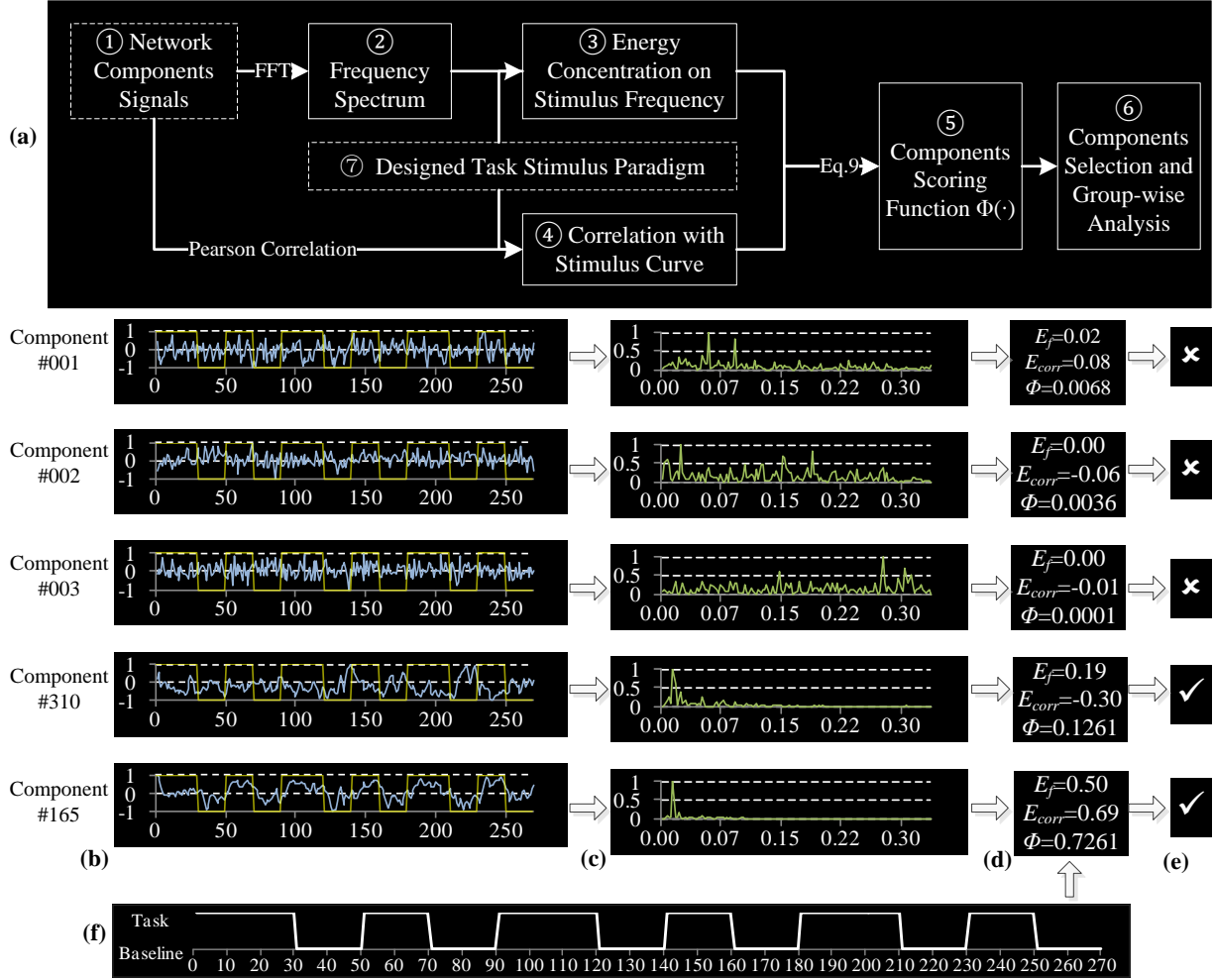


Figure 4. (a) A computational pipeline of the temporal-frequency analysis of network components, which is composed of seven steps. In this framework, the input is the learned dictionary components (D in Fig. 3b) and the output is the selected well-characterized components with their group-wise correspondence. More details of the seven steps are explained as follows. (b) Examples of time series signals of five exemplar network components that are visualized as blue curves, which correspond to the step 1 in the pipeline in (a). The task stimulus curve (yellow, the same as (f)) is overlaid on the component signal for visualization purpose. The x-axis (horizontal) is the temporal points (in volumes), and the y-axis (vertical) is the fMRI BOLD signal normalized to (-1, 1) for visualization. (c) The frequency spectrum of the five network components visualized as green curves, which correspond to the step 2 in (a). The x-axis is the frequency, and the y-axis is the corresponding power normalized to (0, 1). (d) The values of energy

concentration E_f , corresponding to the step 3 in (a); Correlation E_{corr} corresponds to the step 4 in (a); The component score Φ corresponds to the step 5 in (a) of each component. (e) Component selection result, where "✓" means the component is selected as COI by our algorithmic pipeline for further analysis in the next step, which correspond to the step 6 in (a). (f) The stimulus curve of the task paradigm of dataset 1, corresponding to the step 7 in (a). The x-axis is the temporal points (in volumes), and the y-axis is the alternation between task and base-line blocks.

In the diagram in Fig. 4a, the "Network Components Signals" D is the $t \times m$ matrix from the last section as the model input, where m is the number of learned dictionary atoms (network components) and t is the length of the fMRI time series signal. Thus, the signal of the j -th network component is D_j . Another model input is the "Task Stimulus Paradigm" curve TS , which is a vector of length t based on the block-based task design (Faraco et al., 2011; Zhu et al., 2012), as shown in Fig. 4f. For instance, for the working memory task, it can be calculated from the curve (Fig. 4f) that the frequency of a cycle between the task and the baseline is:

$$\frac{1}{\text{average length of task} + \text{average length of rest}} * \frac{1}{TR} = \frac{1}{(20+30)/2+20} * \frac{1}{1.5} = 0.0148Hz \quad (6)$$

which is defined as the stimulus frequency $Fr_{stimulus}$. For other task paradigm (e.g., that in Section 3.5), the stimulus frequency can be calculated in a similar fashion, which is $1/(\text{length of the full paradigm cycle})$. Then, for the j -th network component signal D_j , we can obtain its frequency spectrum FD_j by using the fast Fourier transform on its signal, and calculate the energy concentration $E_{f,j}$ of the stimulus curve frequency over all frequency ranges:

$$E_{f,j} = FD_{Fr_{stimulus,j}} / \sum_i FD_{i,j} \quad (7)$$

where $FD_{Fr_{stimulus,j}}$ denotes the energy of the stimulus frequency in the spectrum of the j -th network component, and $FD_{i,j}$ denotes the energy of the i -th position in the spectrum of the j -th network component. Intuitively, a larger $E_{f,j}$ suggests that this network component is more likely to be responsive (either positively or negatively) to the task stimulus and should be considered as the task related or anti-

task related network. Also, we can obtain the Pearson correlation between the signal of each network component (Fig. 4b) with the stimulus curve (Fig. 4f), which is defined as $E_{corr,j}$:

$$E_{corr,j} = corr(D_j, TS) \quad (8)$$

Essentially, $E_{corr,j}$ measures the temporal similarity between the component's time series and the stimulus curve which is convolved with hemodynamic response function (HRF). A larger value of $E_{corr,j}$ indicates better correspondence between the component and the stimulus. Notably, the widely used GLM model (Friston et al., 1994; Worsley et al., 1997) in the fMRI community uses a similar principle in detecting activated brain regions during a task. Also, the sign of $E_{corr,j}$ can tell whether the network component is positively or negatively correlated with the stimulus curve, which will be used to differentiate task related or anti-task related network components later.

As mentioned in Section 2.1, at the current stage, our work focuses on the network components that are either correlated or anti-correlated with the task paradigm. Therefore, we designed a straightforward, yet effective approach to selecting the components of interests based on both E_f and E_{corr} , and a component scoring function $\Phi(\cdot)$ of the j -th network component is then defined as:

$$\begin{aligned} \Phi(D_j)^+ &= E_{f,j}^2 + E_{corr,j}^2, \text{ if } E_{corr,j} > 0 \\ \Phi(D_j)^- &= E_{f,j}^2 + E_{corr,j}^2, \text{ if } E_{corr,j} < 0 \end{aligned} \quad (9)$$

Here, both $E_{f,j}$ and $E_{corr,j}$ are within the range of (0, 1) and a larger value of $\Phi(\cdot)^+$ or $\Phi(\cdot)^-$ is desired to select the COIs. It should be noted that we defined the scoring function separately for correlated and anti-correlated network components, and thus each component of the learned dictionary will be either in the set $\Phi(\cdot)^+$ or in the set $\Phi(\cdot)^-$. As the positively correlated components were found to have higher scores than anti-correlated components, defining them separately will enable us to select both types of components in a more flexible and reliable manner. A sample illustration of the distributions of components scores in two subjects is shown in Fig. 5.

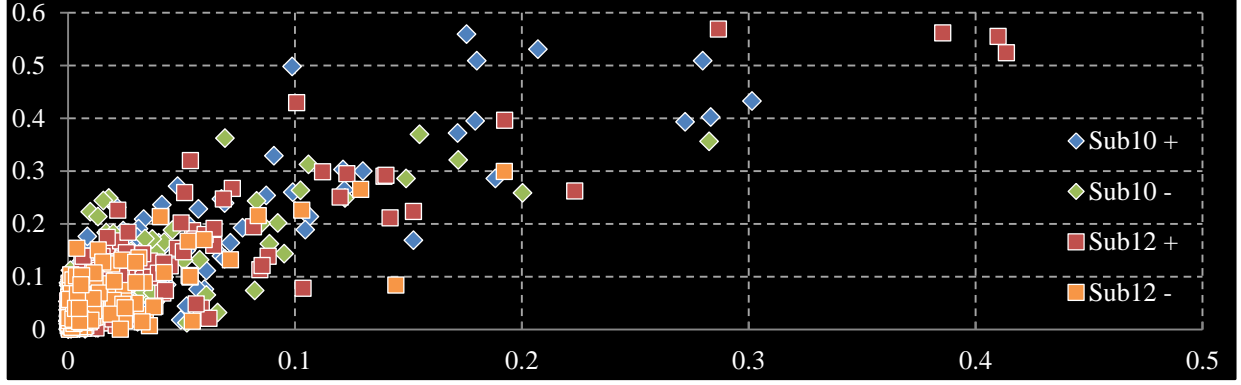


Figure 5. Distribution of $E_{f,j}$ (on the horizontal x-axis) and absolute value of $E_{corr,j}$ (on the vertical y-axis) of the task-related and anti-task components from two randomly selected subjects (subject #10 and #12). “Sub10+” indicates the components from subject #10 that are positively-correlated with the stimulus curve, while “Sub10-” indicates the components from subject 10 that are negatively-correlated with the stimulus curve. We examined these distributions in all of the 15 subjects and observed similar patterns.

In Fig. 5, each icon is a network component, and the components residing in the top-right region (with both large E_f and E_{corr}) are what we aim to select, since we are currently interested in those most responsive components to the stimulus curve. However, as shown in Fig. 5, the distribution of the scores across different types of components and across different subjects is highly variable. Thus, it is more reasonable to individually and adaptively select the best components from each type in each individual subject. Thus, in this work, we designed and applied a greedy iterative searching algorithm to best partition the whole components space into the “selected” and “unselected” groups. For each type (task related/anti-task related) of the components in each subject, we define the “selected” group starting from the component with the highest score $\Phi(\cdot)$, e.g., the top right ones in Fig. 5. We then iterate through all components which are sorted by their scores, and at each step k , we add the new components into the “selected” group, thus forming two partitions $[1..k]$ and $[k+1..m]$ of the total network components. During the greedy iterative searching, as long as the following criterion is decreasing, the iteration will be continued:

$$C_{([1\dots k],[k+1\dots m])} = \frac{1}{k} \sum_{j=1}^k (E_{f,j} - \overline{E_{f,[1\dots k]}})^2 + \frac{1}{m-k} \sum_{j=k+1}^m (E_{f,j} - \overline{E_{f,[k+1\dots m]}})^2 - (\overline{E_{f,[1\dots k]}} - \overline{E_{f,[k+1\dots m]}})^2 \quad (10)$$

In other words, we aim to select the most suitable network components by minimizing the intra-group distance while maximizing the inter-group distance, where the groups are defined by partitioning the sorted components at k -th index.

2.5 Spatial Pattern Analysis of Network Components

The frequency and temporal characteristics of the task related and anti-task related network components in the learned dictionary can be quantitatively described by Eqs. (6)-(9). In addition, the reference weight parameter in each row of the matrix in Fig. 3c for each network component can be projected back to the volumetric fMRI image space (e.g., Fig. 3c) for the interpretation of their spatial distributions. In this way, the spatial distributions of network components in different brains can be compared within a template image space to verify their spatial overlaps, as well as to further determine their spatial correspondences (more details in Section 3.2).

In addition to the task related and anti-task related network components that are characterized in the above Section 2.4, it is interesting that there are also a variety of intrinsic networks (e.g., Fox and Raichle, 2007; Cohen et al., 2008; van den Heuvel et al., 2008) that are identifiable in task-based fMRI data. For instance, there is a network component that clearly corresponds to the DMN (Raichle and Snyder, 2007), as shown in Fig. 2d. Since the temporal and frequency characteristics of the DMN have not been well quantitatively described, we more rely on the spatial distribution patterns of the peak activities of DMN on a template brain space (Fox and Raichle, 2007; Cohen et al., 2008; van den Heuvel et al., 2008), as shown in Supplemental Figure 1. We then use a spatial overlap metric to determine the corresponding DMN components across individual brains.

3. RESULTS

In this section, we designed a series of experiments to evaluate and validate the novel computational pipeline for identification of functional networks via sparse representation of whole-brain fMRI signals. First, the temporal and frequency properties of selected task related and anti-task related COIs from 15 subjects in the dataset 1 are presented in Section 3.1. Afterwards, the spatial distribution patterns of these COIs are detailed and interpreted in Section 3.2. Then the framework is extensively evaluated and validated by comparisons with the ICA method (Section 3.3), by simulation studies with ground-truth (Section 3.4), and by an independent reproducibility studies in a separate dataset 2 (Section 3.5). An additional application of our method on event-related fMRI data is explored in Section 3.6

3.1 Temporal and Frequency Properties of COIs from 15 Subjects

Based on the methods and criteria in Section 2.4, we have obtained 29 task related and 25 anti-task related network components from the learned dictionaries of all the 15 subjects in dataset 1. On average, two network components of each type (task related or anti-task related) were selected for each subject, which correspond to the best-matched functional response to the task stimulus in terms of frequency spectrum and temporal correlation (Eqs. (7)-(10)). The time series component signals, the frequency spectra and the scores of the selected COIs of five randomly-chosen subjects are listed in Figs. 6-7. The results of other ten subjects are shown in Supplemental Figs. 2-3. Quantitatively, the average correlation of the signals of task related components with the stimulus curve (Eq. (8)) over all 15 subjects is 0.585 (with the standard deviation of 0.115), and their average energy concentration on the frequency spectra (Eq. (7)) is 40.9% (with standard deviation of 7%). The relatively high correlations and energy concentrations suggest that these selected COIs are well responsive to the stimulus curve, which is also evident in the second columns of Fig. 6 and Supplemental Fig. 2. It is thus natural to conjecture that these COIs correspond to the

functional networks that are responsive to the working memory task and are potentially equivalent to the traditional activated brain regions detected by the GLM method, which will be verified in Section 3.2.

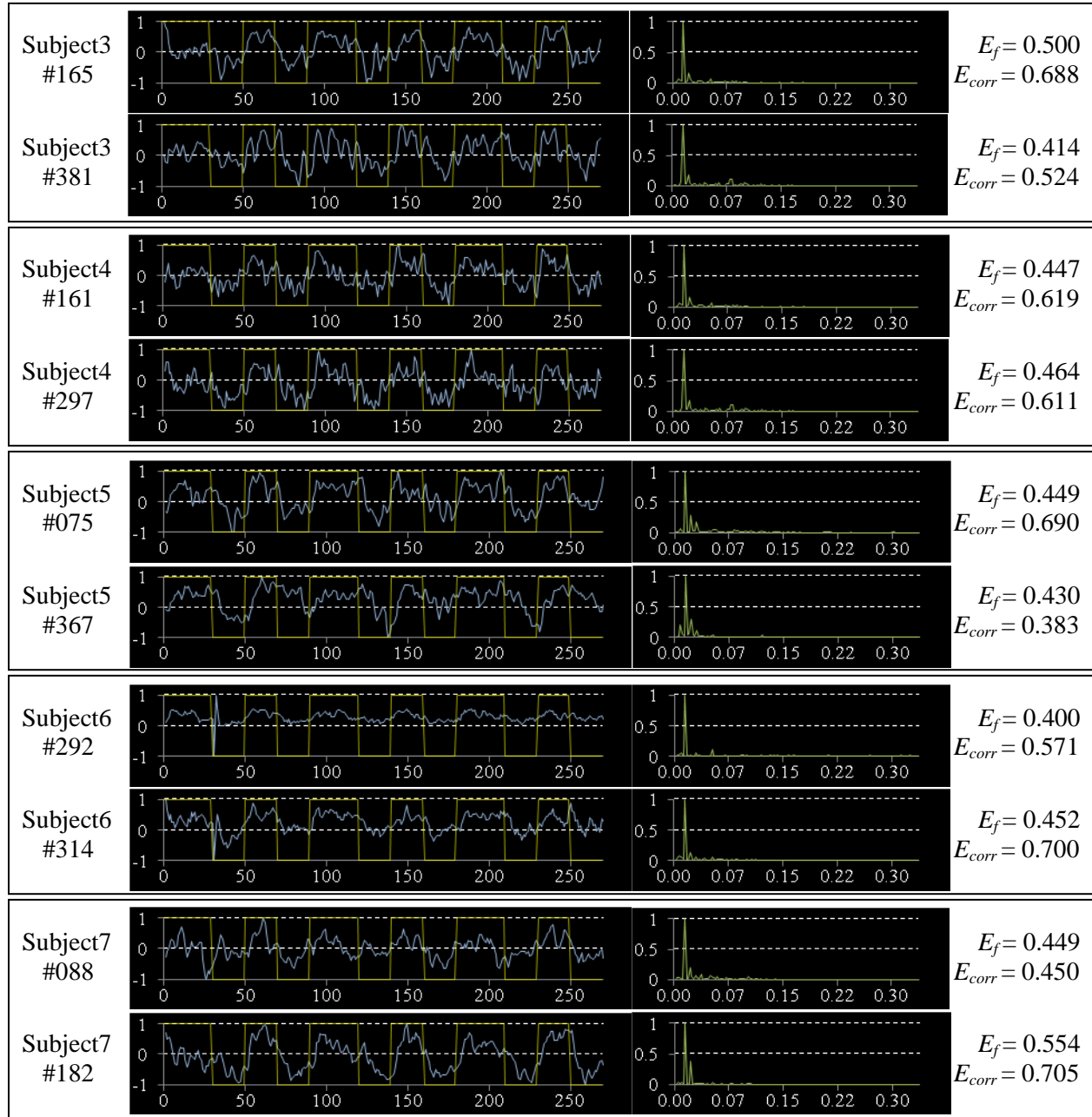


Figure 6. The selected task related network components from five randomly-chosen subjects with a total of 10 components. For each row in the figure, from the left to the right are: subject index and component index, time series signal of that component with overlaid stimulus curve (in yellow), the frequency

spectrum of that component, and the value of component scores, respectively. It is evident that the COI component time series signals are well correlated with the stimulus curve.

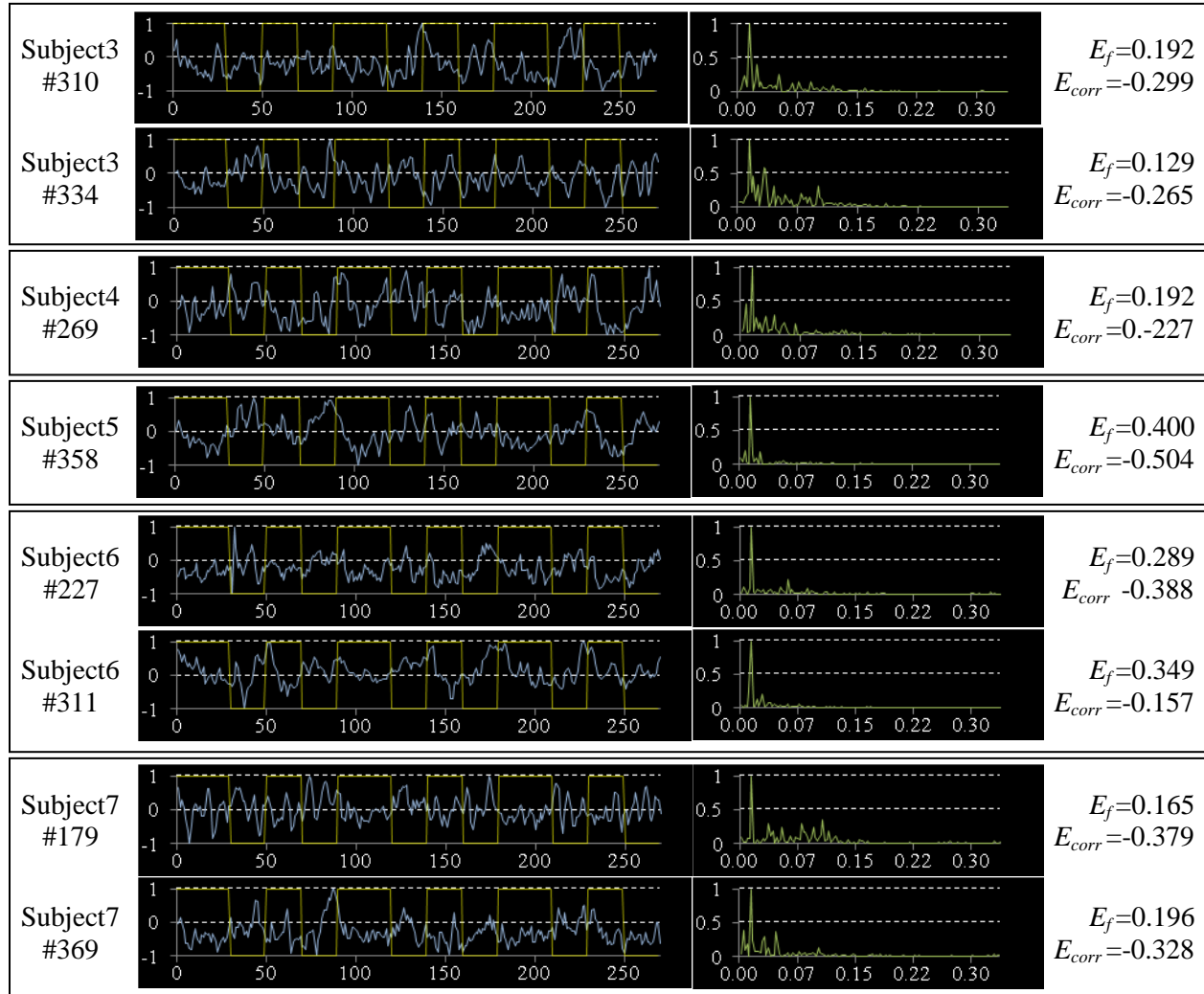


Figure 7. The selected anti-task network components from the same five subjects, with a total of 8 components. For each row in the figure, from the left to the right are: subject index and component index, time series signal of that component with overlaid stimulus curve (in yellow), the frequency spectrum of that component, and the value of component scores, respectively. It is evident that the COI component time series signals are well anti-correlated with the stimulus curve.

Quantitatively, the average correlation of the signal of anti-task component with the stimulus curve (Eq. (8)) over all 15 subjects is -0.348 (with standard deviation of 0.014), and their average energy

concentration on the frequency spectra (Eq. (7)) is 23.1% (with standard deviation of 8%). It can be seen in Fig. 7 and Supplemental Fig. 3 that all the 15 subjects have well-matched anti-task related functional network components, suggesting that our methods can identify common anti-task networks in the response to stimulus paradigm from individual subjects. The relatively high anti-correlations and energy concentrations suggest that these selected COIs are highly anti-responsive to the stimulus curve, which is also evident in the second columns of Fig. 7 and Supplemental Fig. 3. We therefore conjecture that these COIs potentially correspond to the traditional de-activated brain regions detected by the GLM method, which will be evaluated in Section 3.2.

3.2 Spatial Distribution Patterns of COIs

In this section, the identified COIs in Section 3.1 will be further analyzed to elucidate their spatial distributions based on the methods in Section 2.5. Specifically, the 29 task related network components from the learned dictionaries of all the 15 subjects in dataset 1 are mapping to the volumetric images. Specifically, as the learning of coefficient matrix is constrained non-negative and the network region size and scale are controlled by the parameter λ , in our experiment, we simply mapped the coefficients which are “>0” without setting additional threshold. This also applies to the following overlap analysis. As an example, in Figs. 8a-8d, we show two selected task related COIs of subject #1. The results for additional six different subjects are shown in Supplemental Figs. 4-5. In Figs. 8a-8b, the two COIs are color-coded with the reference weights of whole-brain voxels. We can see that each network component is composed of several Gaussian-shaped patterns of reference weights. This distribution pattern is consistent with previous observations of fMRI activation foci patterns (Faraco et al., 2011). From Figs. 8c-8d, we can observe that the signals of the selected networks have high correlation (around 0.6~0.7) with the stimulus curve (Eq. (8)), and its energies in the frequency spectra are dominantly concentrated on the frequency of 0.0148Hz. This result supports our hypothesis in Eq. (6) and demonstrates the effectiveness and accuracy of the data-driven online dictionary learning methods (Mairal et al., 2010) in extracting meaningful basis patterns for sparse representation of whole-brain fMRI signals. Our results also provide additional

supporting evidence to the widely-used GLM methods (Friston et al., 1994; Worsley et al., 1997) that the brain's functional activities could be very responsive to the specific task paradigm, e.g., the exactly matched frequency.

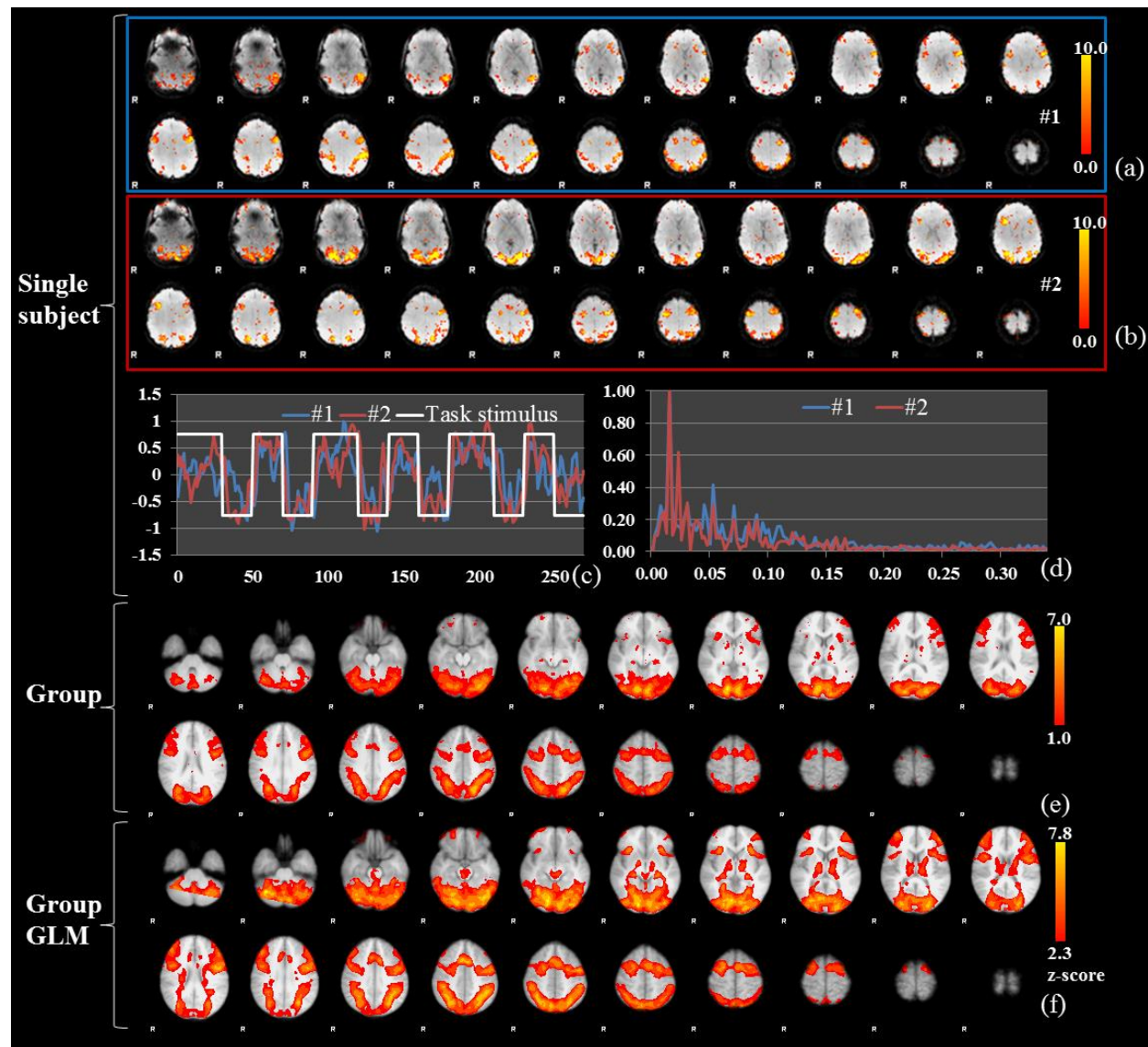


Figure 8. (a)-(b) Two selected task related COIs of subject #1. (c) The corresponding temporal patterns of the two components in (a) and (b). (d) The corresponding frequency distribution of the two components in (a) and (b). (e) The group-wise statistical map of all task related components from 15 subjects of dataset 1 in the MNI space. (f) Group-wise activation foci detected by FSL FEAT.

Furthermore, for each subject, since its task related network components share quite similar temporal and frequency characteristics (Fig. 6), we merged them (the reference weight matrix of α , Fig. 3c) into one volumetric map in order to comprehensively elucidate their spatial distribution patterns. After registering and warping them into the Montreal Neurologic Institute (MNI) template space by the FSL FLIRT, we averaged the complete task related networks from a group of 15 subjects and visualized the averaged statistical atlas in Fig. 8e. For comparison purpose, the group-wise activation map obtained by applying the FSL FEAT on the same working memory task-based fMRI data is also visualized in Fig. 8f. We can see that the spatial distributions of task related network by our methods and those of the activation foci by FSL FEAT are quite similar. Quantitatively, the overlap of color regions in Figs. 8e-8f account for 86.8% of the result by our method (Fig. 8e) and 66.6% of result by FSL FEAT (Fig. 8f). This relatively high overlap demonstrates that the task related functional network detected by our method is quite meaningful and consistent with that by FSL FEAT, suggesting the validity and effectiveness of the dictionary learning and sparse representation methods described in Section 2.3 in uncovering meaningful functional activity patterns from whole-brain fMRI data. Furthermore, the reasonably consistent task-related functional networks in individual brains in Figs. 8a-8b and Supplemental Fig. 4-5, as well as the comparable group-wise activity patterns in Figs. 8e-8f, suggest that our COIs selection methods in Section 2.4 could potentially serve as a novel, alternative approach to detecting task-based fMRI activations. This important issue will be further explored in the Section 3.5.

Similarly, the reference weight matrices (α , Fig. 3c) of 25 anti-task related network components from the learned dictionaries of all the 15 subjects in dataset 1 are mapped and examined on volumetric images. Specifically, in Figs. 9a-9d, we show the two selected anti-task related networks of subject #6. The results of additional six subjects are shown in Supplemental Figs. 6-7. Similar to those in Fig. 8, their spatial distributions are multiple Gaussian-shaped foci. The temporal time series signals of these anti-task components have relatively strong Pearson correlations (-0.4~-0.5) with the block-design stimulus curve,

as shown in Fig. 9c. Also, their energies in the frequency domains are dominantly concentrated on 0.0148Hz, as shown in Fig. 9d. Again, this result further supports our hypothesis in Eq. (6) and demonstrates the validity and reliability of the data-driven online dictionary learning methods (Mairal et al., 2010) in extracting not only task related but also anti-task related basis patterns for sparse representation of whole-brain fMRI signals.

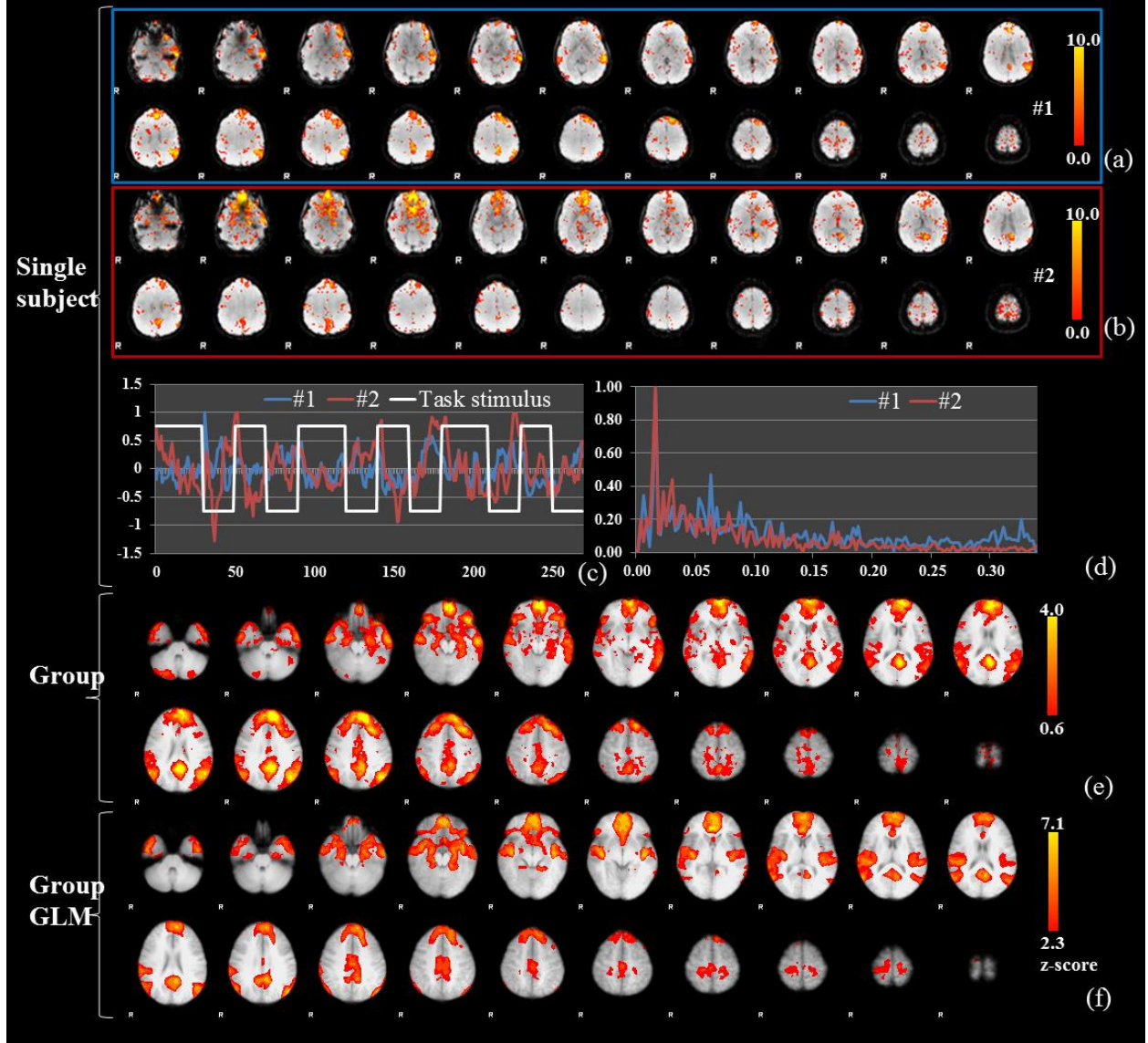


Figure 9. (a-b) Two identified anti-task COIs of subject #6. (c) The corresponding time series patterns of the two components in (a) and (b). (d) The corresponding frequency distribution of the two components in

(a) and (b). (e) The group-wise averaged statistical atlas of all anti-task components from 15 subjects of dataset 1 in the MNI space. (f) Group-wise de-activation foci detected by FSL FEAT.

Additionally, for each subject, given that its anti-task related components exhibit similar temporal and frequency characteristics (Fig. 7), we merged their reference weight matrices (α , Fig. 3c) into one volumetric map in order to better examine their spatial distributions in a similar way as in Fig. 8e. For comparison purpose, the group-wise de-activation map obtained by applying FSL FEAT is visualized in Fig. 9f. It is evident that the spatial distributions of anti-task related network by our methods and those of the de-activation foci by FSL FEAT are similar. Quantitatively, the overlap of color regions in Figs. 9e-9f account for 46.6% of the result by our method (Fig. 9e) and 72.1% of result by FSL FEAT (Fig. 9f). This relatively high overlap suggests that the anti-task related functional network identified by our method is quite meaningful and consistent with that by FSL FEAT, further demonstrating the validity and effectiveness of the dictionary learning and sparse representation methods described in Section 2.3 in uncovering meaningful functional patterns from whole-brain fMRI data. Similarly, the consistent anti-task related functional networks in individual brains in Figs. 9a-9b and Supplemental Fig. 6-7 and the consistent group-wise activity patterns in Figs. 9e-9f indicate that our COIs selection methods in Section 2.4 could potentially serve as a novel, alternative approach to detecting fMRI de-activations, which will be further investigated in the future.

The temporal-frequency analysis framework in Section 2.4 have successfully uncovered the task and anti-task related network components as shown in Figs. 8-9 and Supplemental Figs. 2-7. Then, based on the spatial pattern analysis methods in Section 2.5, we measured the spatial overlaps of the dictionary components with the DMN template in Supplemental Fig. 1. It is interesting that we can successfully identify the DMNs in all of the 15 subjects in dataset 1, as shown in Fig. 10 and Supplemental Fig. 8. The group-wise averaged statistical map of the DMN components by our methods (Fig. 10d) is also visually

and quantitatively (the overlapped area accounts for 42.7% of our result and 56.1% of the template) similar with the template in the MNI space (Supplemental Fig. 1). This result further demonstrates that our methods are effective in uncovering meaningful network components from task-based fMRI data, even though the DMN is a intrinsic network and its temporal and frequency characteristics are much more complex and variable than the task and anti-task components, as shown in Supplemental Fig. 9. Also, an important neuroscience insight obtained from the results here is that intrinsic networks such as the DMN (Fox and Raichle, 2007; Cohen et al., 2008; van den Heuvel et al., 2008) are active in task performance state and are clearly identifiable. This observation and the methods developed in this paper might open a new window to examine the functional interactions among intrinsic networks and task/anti-task related networks in the future.

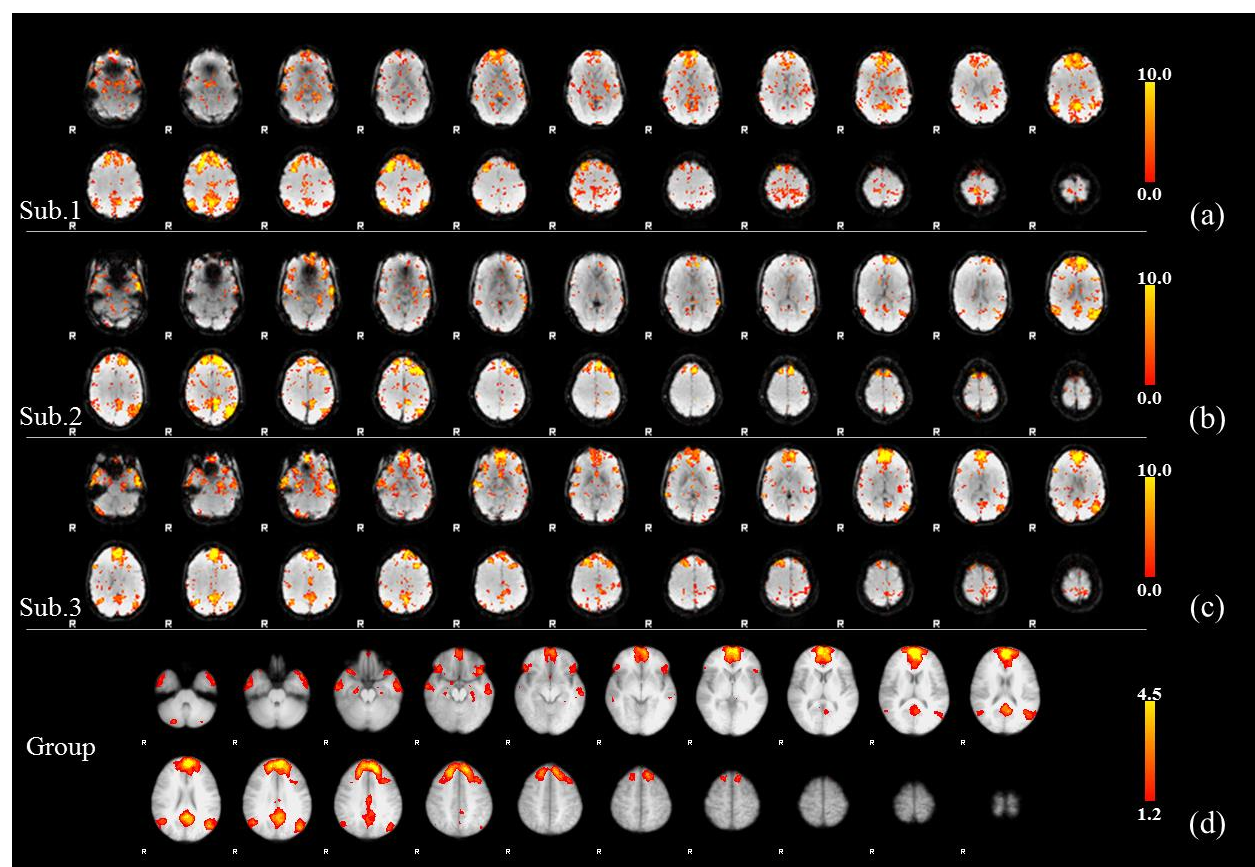


Figure 10. (a-c) Identified default mode network components of subject #1, #2, and #10. Additional examples are shown in Supplemental Fig. 8. (d) The group-wise averaged statistical map of all DMN components from 15 subjects of dataset 1 in the MNI space.

Based on the identified task, anti-task and DMN components in Figs. 8-10 in the 15 subjects, we quantitatively measured the percentages of their volumes and the overlapped regions among these three components, as illustrated in the bottom panels of Supplemental Fig. 10. The percentages for all 15 subjects in dataset 1 are shown in Table 1, and the visualizations of these percentages are shown in Supplemental Fig. 10. From Table 1 and Supplemental Fig. 10, we can clearly see that these three network components are substantially overlapping with each other in the spatial domain, suggesting that functional brain networks do not necessarily work independently, but instead they interact with each other on the overlapped brain areas. These results also demonstrate that one cortical region could potentially participate in multiple functional roles, as widely reported in the literature (e.g., Bisley and Pasternak, 2000; Lalonde et al., 2002; Fogassi et al., 2005; Zaksas et al., 2006; Fischera et al., 2008). It is interesting that the dictionary learning and sparse representation methods can not only uncover and characterize those separate network components, but also reveal how they contribute to the compositions of dozens of thousands of fMRI signals within the whole brain.

Table 1. The overlap percentages of three detected networks from 15 subjects in dataset 1. T: Task network; A: Anti-task network; D: Default mode network. SD stands for standard deviation.

Network Overlap	T(%)	A(%)	D(%)	T&A(%)	A&D(%)	T&D(%)	T&A&D(%)
Sub.1	40.76	62.20	36.39	3.05	36.26	2.04	2.00
Sub.2	46.16	57.07	27.53	3.45	27.25	1.53	1.47
Sub.3	40.43	48.45	27.16	6.66	7.51	2.51	0.64
Sub.4	48.75	25.44	33.99	2.50	3.25	2.52	0.09
Sub.5	67.45	35.92	35.92	3.36	35.92	3.36	3.36
Sub.6	37.12	52.41	22.40	3.53	7.12	1.64	0.36

Sub.7	42.52	41.73	25.93	3.57	4.94	1.97	0.30
Sub.8	43.92	53.33	17.83	9.52	3.84	2.16	0.45
Sub.9	60.71	23.39	24.02	2.05	3.71	2.48	0.13
Sub.10	52.85	52.89	33.75	5.78	33.72	2.21	2.21
Sub.11	52.61	21.97	37.23	2.31	4.66	5.19	0.34
Sub.12	36.61	52.54	26.28	5.01	8.55	2.43	0.56
Sub.13	60.55	13.70	33.99	1.31	3.76	3.28	0.11
Sub.14	44.83	36.81	29.32	1.71	5.83	3.60	0.18
Sub.15	56.25	50.38	36.16	6.67	36.10	4.53	4.51
Average	48.77	41.88	29.86	4.03	14.83	2.76	1.11
SD	9.33	14.88	5.97	2.27	14.15	1.04	1.35

3.3 Comparisons with ICA Method

In this section, we performed independent component analysis (ICA) of whole-brain fMRI signals via the FSL MELODIC toolkit (Beckmann et al. 2005) as an independent source to compare and evaluate the identified functional networks via sparse representation in Section 3.2. Specifically, we set the MELODIC-ICA to automatically estimate the optimal dimensionality of the data to achieve convergence stability. First, we identified and examined the DMN via the methods in Section 2.5 and defined the true positive rate as:

$$R(X, T) = \frac{|X \cap T|}{|T|} \quad (11)$$

where X is the component's spatial map and T is the DMN template (Supplemental Fig. 1). The true positive rate was applied to measure the similarity between the ICA-derived spatial map and the DMN template (Supplemental Fig. 1). For both sparse representation and ICA methods, the spatial map with the highest true positive rate with the DMN template was selected as the DMN, and the results are shown in Fig. 11. The mean true positive rate of identifying DMN in our sparse representation of all 15 subjects is 0.36 (0.30-0.49), while the average true positive rate for ICA method is 0.27 (0.24-0.29). The detailed results for each subject are shown in Table 2. Therefore, both qualitative (Fig. 11) and quantitative (Table 2) results indicate that our sparse representation methods can more consistently and reliably identify

DMN, compared with the commonly-used ICA approach. Also, the sparse representation method can identify a more complete map of DMN than ICA. Our interpretation is as follows. As shown in Table 1, those additional brain regions in the DMN mapped by our sparse representation method are also involved in other network components such as task related and anti-task related components, all of which interact with each other and are not necessarily spatially independent (Daubechies et al., 2009; Lee et al., 2011). Thus, those additional overlapped regions in the DMN are difficult to be identified via the ICA method that assumes spatial independence of the network components. The results in Fig. 11 and Table 2 demonstrated the effectiveness and accuracy of the proposed sparse representation methods in uncovering intrinsic networks (DMN in this work) in task-based fMRI data.

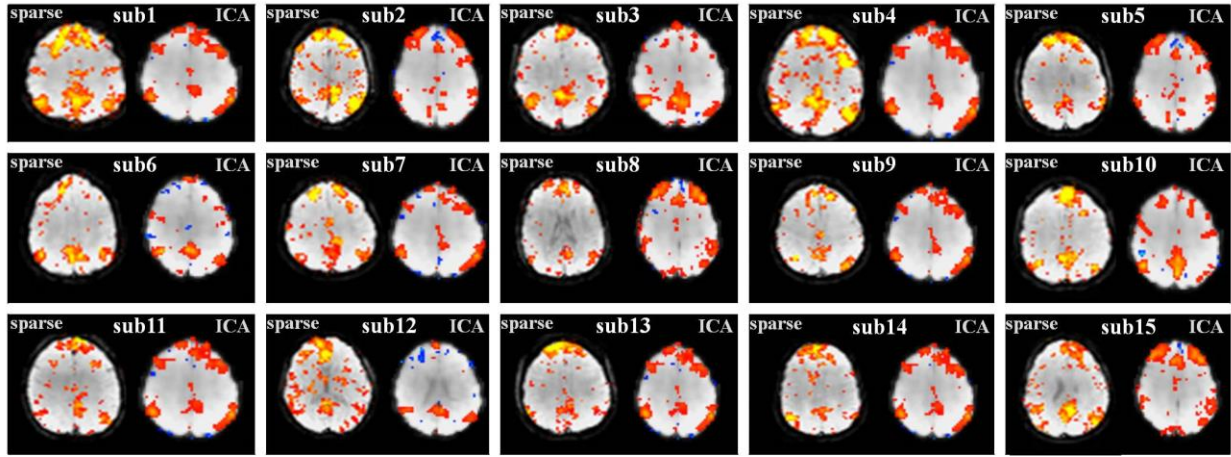


Figure 11. The spatial maps of DMN obtained by the sparse representation and ICA methods for all 15 subjects in dataset 1, respectively. For each subject, the most informative slice which was superimposed on the mean fMRI image of each subject is shown (left: sparse representation; right: ICA). The spatial maps were selected by calculating and sorting the true positive rate with the DMN template provided in GIFT toolbox (<http://mialab.mrn.org/software/gift/index.html>), with a mean rate of 0.36 for sparse representation and a mean rate of 0.27 for ICA of all 15 subjects. All ICA spatial maps were converted to Z-transformed statistic maps using the default threshold value 0.5 (Beckmann et al., 2005). The color scale of spatial maps in sparse representation ranges from 0 to 10.

Table 2. The mean overlap rate of DMN in all 15 subjects of dataset 1 by sparse representation and ICA.

Sub	#1	#2	#3	#4	#5	#6	#7	#8
Sparse	0.49	0.37	0.36	0.41	0.40	0.34	0.31	0.33
ICA	0.27	0.27	0.29	0.27	0.28	0.25	0.27	0.24
Sub	#9	#10	#11	#12	#13	#14	#15	<u>Mean</u>
Sparse	0.34	0.47	0.30	0.35	0.30	0.31	0.37	<u>0.36</u>
ICA	0.26	0.27	0.29	0.25	0.28	0.28	0.28	<u>0.27</u>

3.4 Validation by Simulated Data

In this section, the proposed sparse representation framework is applied on simulated data with ground-truth to examine its reliability, robustness and reproducibility. In our sparse representation framework (Fig. 3), the whole-brain fMRI signals are factorized into multiple network components with corresponding basis time series signals. Thus, we adopt the previously factorized time series basis of components as benchmark, and aggregate them together with a chosen A matrix to generate simulated fMRI signals in the brain. Specifically, we choose the A matrix from the factorization of one model subject as the coefficient map ground truth. The 400 basis signal components will be randomly selected from a total of 6000 trained component signals of 15 subjects and then be used to compose the network components of the simulated subject. Thus, we can generate the simulated whole-brain fMRI signals by:

$$Data^* = \sum_{k=1}^{400} D_{i_k} A_k \quad (12)$$

where A_k are the reference weight matrices of network components from the certain A matrix we chose. D_{i_k} is a randomly picked signal such that $i_k = random(\{1 \dots 6000\})$ and $i_p \neq i_q$ for $p \neq q$. Such simulation was performed 60 times with 4 times on the components of each subject. By using the framework in Fig. 3, 400 dictionary network components are obtained and then compared with the 400 ground-truth components that were used to generate the simulated whole-brain fMRI data. Specifically,

the Jaccard similarity coefficient is used to measure the similarity between the factorized reference weight matrices as below:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (13)$$

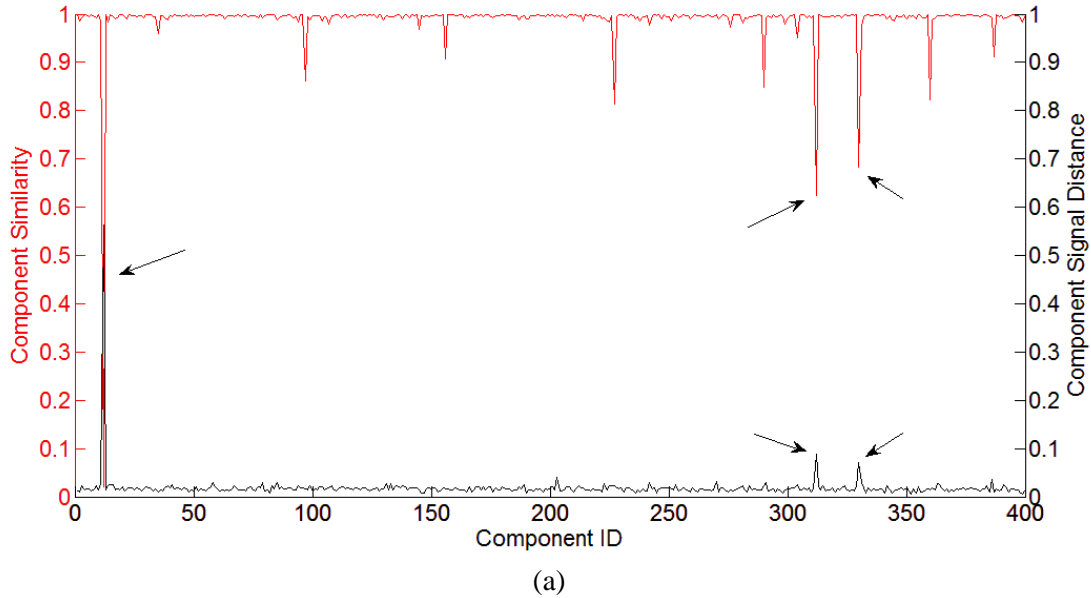
where A, B are two spatial maps. The Sørensen distance (Cha, 2007) is employed to measure the similarity between network component time series signals.

$$Dis(p, q) = \frac{\sum_{i=1}^n |p_i - q_i|}{\sum_{i=1}^n |p_i| + |q_i|} \quad (14)$$

where $p, q \in \mathbb{R}^{1 \times N}$ are two component signal vectors.

Afterwards, each newly obtained network component is compared with the ground-truth components. The pair with the highest Jaccard similarity coefficient or the lowest Sørensen distance is considered as the corresponding components and the similarity/distance values are recorded for further statistical analysis. The comparison result of one simulation is shown in Fig. 12a. More simulation results are provided in Supplemental Figure 11. It is evident that most of the pairs between uncovered components and ground-truth have relatively high similarity (close to 1). Also, the distance between corresponding signals are relatively low (close to 0). Notably, as highlighted by the black arrows in Fig. 12a, components #12, #312, and #330 in this example have lower component similarities and higher signal distances, meaning that the online dictionary learning algorithm (Mairal et al., 2010) might have difficulty in uncovering a very small portion of the network components. For the 60 simulations, we counted the numbers of obtained network components with different component similarities/signal distances in Table 3. By taking those network components with similarity to ground-truth lower than 0.8 or signal distance to ground-truth higher than 0.05 as unsuccessful ones, 99.18% network components are successfully uncovered from the simulated data, which is very high. This result suggests that the online dictionary learning algorithm and the sparse representation framework are reliable and robust in decomposing the aggregated whole-brain fMRI signals into meaningful basis signals and reference weight matrices.

However, the above simulation is based on the ideal assumption. In real data, noise should be taken into consideration. Hence, we added Gaussian random noises to the above simulation to investigate at which level of signal-noise ratio (SNR), the decomposing is effective and stable. In our experiment, the meaningful signals are the ones reconstructed with D and A, and noises are the reconstruction residuals. Bases on the example in Fig. 12a (without noise), we simulated data with different levels of SNRs as shown in Fig. 12b. The comparison with ground-truth is also based on the component similarity and signal distance. As we can see in Fig. 12b, with $\text{SNR} > 10$ DB, the reconstruction is quite akin to the one without noise. While $\text{SNR} < 10$ DB, the reconstruction error becomes gradually dramatic. However, at the noise level $\text{SNR} = 8$ DB, the component similarity is around 0.8 and signal distance is around 0.15, which is still acceptable. In our real data of Dataset 1, based on the settings in this paper, the SNR is around 10 on average. Thus this simulation provides evidence that our settings are reliable in reconstructing stable networks.



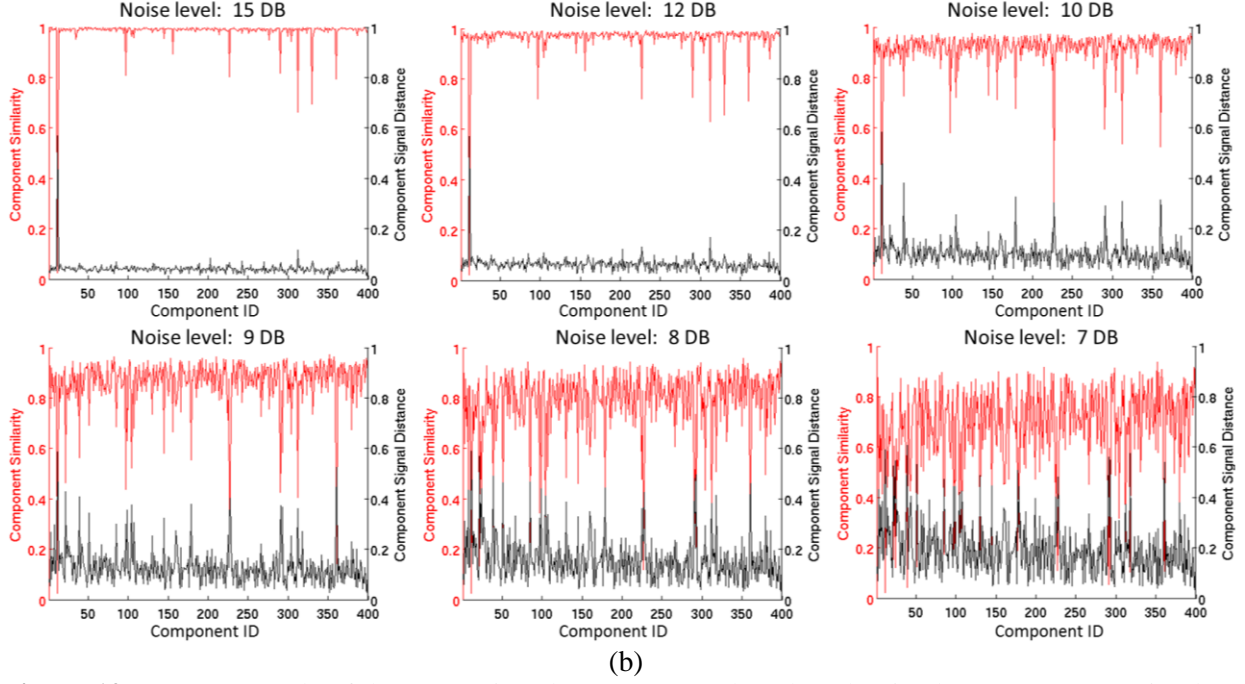


Figure 12. (a) An example of the comparison between ground-truth and trained components on simulated data in terms of component similarity (Left y-axis) and signal distance (Right y-axis). The x-axis indicates component IDs. (b) The same comparison with a series of levels of noise based on the same example in (a). The noise is added with measures of SNR of 15 DB, 12 DB, 10 DB, 9 DB, 8 DB, and 7 DB.

Table 3. The histogram of numbers of network components among 24000 candidates in the simulation.

Component Similarity	1~0.99	0.99~0.95	0.95~0.9	0.9~0.8	0.8~0.6	0.6~0
Component Number	22255	1421	117	88	33	86
Signal Distance	0~0.02	0.02~0.05	0.05~0.1	0.1~0.2	0.2~0.4	0.4~1
Component Number	20791	3037	70	22	23	57

3.5 Reproducibility Study

A key parameter in the sparse representation framework in Section 2.3 is the dictionary size (m in Fig. 3). In this section, we first examine if/how the setting of dictionary size while performing the online dictionary learning would affect the experimental results. As an example, we repeated the methods in

Section 2.3-2.4 on one randomly selected subject in dataset 1 with different dictionary sizes ranging from 300 to 500 with the interval of 10, and the selected dictionary items corresponding to the task related network components are listed in Table 4. We can see that the component #165 was consistently selected among all of these experiments and the component #381 was consistently selected among all the experiments with dictionary size larger than 380. For further verification, in Figs. 13-14, we visualized the temporal, frequency and spatial characteristics of the selected corresponding component #165, while the dictionary size is 300, 350, 400, 450 and 500, respectively. We can see that the selected component #165 is consistent and reproducible across different parameter settings with quite similar temporal, frequency and spatial patterns. Notably, when the dictionary size is lower than 380, our method only selected #165 as the COI. But when the size is higher than 380, our method can consistently detect #381 as COI. This is because the online dictionary learning method considers dictionary components accumulatively (Mairal et al., 2010). Supplemental Fig. 12 visualized the selected anti-task component #310 while the dictionary size is 350~500. Their temporal, frequency and spatial patterns are also quite consistent.

Selection of the dictionary size is still an open question in the machine learning field. Based on our experience, firstly, the dictionary size should be larger than the lowest dimension size of the training data and much smaller than the highest dimension size, e.g., in our experiment $m > t$, $m < n$. This guarantees that the dictionary is over complete to reconstruct the data. Secondly, the dictionary size determines the reconstruction residual, as well as the SNR. As discussed in Section 3.4, if the SNR is not big enough, the reconstruction could not be stable. So the dictionary size should be big enough to satisfy certain level of SNR. But the dictionary could neither be too big, which will contain redundant information. Thus, our solution is to set the dictionary size which satisfy $t < m < 2t$. As discussed in the last paragraph, the interested components can be stably reconstructed in the certain range.

Table 4. The selected task related component items (atom IDs shown here) of one subject using different settings of dictionary sizes in the dictionary learning procedure.

Dictionary size	300	310	320	330	340	350	360	370	380	390	400
Selected Component IDs	#165	#165	#165	#165	#165	#165	#165	#165	#165	#165 #381	#165 #381

Dictionary size	410	420	430	440	450	460	470	480	490	500
Selected Component IDs	#165 #381	#165 #381	#165 #381	#165 #381	#165 #381	#165 #381	#165 #381	#165 #381	#165 #381	#165 #381

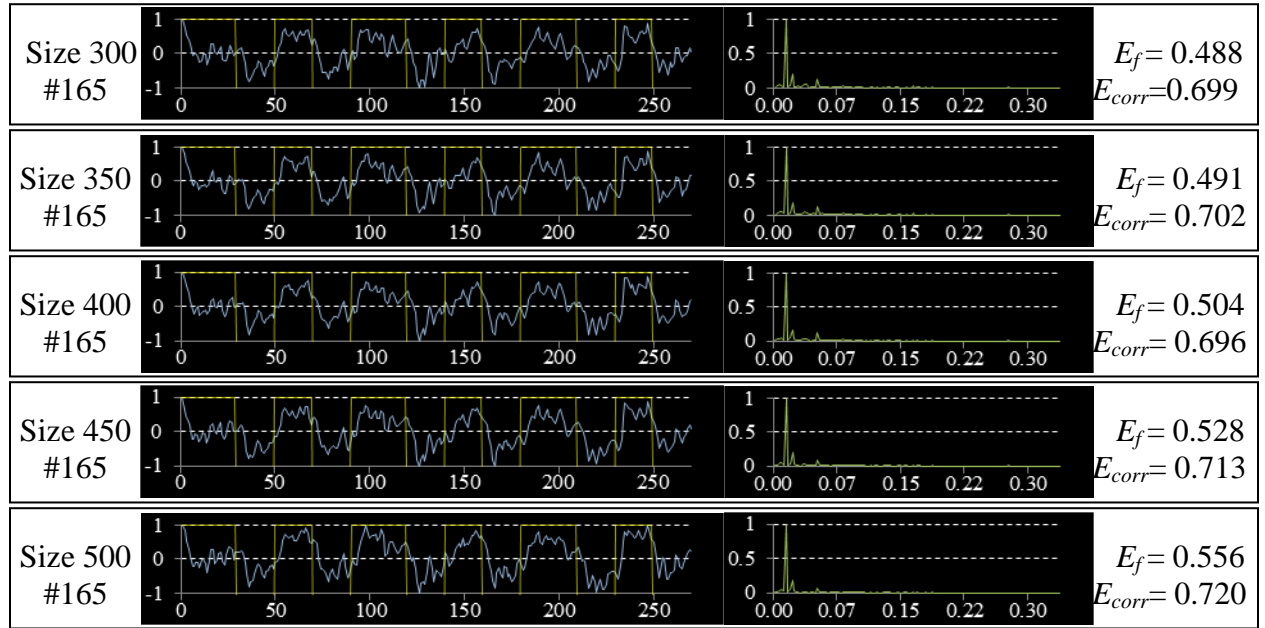
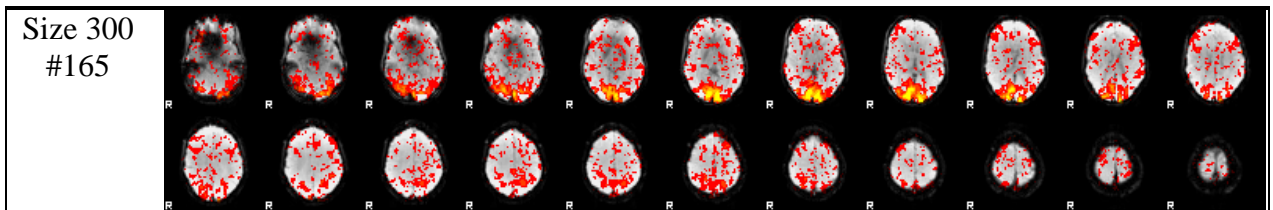


Figure 13. The temporal and frequency characteristics of the network component #165 with different dictionary sizes (300, 350, 400, 450 and 500). For each row in the figure, from the left to the right are: dictionary size and component index, time series signal of that component with overlaid stimulus curve (in yellow), the frequency spectrum of that component, and the value of component scores, respectively.



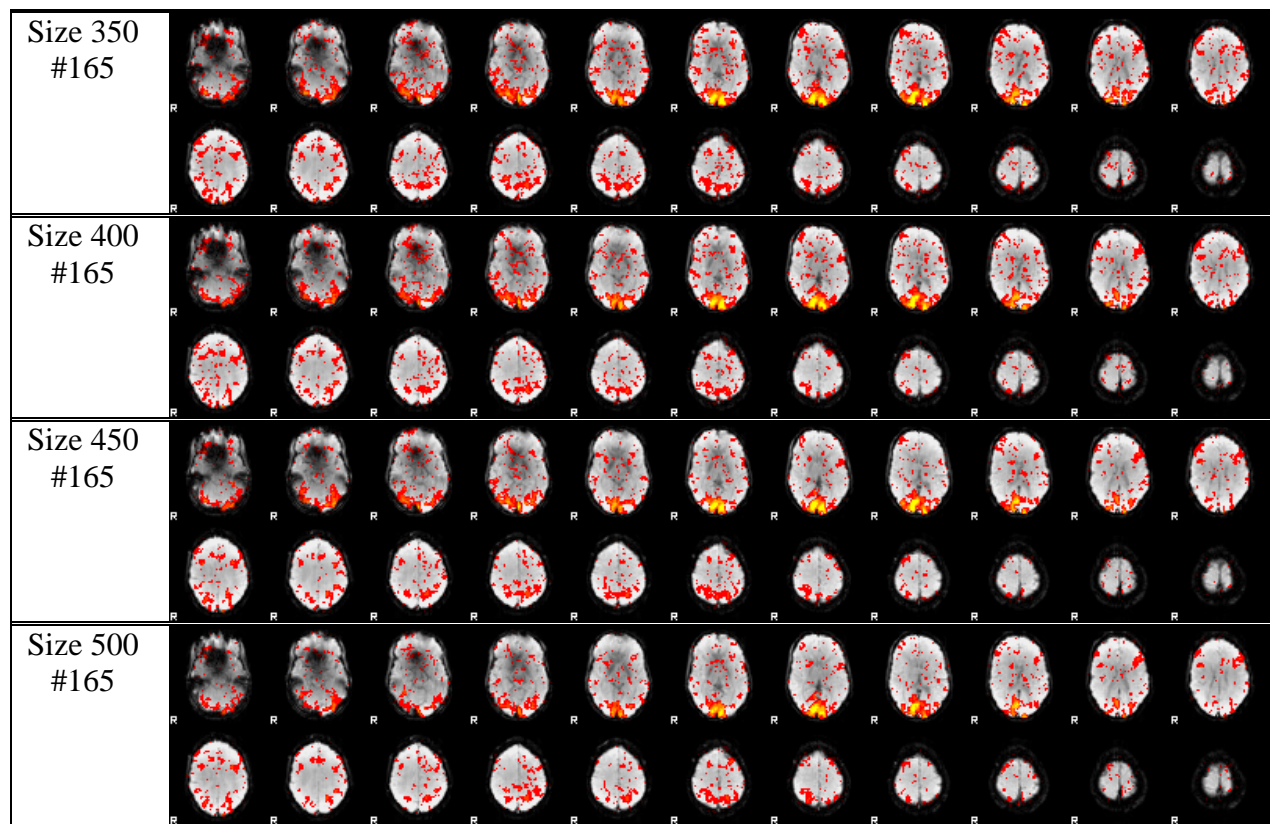


Figure 14. The spatial distribution patterns of network components #165 with different dictionary sizes (300, 350, 400, 450 and 500).

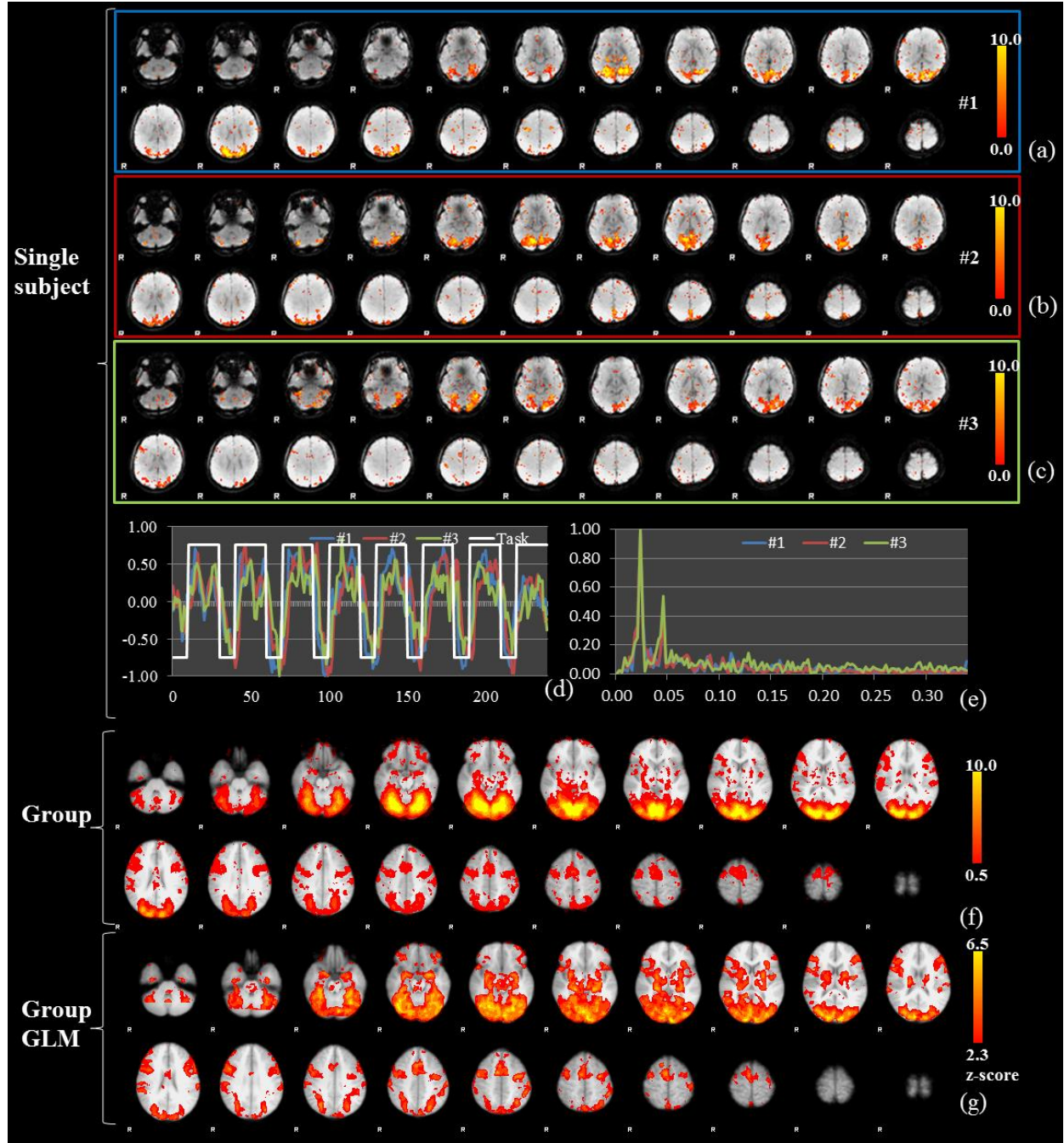


Figure 15. (a)-(c) Three identified task-related network components of a randomly selected subject in dataset 2. (d) The corresponding temporal time series patterns of the three components in (a)-(c). (e) The corresponding frequency distribution of the components in (a)-(c). (f) The group-wise averaged statistical map of all task components from 10 subjects in the MNI space. (g) Group-wise activation detected by FSL FEAT.

To further evaluate and validate the proposed methods, we applied the sparse representation framework in Section 2.3 on the dataset 2 in Section 2.2, and selected those task related components from the learned dictionary based on the criteria in Section 2.4. As an example, the spatial distribution patterns of three selected task related components of one subject are visualized in Figs. 15a-15c. The group-wise averaged map of the selected task related networks (Fig. 15f) among the 10 subjects is highly consistent (the overlapped area accounts for 67.8% of our result and 73.6% of the results by FSL FEAT) with the group activation detection results (Fig. 15g) obtained by FSL FEAT that is based on GLM. This result further demonstrates that the dictionary learning and sparse representation methods in this paper can reliably uncover meaningful brain networks and that this framework could potentially serve as a novel, alternative approach to detecting fMRI activation, as mentioned in Section 3.2.

Also, the corresponding temporal and frequency characteristics of these selected components in dataset 2 are shown in Figs. 15d-15e. It is apparent that the time series of these network components well follow the external task stimulus curve (the white curve in Fig. 15d). Also, the peak of the energy concentrations of these components, which is around 0.022Hz based on the frequency domain analysis, is exactly the same as the theoretic input frequency of the external stimulus, as calculated by the equation below:

$$\frac{1}{length\ of\ task + length\ of\ rest} * \frac{1}{TR} = \frac{1}{20+10} * \frac{1}{1.5} = 0.0222Hz \quad (15)$$

This result further demonstrates the effectiveness and accuracy of the online dictionary learning methods (Mairal et al., 2010) in extracting meaningful basis patterns for sparse representation of whole-brain fMRI data. Quantitatively, the temporal and frequency characteristics of the 20 selected example COIs from 10 subjects are shown in Table 5. These experimental results further showed that our sparse representation methods are robust and reproducible across independent datasets with different paradigm designs.

Table 5. The E_f and E_{corr} of selected COIs of 10 subjects in the semantic decision making task fMRI data.

Selected COIs	Sub.1 #157	Sub.1 #264	Sub.2 #184	Sub.2 #386	Sub.3 #235	Sub.3 #305	Sub.4 #164	Sub.4 #243	Sub.5 #253	Sub.5 #380
E_f	0.482	0.491	0.592	0.535	0.534	0.555	0.590	0.565	0.539	0.514
E_{corr}	0.473	0.503	0.659	0.547	0.586	0.590	0.601	0.704	0.607	0.611
Selected COIs	Sub.6 #31	Sub.6 #194	Sub.7 #223	Sub.7 #384	Sub.8 #170	Sub.8 #196	Sub.9 #368	Sub.9 #374	Sub.10 #205	Sub.10 #388
E_f	0.523	0.396	0.589	0.580	0.685	0.587	0.527	0.576	0.516	0.670
E_{corr}	0.653	0.478	0.580	0.599	0.727	0.474	0.527	0.635	0.721	0.723

3.6 Extended application on event-related fMRI data

In the field of neuroscience, event-related fMRI is another popular methodology, other than block design task fMRI, to analyze brain activations or networks. There are challenges in analyzing event-related fMRI data because neither the temporal pattern nor the frequency distribution of the stimulus is designed in a fixed fashion. Especially the frequency distribution of the event time series could be more complicated. In this section, we extend the application of our method to an open event-related fMRI dataset as detailed in Dataset 3 in Section 2.2. In our application, the $E_{corr,j}$ still applies, but considering the complex frequency distribution, we modified the energy function as:

$$E_{f,j} = corr(FS_{stimulus}, FD_j)$$

where $FD_{stimulus}$ is the frequency spectrum distribution curve of the stimulus which is obtained by applying Fourier transform to the stimulus time series and FD_j is the frequency spectrum distribution curve of the j^{th} atom in dictionary D. In other words, we use the correlation of the frequency spectrum to measure their similarity in the frequency domain.

With this modified approach, the identified task-related networks from the event-related fMRI data are presented in Fig. 16. Fig. 16a and 16b show the spatial distributions of two selected networks from one

single subject. Their temporal patterns are quite similar with the task event curve as shown in Fig. 16c. Meanwhile the frequency spectra of the networks are also akin to the frequency distribution of the stimulus (Fig. 16d). The group average of the networks from 26 subjects is shown in Fig. 16e, which agrees with the group-wise GLM result (Fig. 16f). In addition, the selected anti-task networks and DMN networks are shown in Fig. 17 and Fig. 18. They are also meaningful and reliable. In particular, the task-related networks and anti-task networks also agree with the results reported in Kelly et al., 2008, Mennes et al., 2010 and Mennes et al., 2011.

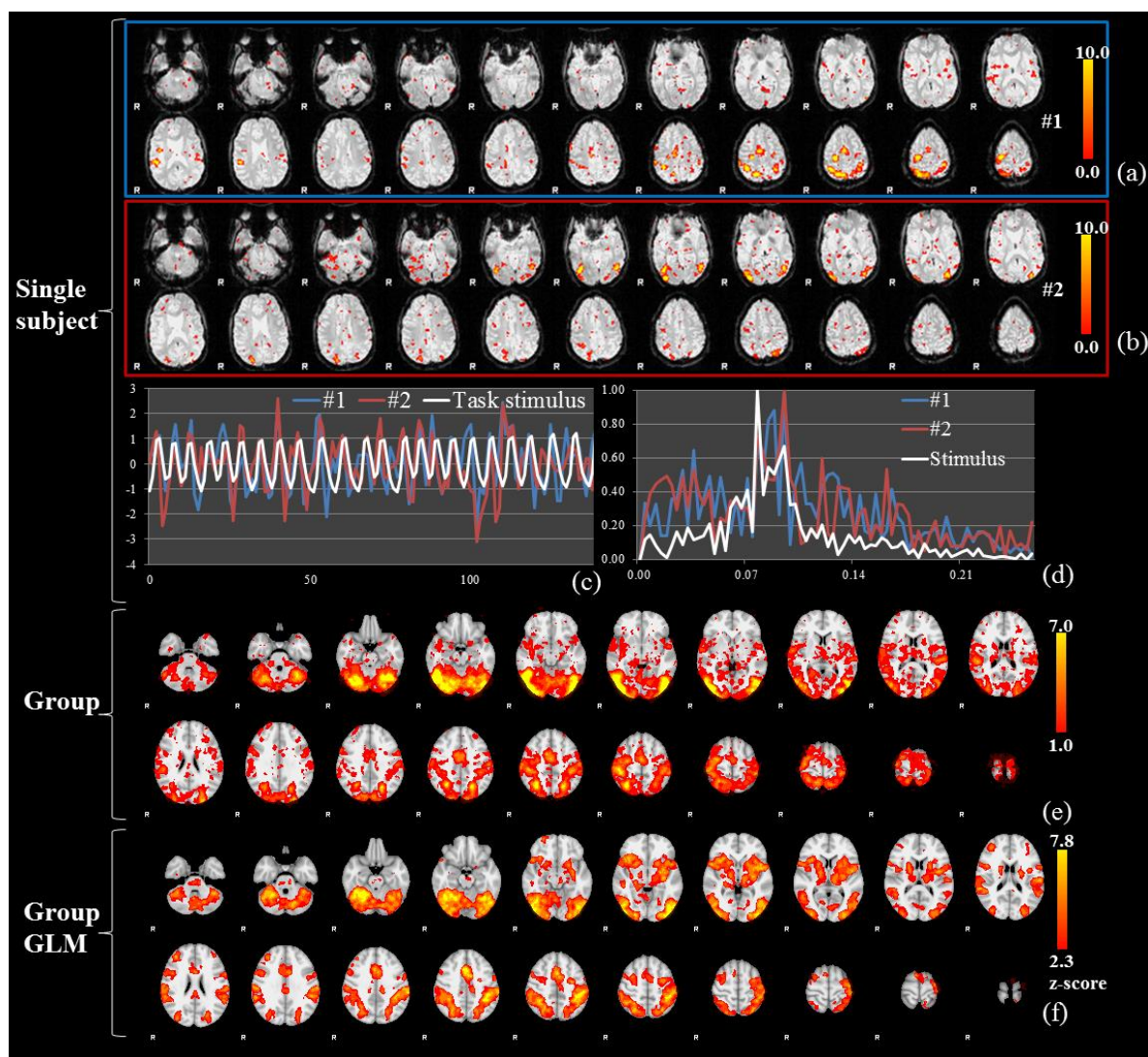


Figure 16. (a)-(b) Two identified task-related network components of a randomly selected subject in dataset 3. (c) The corresponding temporal time series patterns of the two components in (a)-(b). (d) The

corresponding frequency distributions of the components in (a)-(b). The white curves in (c) and (d) are temporal and frequency patterns of stimulus respectively. (e) The group-wise averaged statistical map of all task components from 26 subjects in the MNI space. (f) Group-wise activation detected by FSL FEAT.

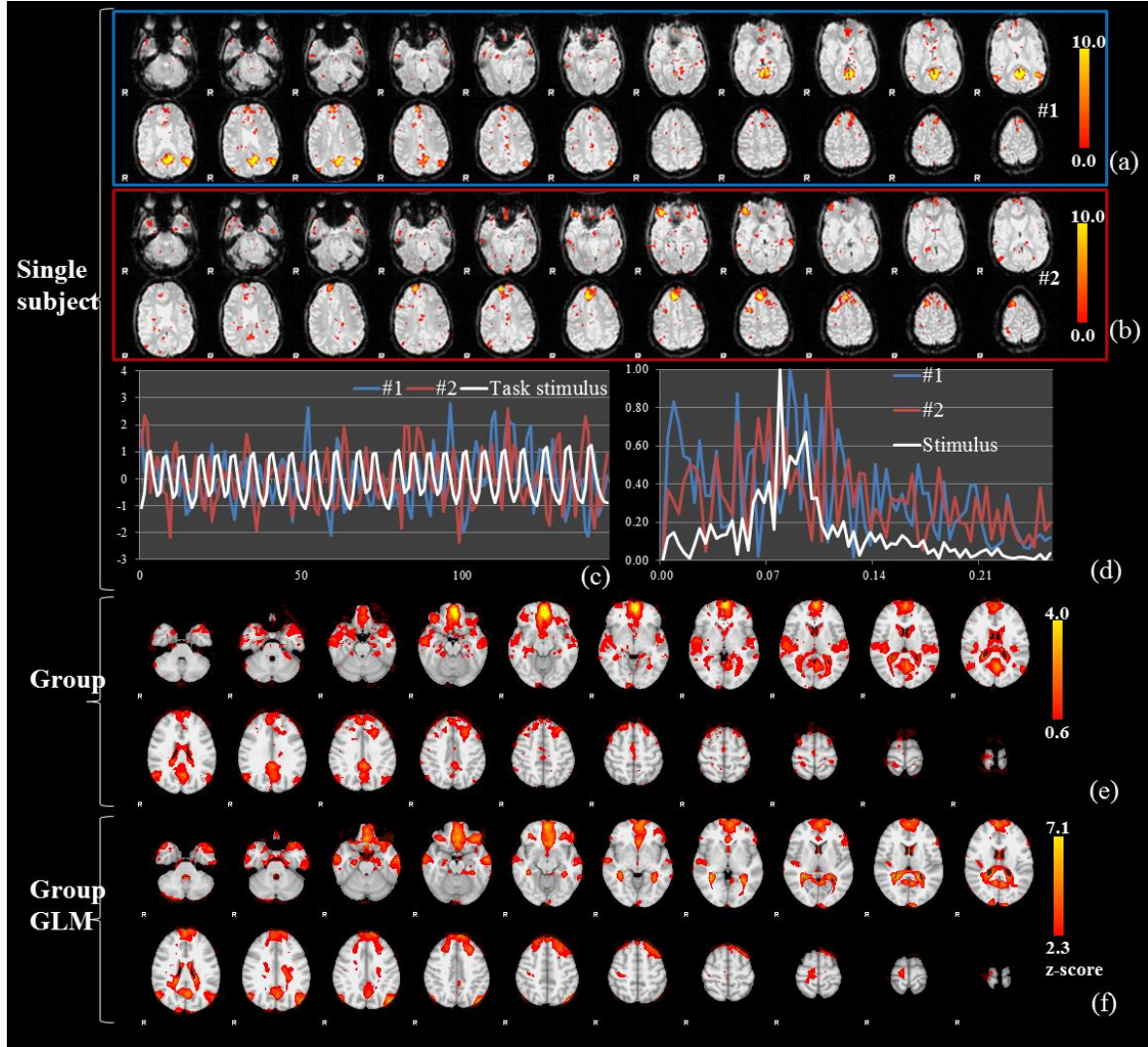


Figure 17. (a)-(b) Two identified anti-task network components of a randomly selected subject in dataset 3. (c) The corresponding temporal time series patterns of the two components in (a)-(b). (d) The corresponding frequency distributions of the components in (a)-(b). The white curves in (c) and (d) are temporal and frequency patterns of stimulus respectively. (e) The group-wise averaged statistical map of all task components from 26 subjects in the MNI space. (f) Group-wise activation detected by FSL FEAT.

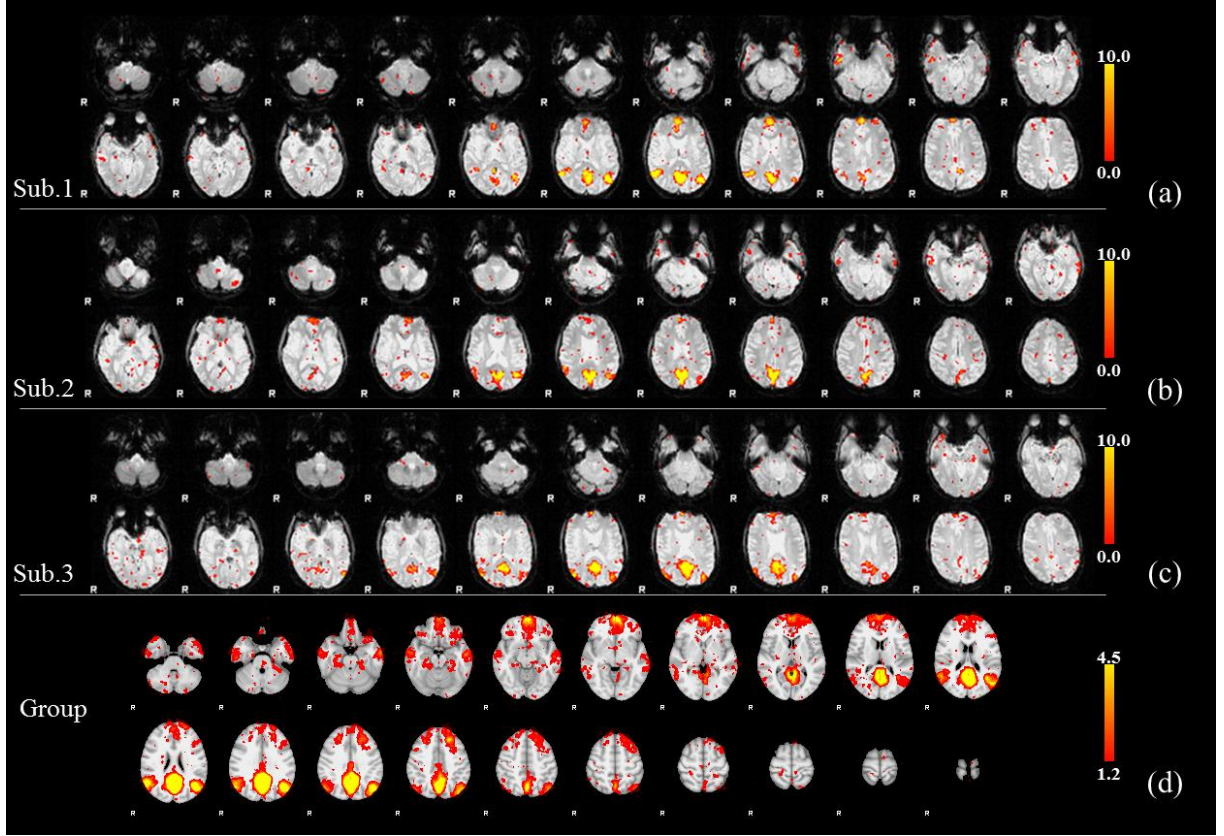


Figure 18. (a-c) Identified default mode network components of three random selected subjects in Dataset 3. (d) The group-wise averaged statistical map of all DMN components from 26 subjects of dataset 3 in the MNI space.

4. DISCUSSION AND CONCLUSION

4.1 Co-activated Networks

Brain regions or networks that are evoked by external stimulus may react in different patterns even though they are all highly correlated with task design. This may be attributed to physical variations, e.g., different HRFs of different regions, however, it is also likely that these variations across networks stem from their different streams or different levels in the brain information flow, as well as their interactions or communications. Thus, modelling brain activations with a uniformed task design in traditional methods is coarse. In our experiments, based on the temporal and frequency criteria, multiple task-related networks

are selected for most of the subjects. From the inspection on Fig. 8, Supplemental Fig. 2 and Supplemental Fig. 4-5, we can find that the temporal and frequency patterns of the multiple networks are all in agreement with the stimulus design, though noticeable difference can also be found among them. Spatially, they perform distinct patterns with ratios of overlaps in each single subject, however, they are all sub-regions of the activations and their aggregation is very similar to the activation pattern from the GLM method. Here we call them co-activated networks. Comparing with the model-driven method which only produces strength of activation, our data-driven framework provides a new window to investigate how sub-networks could interact or cooperate to complete a task in the brain. Although some of them are not stable due to reasons such as parameter selection, but we can still find very stable and dominant networks at the visual cortex, superior/middle frontal gyrus, precentral gyrus, and superior parietal lobe, which are associate with working memory processes.

4.2 Anti-task Network and DMN

Previous studies (Fox et al., 2005; Fransson et al., 2005) have reported that DMN tends to exhibit anti-task performance in task fMRI scans. This point is supported in our study. As observed in Table 1 and Supplemental Fig. 10, anti-task networks and DMNs exhibit substantially high overlap rates. For some subjects, they even have overlaid component IDs, e.g., for subject 1 anti-task networks include component #330 and #387, and the component #387 is also identified as DMN. Meanwhile, in our experiments based on three datasets, the anti-task networks are not limited to DMN, i.e., there are more regions or sub-networks that exhibit anti-task performance, especially when comparing group-wise results in Fig. 9, Fig. 10, Fig. 17 and Fig. 18. From the perspectives of temporal and frequency domains (Supplemental Fig. 9), our detected components of DMNs have negative correlation but the strength is not as high as the task ones. We can observe that the signal patterns of DMN may involve multiple frequencies of fluctuation. It is essential to point out that the selection of anti-task network is based on temporal frequency analysis, while the identification of DMN is based on spatial similarity. The final

overlap of component IDs suggests the effectiveness of both approaches. Also, this result provides a clue that a brain network could be profiled or characterized in multi-domains, suggesting the data-driven dictionary learning and sparse coding methods are effective in modeling brain networks.

4.3 Networks on Thalamus

Based on the analyses of the three datasets, especially when comparing the group-wise networks and group-wise GLM activations in Fig. 8, Fig. 15 and Fig. 16, it is easy to observe a common phenomenon that the task-related networks in our method do not include the thalamus areas. In contrast, across all the three datasets, the GLM based method determined the thalamus or part of it as activation. To explore this question, we used the similar method as that for detecting DMN, that is, the thalamus template was employed to filter all the components of each subject. Finally the most spatially similar network was selected for each subject, as shown in Figs. 19a-19c and Supplemental Fig. 14 (I-II) (a-c). From the figures, we can see that in our method the thalamus was learned into a single network, and the spatial patterns are quite consistent across subjects and across datasets (Fig. 19d, Supplemental Fig. 14 (I-II) d). We further inspected the temporal and frequency patterns of these selected thalamus networks. As visualized in Fig. 20 and Supplemental Fig. 15, the temporal patterns and frequency patterns are complicated and individualized. In particular, the energies of signals are distributed on multiple frequency bands. This is in agreement with the thalamus' complex functions of relaying sensory and motor signals to the cerebral cortex. It is meaningful that the GLM could detect task activations in the thalamus, however, the activation strength is not as strong as other task-evoked areas. This may also be attributed to the complex functions that the thalamus plays. In our view, the functions of thalamus could not be straightforwardly described as strength of activations or how much its fMRI signals follow the stimulus curve. More detailed analysis should be carried out based on our data-driven framework in the future.

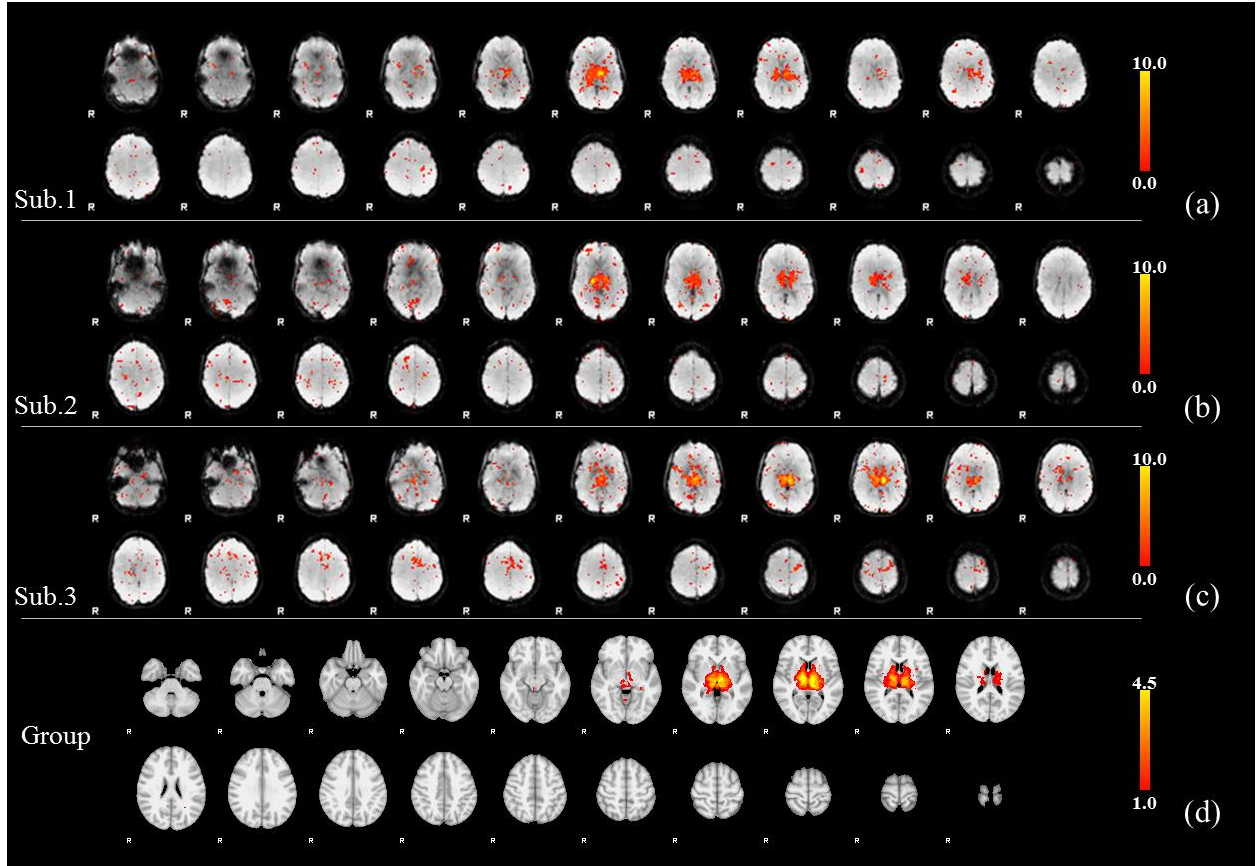


Figure 19. (a-c) Identified thalamus network components of three random selected subjects in Dataset 1. (d) The group-wise averaged statistical map of all thalamus components from 15 subjects of Dataset 1 in the MNI space.

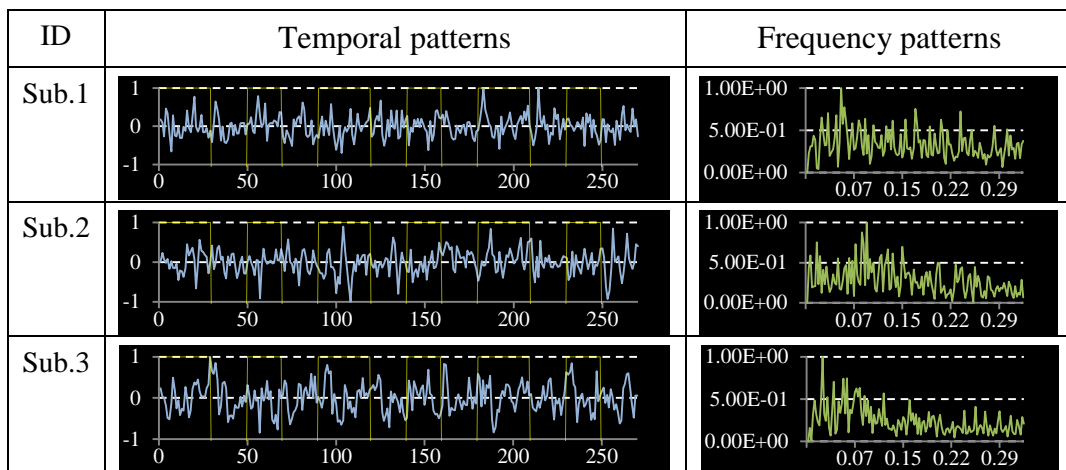


Figure 20. The temporal and frequency patterns of the selected thalamus networks in Fig.19a-19c from Dataset 1.

4.4 Other Detected Networks

Brain function is complex, and thus simple models are not likely sufficient in modeling all the networks that are active or idle, directly or indirectly participating in the task performance. With our data-driven strategy, the whole brain can be decomposed into hundreds of distinct networks with specific activity patterns. These networks are not limited to task-related networks, anti-task networks, DMN and the thalamus. There are also other networks that call for further new methods to characterize them. For example, in the working memory task in which the task is presented as vision stimulus, in addition to the task-related networks we can also find auditory network and motor network among the hundreds of networks as shown in Fig. 21a and Fig. 21b. Their temporal activities are shown in Fig.22a and Fig.22b. Some neuroanatomic areas can be also grouped into a single network, such as the ventricle areas as shown in Fig. 21c and Supplemental Fig.16, the time series of which are shown in Fig.22c and Supplemental Fig.17, respectively. Although we still do not know if the BOLD activity of the ventricle network is caused by motion noise or it is physically meaningful, our method provides a new window to explore in the future. In summary, our premise is that all these networks intrinsically exist in the brain no matter what task the brain is performing; however, in different tasks these networks may exhibit different patterns of activities and interactions. In our future work, new methods will be developed to characterize all these networks and to model their interactions.

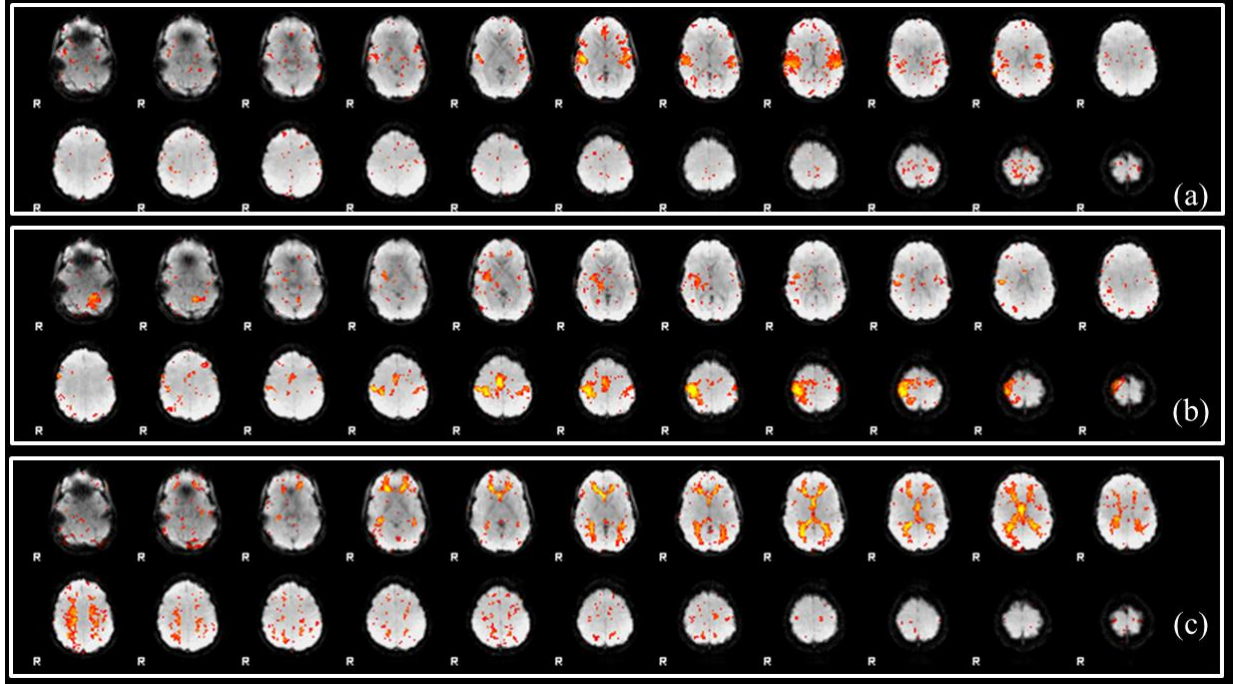


Figure 21. Other networks detected in the working memory task dataset (for one randomly selected subject). (a) Auditory network. (b) Motor network. (c) Ventricle network.

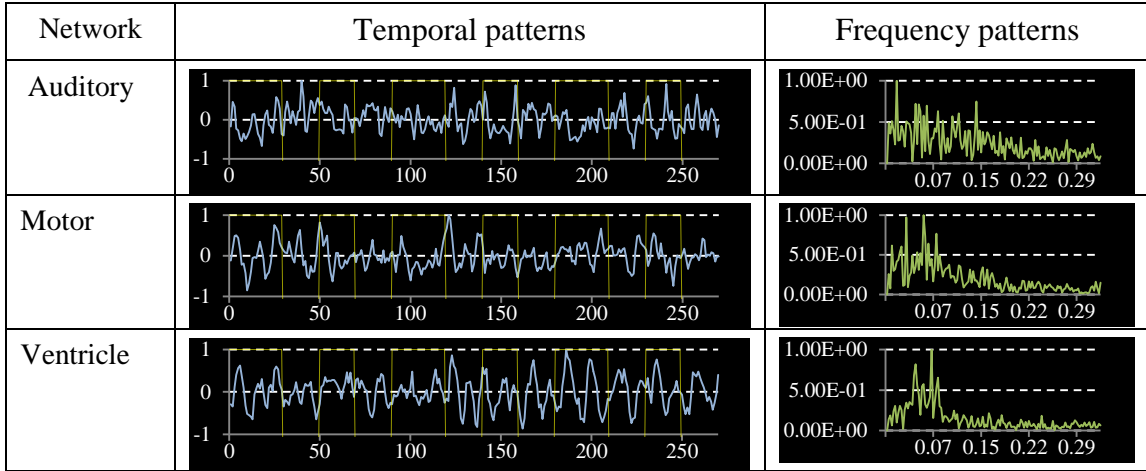


Figure 22. The corresponding temporal and frequency patterns of the networks in Fig.21a-21c from Dataset 1.

4.5 Conclusion

We have presented a novel computational methodology of representing whole-brain fMRI signals via sparse coding and dictionary learning. The basic idea is to aggregate all of fMRI signals within the whole

brain of one subject into a big data matrix, which is factorized into an over-complete dictionary basis matrix and a reference weight matrix via an online dictionary learning algorithm. We then designed a computational framework for quantitative characterization of the dictionary components in temporal, frequency and spatial domains. The interesting result from this work is that the decomposed atomic dictionary components in a working memory task-based fMRI dataset, a semantic decision making task-based fMRI dataset and a flanker event-related fMRI dataset exhibit functionally meaningful spatial, temporal and frequency patterns, as shown in the results sections. Our results not only demonstrated the effectiveness of data-driven sparse representation for task-based fMRI signals in identifying multiple functional networks, but also revealed how these functional networks contribute to the compositions of dozens of thousands of fMRI signals within the whole brain. The comparisons of the results by our method with those by group-wise GLM and ICA methods, as well as with the synthesized data with ground-truth, demonstrated the validity, robustness, reproducibility and effectiveness of our methods. In general, our work potentially provides a novel, alternative window to examine the holistic functional activities of the brain.

Meanwhile, the methods and work in this paper can be further enhanced and expanded in the following directions in the future. In this paper, we have focused on the quantitative characterization of spatial, temporal and frequency patterns of three well-known categories of dictionary atoms including the task-related, anti-task related and default mode network components, as shown in Section 3.2. It should be pointed out that there are many other potentially important and meaningful dictionary network components such as auditory network, motor network, and Thalamus that should be examined and characterized in the future. For instance, Fig. 23 shows the spatial distributions of three other dictionary components (the version on the original surface of Fig. 23 can be found in Supplemental Fig. 13(II)). It is clear that those components are spatially overlapping and temporally interacting with each other. Quantitative description and characterization of those hundreds of dictionary components across multiple

individuals are warranted to comprehensively understand and represent the functional activities of the brain in the task-based fMRI data in the future. Once successful, these well-characterized components can be potentially used as functional network atlases for other brain mapping applications in the future.

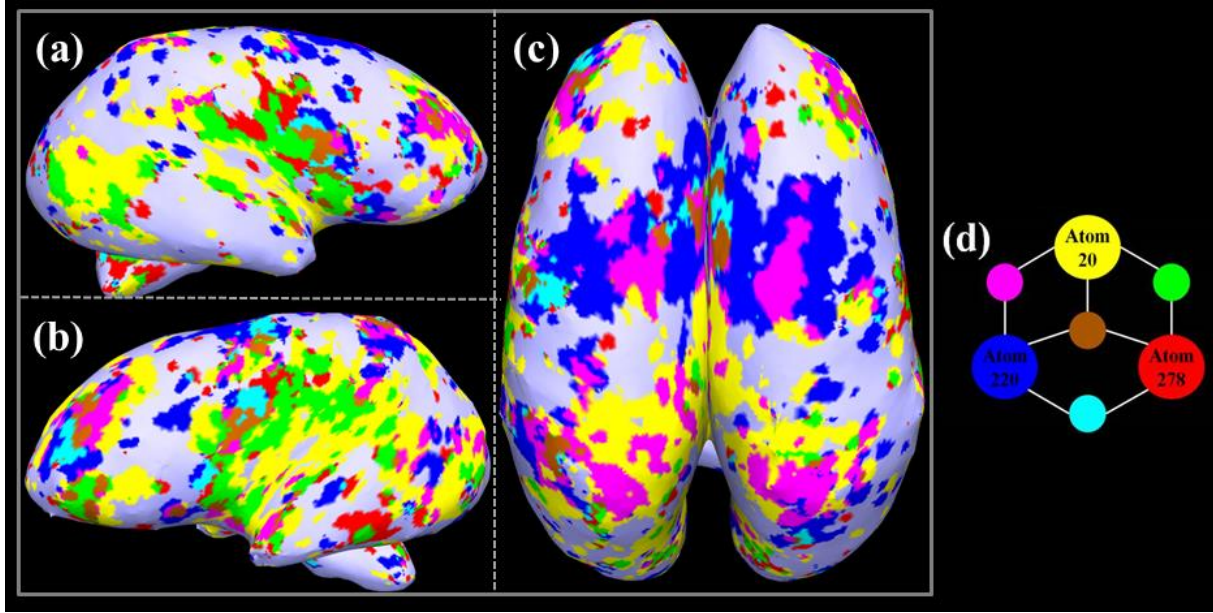


Figure 23. Visualization of three other network components (in addition to the three components in Fig. 1) on the inflated cortical surface. They are atom #20 (yellow), atom #220 (blue) and atom #278 (red) extracted from the learned dictionary of the same subject in Fig. 1. (a-c) show different views of representing the spatial distribution patterns of the three components. (d) demonstrates the color scheme of different components fusion. For example, the regions belonging to both atoms #220 and #278 are represented as cyan, and green stands for the overlapped areas of atoms #20 and #278.

At the current stage, the sparse representation framework was applied on each individual subject independently. Our experimental results have demonstrated the promising reproducibility and robustness of the framework in which those COIs identified in different subjects can find reasonably good correspondences based on their spatial, temporal and frequency characteristics (Section 3.2). In the future, we plan to explore if/how the whole-brain fMRI signals from a group of subjects can be aggregated into one big data matrix so that the dictionary atoms can be learned together and interpreted across a

population. In this case, the commonly shared dictionary network components, such as the task related and anti-task related components, might exhibit consistent functional responses to the same stimulus and possess intrinsically established correspondences across individuals. However, it would be more challenging to interpret those intrinsic network components since their temporal and/or frequency properties might exhibit much more variable and heterogeneous patterns across individuals, as already shown in Supplemental Fig. 9. It is expected that the spatial distribution templates, e.g., that in Supplemental Fig. 1, might be an effective constraint to map those intrinsic network components and we plan to explore such possibility for other intrinsic networks (Fox and Raichle, 2007; Cohen et al., 2008; van den Heuvel et al., 2008).

Our results in Section 3.2 has demonstrated that intrinsic networks such as the DMN (Fox and Raichle, 2007; Cohen et al., 2008; van den Heuvel et al., 2008) are active in task performance state and they can be successfully identified by our sparse representation methods. Our results also demonstrated that the intrinsic networks are spatially overlapping with other task related and anti-task related network components. However, the functional roles of these intrinsic networks and the interactions among them (e.g., Fig. 10) and task/anti-task related networks (e.g., Figs. 8-9), as well as their temporal dynamics, during task performance are not sufficiently characterized at the current stage, which should be extensively explored in the future. In this sense, novel computational and statistical approaches to modeling network-level functional interactions and dynamics (Li et al. 2013) should be developed, validated and employed. We envision that the elucidation of such complex functional interaction and dynamics patterns within large scale networks, such as the hundreds of basis components in the learned dictionary in this work, would significantly advance our understanding of the functioning mechanisms of the human brain.

At the current stage, our methods were applied on the working memory task fMRI dataset, semantic decision making task fMRI dataset and the flanker event-related fMRI dataset. In the future, we plan to apply the methods on other publicly available task-based fMRI datasets such as the recently released Human Connectome Project datasets. Then, we will be able to represent and characterize potentially many other functional networks under different task performances. In addition, we plan to investigate the possibility of applying the proposed sparse representation framework on resting state fMRI datasets and potentially characterize those uncovered intrinsic networks. We predict that once the collection of those well-characterized functional networks are replicated and validated in independent datasets by different research labs, they can be used as comprehensive atlases of functional brain activities for many brain mapping applications, such as measuring functional connectivities and interactions during different task performances, in healthy human brains. In a longer term, those methodologies can be then applied in different brain disorders and conditions to potentially reveal the functions and dysfunctions of many brain disorders such as Alzheimer's disease and Schizophrenia.

Acknowledgements

T Liu was supported by NSF CAREER Award (IIS-1149260), NIH R01 DA-033393, NIH R01 AG-042599, NSF CBET-1302089, and NSF BCS-143905. J Lv was supported by the Doctorate Foundation of Northwestern Polytechnical University. L Guo was supported by the NSFC #61273362. J Zhang was supported by start-up funding and Sesseel Award from Yale University. L Guo was supported by the National Science Foundation of China under Grant 61273362. J Han was supported by the National Science Foundation of China under Grant 61005018 and 91120005, NPU-FFR-JC20120237 and Program for New Century Excellent Talents in University under grant NCET-10-0079. X Hu was supported by the National Science Foundation of China under Grant 61103061, China Postdoctoral Science Foundation under Grant 20110490174 and 2012T50819.

References

- Abolghasemi V., Ferdowsi S., Sanei S., Fast and incoherent dictionary learning algorithms with application to fMRI, *Signal, Image and Video Processing*, (2013). Online: <http://link.springer.com/article/10.1007%2Fs11760-013-0429-2>.
- Andersen, A.H., Gash D.M. and Avison M.J., Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework. *Magnetic Resonance Imaging*, 17(6): p. 795-815 (1999).
- Archer JS, Abbott DF, Waites AB, Jackson GD. fMRI "deactivation" of the posterior cingulate during generalized spike and wave. *NeuroImage*. Dec;20(4):1915-22. (2003).
- Bandettini, P.A., Jesmanowicz A., et al., Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med* 30(2): 161-173 (1993).
- Beckmann C.F., De Luca M., Devlin J.T., Smith S.M., Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc London* 360(1457):1001–1013 (2005).
- Bisley J.W, Pasternak T., The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion, *Cereb Cortex* 10: 1053–1065, 2000.
- Bullmore E., Fadili J., Breakspear M., Salvador R., Suckling J., and Brammer M., Wavelets and statistical analysis of functional magnetic resonance images of the human brain, *Stat. Methods Med. Res.*, vol. 12, no. 5, pp. 375-399, (2003).
- Cohen A.L., Fair D.A., Dosenbach N.U., Miezin F.M., Dierker D., Van Essen D.C., Schlaggar B.L., and Petersen S.E., Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage*, 41(1):45-57, 2008.
- Cha, S.H., Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 1(4), pp.300-307 (2007).

- Daubechies, I., et al., Independent component analysis for brain fMRI does not select for independence. Proceedings of the National Academy of Sciences, 106(26), 10415-10422 (2009).
- Deng F., Zhu D., Lv J., Guo L. and Liu T., FMRI Signal Analysis Using Empirical Mean Curve Decomposition, in press, IEEE Transactions on Biomedical Engineering (2012).
- Descombes X., Kruggel F., and von Cramon D.Y., fMRI signal restoration using a spatio-temporal Markov random field preserving transitions, NeuroImage, vol. 8, no. 4, pp. 340-349, (1998).
- Donoho D.L., Compressed sensing. IEEE Trans. Inf. Theory, 52:1289-1306, (2006).
- DuBois Bowman F., Caffo B., Bassett S.S., Kilts C., Bayesian hierarchical framework for spatial modeling of fmri data. NeuroImage, 39, 146–156 (2008).
- Faraco C.C., Unsworth N, Langley J., Terry D., Li K., Zhang D, Liu T, Miller LS, Complex span tasks and hippocampal recruitment during working memory. Neuroimage. 55(2):773-87 (2011).
- Fischera M.H., Zwaan R.A., Embodied language: A review of the role of the motor system in language comprehension, The Quarterly Journal of Experimental Psychology 61(6), 825-850. (2008).
- Fogassi L., Ferrari P.F., Gesierich B., Rozzi S., Chersi F., Rizzolatti G., Parietal Lobe: From Action Organization to Intention Understanding, Science, 308 (5722): 662-667. (2005).
- Fox, M.D., Raichle, M.E.: Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nature reviews. Neuroscience 8, 700-711 (2007).
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences of the United States of America, 102(27), 9673-9678 (2005).
- Fransson, P.: Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. Human brain mapping, 26(1), 15-29 (2005).
- Friston, K.J., Modalities, Modes, and Models in Functional Neuroimaging. Science 326, 399-403. 2009.

- Friston, K.J., et al., Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, 1994. 2(4): p. 189—210 (1994).
- Hartvig N.V. and Jensen J.L., Spatial mixture modeling of fMRI data, *Hum. Brain Mapp.*, vol. 11, no. 4, pp. 233-248, (2000).
- Huang K. and Aviyente S., Sparse representation for signal classification, *NIPS*, (2006).
- Jaccard, P., Étude comparative de la distribution florale dans une portion des Alpes et des Jura, *Bulletin de la Société Vaudoise des Sciences Naturelles* 37: 547–579 (1901).
- Kelly, A. M., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. Competition between functional brain networks mediates behavioral variability. *Neuroimage*, 39(1), 527-537, (2008).
- Lalonde J., Chaudhuri A., Task-dependent transfer of perceptual to memory representations during delayed spatial frequency discrimination, *Vision Research* 42(14): 1759–1769, (2002).
- Lee, H., Battle, A., Raina, R., & Ng, A. Y. Efficient sparse coding algorithms. *Advances in Neural Information Processing Systems* , 19 , 801–808. (2007).
- Lee K., Tak S., Ye J.C., A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion, *IEEE Trans Med Imaging*. 30(5):1076-89 (2011).
- Lee J., Jeong Y. and Ye J.C., Group sparse dictionary learning and inference for resting-state fMRI analysis of Alzheimer's disease, *ISBI*, (2013).
- Li Y., Namburi P., Yu Z., Guan C., Feng J., Gu Z., Voxel selection in FMRI data analysis based on sparse representation, *IEEE Trans Biomed Eng.* 56(10):2439-51(2009).
- Li Y., Long J., He L., Lu H., Gu Z., et al., A Sparse Representation-Based Algorithm for Pattern Localization in Brain Imaging Data Analysis, *PLoS ONE* 7(12): e50332 (2012).

- Li X., Zhu D., Jiang X., Jin C., Zhang X., Guo L., Zhang J., Hu X., Li J., Liu T., Dynamic Functional Connectomics Signatures for Characterization and Differentiation of PTSD Patients, in press, Human Brain Mapping, (2013).
- Logothetis N.K., What we can do and what we cannot do with fMRI, *Nature*, 453, 869-878 (2008).
- Lu Jinglei, Li Xiang, Zhu Dajiang, Jiang Xi, Zhang Xin, Guo Lei, Liu Tianming, Sparse Representation of Group-wise FMRI Signals, in press, MICCAI 2013, (2013).
- Mairal, J., Bach, F., Ponce, J., & Sapiro, G. Online learning for matrix factorization and sparse coding. *The Journal of Machine Learning Research*, 11, 19-60 (2010).
- Mallat, S. G., & Zhang, Z. Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions on*, 41(12), 3397-3415 (1993).
- Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., and Byers A. H. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. May (2011).
- McKeown, M.J., et al., Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. *PNAS*, 1998. 95(3): p. 803(1998).
- Mennes, M., Kelly, C., Zuo, X. N., Di Martino, A., Biswal, B. B., Castellanos, F. X., & Milham, M. P.: Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. *Neuroimage*, 50(4), 1690-1701, (2010).
- Mennes, M., Zuo, X. N., Kelly, C., Di Martino, A., Zang, Y. F., Biswal, B., ... & Milham, M. P.: Linking inter-individual differences in neural activation and behavior to intrinsic brain dynamics. *Neuroimage*, 54(4), 2950-2959, (2011).
- Oikonomou V.P., Blekas K., Astrakas L., A sparse and spatially constrained generative regression model for fMRI data analysis, *IEEE Trans Biomed Eng.* 59(1):58-67(2012). doi: 10.1109/TBME.2010.2104321.

- Olshausen, B.A. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607-609 (1996).
- Olshausen, B.A., & Field, D.J., Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4), 481-487 (2004).
- Quiroga, R.Q., Kreiman, G., Koch, C., & Fried, I., Sparse but not ‘grandmother-cell’ coding in the medial temporal lobe, *Trends in cognitive sciences*, 12(3), 87-91(2008).
- Raichle, M.E., Snyder, A.Z., A default mode of brain function: A brief history of an evolving idea. *NeuroImage* 37 (4): 1083–90 (2007).
- Shimizu Y., Barth M., Windischberger C., Moser E., and Thurner S., Wavelet-based multifractal analysis of fMRI time series, *Neuroimage*, vol. 22, pp. 1195-1202, (2004).
- Tomasi D, Ernst T, Caparelli EC, Chang L. Common deactivation patterns during working memory and visual attention tasks: an intra-subject fMRI study at 4 Tesla. *Hum Brain Mapp*. 27(8):694-705. (2006).
- Van Den Heuvel M., Mandl R., and Hulshoff Pol H., Normalized cut group clustering of resting-state fMRI data. *PLoS One*, 3(4): e2001, (2008).
- Woolrich, M., Ripley, B., Brady, J., Smith, S., Temporal autocorrelation in univariate linear modelling of FMRI data, *NeuroImage* 14 (6), 1370– 1386 (2001).
- Worsley, K.J. An overview and some new developments in the statistical analysis of PET and fMRI data. *Hum Brain Mapp* 5(4): 254-258 (1997).
- Wright J., Yang A.Y., Ganesh A., Sastry S.S., and Ma Y., Robust face recognition via sparse representation, *IEEE Trans. PAMI*, (2008).
- Wright, J., et al., Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 98(6), 1031-1044 (2010).

- Yang M., Feng X., Zhang D., Fisher discrimination dictionary learning for sparse representation. ICCV (2011).
- Zaksas D., Pasternak T.J., Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task, *Neuroscience* 26(45):11726-42. (2006).
- Zhu D., Li K., Faraco C.C., Deng F., Zhang D., Jiang X., Chen H., Guo L., Miller S., Liu T., Optimization of Functional Brain ROIs via Maximization of Consistency of Structural Connectivity Profiles, 59(2):1382-93., *NeuroImage* (2012).
- Zhu D., Li K., Guo L., Jiang X., Zhang T., Zhang D., Chen H., Deng F., Faraco C., Jin C., Wee C.Y., Yuan Y., Lv P., Yin Y., Hu X., Duan L., Hu X., Han J., Wang L., Shen D., Miller L.S., Li L., Liu T., DICCOL: dense individualized and common connectivity-based cortical landmarks. *Cereb Cortex* 23: 786-800 (2013).