IDENTIFYING FMRI DYNAMIC CONNECTIVITY STATES USING AFFINITY PROPAGATION CLUSTERING METHOD: APPLICATION TO SCHIZOPHRENIA

Mustafa S Salman[1,2], Yuhui Du[2,3] and Vince D Calhoun[1,2*]*

[1]University of New Mexico, Dept. of ECE, Albuquerque, NM, USA
[2]The Mind Research Network, Albuquerque, NM, USA
[3] Shanxi University, School of Computer & Information Technology, Taiyuan, China
ydu@mrn.org, vcalhoun@mrn.org

ABSTRACT

Numerous studies have shown that brain functional connectivity patterns can be time-varying over periods of tens of seconds. It is important to capture inherent non-stationary connectivity states for a better understanding of the influence of disease on brain connectivity. K-means has been widely used to extract the connectivity states from dynamic functional connectivity. However, K-means is dependent on initialization and can be exponentially slow in converging due to extensive noise in dynamic functional connectivity. In this work, we propose to use an affinity propagation clustering method to estimate the connectivity states. By applying K-means and the new method separately, we analyzed dynamic functional connectivity of 82 healthy controls and 82 schizophrenia patients, and then explored group differences between schizophrenia patients and healthy controls in the identified connectivity states. Both methods revealed that group differences mainly lay in visual, sensorimotor and frontal cortices. However, the new approach found more meaningful group differences than K-means. Our finding supports that our method is promising in exploring biomarkers of mental disorders.

Index Terms— functional MRI, dynamic connectivity, affinity propagation, schizophrenia

1. INTRODUCTION

Whole-brain functional connectivity (FC) derived from functional magnetic resonance imaging (fMRI) data has shown its power in the study of healthy and diseased brain. Different from traditional static FC analysis, dynamic functional network connectivity (dFNC) analysis can capture time-varying FCs among networks over tens of seconds [1] [2] [3] [4]. It is expected that the inherent non-stationary connectivity states extracted from dFNC can provide informative biomarkers for distinguishing mental disorders.

So far K-means clustering [1], principal component analysis (PCA) [5], as well as spatial and temporal independent component analysis (ICA) [6] [7] [8] have been used to estimate the reoccurring connectivity states. Among these approaches, K-means clustering, built into the group ICA of fMRI toolbox (GIFT) (http://mialab.mrn.org/software/gift/), is commonly used. K-means initializes cluster centroids by random sampling and iteratively refines them to minimize error. However we found that K-means is quite sensitive to the initial choice of centroids though this problem can be improved somewhat by multiple K-means runs with different initialization. Also, K-means can fail to converge within maximum number of iterations defined as termination criteria when clustering the time-varying FC patterns, probably due to extensive noise in dFNC. Therefore, the resulting connectivity states from K-means may be inaccurate, which influences the effectiveness of the subsequent biomarker identification.

Affinity propagation (AP) clustering provides an alternative approach with much lower error rates [9]. This method performs clustering by using similarity measures between pairs of samples and propagating information until a high-quality set of exemplars and corresponding clusters gradually emerge. There are several advantages of using affinity propagation (AP) over K-means. K-means is susceptible to local minima caused by poor initialization especially when dealing with dFNC data which has high dimensionality and may be influenced by noise. In contrast, AP clustering simultaneously considers information of all samples and thus avoids the problem caused by poor initialization. Just as K-means, different distance measures can be used in AP clustering to build the similarity matrix. However, AP clustering does not require users to specify the number of clusters as input; rather it can identify the number automatically based on an input preference value, which can be chosen based on prior information in order to estimate the connectivity states.

Recently AP has been used to cluster time-averaged connectivity patterns from a longitudinal resting-state fMRI dataset [10] and to identify networks from resting-state fMRI data [11], [12]. To our knowledge, no studies used AP to analyze dynamic connectivity. In this paper we propose to use AP clustering to estimate connectivity states from dynamic connectivity.

2. MATERIALS & METHODS

We investigated the group differences between schizophrenia patients (SZs) and healthy controls (HCs) in connectivity states that were identified using K-means or AP clustering method. Fig. 1 is an outline of our work.

2.1. Materials

Resting-state fMRI data was collected from 82 HCs (age: 37.7 ± 10.8, 19 females) and 82 SZs (age: 38.0 ± 14.0, 17 females) scanned on a 3-Tesla Siemens Trio scanner with a 12-channel radio frequency coil at the Mind Research Network (MRN) (see our previous work [4] for details). The functional scans were acquired using gradient echo planar imaging (EPI) with the following parameters: echo time (TE) $= 29ms$, repeat time (TR) $= 2s$, flip angle = 75°, slice thickness $= 3.5mm$, slice gap $= 1.05mm$, field of view = $240mm$, ma-

This work was in part funded by NIH grants P20GM103472 and R01EB020407 (to Calhoun VD), and NSF grant 1539067 (to Calhoun VD) and by natural science foundation of Shanxi (2016021077, to Du YH).

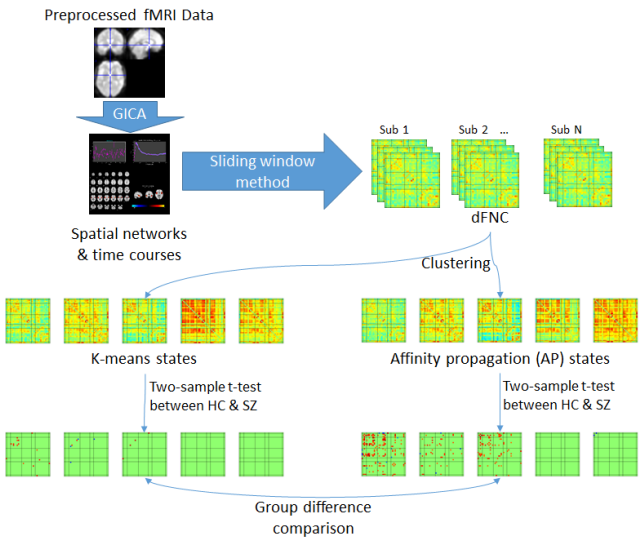

Fig. 1. Overview of methods. Preprocessed fMRI data was subject to group ICA (GICA) in order to parcellate whole brain into small networks. Then dFNC was estimated based on the time series of those networks using a sliding window approach. Afterwards, connectivity states were estimated by performing K-means and affinity propagation clustering methods on dFNC separately. Finally, group differences between HC and SZ in connectivity states were identified for each method.

trix size $= 64 \times 64$, voxel size $= 3.75mm \times 3.75mm \times 4.55mm$. Resting state scans consisted of 150 whole brain images. During data acquisition, subjects were asked to remain alert with eyes open and keep their head still.

A preprocessing pipeline developed at the MRN was used to preprocess the fMRI data [13]. The first 6 volumes from each scan were discarded to allow T1 equilibration. INRIAlign was used to realign the images [14]. Then the data was spatially normalized to the standard Montreal Neurological Institute (MNI) space, resampled to $3mm \times 3mm \times 3mm$ voxels using the nonlinear (affine + low frequency direct cosine transform basis functions) registration implemented in the SPM12 toolbox (http://www.fil.ion.ucl.ac.uk/spm), and smoothed using a Gaussian kernel with a full-width at half-maximum of $8mm$.

2.2. Group independent component analysis (ICA) & dFNC

We performed a group ICA method in GIFT on the preprocessed fMRI data to obtain individual functional networks and their associated time series. First, a group-level spatial ICA was performed using Infomax algorithm to obtain group-level independent components [15]. Then subject-specific components and corresponding time courses (TCs) were calculated based on the group-level components using GICA1 back-reconstruction [16]. After discarding artifact-related components [18], the remaining 36 independent components (ICs) of each subject were characterized as functional networks (Fig. 2) [17]. We post-processed the TCs of the 36 networks by detrending, regressing out head motion, despiking and performing low-pass filtering ($< 0.15Hz$).

A sliding window method is the most commonly used strategy for analyzing dynamics [1] [3] [18] [19]. In our work, a window with size of 26 TR ($52s$) and step of 1 TR was used to separate

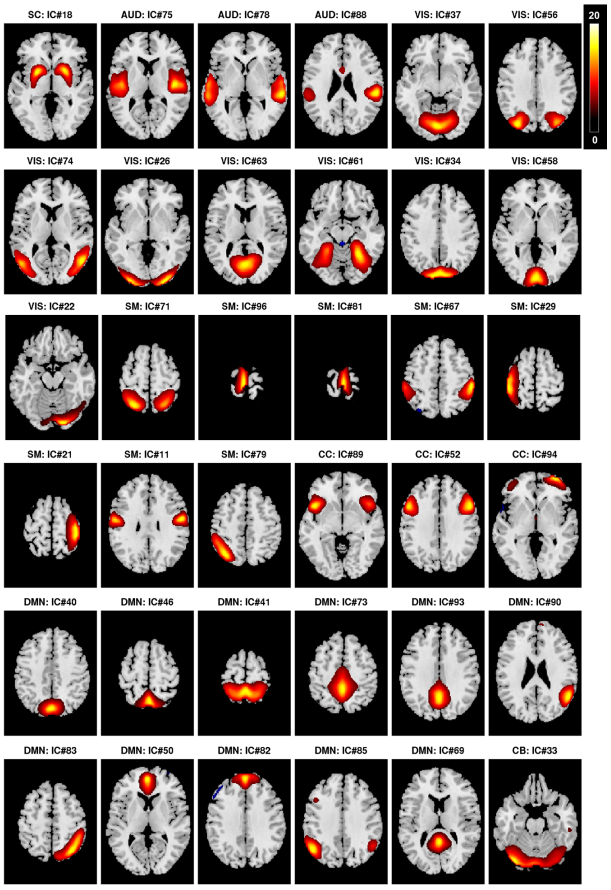

Fig. 2. Axial view of 36 networks obtained from group ICA, categorized into subcortical (SC), auditory (AUD), visual (VIS), sensorimotor (SM), cognitive control (CC), default mode network (DMN) and cerebellar (CB) networks.

each TC into 118 short TCs. Each window was convolved with a Gaussian of $\sigma = 3\,TR$ to obtain tapering along the edges. Then, regarding each window, $36 \times (36-1)/2 = 630$ FCs among networks were estimated based on 36 networks' short TCs in this window from a regularized inverse covariance matrix using graphical LASSO framework [20] [21]. So, for each subject, its dFNC can be represented by a 118×630 matrix. The connectivity values were then Fisher-Z transformed [3].

2.3. Extracting connectivity states from dFNC

2.3.1. K-means clustering

K-means has been widely used to extract connectivity states from dFNC [1] [3]. In our work, each window of each subject corresponds to a vector of 630 FC values. At first a subset of windows were identified as exemplars based on the local maxima of standard deviation of the 630 FCs across all windows and subjects. The number of optimal clusters was determined as five using the gap statistic [22]. K-means was performed on the exemplar windows-related FCs using cityblock distance as the distance measure. These resulting centroids were then used to initialize another K-means to cluster the FCs from all windows of all subjects. K-means was replicated 150 times to avoid local minima. Finally, we estimated the subject-specific

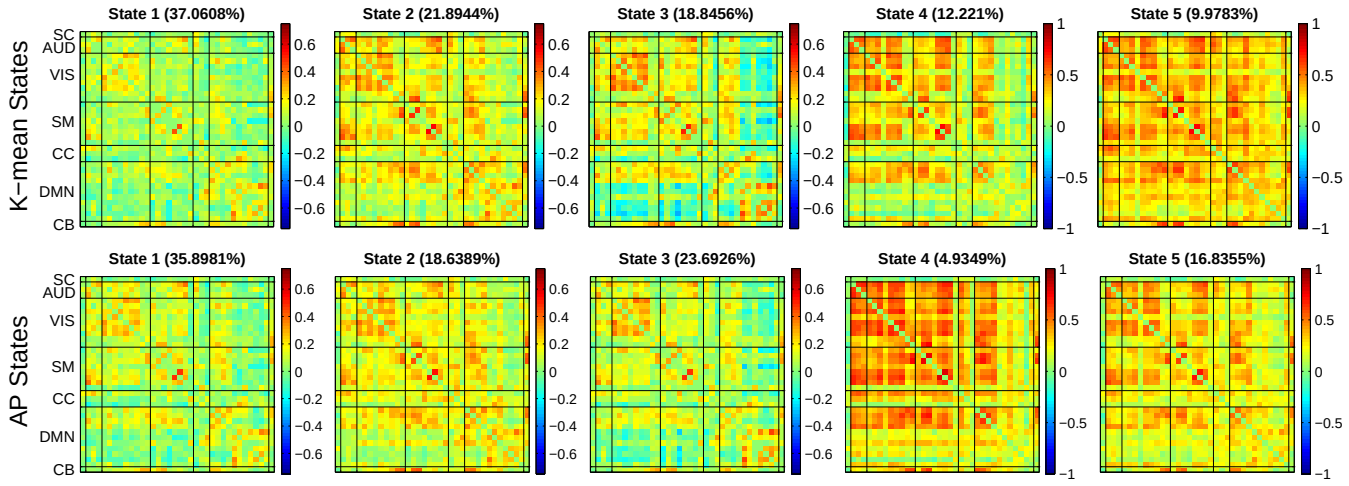

Fig. 3. The mean of individual subject's corresponding connectivity states across subjects for K-means and AP methods, respectively. The occupancy (ratio of dFNC windows belonging to any state and total number of dFNC windows across all subjects, expressed as a percentage) is also shown. States are sorted according to similarity between connectivity states estimated by two approaches.

connectivity states for each subject by averaging the associated FCs which are in windows with the same label [4].

2.3.2. Affinity propagation clustering

Affinity propagation performs clustering by using similarity measures between pairs of samples and propagating information until a high-quality set of exemplars and corresponding clusters gradually emerge [9]. It takes a collection of real-valued similarities between the data points as an input. The similarity $s(i,k)$ between two data-points i and k indicate how similar they are. The similarity criterion can be general, e.g. if the goal is to minimize the squared error, the similarity is set to a negative Euclidean distance. Rather than specifying a required number of clusters, an input real number $s(k,k)$ (preference) is specified for each data point k so that data points with high $s(k,k)$ are more likely to be identified as exemplars. If the preference is set as a real-valued scalar, then all data-points are treated as equally suitable exemplars. A high value of preference results in a high number of clusters being identified and vice versa. But preference can also be specified as a vector of the same length as the number of data points where the preference of each data point to be chosen as an exemplar is based on prior information. Also, the clustering can be performed with only a small number of known similarities between the data points.

There are two kinds of messages exchanged between samples. The responsibility, $r(i,k)$ is sent from data point i to candidate exemplar k as a measure of how suitable point k is as the exemplar of point i, taking the other candidate exemplars into account. The availability, $a(i,k)$ sent from candidate exemplar k to point i measuring how suitable k is for $i$to choose it as the exemplar, given support information from other points to k. The update equation for responsibilities is,

$$r(i,k) \leftarrow s(i,k) - \max_{k' \ s.t.\ k' \neq k} \{a(i,k') + s(i,k')\} \qquad (1)$$

Here, $s(i,k)$ is the input similarity between points i and k. For $k = i$, the responsibility $r(k,k)$ is set based on the input preference $s(k,k)$ minus the largest of the similarities between the point and all other candidate exemplars reflecting how ill-suited it is to be assigned to another exemplar.

The update equation for availability is,

$$a(i,k) \leftarrow min\{0, r(k,k) + \max_{i' \ s.t.\ i' \notin \{i,k\}} \{0, r(i,k')\}\} \qquad (2)$$

For the first iteration, the responsibilities are set to the input similarities and the availabilities are set to zero. In later iterations, the availabilities of some of the points drop below zero, indicating those being assigned to other exemplars. At any point during propagation, exemplars can be identified by combining responsibilities and availabilities i.e. $max\{a(i,k) + r(i,k)\}$. The termination criteria may be a fixed number of iterations, incremental changes in the messages falling below a threshold or decisions staying constant for certain number of iterations.

To apply AP to dFNC data, we first computed the similarity matrix between dFNC from all windows of all subjects based on city-block distance. For the sake of comparison we attempted to estimate the same number of clusters (five) from AP as from K-means. AP clustering was tried several times until an appropriate preference value was found using bisection method for which the algorithm produced five clusters. Once the clusters were found, the connectivity states were computed as the average connectivity patterns of windows with the same label.

2.4. Investigation of group differences in connectivity states

For each method of estimating connectivity states, we investigated the difference in each functional connectivity (FC) strength between healthy controls (HCs) and schizophrenia patients (SZs) using a two-sample t-test based on the corresponding subject-specific states.

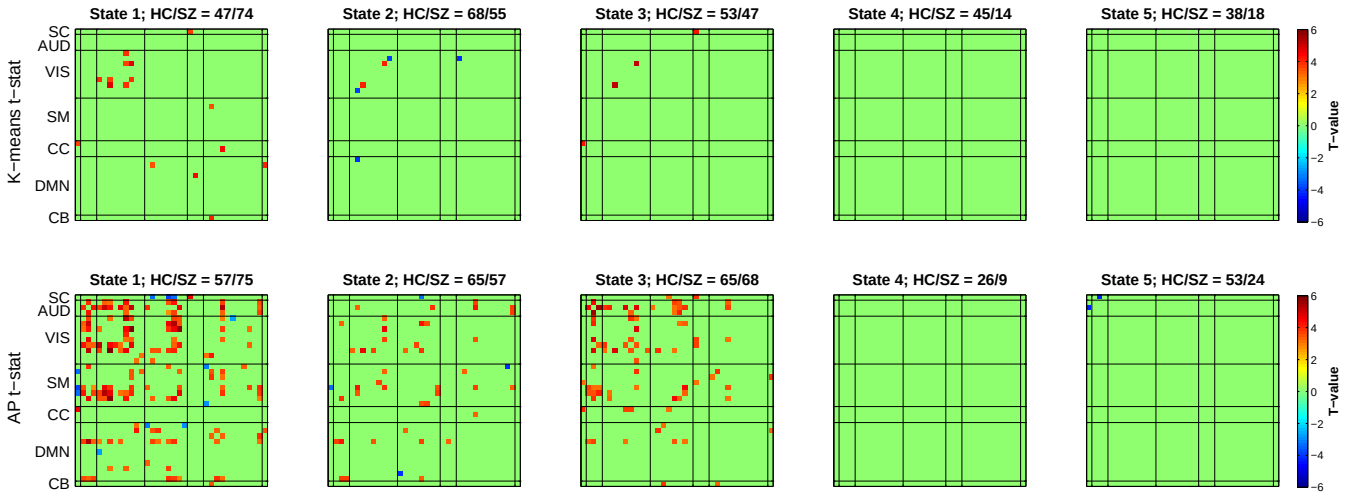

Fig. 4. HC vs SZ group difference in connectivity states estimated by two methods (K-means and AP approach). Two-sample t-test was performed on each connectivity of each state between HCs and SZs to investigate group difference. T-values are shown for connectivity where $p < 0.05$ (FDR corrected). The count of subjects from each group that had at least one window in each state is also shown. HCs show significantly higher connectivity strengths than SZs where t-values are positive (red) and the opposite when it is negative (blue).

3. RESULTS

Two groups of states (five states in each group) obtained from K-means and AP clustering are shown in Fig. 3. In all states, there is higher (positive) correlation within VIS and SM domain. There is also close to zero or negative correlation of VIS and SM areas with DMN cingulate and frontal cortices. Compared to the other states, state 1 estimated by both approach has the highest occupancy and therefore may be the most crucial area when comparing the two approaches.

Fig. 4 shows the t-values obtained from the two-sample t-tests on the connectivity passing a significance level of $p < 0.05$ with false discovery rate (FDR) correction for multiple comparisons [23]. The positive (red) t-values indicate that SZ showed lower FC strength than HC, and the negative (blue) t-values indicate that SZ had increased FC compared to HC. Results show that although the average states across subjects were noticeably similar between the two methods, we found more and interesting group differences using the AP clustering. In AP states 1 and 2, SZ group had significant increased connectivity than HC group in subcortical and sensorimotor networks. The finding is supported by results from previous studies [3] but was absent in any of the K-means states in our analysis. Also, across states 1-3, reduced FC strengths in SZ group between AUD, VIS and SM networks was highly conspicuous in AP results, also supported by previous studies [3] [24]. Furthermore, AP identified SZ group's abnormal connectivity in DMN regions in state 1, which was absent in the K-means results [4].

4. DISCUSSION

Taken together, affinity propagation is a promising approach for dFNC analysis. It does not suffer from the shortcomings of K-means. Our work clearly indicates that the AP approach can estimate more meaningful connectivity states than the traditional K-means and provide more informative measures for differentiating SZ and HC. We did not take full advantage of the semi-supervised features of AP algorithm such as specifying preference based on prior information. Defining an AP clustering framework applicable to dFNC data by specifying optimal parameters such as distance measure for similarity matrix and preference seems to be the way forward. It will facilitate higher dimensional analysis of brain organization and advance the study of healthy and diseased brain.

5. REFERENCES

[1] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking Whole-Brain Connectivity Dynamics in the Resting State," *Cerebral Cortex*, vol. 24, no. 3, pp. 663–676, Mar. 2014.

[2] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adal, "The Chronnectome: Time-Varying Connectivity Networks as the Next Frontier in fMRI Data Discovery," *Neuron*, vol. 84, no. 2, pp. 262–274, Oct. 2014.

[3] E. Damaraju, E. A. Allen, A. Belger, J. M. Ford, S. McEwen, D. H. Mathalon, B. A. Mueller, G. D. Pearlson, S. G. Potkin, A. Preda, J. A. Turner, J. G. Vaidya, T. G. van Erp, and V. D. Calhoun, "Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia," *NeuroImage: Clinical*, vol. 5, pp. 298–308, 2014.

[4] Y. Du, G. D. Pearlson, Q. Yu, H. He, D. Lin, J. Sui, L. Wu, and V. D. Calhoun, "Interaction among subsystems within default mode network diminished in schizophrenia patients: A dynamic connectivity approach," *Schizophrenia Research*, vol. 170, no. 1, pp. 55–65, Jan. 2016.

[5] N. Leonardi, W. R. Shirer, M. D. Greicius, and D. Van De Ville, "Disentangling dynamic networks: Separated and joint expressions of functional connectivity patterns in time: Disentangling Dynamic Networks," *Human Brain Mapping*, vol. 35, no. 12, pp. 5984–5995, Dec. 2014.

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

[7] Y. Du, G. D. Pearlson, H. He, L. Wu, J. Chen, and V. D. Calhoun, "Identifying brain dynamic network states via GIG-ICA: Application to schizophrenia, bipolar and schizoaffective disorders," in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, Apr. 2015, pp. 478–481.

[8] R. L. Miller, M. Yaesoubi, J. A. Turner, D. Mathalon, A. Preda, G. Pearlson, T. Adali, and V. D. Calhoun, "Higher Dimensional Meta-State Analysis Reveals Reduced Resting fMRI Connectivity Dynamism in Schizophrenia Patients," *PLOS ONE*, vol. 11, no. 3, pp. e0149849, Mar. 2016.

[9] B. J. Frey and D. Dueck, "Clustering by Passing Messages Between Data Points," *Science*, vol. 315, no. 5814, pp. 972–976, Feb. 2007.

[10] J. M. Shine, O. Koyejo, and R. A. Poldrack, "Temporal metastates are associated with differential patterns of time-resolved connectivity, network topology, and attention," *Proceedings of the National Academy of Sciences*, p. 201604898, Aug. 2016.

[11] T. Ren, W. Zeng, N. Wang, L. Chen, and C. Wang, "A novel approach for fMRI data analysis based on the combination of sparse approximation and affinity propagation clustering," *Magnetic Resonance Imaging*, vol. 32, no. 6, pp. 736–746, July 2014.

[12] J. Zhang, X. Tuo, Z. Yuan, W. Liao, and H. Chen, "Analysis of fMRI Data Using an Integrated Principal Component Analysis and Supervised Affinity Propagation Clustering Approach," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 11, pp. 3184–3196, Nov. 2011.

[13] H. J. Bockholt, M. Scully, W. Courtney, S. Rachakonda, A. Scott, A. Caprihan, J. Fries, R. Kalyanam, J. M. Segall, R. de la Garza, S. Lane, and V. D. Calhoun, "Mining the mind research network: a novel framework for exploring large scale, heterogeneous translational neuroscience research data sources," *Frontiers in Neuroinformatics*, vol. 3, pp. 36, 2010.

[14] L. Freire, A. Roche, and J. F. Mangin, "What is the best similarity measure for motion correction in fMRI time series?," *IEEE transactions on medical imaging*, vol. 21, no. 5, pp. 470–484, May 2002.

[15] A. J. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, Nov. 1995.

[16] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Human brain mapping*, vol. 13, no. 1, pp. 43–53, 2001.

[17] E. A. Allen, E. B. Erhardt, E. Damaraju, W. Gruner, J. M. Segall, R. F. Silva, M. Havlicek, S. Rachakonda, J. Fries, R. Kalyanam, A. M. Michael, A. Caprihan, J. A. Turner, T. Eichele, S. Adelsheim, A. D. Bryan, J. Bustillo, V. P. Clark, Feldstein E., Sarah W., F. Filbey, C. C. Ford, K. Hutchison, R. E. Jung, K. A. Kiehl, P. Kodituwakku, Y. M. Komesu, A. R. Mayer, G. D. Pearlson, J. P. Phillips, J. R. Sadek, M. Stevens, U. Teuscher, R. J. Thoma, and V. D. Calhoun, "A baseline for the multivariate comparison of resting-state networks," *Frontiers in Systems Neuroscience*, vol. 5, pp. 2, 2011.

[18] U. Sakoglu, G. D. Pearlson, K. A. Kiehl, Y. M. Wang, A. M. Michael, and V. D. Calhoun, "A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 23, no. 5-6, pp. 351–366, Dec. 2010.

[19] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, Oct. 2013.

[20] G. Varoquaux, A. Gramfort, J. Poline, and B. Thirion, "Brain covariance selection: better individual functional connectivity models using population prior," in *Advances in Neural Information Processing Systems*, 2010, pp. 2334–2342.

[21] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, July 2008.

[22] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 411–423, Jan. 2001.

[23] Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[24] M Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore, "Functional connectivity and brain networks in schizophrenia," *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, vol. 30, no. 28, pp. 9477–9487, July 2010.