Sparse SPM: Sparse-Dictionary Learning for Resting-state Functional Connectivity MRI Analysis

Article *in* NeuroImage · November 2015

DOI: 10.1016/j.neuroimage.2015.10.081

8 authors, including:

Kangjoo Lee
McGill University
7 PUBLICATIONS **107** CITATIONS
SEE PROFILE

Duk L Na
Sungkyunkwan University
485 PUBLICATIONS **5,132** CITATIONS
SEE PROFILE

Yong Jeong
Korea Advanced Institute of Science and Tec…
104 PUBLICATIONS **1,311** CITATIONS
SEE PROFILE

Jong Chul Ye
Korea Advanced Institute of Science and Tec…
169 PUBLICATIONS **2,309** CITATIONS
SEE PROFILE

NeuroImage 125 (2016) 1032–1045

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Sparse SPM: Group Sparse-dictionary learning in SPM framework for resting-state functional connectivity MRI analysis

Young-Beom Lee [a], Jeonghyeon Lee [b,1], Sungho Tak [d], Kangjoo Lee [e], Duk L. Na [c], Sang Won Seo [c], Yong Jeong [a,*], Jong Chul Ye [b,*], The Alzheimer's Disease Neuroimaging Initiative [2]

[a] Laboratory for Cognitive Neuroscience and NeuroImaging, Dept. of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), 373–1 Guseong-dong Yuseong-gu, Daejeon 305–701, South Korea
[b] Bio Imaging and Signal Processing Lab., Dept. of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), 373–1 Guseong-dong Yuseong-gu, Daejeon 305–701, South Korea
[c] Department of Neurology, and Neuroscience Center Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea
[d] Wellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London WC1N 3BG, UK
[e] Montreal Neurological Institute, McGill University, Montreal, QC H3A 2B4, Canada
[1] Currently at Korea Hydro and Nuclear Power Co., Ltd, Gyeongju, South Korea

ARTICLE INFO

Article history:
Received 26 June 2015
Accepted 27 October 2015
Available online 31 October 2015

Keywords:
Resting-state fMRI analysis
Functional connectivity
Sparse graph
K-SVD
Sparse dictionary learning
Alzheimer's disease

ABSTRACT

Recent studies of functional connectivity MR imaging have revealed that the default-mode network activity is disrupted in diseases such as Alzheimer's disease (AD). However, there is not yet a consensus on the preferred method for resting-state analysis. Because the brain is reported to have complex interconnected networks according to graph theoretical analysis, the independency assumption, as in the popular independent component analysis (ICA) approach, often does not hold. Here, rather than using the independency assumption, we present a new statistical parameter mapping (SPM)-type analysis method based on a sparse graph model where temporal dynamics at each voxel position are described as a sparse combination of global brain dynamics. In particular, a new concept of a spatially adaptive design matrix has been proposed to represent local connectivity that shares the same temporal dynamics. If we further assume that local network structures within a group are similar, the estimation problem of global and local dynamics can be solved using sparse dictionary learning for the concatenated temporal data across subjects. Moreover, under the homoscedasticity variance assumption across subjects and groups that is often used in SPM analysis, the aforementioned individual and group analyses using sparse dictionary learning can be accurately modeled by a mixed-effect model, which also facilitates a standard SPM-type group-level inference using summary statistics. Using an extensive resting fMRI data set obtained from normal, mild cognitive impairment (MCI), and Alzheimer's disease patient groups, we demonstrated that the changes in the default mode network extracted by the proposed method are more closely correlated with the progression of Alzheimer's disease.

Introduction

Spontaneous low-frequency fluctuations (<0.1 Hz) of blood oxygen level-dependent (BOLD) signals during resting states have been shown to represent cognitive functions and neural physiology (Biswal et al., 1995; Cordes et al., 2001; Damoiseaux et al., 2006). Spatiotemporally distinct resting-state networks have been consistently identified in the primary visual network, default mode network, salience network, fronto-parietal network, and sensory motor network, among others (De Luca et al., 2006). Among the various resting-state subnetworks, the default mode network (DMN), which significantly deactivates during cognitive task-related experiments, has been studied extensively in functional connectivity analyses. It has been shown that the DMN is closely involved with episodic memory processing (Lustig et al., 2003; Greicius et al., 2004). Furthermore, previous works have provided evidence that the PCC, which shows a neural deactivation in early Alzheimer's disease (AD), is the first brain region to exhibit decreased metabolism in AD patients (Minoshima et al., 1997).

Seed-based approaches (Lowe et al., 1998; Rombouts et al., 2003; Fransson, 2005; Fox et al., 2009) and independent component analysis (ICA)-based approaches (van de Ven et al., 2004; Beckmann et al., 2005) are the most commonly used analysis methods in resting-state

* Corresponding authors.
E-mail addresses: yong@kaist.ac.kr (Y. Jeong), jong.ye@kaist.ac.kr (J.C. Ye).

[2] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://www.loni.usc.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators include (complete listing available at http://www.loni.usc.edu/ADNI/Collaboration/ADNIAuthorshiplist.pdf).

http://dx.doi.org/10.1016/j.neuroimage.2015.10.081
1053-8119/

functional connectivity studies. The seed-based approach extracts BOLD signal time courses from a region of interest (ROI), called a "seed" region, and computes the cross-correlation between time course signals from the ROI and all other voxels in the brain to obtain a map of neuronal connectivity (Fox and Raichle, 2007). Despite their popularity, seed-based correlation analyses have limitations such that they require a prior determination of the seed's location. On the other hand, ICA automatically decomposes the entire BOLD dataset into maximally independent components. However, the brain networks are not independent of each other due to their complex, interconnected regions. Another issue in using ICA is that the individual analysis is usually not sensitive in detecting networks compared to seed-based analysis. Moreover, the unified theory that links the individual analysis results to group analysis is still not fully established. Additionally, graph theory-based quantitative analyses of brain connectivity have been developed to study structural and functional brain networks and their interactions (Bullmore and Sporns, 2009). However, graph theory-based analysis is dependent on pre-defined parcellations. Therefore, parcellation-independent graph theoretical analyses are required.

Unlike the conventional approaches, here we present a novel parcellation-free functional connectivity analysis that is inspired by the graph theoretical approach for brain networks. More specifically, our method is derived from signal decomposition based on a sparse graph model that regards the temporal dynamics at each voxel as a *sparse* combination of *unknown* global information flow. Interestingly, we can show that the sparse dictionary learning algorithm and the concept of a spatially adaptive design matrix used for our fMRI analysis in Lee et al. (2011b) can be used to represent local connectivities based on the sparse graph model. However, one of the technical difficulties of using Lee et al. (2011b) for functional connectivity fMRI analysis is that the extracted temporal dynamics corresponding to each network highly depend on the individual. Moreover, subject-dependent regressors should be estimated, after which the group-level statistical inferences should be performed using group average activation maps that are extracted using the subject-specific regressors. This complicates the group sparse learning and statistical inference. Similar difficulties have been observed in other data-driven approaches, such as ICA. In group ICA, the problem has been addressed by concatenating the data or by using tensor factorization. However, even though group-wise activation maps can be detected using these types of approaches, more advanced group analyses, such as a two-sample *t*-test, or an analysis of variance (ANOVA), are often difficult. There are some recent methods for ICA to obtain such components, such as dual regression (Zuo et al., 2010), and GRAICAR (generalized ranking and averaging independent component analysis by reproducibility) (Yang et al., 2012). However, a unified framework from individual to group level using standard statistical analysis tools still appears to be lacking.

To overcome such technical difficulties in group analysis, one of the main contributions of this paper is to propose a novel *unified mixed-effect model* framework where group-level sparse dictionary learning and group inference can be performed in a unified linear mixed model and the restricted maximum likelihood (ReML) variance estimation framework. More specifically, to estimate the unknown global dynamics and local network structures at a group level, we first concatenated the time series across the subject and performed a group sparse dictionary learning for the concatenated temporal data. We showed that the sparse learning for the concatenated time series is equivalent to imposing a constraint that the network structures within a group are similar. Using this constraint, a global dictionary was estimated from the concatenated data, after which the dictionaries from the concatenated time series were separated to obtain each subject-level sparse dictionary. Then, the SPM-type analysis was performed using individualized dictionaries. Under the homoscedasticity variance assumption, we showed that the aforementioned group sparse dictionary learning and inference can be rigorously derived using the unified linear mixed model framework and the restricted maximum likelihood (ReML) variance estimation.

As the mathematical framework for inference turns out to be similar to that of a standard statistical parameter mapping (SPM) analysis with only the exception of a *spatially* adaptive design matrix (which still retains the homogeneous degree of the freedom), rich statistical analysis tools, such as *p*-value correction using random field theory and hypothesis-driven inference, can be used. Accordingly, we call the proposed method as *sparse SPM* (*SSPM*).

To confirm the validity of the proposed method, we provide extensive comparisons using group data from normal, MCI, and Alzheimer subjects from both our clinical data set and the ADNI (Alzheimer's Disease Neuroimaging Initiative) data set (http://www.loni.usc.edu/ADNI).

Theory

Throughout the paper, $\mathbf{x}^i$ and $\mathbf{x}_j$ correspond to the i-th row and the j-th column of matrix X, respectively. When S is an index set, X^S and A_S correspond to a submatrix collecting the corresponding rows of X and columns of A, respectively; $\mathbf{x}_S$ denotes a subvector collecting the corresponding elements of X. The superscripts ' and † denote the adjoint operator and pseudo-inverse, respectively. A vector $\mathbf{1}_L$ denotes a L-dimensional vector with elements of ones, and $I_{k\times k}$ is a $k\times k$ identity matrix.

Sparse graph model for functional connectivity analysis

In the proposed method, the interactions of neural signal between the brain's functional systems are modeled by a set of nodes (voxels) linked by connections as shown in Fig. 1. Specifically, each circle denotes a voxel or node where a temporal dynamic of BOLD signal is measured continuously. A set of nodes in a functional brain network is defined as a community if they share the same information flows.

In Fig. 1, the sparsity level (which determines the variety of information flows) is $k = 3$. For example, a length m vectorized time series at the nodes 1 and 2, denoted by $\mathbf{y}_1 \in \mathbb{R}^m$ and $\mathbf{y}_2 \in \mathbb{R}^m$, is given by

$$\mathbf{y}_1 = 2\mathbf{d}_1 + \mathbf{d}_3 + \mathbf{d}_4, \quad \mathbf{y}_2 = 3\mathbf{d}_2 + \mathbf{d}_4 + \mathbf{d}_5 ,$$

which shows that the information at the nodes 1 and 2 is composed of three distinct information flows, i.e. $k = 3$. Moreover, in Fig. 1, the dotted lines represent information flows other than $d_1, \cdots, d_5$ so that at each node the time trace is composed of $k = 3$ distinct information flows. If we define a global dictionary D by collecting all local or long-range information flows as

$$D = \left[\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_5, \mathbf{d}_6 \cdots, \mathbf{d}_Q\right] ,$$

where $\mathbf{d}_6, \cdots, \mathbf{d}_Q$ denote the information flow on the dotted lines in Fig. 1, then we can easily see that the temporal dynamics at nodes 1 and 2 are described as a *sparse* linear combination of the atoms in the global design matrix. In general, we have

$$\mathbf{y}_n = D_{I_n}\boldsymbol{\alpha}_{I_n} + \boldsymbol{\varepsilon}_n,\ n = 1, \cdots, N \tag{1}$$

where N is the number of voxels, ϵ denotes noise, $D_{I_n} \in \mathbb{R}^{m\times k}$ is a submatrix of $D \in \mathbb{R}^{m\times Q}$ composed of elements in the index set I_n, and $\boldsymbol{\alpha}_{I_n} \in \mathbb{R}^k$ denotes the corresponding weight vectors. Note that a local subset index I_n represents a local network structure at the n-th voxel, and a regressor is a representative dynamic in a network module or community that shares the same information flow. We are aware that model (1) or its variations are frequently used. Structural equation modeling (SEM) (Schlösser et al., 2003), and sparse partial correlations (Ryali et al., 2012), among many others, all use linear regression models between nodes. The only differences in all of these linear models are the additional constraints put into the regression. The main contribution of our model (1) over the existing ones is that the *spatially adaptive* design matrix is used by selecting a subset of regressors from a dictionary

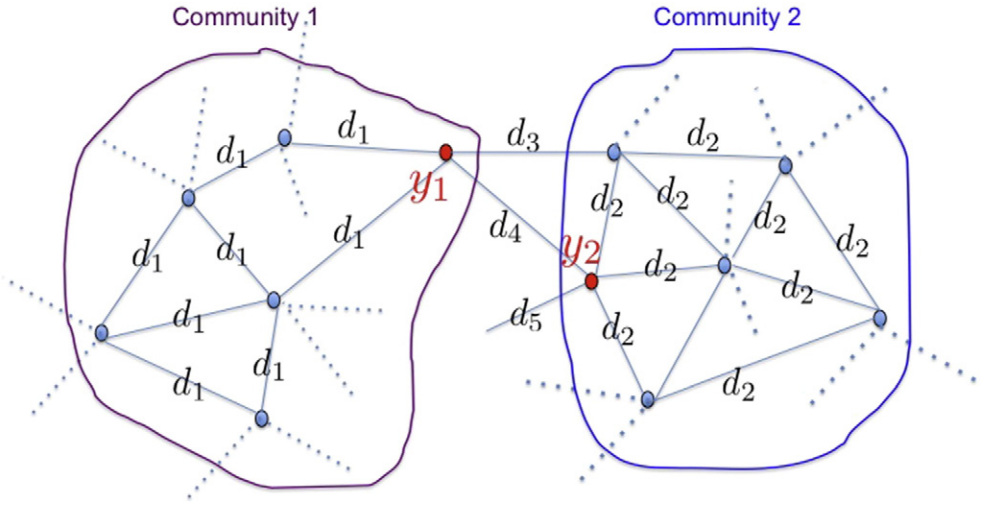

Fig. 1. A graph theoretical model for sparse SPM.

across all brain areas so that the subset index represents the topology of local connectivity.

In Fig. 1, it is important to note that all the voxels, which are connected through k distinct information flows, share the information flow from other communities. A community that a voxel belongs to is determined by statistical testing. Specifically, to extract a *statistically significant* community that shares the same information flow, a hypothesis test is conducted to test whether a specific information flow of interest ($\mathbf{z} \in D$) is presented in each voxel. This test can be formulated as follows: a null hypothesis $H_0 : \theta_n = 0$ is tested against an alternative hypothesis $H_1 : \theta_n \neq 0$, where θ_n is a regression coefficient or an *effect* for the regressor $\mathbf{z}$:

$$\mathbf{y}_n = \mathbf{z}\theta_n + D_{I_n \backslash \mathbf{z}}\boldsymbol{\alpha}_{n\,\backslash \mathbf{z}} + \boldsymbol{\varepsilon}_n \tag{2}$$

where $\backslash\mathbf{z}$ denotes a reduced size matrix or vector made by removing the elements corresponding to the atom $\mathbf{z}$. When $\mathbf{z}$ belongs to the set of atoms from D_{I_n}, we have the following F statistics:

$$\mathcal{F}_n = \frac{\mathbf{y}_n'\left(P^{\perp}_{D_{I_n \backslash \mathbf{z}}} - P^{\perp}_{D_{I_n}}\right)\mathbf{y}_n}{\mathbf{y}_n' P^{\perp}_{D_{I_n}} \mathbf{y}_n} \frac{m-k}{1} \tag{3}$$

where $P^{\perp}_{D_{I_n}}$ and $P^{\perp}_{D_{I_n \backslash \mathbf{z}}}$ denote the projection on the orthogonal complement on the range space of D_n and $D_{I_n \backslash \mathbf{z}}$, respectively. Note that the degree of freedom (df) for (3) is 1, because we test the effect for a single regressor. Hence, this is equivalent to the two-side t-test. The reason we still prefer to call this as F test with $df = 1$ is to extend this later for a group-level F test for ANOVA.

If $\mathbf{z}$ is not presented in the local dictionary D_{I_n}, then $\theta_n = 0$ and the signal at the voxel are irrelevant to the network associated with the atom $\mathbf{z}$, so H_0 holds and $F_n = 0$. For example, in Fig. 1, if we test the presence of $\mathbf{z} = \mathbf{d}_1$ (or $\mathbf{z} = \mathbf{d}_2$), we can obtain the nodes in community 1 (or community 2). By testing $\mathbf{z} = \mathbf{d}_3$ or $\mathbf{z} = \mathbf{d}_4$, the connector hubs between two community can be obtained, which belong to two communities at the same time. This implies that a voxel can belong to multiple communities as long as the memberships for those communities are statistically significant.

For a group of L subjects, we perform a similar connectivity analysis using the concatenated times series from multiple subjects. More specifically, if $Y^{(l)} = [\mathbf{y}_1^{(l)}, \cdots, \mathbf{y}_N^{(l)}]$ denotes a collection of time traces for the l-th subject across all N voxels, then we construct a concatenated temporal time trace from L subjects across all voxels $Y = [Y^{(1)'}, \cdots, Y^{(L)'}]'$. Assuming that the local design matrix index I_n is same across all subjects, we have the following decomposition of Y:

$$Y \equiv \begin{bmatrix} Y^{(1)} \\ \vdots \\ Y^{(L)} \end{bmatrix} = DW = \begin{bmatrix} D^{(1)} \\ \vdots \\ D^{(L)} \end{bmatrix} [\mathbf{w}_1, \cdots, \mathbf{w}_N] \tag{4}$$

where D denotes the concatenated global dictionary, $D^{(l)}$ denotes the corresponding l-th subject individual sparse dictionary, and W is the corresponding coefficients whose non-zero elements are sparse. In our model, the individual dictionary $D^{(l)}$ is assumed to be normalized so that each column has unit norm.

Now, the i-th column of D represents the representative temporal dynamics at the voxel locations that have non-zero coefficients in the i-th row of W. More specifically, the i-th column of the resulting individual dictionary $\{D^{(l)}\}_{l=1}^{L}$ shares the same geometric connectivity. Hence, our assumption that the local design matrix index I_n is same across all subjects imposes the constraint that the *local network structure within a group is the same*. This learning rule is very advantageous to identify the group differences, because the network connectivity changes between groups are believed to be one of the main biomarkers in resting-state fMRI analysis.

ReML estimation of group-wise sparse graph

Mathematically, the temporal dynamics at the n voxel of the subject l can be modeled as

$$\mathbf{y}_n^{(l)} = D_{I_n}^{(l)}\mathbf{w}_{I_n}^{(l)} + \boldsymbol{\epsilon}_n^{(l)}, \quad \boldsymbol{\epsilon}_n^{(l)} \sim \mathcal{N}\left(\mathbf{0}, R_n^{(l)}\right) \tag{5}$$

where $l = 1, \cdots, L$ denote the subject index and $R_n^{(l)}$ denotes the l-th subject temporal correlation matrix at the n voxel. The subject differences within a group can now be modeled as random-effects:

$$\mathbf{w}_{I_n}^{(l)} = \boldsymbol{\alpha}_n + \boldsymbol{\beta}_n^{(l)}, \; \boldsymbol{\beta}_n^{(l)} \sim N(0, G_n), \tag{6}$$

where $\boldsymbol{\alpha}_n$ denote a group mean and G_n is a random-effect covariance matrix. Then, for the concatenated individual data, we can obtain the following mixed model:

$$\mathbf{y}_n = X_n\boldsymbol{\alpha}_n + Z_n\boldsymbol{\beta}_n + \boldsymbol{\epsilon}_n\,, \tag{7}$$

where

$$\mathbf{y}_n = \left[\mathbf{y}_n^{(1)'}, \cdots, \mathbf{y}_n^{(L)'}\right]' \in \mathbb{R}^M$$
$$\boldsymbol{\beta}_n = \left[\boldsymbol{\beta}_n^{(1)'}, \cdots, \boldsymbol{\beta}_n^{(L)'}\right]' \in \mathbb{R}^q$$

where $M = mL$, $q = kL$, and the fixed-effect and random-effect matrix X_n and Z_n are given by

$$X_n = \begin{bmatrix} D_{I_n}^{(1)} \\ D_{I_n}^{(2)} \\ \vdots \\ D_{I_n}^{(L)} \end{bmatrix}, \; Z_n = \begin{bmatrix} D_{I_n}^{(1)} & 0 & \cdots & 0 \\ 0 & D_{I_n}^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & ND_{I_n}^{(L)} \end{bmatrix},$$

or equivalently,

$$X_n = Z_n X_G, \quad X_G = 1_L \otimes I_{k\times k},$$

where $\otimes$ represents the Kronecker product.

Note that in our sparse graph estimation problem, there are several unknowns that need to be estimated, including the global dictionary D, local network connectivity $\{I_n\}_{n=1}^N$, fixed and random-effect regression coefficients $\{\boldsymbol{\alpha}_n, \boldsymbol{\beta}_n\}_{n=1}^N$, and covariance matrices $\{R_n^{(l)}\}_{n,l=1}^{N,L}$ and $\{G_n\}_{n=1}^N$. The maximum likelihood (ML) principle is perhaps the most popular parameter estimation procedure in statistics because of its many optimal properties, such as consistency (Searle, 1979). However, one of the main problems of the ML approach for covariance estimation is that it does not take into account the loss in degrees of freedom resulting from the estimation of the model's fixed-effects (Searle, 1979; Patterson and Thompson, 1971). To address this issue, Patterson and Thompson (1971) proposed a method that takes into account the loss in degrees of freedom resulting from estimating fixed-effects. More specifically, they restricted their attention to a set of such contrasts that are invariant to the fixed-effect parameter and then estimated the covariance for such *restricted* cases of maximum likelihood. This idea is known as the *restricted maximum likelihood* (*ReML*) method (Searle, 1979; Harville, 1977; Kenward and Roger, 1997; Graser et al., 1987; Friston et al., 2011). More specifically, if we are interested in finding the restricted maximum likelihood that is invariant to the fixed-effect parameter, we need to minimize the following form of the ReML equation (Searle, 1979; SAS Institute, 1985):

$$C_{ReML}\left(\left\{R_n^{(l)}\right\}, \{G_n\}, \{I_n\}, \{\boldsymbol{\alpha}_n\}\right) \tag{8}$$
$$= -\sum_n \left(\frac{1}{2}\log|V_n| + \frac{1}{2}\log|X_n' V_n^{-1} X_n| + \frac{1}{2}(\mathbf{y}_n - X_n\boldsymbol{\alpha}_n)' V_n^{-1}(\mathbf{y}_n - X_n\boldsymbol{\alpha}_n)\right),$$

where V_n is given by

$$V_n = Z_n G_n Z_n' + R_n, \tag{9}$$

$$R_n = \begin{bmatrix} R_n^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R_n^{(L)} \end{bmatrix} \tag{10}$$

However, ReML is usually computationally expensive. To simplify the covariance estimation and the resulting inference, we first approximate that $D_{I_n}^{(l)'} D_{I_n}^{(l)} \simeq I$. Because each individual dictionary is normalized to have unit norm, this implies that each regressor is nearly orthogonal, which is often used in SPM analysis. Next, we assume that the noise in individual temporal dynamics is white, i.e. $R_n^{(l)} = \sigma_n^{(l)2} \simeq I$, and the random-effect variance is spherical, i.e. $G_n = g_n^2 I$. This can be easily satisfied using a prewhitening procedure. Finally, we employ the homoscedasticity variance assumption across subjects and groups, i.e. $\sigma_n^{(l)2} = \sigma_n^2$, which has also been used frequently in SPM. Then, the resulting cost function becomes

$$C_{ReML}(\{\sigma_n^2\}, \{g_n^2\}, \{I_n\}, \{\boldsymbol{\alpha}_n\})$$
$$= -\sum_n \left(\frac{L(M-k)}{2}\ln(g_n^2 + \sigma_n^2) + \frac{1}{2(g_n^2+\sigma_n^2)}\|\mathbf{y}_n - X_n\boldsymbol{\alpha}_n\|^2\right).$$

The problem can now be solved using an alternating minimization. More specifically, for the estimated local connectivity I_n and the fixed-effect parameter estimation $\hat{\boldsymbol{\alpha}}_n$, we have

$$\hat{g}_n^2 + \hat{\sigma}_n^2 = \frac{1}{L(M-K)}\|\mathbf{y}_n - X_n\hat{\boldsymbol{\alpha}}_n\|^2$$
$$= \frac{1}{L(M-k)}\sum_{l=1}^{L}\left\|\mathbf{y}_n^{(l)} - D_{I_n}^{(l)}\hat{\boldsymbol{\alpha}}_n\right\|^2 \tag{11}$$

Next, for a given sparsity level k for the maximum number of local connections, we claim that the estimation problem of local connectivity I_n, global dictionary D, and the fixed-effect parameter estimation $\hat{\boldsymbol{\alpha}}_n$ can be equivalently written as a minimization problem with respect to D, W

$$\min_{D,W}\|Y - DW\|_F, \text{ subject to } \|\mathbf{w}_n\|_0 \le k, \; for \; all \; n. \tag{12}$$

To see the equivalence, we consider a sparse regression problem for a given global dictionary D:

$$\min_{\mathbf{w}_n}\|\mathbf{y}_n - D\mathbf{w}_n\|_2^2, \text{ subject to } \|\mathbf{w}_n\|_0 \le k. \tag{13}$$

which is equivalent to

$$\min_{I_n, \boldsymbol{\alpha}_n}\|\mathbf{y}_n - D_{I_n}\boldsymbol{\alpha}_n\|_2^2, \text{ subject to } |I_n| \le k. \tag{14}$$

Therefore, by solving (13) and obtaining the non-zero index of $\hat{\mathbf{w}}_n$ as I_n, we can obtain D_{I_n} and $\hat{\boldsymbol{\alpha}}_n$.

The simultaneous estimation of D and W in (12) can be addressed by using a sparse dictionary learning algorithm called K-SVD (Aharon et al., 2006), which attempts to find the best possible dictionary D for a sparse representation of Y in a greedy manner using a sparse coding step and dictionary update step. More specifically, for a given dictionary D, the sparse coding solves (13) for $n = 1, \cdots, N$. The active index set I_n is then estimated by collecting indices corresponding to the k-largest correlation between $\mathbf{y}_n$ and $\mathbf{d}_j$ as:

$$C_{\mathbf{y}_n}(j) = \frac{\|\mathbf{y}_n'\mathbf{d}_j\|_2^2}{\|\mathbf{d}_j\|_2^2}, \quad j = 1, \cdots, Q, \tag{15}$$

where Q denotes the number of atoms in a global dictionary D. The fixed-effect parameter estimate $\hat{\boldsymbol{\alpha}}_n$ is then given by

$$\hat{\boldsymbol{\alpha}}_n = \left(D_{I_n}' D_{I_n}\right)^{-1} D_{I_n}'\mathbf{y}_n. \tag{16}$$

The dictionary update step then refines each column vector $\mathbf{d}_j (j = 1, \ldots, Q)$ and the corresponding coefficient row vector $\mathbf{w}^j$ sequentially by a rank-1 approximation using singular value decomposition. More specifically, with estimated W, K-SVD puts into question only one column in the dictionary, $\mathbf{d}_j$, and the corresponding coefficient $\tilde{\mathbf{w}}^j$, the j-th row of $\tilde{W}$. This can be solved using singular value decomposition (SVD) with sparsity constraint. Then, for each $j = 1, 2, \cdots, Q$, K-SVD does the following: (*i*) defines the index set ω^j corresponding to non-zero indices of $\tilde{\mathbf{w}}^j$; (*ii*) computes $E_j = Y - \sum_{p\neq j}\mathbf{d}_j\tilde{\mathbf{w}}^j$; (*iii*) defines Ω_j as a diagonal matrix with ones for the indices corresponding to ω^j and

zeros elsewhere; (*iv*) chooses a subset $E_j^{\mathscr{R}} = E_j\Omega_j$; (*v*) takes SVD to the restricted $E_j^{\mathscr{R}}$,

$$E_j^{\mathscr{R}} = U\Lambda V' = \sum_{p=1}^{p} \sigma_p \mathbf{u}_p \mathbf{v}_p' \tag{17}$$

and (*vi*) updates $\hat{\mathbf{d}}_j = \mathbf{u}_1$, $\hat{\mathbf{w}}^j_{\mathscr{R}} = \sigma_1 \mathbf{v}_1'$.

Note that the K-SVD procedure has very interesting interpretations for resting-state fMRI analysis. First, as shown in Fig. 2, an update stage for an atom $\hat{\mathbf{d}}_j$ in the dictionary update stage of K-SVD is equivalent to finding the largest principal component from a set of raw data $\{\mathbf{y}_n\}_{n\in\omega^j}$, where ω^j denotes the voxel indices that have non-zero coefficients from $\tilde{\mathbf{w}}^j$. More specifically, ω^j denotes the voxel indices for a community that shares the dynamics $\hat{\mathbf{d}}_j$ as shown in Fig. 2. Hence, the dictionary update stage is basically to update a representative temporal dynamic that constitutes a specific network structure by removing noise using principal component analysis. Hence, this corresponds to a *filtering* procedure.

Second, because a sparse coding step estimates the non-zero coefficients of the regression coefficients, it updates the local network structures by which each voxel is connected to other voxels in the brain. For example, as shown in Fig. 2, the resulting non-zero voxels that correspond to the $\mathbf{d}_1$ constitute a local community that shares the same temporal dynamic $\mathbf{d}_1$. Therefore, this procedure corresponds to a clustering procedure that identifies the voxels with similar dynamics. In particular, for a sparse coding stage, if we use the correlation (15) to update I_n, then the clustering is based on the correlation with respect to raw temporal dynamics $\mathbf{y}_n$ and the updated local information flow $\mathbf{d}_j$, which is very similar to seed-based analysis. Hence, K-SVD sparse coding with (15) can be considered *seed-based clustering*. However, the main difference between the conventional seed-based connectivity analysis and the proposed method is that in the proposed method, the seed $\mathbf{d}_j$ is adaptively updated and denoised during the dictionary estimation stage, which is not the case in the conventional approach.

Therefore, such repeated applications of filtering and adaptive seed-based clustering procedures allow our method to take advantage of both the conventional seed-based approach and the ICA type of clustering approaches which makes the proposed algorithm very powerful.

Inference on group differences

Suppose we are interested in comparing two different groups. Once the group sparse dictionary learning is performed for each set of group data, the temporal dynamics at the n voxel of the subject l in the i-th group model are modeled as

$$\mathbf{y}_n^{(il)} = D^{(il)}_{I_n^{(i)}} \mathbf{x}^{(il)}_{I_n^{(i)}} + \boldsymbol{\epsilon}_n^{(il)}, \quad \boldsymbol{\epsilon}_n^{(il)} \sim \mathcal{N}\left(\mathbf{0}, R_n^{(il)}\right) \tag{18}$$

where $i = 1, 2$ and $l = 1, \cdots, L_i$ denote the group and subject index for each group, respectively, and L_i is the number of subjects for the group i. The subject differences within a group can now be modeled as random-effects:

$$\mathbf{x}^{(il)}_{I_n} = \boldsymbol{\alpha}_n^{(i)} + \boldsymbol{\beta}_n^{(il)}, \quad \boldsymbol{\beta}_n^{(il)} \sim \mathcal{N}(\mathbf{0}, G_n), \tag{19}$$

where $\boldsymbol{\alpha}_n^{(i)}$ denotes a group mean. Therefore, if we stack the data together (after ignoring the voxel-dependent index n for simplicity), we have the following mixed model:

$$\mathbf{y} = X\boldsymbol{\alpha} + Z\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{20}$$

where

$$\mathbf{y} = \left[\mathbf{y}_n^{(11)'}, \mathbf{y}_n^{(12)'}, \cdots, \mathbf{y}_n^{(1L_1)'}, \cdots, \mathbf{y}_n^{(2L_2)'}\right]' \in \mathbb{R}^M$$
$$\boldsymbol{\alpha} = \left[\boldsymbol{\alpha}^{(1)'}, \boldsymbol{\alpha}^{(2)'}\right]' \in \mathbb{R}^p$$
$$\boldsymbol{\beta} = \left[\boldsymbol{\beta}_n^{(11)'}, \boldsymbol{\beta}_n^{(12)'}, \cdots, \boldsymbol{\beta}_n^{(1L_1)'}, \cdots, \boldsymbol{\beta}_n^{(2L_2)'}\right]' \in \mathbb{R}^q$$

where $M = m\sum_i L_i$ and $q = k\sum_i L_i$ and the random-effect matrix Z is given by

$$Z = \begin{bmatrix} D^{(11)}_{I_n^{(1)}} & 0 & \cdots & 0 \\ 0 & D^{(12)}_{I_n^{(1)}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D^{(2L_2)}_{I_n^{(2)}} \end{bmatrix},$$

and the fixed-effect matrix is

$$X = ZX_G, \quad X_G = \begin{bmatrix} 1_{L_1} \otimes I & 0 \\ 0 & 1_{L_2} \otimes I \end{bmatrix}.$$

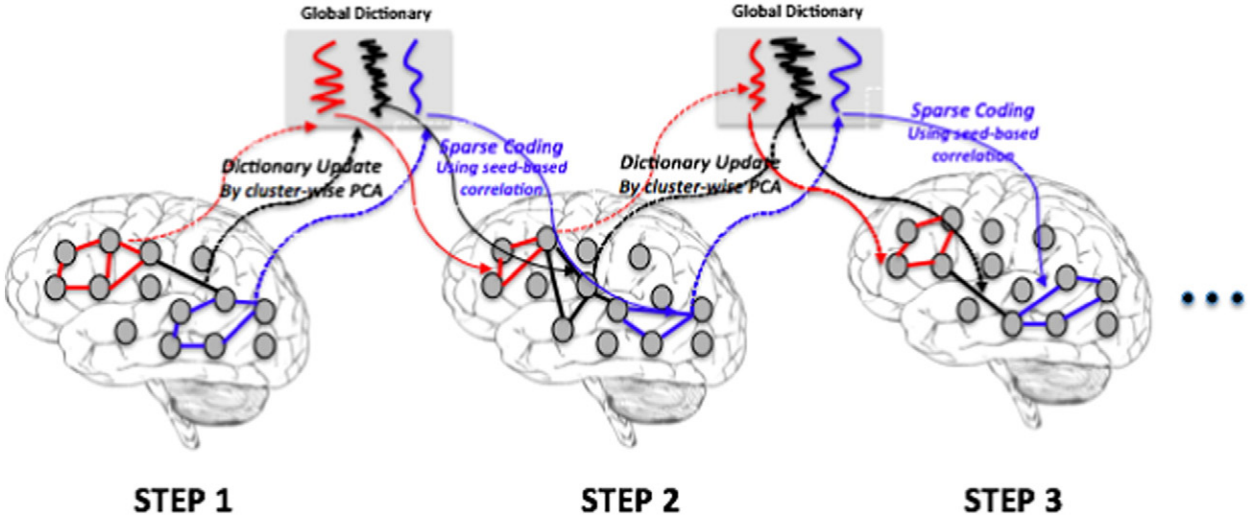

Fig. 2. Dictionary update step: for a given set of voxels within a community, the corresponding dictionary atoms are updated as a principal component analysis of the temporal dynamics. Sparse coding step: for each dictionary atom that represents the common temporal dynamics, the most correlated voxels are statistically tested to define their membership for each community.

Then, using the standard results for the mixed-effect model, the fixed-effect parameter estimate $\hat{\boldsymbol{\alpha}}$ from the mixed model can be equivalently represented as the second-level GLM using the summary statistics

$$\hat{\boldsymbol{\chi}} = X_G\boldsymbol{\alpha} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, V_G) \tag{21}$$

where $\hat{\boldsymbol{\chi}} = [\hat{\mathbf{w}}_{I_n^{(1)}}^{(11)'}, \hat{\mathbf{w}}_{I_n^{(1)}}^{(12)'}, \cdots, \hat{\mathbf{w}}_{I_n^{(1)}}^{(1L_1)'}, \cdots, \hat{\mathbf{w}}_{I_n^{(2)}}^{(2L_2)'}]'$ and

$$\hat{\mathbf{w}}_{I_n^{(i)}}^{(il)} = \left(D_{I_n^{(i)}}^{(il)'} R^{-(il)} D_{I_n^{(i)}}^{(il)}\right)^{-1} D_{I_n^{(i)}}^{(il)'} R^{-(il)} \mathbf{y}_n^{(il)} \tag{22}$$

Here, for $G_n = g_n^2 I$, V_G is a block diagonal matrix whose (il)-th block is composed of

$$\left(D_{I_n^{(i)}}^{(il)T} R^{-(il)} D_{I_n^{(i)}}^{(il)}\right)^{-1} + g_n^2 I. \tag{23}$$

For the group inference on group mean activation or differences, we are interested in testing the following null hypothesis:

$$H_0 : C\boldsymbol{\alpha} = \mathbf{0},$$

where $C \in \mathbb{R}^{p_1 \times k}$ denotes the contrast matrix. Since $(X^T V^{-1} X)^{-1} = (X_G^T V_G^{-1} X_G)^{-1}$ as shown in Appendix A, the test statistics for the mixed model are equivalent to the second-level inference statistics:

$$S = = \frac{\hat{\boldsymbol{\alpha}}' C' \left(C\left(X'\hat{V}^{-1}X\right)^{-1} C'\right)^{-1} C\hat{\boldsymbol{\alpha}}}{p_1} = \frac{\hat{\boldsymbol{\alpha}}' C' \left(C\left(X_G'\hat{V}_G^{-1}X_G\right)^{-1} C'\right)^{-1} C\hat{\boldsymbol{\alpha}}}{p_1}.$$

Under the aforementioned assumptions of homoscedasticity variances and orthogonal regressors, we can now show that $\hat{V}_G$ is a diagonal matrix. Therefore, using the equivalence relationship in Appendix B, we can show that

$$\mathcal{S} = \frac{\hat{\boldsymbol{\chi}}'\left(P_{X_{G,0}}^{\perp} - P_{X_G}^{\perp}\right)\hat{\boldsymbol{\chi}}}{\hat{\boldsymbol{\chi}}' P_{X_G}^{\perp} \hat{\boldsymbol{\chi}}} \frac{M-k}{p_1} \sim F_{p_1,u}$$

where $X_{G,0}$ denotes the reduced model by excluding the effect estimated by contrast C, $u = m\sum_{il} L_{il} - k$ and $\hat{\boldsymbol{\chi}}$ are summary statistics in (21).

Because this F statistics are standard statistics for ANOVA, the results indicate that we can perform classical ANOVA using the summary statistics, and such analysis is equivalent to the inference in the mixed model as long as our assumption holds. Moreover, we do not need to perform a computationally expensive ReML covariance estimation, because the parts of the ReML variance estimation are already built in within the resulting F statistics. Note that the degree of freedom is same across all voxels, so it is easy to apply a random field-based p-value correction (Friston et al., 2011) for the resulting F-maps, which is another advantage of the proposed method compared to ICA approach, which addresses family-wise error correction in a heuristic way.

Method

Data acquisition

We collected three groups of resting-state fMRI data: 1) 22 normal subjects (8 male, mean age 70 years), 2) 37 MCI patients (21 male, mean age 72 years), and 3) 20 AD patients with CDR 0.5 (5 male, mean age 72 years). During the task period, subjects were instructed to be awake and alert but not actively involved in a task and with eye closed. A 3.0 T fMRI system (Philips, Netherlands) was used to measure the BOLD response. The echo planar imaging (EPI) sequence was used with TR/TE = 3000/35 *ms*, flip angle = 90°. Each acquisition consisted of 35 continuous slices, FOV (RL, AP, FH) = 220 × 140 × 220 mm, and a voxel size of (RL, AP) = 2.875 × 2.875 mm. In the subsequent anatomical scanning session, T1-weighted structural images were acquired. A total of 100 acquisitions are obtained for each subject, and the total recording time was 300 sec. The experiments were approved by the Institutional Review Board of the Samsung Medical Center in South Korea.

We also used the data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://www.loni.usc.edu/ADNI). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations as a $60 million, 5-year public/private partnership.

Data analysis using conventional methods

We used two conventional methods for resting-state fMRI analysis and compared them with the proposed method. First, a multi-session temporal concatenation of the Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC v3.0) within FMRIB's Software Library (FSL) was used as an ICA method. We began with a brain extraction tool (BET v2.1) for the anatomy data with a fractional intensity threshold of 0.5 and the option of "robust brain center estimation (iterates BET2 several times)" to obtain brain images from the whole anatomy image. The following parameters were applied in the MELODIC analysis: 128 sec for the high pass filter cutoff, motion correction, smoothing using a Gaussian kernel of FWHM 8 mm, normalization into MNI coordinates with a resampling resolution of 2 mm, variance-normalized timecourses, and 20 independent components for the dictionary regressor to make the analysis conditions the same as those in our proposed method. Because MELODIC provides multiple subnetworks, we manually chose the DMN with visual inspection.

Second, we used the functional connectivity toolbox (CONN) based on statistical parametric mapping (SPM) for seed-based analysis. In the CONN toolbox, there are pre-defined ROIs, including the posterior cingulate cortex (PCC) region. The coordinates for the PCC region are provided by the CONN tool, so we used this to extract the PCC seed region. The region posterior cingulate cortex (PCC) was used for the ROI in the seed-based analysis, where the following procedures were applied: realignment, segmentation, normalization, smoothing using a Gaussian kernel of FWHM 8 mm, band pass filter with a cutoff frequency of 0.008–0.09 Hz, and threshold p-value 0.001.

Data analysis using sparse SPM

The images were first spatially realigned to remove movement artifacts in the fMRI time series. The images were then spatially normalized to a Montreal Neurological Institute (MNI) standard space and resampled with voxel size 2 × 2 × 2 mm. Spatial smoothing was then applied with full-width at half-maximum (FWHM) Gaussian kernel size 8 × 8 × 8 mm. The brain regions of functional data were extracted using a segmented anatomy data as a mask image with respect to gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). To filter the temporal time series at each voxel, we used a discrete cosine transform (DCT) filter with a cutoff frequency of 1/128 Hz, an appropriate range of frequency for resting-state data, which show low-frequency oscillations in the range of 0.0–0.1 Hz. From each time series, temporal DC components were removed. These preprocessing steps were conducted using SPM8.

For group analysis, 3-D coordinates for the segmented GM, WM, and CSF from each set of individual data were compared, and the corresponding voxel coordinates common across all subjects in a group

were extracted to build reference masks for group maps of GM, WM, and CSF, respectively (see Fig. 3 for an example of training data from GM data). The time series that corresponded to the group GM maps were then collected from each set of individual data and concatenated together to build a group training data set for sparse dictionary learning. Considering that neural activation occurs mostly on gray matter that contains neural cell bodies, gray matter is the basis of all brain functions. Therefore, we focused on the gray matter rather than white matter and CSF. In particular, the aging procedure is associated with general white and gray matter loss. Consequently, if we use data from an aging group, it can increase CSF partial volume artifacts, which can cause problems in regions prone to CSF contamination, such as the genu of corpus callosum that pass through or close to the ventricles. Accordingly, training using GM can extract the real neural dynamics rather than other physiological noise. To reduce the computational burden of a sparse dictionary learning algorithm, rather than initializing the dictionary using all of the data, we choose a subset of the voxels by downsampling by a factor of 4.

Once the group sparse dictionary learning was completed for each group, we used the estimated global dictionaries to identify the local network structures across the whole voxels in all brain areas. This was an additional necessary step because the group sparse dictionary learning was performed using the downsampled data set to reduce the computational complexity, but we needed to identify the network structures across all voxels in the brain. This step was performed as follows. First, we identified the group-wise coordinates for the brain area and then applied one more sparse coding step using the estimated global dictionary, as shown in Fig. 3. This provided a spatially varying design matrix $\{D_{I_n}\}_{n=1}^{N}$. Then, the spatially varying design matrix $\{D_{I_n}\}$ was decomposed into an individual design matrix $\{D^{(il)}_{I^{(i)}_n}\}_{n,i,l}$, whose regressors were normalized to have the same magnitude value of 1. The final regression coefficients were then estimated by using the least square method with respect to the design matrices. The set of individual summary statistics were then combined to obtain a group activation map. These procedures are summarized in Fig. 4.

Experimental results

Parameter selection

The determination of the number of atoms in a dictionary is an important issue, because it represents the number of linearly independent temporal dynamics across the whole brain. Group results with various dictionary numbers are shown in Fig. 5. In general, when a small number of dictionary components are chosen, they tend to aggregate different networks. On the other hand, a large dictionary size tends to segregate a subnetwork into multiple fragments. We tested various dictionary component numbers. As shown in Fig. 5, with fifteen dictionary components, there were noises in the mid-brain regions. Specifically, we could also see the activation in the region of pons even though this region is not the main ROI of the DMN. In addition, the frontal region of the DMN could not be identified with twenty-five dictionary components. Thus, we concluded that twenty was the optimal number of dictionary components to extract.

Another important hyper-parameter that we needed to estimate for our method was the sparsity level $k = |I_n|$. In our previous work for individual analysis (Lee et al., 2011b), we proposed the minimum description length (MDL) as a criterion in deciding the sparsity level. In general, a higher sparsity level provides more fragmented DMN structures, so the different sparsity level for each group may provide different topological structures of the DMN network. However, in group analysis, we found that imposing a fixed sparsity level is preferable to retain the same subnetwork structures. Based on empirical results, we decided to use the sparsity level $k = 3$ across all groups because it provided the least fragmented DMN across all groups.

Because the data across subjects were concatenated for group analysis, we verified whether SSPM was sensitive to the concatenation order. We processed SSPM with different orders, but the results were identical (results not shown). This is as expected because the dictionary learning step using K-SVD is invariant to the permutation of the data set.

A pre-defined number of five iterations was used in the dictionary learning step. We also applied different numbers of iteration steps

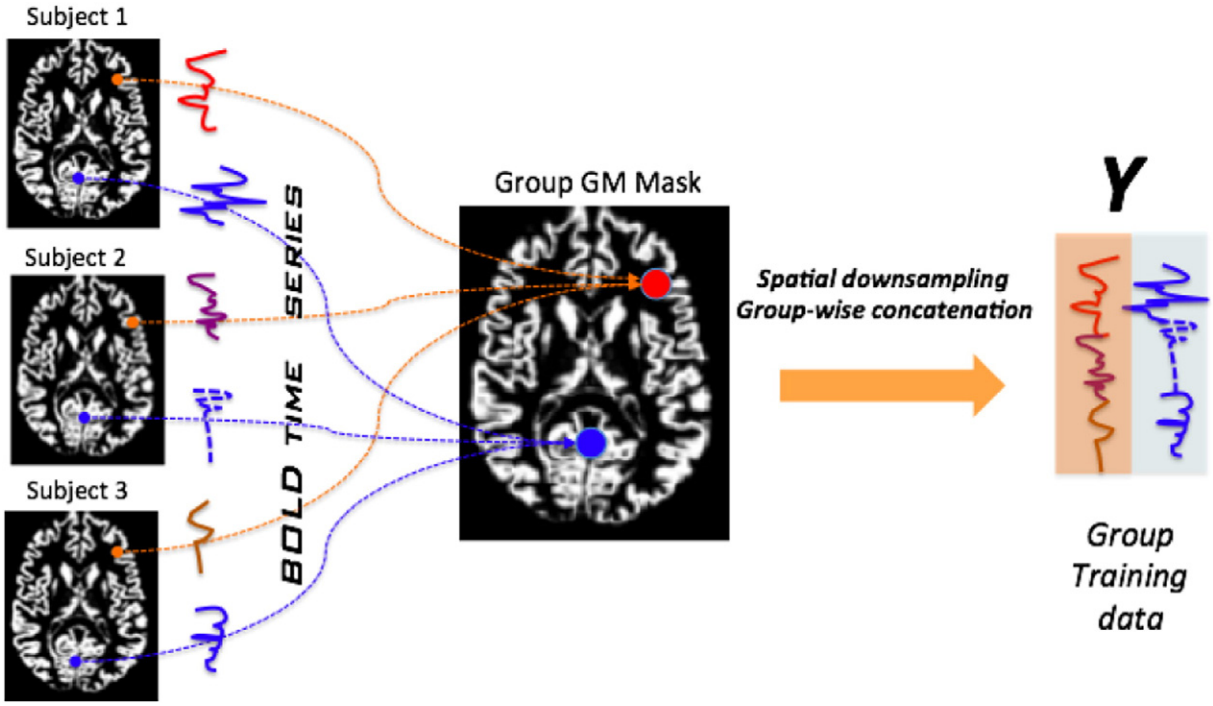

Fig. 3. Training data extraction for sparse dictionary learning from individual data sharing a same group gray matter mask.

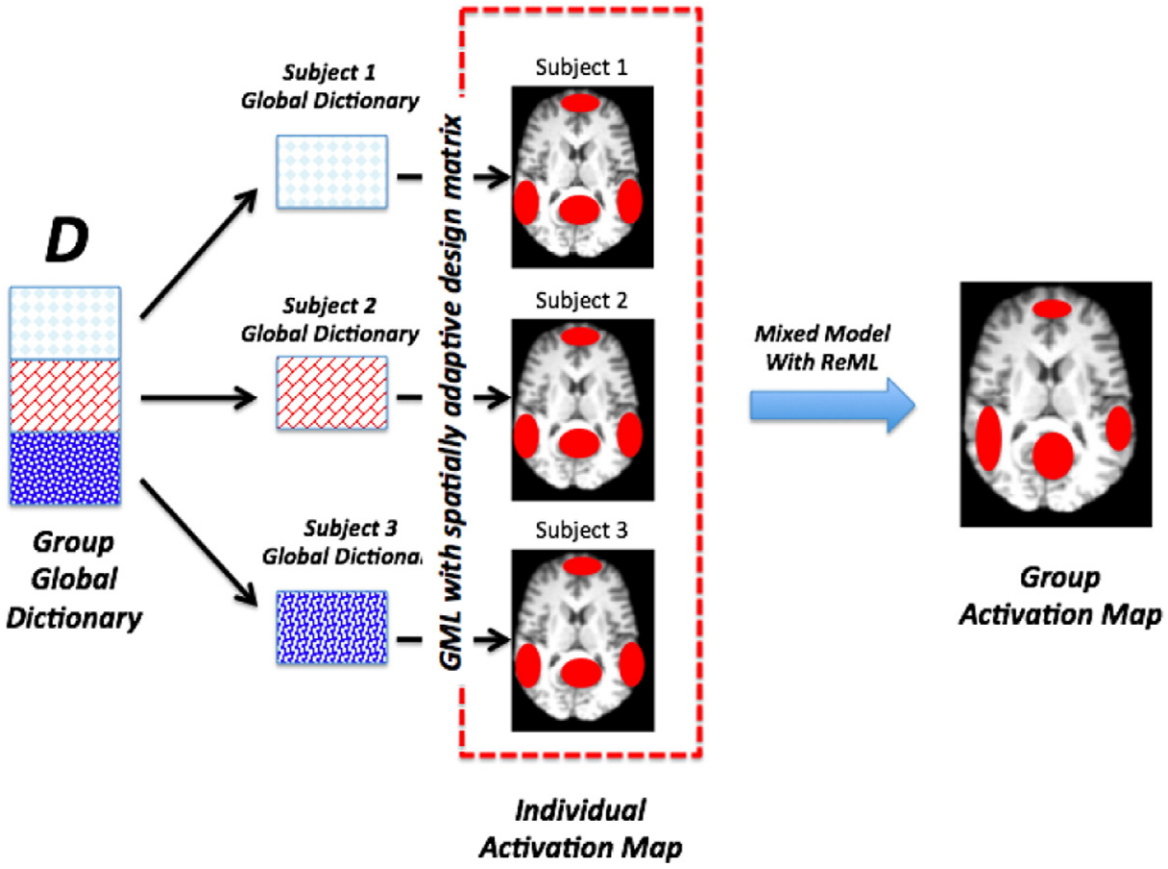

Fig. 4. Group activation detection using a learned group dictionary.

(from 1 to 10) with the normal subject group. We calculated the spatial correlation between DMNs from adjacent iteration numbers. The results in Fig. 6 show that with a small number of iterations ($n < 3$), there were some topological errors, which diminished as the iteration number increased. However, for more than three iterations, there were only small differences.

Network extraction

To verify that SSPM analysis can extract other networks simultaneously, Fig. 7 illustrates all 20 subnetworks extracted by SSPM analysis using normal subject data. We found other well-known networks, such as the salience network, primary visual network, fronto-temporal network, and sensory motor network. It is noteworthy that some of these well-known networks in Fig. 7 could not be observed using ICA analysis with the same data set (results not shown). This indicates that SSPM is quite effective in extracting other networks, which are also useful for resting-state data analysis.

If the noise is purely random, then SSPM does not extract the noise components, because the random noise is considered as background noise in the mixed-effect model. However, if the noise comes from head motion, our previous work (Lee et al., 2011b) for individual data showed that the motion could be decomposed as a separate component. In this data set, the experimental conditions were well controlled and

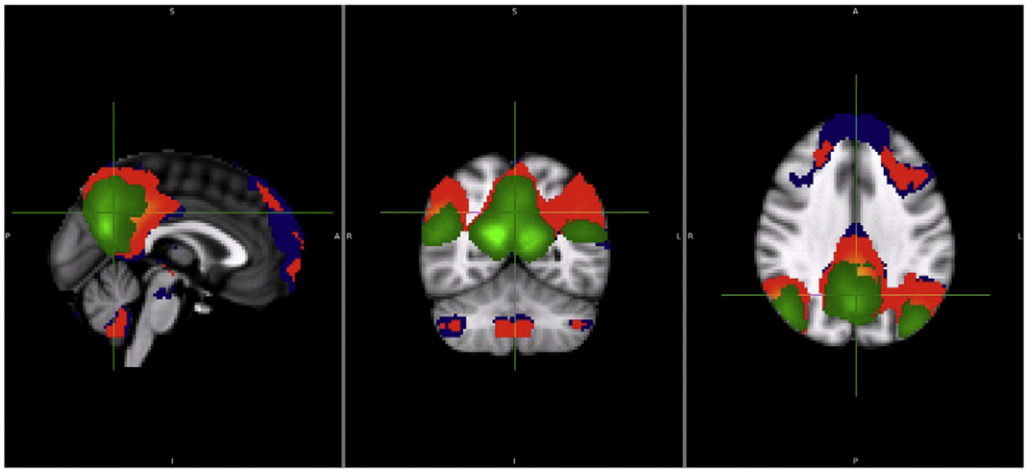

Fig. 5. The choice of the number of dictionary atoms of normal group data: (Blue) $N = 15$, (Red) $N = 20$, (Green) $N = 25$.

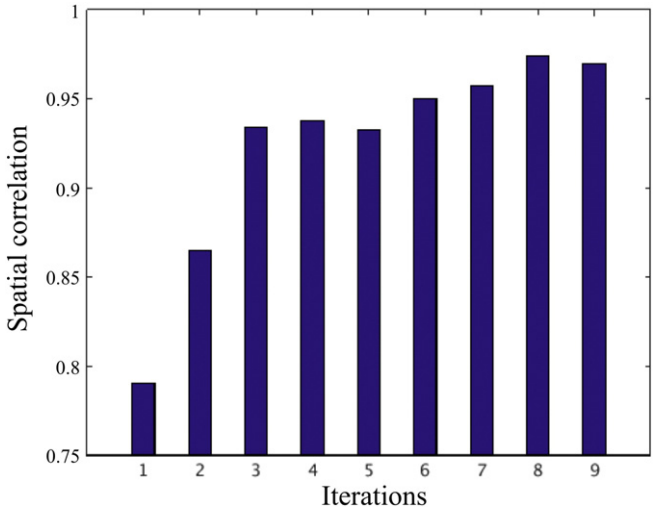

Fig. 6. Spatial correlation between DMNs from adjacent iteration numbers.

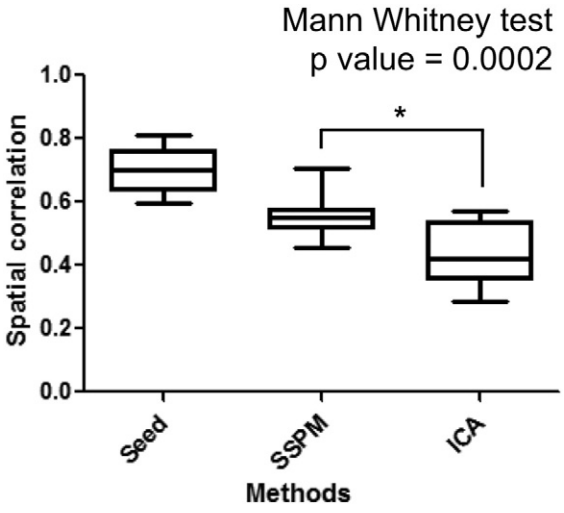

Fig. 8. Spatial correlation between individual DMN maps.

there were no perceivable motions from each subject; thus, we could not find any component that represented the motion. Consequently, as shown in Fig. 7, the 20 networks extracted by SSPM mostly included well-known networks, such as the DMN, sensory motor network, fronto-parietal network, and visual network.

Next, we calculated the spatial correlation of the DMN between individual and group data. The DMN component was identified by selecting the component with the highest spatial correlation with the result from the group data (see Fig. 8 and Table 1). Results showed that single-subject analysis of SSPM had a large spatial correlation value compared with the results from ICA method ($p = 0.0002$). Bartlett's test for equal variance showed that the difference in variance between SSPM and ICA is statistically significant ($p < 0.0147$), whereas the variance difference between SSPM and the seed-based approach is not ($p < 0.347$). Similarly, the difference in variance between ICA and the seed-based approach is not statistically significant ($p < 0.1251$). This indicates that SSPM also had a smaller variance compared to ICA, which is another advantage of SSPM over ICA analysis.

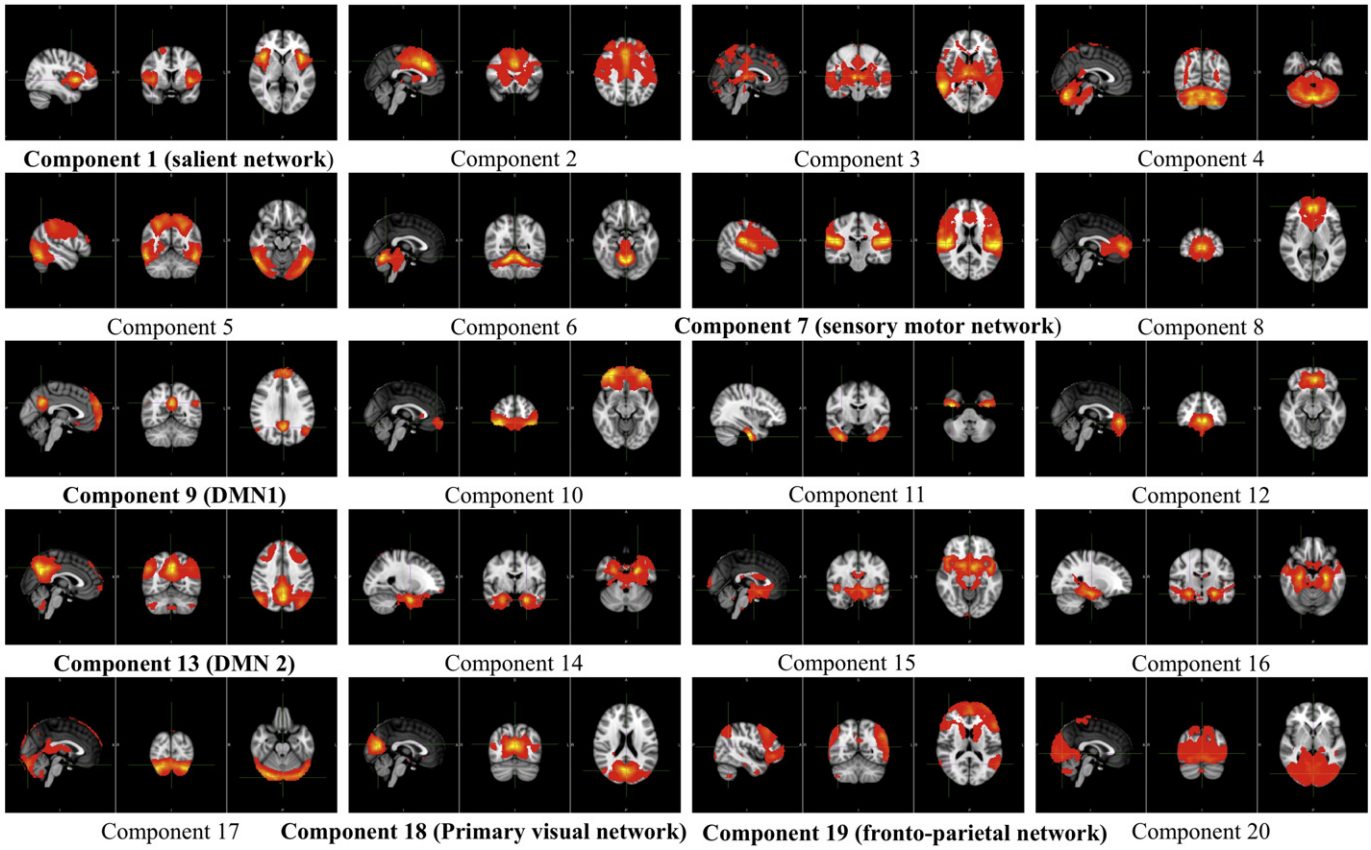

Fig. 7. All networks extracted by SSPM.

Table 1
Spatial correlation between individual maps generated by different methods.

	Seed-based	SSPM	ICA
Spatial Correlation	0.6991 ±0.0672[a]	0.5518 ±0.0545[a]	0.4366 ±0.0946[a]

[a] Mean ± standard deviation format.

Effect of disease progress

Because the proposed SSPM framework can conduct a group-level ANOVA to identify regions that are affected by disease progression, we performed such an analysis. Specifically, after performing an omnibus *F* test to find the statistically significant group effect on the DMN, we performed a 1 × 2 ANOVA to find the differences between adjacent groups. Fig. 9 illustrates representative slices. As shown in that figure, statistically significant changes occurs at the PCC area when the disease progress from normal to MCI. In addition to the PCC area, the differences at the MPFC and IPL became statistically significant when we compared normal versus Alzheimer's patient groups. Only the PCC area had a significant difference between the MCI and AD groups. We calculated the mean value of the *t*-maps across all activated voxels for each group to investigate whether the activation strength changes according to the disease progression. Results showed that the mean value of normal data was higher than in other groups (NL = 0.1122, MCI = 0.0854, AD = 0.1050). The mean value slightly increased as the disease progressed from MCI to AD. This coincided with the finding of Sohn et al. (2014), which used an identical data set. In particular, the results in Sohn et al. (2014) were obtained using the standard resting-state analysis tools, which showed an interesting finding that there can be compensatory mechanisms to make up for impaired cognitive function in the AD stage.

ADNI data results

The data on 20 normal subjects and 22 AD patients were collected from ADNI data. We then applied group SSPM to these data sets. Fig. 10(a) clearly shows the DMN network from the normal group. On the other hand, the DMN network from the AD data appeared more diffuse, as shown Fig. 10(b). This suggests that the DMN in the AD group may have deteriorated. The mean values of the SPM map (Normal: 0.1286; AD: 0.1069) also confirmed the weakening of the activations. Group differences between two groups from SSPM in Fig. 10(c) also showed that there were changes in the PCC and IPL areas between the groups.

Discussion

The ANOVA results in Fig. 9 clearly indicate progression of Alzheimer's disease. The results imply that the PCC is the first area to deteriorate and is followed by MPFC and IPL areas. These findings with SSPM coincide with the biological findings that the posterior components of the default network, including the precuneus and posterior cingulate, are particularly vulnerable to early deposition of amyloid β-protein, one of the hallmark pathologies of AD (Sperling et al., 2009). This clearly indicates that SSPM is a valuable tool for resting-state analysis.

One of the weaknesses of seed-based analysis compared to ICA and SSPM analysis is that other well-known networks should be recalculated with additional seeding, which is not the case for ICA and SSPM, which simultaneously extract the other subnetworks in addition to the DMN. In addition, even though the seed-based analysis has the highest spatial correlation among individual subjects (see Fig. 8), the seed-based approach may cause problems. For example, as shown in Fig. 11, when we defined a seed in the MPFC region, the seed-based approach failed to extract the fronto-parietal network, whereas the fronto-parietal network was intact in the SSPM analysis.

Single-subject analyses are more sensitive to noise, because a single-subject dataset has a low signal-to-noise ratio from insufficient information. The results can also be biased according to the nature of each individual. Thus, many brain network extraction methods only take the simple average to overcome these issues. To show that these analysis have systematic drawbacks, we used normal subject data (n = 22) to extract the averaged DMN from each individual DMN. Parameters used in the single-subject approach were the same. Among various subnetworks for each subject, we chose a single-subject DMN that has the highest spatial correlation with that of group DMN. Then, we averaged them and applied thresholding, assuming that the individual map followed the Gaussian distribution (Fig. 12). The results in Fig. 12 appear similar to those of the group SSPM analysis, but the activation patterns looks more diffuse, and more false activations were observed compared to those from the SSPM group analysis. Furthermore, there are additional limitations in this simple averaging approach. First, there are many subjects in whom the individual DMN was not clear, so finding individual DMNs using spatial correlation can be erroneous. Second, the statistical process for calculating the threshold value is based on the specific statistical assumption (such as Gaussian) of each individual DMN map. However, this assumption is not true considering the complicated procedure of calculating SPM maps, so we cannot quantify the *p*-value accurately. Third, to extract other subnetwork beyond the DMN, the selection of corresponding multiple subnetworks from each individual result is very complicated, which makes the simultaneous extraction of all subnetworks difficult. On the other hand, the group SSPM

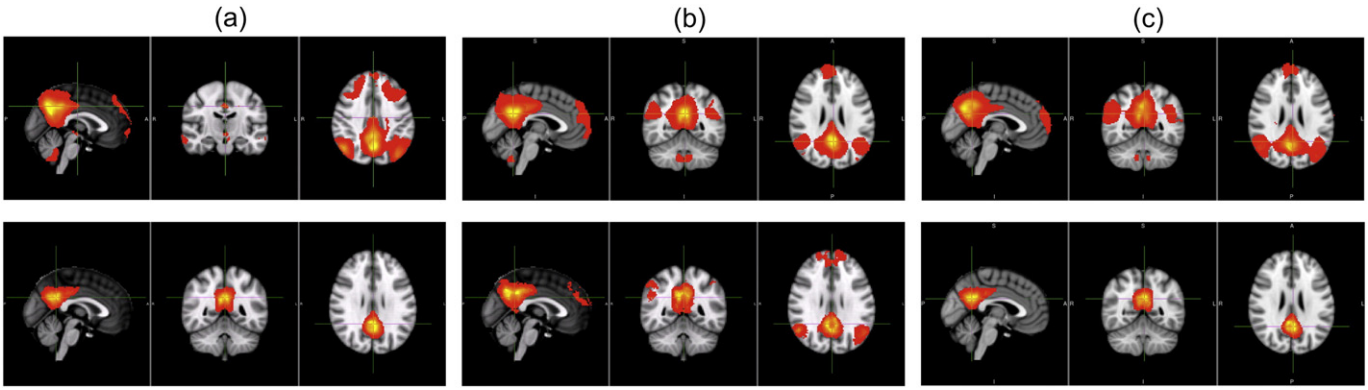

Fig. 9. (Upper) DMN of each group (a) normal, (b) MCI, and (c) AD. (Lower) 1 × 2 ANOVA maps with $p < 0.05$ (uncorrected) (a) normal vs. MCI, (b) normal vs. AD and (c) MCI vs. AD.

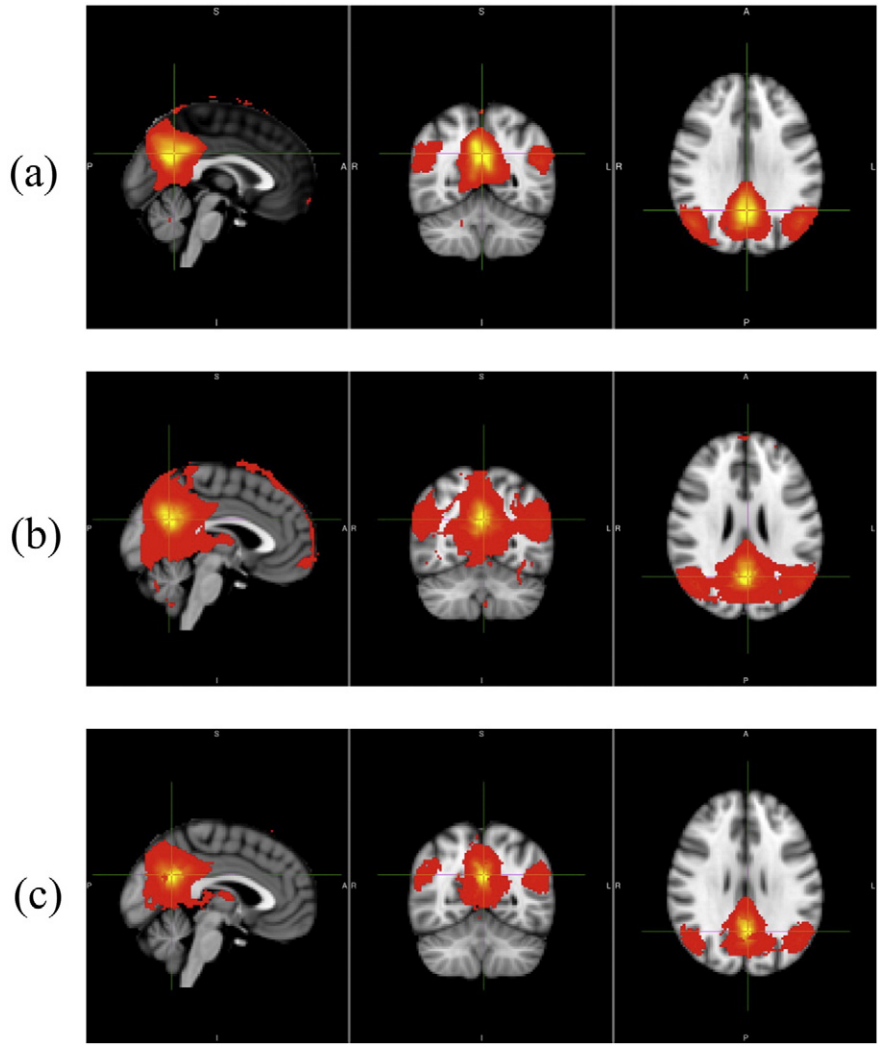

Fig. 10. Results from the ADNI dataset. (a) normal data, (b) AD data, (c) 1×2 ANOVA maps with $p < 0.05$ (uncorrected).

approach imposes the constraint of local connectivity, and therefore, the individual maps from the group SSPM are usually much clearer. Moreover, due to our ReML approach with a mixed-effect model, the quantification of the *p*-value is rigorous. Moreover, all the subnetworks can be estimated simultaneously.

To verify the stability of SSPM, we extracted the DMN from the subset of the whole data. Even and odd indexed datasets were split into two groups, and we applied the SSPM method to each group. Fig. 13 shows the extracted networks from the two sub-groups. They maintained a high spatial correlation with the original group result. In the DMN, the spatial correlation value between the whole dataset and the even order sub-dataset was 0.75, and it was 0.81 between the whole and odd order sub-datasets. This confirms that SSPM is not sensitive to the choice of data.

The sparse brain network model has been extensively investigated for functional brain connectivity analysis (Huang et al., 2010; Lee et al., 2011a). For example, Huang et al. (2010) used the sparse inverse covariance matrix to estimate brain connectivity, whereas Lee et al. (2011a) utilized the compressed sensing framework to estimate the sparse brain network. As an extension of these sparse brain network models, there are several existing studies that used the sparse dictionary learning (Lee et al., 2011b; Eavani et al., 2012; Abolghasemi et al., 2015; Zhang et al., 2015; Leonardi et al., 2014; Lv et al., 2014; Khalid and Seghouane, 2014), or sparse PCA (principal component analysis)

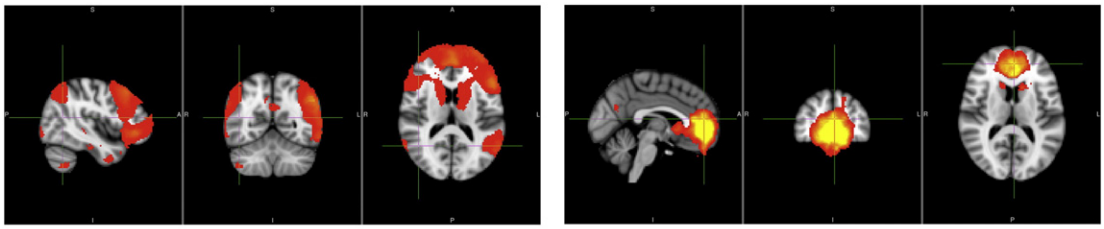

Fig. 11. Fronto-parietal network from SSPM (left), and the seed-based method (seed = MPFC) (right).

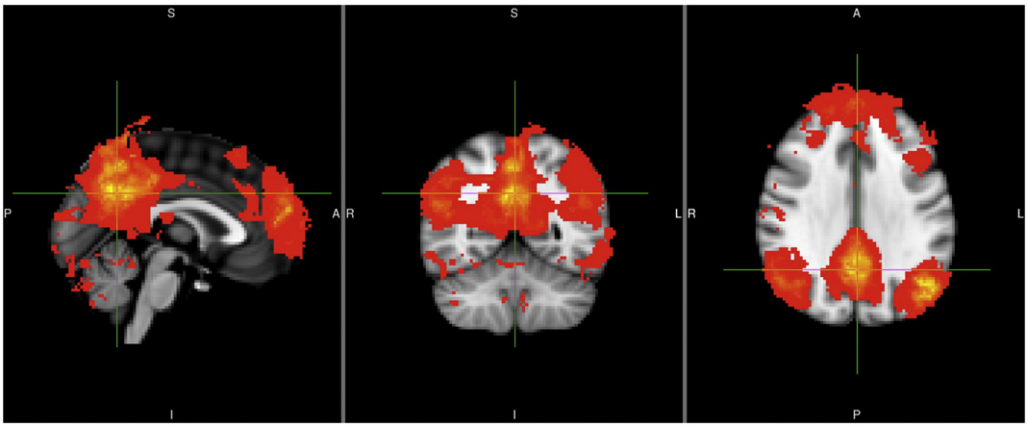

Fig. 12. Averaged DMN from each subject's DMN.

(Ulfarsson and Solo, 2007), which exploit the "sparsity" of temporal dynamics in the same way of our method. Specifically, the work by Lee et al. (2011b) was one of the earliest works that used the sparse dictionary learning to estimate the temporal dynamics in the sparse brain network. However, as described previously, the extension to the group-level analysis was lacking in Lee et al. (2011b). Similarly, the main focus of most other works was to develop different types of sparse dictionary learning algorithms (Abolghasemi et al., 2015; Lv et al., 2014) or sparse PCA (Ulfarsson and Solo, 2007, 2008). In another noteworthy approach in Leonardi et al. (2014), the authors were interested in extracting dynamic functional connectivity (dFC) using the sparse matrix factorization of a spatio-temporal times series, which is different from the static resting fMRI analysis we focus on in this paper. In Lv et al. (2014), the authors applied sparse dictionary learning for each

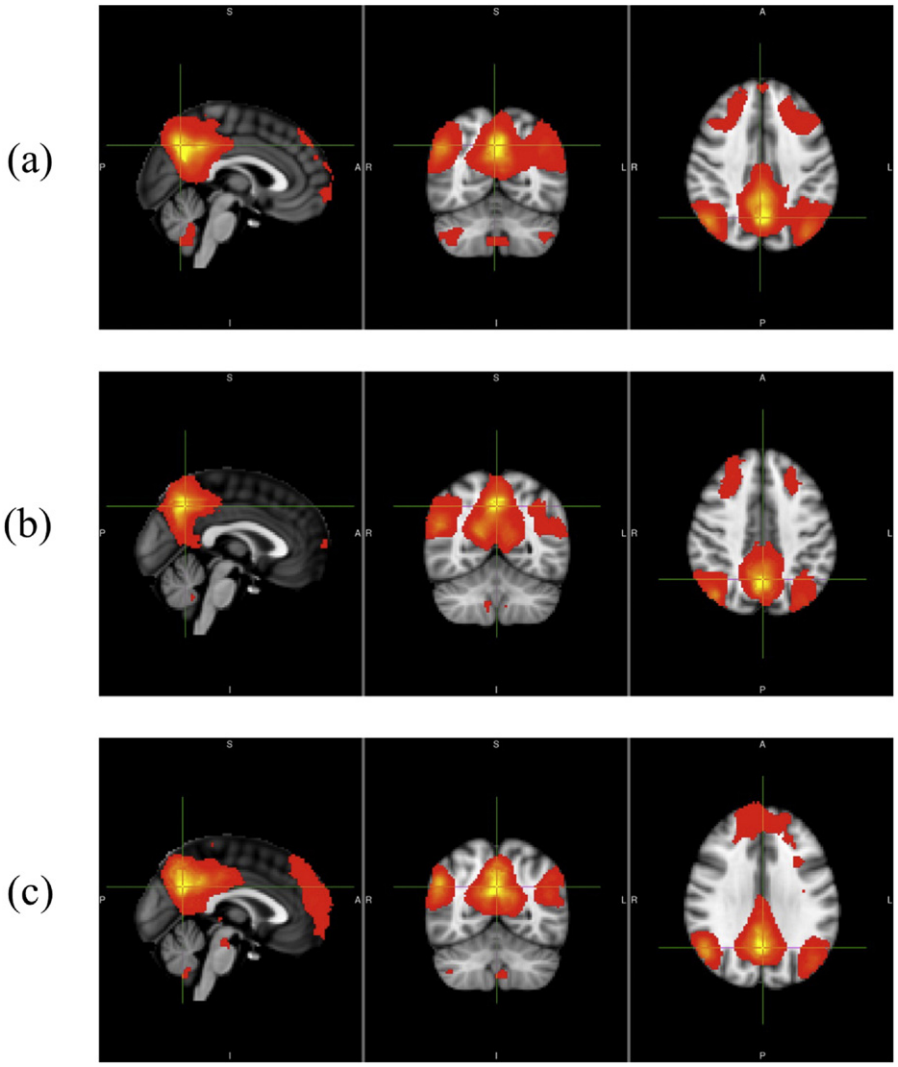

Fig. 13. DMN extracted by SSPM from the sub-dataset. (a) whole dataset, (b) sub-dataset from odd order, (c) sub-dataset from even order.

individual and obtained the group activating maps by simply averaging the individual activation patterns that were manually selected based on visual inspection, whereas in Zhang et al. (2015), the second-step dictionary learning was performed from individualized sparse dictionaries. However, these types of group analysis were developed in an ad hoc manner, the rigorous analysis of the impact to statistical analysis is lacking, and there is no systematic way of calculating the *p*-value from the standard statistical analysis. On the other hand, our group analysis was developed based on the standard mixed-effect model with a ReML covariance estimation, so accurate *p*-value calculation is possible. Moreover, the group-level local connectivity can be obtained automatically rather than by visual inspection, which makes the analysis very practical.

One of the limitations of the current work is that the analysis was performed based on the assumption that the functional connectivity during the resting-state does not change. Growing evidence suggests the importance of dynamical features of resting-state fMRI data to discover relevant organizations of brain function. The extension of the sparse SPM in this regard will be very important, and we believe that the recent work using the sparse innovation model (Karahanoğlu and Van De Ville, 2015) gives an important clue for this future work.

Conclusion

In this paper, we developed a unified mixed model called sparse SPM for group sparse dictionary learning and inference for resting-state fMRI analysis. Unlike ICA methods, the new algorithm exploits the fact that temporal dynamics at each voxel can be represented as a sparse combination of global dynamics because of the property of small-worldness of brain networks. In addition, the sparse coding step in the sparse dictionary learning step of our proposed method enabled SSPM for pacellation-free brain functional connectivity analysis solely based on the graph theoretical approach.

Under the reasonable assumption that the individual network structures in the same group are similar, we developed a group sparse dictionary learning algorithm that determines the subnetwork structures and mixed-effect parameters from a unified individual- and group-level statistical analysis and inference. This framework also enabled ANOVA at a group level, which provided systematic tools to investigate the disease progression.

We compared and validated our tools with the existing seed-based and ICA approaches for normal, MCI, and Alzheimer's disease patients. The results indicate that the DMN network extracted with our method shows a clear correlation with the progression of disease. Moreover, the SSPM tools provided more spatially correlated individual maps compared to the ICA analysis, which is another indication of the robustness of the algorithm. We also showed that SSPM can extract other subnetworks, which was not feasible using seed-based analysis.

Results indicate that extracted DMNs using the proposed method exhibit excellent correlation with disease progression. Moreover, the analysis of individual data showed much stronger spatial correlations between individual data compared to ICA analysis. Compared to taking the simple mean of the individualized sparse dictionary learning, the proposed ReML framework provides more rigorous and robust group network structures that can be utilized for comparing normal and patient groups. In addition, we demonstrate that other subnetworks such as the salience network, fronto-parietal network, sensory motor network, and primary visual network can be simultaneously detected using the proposed method, which is not the case in seed-based analysis.

Based on the results and unique theoretical advantages, we believe that the results in this paper provide strong evidence that the proposed SSPM is a powerful tool for resting-state fMRI analysis. The whole package of the SSPM will be available on the authors homepage (http://bispl.weebly.com).

Acknowledgment

JCY was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (NRF-2014R1A2A1A11052491). YJ was supported by the Brain Research Program (NRF-2010-0018843), Basic Science Research Program (NRF-2012R1A1A2044776) through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning.

Appendix A. Matrix equalities.

Using the matrix inversion lemma, we have

$$V^{-1} = (ZGZ' + R)^{-1} = R^{-1} - R^{-1}Z\left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1}.$$

Hence,

$$\begin{aligned} Z'V^{-1} &= Z'(R^{-1} - R^{-1}Z\left(G^{-1} + Z'R^{-1}Z)^{-1}Z'R^{-1}\right) \\ &= (I - Z'R^{-1}Z\left(G^{-1} + Z'R^{-1}Z)^{-1}\right)Z'R^{-1} \\ &= (I - \left(Z'R^{-1}Z + G^{-1} - G^{-1}\right)\left(G^{-1} + Z'R^{-1}Z)^{-1}\right)Z'R^{-1} \\ &= G^{-1}\left(G^{-1} + Z'R^{-1}Z\right)^{-1}Z'R^{-1} \\ &= \left(\left(Z'R^{-1}Z\right)^{-1} + G\right)^{-1}\left(Z'R^{-1}Z\right)^{-1}Z'R^{-1}\ , \end{aligned}$$

which leads us to

$$Z'V^{-1}Z = \left(\left(Z'R^{-1}Z\right)^{-1} + G\right)^{-1}.$$

Therefore, for $X = ZX_G$, we have

$$X'V^{-1}X = X_G'V_G^{-1}X_G \tag{24}$$

where

$$V_G = \left(Z'V^{-1}Z\right)^{-1} = \left(Z'R^{-1}Z\right)^{-1} + G \tag{25}$$

B. Equivalence

Consider the following form of covariance matrices:

$$V = ZGZ' + R, \tag{26}$$

$$\Omega = V^{-1} - V^{-1}X\left(X'V^{-1}X\right)^{\dagger}X'V^{-1} \tag{27}$$

where R, G, and V are assumed invertible. Using the definition of V and Ω in Eqs. (26) and (27), it is easy to show

$$\Omega X = 0, \quad X'\Omega = 0, \quad P_X^{\perp}\Omega = \Omega, \quad P_X^{\perp}V\Omega = I \tag{28}$$

Now, we show that $\Omega = (P_X^{\perp}VP_X^{\perp})^{\dagger}$. To this, we need to show the following

$$\begin{aligned} P_X^{\perp}VP_X^{\perp}\Omega P_X^{\perp}VP_X^{\perp} &= P_X^{\perp}VP_X^{\perp} \\ \Omega P_X^{\perp}VP_X^{\perp}\Omega &= \Omega \\ \left(P_X^{\perp}VP_X^{\perp}\Omega\right)' &= P_X^{\perp}VP_X^{\perp}\Omega \\ \left(\Omega P_X^{\perp}VP_X^{\perp}\right)' &= \Omega P_X^{\perp}VP_X^{\perp} \end{aligned}$$

These are straightforward results using the properties in Eq. (28). Now, as $P_X^{\perp} = P_Q$, we have $(P_X^{\perp}VP_X^{\perp})^{\dagger} = (P_QVP_Q)^{\dagger} = Q(Q'VQ)^{-1}Q'$.

Therefore, we have

$$Q(Q'VQ)^{-1}Q' = (P_X^{\perp}VP_X^{\perp})^{\dagger}$$

Now, we are ready to prove the main result on the equivalence. As we are only interested in estimable function, we assume $L' \subset R(X')$ (Rao and Toutenburg, 1999). Then, the test statistics is

$$S = \frac{\hat{\boldsymbol{\alpha}}'C'\left(C'(X'X)^{\dagger}C\right)^{-1}C\hat{\boldsymbol{\alpha}}}{\hat{\sigma}^2\mathrm{rank}(C)} = \frac{\mathbf{y}'X'(X'X)^{\dagger}C'\left(C'(X'X)^{\dagger}C\right)^{-1}C(X'X)^{\dagger}X'\mathbf{y}}{\hat{\sigma}^2 p_1}$$

From the definition of generalized inverse, we can find the full rank matrix X^* such that $R(X^*) = R(X)$ and $(X'X)^{\dagger} = (X^{*'}X^*)^{-1}$. Therefore, using the definition Eq. (27), we can define Ω such that

$$\Omega = \left(X^{*'}X^*\right)^{-1} - \left(X^{*'}X^*\right)^{-1}C'\left(C'(X'X)^{\dagger}C\right)^{-1}C\left(X^{*'}X^*\right)^{-1} = \left(P_{C'}^{\perp}\left(X^{*'}X^*\right)^{-1}P_{C'}^{\perp}\right)^{\dagger}$$

Then, we have

$$\begin{aligned}\mathcal{S} &= \frac{\mathbf{y}'X'\left(\left(X^{*'}X^*\right)^{-1} - \Omega_{X^{*'}X^*}\right)X'\mathbf{y}}{\hat{\sigma}^2 p_1} \\ &= \frac{\mathbf{y}'X'\left((X'X)^{\dagger} - (P_{C'}^{\perp}(X'X)P_{C'}^{\perp})^{\dagger}\right)X'\mathbf{y}}{\hat{\sigma}^2 p_1} \\ &= \frac{\mathbf{y}'X'\left((X'X)^{\dagger} - Q(Q'X'XQ)^{-1}Q\right)X'\mathbf{y}}{\hat{\sigma}^2 p_1} \\ &= \frac{\mathbf{y}'(P_X - P_{X_0})\mathbf{y}}{\hat{\sigma}^2 p_1} \\ &= \frac{\mathbf{y}'\left(P_{X_0}^{\perp} - P_X^{\perp}\right)\mathbf{y}}{\mathbf{y}'P_X^{\perp}\mathbf{y}}\frac{N-p}{p_1}\end{aligned}$$

where $X_0 = XQ$ denotes the reduced model by excluding the effect of the contrast matrix C. This concludes the proof.

References

Abolghasemi, V., Ferdowsi, S., Sanei, S., 2015. Fast and incoherent dictionary learning algorithms with application to fmri. SIViP 9 (1), 147–158.

Aharon, M., Elad, M., Bruckstein, A., 2006. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process. 54 (11), 4311–4322.

Beckmann, C., DeLuca, M., Devlin, J., Smith, S., 2005. Investigations into resting-state connectivity using independent component analysis. Philos. Trans. R. Soc. B Biol. Sci. 360 (1457), 1001.

Biswal, B., Yetkin, F., Haughton, V., Hyde, J., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magn. Reson. Med. 34 (4), 537–541.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186–198.

Cordes, D., Haughton, V., Arfanakis, K., Carew, J., Turski, P., Moritz, C., Quigley, M., Meyerand, M., 2001. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. Am. J. Neuroradiol. 22 (7), 1326.

Damoiseaux, J., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C., Smith, S., Beckmann, C., 2006. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. 103 (37), 13848.

De Luca, M., Beckmann, C., De Stefano, N., Matthews, P., Smith, S., 2006. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. NeuroImage 29 (4), 1359–1367.

Eavani, H., Filipovych, R., Davatzikos, C., Satterthwaite, T.D., Gur, R.E., Gur, R.C., 2012. Sparse dictionary learning of resting state fMRI networks. 2012 International Workshop on Pattern Recognition in NeuroImaging (PRNI). IEEE, pp. 73–76.

Fox, M., Raichle, M., 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat. Rev. Neurosci. 8 (9), 700.

Fox, M., Zhang, D., Snyder, A., Raichle, M., 2009. The global signal and observed anticorrelated resting state brain networks. J. Neurophysiol. 101 (6), 3270.

Fransson, P., 2005. Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. Hum. Brain Mapp. 26 (1), 15–29.

Friston, K., Ashburner, J., Kiebel, S., Nichols, T., Penny, W., 2011. Statistical Parametric Mapping: The Analysis of Functional Brain Images: The Analysis of Functional Brain Images. Academic Press.

Graser, H., Smith, S., Tier, B., 1987. A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood. J. Anim. Sci. 64 (5), 1362–1370.

Greicius, M., Srivastava, G., Reiss, A., Menon, V., 2004. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional mri. Proc. Natl. Acad. Sci. U. S. A. 101 (13), 4637.

Harville, D., 1977. Maximum likelihood approaches to variance component estimation and to related problems. J. Am. Stat. Assoc. 72 (358), 320–338.

Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., Chen, K., Reiman, E., A. D. N. Initiative, et al., 2010. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. NeuroImage 50 (3), 935–949.

Karahanoğlu, F.I., Van De Ville, D., 2015. Transient brain activity disentangles fmri resting-state dynamics in terms of spatially and temporally overlapping networks. Nat. Commun. 6.

Kenward, M., Roger, J., 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 983–997.

Khalid, M.U., Seghouane, A.-K., 2014. A single svd sparse dictionary learning algorithm for fmri data analysis. 2014 IEEE Workshop on Statistical Signal Processing (SSP). IEEE, pp. 65–68.

Lee, H., Lee, D.S., Kang, H., Kim, B.-N., Chung, M.K., 2011a. Sparse brain network recovery under compressed sensing. IEEE Trans. Med. Imaging 30 (5), 1154–1165.

Lee, K., Tak, S., Ye, J., 2011b. A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion. IEEE Trans. Med. Imaging 30 (5), 1076–1089.

Leonardi, N., Shirer, W.R., Greicius, M.D., Van De Ville, D., 2014. Disentangling dynamic networks: separated and joint expressions of functional connectivity patterns in time. Hum. Brain Mapp. 35 (12), 5984–5995.

Lowe, M., Mock, B., Sorenson, J., 1998. Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. NeuroImage 7 (2), 119–132.

Lustig, C., Snyder, A., Bhakta, M., O'Brien, K., McAvoy, M., Raichle, M., Morris, J., Buckner, R., 2003. Functional deactivations: change with age and dementia of the alzheimer type. Proc. Natl. Acad. Sci. U. S. A. 100 (24), 14504.

Lv, J., Jiang, X., Li, X., Zhu, D., Chen, H., Zhang, T., Zhang, S., Hu, X., Han, J., Huang, H., et al., 2014. Sparse representation of whole-brain fMRI signals for identification of functional networks. Medical image analysis.

Minoshima, S., Giordani, B., Berent, S., Frey, K., Foster, N., Kuhl, D., 1997. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Ann. Neurol. 42 (1), 85.

Patterson, H., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58 (3), 545–554.

Rao, C., Toutenburg, H., 1999. Linear models: least squares and alternatives. Springer.

Rombouts, S., Stam, C., Kuijer, J., Scheltens, P., Barkhof, F., 2003. Identifying confounds to increase specificity during a "no task condition". NeuroImage 20 (2), 1236–1245.

Ryali, S., Chen, T., Supekar, K., Menon, V., 2012. Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty. NeuroImage 59 (4), 3852–3861.

SAS Institute, 1985. SAS user's guide: statistics vol. 2. Sas Inst.

Schlösser, R., Gesierich, T., Kaufmann, B., Vucurevic, G., Hunsche, S., Gawehn, J., Stoeter, P., 2003. Altered effective connectivity during working memory performance in schizophrenia: a study with fmri and structural equation modeling. NeuroImage 19 (3), 751–763.

Searle, S., 1979. Notes on variance component estimation: a detailed account of maximum likelihood and kindred methodology. Biometrics Unit, Cornell University.

Sohn, W.S., Yoo, K., Na, D.L., Jeong, Y., 2014. Progressive changes in hippocampal resting-state connectivity across cognitive impairment: a cross-sectional study from normal to alzheimer disease. Alzheimer Dis. Assoc. Disord. 28 (3), 239–246.

Sperling, R.A., LaViolette, P.S., O'Keefe, K., O'Brien, J., Rentz, D.M., Pihlajamaki, M., Marshall, G., Hyman, B.T., Selkoe, D.J., Hedden, T., et al., 2009. Amyloid deposition is associated with impaired default network function in older persons without dementia. Neuron 63 (2), 178–188.

Ulfarsson, M.O., Solo, V., 2007. Sparse variable principal component analysis with application to fMRI. ISBI 2007. 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007. IEEE, pp. 460–463.

Ulfarsson, M.O., Solo, V., 2008. Sparse variable PCA using geodesic steepest descent. IEEE Trans. Signal Process. 56 (12), 5823–5832.

van de Ven, V., Formisano, E., Prvulovic, D., Roeder, C., Linden, D., 2004. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum. Brain Mapp. 22 (3), 165–178.

Yang, Z., Zuo, X.-N., Wang, P., Li, Z., LaConte, S.M., Bandettini, P.A., Hu, X.P., 2012. Generalized RAICAR: discover homogeneous subject (sub) groups by reproducibility of their intrinsic connectivity networks. NeuroImage 63 (1), 403–414.

Zhang, S., Li, X., Lv, J., Jiang, X., Guo, L., Liu, T., 2015. Characterizing and differentiating task-based and resting state fMRI signals via two-stage sparse representations. Brain Imaging Behav. 1–12.

Zuo, X.-N., Kelly, C., Adelstein, J.S., Klein, D.F., Castellanos, F.X., Milham, M.P., 2010. Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. NeuroImage 49 (3), 2163–2177.