



# Simultaneous Estimation of Low- and High-Order Functional Connectivity for Identifying Mild Cognitive Impairment

Yueying Zhou<sup>1</sup>, Lishan Qiao<sup>1\*</sup>, Weikai Li<sup>1,2</sup>, Limei Zhang<sup>1</sup> and Dinggang Shen<sup>3,4\*</sup>

<sup>1</sup> School of Mathematics, Liaocheng University, Liaocheng, China, <sup>2</sup> College of Information Science and Engineering, Chongqing Jiaotong University, Chongqing, China, <sup>3</sup> Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, <sup>4</sup> Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea

## OPEN ACCESS

### Edited by:

Pedro Antonio Valdes-Sosa,  
Clinical Hospital of Chengdu Brain  
Science Institute, China

### Reviewed by:

Feng Liu,  
Tianjin Medical University General  
Hospital, China  
Yifeng Wang,  
University of Electronic Science and  
Technology of China, China

### \*Correspondence:

Lishan Qiao  
qiaolishan@lcu.edu.cn  
Dinggang Shen  
dgshen@med.unc.edu

**Received:** 04 December 2017

**Accepted:** 22 January 2018

**Published:** 06 February 2018

### Citation:

Zhou Y, Qiao L, Li W, Zhang L and  
Shen D (2018) Simultaneous  
Estimation of Low- and High-Order  
Functional Connectivity for Identifying  
Mild Cognitive Impairment.  
*Front. Neuroinform.* 12:3.  
doi: 10.3389/fninf.2018.00003

Functional connectivity (FC) network has been becoming an increasingly useful tool for understanding the cerebral working mechanism and mining sensitive biomarkers for neural/mental disease diagnosis. Currently, Pearson's Correlation (PC) is the simplest and most commonly used scheme in FC estimation. Despite its empirical effectiveness, PC only encodes the low-order (i.e., second-order) statistics by calculating the pairwise correlations between network nodes (brain regions), which fails to capture the high-order information involved in FC (e.g., the correlations among different edges in a network). To address this issue, we propose a novel FC estimation method based on Matrix Variate Normal Distribution (MVND), which can capture both low- and high-order correlations simultaneously with a clear mathematical interpretability. Specifically, we first generate a set of BOLD subseries by the sliding window scheme, and for each subseries we construct a temporal FC network by PC. Then, we employ the constructed FC networks as samples to estimate the final low- and high-order FC networks by maximizing the likelihood of MVND. To illustrate the effectiveness of the proposed method, we conduct experiments to identify subjects with Mild Cognitive Impairment (MCI) from Normal Controls (NCs). Experimental results show that the fusion of low- and high-order FCs can generally help to improve the final classification performance, even though the high-order FC may contain less discriminative information than its low-order counterpart. Importantly, the proposed method for simultaneous estimation of low- and high-order FCs can achieve better classification performance than the two baseline methods, i.e., the original PC method and a recent high-order FC estimation method.

**Keywords:** functional connectivity, high-order network, matrix variate normal distribution, mild cognitive impairment, disease diagnosis

## INTRODUCTION

Functional connectivity (FC) network, calculated by resting-state functional magnetic resonance imaging (rs-fMRI) (Liu et al., 2008), has become an increasingly useful tool for understanding the working mechanism of the brain and providing informative biomarkers for diagnosing some neural/mental disorders, such as autism spectrum disorder (Wee et al., 2016b; Li et al., 2017),

major depressive disorder (Greicius et al., 2007; He et al., 2016), obsessive compulsive disorder (Admon et al., 2012), schizophrenia (Zhou et al., 2007; Ganella et al., 2016), social anxiety disorder (Liu et al., 2015a,b), Alzheimer's disease (Zhu et al., 2015; Wang et al., 2017), and its early stage, i.e., mild cognitive impairment (MCI) (Wee et al., 2012; Yu et al., 2017).

In view of its great potential, how to construct high-quality FC networks comes to a key issue. Theoretically, we can treat the FC network as a graph, where the nodes correspond to different brain regions or, more generally, the regions-of-interest (ROIs), while the edges correspond to the pairwise FCs between these nodes. In other words, FC network can be seen as a combination of the node set and the edge set. Currently, researchers have proposed a series of FC network modeling methods (Smith et al., 2011, 2013), among which Pearson's Correlation (PC) is the simplest and the most popular way. Although it has been successfully applied in FC estimation, PC can only capture the low-order (or second-order) statistical information by calculating the pairwise correlations between the network nodes (e.g., ROIs in this paper).

In practice, some high-order statistics (e.g., the correlations among different edges) may also offer additional and useful information for FC analysis (Plis et al., 2014; Chen et al., 2016; Zhang et al., 2016). To make this easy to understand, we consider the traffic network as an analogy, where the cities are regarded as nodes, the roads are edges between the city nodes, and the traffic flow can be used as a measure of the dynamic weight on each edge. In this example, the road (or traffic flow) provides the low-order connection information of the network. On the other hand, however, there may exist some relationships among different roads/edges. For instance, a traffic jam on one road tends to affect the traffic flow on another road. Compared to the edges that measuring the relationship between nodes, the relationship between edges is expected to provide some high-order connection information for a network system.

Recently, some high-order methods have been developed for estimating FCs (Plis et al., 2014; Chen et al., 2016, 2017; Zhang et al., 2016). For example, Plis et al. used mutual information to investigate the nonlinear interactions among brain regions (Plis et al., 2014); Zhang et al. proposed a high-order FC network estimation method based on the topographical information and applied it to MCI detection (Zhang et al., 2016); Chen et al. proposed a scheme for estimating high-order FCs based on sliding windows, and empirically verified that it can achieve better accuracy than the low-order counterparts for early MCI identification (Chen et al., 2016). However, modeling such a high-order FC in Chen et al. (2016) involves a consistent amount of parameters, thus easily leading to an over-fitting problem given limited training data. In addition, the work of Chen et al. (2016) is heuristic without support from mathematical models, and thus cannot make it clear where the high-order relationship (i.e., correlation's correlation) is theoretically derived from.

To address these problems, in this paper we put forward a novel high-order FC network estimation method based on Matrix Variate Normal Distribution (MVND) (Gupta and Nagar, 2000). MVND *not only* provides a clear theoretical explanation for high-order FC networks, *but also* achieves both low- and high-order networks simultaneously in a unified framework. In

particular, a sliding window approach is first used to generate a sequence of overlapping time subseries. For each subseries, the traditional low-order FC network is constructed by PC. Then, the so-constructed FC networks are used as samples to estimate the final low- and high-order FC networks by maximizing the likelihood function of MVND. Especially, we adopt Kronecker product as a novel way to reduce the number of parameters (see section The Proposed Method for details). To validate the effectiveness of the proposed method, we perform experiments to identify MCI subjects from Normal Controls (NCs), based on the estimated low- and high-order FC networks. The experimental results illustrate that the proposed approach can achieve better performance than the baseline method, and both low- and high-order information are generally helpful for classification. Both preprocessed data and codes can be downloaded in <https://github.com/Zhouyy92/high-order-based-on-MVND/>. More specifically, the main contributions of our work can be summarized as follows:

- 1) To our best knowledge, this is the first work that adopts MVND for estimating high-order FC networks. MVND *not only* provides a clear mathematical definition of high-order network, *but also* can simultaneously estimate low- and high-order FC networks in a single model.
- 2) A new finding of this paper is that the low-order FC tends to contain more discriminative information than its high-order counterpart, which is exactly the opposite conclusion of Chen et al. (2016). However, similar to Chen et al.'s work, we also note that the fusion of low- and high-order FCs can generally improve the identification accuracy to some extent, indicating that there is some useful information in the high-order FCs for discrimination.

The rest of this paper is organized as follows. In section Materials and Methods, we introduce the materials and propose our method. In section Results, we evaluate the proposed method in identifying MCI subjects from NCs. In section Discussion, we discuss our findings based on the experimental results. In section Conclusion, we conclude the whole paper.

## MATERIALS AND METHODS

### Data Acquisitions and Processing

Totally, 137 participants, including 68 MCI patients and 69 NCs from Alzheimer's Disease Neuroimaging Initiative (ADNI)<sup>1</sup> dataset (Jack et al., 2008), are used in this study. The observed rs-fMRI images were scanned by 3.0T Philips scanners with the following parameters: TR/TE is 3,000/30 mm, flip angle is 80° , imaging matrix size is 64 × 64 with 48 slices and 140 volumes, and voxel thickness is 3.3 mm.

The acquired rs-fMRI data was processed by SPM8<sup>2</sup> toolbox based on the well-accepted pipeline. The first three volumes of each subject were removed for signal stabilization. Then, the remaining 137 volumes were corrected for different slice acquisition timing and head motion (Poldrack et al., 2011).

<sup>1</sup><http://adni.loni.ucla.edu>

<sup>2</sup><http://www.fil.ion.ucl.ac.uk/spm/>

To further reduce the influences of nuisance signals, regression of ventricle and white matter signals as well as six head-motion profiles were conducted. Based on the Automated Anatomical Labeling (AAL) template atlas (Tzourio-Mazoyer et al., 2002), the pre-processed Blood Oxygen Level Dependent (BOLD) time series signals were partitioned into 116 ROIs. Prior to FC estimation, the mean rs-fMRI time series of each ROI was band-pass filtered from 0.01 to 0.08 Hz. Finally, the mean time series was put into a data matrix  $\mathbf{X} \in R^{137 \times 116}$ , which will be used for FC network estimation.

## Functional Connectivity Network Estimation

### Baseline Method

According to a recent review (Smith et al., 2013), PC is the most widely-used method for estimating FC networks. Therefore, in this paper, we select PC as the baseline and building block for developing our method, even though the idea can *in principle* be used in any correlation-based FC construction methods. The PC-based FC is defined as follows:

$$W_{ij}^{(PC)} = \frac{(x_i - \bar{x}_i)^T(x_j - \bar{x}_j)}{\sqrt{(x_i - \bar{x}_i)^T(x_i - \bar{x}_i)}\sqrt{(x_j - \bar{x}_j)^T(x_j - \bar{x}_j)}} \quad (1)$$

where  $x_i \in R^V (i = 1, 2, \dots, P)$  is the time series associated with the  $i$ th ROI,  $V = 137$  is the total number of temporal image volumes,  $P = 116$  is the number of ROIs, and  $\bar{x}_i \in R^V$  is the corresponding mean vector of  $x_i$ . Without loss of generality, in this paper we suppose that  $x_i$  is centralized by  $x_i - \bar{x}_i$  and normalized by  $\sqrt{(x_i - \bar{x}_i)^T(x_i - \bar{x}_i)}$ . Then, PC can be simply expressed as  $W_{ij}^{(PC)} = x_i^T x_j$ , or, an equivalent matrix form,

$$\mathbf{W}^{PC} = \mathbf{X}^T \mathbf{X} \quad (2)$$

where  $\mathbf{X} = [x_1, x_2, \dots, x_P] \in R^{V \times P}$  is the data matrix.

Given the fact that the BOLD time series signals commonly contain noises, the original PC-based FC network tends to be dense (Fornito et al., 2016). To alleviate this problem, the thresholding operation is generally employed to filter out the noisy or weak connections. Please refer to, for example, section 3.2.1 in Fornito et al. (2016) for a detailed discussion on different thresholding schemes.

### The Proposed Method

In this section, we introduce the new FC estimation scheme based on MVND that can encode both low- and the high-order correlations in a single framework. As a result, we can model FC from two different views.

In particular, we suppose the low-order FC between the  $i$ th and  $j$ th ROIs is a random variable  $w_{ij}$  that follows the normal distribution, and thus the corresponding FC network is a random matrix  $\mathbf{W} = (w_{ij})_{P \times P}$  that has the multivariate normal distribution. That is,

$$\mathbf{W} \sim N(\mathbf{M}, \mathbf{\Sigma}) \quad (3)$$

where  $\mathbf{M} \in R^{P \times P}$  is the population mean or mathematical expectation of  $\mathbf{W}$ , and  $\mathbf{\Sigma} \in R^{P^2 \times P^2}$  is the population covariance

matrix of  $\mathbf{W}$ . Note that the entries in  $\mathbf{M}$  measure the relationship between the network nodes (i.e., ROIs), still corresponding to the low-order FC; while the entries in  $\mathbf{\Sigma}$  describe the relationship between the edges, corresponding to the higher-order FC, as shown in **Figure 1**.

Although the  $\mathbf{\Sigma}$  in Equation (3) gives a clear definition of the high-order FC, the estimation of a  $P^2 \times P^2$  matrix is challenging, since it contains a consistent amount of parameters. For example,  $P = 116$  in this study can result in  $P^2 \times (P^2 - 1) / 2 \approx 9 \times 10^7$  free parameters. In Chen et al. (2016), they first reduce the number of the edges by clustering them into some groups, and then consider these groups as new “nodes” for constructing the high-order FC network. However, such a scheme has no clear mathematical explanation or principled way to determine the cluster size.

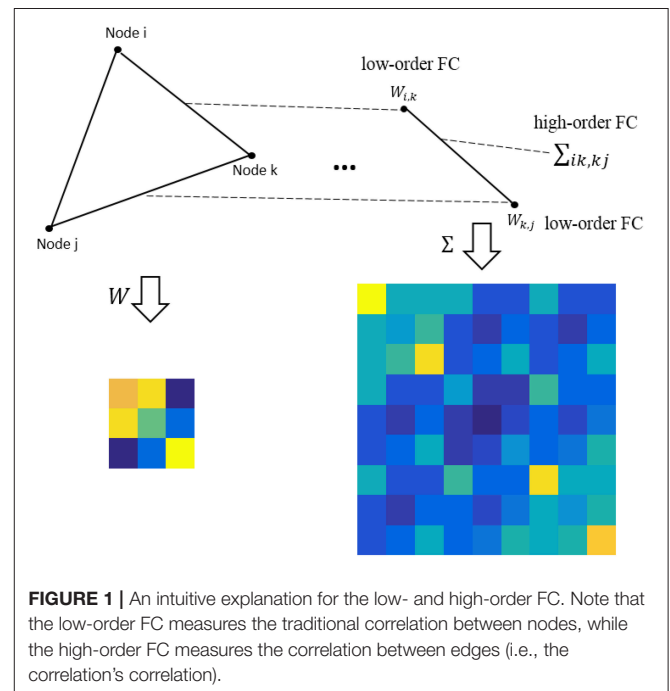
Therefore, in this paper we propose a new strategy to eliminate the difficulty of estimating  $\mathbf{\Sigma}$  by assuming that it has a form of Kronecker product decomposition (Gupta and Nagar, 2000), i.e.,  $\mathbf{\Sigma} = \mathbf{\Omega}_1 \otimes \mathbf{\Omega}_2$ . In other words, the random network matrix  $\mathbf{W}$  follows the MVND:

$$\mathbf{W} \sim N(\mathbf{M}, \mathbf{\Omega}_1 \otimes \mathbf{\Omega}_2) \quad (4)$$

More specifically, the probability density function of MVND is defined (Gupta and Nagar, 2000) as follows:

$$f(\mathbf{W}) = (2\pi)^{-\frac{P^2}{2}} \det(\mathbf{\Omega}_1)^{-\frac{P}{2}} \det(\mathbf{\Omega}_2)^{-\frac{P}{2}} \times e^{-tr\left(\frac{1}{2}\mathbf{\Omega}_1^{-1}(\mathbf{W}-\mathbf{M})\mathbf{\Omega}_2^{-1}(\mathbf{W}-\mathbf{M})^T\right)} \quad (5)$$

where  $\mathbf{\Omega}_1, \mathbf{\Omega}_2 \in R^{P \times P}$  are positive semi-definite, indicating the column and row covariance matrices of  $\mathbf{W}$ , respectively. The  $\det(\cdot)$  and  $tr(\cdot)$  denote the determinant and trace operators of a matrix, respectively. In section Discussion, we will give a



**FIGURE 1** | An intuitive explanation for the low- and high-order FC. Note that the low-order FC measures the traditional correlation between nodes, while the high-order FC measures the correlation between edges (i.e., the correlation's correlation).

further discussion on the meaning of the Kronecker product decomposition in MVND by comparing with clustering scheme in Chen et al. (2016).

In this study, we mainly focus on the *undirected* FC network, meaning that the edge weight matrix  $\mathbf{W}$  is symmetric, and therefore the corresponding column covariance matrix is equal to the row covariance matrix, i.e.,  $\mathbf{\Omega}_1 = \mathbf{\Omega}_2$ . Without loss of generality, we let  $\mathbf{\Omega} = \mathbf{\Omega}_1 = \mathbf{\Omega}_2$  for simplifying the mathematical expression. Since the  $\mathbf{\Omega}$  has the size of  $P \times P$ , the number of the parameters need to be estimated in  $\mathbf{\Sigma}$  reduces from  $P^2 \times (P^2 - 1) / 2$  to  $P \times (P - 1) / 2$ , and the  $\mathbf{\Omega}$  includes all information for tracking back  $\mathbf{\Sigma}$  under the MVND assumption. In what follows, we develop a two-step scheme for estimating the parametric matrices  $\mathbf{M}$  (corresponding to the low-order FC network) and  $\mathbf{\Omega}$  (corresponding to the high-order FC network).

**Steps 1 Sampling: Construct FC network series by sliding windows**

In order to estimate  $\mathbf{M}$  and  $\mathbf{\Omega}$ , we first use the sliding window approach to generate samples, as shown in the top two panels of **Figure 2**. In particular, we suppose  $N$  is the width of the sliding windows and  $s$  is the step size of two adjacent windows, and thus we can obtain  $K = \lfloor (V - N) / s \rfloor + 1$  windows, where  $V$  is the number of image volumes. Based on the subseries in each window, we calculate the FC networks via PC. More rigorously, we define  $\mathbf{x}_i^k \in R^N, (k = 1, \dots, K)$ , to denote the  $k$ th subseries associated with the  $i$ th ROI. Then, we concatenate  $\mathbf{x}_i^k$  together to obtain a data matrix  $\mathbf{X}^{(k)} = [\mathbf{x}_1^k, \mathbf{x}_2^k, \dots, \mathbf{x}_P^k] \in R^{N \times P}$  corresponding to the  $k$ th window. Similar to PC defined in Equation (2) with centralized and normalized procedures, the  $k$ th temporal FC network  $\mathbf{W}^{(k)}$  is constructed as follows:

$$\mathbf{W}^{(k)} = (\mathbf{X}^{(k)})^T \mathbf{X}^{(k)} \tag{6}$$

As a result, we can get  $K$  samples for estimating  $\mathbf{M}$  and  $\mathbf{\Omega}$ .

**Steps 2 MLE: Estimate  $\mathbf{M}$  and  $\mathbf{\Omega}$  for low- and high-order FC networks**

Based on the  $K$  temporal FC networks as samples, we estimate the low- and high-order FC networks (corresponding to  $\mathbf{M}$  and  $\mathbf{\Omega}$ , respectively) by the maximum likelihood estimation (MLE) theory of MVND (Dutilleul, 1999; Zhang and Schneider, 2010). More specifically, the MLE of the mean  $\mathbf{M}$  is

$$\mathbf{M} = \frac{1}{K} \sum_{k=1}^K \mathbf{W}^{(k)} \tag{7}$$

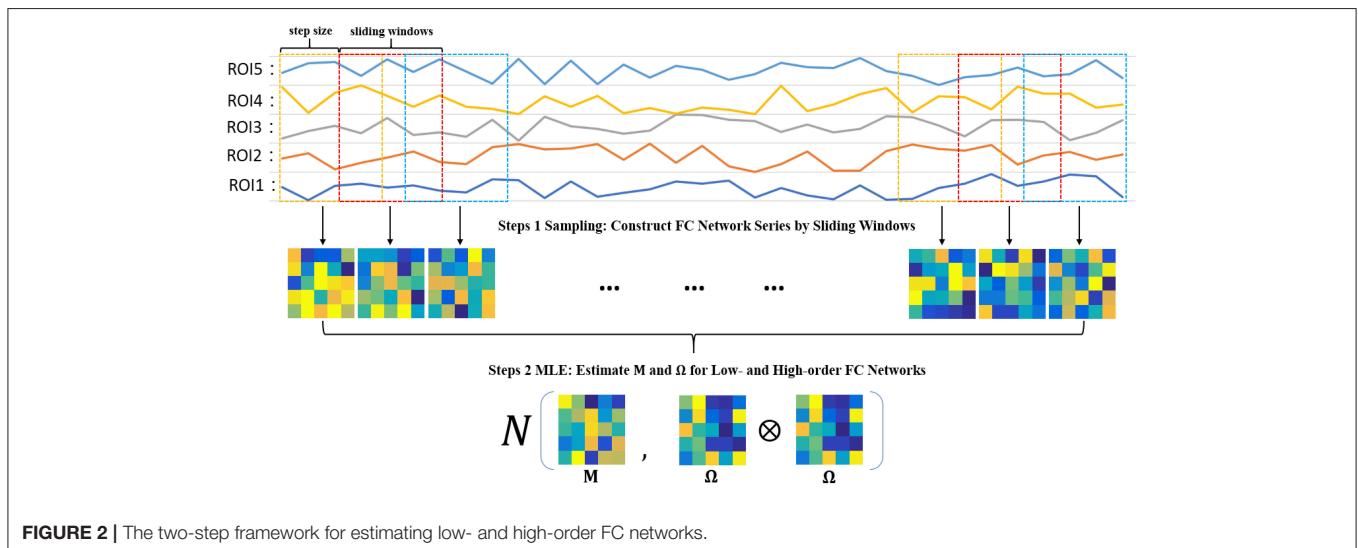
and the MLE of  $\mathbf{\Omega}$  can be achieved by the following iteration formula,

$$\mathbf{\Omega} = \frac{1}{KP} \sum_{k=1}^K (\mathbf{W}^{(k)} - \mathbf{M}) \mathbf{\Omega}^{-1} (\mathbf{W}^{(k)} - \mathbf{M})^T \tag{8}$$

Note that, however, the estimation of  $\mathbf{\Omega}$  represents the covariance matrix rather than the correlation matrix, and therefore we employ the normalized  $\mathbf{\Omega}$  as the high-order FC network in our study. As such, the algorithm for solving Equations (7, 8) is summarized in **Table 1**.

**FC Network Evaluation**

In order to verify the performance of the estimated FCs, we utilize the low-order FC network  $\mathbf{M}$  (LoM), high-order FC network  $\mathbf{\Omega}$  (HiO), and their fusion (FuMO), respectively, to train classifiers for MCI diagnosis. Of note, FuMO is fused by a linear combination of LoM and HiO. Since it is hard to determine the combination coefficient in practice, in this paper, we simply fuse them by  $0.5 \times (\mathbf{M} + \mathbf{\Omega})$ . In addition, we select the original PC and recent high-order method called HON in Chen et al. (2016), as baseline methods to make a comparison. Based on the estimated FC networks from different methods, we employ the  $t$ -test ( $p < 0.05$ ) for feature selection, and the linear Support Vector Machine (SVM) (Chang and Lin, 2011) with default  $C = 1$  for classification.



**FIGURE 2 |** The two-step framework for estimating low- and high-order FC networks.



**TABLE 1** | Algorithm of MVND-based low- and high-order FC estimation.

Input: $X$ //observed data
Output: $M$ and $\Omega$ //low- and high-order FC
Apply sliding windows to obtain more samples $X^{(k)}$ and PC to construct temporal low-order FC $W^{(k)} = (X^{(k)})^T X^{(k)}$ ;
$M = \frac{1}{K} \sum_{k=1}^K W^{(k)}$
Initialize $\Omega = I$ // identity matrix
while not converge
$\Omega \leftarrow \frac{1}{KP} \sum_{k=1}^K (W^{(k)} - M)\Omega^{-1}(W^{(k)} - M)^T$
end

Due to the limited subjects, in this paper, we use the nested leave-one-out cross validation (LOOCV) to estimate the classification performance, in which only one participant is left out for testing while the others are adopted for training a classifier and obtaining the optimal parameters. In terms of the thresholding parameter of FC networks, we empirically employ 11 sparsity levels ranging in [1%, 10%, ..., 90%, 100%] for all the methods. For instance, 10% means that 90% of the weak edges are filtered out from the FC networks, while 100% means all the edges are reserved. We determine the optimal thresholding parametric value using an inner LOOCV procedure on the training dataset.

## RESULTS

### Classification Performance

In our experiments, we adopt accuracy, sensitivity and specificity (Sokolova et al., 2006) as performance metrics to evaluate different FC estimation methods. Their mathematical definitions are given as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (11)$$

where  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  indicate true positive, true negative, false positive and false negative, respectively.

In **Table 2**, we first report the best results for MCI identification with specific parametric values, and then, in the next section Sensitivity to Network Modeling Parameters, we discuss the influence of different parametric values on the final classification performance. To be specific, for PC, no free parameter is involved; for the proposed method, the width of the sliding window is  $N = 50$ , and the step size is  $s = 8$ ; for HON method,  $N = 110$ ,  $s = 1$ , and the cluster size is 300. As can be seen from **Table 2**, the proposed method significantly outperforms the two baseline methods.

**TABLE 2** | Comparison on MCI classification performance with different methods.

Method	Accuracy	Sensitivity	Specificity
PC	0.7956	0.7647	0.8261
HON (Chen et al., 2016)	0.8207	0.8194	0.8377
LoM	0.9051	0.9118	0.8986
HiO	0.8394	0.8235	0.8551
FuMO	0.8905	0.8676	0.9130

### Sensitivity to Network Modeling Parameters

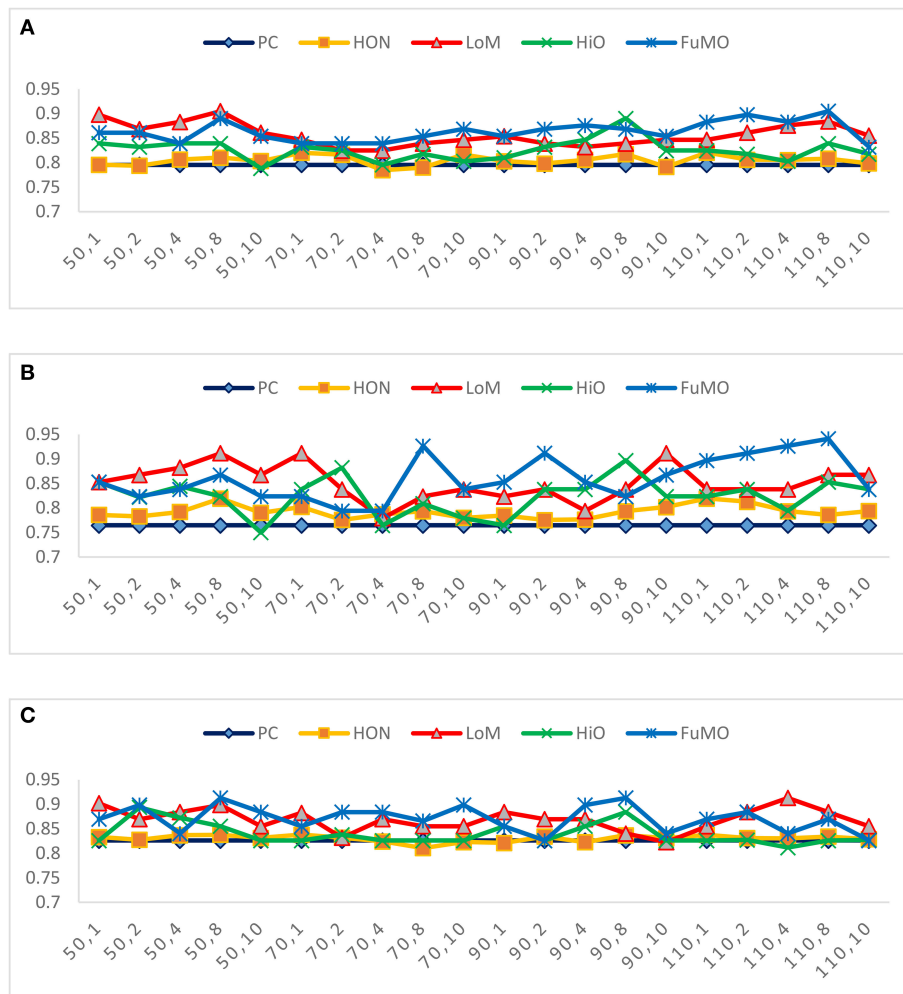
In general, the free parameters involved in the FC network estimation methods have a big influence on the ultimate classification performance. In the proposed framework, there are two free parameters, including the width of sliding windows ( $N$ ) and the step size ( $s$ ). To evaluate the sensitivity of the proposed framework with respect to  $N$  and  $s$ , we repeat MCI identification experiments based on different combinations of window width ( $N = 50, 70, 90, 110$ ) and step size ( $s = 1, 2, 4, 8, 10$ ). **Figure 3** reports the accuracy, sensitivity and specificity of each involved method *with regard to* different combinations of  $N$  and  $s$  shown on the horizontal axis. For example, the “50, 1” denotes that the window width is 50, and the step size is 1.

In **Figure 3**, we observe that the proposed method consistently outperforms the two baseline methods on most of the parameter combinations. Additionally, compared with conclusion of Chen et al. (2016) that high-order FC networks work better than low-order FC networks, we actually have an opposite finding that low-order FC networks are instead more discriminating than the high-order counterparts under most cases. In the next section, we will further discuss the possible reasons for this new finding.

## DISCUSSION

In this paper, we propose a new FC network estimation framework based on MVND that can simultaneously capture low- and high-order correlation information in data. The proposed method is validated on ADNI dataset by an MCI identification task. According to the experimental results, we have the following discussions:

- 1) In general, the performance of the proposed method outperforms the baseline methods, including the original PC method and the recently proposed high-order method (HON) in Chen et al. (2016). Compared with PC, the proposed method encodes the temporal information by the sliding window scheme, and integrates the information based on MVND. Since the brain has an inherent dynamic property, an effective modeling of the dynamic temporal information in FC network is expected to improve the final performance (Marusak et al., 2016; Wee et al., 2016a; Liu et al., 2017; Park et al., 2017). In this paper, we simply employ the sliding windows, the most popular method (Hutchison et al., 2013; Preti et al., 2017), to capture dynamic temporal information. However, signals in the sliding windows can



**FIGURE 3 |** The (A) accuracy, (B) sensitivity, and (C) specificity of each involved method with respect to different combinations of window width and step size.

be easily influenced by some non-stationary noise, thus resulting in sudden changes of FC across the brain regions (Hutchison et al., 2013); on the other hand, it is a dilemma to choose an appropriate window width or step size, since the window width should be short enough to capture short-term fluctuations while long enough to allow robust FC estimation (Sakoglu et al., 2010). Especially, in terms of short windows, it may contain less cycles of lower frequency, resulting in less stable measures (Wang et al., 2018). Therefore, in our experiments, we investigate different window sizes in a large range, and empirically found that most methods are sensitive to this parameter. In practice, we need to select window width carefully toward better understanding of dynamics in brains.

- 2) Interestingly, we find that the proposed low-order FC is generally more discriminative than its high-order counterpart, which is contrary to the conclusion in Chen et al. (2016). We argue that our result is relatively reasonable, since the low-order FC (corresponding to the mean of the random variables) determine the main tendency of the data,

while the high-order FC (corresponding to the covariance of the random variables) only capture the spread of the data and could be more noisy as well. Of course, the high-order FCs tend to include some informative structures for MCI identification, because, as reported in **Table 2** and **Figure 3**, HiO generally performs better than PC. In addition, the simple combination of low- and high-order FC networks (i.e., FuMO) can achieve better performance than a single-view network. More specifically, as shown in **Figure 3**, it has 60% possibility for FuMO to obtain the best accuracy, while 55 and 60% possibilities for sensitivity and specificity, respectively.

- 3) From the experimental results, we also find that the estimated FC networks (including the low-order, high-order, and their combination) consistently outperform HON in Chen et al. (2016). In what follows, we discuss the similarities and differences between HON and our method, aiming to help understand why their method does not work well. In particular, the two methods share the same sliding window step for generating network samples. For the high-order FC

network estimation step, however, HON method applied the clustering algorithm to divide the edges of a network into different groups for reducing the computation/estimation cost. Such a scheme actually destroys the original structure of the network, and also introduces a new free parameter (i.e., cluster size) that can cause over-fitting problem. Additionally, HON is heuristic without a support of mathematical theory, and thus cannot provide a clear definition of the high-order correlation. In contrast, based on the MVND assumption, our proposed method actually takes the edges linked to the same node (i.e., ROI) as a cluster, i.e., 116 clusters in this study, which *not only* keep the natural structure of the network, *but also* avoid the issue of selecting cluster size.

## CONCLUSION

In this paper, we develop a novel PC-based FC estimation framework with the assumption that the network edge weights follow the MVND. The proposed method is simple, has a rigorous mathematical model, and can capture both low- and high-order FCs of the brain network simultaneously. The experiments on MCI identification show that our proposed method outperforms the original PC method and a recent

high-order FC estimation method. On the other hand, although we design our method based on PC, the idea can be used in any correlation-based FC estimation methods. In the future, we plan to generalize the MVND-based scheme to the partial correlation-based FC estimation problem.

## AUTHOR CONTRIBUTIONS

DS: Proposed the idea of high-order FC and provided the preprocessed rs-fMRI data; LQ: Proposed the mathematical models for simultaneously estimating low- and high-order FC networks; YZ, LZ, and WL: Designed the procedures of MCI evaluation experiments; All authors developed the estimation algorithm and contributed to the preparation of the article, figures, and charts.

## ACKNOWLEDGMENTS

This work is partly supported by National Natural Science Foundation of China (61300154, 61402215), Chongqing Graduate Student Research Innovation Project (CYS16183), and NIH grants (EB022880, AG041721, AG049371, and AG042599).

## REFERENCES

- Admon, R., Bleich-Cohen, M., Weizmant, R., Poyurovsky, M., Faragian, S., and Hendler, T. (2012). Functional and structural neural indices of risk aversion in obsessive-compulsive disorder (OCD). *Psychiatry Res.* 203, 207–213. doi: 10.1016/j.psychres.2012.02.002
- Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199
- Chen, X., Zhang, H., Gao, Y., Wee, C. Y., Li, G., Shen, D., et al. (2016). High-order resting-state functional connectivity network for MCI classification. *Hum. Brain Mapp.* 37, 3282–3296. doi: 10.1002/hbm.23240
- Chen, X., Zhang, H., Lee, S. W., and Shen, D. (2017). Hierarchical high-order functional connectivity networks and selective feature fusion for MCI classification. *Neuroinformatics* 15, 271–284. doi: 10.1007/12021-017-9330-4
- Dutilleul, P. (1999). The mle algorithm for the matrix normal distribution. *J. Stat. Comput. Simul.* 64, 105–123. doi: 10.1080/00949659908811970
- Fornito, A., Zalesky, A., and Bullmore, E. (2016). *Fundamentals of Brain Network Analysis*. London; San Diego, CA; Cambridge: Academic Press.
- Ganella, E. P., Bartholomeusz, C. F., Seguin, C., Whittle, S., Bousman, C., Phassouliotis, C., et al. (2016). Functional brain networks in treatment-resistant schizophrenia. *Schizophr Res.* 184, 73–81. doi: 10.1016/j.schres.2016.12.008
- Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., et al. (2007). Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol. Psychiatry* 62, 429–437. doi: 10.1016/j.biopsych.2006.09.020
- Gupta, A. K., and Nagar, D. K. (2000). *Matrix Variate Distributions*. Boca Raton, FL; London; New York, NY; Washington, DC: Chapman & Hall/CRC.
- He, H., Yu, Q., Du, Y., Vergara, V., Victor, T. A., Drevets, W. C., et al. (2016). Resting-state functional network connectivity in prefrontal regions differs between unmedicated patients with bipolar and major depressive disorders. *J. Affect. Disord.* 190, 483–493. doi: 10.1016/j.jad.2015.10.042
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079
- Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691. doi: 10.1002/jmri.21049
- Li, W., Zhang, L., Qiao, L., and Shen, D. (2017). Remodeling Pearson's correlation for functional brain network estimation and autism spectrum disorder identification. *Front. Neuroinform.* 11:55. doi: 10.3389/fninf.2017.00055
- Liu, F., Guo, W., Fouché, J. P., Wang, Y., Wang, W., Ding, J., et al. (2015a). Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Struct. Funct.* 220, 101. doi: 10.1007/s00429-013-0641-4
- Liu, F., Wang, Y., Li, M., Wang, W., Li, R., Zhang, Z., et al. (2017). Dynamic functional network connectivity in idiopathic generalized epilepsy with generalized tonic-clonic seizure. *Hum. Brain Mapp.* 38, 957–973. doi: 10.1002/hbm.23430
- Liu, F., Zhu, C., Wang, Y., Guo, W., Li, M., Wang, W., et al. (2015b). Disrupted cortical hubs in functional brain networks in social anxiety disorder. *Clin. Neurophysiol.* 126, 1711–1716. doi: 10.1016/j.clinph.2014.11.014
- Liu, Y., Wang, K., Chunshui, Y. U., He, Y., Zhou, Y., Liang, M., et al. (2008). Regional homogeneity, functional connectivity and imaging markers of Alzheimer's disease: a review of resting-state fMRI studies. *Neuropsychologia* 46:1648. doi: 10.1016/j.neuropsychologia.2008.01.027
- Marusak, H. A., Calhoun, V. D., Brown, S., Crespo, L. M., Sala-Hamrick, K., Gotlib, I. H., et al. (2016). Dynamic functional connectivity of neurocognitive networks in children. *Hum. Brain Mapp.* 38, 97–108. doi: 10.1002/hbm.23346
- Park, B. Y., Moon, T., and Park, H. (2017). Dynamic functional connectivity analysis reveals improved association between brain networks and eating behaviors compared to static analysis. *Behav. Brain Res.* 337:114. doi: 10.1016/j.bbr.2017.10.001
- Plis, S. M., Jing, S., Lane, T., Roy, S., Clark, V. P., Potluru, V. K., et al. (2014). High-order interactions observed in multi-task intrinsic networks are dominant indicators of aberrant brain function in schizophrenia. *Neuroimage* 102, 35–48. doi: 10.1016/j.neuroimage.2013.07.041
- Poldrack, R., Mumford, J., and Nichols, T. (2011). *Handbook of Functional MRI Data Analysis*. New York, NY: Cambridge University Press.

- Preti, M. G., Bolton, T. A. W., and Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art and perspectives. *Neuroimage* 160, 41–54. doi: 10.1016/j.neuroimage.2016.12.061
- Sakoglu, Ü., Pearson, G. D., Kiehl, K. A., Wang, Y. M., Michael, A. M., and Calhoun, V. D. (2010). A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia. *Magn. Reson. Mater. Phys. Biol. Med.* 23, 351–366. doi: 10.1007/s10334-010-0197-8
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for FMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., et al. (2013). Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* 17, 666–682. doi: 10.1016/j.tics.2013.09.016
- Sokolova, M., Japkowicz, N., and Szpakowicz, S. (2006). “Beyond accuracy, F-Score and ROC: a family of discriminant measures for performance evaluation,” in *AI 2006: Advances in Artificial Intelligence: 19th Australian Joint Conference on Artificial Intelligence*, eds A. Sattar and B.-H. Kang (Hobart, TAS; Berlin; Heidelberg: Springer Berlin Heidelberg), 1015–1021.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Wang, Y., Zhu, L., Zou, Q., Cui, Q., Liao, W., Duan, X., et al. (2018). Frequency dependent hub role of the dorsal and ventral right anterior insula. *Neuroimage* 165, 112–117. doi: 10.1016/j.neuroimage.2017.10.004
- Wang, Z., Zhu, X., Adeli, E., Zhu, Y., Nie, F., Munsell, B., et al. (2017). Multi-modal classification of neurodegenerative disease by progressive graph-based transductive learning. *Med. Image Anal.* 39:218. doi: 10.1016/j.media.2017.05.003
- Wee, C. Y., Yang, S., Yap, P. T., and Shen, D. (2016a). Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging Behav.* 10, 342–356. doi: 10.1007/s11682-015-9408-2
- Wee, C. Y., Yap, P. T., and Shen, D. (2016b). Diagnosis of autism spectrum disorders using temporally distinct resting-state functional connectivity networks. *CNS Neurosci. Ther.* 22, 212–219. doi: 10.1111/cns.12499
- Wee, C.-Y., Yap, P.-T., Zhang, D., Wang, L., and Shen, D. (2012). “Constrained sparse functional connectivity networks for MCI classification,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012: 15th International Conference*, eds N. Ayache, H. Delingette, P. Golland, and K. Mori (Nice; Berlin; Heidelberg: Springer Berlin Heidelberg), 212–219. doi: 10.1007/978-3-642-33418-4\_27
- Yu, R., Zhang, H., An, L., Chen, X., Wei, Z., and Shen, D. (2017). Connectivity strength-weighted sparse group representation-based brain network construction for MCI classification. *Hum. Brain Mapp.* 38, 2370–2383. doi: 10.1002/hbm.23524
- Zhang, H., Chen, X., Shi, F., Li, G., Kim, M., Giannakopoulos, P., et al. (2016). Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment. *J. Alzheimers Dis.* 54, 1095–1112. doi: 10.3233/JAD-160092
- Zhang, Y., and Schneider, J. G. (2010). “Learning multiple tasks with a sparse matrix-normal penalty,” in *Conference on Neural Information Processing Systems 2010* (Vancouver, BC: DBLP).
- Zhou, Y., Liang, M., Tian, L., Wang, K., Hao, Y., Liu, H., et al. (2007). Functional disintegration in paranoid schizophrenia using resting-state fMRI. *Schizophr. Res.* 97:194. doi: 10.1016/j.schres.2007.05.029
- Zhu, X., Suk, H. I., Wang, L., Lee, S. W., and Shen, D., (2015). A novel relational regularization feature selection method for joint regression and classification in AD diagnosis. *Med. Image Anal.* 75, 570–577. doi: 10.1016/j.media.2015.10.008

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Zhou, Qiao, Li, Zhang and Shen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.