Low Rank plus Sparse Decomposition of fMRI Data with Applications in Alzheimer's Disease

by

Fangfang Fu

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Statistics

Department of Mathematical and Statistical Sciences

University of Alberta

Abstract

Functional brain connectivity plays an important role on understanding how human brain functions and neuropsychological diseases such as Autism, Attention-deficit hyperactivity disorder (ADHD), and Alzheimer's disease (AD). Functional Magnetic Resonance Imaging (fMRI) is one of the most powerful techniques to construct functional brain connectivity. However, the presence of potential outliers in fMRI BOLD signals might lead to unreliable results in construction. Another challenge of most existing connectivity construction methods is the results might not be stable. In this thesis, we propose a framework which is able to provide robust and stable connectivity. In particular, a low-rank plus sparse ($L + S$) matrix decomposition technique is adapted to decompose the resting state fMRI BOLD signals, where the low-rank matrix L recovers the essential common features from regions of interest (ROIs), and the sparse matrix S catches the sparse individual variability and potential outliers. Based on decomposition, various approaches can be applied to construct functional brain connectivity, such as correlation, partial correlation, Graphical Lasso and others. However, we propose to use the recently developed novel sparse matrix estimation based on concentration inequality. Statistical test for each connection is implemented for differentiating group difference. Through bootstrapping afterwards, we verify whether performing low-rank plus sparse matrix decomposition can achieve more stable and robust results.

We apply our method on Alzheimer's Disease Neuroimaging Initiative (ADNI) data, and compare the results with those based on original BOLD signals. We discover that the methods for building connectivities based on low rank ma-

trices behave better than based on original BOLD signals, in the sense that the former can reveal and identify more significant ROI connections. We find that the recently proposed concentration inequality based method performs better overall compared with correlation, partial correlation, and Graphical Lasso method. We also obtain the first ten most significant connections for differentiating group differences for Alzheimer's disease. Among them, the left hippocampus region and the left cerebellum 7 region is the most significant one, with p value smaller than 0.0005, which is consistent with existing literatures' findings. Through bootstrapping, we verify that performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing functional brain connectivities.

Acknowledgements

First and foremost, I am deeply and sincerely grateful to my supervisor Dr. Linglong Kong. I have been always having his support, both mentally and academically. He is the one who introduced me to the neuroimaging field which I found fascinating. He is the one who always encouraged me to try new things and mine my potential. I learned so many things from him, both in the academical way and in the life attitude way. I feel extremely lucky that I can be his student from the deep of my heart.

Second, I would love to express my deep gratitude to the group studies' organizers Dr. Ivan Mizera and Dr. Linglong Kong, and also the speaker Dr. Adam Kashlak. I learned large amount of knowledge that I did not learn on class before such as deep learning and asymptotic statistics.

Third, I would like to express my deepest gratitude to Dr. Christoph Frei, Dr. Ivan Mizera, Dr. Adam Kashlak, Dr. Guozhong Zhu for being my thesis examining committee members and for going over my thesis.

Fourth, I would love to thank Dr. Adam Kashlak who kindly shared the R codes of concentration inequality-based estimation method with me, and even added comments in the codes in order to help me understand them better.

Fifth, I would love to thank Wei Tu, a Ph.D. candidate, who gave me large amount of insightful suggestions during my research, and also my dear friends, Dong Yang, Sile Tao, etc., for supplying me useful suggestions on my thesis writing.

A very special word of thanks goes for my parents, Shuying Wang and Ping Fu, my grandparents, Xinzhu Chen and Bangfu Wang, and my partner Nicholas Dombrosky for their love and support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Functional Brain Connectivity

Human brain is a very complex and efficient network, which consists of a great amount of interlinked brain regions. Each of them has their own function, but also constantly shares information with one another. Hence one complex integrative network is formed where information processing and transportation happens among those interlinked brain regions, which is critical for human being's cognitive functioning [53]. The functional integration of specific brain areas has proven hard to evaluate. However, functional brain connectivity is one approach to assess it. Functional brain connectivity is defined as the correlations between measurements of neuronal activity in different brain areas [28]. In practice, function brain connectivity can be evaluated by functional neuroimaging, where functional brain connectivity represents the functionally integrated relationship between the neuronal activation patterns of spatially separated brain regions. It embodies the level of functional information exchanges among brain regions [53].

Scientists and researchers have verified that many diseases are associated with changes in brain connectivity, such as Autism, Attention-deficit hyperactivity disorder (ADHD), and Alzheimer's disease (AD) [23, 37, 52]. Hence studying functional brain connectivity is of great importance and advantage for us to better understand the mechanisms of human brain and related diseases.

The advanced non-invasive functional brain imaging techniques play an important role in understanding large-scale functional brain connectivity. Over the last several decades, functional neuroimaging techniques have greatly developed. In particular, the available methods for today's researchers include Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), Electro-encephalography (EEG), Magneto-encephalography (MEG), etc. [39]. Among those, MRI has proven to be able to supply extremely versatile human body image. The first MRI body scan of a human was performed in 1977 by Raymond Damadian along with Larry Minkoff and Michael Goldsmith, which took five hours with low-quality image. Since then, the resolution and speed of MRI have been significantly boosted. The developments in scanner techniques have enabled the acquisition of very fast and approximate real-time MRI imaging. Recently a few new MRI methods have evolved, including functional MRI (fMRI), diffusion tensor imaging (DTI), and magnetic resonance spectroscopy (MRS) to provide useful functional information for clinical and research purposes [18].

Functional Magnetic Resonance Imaging (fMRI) is a promising development in MRI, and has become increasingly popular over recent years. It utilizes magnetic changes between oxygenated and deoxygenated hemoglobin during neuronal activation and rest in the brain [29]. The neural activity in the brain is observed through imaging blood flow changes which results in a so-called

"blood-oxygen-level dependent" (BOLD) signal change measured by fMRI. BOLD contrast method, discovered firstly by Seiji Ogawa, is the most common used measure of neural activity in the brain so far [5, 18]. These small changes are detected as minute distortions in the magnetic field by fMRI and employed to generate functional images reflecting brain activity [29]. During the fMRI scanning, the subject is placed into a cylindrical magnet machine with strong and homogeneous magnetic field. The proton nucleus of the hydrogen atom align themselves under the field and achieve a thermal equilibrium, hence the subject is magnetized. Then the machine applies a brief radio frequency electromagnetic pulse for disturbing the equilibrium, such that the proton nuclei become out of alignment. After ending the radio pulse, the proton nuclei fall back in line and release radio signals, which are detected and hence formed an image [17].

fMRI has two categories, task-based and resting state. Resting state functional magnetic resonance imaging (rs-fMRI), or taskless fMRI, is a method of functional magnetic resonance imaging that allows investigators to assess brain functional connectivity in a resting or task-negative state in the clinical setting. Doctor Biswal and his colleagues firstly stated that human beings' left and right hemispheric regions are not silent under resting states, instead, they illustrate a high correlation between functional MRI BOLD time-series [11, 12]. During resting state fMRI scan, the subjects were asked to relax and think nothing whilst their spontaneous brain activities were evaluated, thus the spontaneous functional MRI time-series were generated [53].

Figure 1.1 illustrates the fMRI device (model: Siemens Skyra) [2]. During the fMRI scanning, the subject is placed into the cylindrical magnet machine with strong and homogeneous magnetic field. The neural activity in the brain

Figure 1.1: fMRI Device (Model: Siemens Skyra). During the fMRI scanning, the subject is placed into the cylindrical magnet machine with strong and homogeneous magnetic field. The neural activity in the brain is observed through imaging blood flow changes. These small changes are detected as minute distortions in the magnetic field by fMRI and employed to generate functional images.

is observed through imaging blood flow changes. These small changes are detected in the magnetic field by fMRI and employed to generate functional images. Figure 1.2 and Figure 1.3 show task-based and resting state BOLD signals that fMRI machine picks up, respectively [26].

In order to build functional brain connectivities based on fMRI data, normally we first identify functional "nodes". Spatial regions of interest (ROIs) method is one simple way to define nodes for the case of fMRI. When we perform brain connectivity studies, we commonly employ clusters of neighboring voxels or ROIs. An average BOLD signal time series is normally treated as representative of each ROI as for the most brain connectivity studies we only consider one signal for each region of input [46, 49]. Graph analysis enables us to explore complex systems described by pairwise connections between brain regions and the structure of functional connectivity [48]. In the field of graph

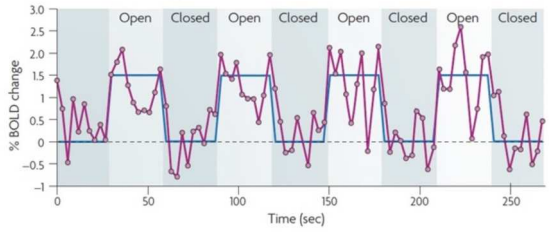


Figure 1.2: Task-based BOLD Signals. A simple task paradigm requires the subject to open and close eyes, and the resulting changes in neuronal activity are measured. The paradigm is shown in blue (delayed to account for the haemodynamic response).

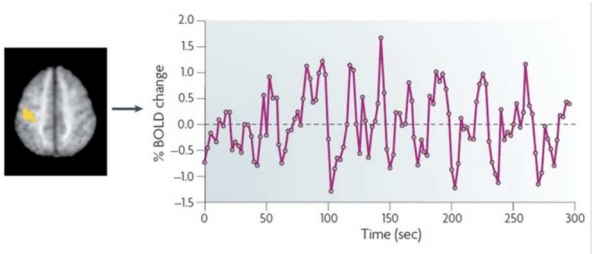
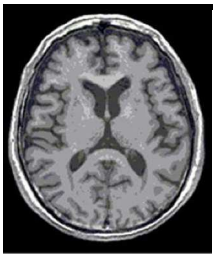


Figure 1.3: Resting-state BOLD Signals. During resting state fMRI scan, the subject is asked to relax and think nothing whilst their spontaneous brain activities are evaluated, thus the spontaneous functional MRI time-series are generated.

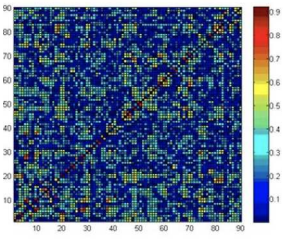
theory [13], a graph consists of the nodes and edges. To study the brain connectivity networks' structure, ROIs can be treated as nodes and connections as edges. No connection means the two ROIs are conditionally independent with each other. Once the nodes are defined, then the connections between two nodes can be treated as edges. Thus we can conduct a connectivity analysis among ROIs [50].

Figure 1.4 shows us generally there are four steps to construct functional brain connectivity. For each subject, the brain fMRI images (Figure 1.4 a) were parcelled in to ROIs using Automated Anatomical Labeling (AAL) template (Figure 1.4 b). Time courses in each region of interest were then extracted. Various methods can be applied to build functional brain connectivity, for example correlation, partial correlation, and Graphical Lasso. If we choose

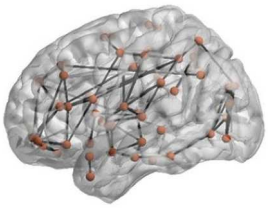


(a) fMRI Image

(b) AAL atlas



(c) Correlation matrix



(d) Functional Brain Connectivity

Figure 1.4: Functional Brain Connectivity Construction Process. When constructing functional brain connectivity, the brain fMRI images (a) were first parcelled in to ROIs using Automated Anatomical Labeling (AAL) template (b). After building the function brain connectivity, the correlation matrix for each subject was derived (c). By enforcing sparsity using threshold, the functional brain connectivity was obtained (d).

correlation method, then we calculated the Pearson correlation coefficient between each ROI pair of different time courses, which stands for the strength of the corresponding connectivities. Hence we derived correlation matrix for each subject (Figure 1.4 c). Through setting up the reasonable hard threshold value for the correlation matrix, we obtained the functional brain connectivity (Figure 1.4 d) [41].

Once we build the functional brain connectivities for both normal group and disease group, we can conduct the comparisons in order to identify the group differences, which is essential for uncovering underlying neurological processes associated with the corresponding disease. Recently, there are lots

of studying focusing on studying group differences in functional brain connectivity such as [36, 45, 56]. Detecting group differences for specific diseases associated with functional brain connectivities is critical for both research and clinical uses.

Due to scanner stabilities, acquisition, or issues in the underlying biomedical experimental protocol, however, fMRI BOLD signals might contain outliers [44]. The presence of potential outliers might lead to unreliable results in constructing functional brain connectivities. Another challenge of most existing connectivity construction methods is that the results could be unstable, for example, small changes in the sample will result in significantly different connectivities. In order to deal with these challenges, we proposed a framework which enables us to provide robust and stable connectivities. In particular, a low-rank plus sparse ($L+S$) matrix decomposition technique was adapted to decompose the resting state fMRI BOLD signals, where the low-rank matrix L recovers the essential common features from regions of interest (ROIs), and the sparse matrix S catches the sparse individual variability and potential outliers. Based on the decomposition, various methods can be applied to construct functional brain connectivities, such as correlation, partial correlation, Graphical Lasso and others. In this thesis, we proposed to use the recently developed novel sparse matrix estimation based on concentration inequality, which is proved to have superior performance than other methods. Statistical test for each ROI connection was implemented for differentiating group difference. Through bootstrapping afterwards, we verified that performing low-rank plus sparse matrix decomposition can achieve more stable results.

In our study, in order to implement low-rank plus sparse matrix decom-

position method, we employed Alzheimer's Disease Neuroimaging Initiative (ADNI) data for Alzheimer's disease (AD). The recent research has shown that Resting-state functional magnetic resonance imaging (rs-fMRI) can supply the evidence that indicates functional brain connectivities are associated with Alzheimer's disease [35]. Therefore in our study we utilized resting-state functional magnetic resonance imaging data from ADNI study.

1.2 Contributions

Functional brain connectivity plays an important role on understanding how human brain functions and paving a path to reveal the mechanisms of neuropsychological diseases such as Autism, Attention-deficit hyperactivity disorder (ADHD), and Alzheimer's disease (AD). Functional Magnetic Resonance Imaging (fMRI) is one of the most powerful techniques to construct functional brain connectivity. However, the presence of potential outliers in fMRI BOLD signals might lead to unreliable results in construction. Another challenge of most existing connectivity construction methods is that the results might not be stable, for example, small changes in the sample will result in significantly different connectivities.

To deal with challenges, we proposed a framework which enables us to provide robust and stable connectivity. In particular, a low-rank plus sparse $(L+S)$ matrix decomposition technique was adapted to decompose the resting state fMRI BOLD signals, where the low-rank matrix L recovers the essential common features from regions of interest (ROIs), and the sparse matrix S catches the sparse individual variability and potential outliers. Based on decomposition, various methods can be applied to construct functional brain

connectivity, such as correlation, partial correlation, Graphical Lasso and others. However, we proposed to use the recently developed novel sparse matrix estimation based on concentration inequality. Statistical test for each ROI connection was implemented for differentiating group difference. Through bootstrapping afterwards, we verified whether performing low-rank plus sparse matrix decomposition can achieve more stable results.

We applied our method on Alzheimer's Disease Neuroimaging Initiative (ADNI) data, and compared the results with those based on original BOLD signals. We discovered that the methods for building connectivities based on low rank matrices behave better than based on original BOLD signals, in the sense that the former can reveal and identify more significant ROI connections. Hence this suggests that when we study group difference for functional brain connectivity, performing low-rank plus sparse matrix decomposition can achieve more differentiable results, which can contribute to uncovering underlying neurological processes associated with the disease for clinical use. We found that the recently proposed concentration inequality based method performs better overall compared with correlation, partial correlation, and Graphical Lasso method. We also obtained the first ten most significant connections for differentiating group differences for Alzheimer's disease. Among them, the left hippocampus region and the left cerebellum 7 region is the most significant one, with p value smaller than 0.0005, which is consistent with existing literatures' findings. Through bootstrapping, we verified that performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing functional brain connectivities.

The thesis is organized as below. In Chapter 2, we introduced low rank plus sparse matrix decomposition first. Then we introduced the sparse matrix

estimation based on concentration inequality for constructing functional brain connectivity. To facilitate the comparisons, we introduced correlation, partial correlation and Graphical Lasso for construction as well. Subsequently, we elaborated the methods for detecting group differences. In Chapter 3, we first introduced data description. We then applied our method on ADNI data and compared the results with those based on original BOLD signals. Through bootstrapping we analyzed stability property. The conclusion and potential future work were given in Chapter 4.

Chapter 2

Methodology

In our study, we first performed low-rank plus sparse matrix decomposition to our original fMRI BOLD signals. Then we constructed the functional brain connectivities using the recently proposed novel concentration inequality based estimation method. To facilitate the comparisons, we utilized correlation, partial correlation and Graphical Lasso methods for construction as well. Subsequently, we implemented statistical test for each ROI connection to reveal and identify the underlying group differences. After getting the p value matrices based on both original BOLD signals and low rank matrices, we calculated the percentage of significant connections and compared the results. We also obtained the first ten most significant connections for differentiating group differences for Alzheimer's disease. Then we defined and calculated the overlap rate in order to see whether performing low-rank plus sparse matrix decomposition can keep some level of consistency compared with when using original BOLD signals. Lastly we implemented bootstrapping. By calculating the variances for percentages of significant connections in p value matrices for 50 times resampling, we explored whether performing low-rank plus sparse ma-

trix decomposition can achieve more stable results for constructing functional brain connectivities.

2.1 Low Rank plus Sparse Matrix Decompositions

Matrix decomposition is an important topic and has wide applications in many scientific and engineering problems. Various matrix decomposition techniques have been developed, such as LU Decomposition, QR decomposition, singular-value decomposition (SVD) [6]. Low-rank plus sparse matrix decomposition is one of them. It originates from robust principal component analysis. Principal component analysis (PCA) is a powerful data analysis tool for drastically reducing the dimension of large datasets, which tries to find a low rank subspace that approximates the original data matrix best and increase the interpretability, but at the same time preserve the information from the data as much as possible. However, the nature of PCA means it is sensitive to the outliers. Hence we would like to develop methods which are able to extract the principal components even in the presence of outliers in the data, therefore enhance the robustness to outliers. This led researchers to explore various robust PCA methods [32].

Robust principal component analysis (RPCA) is a promising framework for dealing with this task. Early attempts to solve the RPCA problem have been conducted [19, 20, 21, 22, 59]. However they could not achieve polynomial time solutions with high performance. A more recent version of Robust PCA becomes increasingly popular [33]. Low-rank plus sparse ($L + S$) matrix de-

composition is sometimes also referred to Robust PCA. The recovery of both low-rank and sparse matrices is of great interest in many applications, such as imaging processing and bioinformatic data analysis [42].

The straightforward formulation for this problem is to employ l_0-norm to solve

$$\min \ \text{rank}(L) + \lambda \parallel S \parallel_0,$$

$$\text{subject to } L + S = M,$$

where M is a $m \times n$ matrix to be decomposed, L is a low-rank matrix, S is a sparse matrix, $\|\cdot\|_0$ is the l_0-norm and λ is a nonnegative tuning parameter. But this optimization problem is not convex and NP-hard. Hence it was suggested [16] that we can approximate this problem by a convex optimization problem, which is to minimize a combination of the $l1$ norm of S and the nuclear norm of L [57]. This essentially is the low-rank plus sparse $(L + S)$ matrix decomposition. Specifically, the low-rank plus sparse matrix decomposition can be expressed as

$$\min \ \parallel L \parallel_* + \lambda \parallel S \parallel_1, \tag{2.1}$$

$$\text{subject to } L + S = M,$$

where $\|\cdot\|_*$ is the nuclear norm (sum of all singular values), $\|\cdot\|_1$ is the $l1$-norm, and λ is a nonnegative tuning parameter. The choice of λ is a trade-off between the low rank and the sparsity. For example, smaller λ can enforce lower rank for L, but relax the sparsity for S. It was shown that mathematically a whole range of λ values can ensure the exact recovery of both low-rank and sparse matrices [16].

There are some existing algorithms which can solve this problem computationally, such as Accelerated Proximal Gradient (APG), Augmented Direction Method (ADM), Augmented Lagrange Multiplier (ALM), etc. [14]. In this thesis we used ALM method which achieves the complexity as $O(mn\,min(m,n))$, where m, n are the number of rows and columns of the decomposed matrix [42]. The R package for ALM is available online.

In the framework of ALM, the target is to solve the following optimization problem with the constrain,

$$\min f(X), \text{ s.t. } h(X) = 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$. In ALM, an augmented Lagrangian function is needed, which is defined as

$$L(X, Y, \mu) = f(X) + < Y, h(X) > + \frac{\mu}{2} \parallel h(X) \parallel_F^2 \tag{2.2}$$

where μ is a positive constant. Under the framework of RPCA,

$$X = (L, S),\ f(X) = \parallel L \parallel_* + \lambda \parallel S \parallel_1, \text{ and } h(X) = M - L - S.$$

The corresponding Lagrangian function is

$$L(L, S, Y, \mu) = \parallel L \parallel_* + \lambda \parallel S \parallel_1 + < Y,\ M - L - S > + \frac{\mu}{2} \parallel M - L - S \parallel_F^2 .$$

The ALM algorithm for solving RPCA problem in (2.1) is shown in Algorithm 1 [42].

In this thesis, we applied the low-rank plus sparse $(L+S)$ matrix decomposition to the original fMRI BOLD signals. More specifically, let $X_{N\times T\times J} = x_{itj}$ be a 3 dimensional matrix where i represents the i-th subject $(i = 1, ..., N)$, t

Algorithm 1 Augmented Lagrange Multiplier (ALM)

Input: Observation matrix $M \in \mathbb{R}^{m \times n}$, λ.

$Y_0^* = \text{sgn}(M)/J(\text{sgn}(M))$; $\mu_0 > 0$; $\rho > 1$; $k = 0$.
while not converged **do**
  //Lines 4-12 solve $(L_{k+1}^*, S_{k+1}^*) = \underset{L,S}{\text{argmin}} L(L, S, Y_k^*, \mu_k)$.
  $L_{k+1}^0 = L_k^*, S_{k+1}^0 = S_k^*, j = 0$;
  while not converged **do**
    //Lines 7-8 solve $L_{k+1}^{j+1} = \underset{L}{\text{argmin}} L(L, S_{k+1}^j, Y_k^*, \mu_k)$.
    $(U, D, V) = svd(M - S_{k+1}^j + \mu_k^{-1} Y_k^*)$;
    $L_{k+1}^{j+1} = U D_{\mu_k^{-1}}[D] V^T$;
    //Lines 10 solve $S_{k+1}^{j+1} = \underset{S}{\text{argmin}} L(L_{k+1}^{j+1}, S, Y_k^*, \mu_k)$.
    $S_{k+1}^{j+1} = D_{\lambda \mu_k^{-1}}[M - L_{k+1}^{j+1} + \mu_k^{-1} Y_k^*]$;
    $j \leftarrow j + 1$.
  end while
  $Y_{k+1}^* = Y_k^* + \mu_k (M - L_{k+1}^* - S_{k+1}^*)$.
  update μ_k to μ_{k+1}.
  $k \leftarrow k + 1$.
end while

Output: (L_k^*, S_k^*).

is the t-th time course($t = 1, ..., T$), and j is the j-th ROI($j = 1, ..., J$). Then for each fixed ROI, we conducted low-rank plus sparse matrix decomposition to each $X_{N\times T}$ matrix. Thus we derived a new 3 dimensional matrix denoted as $L_{N\times T\times J} = l_{itj}$. Then we built functional brain connectivity based on each $L_{T\times J}$. Here $X_{N\times T}$ matrix is equivalent to M matrix, $L_{T\times J}$ is equivalent to L matrix and N and T are corresponding to m and n in (2.1) respectively.

2.2 Functional Brain Connectivity

In the thesis, we constructed the functional brain connectivities by utilizing the recently proposed novel concentration inequality based estimation method of sparse covariance matrices. In order to facilitate the comparisons, we also employed correlation, partial correlation and Graphical Lasso methods for construction.

Estimation of covariance matrices is an important topic. It can be utilized on lots of applications for genetics, imaging and other types of data. Many estimators of covariance matrix have been explored working under the assumption of sparsity, which is desirable and applicable in real data analysis especially in high dimension settings since many variable pairings might be considered uncorrelated. Shrinkage estimators, and thresholding estimators, for example, are two classes for sparse estimators of covariance matrix.

Let vectors $X_1, \cdots X_T \in \mathbb{R}^J$ be a sample with mean zero and covariance matrix Σ, and S is the empirical estimate of Σ for each subject. When the dimension J is large and Σ is sparse, S might not be a good estimator. To construct a better estimator, a novel approach is recently proposed making use of confidence sets constructed from concentration inequalities for nonasymptotic

covariance matrix estimation. Let $\hat{\Sigma}^{sp}$ be the sparse estimator for Σ. This method chooses a $\hat{\Sigma}^{sp}$ such that it is close enough to S while it lies far enough away to result in a sparse estimator. This novel concentration inequality based method supplies finite sample guarantees and a much faster computing time compared with costly optimization and cross validation methods [34].

The sparse estimation procedure aims at constructing a sparse estimator $\hat{\Sigma}^{sp}$ for Σ by constructing a nonasymptotic confidence set first employing concentration inequalities for Σ based on S, and searching this set in order to obtain the sparsest estimator. The general form for the concentration inequalities is

$$\mathsf{P}(d(\Sigma, S) \geq \mathsf{E}d(\Sigma, S) + r) \leq \mathrm{e}^{-\psi(r)},$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is a monotonically increasing function, and $d(\cdot)$ is some metric measuring the distance of two covariance matrices. To construct a $1 - \alpha$ confidence set, $r = r_\alpha$ is chosen such that

$$\exp(-\psi(r_\alpha)) = \alpha.$$

The smaller α is, the larger r is. Our sparse estimator $\hat{\Sigma}^{sp}$ is expected to be close to Σ in the sense of the above confidence set, and thus we focus on choosing a $\hat{\Sigma}^{sp}$ such that

$$d(\hat{\Sigma}^{sp}, S) \leq r_\alpha.$$

It begins with S and attempts to threshold it as much as possible while still remaining in the confidence set.

First, we define a generalized thresholding operator [47] as $s_\lambda(\cdot) : \mathbb{R} \to \mathbb{R}$ such that

$$| s_\lambda(z) | \leq z,\ s_\lambda(z) = 0 \text{ for } | z | \leq \lambda, \text{ and } | s_\lambda(z) - z | \leq \lambda,$$

which will apply to each single element of a matrix. In the past this estimator was applied to S for some λ chosen by cross validation. Instead of choosing the threshold λ, this approach tries to choose a confidence level $1-\alpha$ and then seek for the largest λ such that $d(s_\lambda(S), S) \leq r_\alpha$ [34].

The algorithm for how to derive the sparse covariance matrix estimation is shown in Algorithm 2 [34].

Algorithm 2 Concentration Inequality Based Estimation of Sparse Covariance Matrices

0. Set $\hat{\Sigma}_0^{sp} = (\hat{\Sigma}^{diag})^{-1/2} S (\hat{\Sigma}^{diag})^{-1/2}$, $\lambda = 0.5$ and write $\hat{\Sigma}_\lambda^{sp} = s_\lambda(S)$. Define $k = 1$ as the number of the recursion's steps. Choose an α and compute r_α.
1. Increase $k \leftarrow k + 1$, then update the threshold λ as below:
if $d(\hat{\Sigma}_\lambda^{sp} - S) \leq r_\alpha$, **then**
    let $\lambda \leftarrow \lambda + 2^{-k-1}$.
else
    let $\lambda \leftarrow \lambda - 2^{-k-1}$.
end if
2. Repeat step 1 until k has gotten to the desired number of iterations. Usually $k = 10$ would be enough for us.
3. The resulting sparse estimator is $\hat{\Sigma}^{sp} = (\hat{\Sigma}^{diag})^{1/2} (\hat{\Sigma}_\lambda^{sp}) (\hat{\Sigma}^{diag})^{1/2}$, where $\hat{\Sigma}_\lambda^{sp}$ is our final sparse estimator.

In our study, we employed the operator norm $\| \hat{\Sigma}_\lambda^{sp} - S \|_\infty$ as the distance metric $d(\cdot, \cdot)$. We chose reasonable false positive rate α in order to get the reasonable sparsity for the estimation of sparse covariance matrices. Once we obtain the sparse covariance estimation, we can calculate the corresponding correlation matrix as functional brain connectivity.

To facilitate the comparisons, we also employed correlation, partial correlation and Graphical Lasso methods for constructing the connectivities in our study. Pearson's correlation between two ROIs is used in lots of studies for building functional brain networks and connectivity, due to the fact that it is

not hard to compute using a sample covariance matrix [36]. One disadvantage for Pearson's correlation, however, is that if the correlation between two nodes is caused by a third node, we would not be able to see if the direct functional connection between these two nodes exists. To overcome this limitation, partial correlation has been adopted [30]. The partial correlation conditions on the other regions. Specially, two regions of interest are considered conditionally independent under the Gaussian assumption if partial correlation equals zero. When estimating partial correlations, a precision matrix (or inverse covariance matrix) can be used [36]. After obtaining correlation or partial correlation matrices, we select hard threshold values such that these matrices can be as sparse as we expect. Thus the functional brain connectivities are estimated by the thresholded correlation or partial correlation matrices.

The correlation between two ROIs can be calculated based on a sample covariance matrix. Specifically, let the matrix $M_{J\times T}$ be the i-th subject's BOLD signals or low-rank matrix from decomposition, and each column $M_1, \cdots M_T \in \mathbb{R}^J$. Then the sample covariance matrix can be derived from $S = s_{pq} = (T-1)^{-1}\sum_{i=1}^{T}(M_i - \overline{M})(M_i - \overline{M})^T$, where $\overline{M} = T^{-1}\sum_{i=1}^{T} M_i$. Then the full correlation between the p-th ROI and the q-th ROI can be estimated as

$$r_{pq} = s_{pq}/(s_{pp}s_{qq})^{\frac{1}{2}}.$$

Define the precision matrix $\Theta_{pq} = (\theta_{pq}) = \Sigma^{-1}$, where Σ is covariance matrix, then the partial correlation between the p-th ROI and the q-th ROI is

$$\rho_{pq} = -\theta_{pq}/(\theta_{pp}\theta_{qq})^{\frac{1}{2}}.$$

If the number of ROIs is relatively large, then our derived correlation and partial correlation matrices would be also relatively with high dimensions.

In practice, the fact is for the most of time we might only be interested in selecting those connection pairs with larger correlation coefficient values, which means they have stronger connections compared to the others. In order to achieve this goal, we apply thresholding values to both correlation and partial correlation matrices [15, 24]. In this way we would get sparse correlation and partial correlation matrices, which would help us focusing on the relatively more important connections among those ROIs.

Let $R = (r_{ij})_{J \times J}$ be the sample correlation matrix, and let τ be the reasonable pre-specified thresholding value. Then we enforce the thresholding value to all the off diagonal elements of our correlation matrix to get the corresponding sparse correlation matrix R^{sp}, i.e. for the (i, j)-th element of R^{sp},

$$R^{sp}_{(i,j)} = \begin{cases} 1, & i = j \\ r_{ij} 1\{\mid r_{ij} \mid > \tau\}, & i \neq j \end{cases}.$$

When $\tau = 1$, it is an identity matrix, while when 0, it is the original sample correlation matrix as we do not apply any thresholding. The same thresholding method was also applied to the estimated partial correlation matrices in the thesis.

We also utilized Graphical Lasso for construction. Denote $\varsigma(\upsilon, \varepsilon)$ as a Gaussian graphical model, which is composed of nodes $\upsilon = \{1, 2, ..., p\}$ and edges $\varepsilon \subset \upsilon \times \upsilon$. Any edge $(i, j) \in \varepsilon$ illustrates that the i-th ROI and j-th ROI are statistically dependent conditional on all the other nodes. Then any nonzero entry in the precision matrix Θ implies that these two ROIs have direct relationship while building functional brain connectivity [45].

A precision matrix can be used for estimating partial correlation. Natu-

rally if we take the inverse of sample covariance matrix, the estimated precision matrix will be not sparse. Also, the estimate of inverse covariance in a high dimensional and relatively small number of observations setting can have singularity problem. Graphical Lasso (GLasso) [27] is a regularized method which enables us to obtain sparse precision matrix estimation by $l1$ norm regularization. We can achieve desired sparsity through tuning parameter ρ.

The Graphical Lasso problem is to maximize the $l1$-penalized Gaussian log-likelihood

$$\log(\det \Theta) - \mathrm{tr}(S\Theta) - \rho \left\|\Theta\right\|_1, \tag{2.3}$$

over nonnegative definite matrices Θ, where $\Theta = \Sigma^{-1}$, tr denotes the trace, S is the sample covariance matrix, ρ is a nonnegative tuning parameter, and $\left\|\Theta\right\|_1$ is the l_1 norm, the sum of the absolute values of all the entries of Σ^{-1} [27]. When ρ is 0, then there is no penalty. When ρ is sufficiently large, the estimate $\hat{\Theta}$ will be sparse due to the lasso-type penalty. The problem (2.3) was shown to be convex [10].

In our study, Graphical Lasso was conducted to obtain the estimated sparse precision matrix automatically due to the method as functional brain connectivity. The value ρ in (2.3) was chosen to get the reasonable sparsity.

2.3 Detecting Group Differences

In our study, in order to detect group differences in functional brain connectivity, we conducted two sample T statistical test to test whether there is any difference for each single ROI connection between control group and

disease group. Theoretically, if we have J ROIs, then we have $J(J-1)/2$ total connections. Suppose a connectivity matrix $C^{(i)}_{J\times J}$ for i-th subject can be denoted as $(c^{(i)}_{p,q})$, where $c^{(i)}_{p,q}$ represents the connectivity metric of the p-th ROI and q-th ROI, $p = 1, 2, \cdots, J$, and $q = 1, 2, \cdots, J$ (for example, for correlation method, it is the correlation coefficient). Then for each single ROI pair $c^{(i)}_{p,q}$ of matrix $C^{(i)}_{J\times J}$, we will have one group of values: $c^{(1)}_{p,q}, c^{(2)}_{p,q}, \cdots c^{(N_1)}_{p,q}$ from control group and another group of values: $c^{(1)}_{p,q}, c^{(2)}_{p,q}, \cdots c^{(N_2)}_{p,q}$ from disease group. Then we can conduct two sample T statistical test for each single connection $c_{p,q}$ of p-th ROI and q-th ROI, to test whether there is any difference for each single connection between control group and disease group. Afterwards we can generate a p value matrix $P_{J\times J} = (p_{p,q})$, where $p_{p,q}$ is the corresponding p value for T statistical test for the connection $c_{p,q}$ between control group and disease group, $p = 1, 2, \cdots, J$, and $q = 1, 2, \cdots, J$. In the following analysis, we ignored the diagonal elements since they do not make any sense, and also only focused on the elements in upper triangle since it is symmetric.

Chapter 3

Numerical Study on ADNI Dataset

In this chapter, we adapted the low-rank plus sparse matrix decomposition technique to decompose the resting state fMRI BOLD signals from Alzheimer's Disease Neuroimaging Initiative (ADNI) data. Then we constructed the functional brain connectivities using correlation, partial correlation, Graphical Lasso, and the concentration inequality based estimation method, based on both original BOLD signals and low rank matrices from decomposition, respectively. Subsequently, we implemented two sample T statistical test for each ROI connection to get the p value matrices for both original BOLD signals and low rank matrices. We also obtained the first ten most significant connections for differentiating group differences for Alzheimer's disease. We then defined and calculated the percentage of significant connections and compared the results. After that, we defined and calculated the overlap rate in order to see whether performing low rank plus sparse matrix decomposition can keep some level of consistency compared with when using original BOLD

signals. Lastly we implemented bootstrapping and then calculated the variances for percentages of significant connections in p value matrices in order to analyze stability property of our method.

3.1 Data Description

We applied our method on Alzheimer's Disease Neuroimaging Initiative (ADNI) data. Hence we first reviewed the Alzheimer's Disease and ADNI. In history, Alzheimer's disease (AD) was first diagnosed by Dr. Alois Alzheimer in 1906 [38]. Alzheimer's disease is a chronic irreversible neurodegenerative disease which can result in a loss of cognitive functions. It is the most common cause of late life dementia that causes problems such as memory loss, slowed thinking, personality changes and behavior changes, which accounts for 60% to 80% of dementia cases. Alzheimer's disease usually progresses very slowly and gradually worsen over a number of years, becoming serious enough to interfere with people's daily life. It typically develops gradually in three stages: mild (early stage), moderate (middle stage), and severe (late stage) [4]. Alzheimer's disease is not a part of normal aging, but the most significant known risk factor for Alzheimer's is increasing age. The majority of people diagnosed with Alzheimer's disease are older than age 65. However it can occur in people younger than 65, which is much more rare [9].

Alzheimer's disease is recognized as a major public health problem, as it eventually affects every aspect of people's life. In 2006, 26.6 million people in the world suffered from Alzheimer's disease [43].The World Health Organization estimated that the prevalence rate of Alzheimer's disease in the overall worldwide population would increase to 0.556% in 2030. Another study reck-

oned in 2006, there were 0.4% of the worldwide population suffering from the Alzheimer's disease, and this number would triple by 2050 [55].

The Alzheimer's Disease Neuroimaging Initiative (ADNI) [1] *is a global longitudinal study for AD through the enrollment and follow-up of cohorts of individuals who have mild cognitive impairment (MCI) and mild Alzheimer's disease. It is jointly funded by FNIH (Foundation for the National Institutes of Health), alzheimer's & association, abbvie, etc. The study is designed for the detection at the earliest possible stage and tracking the progression of Alzheimer's disease with biomarkers to assess the brain structure and the brain function. The participants enrolled by ADNI were between 55 to 90 years of age, selected based on the particular criteria, and recruited at the 57 ADNI acquisition sites located in the United States and Canada. The five cohorts in this study are Normal Control (CN), Significant Memory Concern (SMC), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), and Alzheimer's Disease (AD), respectively. The selected subjects undergo clinical, imaging, genetic, and also biochemical biomarkers at multiple time points [1, 7].

The general goals of the ADNI study are for validation of biomarkers' data for the trials' use in the illness clinical treatment and assessments, for exploring methods for obtaining data and analyzing neuroimaging data in longitudinal studies for clinical trials on patients with normal controls, mild cognitive impairment, and Alzheimer's disease, for making data repository

*Data used in preparation of this thesis were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

accessible for other researchers and communities, and for developing technical standards of imaging in longitudinal studies [1, 31].

In this study, we employed a subset of the resting-state fMRI ADNI data which includes 57 subjects from two cohorts, one is 33 subjects from Normal Control group, and the other one is 24 subjects from Alzheimer's Disease group. The resting-state fMRI data were preprocessed using Automated Anatomical Labeling (AAL) template [51]. AAL is an anatomical atlas of total regions of interest obtained on one subject. The AAL template is broadly employed in functional neuroimaging research, including resting-state fMRI. It aims at deriving neuroanatomical labels in a space where the measurements of brain function were captured [51]. Figure 3.1 shows us an AAL atlas [40]. The non-overlapping regions of interest were then extracted for each subject. For each subject, each time-series and ROI were computed through averaging all the voxels' time series within the ROIs [48]. Hence each subject has BOLD signal data at 116 ROIs through 134 equal spaced time courses. Also the demographic and clinical information of 57 subjects were collected, which consist of ID, gender, age and diagnostic information. For the variable gender, 0 represents female and 1 male. For diagnosis (DX), 0 stands for Normal Control(CN), and 4 Alzheimer's Disease (AD). All subjects had 1.5 Tesla and 3 Tesla scans by Philips scanners, having their eyes open when receiving the scanning [3].

3.2 ADNI Data Results

In this paper we utilized Augmented Lagrange Multiplier (ALM) [16] method for our low-rank plus sparse ($L+S$) matrix decomposition. We conducted low-

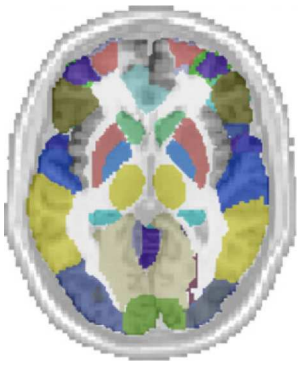


Figure 3.1: Automated Anatomical Labeling Atlas. It is broadly employed in functional neuroimaging research, including resting-state fMRI. Different colors represent different regions.

rank plus sparse decomposition to the original BOLD signals for each ROI. Specifically we employed $\lambda = 0.086$ in our work such that the rank of the matrices $L_{N\times T}$ was reduced to around 30, approximately half as much as the original rank 57 of the decomposed matrix $X_{N\times T}$. Then we constructed functional brain connectivities using correlation, partial correlation, Graphical Lasso, and the concentration inequality based estimation method, based on both original BOLD signals and low rank matrices respectively. The process is illustrated in Figure 3.2.

We first conducted correlation and partial correlation based on both the original BOLD signals and low rank matrices. We set the reasonable hard threshold value 0.4 for correlation for both cases, 0.8 for partial correlation based on original data matrices, and 0.2 for partial correlation based on low rank matrices, in order to derive the sparse connectivity selection results. Figure 3.3 and Figure 3.4 are connectivity selection results for correlation, and Figure 3.5 and Figure 3.6 for partial correlation. From here onwards, in all the graphs for indicating connectivity matrices, dark blue dots represent the nonzero elements of connectivity matrices and white dots mean entries with

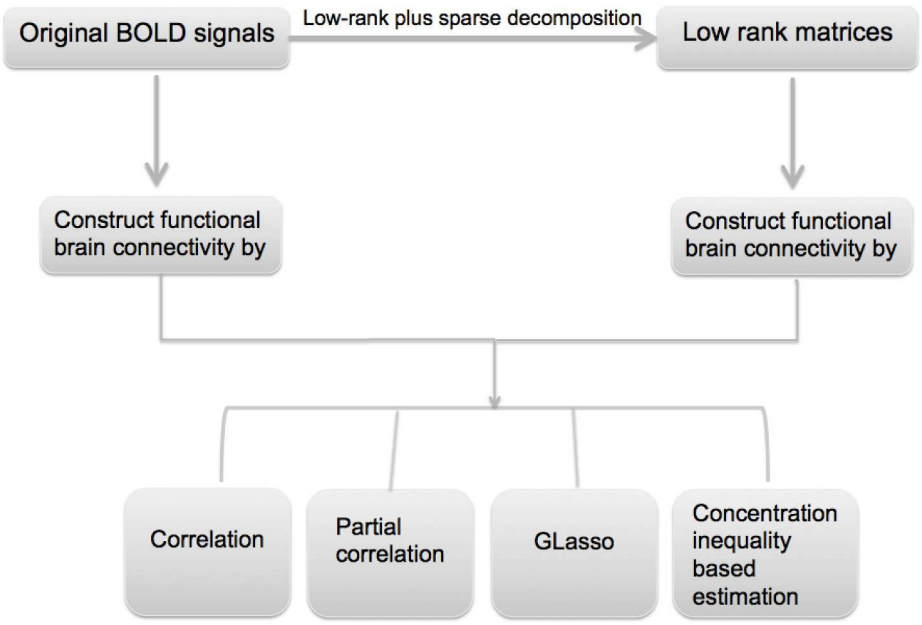
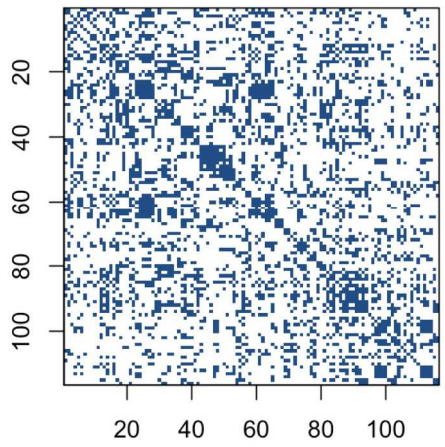


Figure 3.2: Flow Process for Functional Brain Connectivity Construction in Our Study. We conducted low-rank plus sparse decomposition to the original BOLD signals, then constructed functional brain connectivities based on both original BOLD signals and low rank matrices respectively.

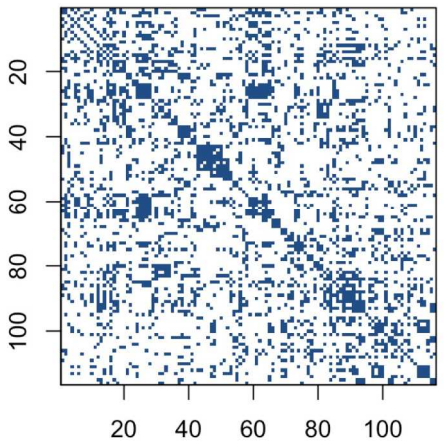
zero values.

From Figure 3.3 and Figure 3.4 we can see that the patterns of connectivity matrices based on both the original BOLD signals and low rank matrices for either control group subjects or AD group subjects are quite similar. Though we can also see that for both groups, the correlation matrices based on low rank matrices after employing thresholding values are slightly sparser than the ones based on original BOLD signals. This suggests that low rank plus sparse decomposition method can conservatively pick up those more important ROIs' connectivities for sparse correlation method. In other words, low rank plus sparse decomposition method might still identify the essential common features while still retaining most of the features.

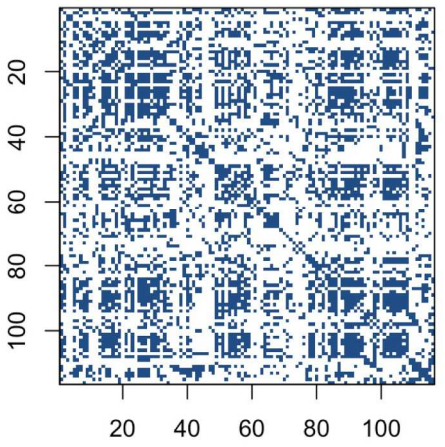
From Figure 3.5 and Figure 3.6, we can see that for partial correlation with hard threshold value, the output patterns based on low rank matrices



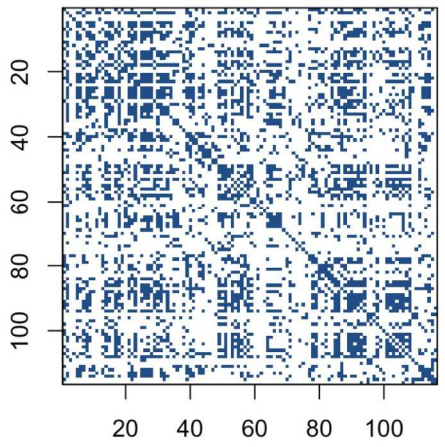
(a) Control ID 002-4225 based on original BOLD signals



(b) Control ID 002-4225 based on low-rank matrix

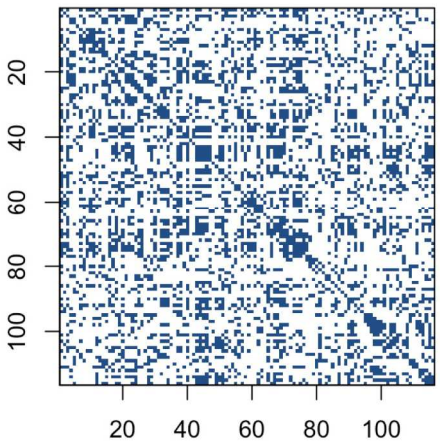


(c) Control ID 006-4357 based on original BOLD signals

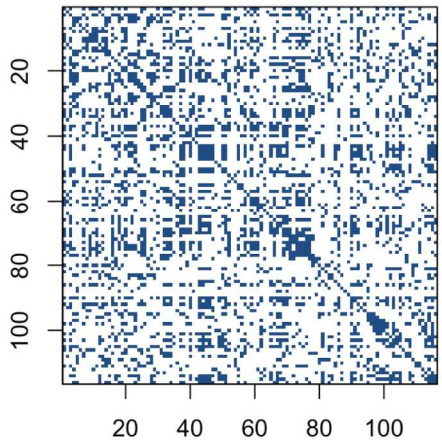


(d) Control ID 006-4357 based on low-rank matrix

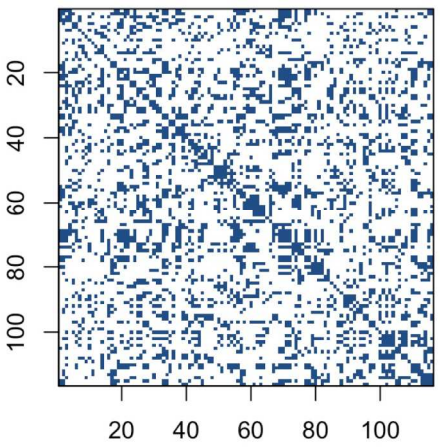
Figure 3.3: Connectivity Selection Results for Control Group from Correlation with Thresholding. The patterns of connectivity matrices based on both original BOLD signals and low rank matrices for control group subjects are quite similar. The correlation matrices based on low rank matrices after employing thresholding values are slightly sparser than the ones based on original BOLD signals.



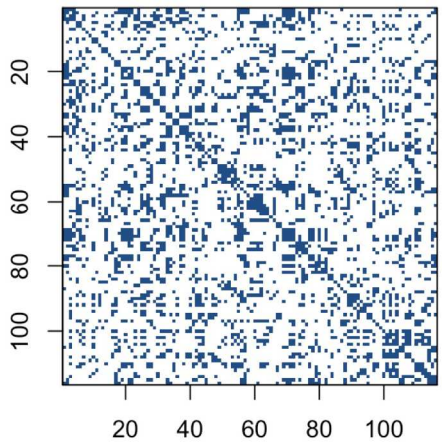
(a) AD ID 019-4477 based on original BOLD



(b) AD ID 019-4477 based on low-rank matrix signals

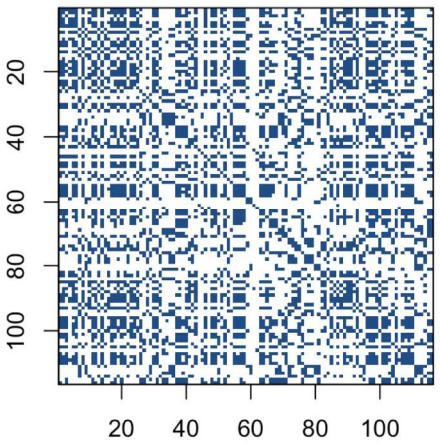


(c) AD ID 130-5006 based on original BOLD

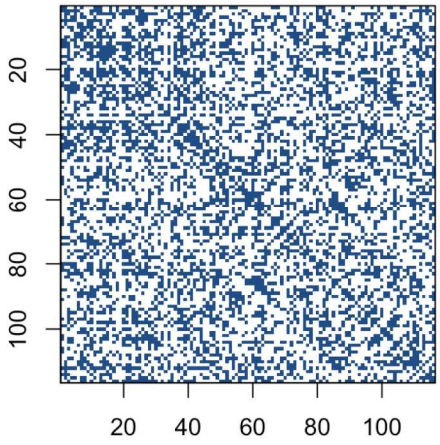


(d) AD ID 130-5006 based on low-rank matrix signals

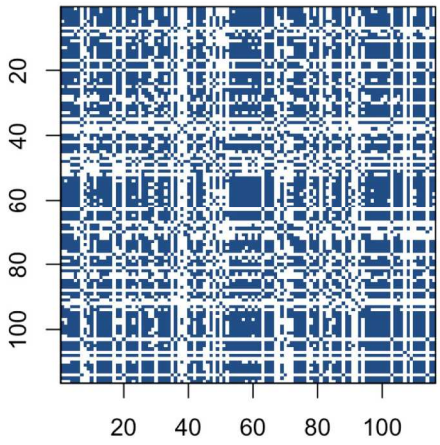
Figure 3.4: Connectivity Selection Results for AD Group from Correlation with Thresholding. The patterns of connectivity matrices based on both original BOLD signals and low rank matrices for AD group subjects are quite similar. The correlation matrices based on low rank matrices after employing thresholding values are slightly sparser than the ones based on original BOLD signals.



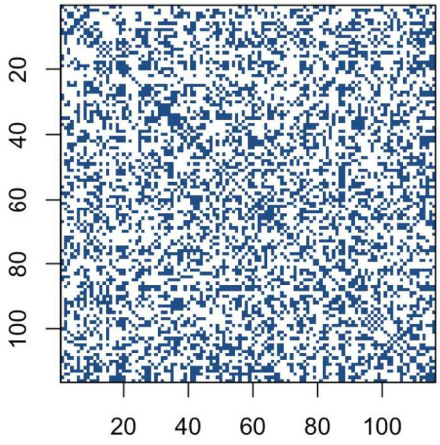
(a) Control ID 002-4225 based on original BOLD signals



(b) Control ID 002-4225 based on low-rank matrix

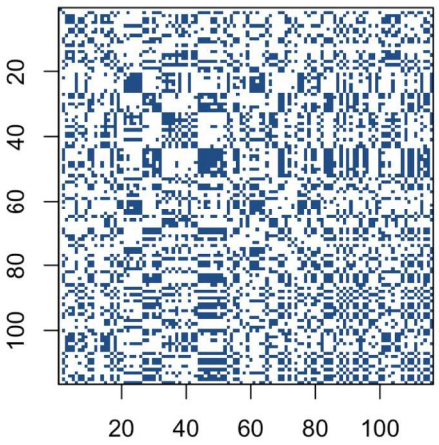


(c) Control ID 006-4357 based on original BOLD signals

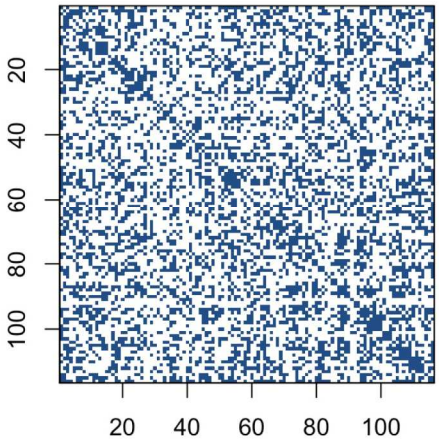


(d) Control ID 006-4357 based on low-rank matrix

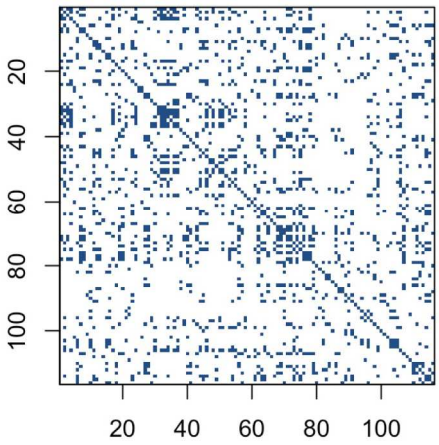
Figure 3.5: Connectivity Selection Results for Control Group from Partial Correlation with Thresholding. For partial correlation with hard threshold value, the output patterns based on low rank matrices keep decent similarity levels as the output based on original BOLD signals, but less similarity compared with sparse correlation method outputs.



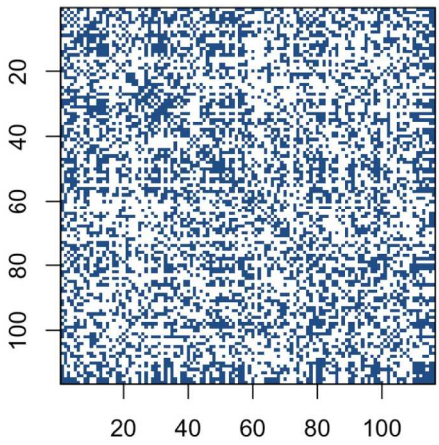
(a) AD ID 019-4477 based on original BOLD signals



(b) AD ID 019-4477 based on low-rank matrix



(c) AD ID 130-5006 based on original BOLD signals



(d) AD ID 130-5006 based on low-rank matrix

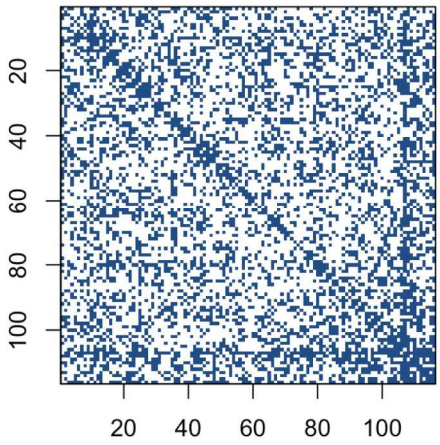
Figure 3.6: Connectivity Selection Results for AD Group from Partial Correlation with Thresholding. For partial correlation with hard threshold value, the output patterns based on low rank matrices keep decent similarity levels as the output based on original BOLD signals, but less similarity compared with sparse correlation method outputs.

keep decent similarity levels as the output based on original BOLD signals, but less similarity compared with sparse correlation method outputs as shown in Figure 3.3 and Figure 3.4.

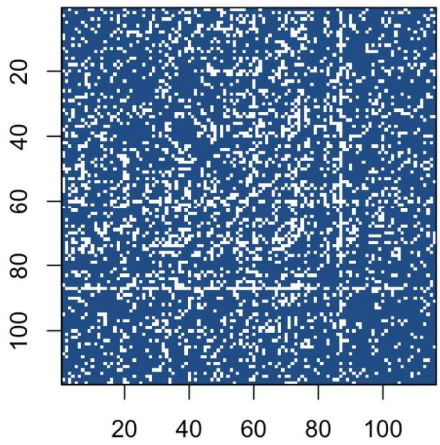
Graphical Lasso was also conducted to obtain the estimated sparse precision matrix automatically due to the method. In our study the value ρ in (1) was chosen as 0.1 in order to get the reasonable sparsity, and also to achieve the comparable sparsity with other methods for construction. The results are shown in Figure 3.7 and Figure 3.8.

From Figure 3.7 and Figure 3.8, we can see that for Graphical Lasso method, the sparse precision matrices obtained based on low-rank matrices are less sparser than based on original BOLD signals, if we adopt the same parameter value when conducting Graphical Lasso. Furthermore, due to the relationship between the precision matrix and the partial correlation matrix that was mentioned in section 2.2, we know that if we draw a graph for the derived sparse partial correlation output based on Graphical Lasso method, it would be exactly the same as sparse precision matrix shown in Figure 3.7 and Figure 3.8. Hence we did not demonstrate the corresponding output graphs for sparse partial correlation based on Graphical Lasso.

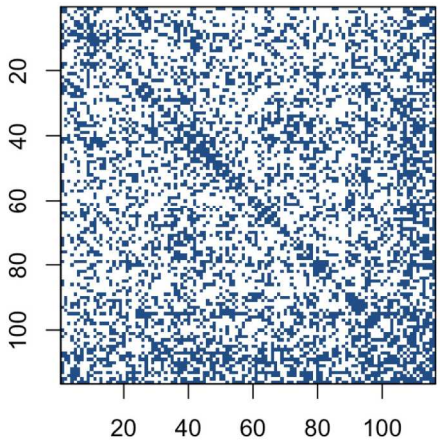
We then built the functional brain connectivities using the recently proposed novel concentration inequality based estimation method of sparse covariance matrices. We chose false positive rate α=0.35 in order to get the reasonable sparsity for the estimation of sparse covariance matrices, and also to achieve the comparable sparsity with other methods for construction. The sparsity of connectivity matrices for the concentration inequality based estimation method is around 64%. Our sparse estimators $\hat{\Sigma}^{sp}$ results using concentration inequality based method for both original BOLD signals and low



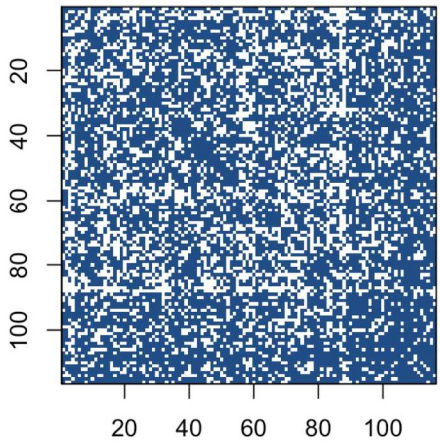
(a) Control ID 002-4225 based on original BOLD signals



(b) Control ID 002-4225 based on low-rank matrix

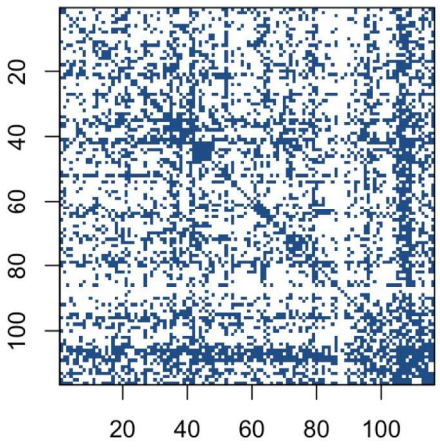


(c) Control ID 006-4357 based on original BOLD signals

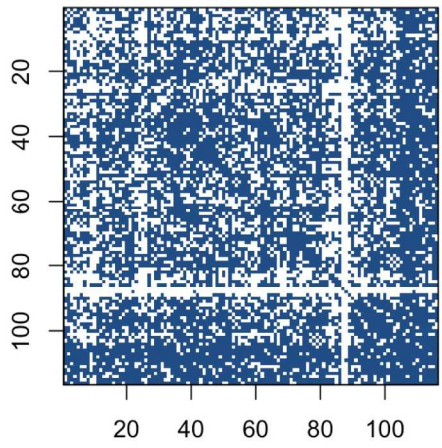


(d) Control ID 006-4357 based on low-rank matrix

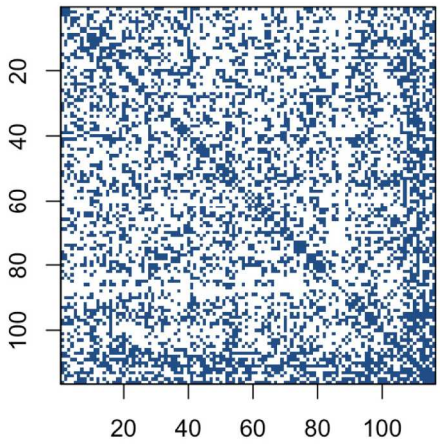
Figure 3.7: Connectivity Selection Results for Control Group from GLasso. The sparse precision matrices obtained based on low-rank matrices for control group are less sparser than based on original BOLD signals.



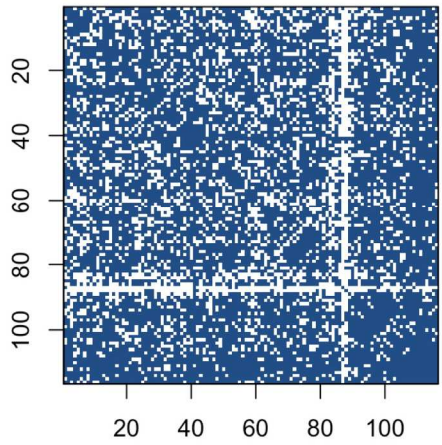
(a) AD ID 019-4477 based on original BOLD signals



(b) AD ID 019-4477 based on low-rank matrix



(c) AD ID 130-5006 based on original BOLD signals



(d) AD ID 130-5006 based on low-rank matrix

Figure 3.8: Connectivity Selection Results for AD Group from GLasso. The sparse precision matrices obtained based on low-rank matrices for AD group are less sparser than based on original BOLD signals.

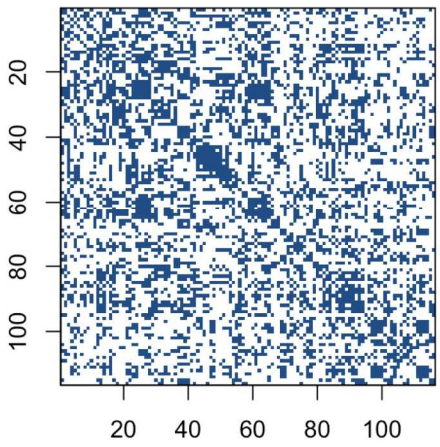
rank matrices are shown in Figure 3.9 and Figure 3.10.

As is evident from the comparison of Figure 3.3-3.4 and Figure 3.9-3.10, the sparse correlation method and this novel sparse covariance estimation method supply consistent connectivity selection outputs for both control group subjects and AD group subjects. This indicates that the recently proposed novel concentration inequality based estimation method performs well in terms of the estimation of sparse covariance matrices, for both cases based on original BOLD signals and low rank matrices. Therefore we declared that this method supplies us with a novel and efficient way for studying functional brain connectivities. Through controlling the parameter false positive rate α in this method, we can achieve the desired sparsity for our estimation of sparse covariance matrices.

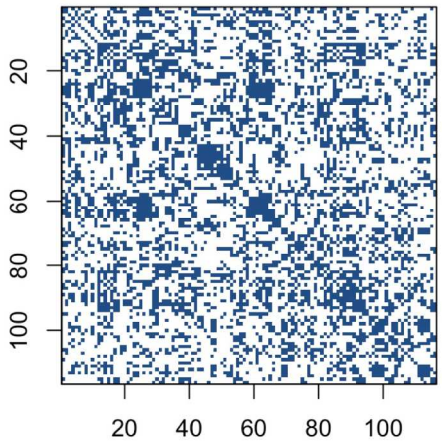
Furthermore, as can be seen in Figure 3.9 and Figure 3.10, the patterns of connectivity matrices based on both the original BOLD signals and low rank matrices for both groups are still quite similar, which indicates that low rank plus sparse decomposition method can identify the essential common features while still retaining most of features for concentration inequality-based estimation method.

3.3 Comparison and Discussion

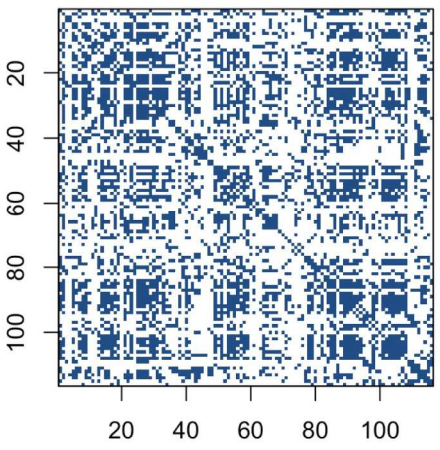
In this section, we first implemented statistical test for each ROI connection to reveal and identify the underlying group differences. After getting the p value matrices based on both original BOLD signals and low rank matrices, we calculated the percentage of significant connections and compared the results. We also obtained the first ten most significant connections for differentiating



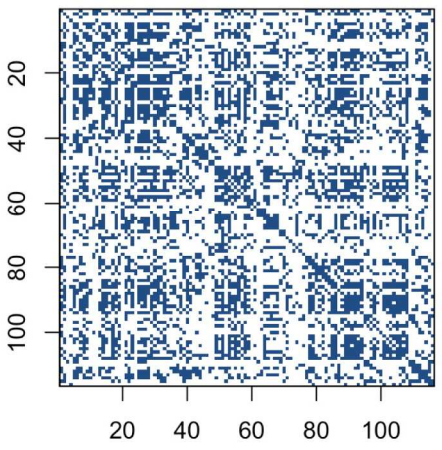
(a) Control ID 002-4225 based on original BOLD signals



(b) Control ID 002-4225 based on low-rank matrix

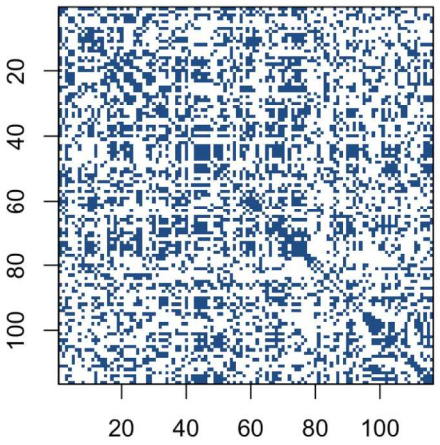


(c) Control ID 006-4357 based on original BOLD signals

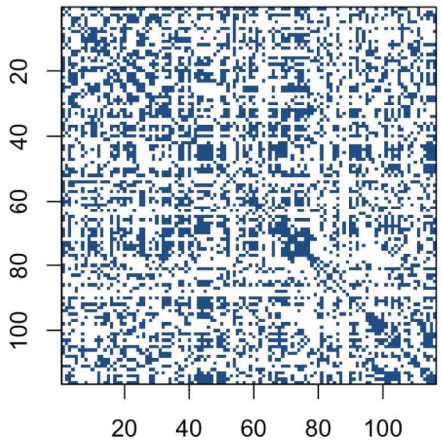


(d) Control ID 006-4357 based on low-rank matrix

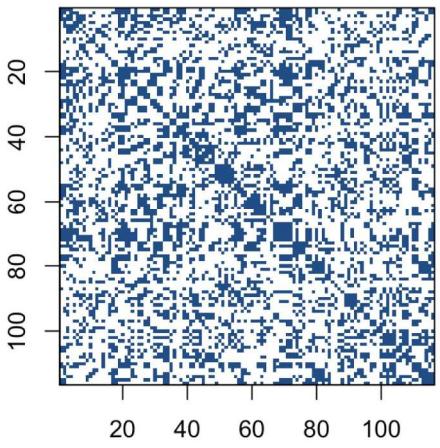
Figure 3.9: Connectivity Selection Results for Control Group from Sparse Covariance Estimation Method. From the comparison with Figure 3.3, the sparse correlation method and this novel sparse covariance estimation method supply consistent connectivity selection outputs for control group subjects. Furthermore, the patterns of connectivity matrices based on both original BOLD signals and low rank matrices are still quite similar.



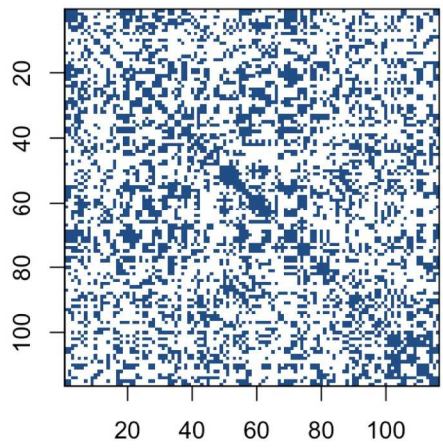
(a) AD ID 019-4477 based on original BOLD



(b) AD ID 019-4477 based on low-rank matrix signals



(c) AD ID 130-5006 based on original BOLD



(d) AD ID 130-5006 based on low-rank matrix signals

Figure 3.10: Connectivity Selection Results for AD Group from Sparse Covariance Estimation Method. From the comparison with Figure 3.4, the sparse correlation method and this novel sparse covariance estimation method supply consistent connectivity selection outputs for AD group subjects. Furthermore, the patterns of connectivity matrices based on both original BOLD signals and low rank matrices are still quite similar.

group differences for Alzheimer's disease. Then we defined and calculated the overlap rate in order to see whether performing low-rank plus sparse matrix decomposition can keep some level of consistency compared with when using original BOLD signals. Lastly we implemented bootstrapping. By calculating the variances for percentages of significant connections in p value matrices for 50 times bootstrapping, we explored whether performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing functional brain connectivities.

From the previous outputs and analyses, we have seen that for each single subject and for each method we employed for construction, we already derived connectivity results based on both original BOLD signals and low rank matrices. Subsequently we conducted two sample T statistical test for each single ROI connection, in order to reveal and identify the underlying group differences between control group and AD group. In our study, we have 6670 total connections $((116 \times (116-1)/2))$. Two sample T statistical test was implemented to each single connection $c_{p,q}$ of p-th ROI and q-th ROI. Then we derived the p value matrix with 116×116 dimension for each single method we used, based on both original BOLD signals and low rank matrices. In the following analysis, we ignored the diagonal elements since the diagonal elements do not make sense, and also only focused on the elements in upper triangle since it is symmetric. Once we get the p value matrix, we need a reasonable significant level α. The adjusted p values are normally used in multiple comparisons [58]. But in our study we did not employ adjusted p values. The reason is the overlap rates for significant connection locations based on original BOLD signals and low rank matrices start getting stable from threshold value 0.04 onwards, which we illustrated more in later analysis. Hence, we

chose 0.05 as the significant level. Figure 3.11 and 3.12 are examples for p value matrices, where dark blue dots represent the locations for the p values less than 0.05, which indicates the significant connection locations. The p value matrices supply us the significant connection locations for differentiating control group and AD group, which can intensely contribute to uncovering underlying neurological processes associated with the Alzheimer's disease for clinical use.

After getting the p value matrices, we firstly focused on the quantities of significant ROI connections out of 6670 total connections. That is to say, we explored how many connections can be considered significant in terms of distinguishing the differences between normal group and AD group. We defined and calculated the percentage of significant connections and compared the results based on both original BOLD signals and low rank matrices. Under significant level 0.05, we define the percentage of significant connections of p value matrix as the number of elements smaller than 0.05 / 6670. Then we have the following comparison results based on both original BOLD signals and low rank matrices for all the methods we employed in our study, shown in Table 3.1.

From Table 3.1, we can see that except the correlation method and Glasso partial correlation method, all the other methods for construction based on low rank matrices have higher percentage of significant connections than based on the original BOLD signals. This suggests that the former behaves better than based on the original BOLD signals, in the sense that it can reveal and identify more significant ROI connections between the control group and AD group. Therefore, we verified performing low-rank plus sparse matrix decomposition can help us achieve more differentiable results for functional brain

Methods	Original BOLD Signals	Low Rank Matrix
Correlation	6.3868	6.3718
Sparse Correlation	5.7121	5.9970
Partial Correlation	3.2684	4.5577
Sparse Partial Correlation	3.6732	4.7826
Glasso Precision	4.6927	4.9775
Glasso Partial Correlation	4.7077	4.7077
Sparse Covariance Estimation	5.8321	5.9220

Table 3.1: Percentage of Significant Connections. Almost all the methods for construction based on low rank matrices have higher percentage of significant connections than based on original BOLD signals. This suggests that the former behaves better than based on original BOLD signals, in the sense that it can reveal and identify more significant ROI connections between control group and AD group. Furthermore, the recently proposed concentration inequality based method performs better overall compared with other methods.

connectivity. Furthermore, we can also see that the recently proposed concentration inequality based method performs better overall compared with other methods.

Subsequently, we demonstrated the first ten most significant pairs of ROIs with first ten smallest p value for both sparse correlation method and concentration inequality based estimation, both based on low rank matrices, shown as Table 3.2 and Table 3.3. These two tables give us the most important ROI pairs for distinguishing normal subject and AD subject we should pay more attention to in clinical use. We can also see that the connection between left Hippocampus region and left Cerebellum 7 region is the most significant ROI connection for differentiating the normal group and AD group for Alzheimer's disease, with p value smaller than 0.0005. Researchers have shown that the cerebellum has a strong role in higher cognitive functions which include memory processes, and possibly serves long-term memory encoding and information storage [25]. It has been also demonstrated that AD patients showed abnor-

Pair	Region 1	Classification	Region 2	Classification	p value
1	L.HIP	Limbic lobe	L.CER7	Cerebellum	0.00010
2	L.CER45	Cerebellum	VER7	Vermis	0.00045
3	L.MFG	Frontal	L.IFGtriang	Frontal	0.00049
4	R.SFGdor	Frontal	L.IFGtriang	Frontal	0.00062
5	R.ANG	Parietal	VER8	Vermis	0.00096
6	R.INS	Insula	R.SMG	Parietal	0.00101
7	R.ORBsupmed	Frontal	L.ITG	Temporal	0.00104
8	L.SMA	Frontal	R.CUN	Occipital	0.00110
9	L.IFGoperc	Frontal	R.IFGtriang	Frontal	0.00113
10	L.SFGdor	Frontal	VER9	Vermis	0.00123

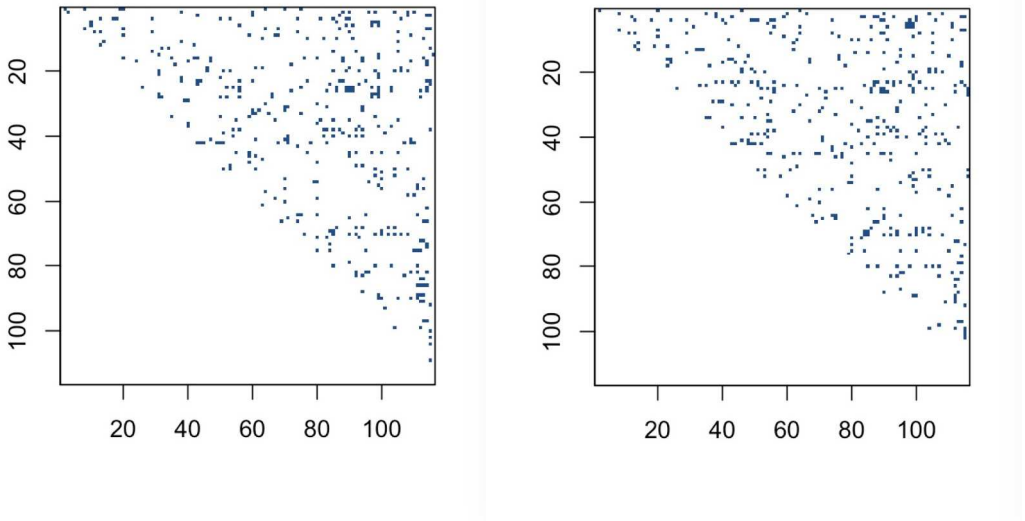
Table 3.2: The First 10 Most Significant Pairs for Sparse Correlation Based on Low-rank Matrices. The connection between left Hippocampus region and left Cerebellum 7 region is the most significant ROI connection for differentiating normal group and AD group for Alzheimer's disease, with p value 0.00010, which is consistent with existing literatures' findings.

mal hippocampal connectivity during resting state [54]. Other research has illustrated that the connectivity between hippocampus and cerebellum area is significantly different for control group and AD group [8]. Therefore our finding here is consistent with existing literatures' findings.

We then explored whether performing low-rank plus sparse matrix decomposition when studying functional brain connectivity can keep some level of consistency compared with when using original data matrices. In order to see qualitatively the distribution comparison of those significant connection locations based on original BOLD signals and low rank matrices, we drew significant connection location graphs for sparse correlation method and sparse covariance estimation method, which are shown in Figure 3.11 and Figure 3.12. The dark blue dots represent the significant connection locations, or the locations for the p values less than 0.05 in the p value matrix.

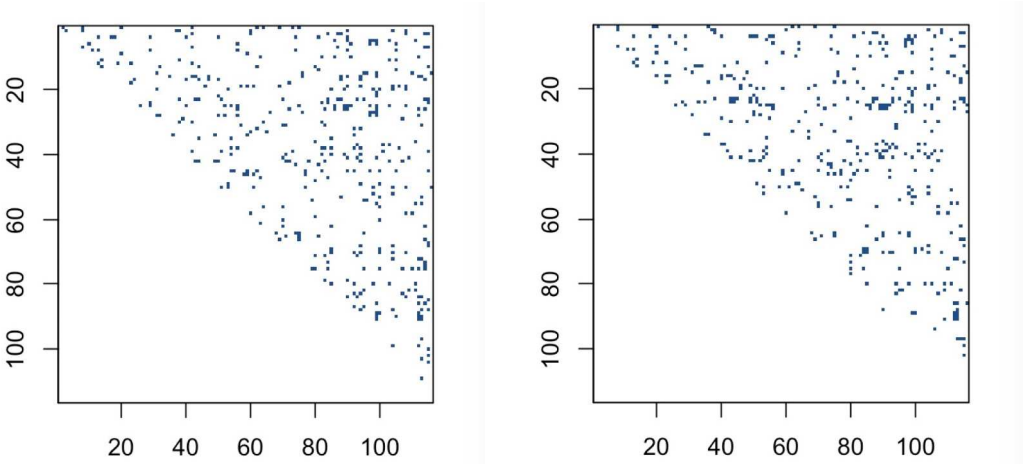
Pair	Region 1	Classification	Region 2	Classification	p value
1	L.HIP	Limbic lobe	L.CER7	Cerebellum	0.00015
2	L.MFG	Frontal	L.IFGtriang	Frontal	0.00020
3	R.REC	Frontal	R.SOG	Occipital	0.00028
4	L.SMA	Frontal	R.CUN	Occipital	0.00033
5	L.CAU	Corpus striatum	L.TPOsup	Limbic	0.00062
6	R.SFGdor	Frontal	L.IFGtriang	Frontal	0.00085
7	L.ITG	Temporal	VER6	Vermis	0.00101
8	R.ORBsupmed	Frontal	R.ITG	Temporal	0.00121
9	L.OLF	Frontal	L.CER6	Cerebellum	0.00140
10	R.ANG	Parietal	VER8	Vermis	0.00163

Table 3.3: The First 10 Most Significant Pairs for Concentration Inequality Based Estimation Method Based on Low-rank Matrices. The connection between left Hippocampus region and left Cerebellum 7 region is the most significant ROI connection for differentiating the normal group and AD group for Alzheimer's disease, with p value 0.00015, which is consistent with existing literatures' findings.



(a) Based on original BOLD signals (b) Based on low-rank matrices

Figure 3.11: Significant Connection Location Detecting for Sparse Correlation Method. The significant connection locations detected for sparse correlation method have visually similar distribution, which indicates decent overlap, for both based on original BOLD signals and low rank matrices.



(a) Based on original BOLD signals

(b) Based on low-rank matrices

Figure 3.12: Significant Connection Location Detecting for Sparse Covariance Estimation Method. The significant connection locations detected for sparse covariance estimation method have visually similar distribution, which indicates decent overlap, for both based on original BOLD signals and low rank matrices.

As shown in Figure 3.11 and Figure 3.12, the significant connection locations detected have visually similar distribution, which indicates decent overlap, for both based on original BOLD signals and low rank matrices, for sparse correlation method and also for sparse covariance estimation method.

Furthermore, in order to go a step further to quantitatively check significant connection location distribution to see the overlap status for both original BOLD signals and low rank matrices, we defined and calculated the overlap rate. Denote $p^{original}$ as the p value matrix based on original BOLD signals, and $p^{lowrank}$ as the p value matrix based on low rank matrices. Here we only focused on the upper triangle of the symmetric p value matrices. For the counterpart elements $p^{original}_{(i,j)}$ and $p^{lowrank}_{(i,j)}$, where $i < j$, we denote n_1 as the number of elements of upper triangle which satisfy $0.05 > p^{original}_{(i,j)} > p^{lowrank}_{(i,j)}$, n_2 as the number of elements of upper triangle which satisfy $p^{original}_{(i,j)} < p^{lowrank}_{(i,j)} < 0.05$,

and $n_{nonzero}^{original}$ the number of nonzero of the upper triangle of p value matrix based on original data. Then we define overlap rate as

$$(n_1 + n_2)/n_{nonzero}^{original}.$$

Denote n_3 as the number of elements which satisfy $p_{(i,j)}^{original} > 0.05$ and $p_{(i,j)}^{lowrank} < 0.05$, and $n_{nonzero}^{lowrank}$ the number of nonzero of the upper triangle of p value matrix based on low rank matrices, then we define a smaller p value rate as

$$(n_1 + n_3)/n_{nonzero}^{lowrank}.$$

The results for overlap rate and smaller p value rate based on different methods we utilized in our study are shown in Table 3.4. From Table 3.4 we can see that the overlap rates for correlation method, sparse correlation method, and sparse covariance estimation method are relatively large enough, whilst the other methods are not. Furthermore, the concentration inequality based estimation has slightly higher overlap rate and smaller p value rate than sparse correlation method. These findings indicate that building brain connectivity based on the low rank matrices when using correlation method, sparse correlation method, and sparse covariance estimation method, can achieve decent level of consistency, in the sense of the overlap status compared with the outputs based on original BOLD signals. And the concentration inequality based estimation has better consistency result than sparse correlation. Hence we verified that performing low-rank plus sparse matrix decomposition when we study functional brain connectivity can keep decent level of consistency with when using original BOLD signals.

Furthermore, we also calculated the overlap rates under different threshold values for correlation method, sparse correlation method, and sparse covari-

Methods	Overlap Rate	Smaller p Value Rate
Correlation	77.93	60.24
Sparse Correlation	58.53	69.75
Partial Correlation	4.13	99.34
Sparse Partial Correlation	5.31	98.12
Glasso Precision	6.39	96.39
Glasso Partial Correlation	5.41	96.82
Sparse Covariance Estimation	59.90	72.15

Table 3.4: Overlap Rate (%). The overlap rates for correlation method, sparse correlation method, and sparse covariance estimation method are relatively large enough, whilst the other methods are not. Furthermore, the concentration inequality based estimation has slightly higher overlap rate and smaller p value rate than sparse correlation method. These findings indicate that building brain connectivity based on low rank matrices when using these three methods can achieve decent level of consistency. Moreover, the concentration inequality based estimation has better consistency result than the sparse correlation.

ance estimation method. The results are shown in Table 3.5, Table 3.6, Table 3.7. We further drew a line chart for overlap rate to better demonstrate the results included in Table 3.5-3.7, shown as Figure 3.13. From the line chart we can see that the significant level 0.05 chosen in our study is reasonable as the overlap rates start getting more stable from threshold value 0.04 onwards.

In order to verify whether performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing functional brain connectivities, we conducted bootstrapping. Specifically, we implemented bootstrapping for 50 times, and each time we sampled 33 subjects out of 33 subjects in control group and 24 subjects out of 24 subjects in AD group, both with replacement. For each resampling, we conducted two sample T statistical tests to derive the p value matrices based on both original BOLD signals and low rank matrices, as shown in Table 3.8. We implemented this process for all the methods we employed for construction. As we stated above, for each p value matrix we have a percentage of significant connections. Then for 50

Threshold	Overlap Rate	# Of Connections
0.001	50.00	10,9
0.002	64.00	25,26
0.003	75.00	32,33
0.004	72.50	40,46
0.005	69.23	52,52
0.006	68.85	61,63
0.007	68.92	74,68
0.008	69.05	84,75
0.009	68.09	94,84
0.01	66.99	103,92
0.02	76.44	191,190
0.03	77.78	279,265
0.04	77.30	348,344
0.05	77.93	426,425
0.06	75.30	498,489
0.07	76.80	569,556
0.08	76.32	642,620
0.09	77.29	709,689
0.1	78.66	792,762

Table 3.5: Overlap Rate for Different Threshold Value for Correlation Method.

Threshold	Overlap Rate	# Of Connections
0.001	25.00	4,4
0.002	37.50	8,10
0.003	35.00	20,14
0.004	39.29	28,22
0.005	37.21	43,31
0.006	40.00	50,40
0.007	38.98	59,47
0.008	39.71	68,51
0.009	40.54	74,54
0.01	37.50	88,60
0.02	48.39	155,152
0.03	52.59	232,227
0.04	57.24	304,313
0.05	58.53	381,400
0.06	59.47	449,459
0.07	59.16	524,527
0.08	61.41	596,602
0.09	61.14	682,679
0.1	62.25	747,748

Table 3.6: Overlap Rate for Different Threshold Value for Sparse Correlation Method.

Threshold	Overlap rate	# Of Connections
0.001	50.00	4,6
0.002	42.86	7,12
0.003	21.43	14,18
0.004	36.00	25,23
0.005	35.90	39,31
0.006	39.53	43,37
0.007	36.00	50,45
0.008	35.09	57,53
0.009	38.71	62,64
0.01	36.62	71,70
0.02	49.69	161,159
0.03	57.56	238,248
0.04	58.20	311,316
0.05	59.90	389,395
0.06	58.57	461,459
0.07	60.84	526,543
0.08	61.69	603,615
0.09	63.68	691,691
0.1	64.90	755,758

Table 3.7: Overlap Rate for Different Threshold Value for Sparse Covariance Estimation.

Figure 3.13: Overlap Rate for Different Threshold Values. The significant level 0.05 chosen in our study is reasonable as we can see the overlap rates start getting more stable from the threshold value 0.04 onwards.

Bootstrapping	Based on Original BOLD Signals	Based on Low Rank Matrix
1	$p_{1,1}$	$p_{2,1}$
2	$p_{1,2}$	$p_{2,2}$
3	$p_{1,3}$	$p_{2,3}$
$\vdots$	$\vdots$	$\vdots$
50	$p_{1,50}$	$p_{2,50}$

Table 3.8: p Value Matrices Scheme. For each time resampling, we conducted two sample T statistical tests to derive the p value matrices based on both original BOLD signals and low rank matrices. We implemented this process for all the methods we employed for construction.

Methods	Original BOLD Signals	Low Rank Matrix
Correlation	5.6443	4.7270
Sparse Correlation	5.6554	5.0379
Partial Correlation	3.6149	3.4020
Sparse Partial Correlation	4.4523	3.6015
Glasso Precision	4.0707	3.8987
Glasso Partial Correlation	3.0966	3.0435
Sparse Covariance Estimation	5.1630	4.0835

Table 3.9: Variance Comparisons of Percentages of Significant Connections for 50 Times Bootstrapping (unit: $\times 10^{-4}$). The variances of percentages of significant connections based on low rank matrices are all smaller than those based on original BOLD signals, for all the methods we employed for construction. This result indicates that performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing brain functional connectivities.

times bootstrapping, we have 50 percentages of significant connections. Thus we calculated the variance for percentages of significant connections, and the results are demonstrated in Table 3.9. As seen in Table 3.9, the variances of percentages of significant connections based on low rank matrices are all smaller than those based on original BOLD signals, for all the methods we employed for construction in our study. This result indicates that performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing functional brain connectivities.

Chapter 4

Conclusion and Future Work

In this thesis, we proposed a framework which enables us to provide robust and stable functional brain connectivity. In particular, a low-rank plus sparse matrix decomposition technique was adapted to decompose the resting state fMRI BOLD signals from ADNI data, where the low-rank matrix L recovers the essential common features from regions of interest (ROIs), and the sparse matrix S catches the sparse individual variability and potential outliers. We applied our construction methods based on low rank matrices from decomposition and compared the results with those based on original BOLD signals. We discovered that the methods for building connectivities based on low rank matrices behave better than based on original BOLD signals, in the sense that the methods we employed for constructing connectivities based on low rank matrices can reveal and identify more significant ROI connections for group differences. Hence this suggests that when we study group difference for functional brain connectivity, performing low-rank plus sparse matrix decomposition can achieve more differentiable results, which can contribute to uncovering underlying neurological processes associated with the disease for

clinical use.

We found that the recently proposed concentration inequality based method performs better overall compared with correlation, partial correlation, and Graphical Lasso method. It supplies us a novel and efficient way to explore functional brain connectivity.

We also obtained the first ten most significant connections for differentiating group differences for Alzheimer's disease. Among them, the left hippocampus region and the left cerebellum 7 region is the most significant one, with p value smaller than 0.0005, which is consistent with existing literatures' findings.

Furthermore, through bootstrapping, we verified that performing low-rank plus sparse matrix decomposition can achieve more stable results for constructing functional brain connectivities.

In the future we will repeat our whole study process using simulation data. We will try to verify the robustness property of our proposed framework by comparing the outputs based on both before and after removing outliers. Additionally, the sparsity levels of the connection matrices might influence the results of statistical test for group differences, hence we will look into this interesting question in the future study.

Bibliography

[1] Alzheimer's disease neuroimaging initiative. http://adni.loni.usc.edu/.

[2] Monash biomedical imaging. https://www.monash.edu/researchinfrastructure/mbi/facilities/human/3t-mri.

[3] MRI scanner protocols. http://adni.loni.usc.edu/methods/documents/mri-protocols/.

[4] Stages of Alzheimer's. https://www.alz.org/alzheimers-dementia/stages.

[5] What is fMRI? http://fmri.ucsd.edu/Research/whatisfmri.html.

[6] Agarwal, M. and R. Mehra (2014). Review of matrix decomposition techniques for signal processing applications. *Int. Journal of Engineering Research and Applications 4*(1), 90–93.

[7] Aisen, P. S., R. C. Petersen, M. C. Donohue, A. Gamst, R. Raman, R. G. Thomas, S. Walter, J. Q. Trojanowski, L. M. Shaw, L. A. Beckett, et al. (2010). Clinical core of the Alzheimer's disease neuroimaging initiative: progress and plans. *Alzheimer's and Dementia 6*(3), 239–246.

[8] Allen, G., H. Barnard, R. McColl, A. L. Hester, J. A. Fields, M. F. Weiner, W. K. Ringe, A. M. Lipton, M. Brooker, E. McDonald, et al. (2007). Re-

duced hippocampal functional connectivity in Alzheimer disease. *Archives of Neurology 64*(10), 1482–1487.

[9] Alzheimer's Association (2015). 2015 Alzheimer's disease facts and figures. *Alzheimer's & dementia: the Journal of the Alzheimer's Association 11*(3), 332.

[10] Banerjee, O., L. E. Ghaoui, and A. d'Aspremont (2008). Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *Journal of Machine Learning Research 9*(Mar), 485–516.

[11] Biswal, B., F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine 34*(4), 537–541.

[12] Biswal, B. B., J. V. Kylen, and J. S. Hyde (1997). Simultaneous assessment of flow and bold signals in resting-state functional connectivity maps. *NMR in Biomedicine 10*(4-5), 165–170.

[13] Bollobás, B. (2013). *Modern graph theory*, Volume 184. Springer Science & Business Media.

[14] Bouwmans, T. and E. H. Zahzah (2014). Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance. *Computer Vision and Image Understanding 122*, 22–34.

[15] Cai, T. and W. Liu (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association 106*(494), 672–684.

[16] Candès, E. J., X. Li, Y. Ma, and J. Wright (2011). Robust principal component analysis? *Journal of the ACM (JACM) 58*(3), 11.

[17] Cohen, M. S. and S. Y. Bookheimer (1994). Localization of brain function using magnetic resonance imaging. *Trends in Neurosciences 17*(7), 268–277.

[18] Coran, A. G., A. Caldamone, N. S. Adzick, T. M. Krummel, J.-M. Laberge, and R. Shamberger (2012). *Pediatric Surgery E-Book*, Volume 2. Elsevier Health Sciences.

[19] Croux, C. and G. Haesbroeck (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. *Biometrika 87*(3), 603–618.

[20] Croux, C. and A. Ruiz-Gazen (2005). High breakdown estimators for principal components: the projection-pursuit approach revisited. *Journal of Multivariate Analysis 95*(1), 206–226.

[21] De la Torre, F. and M. J. Black (2001). Robust principal component analysis for computer vision. *In International Conference on Computer Vision*, 362–369.

[22] De La Torre, F. and M. J. Black (2003). A framework for robust subspace learning. *International Journal of Computer Vision 54*(1-3), 117–142.

[23] Dennis, E. L. and P. M. Thompson (2014). Functional brain connectivity using fMRI in aging and Alzheimer's disease. *Neuropsychology Review 24*(1), 49–62.

[24] Fan, J., Y. Liao, and H. Liu (2016). An overview of the estimation of

large covariance and precision matrices. *The Econometrics Journal 19*(1), C1–C32.

[25] Filippini, N., B. J. MacIntosh, M. G. Hough, G. M. Goodwin, G. B. Frisoni, S. M. Smith, P. M. Matthews, C. F. Beckmann, and C. E. Mackay (2009). Distinct patterns of brain activity in young carriers of the apoe-ε4 allele. *Proceedings of the National Academy of Sciences 106*(17), 7209–7214.

[26] Fox, M. D. and M. E. Raichle (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience 8*(9), 700.

[27] Friedman, J., T. Hastie, and R. Tibshirani (2008). Sparse inverse covariance estimation with the graphical Lasso. *Biostatistics 9*(3), 432–441.

[28] Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connectivity 1*(1), 13–36.

[29] Holdsworth, S. J. and R. Bammer (2008). Magnetic resonance imaging techniques: fMRI, DWI, and PWI. *Seminars in Neurology 28*(4), 395.

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

[31] Jack Jr, C. R., M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, et al. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI meth-

ods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine 27*(4), 685–691.

[32] Jolliffe, I. T. and J. Cadima (2016). Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A 374*(2065), 20150202.

[33] Kang, Z., C. Peng, and Q. Cheng (2015). Robust PCA via nonconvex rank approximation. *arXiv preprint arXiv:1511.05261*.

[34] Kashlak, A. B. and L. Kong (2017). Nonasymptotic estimation and support recovery for high dimensional sparse covariance matrices. *arXiv preprint arXiv:1705.02679*.

[35] Kim, J., W. Pan, Alzheimer's Disease Neuroimaging Initiative, et al. (2015). Highly adaptive tests for group differences in brain functional connectivity. *NeuroImage: Clinical 9*, 625–639.

[36] Kim, J., J. R. Wozniak, B. A. Mueller, and W. Pan (2015). Testing group differences in brain functional connectivity: using correlations or partial correlations? *Brain Connectivity 5*(4), 214–231.

[37] Konrad, K. and S. B. Eickhoff (2010). Is the ADHD brain wired differently? a review on structural and functional connectivity in attention deficit hyperactivity disorder. *Human Brain Mapping 31*(6), 904–916.

[38] Lage, J. M. M. (2006). 100 years of Alzheimer's disease (1906–2006). *Journal of Alzheimer's Disease 9*(s3), 15–26.

[39] Landini, L., V. Positano, and M. Santarelli (2005). *Advanced image processing in magnetic resonance imaging*. CRC press.

[40] Li, L., M. Cazzell, O. M. Babawale, and H. Liu (2016). Automated voxel classification used with atlas-guided diffuse optical tomography for assessment of functional brain networks in young and older adults. *Neurophotonics 3*(4), 045002.

[41] Li, W., M. Wang, W. Zhu, Y. Qin, Y. Huang, and X. Chen (2016). Simulating the evolution of functional brain networks in Alzheimer's disease: exploring disease dynamics from the perspective of global activity. *Scientific Reports 6*.

[42] Lin, Z., M. Chen, and Y. Ma (2010). The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv preprint arXiv:1009.5055*.

[43] MacDonald, J. P., D. E. Barnes, and L. E. Middleton (2015). Implications of risk factors for Alzheimer's disease in canada's indigenous population. *Canadian Geriatrics Journal 18*(3), 152.

[44] Magnotti, J. F. and N. Billor (2014). Finding multivariate outliers in fMRI time-series data. *Computers in Biology and Medicine 53*, 115–124.

[45] Narayan, M. and G. I. Allen (2016). Mixed effects models for resampled network statistics improves statistical power to find differences in multi-subject functional connectivity. *Frontiers in Neuroscience 10*, 108.

[46] Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience 2*(1), 67–70.

[47] Rothman, A. J., E. Levina, and J. Zhu (2009). Generalized thresholding

of large covariance matrices. *Journal of the American Statistical Association 104*(485), 177–186.

[48] Sanz-Arigita, E. J., M. M. Schoonheim, J. S. Damoiseaux, S. A. Rombouts, E. Maris, F. Barkhof, P. Scheltens, and C. J. Stam (2010). Loss of 'small-world' networks in Alzheimer's disease: graph analysis of fMRI resting-state functional connectivity. *PloS one 5*(11), e13788.

[49] Sato, J. R., A. Fujita, E. F. Cardoso, C. E. Thomaz, M. J. Brammer, and E. Amaro Jr (2010). Analyzing the connectivity between regions of interest: an approach based on cluster granger causality for fMRI data analysis. *Neuroimage 52*(4), 1444–1455.

[50] Smith, S. M., K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich (2011). Network modelling methods for fMRI. *Neuroimage 54*(2), 875–891.

[51] Tzourio-Mazoyer, N., B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage 15*(1), 273–289.

[52] Uddin, L. Q., K. Supekar, and V. Menon (2013). Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Frontiers in Human Neuroscience 7*, 458.

[53] Van Den Heuvel, M. P. and H. E. H. Pol (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology 20*(8), 519–534.

[54] Wang, L., Y. Zang, Y. He, M. Liang, X. Zhang, L. Tian, T. Wu, T. Jiang, and K. Li (2006). Changes in hippocampal connectivity in the early stages of Alzheimer's disease: evidence from resting state fMRI. *Neuroimage 31*(2), 496–504.

[55] World Health Organization (2006). *Neurological disorders: public health challenges*. World Health Organization.

[56] Wozniak, J. R., B. A. Mueller, C. J. Bell, R. L. Muetzel, H. L. Hoecker, C. J. Boys, and K. O. Lim (2013). Global functional connectivity abnormalities in children with fetal alcohol spectrum disorders. *Alcoholism: Clinical and Experimental Research 37*(5), 748–756.

[57] Wright, J., A. Ganesh, S. Rao, Y. Peng, and Y. Ma (2009). Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. *Advances in Neural Information Processing Systems*, 2080–2088.

[58] Wright, S. P. (1992). Adjusted p-values for simultaneous inference. *Biometrics*, 1005–1013.

[59] Xu, L. and A. L. Yuille (1995). Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE Transactions on Neural Networks 6*(1), 131–143.

Appendix

Anatomical parcellation of the brain and their abbreviations

Region	Abbreviation	Region	Abbreviation
Central region		**Limbic lobe**	
Precental gyrus	PreCG	Anterior cingulate and paracingulate gyri	ACG
Rolandic operculum	ROL	Median cingulate and paracingulate gyri	DCG
Postcentral gyrus	PoCG	Posterior cingulate gyrus	PCG
Frontal lobe		Hippocampus	HIP
Superior frontal gyrus, dorsolateral	SFGdor	Parahippocampal gyrus	PHG
Superior frontal gyrus, orbital part	ORBsup	Temporal pole: superior temporal gyrus	TPOsup
Middle frontal gyrus	MFG	Temporal pole: middle temporal gyrus	TPOmid
Middle frontal gyrus, orbital part	ORBmid	**Corpus striatum**	
Inferior frontal gyrus, opercular part	IFGoperc	Amygdala	AMYG
Inferior frontal gyrus, triangular part	IFGtriang	Caudate nucleus	CAU
Inferior frontal gyrus, orbital part	ORBinf	Lenticular nucleus, putamen	PUT
Supplementary motor area	SMA	Lenticular nucleus, pallidum	PAL
Olfactory cortex	OLF	Thalamus	THA
Superior frontal gyrus, medial	SFGmed	**Insula**	INS
Superior frontal gyrus, medial orbital	ORBsupmed	**Cerebellum**	
Gyrus rectus	REC	Cerebellum crus 1	CERC1
Paracentral lobule	PCL	Cerebellum crus 2	CERCR2
Temporal lobe		Cerebellum 3	CER3
Heschl gyrus	HES	Cerebellum 4 5	CER45
Superior temporal gyrus	STG	Cerebellum 6	CER6
Middle temporal gyrus	MTG	Cerebellum 7	CER7
Inferior temporal gyrus	ITG	Cerebellum 8	CER8
Parietal lobe		Cerebellum 9	CER9
Superior parietal gyrus	SPG	Cerebellum 10	CER10
Inferior parietal, but supramarginal and angular gyri	IPL	**Vermis**	
Supramarginal gyrus	SMG	Vermis 1 2	VER12
Angular gyrus	ANG	Vermis 3	VER3
Precuneus	PCUN	Vermis 4 5	VER45
Occipital lobe		Vermis 6	VER6
Calcarine fissure and surrounding cortex	CAL	Vermis 7	VER7
Cuneus	CUN	Vermis 8	VER8
Lingual gyrus	LING	Vermis 9	VER9
Superior occipital gyrus	SOG	Vermis 10	VER10
Middle occipital gyrus	MOG		
Inferior occipital gyrus	IOG		
Fusiform gyrus	FFG		