STATISTICAL LEARNING METHODS FOR DIFFUSION MAGNETIC RESONANCE IMAGING

Xifeng Wang

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics in the Gillings School of Global Public Health.

Chapel Hill
2021

Approved by:

Hongtu Zhu

Joseph G. Ibrahim

Pew-Thian Yap

Martin Styner

Quefeng Li

ABSTRACT

Xifeng Wang: Statistical Learning Methods for Diffusion Magnetic Resonance Imaging
(Under the direction of Hongtu Zhu)

Diffusion Magnetic Resonance Imaging (dMRI) is a commonly used imaging technique to reveal white matter (WM) microstructure by probing the diffusion of water molecules. The diffusion of water molecules is constrained by the biological boundaries including nerves and tissues. Thus, quantifying the diffusion process is important to understand the WM microstructure. However, the development of efficient analytical methods for the reconstruction, lifespan structural connectome analysis, and surrogate variable analysis have fallen seriously behind the technological advances. This challenge motivates us to develop new statistical learning methods for dMRI.

In the first project, we propose a two-stage sparse and adaptive smoothing model (TSASM) for two major image denoising tasks in neuroimaging data analysis, including image reconstruction from a series of noisy images within each subject and group analysis of images obtained from different subjects. Our TSASM consists of an initial smoothing stage of applying a penalized $M-$estimator and a refined smoothing stage of applying kernel-based smoothing methods. The key novelties of our TSASM are that it accounts for the sparse structure of imaging signals, while preserving piecewise smooth regions with unknown edges.

In the second project, we develop a scalable analytical method for mapping the lifespan human structural connectome. Specifically, we develop a novel lifespan population-based structural connectome (LPSC) framework that integrates fiber bundle and functional network information for hierarchically guiding the registration. Our LPSC is applicable to several neuroimaging studies of neuropsychiatric disorders as well as normal brain development. An improved understanding of human structural connectome has the potential to inspire new approaches to prevention, diagnosis, and treatment of many illnesses.

In the third project, we propose an eigen-shrinkage projection (ESP) method to perform the surrogate variable analysis and solve the hidden confounders and harmonization problems in the neuroimaging studies. Our ESP can eliminate the signals from primary variable while preserving the eigenvalue-gap between hidden confounder and noises, which enables hidden confounder estimation from the projected data. We then investigate the statistical properties of the estimated hidden confounders and uncover the natural connection with ridge regression. Numerical experiments are used to illustrate the finite-sample performance.

ACKNOWLEDGEMENTS

First of all, I am extremely grateful to my two advisors Drs. Hongtu Zhu and Joseph G. Ibrahim for their guidance, supervision, and patience during my Ph.D. study. Besides my advisors, I would like to thank the rest of my committee members: Drs Pew-Thian Yap, Martin Styner and Quefeng Li for their helpful comments and feedback.

I would like to thank my collaborators: Drs. Qiang Sun, Zhengwu Zhang, Baiguo An, Shangbang Rao, Xianming Tan, Ye Wu, and Jian Cheng for their help and collaborations in my research.

It has been my pleasure to be a member of UNC Biostatistics and Imaging Genomics Analysis Lab (BIG-S2). I would like to thank all the members of BIG-S2 lab for many enjoyable memories. Special thanks to Drs Bingxin Zhao and Tengfei Li for the valuable discussion and wonderful collaborations. It is my pleasure to learn from these passionate people.

I would like to thank my fellow classmates: Jitong Lou, Jiawei Xu, and Meichen Dong for their help and support during the five years at UNC. I also want to acknowledge my friends: Qing Wang, Long Gu, Ronghao Yang, and Chenyu Shen for their priceless friendship, which help me get through many difficult times.

Lastly, many thanks to my parents, Xuemei Mu and Zhaobin Wang and my whole family for their love and supports in my life.

etoolbox

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Diffusion Magnetic Resonance Imaging (dMRI) is the unique non-invasive method to reveal the white matter (WM) microstructure by probing the diffusion of water molecules (Johansen-Berg and Behrens, 2009; Le Bihan and Iima, 2015; Thiebaut de Schotten et al., 2011; Hagmann et al., 2006; Jones, 2010). Since the diffusion of water molecules is restricted by the biological boundaries(Sykov and Nicholson, 2008), such as cells and tissues, recovering the information of this diffusion process is important to further understand the WM geometry.

Current dMRI acquisition can be divided into two categories: Diffusion Tensor Imaging (DTI)(Basser et al., 1994) and High Angular Resolution Diffusion Imaging (HARDI)(Tuch et al., 1999; Descoteaux, 2008; Alexander, 2005). HARDI has recently been of great interest and many HARDI techniques have been proposed since a better sampling scheme than DTI was considered. For instance, the Orientation Distribution Functions (ODFs) derived from either single-shell HARDI (sHARDI) (Tournier et al., 2007; Aganj et al., 2009; Yan et al., 2018; Tristán-Vega et al., 2009; Descoteaux et al., 2007a; Cheng et al., 2010) or multi-shell HARDI (mHARDI) (Aganj et al., 2010; Jeurissen et al., 2014; Cheng et al., 2014) are applied into fiber tracking tasks since it can capture the direction and angular information of the fiber tracts. In contrast, full three-dimensional (3D) Ensemble Average Propagator (EAP) can achieve more information of the diffusion process and represent it in WM more accurately (Tuch, 2002a). Therefore, in Chapter 3, we develop a two-stage sparse and adaptive smoothing model for reconstruction and group analysis of neuroimaging data, which can be utilized in the estimation of EAP and its features.

Recently, with recent development in imaging technologies, many large-scale biomedical studies, such as the UK Biobank (Miller et al., 2016) and Human Connectome Project (HCP) (Sotiropoulos et al., 2013; Van Essen et al., 2013), have collected/are collecting massive dMRI data with high spatial and temporal resolution as well as other complex information (e.g., genomics). However, analyzing such large-scale neuroimaging data poses two major challenges. The first challenge arises from the large intra-subject heterogeneity of WM fibers and the complex inter-subject spatio-temporal patterns of fibers when mapping the structural connectome across the lifespan at the population level. Particularly, the quality of WM tractographic data varies strongly by age, anatomical tract, tractography algorithm, and acquisition parameters (e.g., directions), among others, so it is critically important to accurately segment WM tractographic data (St-Jean et al., 2019; Guevara et al., 2011a; Buchanan et al., 2014; Wakana et al., 2007; Guevara et al., 2020; Schilling et al., 2019). Existing voxel-based (Smith et al., 2006; Schwarz et al., 2014; Snook et al., 2007) and tract-based (Fornito et al., 2013; Zhu et al., 2011; Yeatman et al., 2012; Cousineau et al., 2017; Jin et al., 2014; Heiervang et al., 2006; Ciccarelli et al., 2003; Wang et al., 2016; Wassermann et al., 2010; Garyfallidis et al., 2017; Olivetti et al., 2017; Sharmin et al., 2016) analysis methods may either lose individual fiber tract specificity or rely on fiber clustering methods and parcellation atlas. Therefore, in Chapter 4, we propose a new lifespan population-based structural connectome method to create a parcellation-based tractographic skeleton atlas for accurately mapping dMRI and structural MRI data acquired from the subjects with multiple scans measured at different times on the atlas across the lifespan. The reliable construction of such atlas allows us to quantify the development of WM within individual subjects across time and understand its variations across groups.

The second challenge is that the imaging data may be acquired in multiple imaging sites. Therefore, technical variability across sites, including heterogeneity in the imaging protocol, variations in the scanning parameters and differences in the scanner manufacturers remain to be the batch effect/hidden confounders that would affect the reliability of dMRI images and their imaging features such as Fractional Anisotropy (FA) and Mean Diffusivity (MD)(Zhu et al., 2009, 2011). The research method to adjust for these hidden confounders is termed as *Surrogate Variable Analysis* (SVA) by Leek and Storey (2007). Although this approach is originally proposed for large-scale multiple testing problem in genetics and omics data, it has increasingly gained attention in neuroimaging data(Fortin et al., 2017, 2016; Guillaume et al., 2018). Therefore, in Chapter 5, we develop a new eigen-shrinkage projection based method to estimate and adjust for the hidden confounders, which has the potential use for the harmonization of multi-site dMRI data.

CHAPTER 2: LITERATURE REVIEW

2.1 Diffusion Magnetic Resonance Imaging(dMRI) Acquisition

For dMRI data, we acquire n Diffusion Weighted Imaging (DWI) data with each image to be a 3D volume containing N voxels for each subject. Each DWI data at each voxel $\mathbf{v} \in \mathcal{V}$ denotes as $(S(\mathbf{q}_i;\mathbf{v}), \mathbf{r}_i, b_i))$, where $\mathbf{q}_i$ is associated with the gradient vector $\mathbf{r}_i = (r_{i,1}, r_{i,2}, r_{i,3})^T$ and the b factor for each i. The diffusion weighted signal attenuation $E(\mathbf{q}_i;\mathbf{v})$ is defined through the normalization of DWI data:

$$E(\mathbf{q}_i;\mathbf{v}) = S(\mathbf{q}_i;\mathbf{v})/S(\mathbf{q}_0;\mathbf{v})$$

where $\mathbf{q}_0$ represents for q_i with $b_i = 0$.

Ensemble Average Propagator (EAP) is defined as

$$P(\mathbf{R}) = \mathscr{F}_{3D}\{E(\mathbf{q})\}(\mathbf{R}) = \int_{\mathbb{R}^3} E(\mathbf{q}) \exp\left(-2\pi i \mathbf{q}^T \mathbf{R}\right) \mathrm{d}\mathbf{q} = \int_{\mathbb{R}^3} E(\mathbf{q}) \cos\left(2\pi \mathbf{q}^T \mathbf{R}\right) \mathrm{d}\mathbf{q} \tag{2.1.1}$$

where $\mathscr{F}_{3D}$ and $\mathscr{F}_{3D}^{-1}$, respectively, denote the Fourier transformation and its inverse transformation. The last equation in (2.1.1) is correct since $E(\mathbf{q}) = E(-\mathbf{q})$.

With (2.1.1), $E(\mathbf{q})$ can be rewritten as

$$E(\mathbf{q}) = \int_{\mathbb{R}^3} P(\mathbf{R}) \exp\left(2\pi i \mathbf{q}^T \mathbf{R}\right) \mathrm{d}\mathbf{R} = \mathscr{F}_{3D}^{-1}\{P(\mathbf{R})\}(\mathbf{q}) \tag{2.1.2}$$

Based on the sampling scheme of gradient vector $\mathbf{r}$ in $\mathbf{q}$-space, dMRI can be generally categorized into Diffusion Tensor Imaging (DTI) and High Angular Resolution Diffusion Imaging (HARDI).

2.1.1 Diffusion Tensor Imaging (DTI)

In diffusion tensor imaging(DTI)(Basser et al., 1994), the signal is written using Stejskal-Tanner equation,

$$E(\mathbf{q}) = \exp\left(-b\mathbf{r}^{\mathrm{T}}\mathrm{D}\mathbf{r}\right). \tag{2.1.3}$$

The diffusion tensor matrix D can be decomposed as:

$$\mathrm{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} = \lambda_1 v_1 v_1^{\mathrm{T}} + \lambda_2 v_2 v + \lambda_3 v_3 v_3^{\mathrm{T}} \tag{2.1.4}$$

With diffusion tensor matrix D, Fractional Anisotropy (FA) and Mean Diffusivity (MD)(Pierpaoli and Basser, 1996) can be defined as

$$\mathbf{FA} = \frac{\sqrt{3}\left\|\mathbf{D} - \frac{1}{3}\,\mathrm{Trace}(\mathbf{D})\mathbf{I}\right\|}{\sqrt{2}\|\mathbf{D}\|} = \sqrt{\frac{3}{2}}\sqrt{\frac{\left(\lambda_1 - \bar{\lambda}\right)^2 + \left(\lambda_2 - \bar{\lambda}\right)^2 + \left(\lambda_3 - \bar{\lambda}\right)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.1.5}$$

$$\mathbf{MD} = \bar{\lambda} \tag{2.1.6}$$

where $\bar{\lambda} = \left(\lambda_1 + \lambda_2 + \lambda_3\right)/3$. Linear, planar, and spherical measures are introduced as

$$\mathbf{LA} = \left(\lambda_1 - \lambda_2\right)/(3\bar{\lambda}), \quad \mathbf{PA} = 2\left(\lambda_2 - \lambda_3\right)/(3\bar{\lambda}), \quad \mathbf{SA} = \lambda_3/\bar{\lambda} \tag{2.1.7}$$

2.1.2 High Angular Resolution Diffusion Imaging (HARDI)

High Angular Resolution Diffusion Imaging (HARDI)(Tuch et al., 1999; Tuch, 2002b) was proposed to use a better angular resolution sampling scheme than DTI. Many research works have been proposed to estimate Orientation Distribution Functions (ODF) or EAPs for HARDI.

Diffusion Spectrum Imaging (DSI) Diffusion Spectrum Imaging (DSI) estimate $p(\mathbf{R})$ from (2.1.1) using Fourier transform. EAP is visualized through the EAP profile(Wedeen et al., 2005), which is defined as $p\left(R_0\mathbf{r}\right) = p(R\mathbf{r})|_{R=R_0}$ for a given radius R_0. Usually $R_0 = 15\mu m$ is used in EAP profile to detect the fiber directions since large R may introduce more error. The Orientation Distribution Function (ODF(Wedeen et al., 2005)) is defined through the marginal distribution of $p(\mathbf{R})$

$$\Phi_w(\mathbf{r}) = \int_0^\infty p(\mathbf{R})R^2 dR. \tag{2.1.8}$$

Q-Ball Imaging (QBI) Different from DSI, Q-Ball Imaging (QBI)(Tuch, 2002b) does not require the b values with large range and it is one of the most commonly used HARDI method.

The ODF in QBI is defined as:

$$\Phi_t(\mathrm{r}) = \frac{1}{Z}\int_0^\infty p(R\mathbf{r})dR, \tag{2.1.9}$$

where Z is the normalization factor to make it as a probability density function. Originally, $\Phi_t(\mathrm{r})$ was estimated from samples of $E(\mathbf{q})$ in single shell data based on Funk-Radon Transform (FRT) numerically. For single shell data with $b = 4\pi^2\tau q_0^2$, the FRT of $E(\mathbf{q})$ in direction $\mathbf{r}$ is defined as

$$\mathrm{FRT}\left\{E\left(q_0\mathbf{u}\right)\right\}(\mathbf{r}) = \int_{\Pi_\mathrm{r}} E(q\mathbf{u})\delta\left(q-q_0\right)q\mathrm{d}q\mathrm{d}\mathbf{u} = q_0\int_{\mathbf{u}\in\mathbb{S}^2} E\left(q_0\mathbf{u}\right)\delta\left(\mathbf{u}^T\mathbf{r}\right)\mathrm{d}\mathbf{u} \tag{2.1.10}$$

Spherical Polar Fourier Imaging (SPFI) Spherical Polar Fourier Imaging (SPFI)(Assemlal et al., 2008, 2009) is a model-free method for mHARDI data(Cheng et al., 2010a). It represents the diffusion signal $E(\mathbf{q})$ by Spherical Polar Fourier (SPF) basis, which has both spherical and radial parts. The SPFI represents the image signal by

$$E(\mathbf{q}) = \sum_{n=0}^{N} \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{n,l,m} B_{n,l,m}(\mathbf{q}), \tag{2.1.11}$$

where $B_{n,l,m}(\mathbf{q})$ is the SPF basis and $\{a_{n,l,m}\}$ is the corresponding coefficient. EAP profile $p\left(R_0\mathbf{r}\right)$ for a given R_0 can further be represented by SH basis

$$\begin{aligned} p\left(R_0\mathbf{r}\right) &= \sum_{l=0}^{L} \sum_{m=-l}^{l} \left\{ 4(-1)^{l/2} \frac{\zeta^{0.5l+1.5}\pi^{l+1.5}R_0^l}{\Gamma(l+1.5)} \sum_{k=0}^{K} f_{k,l,m}\left(\zeta, R_0\right) a_{k,l,m} \right\} Y_l^m(\mathbf{u}) \\ &= \sum_{l=0}^{L} \sum_{m=-l}^{l} c_l, m Y_l^m(\mathbf{u}). \end{aligned} \tag{2.1.12}$$

where

$$\begin{aligned} & f_{n,l,m}\left(\zeta, R_0\right) \\ &= \kappa_n(\zeta) \sum_{i=0}^{n} (-1)^i \begin{pmatrix} n+0.5 \\ n-i \end{pmatrix} \frac{1}{i!} 2^{0.5l+i-0.5} \Gamma\left(\frac{2i+l+3}{2}\right) F\left(\frac{2i+l+3}{2}; l+\frac{3}{2}; -2\pi^2 R_0^2 \zeta\right) \end{aligned} \tag{2.1.13}$$

$$\kappa_n(\zeta) = \left[\frac{2}{\zeta^{3/2}} \frac{n!}{\Gamma(n+3/2)} \right] \tag{2.1.14}$$

$$F(a;b;x) = \sum_{k=0}^{\infty} \frac{(a)_k x^k}{(b)_k k!}, \quad (a)_k = (a(a+1)\ldots(a+k-1)), \text{ with } (a)_0 = 1. \tag{2.1.15}$$

2.2 Reconstruction and Estimation Methods

2.2.1 Voxel-wise Estimation Methods

It is assumed that the dMRI signal at voxel $\mathbf{v}$ follows the following model:

$$f\left(E\left(\mathbf{q}_i; \mathbf{v}\right)\right) = \mathbf{x}_i^T \beta(\mathbf{v}) + \epsilon_i(\mathbf{v}) \tag{2.2.1}$$

where $f(\cdot)$ is a transformation function, $\mathbf{x}_i$ is a $p \times 1$ covariate vector associated with $\mathbf{q}_i, \beta(\mathbf{v})$ is a $p \times 1$ coefficient vector, and $\epsilon_i(\mathbf{v})$ is the measurement error following the distribution with mean zero and variance $\sigma_i^2(\mathbf{v})$. We exploit the penalized estimator for $\beta(\mathbf{v})$ as

$$\widehat{\beta}(\mathbf{v}) = \arg\min_{\beta(\mathbf{v})} \rho(\mathbf{y}(\mathbf{v}) - \mathbf{X}\beta(\mathbf{v}); \lambda(\mathbf{v})) \tag{2.2.2}$$

where $\mathbf{y}(\mathbf{v}) = \left(f\left(E\left(\mathbf{q}_1; \mathbf{v}\right)\right), \cdots, f\left(E\left(\mathbf{q}_n; \mathbf{v}\right)\right)\right)^T$, $\mathbf{X}$ is the covariate matrix with $\mathbf{x}_i$ being the i th row and $\rho(\cdot; \lambda(\mathbf{v}))$ is a penalty function with tuning parameter $\lambda(\mathbf{v})$. In the following sections, we focus on Regularized Least-Squares Regression(RLSR) and Robust Regression.

Regularized Least-Squares Regression(RLSR) Regularized Least-Squares Regression(RLSR) reformulates (2.2.2) as

$$\widehat{\beta}(\mathbf{v}) = \arg\min_{\beta(\mathbf{v})} \|\mathbf{y}(\mathbf{v}) - \mathbf{X}\beta(\mathbf{v})\|^2 + \rho_1(\beta(\mathbf{v}); \lambda(\mathbf{v})) \tag{2.2.3}$$

Different methods differs from each other in terms of $\rho_1(\beta(\mathbf{v}); \lambda(\mathbf{v}))$. L_2 regularization assumes

$$\rho_1(\beta(\mathbf{v}); \lambda(\mathbf{v})) = \lambda(\mathbf{v})\beta(\mathbf{v})^T \Lambda \beta(\mathbf{v}) \tag{2.2.4}$$

where $\Lambda = \operatorname{diag}\left(0, 4, 4, 4, \cdots, L^2(L+1)^2, \cdots, L^2(L+1)^2\right)$. With (2.2.4) as penalty function, (2.2.3) can be solved with closed-form $\widehat{\beta}_{LB}(\mathbf{v}) = \left(\mathbf{X}^T\mathbf{X} + \lambda(\mathbf{v})\Lambda\right)^{-1} \mathbf{X}^T\mathbf{y}(\mathbf{v})$.

Additionally, we may consider L_1 regularization. For example, the generalized LASSO(Tibshirani, 1996; Tibshirani and Taylor, 2011) utilizes the penalty function

$$\rho_1(\beta(\mathbf{v}); \lambda(\mathbf{v})) = \lambda(\mathbf{v})\|D\beta(\mathbf{v})\|_1, \tag{2.2.5}$$

where D is the $p \times p$ weight matrix. (2.2.5) is not differentiable at zero thus the solution does not have a close-form in general. Fortunately, (2.2.3) is a convex problem. Several algorithms have been proposed to solve it, such as Alternating Direction Method of Multipliers(ADMM(Boyd et al., 2011)), Least Angle Regression(LARS)(Efron et al., 2004) and Coordinate Descent(Wu and Lange, 2008).

Robust Regression To improve the robustness of the estimator, robust regression uses the huber loss function (Huber, 1964)

$$\rho\left(e_i\right)=\begin{cases} e_i^2 & \text{if } \left|e_i\right|\leq c \\ c\left(2\left|e_i\right|-c\right) & \text{if } \left|e_i\right|>c \end{cases} \tag{2.2.6}$$

where $e_i=y_i-\mathbf{x}_i^T\beta$ is the residual for the ith observation, c is the tuning parameter to adjust the robustness level. Robust regression estimators can be obtained by solving

$$\sum_i \rho'\left(e_i\right)\mathbf{x}_i=0. \tag{2.2.7}$$

Specifically, (2.2.7) is rewritten as $\sum_i \omega_i e_i \mathrm{x}_i=0$ with $\omega_i=\rho'\left(e_i\right)/e_i$. It solves the weighted least squares problem, updates the weights and resolves the problem until convergence.

All these existing voxel-wise estimation methods perform reconstruction independently at each voxel, which essentially ignores the spatial information across all the voxel of DWI signals.

2.2.2 Spatial Regularization Estimation Methods

Many spatial-adaptive reconstruction methods have been proposed to incorporate spatial constraints of dMRI data. The assumption behind these approaches is that EAP/ODF changes smoothly in the same fiber crossing region while there may have sudden jumps at the boundaries of different fiber crossing regions. Incorporating these constraints can improve the reconstruction performance. Current spatial-adaptive methods can be generally categorized into three types, the denoising of raw DWI data, the denoising of the estimated EAP or ODF, and simultaneous smoothing and estimation of DWI data(Cheng and Zhu, 2016).

In the first type, penalization methods and nonparametric methods are commonly used to denoise the raw DWI data $\left\{\left(S\left(\mathbf{q}_i;\mathbf{v}\right):\mathbf{v}\in\mathcal{V}\right)_{i\geq 1}\right\}$ and obtain the denoised data $\left\{\left(S_*\left(\mathbf{q}_i;\mathbf{v}\right):\mathbf{v}\in\mathcal{V}\right)_{i\geq 1}\right\}$ by imposing the spatial information. Penalization methods solve

$S_*\left(\mathbf{q}_i;\mathbf{v}\right)$ by

$$\operatorname{argmin}_{S_*\left(\mathbf{q}_i;\mathbf{v}\right)} \int_{\mathbf{v}\in\mathcal{V}} \left\{\ell\left(S\left(\mathbf{q}_i;\mathbf{v}\right), S_*\left(\mathbf{q}_i;\mathbf{v}\right)\right) + \rho\left(S_*\left(\mathbf{q}_i;\mathbf{v}\right), \lambda(\mathbf{v})\right)\right\} dL(\mathbf{v}) \tag{2.2.8}$$

where $\ell\left(S\left(\mathbf{q}_i;\mathbf{v}\right), S_*\left(\mathbf{q}_i;\mathbf{v}\right)\right)$ is the log-likelihood function of the DWI signal at voxel $\mathbf{v}$, $\rho(\cdot,\cdot)$ is the penalty function, and $L(\mathbf{v})$ is a measure function defined on the sets of $\mathcal{V}$.

Nonparametric statistical methods such as non-local means (NLM), propagation-seperation methods, anisotropic Wiener filtering, the bilateral filter, and the Sigma filter(Arias-Castro et al., 2012; Yaroslavsky, 1985; Katkovnik et al., 2010; Polzehl and Spokoiny, 2000a) aim to denoise the DWI signals through assigning weights to the nearby voxel based on the spatial similarity and taking the average. This strategy can preserve the piecewise smooth feature of imaging data. For example, NLM can be defined as

$$\mathrm{NLM}(S(\mathbf{q},\mathbf{v})) = \sqrt{\sum_{(\mathbf{q}',\mathbf{v}')\in V_{(\mathbf{q},\mathbf{v})}} w\left((\mathbf{q},\mathbf{v}),(\mathbf{q}',\mathbf{v}')\right) S\left(\mathbf{q}',\mathbf{v}'\right)^2 - 2\sigma^2} \tag{2.2.9}$$

where $w\left((\mathbf{q},\mathbf{v}),(\mathbf{q}',\mathbf{v}')\right)$ is the weight defined through the similiarity between the patches centered in $(\mathbf{q},\mathbf{v})$ and $(\mathbf{q}',\mathbf{v}')$ and σ is the variance of noise, which can be estimated from all the DWI signals.

In the second type, most methods perform denoising on the estimation of ODF or EAP. Same with the first category, penalization methods and nonparametric methods are commonly used technique. Penalization methods formulates the problem as:

$$\operatorname{argmin}_{\{\mathrm{D}(\mathrm{v}):\mathrm{v}\in\mathcal{V}\}} \int_{\mathrm{v}\in\mathcal{V}} \left. d(\widehat{\mathrm{D}}(\mathrm{v}), \mathrm{D}(\mathrm{v})) + \rho(\mathrm{D}(\mathrm{v}), \lambda(\mathrm{v}))\right\} dL(\mathrm{v}) \tag{2.2.10}$$

where $d(\widehat{\mathbf{D}}(\mathbf{v}), \mathbf{D}(\mathbf{v}))$ is the distance between $\widehat{\mathbf{D}}(\mathbf{v})$ and $\mathbf{D}(\mathbf{v})$ and $\rho(\cdot,\cdot)$ is the penalty function. Nonparametric methods starts with a set of estimated diffusion tensors

$$\left(\mathbf{v}_1, \widehat{\mathbf{D}}\left(\mathbf{v}_1\right)\right), \ldots, \left(\mathbf{v}_m, \widehat{\mathbf{D}}\left(\mathbf{v}_m\right)\right)$$

at given location $\{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ and formulates the problem as

$$\widetilde{\mathbf{D}}^{\alpha}(\mathbf{v}) = \arg\min_{\mathbf{D}(\mathbf{v})} \sum_{i=1}^{m} w\left(\mathbf{v}_m, \mathbf{v}\right) g\left(\widehat{\mathbf{D}}_m\left(\mathbf{v}_m\right), \mathbf{D}(\mathbf{v})\right)^{\alpha} \tag{2.2.11}$$

where $\alpha \geq 1, w\left(\mathbf{v}, \mathbf{v}'\right)$ is the distance between $\mathbf{v}$ and $\mathbf{v}'$, $g\left(\mathbf{D}(\mathbf{v}), \mathbf{D}\left(\mathbf{v}'\right)\right)$ is the geodesic distance between $\mathbf{D}(\mathbf{v})$ and $\mathbf{D}\left(\mathbf{v}'\right)$ and (2.2.11) can be solved through local polynomial regression estimate (Yuan et al., 2012).

In the third type, a multiscale adaptive regression modelling (MARM) framework can be used to simultaneously smooth and estimate DTI/ODF/EAP(Tabelow et al., 2008a; Li et al., 2011). Specifically, let $B(\mathbf{v}, h)$ be a sphere with radius h centered at voxel $\mathbf{v}$ and $\omega\left(\mathbf{v}, \mathbf{v}'; h\right)$ be a weight function of triple $\left(\mathbf{v}, \mathbf{v}', h\right)$ such that

$$\sum_{\mathbf{v}' \in B(\mathbf{v}, h)} \omega\left(\mathbf{v}, \mathbf{v}'; h\right) = 1 \text{ and } \omega\left(\mathbf{v}, \mathbf{v}'; h\right) \geq 0 \text{ for all } h \geq 0 \tag{2.2.12}$$

MARM starts with building a sequence of nested spheres increasing radii $h_0 = 0 < h_1 < \cdots < h_S = r$ at each voxel v. Then it iteratively solves the following optimization problem as

$$\widehat{\mathbf{D}}\left(\mathbf{v}; h_s\right) = \operatorname{argmax}_{\mathbf{D}(\mathbf{v})} \sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} \omega\left(\mathbf{v}, \mathbf{v}'; h_s\right) \ell_n\left(\left\{S\left(\mathbf{q}_i; \mathbf{v}'\right)\right\}_{i \geq 1}; \mathbf{D}(\mathbf{v})\right). \tag{2.2.13}$$

At the scale $h_0 = 0$, $\widehat{\mathbf{D}}\left(\mathbf{v}; h_0\right) = \widehat{\mathbf{D}}(\mathbf{v})$ without imposing any spatial information. From $h_0 = 0$ to $h_S = r$, a path diagram of MARM is given below:

$$\begin{tikzcd}
w(\mathbf{v}, \mathbf{v}'; h_0) \arrow[d, Rightarrow] & w(\mathbf{v}, \mathbf{v}'; h_1) \arrow[d, Rightarrow] & \cdots & w(\mathbf{v}, \mathbf{v}'; h_S = r) \arrow[d, Rightarrow] \\
\widehat{\mathbf{D}}(\mathbf{v}; h_0) \arrow[ur] & \widehat{\mathbf{D}}(\mathbf{v}; h_1) \arrow[ur] & \cdots & \widehat{\mathbf{D}}(\mathbf{v}; h_S)
\end{tikzcd}$$

2.3 Tractography Algorithms

Many tractography algorithms(Reisert et al., 2011; Mori and van Zijl, 2002; Lazar, 2010; Jbabdi and Behrens, 2013; Fillard et al., 2011) have been developed to construct fiber streamlines through the whole brain based on ODF/EAP. These algorithms can be classified into two categories: local and global methods.

2.3.1 Local Streamline Methods

Local methods use local ODF/EAP information to map fibers and can be further classified into two groups: deterministic and probabilistic algorithm. Deterministic algorithms usually start at seed voxels and at each step it extract a most similar peak from the EAP/ODF of the current voxel. Probabilistic algorithms decide the direction at each voxel by generating the principal directions with Monte Carlo simulations. Mathematically, local tractography methods formulates the fiber tracking problem as a stochastic ordinary differential equation:

$$\frac{d\mathbf{v}(t)}{dt} = \mathbf{e}(\mathbf{v}), \quad t \geq 0 \quad \text{with} \quad \mathbf{v}(0) = \mathbf{v}_0 \tag{2.3.1}$$

where $\mathbf{v}(t)$ is the fiber curves in R^3, $\mathbf{e}(\mathbf{v})$ is the fiber direction at location $\mathbf{v}$ and $\mathbf{v}_0$ is the position of the seed point. Local methods are computational efficiency but are easy to accumulate errors, which may affect the fiber tracking accuracy severely.

2.3.2 Global Streamline Methods

Global methods aim to reconstruct the all the detectable fibers simultaneously. They are more robust to noise components and outliers compared with local methods but are usually computationally intensive. Denote $\mathcal{M}$ as a fiber model in $\mathcal{V}$ and $\mathcal{S}$ the set of all the DWI signals. A Bayesian framework for global methods can be considered. Basically, the likelihood function $p(\mathcal{S}|\mathcal{M})$ and the prior distribution $p(\mathcal{M})$ need to be specified. The likelihood function can be defined through the dMRI models discussed in the previous sections but how to specify $p(\mathcal{M})$ is of worthy attention. For example, Reisert et al. (2011) proposed to use small streamline segments to represent streamline and $p(\mathcal{M})$ was achieved by applying interaction model for all the connected segments. With the prior distribution, the posterior distribution $p(\mathcal{M}|\mathcal{S}) \propto p(\mathcal{S}|\mathcal{M})p(\mathcal{M})$ and the corresponding mode $\widehat{\mathcal{M}} = \text{argmax}_{\mathcal{M}}\, p(\mathcal{M}|\mathcal{S})$ can be obtained.

2.4 Fiber Analysis

Once the streamlines are reconstructed through tractography algorithms, anatomically meaningful fiber bundles/tracks can be recovered by clustering the streamlines. Fiber bundle clustering is essential for understanding and analyzing the white matter structure(Guevara et al., 2012; Siless et al., 2018; Donnell and Westin, 2007). The clustering approaches such as k-means(Li et al., 2010), hierarchical clustering(Donnell and Westin, 2007) and Kernel dictionary learning(Kumar et al., 2019a) have been applied into fiber bundles clustering successfully. Some of them will be reviewed in section 2.4.1. Recently, some deep learning based clustering methods including deep adaptive image clustering(Chang et al., 2017), deep self-evolution clustering(Chang et al., 2018) have been proposed to overcome the shortcomes of traditional methods. For example, the performance of traditional methods is sensitive to the choice of distance metric between two objects. Therefore, these deep learning based clustering methods have potential application on fiber bundle clustering. Some of them will be reviewed in section 2.4.2.

2.4.1 Fiber Clustering Methods

Notations and concepts are declared as follows. Denote $\mathcal{S}$ the set of streamlines and $\mathbf{X} \in \mathcal{S}^n$ is the streamlines generated from tractography. $\mathbf{D} \in \mathcal{S}^m$ is the dictionary of bundle prototype/cluster center and $\mathbf{W} \in \mathbb{R}^{m \times n}$ assigns the streamline to bundle and $\phi : \mathcal{S} \to \mathbb{R}^q$ is a mapping function. Define Φ as mapped streamlines and $\mathbf{K} \in \mathbb{R}^{n \times n}$ to be the kernel matrix for streamlines.

Dictionary Learning and the k-means Let $\mathbf{X}$ be a $d \times n$ matrix such that each column is the feature vector of each streamline. The fiber clustering problem is to assign each streamline to one of the m bundles and can be formulated as

$$\min_{\mathbf{W},\mathbf{D}} \|\mathbf{X}-\mathbf{D}\mathbf{W}\|_F^2.$$

Although solving $\mathbf{D}$ and $\mathbf{W}$ together is NP-hard, optimizing $\mathbf{W}$ or $\mathbf{D}$ iteratively is feasible. Specifically, with fixed dictionary $\mathbf{D}$, the optimal $\mathbf{W}$ assigns each streamline to the prototype with the nearest distance. With fixed $\mathbf{W}$, the optimal dictionary $\mathbf{D}$ is achieved by linear regression estimator. This idea is shared with the well-known k-means algorithm.

Kernel k-means In the kernel k-means, each streamline is projected to a q-dimensional space by $\phi : \mathbb{R}^d \to \mathbb{R}^q$, where $q \gg d$. Denote $\Phi \in \mathbb{R}^{q\times n}$ as $\Phi = \phi(\mathbf{X})$ and the kernel matrix $\mathbf{K} = \mathbf{\Phi}^\top\mathbf{\Phi}$. The clustering problem can be formulated as:

$$\arg\min_{\mathbf{D},\mathbf{W}} \|\mathbf{\Phi}-\mathbf{D}\mathbf{W}\|_F^2 \quad \text{subject to} \quad \mathbf{W}^\top\mathbf{1}_m = \mathbf{1}_n \tag{2.4.1}$$

We further define $\mathbf{D} = \mathbf{\Phi}\mathbf{A}$, where $\mathbf{A} \in \mathbb{R}^{n\times m}$. Then (2.4.1) becomes

$$\arg\min_{\mathbf{W},\mathbf{A}} \|\mathbf{\Phi}-\mathbf{\Phi}\mathbf{A}\mathbf{W}\|_F^2 \quad \text{subject to} \quad \mathbf{W}^\top\mathbf{1}_m = \mathbf{1}_n \tag{2.4.2}$$

Following the similar optimization method in k-means, $\mathbf{W}$ is updated by:

$$\mathbf{w}_{mi} = \begin{cases} 1: & \text{if } m = \arg\min_{m'} \left[\mathbf{A}^\top\mathbf{K}\mathbf{A}\right]_{m'm'} - 2\left[\mathbf{A}^\top\mathbf{k}_i\right]_{m'} \\ 0: & \text{otherwise.} \end{cases} \tag{2.4.3}$$

where $\mathbf{k}_i$ is the ith column of $\mathbf{K}$. $\mathbf{A}$ is then updated by the linear regression solution $\mathbf{A} = \mathbf{W}^\top\left(\mathbf{W}\mathbf{W}^\top\right)^{-1}$. $\mathbf{A}$ and $\mathbf{W}$ are optimized alternatively until convergence or the maximum number of iterations is reached.

Kernel Dictionary Learning In fiber clustering problem, hard clustering approaches such as k-means can be sensitive to streamline outliers. Therefore, kernel dictionary learning is proposed to overcome this drawback. Specifically, the L_0 norm constraints are imposed on the column of $\mathbf{W}$ and the i th column $\mathbf{w}_i$ can be solved independently by the following optimization problem:

$$\arg\min_{\mathbf{W}_i\in\mathbb{R}_+^m} \|\phi(\mathbf{x}_i)-\Phi\mathrm{A}\mathbf{w}_i\|_2^2 \quad \text{subject to} \quad \|\mathbf{w}_i\|_0 \leq S_{\max}, \tag{2.4.4}$$

where $S_{\max}$ is defined to control the sparsity level of $\mathbf{w}_i$.

To solve matrix $\mathbf{W}$ and $\mathbf{A}$, each column of $\mathbf{W}$ is first updated by using kernel orthogonal matching pursuit (kOMP)(Nguyen et al., 2012) approach with non-negativity constrains. Then $\mathbf{A}$ is updated by

$$[\mathbf{A}]_{ij} \leftarrow [\mathbf{A}]_{ij} \cdot \frac{\left[\mathbf{K}\mathbf{W}^\top\right]_{ij}}{\left[\mathbf{K}\mathbf{A}\mathbf{W}\mathbf{W}^\top\right]_{ij}}, \quad i=1,\ldots,n, \quad j=1,\ldots,m \tag{2.4.5}$$

$\mathbf{A}$ and $\mathbf{W}$ are optimized alternatively until convergence or the maximum number of iterations is reached.

2.4.2 Deep Adaptive Image Clustering

The Deep Adaptive Clustering(DAC)(Chang et al., 2017) proposes to consider image clustering as a binary pairwise classification task. Specifically, for the training set $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{x}_j, r_{ij})\}_{i=1,j=1}^n$, define $r_{ij}=1$ if $\mathbf{x}_i, \mathbf{x}_j$ belong to the same cluster and $r_{ij}=0$ otherwise. This cluster problem is formulated as

$$\min_{\mathbf{w}} \mathbf{E}(\mathbf{w}) = \sum_{i,j} L\left(r_{ij}, g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)\right) \tag{2.4.6}$$

where $g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)$ is the similarity function between $\mathbf{x}_i$ and $\mathbf{x}_j$ with parameters $\mathbf{w}$ and $L\left(r_{ij}, g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)\right)$ is the loss function between r_{ij} and $g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)$. The loss function can be further expressed as

$$L\left(r_{ij}, g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)\right) = -r_{ij}\log\left(g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)\right) - \left(1-r_{ij}\right)\log\left(1-g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)\right). \tag{2.4.7}$$

To define the similarity function, the set of label features are first defined as $\mathcal{L} = \left\{\mathbf{l}_i \in \mathbb{R}^k\right\}_{i=1}^n$, where $\mathbf{l}_i$ is the label feature vector of the image $\mathbf{x}_i$ and

$$\|\mathbf{l}_i\|_2 = 1, \text{ and } l_{ih} \geq 0, h = 1, \cdots, k \tag{2.4.8}$$

The similarity $g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right)$ is further defined as

$$g\left(\mathbf{x}_i, \mathbf{x}_j; \mathbf{w}\right) = \mathbf{l}_i \cdot \mathbf{l}_j \tag{2.4.9}$$

With the clustering constraint of $\mathbf{l}_i$, the DAC model can be reformulated as:

$$\min_{\mathbf{w}} \mathbf{E}(\mathbf{w}) = \sum_{i,j} L\left(r_{ij}, \mathbf{l}_i \cdot \mathbf{l}_j\right) \text{ s.t. } \forall i, \|\mathbf{l}_i\|_2 = 1, \text{ and } l_{ih} \geq 0, h = 1, \cdots, k \tag{2.4.10}$$

However, since the true labels of images are unknown in clustering, it is unclear how to select labeled samples for training purpose. A strategy to solve the problem is to use ALL-ConvNets(Springenberg et al., 2014) based on

$$r_{ij} := \begin{cases} 1, \text{ if } 1_i \cdot 1_j \geq u(\lambda) \\ 0, \text{ if } 1_i \cdot 1_j < l(\lambda), \\ \text{None, otherwise} \end{cases}, \quad i, j = 1, \cdots, n \tag{2.4.11}$$

where λ is a tuning parameter and None means the sample is not selected for training. To ensure that the samples are gradually selected in the training process, λ is increasing and $u(\lambda) \propto -\lambda, l(\lambda) \propto \lambda$ and $l(\lambda) \leq u(\lambda)$ are also satisfied.

The DAC model can be finalized as

$$\min_{\mathbf{w},\lambda} \mathbf{E}(\mathbf{w},\lambda) = \sum_{i:j} v_{ij} L\left(r_{ij}, \mathbf{1}_i \cdot \mathbf{1}_j\right) + u(\lambda) - l(\lambda)$$

$$\text{s.t. } l(\lambda) \leq u(\lambda), v_{ij} \in \{0,1\}, i,j = 1,\cdots,n, \forall i, \|\mathbf{l}_i\|_2 = 1, \text{ and } l_{ih} \geq 0, h = 1,\cdots,k$$

$$r_{ij} := \begin{cases} 1, \text{ if } 1_i \cdot 1_j \geq u(\lambda) \\ 0, \text{ if } 1_i \cdot 1_j < l(\lambda), \\ \text{None, otherwise} \end{cases}, \quad i,j = 1,\cdots,n, \tag{2.4.12}$$

where $\mathbf{v}$ satisfies

$$v_{ij} := \begin{cases} 1, \text{ if } r_{ij} \in \{0,1\}, \\ 0, \text{ otherwise} \end{cases} \quad i,j = 1,\cdots,n. \tag{2.4.13}$$

To solve (2.4.12), Adaptive Learning algorithm is utilized to optimize $\mathbf{w}$ and λ alternately. With λ fixed, (2.4.12) degenerates to

$$\min_{\mathbf{w}} \mathbf{E}(\mathbf{w}) = \sum_{i,j} v_{ij} L\left(r_{ij}, f\left(\mathbf{x}_i;\mathbf{w}\right) \cdot f\left(\mathbf{x}_j;\mathbf{w}\right)\right) \tag{2.4.14}$$

and the back-propagation algorithm can be used to update $\mathbf{w}$. Similarly, with $\mathbf{w}$ fixed, (2.4.12) becomes:

$$\min_{\lambda} \mathbf{E}(\lambda) = u(\lambda) - l(\lambda). \tag{2.4.15}$$

λ can be updated by gradient descent algorithm:

$$\lambda := \lambda - \eta \cdot \frac{\partial \mathbf{E}(\lambda)}{\partial \lambda} \tag{2.4.16}$$

CHAPTER 3: A TWO-STAGE SPARSE AND ADAPTIVE SMOOTHING MODEL FOR NEUROIMAGING DATA ANALYSIS

3.1 Introduction

Many large-scale neuroimaging studies, such as the lifespan human connectome studies and the UK biobank, have collected/are collecting massive multi-modal imaging data, including structure magnetic resonance imaging (MRI) data (Brown et al., 2014), diffusion MRI (dMRI) (Johansen-Berg and Behrens, 2009; Le Bihan and Iima, 2015), and functional MRI (fMRI) data (Lazar, 2008b; Penny et al., 2011), with high spatial and/or temporal resolution to map human brain function and structure and assess their variability across time and groups. An improved understanding of human brain function and structure through neuroimaging data analysis has the potential to inspire new and urgently needed approaches to prevention, diagnosis, and treatment of many illnesses (e.g., schizophrenia, and Alzheimer). Two major neuroimage denoising tasks include image reconstruction from multiple noisy image volumes within each subject and group analysis of normalized images obtained from different subjects. These two tasks are highly related with the classical problem of image noise removal.

The problem of image noise removal has received extensive attention for decades (Fan et al., 2019; Buades et al., 2005; Arias-Castro et al., 2012). Among all existing methods, kernel-based methods are extremely popular due to their computational simplicity and theoretical elegance (Nadaraya, 1964; Watson, 1964; Yaroslavsky, 1985; Lee, 1983; Buades et al., 2005; Awate and Whitaker, 2006; Dabov et al., 2007; Dong et al., 2012; Coupé et al., 2008; Arias-Castro et al., 2012). Some well-known examples include linear filtering (Nadaraya, 1964; Watson, 1964), Yaroslavsky's filter (Yaroslavsky, 1985; Lee, 1983), non-local means (Buades et al., 2005), propagation−separation (PS) approach Polzehl and Spokoiny (2000b); Polzehl et al. (2010), unsupervised, information-theoretic and adaptive filter (UINTA)) (Awate and Whitaker, 2006) and many variants of these approaches (Dabov et al., 2007; Dong et al., 2012; Coupé et al., 2008). Those smoothing methods usually solve a weighted loss function by incorporating signals in the neighboring locations of each location. A further refinement consists of building a sequence of increasing scales and then sequentially fitting the weighted loss functions from the small scale to the large scale (Polzehl and Spokoiny, 2000b; Polzehl et al., 2010). However, these kernel-based denoising methods focus on smoothing a single image at a time, so they are not optimal for the two major image denoising tasks discussed above due to additional imaging features for neuroimaging data.

There are at least three key imaging features associated with many neuroimaging modalities, including local low-dimensional representation, low signal-to-noise ratio, and spatial smoothness. The first one is that multiple imaging signals in each spatial location are usually represented as either a parametric model with a few parameters or a sparse representation of a set of basis functions (Zhu et al., 2009; Zhang et al., 2016, 2013; Lindquist et al., 2010; Tabelow et al., 2008b). For instance, in dMRI, spherical polar Fourier imaging (Cheng et al., 2010b; Assemlal et al., 2009) was proposed to approximate diffusion signal attenuation by using the spherical polar Fourier basis (SPF) basis. In fMRI, various basis functions used for modeling a hemodynamic response function include the finite impulse response (FIR) basis set, the principal component basis set, and spline basis set, among others. The second one is that the signal-to-noise ratio of neuroimaging signals in each image volume can be very low partially due to improving acquisition efficiency in many biomedical studies. Therefore, it is very rare to independently apply image denoising methods to each image volume, possibly leading to many artifacts. The third one is that neuroimaging data is usually expected to contain spatially contiguous regions or effect regions with relatively sharp edges due to the physical and biological reasons. For instance, scientists have been trying to subdivide the human brain into anatomically and functionally distinct, spatially contiguous areas (cortical areas and subcortical nuclei), as a prerequisite for understanding how the brain works (Bijsterbosch et al., 2020; Glasser et al., 2016). Accurate parcellation enables efficient comparison of results across studies and communication among investigators and a foundation for illuminating the functional and structural organization of the brain.

Two major image denoising methods include voxel-wise methods and adaptive smoothing methods. Conventional voxel-wise approaches for image reconstruction independently fit a statistical model (e.g., Rician regression or high-dimensional linear regression) to imaging data at each location (Zhu et al., 2009; Zhang et al., 2016, 2013; Lindquist et al., 2010; Tabelow et al., 2008b). Conventional voxel-wise approaches for group analysis involve in Gaussian smoothing imaging data, independently fitting a statistical model to imaging data at each voxel, and generating statistical maps of test statistics and $p-$values (Lazar, 2008a). It is well-known that many voxel-wise methods ignores the spatial information of imaging data and therefore, they are generally not optimal in power (Zhu et al., 2014; Lindquist et al., 2010). Therefore, there is a great interest in developing multiscale adaptive methods to adaptively and simultaneously smooth multiple neuroimaging images for the two neuroimaging denoising tasks (Tabelow et al., 2008a,b; Becker et al., 2014; Polzehl et al., 2010; Zhu et al., 2014; Li et al., 2011). Those multiscale adaptive methods dramatically increase signal-to-noise ratio, while preserving spatial details (e.g., spatial smoothness and edges). Moreover, a multiscale adaptive regression model (Li et al., 2011) was developed for a large class of parametric models by integrating the propagation-separation approach and voxel-wise approach.

The aim of this paper is to develop a two-stage sparse and adaptive smoothing model (TSASM) to solve the two neuroimage denoising tasks in neuroimaging data analysis, while accounting for the key neuroimaging features as discussed above. We focus on the case when multiple imaging signals in each spatial location can be represented a sparse representation of a set of basis functions, in which the number of basis functions can be much larger than the number of observations in each location. Our TSASM involves an initial smoothing stage based on a penalized $M-$estimator and a refined smoothing stage by applying Yaroslavsky's filter and non-local means. Major contributions of the article are as follows.

- In the first stage of TSASM, we fit a penalized linear regression model to select the active set for each voxel/vertex. We study the variable selection consistency for the penalized M-estimator. Since we consider the general loss function and nonconvex penalty function, these results work for several variable selection methods with regularization strategy including SCAD (Fan and Li, 2001), MCP (Zhang, 2010), and many others.

- In the second stage, we apply the adaptive kernel including Yaroslavsky's filter (Yaroslavsky, 1985) and Non-local means method (Buades et al., 2005; Arias-Castro et al., 2012) on the coefficient images estimated from the first stage to adjust for the piecewise-smoothness with discontinuity jump edges feature of neuroimaging data.

- Our two-stage estimation procedure smooth each of coefficient images independently, which is computationally more efficient compared with the existing methods.

- We systematically study the oracle inequalities and derive the theoretical minimax rates based on mean squared error under different noise levels. We also give the optimal choice of bandwidth for the kernel function.

The rest of this paper proceeds as follows. Section 3.2 presents the two-stage imaging data noise removal process and establish the theoretical properties. We examine the finite-sample performance of our proposed methods by simulation studies in Section 3.4. In Section 3.5, we apply our proposed method on a diffusion magnetic resonance imaging dataset. Section 3.6 concludes the article with some discussions. Theoretical assumptions and proofs are given in the Appendix.

We summarize here the notation that will be used throughout the paper. For any vector $\boldsymbol{u} = (u_1, \ldots, u_d)^{\mathrm{T}} \in \mathcal{R}^d$ and $q \geq 1$, $\|\boldsymbol{u}\|_q = \left(\sum_{j=1}^{d} |u_j|^q\right)^{1/q}$ is the ℓ_q norm and $\|\boldsymbol{u}\|_\infty = \max_{1\leq j\leq d} |u_i|$ is the ℓ_∞ norm. For any vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{R}^d$, we write $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$. We use C to denote a generic constant which may change from line to line. For two sequences of real numbers $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$, we write $a_n = O(b_n)$ or $a_n \lesssim b_n$ if $|a_n| \leq C|b_n|$ for some constant $C > 0$ and $a_n \asymp b_n$ if $a_n = O(b_n)$ and $b_n = O(a_n)$. If $\boldsymbol{A}$ is an $m \times n$ matrix, we use $\|\boldsymbol{A}\|_q$ to denote its ℓ_q operator norm, defined by $\|\boldsymbol{A}\|_q = \max_{\mathrm{u} \in \mathcal{R}^n} \|\boldsymbol{A}\mathrm{u}\|_q / \|\mathrm{u}\|_q$. For an $n \times n$ matrix $\boldsymbol{A}$, we use $\lambda_{\max}(\boldsymbol{A})$ and $\lambda_{\min}(\boldsymbol{A})$ to denote the maximum and minimum eigenvalues of $\boldsymbol{A}$. For a function $g : \mathcal{R}^d \to \mathcal{R}$, we use $\nabla g \in \mathcal{R}^d$ to denote its gradient vector.

3.2 Methodology

3.2.1 Model Setup

Suppose that we observe n image volumes (or subjects) defined on a common space $\mathcal{D}_0$ in a compact set of $\mathcal{R}^3$, where $\mathcal{D}_0$ is the set of all voxels $\mathbf{d}_0$. We use N_D to denote the total number of all voxels in $\mathcal{D}_0$. For the i-th volume, we observe an imaging measure $y_i(\mathbf{d}_0)$ at $\mathbf{d}_0 \in \mathcal{D}_0$, and thus we would have a $N_D \times 1$ vector of measurements across $\mathcal{D}_0$. Besides the imaging measurements, suppose we collect a set of predictors, such as age, gender and other clinical variables, and/or basis functions, denoted as a vector $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^{\mathrm{T}} \in \mathcal{R}^p$ for the i th volume, where p is the number of the predictors.

First, it is assumed that the imaging measures at $\mathbf{d}_0$ follow the following model:

$$y_i(\mathbf{d}_0) = f(\boldsymbol{x}_i, \mathbf{d}_0) + \epsilon_i(\mathbf{d}_0) \quad \text{for all} \;\; i = 1, \ldots, n \;\; \text{and} \;\; \mathbf{d}_0 \in \mathcal{D}_0, \tag{3.2.1}$$

where $f(\boldsymbol{x}_i, \mathbf{d}_0) : \mathcal{R}^p \times \mathcal{D} \to \mathcal{R}$ and $\epsilon_i(\mathbf{d}_0)$ is a measurement error. Throughout the paper, it is assumed that $f(\boldsymbol{x}_i, \mathbf{d}_0) = \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}(\mathbf{d}_0)$, where $\boldsymbol{\beta}(\mathbf{d}_0) = (\beta_1(\mathbf{d}_0), \ldots, \beta_p(\mathbf{d}_0))^{\mathrm{T}}$ is a p-dimensional vector of coefficient functions of $\mathbf{d}_0$. For simplicity, $f(\boldsymbol{x}_i, \mathbf{d}_0)$ is denoted by $f_i(\mathbf{d}_0)$. Throughout the paper, p is allowed to be much larger than the sample size n and tends to ∞ with n, but only a few of them are related to $y_i(\mathbf{d}_0)$ with the corresponding elements of $\boldsymbol{\beta}(\mathbf{d}_0)$ being nonzero, and the rest elements of $\boldsymbol{\beta}(\mathbf{d}_0)$ equal to 0 exactly. Let $\boldsymbol{\beta}^*(\mathbf{d}_0) = (\beta_1^*(\mathbf{d_0}), \ldots, \beta_p^*(\mathbf{d}_0))^{\mathrm{T}}$ be the true value of the parameter vector $\boldsymbol{\beta}(\mathbf{d}_0)$ and the true value of $f_i(\mathbf{d}_0)$ be $f_i^*(\mathbf{d}_0) = \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}^*(\mathbf{d})$. We define the support set for every $\mathbf{d} \in \mathcal{D}$ as

$$\mathcal{S}(\mathbf{d}_0) = \{j : 1 \leq j \leq p, \quad \beta_j^*(\mathbf{d}_0) \neq 0\}.$$

We define $s(\mathbf{d}_0) = |\mathcal{S}(\mathbf{d_0})|$ as the number of elements contained in $\mathcal{S}(\mathbf{d})$ and $s = \max_{\mathbf{d_0}} s(\mathbf{d}_0)$.

Second, it is assumed that both f_i^* and β_j^* belong to the Cartoon function class $\mathcal{F}$. This assumption is quite reasonable since $\{y_i(\mathbf{d}_0) : \mathbf{d}_0 \in \mathcal{D}_0\}$ can be regarded as a noisy version of a piecewise-smooth function of $\mathbf{d} \in \mathcal{D}$ with jumps or edges and may inherit the piecewise-smooth feature from imaging data. In many neuroimaging data, those jumps or edges often reflect anatomically and functionally distinct brain subregions, such as major white matter bundles. We define the Cartoon function class and its related Hölder function class as follows.

Definition 3.2.1 (Hölder Function Class)**.** For any $\alpha \in \mathbb{R}_+$, let $[\alpha]$ denote the largest integer strictly less than α. We define $\mathcal{H}_m(\alpha, C_0)$ as the Hölder class of functions g:$[0,1]^m \to [0,1]$ that are $[\alpha]$ times differentiable such that

$$\forall \mathbf{d} = (d_1, \ldots, d_m)^T \in [0,1]^m, \forall \mathbf{s} = (s_1, \ldots, s_m)^T \in \mathbb{N}^m, 1 \leq |\mathbf{s}| \leq [\alpha] : |\mathbf{g}^{(\mathbf{s})}(\mathbf{d})| \leq \mathbf{C_0},$$

$$\forall \mathbf{d} \text{ and } \mathbf{d}^{'} \in [0,1]^m, |\mathbf{s}| := s_1 + \cdots + s_m = [\alpha] : |g^{(\mathbf{s})}(\mathbf{d}) - g^{(\mathbf{s})}(\mathbf{d}^{'})| \leq C_0 \left\| \mathbf{d} - \mathbf{d}^{'} \right\|^{\alpha - [\alpha]},$$

where α, $C_0 > 0$ are some constants, $g^{(\mathbf{s})}(\mathbf{d})$ is the $\mathbf{s}$-derivative of $g(\cdot)$ at $\mathbf{d} \in \mathbb{R}^m$ and given by

$$g^{(\mathbf{s})}(\mathbf{d}) = \frac{\partial^{|\mathbf{s}|}}{\partial_{d_1}^{s_1} \cdots \partial_{d_m}^{s_m}} g(\mathbf{d}).$$

Definition 3.2.2 (Cartoon Function Class)**.** Let $\Omega = \phi(B(0,1))$, where $B(0,1)$ is the unit ball in $\mathbb{R}^3$ and $\phi : \mathcal{R}^3 \to \mathcal{R}^3$ is injective with both ϕ and ϕ^{-1} being C_0-Lipschitz. Let Ω^c and $\partial\Omega$ be the complement of Ω and the boundary of Ω, respectively. For $C_0 > 0$, we define the Cartoon function class $\mathcal{F}$ as the set of functions of the form

$$f(\mathbf{d}) = \mathbf{1}(\{\mathbf{d} \in \Omega\}) f_\Omega(\mathbf{d}) + \mathbf{1}(\{\mathbf{d} \in \Omega^c\}) f_{\Omega^c}(\mathbf{d}),$$

where $\mathbf{1}(A)$ is the indicator function of an event A and $f_\Omega(\mathbf{d})$ and $f_{\Omega^c}(\mathbf{d})$ belong to Hölder function class $\mathcal{H}_3(1, C_0)$, with jump

$$\mu(f) = \inf_{\mathbf{d}\in\partial\Omega} |f_\Omega(\mathbf{d}) - f_{\Omega^c}(\mathbf{d})| \geq 1/C_0.$$

Our purpose here is to identify the active sets, estimate the model coefficient, and finally get a piecewise-smooth estimator for both $f_{*,i}(\cdot)$ and $\beta^*(\cdot)$.

3.2.2 Estimation Procedure

We develop a two-stage estimation procedure, including the first stage for estimating $\boldsymbol{\beta}^*(\mathbf{d}_0)$ and the second stage for estimating $f_{*,i}(\mathbf{d}_0)$, as follows.

First Stage: Estimating $\boldsymbol{\beta}^*(\mathbf{d}_0)$ We first identify the support set of $\boldsymbol{\beta}^*(\mathbf{d}_0)$ and obtain an initial estimator of $\boldsymbol{\beta}^*(\mathbf{d}_0)$ for each voxel $\mathbf{d}_0 \in \mathcal{D}_0$. Specifically, we exploit the penalized M-estimator

$$\widehat{\beta}(\mathbf{d}_0, \lambda) = \arg\min_{\|\boldsymbol{\beta}\|_1 \leq R} \left\{ \mathcal{L}(\boldsymbol{\beta}; Y(\mathbf{d}_0), X) + \sum_{j=1}^{p} p_\lambda(\beta_j) \right\}, \tag{3.2.2}$$

where $\mathcal{L}(\cdot)$ is a convex loss function, $Y(\mathbf{d}_0) = (y_1(\mathbf{d}_0), \ldots, y_n(\mathbf{d}_0))^{\mathrm{T}}$, $X = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^{\mathrm{T}}$, $p_\lambda(\cdot)$ is a penalty function, and $\lambda > 0$ is a regularization parameter. We require R to be large enough such that $\|\boldsymbol{\beta}^*\|_1 \leq R$. Subsequently, the corresponding support set estimator is given by $\widehat{\mathcal{S}}(\mathbf{d}_0, \lambda) = \{j : 1 \leq j \leq p, \widehat{\beta}_j(\mathbf{d}_0, \lambda) \neq 0\}$. Fo simplicity, we write $\mathcal{L}_{\mathbf{d}_0}(\boldsymbol{\beta}) = \mathcal{L}(\boldsymbol{\beta}; Y(\mathbf{d}_0), X)$.

Second Stage: Estimating $f_{*,i}(\mathbf{d}_0)$ We second refine the estimator of $\boldsymbol{\beta}^*(\mathbf{d}_0)$ by accounting for its piecewise-smooth property with discontinuities along smooth hypersurfaces. Specifically, the initial estimator $\widehat{\boldsymbol{\beta}}(\mathbf{d}_0, \lambda)$ does not utilize this information, so it may be suboptimal. We refine this estimator by using kernel-based methods as follows. We apply the local linear regression to get a piecewise smooth estimator $\widehat{\beta}_j(\cdot)$ for $j = 1, \ldots, p$. Let $B_j(\mathbf{d}) = (\beta_j(\mathbf{d}), h(\partial\beta_j(\mathbf{d})/\partial\mathbf{d})^{\mathrm{T}})^{\mathrm{T}} \in \mathcal{R}^4$ and $\mathbf{z}_h(\mathbf{d_m} - \mathbf{d}) = (1, (d_{m,1} - d_1)/h, (d_{m,2} - d_2)/h, (d_{m,3} - d_3)/h)^{\mathrm{T}}$, where $\mathbf{d} = (d_1, d_2, d_3)^T$ and $\mathbf{d_m} = (d_{m,1}, d_{m,2}, d_{m,3})^T \in \mathcal{D}_0$. To estimate $B_j(\mathbf{d})$, we propose to solve

$$\widehat{B}_j(\mathbf{d}) = \arg\min_{B_j} \sum_{\mathbf{d}_m \in \mathcal{D}_0} \left\{\widehat{\beta}_j(\mathbf{d}_m, \lambda) - B_j^{\mathrm{T}} \mathbf{z}_h(\mathbf{d}_m - \mathbf{d})\right\}^2 K_{j,h}(\mathbf{d}_m - \mathbf{d}), \tag{3.2.3}$$

where $K_{j,h}(\mathbf{d}_m - \mathbf{d})$ is a rescaled kernel function with bandwidth h, which will be discussed in Subsection 3.2.3. Taking the derivative with respect to B_j, we get

$$\begin{aligned} \widehat{B}_j(\mathbf{d}) = & \left\{ \sum_{\mathbf{d}_m \in \mathcal{D}_0} \mathbf{z}_h(\mathbf{d}_m - \mathbf{d}) \mathbf{z}_h(\mathbf{d}_m - \mathbf{d})^{\mathrm{T}} K_{j,h}(\mathbf{d}_m - \mathbf{d}) \right\}^{-1} \\ & \cdot \left\{ \sum_{\mathbf{d}_m \in \mathcal{D}_0} \mathbf{z}_h(\mathbf{d}_m - \mathbf{d}) K_{j,h}(\mathbf{d}_m - \mathbf{d}) \widehat{\beta}_j(\mathbf{d}_m, \lambda) \right\}. \end{aligned} \tag{3.2.4}$$

Moreover, $\widehat{\beta}_j(\mathbf{d}) = \widehat{B}_j(\mathbf{d})^{\mathrm{T}} \mathbf{e}_1$ and $\mathbf{e}_1 = (1, 0, 0, 0)^{\mathrm{T}}$. For every $\mathbf{d}_0'$ and $\mathbf{d} \in \mathcal{D}_0$, we define

$$w_{j,h}(\mathbf{d}_0', \mathbf{d}) = \mathbf{z}_h(\mathbf{d}_0' - \mathbf{d})^{\mathrm{T}} [\sum_{\mathbf{d}_m \in \mathcal{D}_0} \mathbf{z}_h(\mathbf{d}_m - \mathbf{d}) \mathbf{z}_h(\mathbf{d}_m - \mathbf{d})^{\mathrm{T}} K_h(\mathbf{d}_m - \mathbf{d})]^{-1} \mathbf{e}_1 K_{j,h}(\mathbf{d}_{\mathbf{0}}' - \mathbf{d}), \tag{3.2.5}$$

then we have $\widehat{\beta}_j(\mathbf{d}) = \sum_{\mathbf{d}_0' \in \mathcal{D}_0} w_{j,h}(\mathbf{d}_0', \mathbf{d}) \widehat{\beta}_j(\mathbf{d}_0', \lambda)$. The piece-wise smooth estimator for $f_{*,i}(\mathbf{d}_0)$ can be further given by $\widehat{f}_i(\mathbf{d}) = \sum_{j=1}^{p} x_{ij} \widehat{\beta}_j(\mathbf{d})$.

3.2.3 Three different kernel functions

Various kernel-based methods that we discuss in this paper differ only in the choice of kernel $K_h(\cdot)$. We consider three different kernels corresponding to oracle kernel, Yaroslavsky's filter (YF), and non-local means (NLM) method, respectively, as follows. The oracle kernel is based on the boundary information $\partial\Omega$, whereas the other two do not assume that.

Oracle Kernel We consider the oracle kernel by using the exact location of Ω and Ω^c. Therefore, we incorporate such location information into the kernel function and substitute it into (3.2.3) to smooth β_j. For example, to get $\widehat{\beta}_{j,\Omega}(\mathbf{d})$ for $\mathbf{d} \in \Omega$, we set the kernel function in as

$$K_h(\mathbf{d}_m - \mathbf{d}) = h^{-3} \prod_{k=1}^{3} K_{\mathrm{loc}}((d_{m,k} - d_k)/h) \mathbf{1}(\{\mathbf{d} \in \Omega \bigcap \mathcal{D}_0\}). \tag{3.2.6}$$

Here we set $K_{\text{loc}}(u) = (1-u)_+ := \max(0, 1-u)$. Similarly, we can obtain the smoothing estimator $\widehat{\beta}_{j,\Omega^c}(\mathbf{d})$ for $\beta_{j,\Omega^c}(\mathbf{d})$. The final estimator for $\beta_j(\mathbf{d})$ is $\widehat{\beta}_j(\mathbf{d}) = 1_{\{\mathbf{d}\in\Omega\}}\widehat{\beta}_{j,\Omega}(\mathbf{d}) + 1_{\{\mathbf{d}\in\Omega^c\}}\widehat{\beta}_{j,\Omega^c}(\mathbf{d})$ and the corresponding estimator for $f_i(\mathbf{d})$ is given by $\widehat{f}_i(\mathbf{d}) = \sum_{j=1}^{p} x_{ij}\widehat{\beta}_j(\mathbf{d})$.

Yaroslavsky's filter The Yaroslavsky's filter explicitly uses the similarity between two voxels based on both their spatial distance and the relative proximity of image intensities in them. Specifically, in order to utilize the information from the image intensity, YF set the kernel function for the j-th coefficient in (3.2.3) as

$$K^{\text{adj}}_{j,h,h_y}(\mathbf{d}_m - \mathbf{d}) = h^{-3}\prod_{k=1}^{3} K_{\text{loc}}((d_{m,k} - d_k)/h) L_{h_y}(\widehat{\beta}_j(\mathbf{d}_m, \lambda), \widehat{\beta}_j(\mathbf{d}, \lambda)), \tag{3.2.7}$$

where $L_{h_y}(\widehat{\beta}_j(\mathbf{d}_m, \lambda), \widehat{\beta}_j(\mathbf{d}, \lambda)) = \mathbf{1}(\{|\widehat{\beta}_j(\mathbf{d}_m, \lambda) - \widehat{\beta}_j(\mathbf{d}, \lambda)| \le h_y\})$ is used to control the photometric proximity. Let $\widehat{\beta}^{\text{YF}}_{j,h,h_y}(\mathbf{d})$ be the estimator of $\beta_j(\mathbf{d})$ obtained from (3.2.3) based on the adjusted kernel function $K^{\text{adj}}_{j,h,h_y}(\mathbf{d}_m - \mathbf{d})$, we calculate the corresponding estimator for $f_i(\mathbf{d})$, given by $\widehat{f}^{\text{YF}}_{i,h,h_y}(\mathbf{d}) = \sum_{j=1}^{p} x_{ij}\widehat{\beta}^{\text{YF}}_{j,h,h_y}(\mathbf{d})$.

Non-local means method Non-local means can be regarded as an extension of YF for handling the unknown boundary case. Specifically, NLM uses the relative proximity of image intensity of two voxels through the discrepancy between patches surrounding the voxels considered, whereas YF only uses the discrepancy between two considered voxels. Define $\mathcal{P}_{\mathbf{d}} = \{\mathbf{d}' : \max_{1\le j\le 3}|d'_j - d_j| \le h_{\mathcal{P}}/2\}$, $m_{\mathcal{P}}$ the number of voxels contained in $\mathcal{P}_{\mathbf{d}_0}$ and $\mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}}} = \{\widehat{\beta}_j(\mathbf{d}_m, \lambda) : \mathbf{d}_m \in \mathcal{D}\bigcap\mathcal{P}_{\mathbf{d}}\}$ for $h_{\mathcal{P}} > 0$ and every $\mathbf{d} \in \mathcal{D}$. Then, the adjusted kernel function in NLM is given by

$$K^{\text{adj}}_{j,h,h_y}(\mathbf{d}_m - \mathbf{d}) = h^{-3}\prod_{k=1}^{3} K_{\text{loc}}((d_{m,k} - d_k)/h) L_{h_f}(\mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}_m}}, \mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}}}), \tag{3.2.8}$$

where $L_{h_f}(\mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}_m}}, \mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}}}) = \mathbf{1}(\{\|\mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}_m}} - \mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}}}\|_2 \le h_y\})$. In practice, it is computationally intensive to calculate $L_{h_f}(\mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}_m}}, \mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}}})$ for large $h_{\mathcal{P}}$. Several methods have been proposed to make a trade-off between computational complexity and accuracy (Azzabou et al., 2007; Tolga, 2009). Throughout this paper, we follow Mahmoudi and Sapiro (2005) and compare patches only by using their means. It leads to the following photometric kernel

$$L_{h_f}(\mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}_m}}, \mathcal{Y}_{j,\mathcal{P}_{\mathbf{d}}}) = \bar{L}_{h_f}(\bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_m}}, \bar{\beta}_{j\mathcal{P}_{\mathbf{d}}}) = \mathbf{1}(\{|\bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}_{i,\mathcal{P}_{\mathbf{d}}}| \leq h_y\}),$$

where $\bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_m}}$ is the average of $\widehat{\beta}_j(\mathbf{d}_m, \lambda)$ within the patch $\mathcal{P}_{\mathbf{d}_m}$. Let $\widehat{\beta}^{\text{NLM}}_{j,h,h_y}(\mathbf{d})$ be the estimator of $\beta_j(\mathbf{d})$ obtained from (3.2.3) based on the adjusted kernel function (3.2.8), we calculate the corresponding estimator for $f_i(\mathbf{d})$, given by $\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}) = \sum_{j=1}^{p} x_{ij}\widehat{\beta}^{\text{NLM}}_{j,h,h_y}(\mathbf{d})$.

3.3 Theoretical Properties

We systematically investigate the theoretical properties of all estimators obtained from the two-stage estimation procedure. Without otherwise stated, we assume that $o_p(1)$ and $O_p(1)$ hold uniformly across all $\mathbf{d}$ in either $\mathcal{D}_0$ or $\mathcal{D}$ throughout the article. Moreover, both the sample-size n and the number of voxels N_D are allowed to diverge to infinity. We define the oracle estimator $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)$ as

$$\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) = \underset{\text{supp}(\boldsymbol{\beta})\subseteq\mathcal{S}(\mathbf{d}_0)}{\text{argmin}} \ \mathcal{L}(\boldsymbol{\beta}; Y(\mathbf{d}_0), X), \tag{3.3.1}$$

in which $\mathcal{S}(\mathbf{d}_0)$ is the true support set of $\boldsymbol{\beta}^*(\mathbf{d}_0)$. We state the following theorems, whose detailed proofs can be found in a supplementary document.

We need the following assumptions to facilitate the technical details, although they may not be the weakest conditions.

Assumption 3.1. *The penalty function satisfies the following conditions: (i) $p_\lambda(\cdot)$ is symmetric on $\mathcal{R}$, non-decreasing and differential on $x > 0$; (ii) $p_\lambda(x) + 0.5\mu\|x\|_2^2$ is convex for some $\mu > 0$; (iii)$\dot{p}_\lambda(x) = dp_\lambda(x)/dx$ is continuous and non-increasing on* $(0, \infty)$ *and* $\lim_{x\to 0+}\dot{p}_\lambda(x) = \lambda$; *and (iv) there exists a constant $\gamma > 0$ such that $\dot{p}_\lambda(x) = 0$ for any* $|\beta| \geq \gamma\lambda > 0$.

Assumption 3.2. *The localized restricted strong convexity (LRSC) assumption holds, that is, there exists a* $\rho > 0$ *such that*

$$\langle \nabla \mathcal{L}(\boldsymbol{\beta}^* + \boldsymbol{\Delta}) - \nabla \mathcal{L}(\boldsymbol{\beta}^*), \boldsymbol{\Delta} \rangle \geq \rho \|\boldsymbol{\Delta}\|_2^2 - \tau \frac{\log p}{n} \|\boldsymbol{\Delta}\|_1^2, \ \forall \ \|\boldsymbol{\Delta}\|_2 \leq r,$$

where $\tau > 0$ *is a tolerance parameter and* $r > 0$ *is a local radius parameter.*

Assumption 3.3. $\epsilon_i(\mathbf{d})$*'s are i.i.d. sub-Gaussian random variables with mean* 0 *and variance proxy* σ^2*, and* $\max_{\mathbf{d}_0 \in M} \|\nabla \mathcal{L}_{\mathbf{d}_0}(\boldsymbol{\beta}^*)\|_\infty \vee \|\nabla \mathcal{L}_{\mathbf{d}_0}(\widehat{\boldsymbol{\beta}}^0)\|_\infty \lesssim \sigma \sqrt{\log(p|M|/\delta)/n}$ *holds with probability at least* $1 - \delta$ *for any set* M*.*

Assumption 3.4. $\mathbf{d}_0 \in \mathcal{D}_0$*,* $\boldsymbol{x}_{i,\mathcal{S}(\mathbf{d}_0)}^{\mathrm{T}} (X_{\mathcal{S}(\mathbf{d}_0)}^{\mathrm{T}} X_{\mathcal{S}(\mathbf{d}_0)}/n)^{-1} \boldsymbol{x}_{i,\mathcal{S}(\mathbf{d}_0)} \asymp s$ *for all* $i = 1, \ldots, n$*.*

Assumption 3.5. *For every* $\mathbf{d}_0 \in \mathcal{D}_0$*,* $\lambda_{\max}(X_{\mathcal{S}(\mathbf{d}_0)}^{\mathrm{T}} X_{\mathcal{S}(\mathbf{d}_0)}/n) \asymp \lambda_{\min}(n^{-1} X_{\mathcal{S}(\mathbf{d}_0)}^{\mathrm{T}} X_{\mathcal{S}(\mathbf{d}_0)}/n) = O(1)$*.*

Assumption 3.6. *Suppose there exists* $C_1 > 0$ *such that* $(1/(pN_D))^{C_1} = o((\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5})$ *and* $\sigma = o(n/(s + \log(N_D)))$*.*

Assumption 3.7. *Suppose* $\lambda \asymp \sigma \sqrt{\log(pN_D)/n}$*,* $s/n = o(1)$ *and* $(\log N_D)/n = o(1)$ *.*

Assumption 3.8. *The bias and the variance of the oracle estimator have the following orders*

$$\|bias(\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0))\|_2 := \|E\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_2 \lesssim \sqrt{\frac{s}{n}\sigma^2},$$

$$var(\boldsymbol{u}^T \widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)) \lesssim \frac{s}{n}\sigma^2$$

for any unit vector $\boldsymbol{u} \in \mathcal{R}^{p \times 1}$*.*

Assumption 3.9. *The tail probability of* $\|\bar{\beta}_{\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{\mathcal{P}_{\mathbf{d}_m}}\|_\infty$ *satisfies*

$$\mathbb{P}\left\{ \max_{\mathbf{d}_m \in B(\mathbf{d}_0, h)} \|\bar{\beta}_{\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{\mathcal{P}_{\mathbf{d}_m}}\|_\infty \gtrsim \sigma(s + \log(m_{\mathcal{P}} N_D))/n + \sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})} \right\}$$
$$= o((\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5}).$$

Remark 1. Assumption 3.1 is satisfied for several folded concave penalty functions such as SCAD, MCP and the capped ℓ_1 penalty. The local ball constraint can be used in the case where the loss function is not quadratic like when the parameters are far away from the underlying true, such as the Huber loss function. This can be seen as a generalization of the restricted strong convexity condition by Loh and Wainwright (2015). Assumption 3.3 illustrate the distribution of measurement error ϵ and the tail probability of the first derivative of loss function. Assumptions 3.4 and 3.5 imply the scale of the covariate matrix $X_{\mathcal{S}(\mathbf{d}_0)}$ for each $\mathbf{d}_0$. Assumptions 3.6 and 3.7 assume the order of tuning parameter λ, stand deviation σ, the ratio of s and n and the ratio of n and p for selecting the support set. Assumptions 3.8 and 3.9 describe the orders of bias and variance of the oracle estimator and the tail probability of oracle estimator average within the patch for NLM respectively. These assumptions are satisfied for the loss function such as L_2 loss and Huber loss. The validation of these assumptions using L_2 loss and Huber loss can be found in Appendix.

We first investigate the consistency property of the estimated support set across all voxels $\mathbf{d}_0$.

Theorem 3.3.1. Assume Assumptions 3.1 and 3.2 hold for loss $\mathcal{L}_{\mathbf{d}_0}$ with $3\mu < 4\rho$ and $r = 2R$. Suppose $\|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_\infty \vee \|\nabla \mathcal{L}_{\mathbf{d}_0}(\boldsymbol{\beta}^*)\|_\infty \vee \|\nabla \mathcal{L}_{\mathbf{d}_0}(\widehat{\boldsymbol{\beta}}^0)\|_\infty \le \lambda/2$, and $\tau s (\log p/n)^{3/2} \le c\lambda$ for some small enough constant c. Suppose $\min_{j \in \mathcal{S}(\mathbf{d}_0)} |\beta_j^*| \ge (\gamma + 1/2)\lambda$. Then

$$\widehat{\mathcal{S}}(\mathbf{d}_0, \lambda) = \mathcal{S}(\mathbf{d}_0),\ \widehat{\boldsymbol{\beta}}(\mathbf{d}_0, \lambda) = \widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0).$$

Further if Assumption 3.2 holds for all losses, $\max_{\mathbf{d}_0 \in \mathcal{D}_0} \|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_\infty \vee \|\nabla \mathcal{L}_{\mathbf{d}_0}(\boldsymbol{\beta}^*)\|_\infty \vee \|\nabla \mathcal{L}_{\mathbf{d}_0}(\widehat{\boldsymbol{\beta}}^0)\|_\infty \le \lambda/2$, and $\min_{\mathbf{d}_0 \in \mathcal{D}_0} \min_{j \in \mathcal{S}(\mathbf{d}_0)} |\beta_j^*(\mathbf{d}_0)| \ge (\gamma + 1/2)\lambda$, then

$$\widehat{\mathcal{S}}(\mathbf{d}_0, \lambda) = \mathcal{S}(\mathbf{d}_0) \quad \text{and} \quad \widehat{\boldsymbol{\beta}}(\mathbf{d}_0, \lambda) = \widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) \text{ for all } \mathbf{d}_0 \in \mathcal{D}_0.$$

Remark 2. Theorem 3.3.1 states that the support set $\mathcal{S}(\mathbf{d}_0)$ for each $\mathbf{d}_0$ can be estimated consistently and the the penalized M-estimator is equal to the oracle estimator with high probability. These results also stand uniformly for all the $\mathbf{d}_0 \in \mathcal{D}_0$ under more strict scaling conditions.

To evaluate the performance of the estimate method, we define

$$\mathrm{MSE}_f(\widehat{f}) = \frac{\sum_{\mathbf{d}_0 \in \mathcal{D}_0} \sum_{i=1}^{n} E(\widehat{f}_i(\mathbf{d}_0) - f_i^*(\mathbf{d}_0))^2}{nN_D} = \frac{\sum_{\mathbf{d}_0 \in \mathcal{D}_0} E\|\widehat{\boldsymbol{f}}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2}{nN_D},$$

where $\widehat{\boldsymbol{f}}(\mathbf{d}_0) = (\widehat{f}_1(\mathbf{d}_0), \ldots, \widehat{f}_n(\mathbf{d}_0))^{\mathrm{T}}$ and $\boldsymbol{f}^*(\mathbf{d}_0) = (f_1^*(\mathbf{d}_0), \ldots, f_n^*(\mathbf{d}_0))^{\mathrm{T}}$. Further define $\mathcal{R}_n(\widehat{f}) = \sup_{f \in \mathcal{F}} \mathrm{MSE}_f(\widehat{f})$ and the minimax risk $\mathcal{R}_n^* = \inf_{\widehat{f}} \sup_{f \in \mathcal{F}} \mathrm{MSE}_f(\widehat{f})$.We say the estimator $\widehat{f}$ achieves the minimax risk if $\mathcal{R}_n(\widehat{f}) = O(\mathcal{R}_n^*)$. Since we already have the value of f_i^* belongs to $[0,1]$, we clip $\widehat{f}_i$ so that it also takes value in $[0,1]$ and this will not increase MSE.

As pointed out in section 3.2.3, oracle kernel knows which voxel belongs to Ω or to Ω^c. This information is sufficient to help us smooth f as if there is no discontinuity. We call the estimator that uses the oracle kernel the oracle estimator. We first focus on the performance of oracle estimator and consider it as benchmark for other adaptive kernels.

Theorem 3.3.2. Under Assumptions 3.1-3.8, the risk of the oracle estimator has the following upper bound

$$\inf_h \mathcal{R}_n(\widehat{f}_h) \lesssim (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \vee \sigma^2 s^2 n^{-1}$$

with the optimal choice of bandwidth $h \asymp (\sigma^2 s^2 n^{-1} N_D^{-1})^{1/5}$. If we further assume that $\epsilon_i(\mathbf{d})$'s are i.i.d. Gaussian random variables, then

$$(\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \lesssim \inf_h \mathcal{R}_n(\widehat{f}_h) \lesssim (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \vee \sigma^2 s^2 n^{-1}.$$

If $N_D \lesssim (\sigma^{-2} s^{-2} n)^{\frac{3}{2}}$, the oracle estimator achieves the minimax risk

$$\inf_h \mathcal{R}_n(\widehat{f}_h) \asymp \mathcal{R}_n^* \asymp (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5}.$$

Remark 3. Theorem 3.3.2 states the lower and upper bound for the oracle estimator. Compared with the Theorem 4.2 in (Arias-Castro et al., 2012), our lower bound is the similar minimax bound (Korostelev and Tsybakov, 2012) and requires the errors have Gaussian distribution while our upper bound has the extra term $\sigma^2 s^2 n^{-1}$, which comes from fitting the linear model voxel-wise in our first stage. This upper bound holds for the general sub-gaussian errors.

We next study the performance of YF. YF usually performs well when the noise level is not high.

Theorem 3.3.3. Under Assumptions 3.1-3.8 and if $\sigma \lesssim \sqrt{n/(s\log(pN_D))}$,then we have

$$\inf_{h,h_y} \mathcal{R}(\widehat{f}^{\mathrm{YF}}_{h,h_y}) \asymp \inf_h \mathcal{R}_n(\widehat{f}_h)$$

with $h \asymp (\sigma^2 s n^{-1} N_D^{-1})^{1/5}, h_y \asymp 1$.

Remark 4. Theorem 3.3.3 states that YF can mimic the oracle kernel and achieve a performance comparable to that of oracle kernel when the noise level is not high. However, the results are no longer true when the noise level is of higher order. This is because under strong noise level YF would smooth the voxel from both Ω and Ω^c, which makes the separation at the boundary region $\partial\Omega$ impossible.

In the previous section, we have discovered that YF can perform as well as the oracle kernel when the noise level is not high. To adapt YF in the case of strong noise level, a natural idea is to use a two-stage algorithm: denoise first and then apply YF. Indeed, this is the rational behind NLM. We first investigate the performance of NLM when the noise level is not high.

Theorem 3.3.4. Under Assumptions 3.1-3.8 and if $\sigma \lesssim \sqrt{n/(s\log(pN_D))}$, then we have

$$\inf_{h,h_y} \mathcal{R}(\widehat{f}^{\mathrm{NLM}}_{h,h_y}) \asymp \inf_h \mathcal{R}_n(\widehat{f}_h),$$

with $h_{\mathcal{P}} = 1/N_D^{1/3}, h \asymp (\sigma^2 s n^{-1} N_D^{-1})^{1/5}, h_y \asymp 1$.

Remark 5. Theorem 3.3.4 states that if the same noise condition in Theorem 3.3.3 is satisfied, the patch with width $h_{\mathcal{P}} = 1/N_D^{1/3}$ only contains one single voxel, which would make NLM degenerate to YF, and we can still achieve the oracle risk bound.

We next consider the performance of NLM method under the general case of noise level.

Theorem 3.3.5. Under Assumptions 3.1-3.9, we have the upper bound of the risk as

$$\inf_{h,h_y} \mathcal{R}(\widehat{f}^{\text{NLM}}_{h,h_y}) \lesssim \frac{(\sigma^2 s n^{-1} \log p N_D)^{1/3}}{N_D^{1/3}} \vee \sigma^2 s^2 n^{-1},$$

with $h_{\mathcal{P}} \asymp \frac{(\sigma^2 n^{-1} \log p N_D)^{1/3}}{N_D^{1/3}}$, $h_y \asymp 1$, $h \asymp (\sigma^2 s n^{-1} N_D^{-1})^{1/5}$.

Remark 6. Theorem 3.3.5 gives an upper bound for the risk of the NLM estimator for the general noise level. Compared with the Theorems 4.4-4.5 in Arias-Castro et al. (2012), which require $\sigma = o(1)$ for YF and $\sigma = O(1)$ for NLM, our Theorems 3.3.3-3.3.5 can allow for stronger noise level. This is because we utilize the cross-image information to reduce the noise level in the first stage and integrate the spatial constraints in the second stage to denoise further. This also illustrates that incorporating the covariates into denoising process is helpful for image noise removal again.

3.4 Simulation Studies

We examined the finite sample performance of our two-stage framework on the reconstruction of ensemble average propagator (EAP) by simulating synthetic high angular resolution diffusion imaging (HARDI) (Tuch et al., 1999). Data were generated from the multi-tensor model (Alexander et al., 2002; Tuch, 2004) given by

$$E(\mathbf{q}_i; \mathbf{d}) = \sqrt{(\sum_{k=1}^{K} p_k e^{-b_i \mathbf{u}_i^{\mathrm{T}} \mathbf{D}_k(\mathbf{d}) \mathbf{u}_i} + \sigma \epsilon_{i1})^2 + (\sigma \epsilon_{i2})^2} \quad \text{for} \quad i = 1, \cdots, n \tag{3.4.1}$$

where $\mathbf{q}_i = q_i \mathbf{u}_i$ with $\mathbf{u}_i$ being a unit vector and $q_i = ||\mathbf{q_i}||_2$, b_i is the b-value for the i-th gradient direction, p_k is the weight and $\mathbf{D}_k(\mathbf{d})$ is the tensor matrix for the k-th fiber, K is the number of fibers, ϵ_{i1} and ϵ_{i2} come from independent standard normal distribution and the signal-to-noise ratio (SNR) is defined as $1/\sigma$. The synthetic data contained four shells ($b = 500, 1000, 2000$, and 3000 s/mm^2) and each shell had 81 sampling directions ($n = 324$) on the hemisphere for the 3-rd order tessellation of the icosahedron.

We generated the voxels with a single fiber and the voxels with isotropic tensor from a single tensor model using $D_1(\mathbf{d})$ with eigenvalues $[1.7, 0.3, 0.3] \times 10^{-3}$mm^2/s and $[1, 1, 1] \times 10^{-3}$mm^2/s respectively. The voxels with two fiber directions were generated by a two-tensor model $E(\mathbf{q}_i; \mathbf{d}) = e^{-b_i \mathbf{u}_i^{\mathrm{T}} \mathbf{D}_1(\mathbf{d}) \mathbf{u}_i}/2 + e^{-b_i \mathbf{u}_i^{\mathrm{T}} \mathbf{D}_2(\mathbf{d}) \mathbf{u}_i}/2$. The direction for one of the two fibers was along $x-$axis and the other one had changing angles with $x-$axis from 45^o , 60^o, 75^o, to 90^o and then from 90^o, 75^o, 60^o to 45^o.

We represented the signals using spherical polar Fourier imaging (SPFI) (Assemlal et al., 2008, 2009), which is a model-free and fast method for multiple-shell HARDI data (Cheng et al., 2010a). It represents the diffusion signal by spherical polar Fourier (SPF) basis

$$E(\mathbf{q}_i; \mathbf{d}) = \sum_{k=0}^{K} \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{k,l,m}(\mathbf{d}) B_{k,l,m}(\mathbf{q_i}) \tag{3.4.2}$$

where $B_{k,l,m}(\mathbf{q_i})$ is SPF basis and $a_{k,l,m}(\mathbf{d})$ is the corresponding coefficient. In this case, the representation can be regarded as a special case of model (3.2.1) if we set $y_i(\mathbf{d}) = E(\mathbf{q}_i; \mathbf{d})$, $\mathbf{x}_i = (B_{0,0,0}(\mathbf{q}_i) \cdots, B_{K,L,L}(\mathbf{q}_i))^{\mathrm{T}}$ and $\beta(\mathbf{d}) = (a_{0,0,0}(\mathbf{d}), \cdots, a_{K,L,L}(\mathbf{d}))^{\mathrm{T}}$. EAP can be further represented using Spherical Harmonic (SH) basis

$$p(R_0 \mathbf{r}; \mathbf{d}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} c_{l,m}(R_0, \mathbf{d}) Y_l^m(\mathbf{r}) \tag{3.4.3}$$

where $Y_l^m(\mathbf{u})$ is the l order m degree SH basis, $\mathbf{r}$ is a unit vector, $R_0 = 15\mu m$ and the coefficients $\{c_{l,m}\}$ can be written as a linear transformation from $\{a_{k,l,m}\}$. As suggested by (Cheng et al., 2011), we chose $K = 8$ and $L = 4$ to get a sparse representation in (3.4.2). We took the L2 loss function $\mathcal{L}(\boldsymbol{\beta}; Y(\mathbf{d}_0), X) = \|Y(\mathbf{d}_0) - X\boldsymbol{\beta}\|_2^2$ and SCAD penalty function (Fan and Li, 2001) in (3.2.2) and the tuning parameter λ in (3.2.2) was changed to $\lambda_{k,l,m} = \lambda_K k^2(k+1)^2 + \lambda_L l^2(l+1)^2$ with $\lambda_K = 2e^{-8}$ and $\lambda_L = 1e^{-6}$ in our simulation studies.

We first estimated the coefficients $\beta = [a_{k,l,m}]$ in (3.4.2) by SCAD estimation and then smoothed the coefficients estimates using YF and NLM respectively. Denote

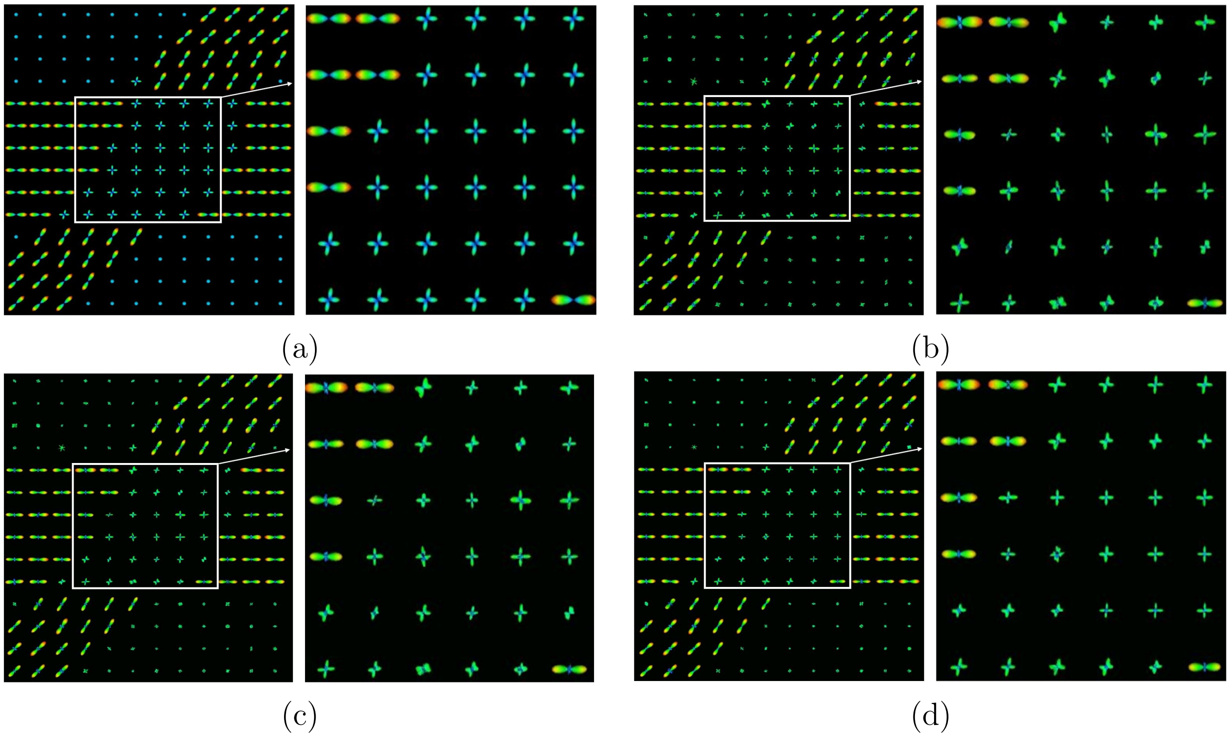
$$\widehat{\sigma}(\mathbf{d}_0)^2 = \frac{(Y(\mathbf{d}_0) - \widehat{\beta}(\mathbf{d}_0, \lambda))^T (Y(\mathbf{d}_0) - \widehat{\beta}(\mathbf{d}_0, \lambda))}{n - |\widehat{\mathcal{S}}(\mathbf{d}_0, \lambda)|}$$

$$v_1(\mathbf{d}_0) = diag((X^T_{\mathcal{S}(\mathbf{d}_0)} X_{\mathcal{S}(\mathbf{d}_0)})^{-1} \widehat{\sigma}(\mathbf{d}_0)^2)$$

The patch size for NLM was $3 \times 3 (i.e. h_{\mathcal{P}} = 3)$ and the photometric bandwidth h_y was $\chi^2_1(0.8)v_0$ for YF and was $\chi^2_1(0.8)v_0/h^2_{\mathcal{P}}$ for NLM, where v_0 is the 75% quantile for the set $\{v_1(\mathbf{d}_0) : \mathbf{d}_0 \in \mathcal{D}_0\}$. For the spatial bandwidth h, they were chosen to minimize the overall MSE $= \sum_{i=1}^n \sum_{\mathbf{d} \in \mathcal{D}_0} (\widehat{f}_i(\mathbf{d}) - f_i(\mathbf{d}))^2$.

Figure 3.1, 3.2 and 3.3 give the comparison of EAP reconstructions of SCAD, YF and NLM on data with SNR = 10, 15 and 20 respectively. We can see that YF and NLM results are closer to ground truth than SCAD result because they can reduce the noise by incorporation useful neighborhood information into the estimation procedure. It is also observed that YF is the closest to ground truth under low noise level (SNR = 20) while NLM performs best under strong noise level(SNR = 10 and 15).

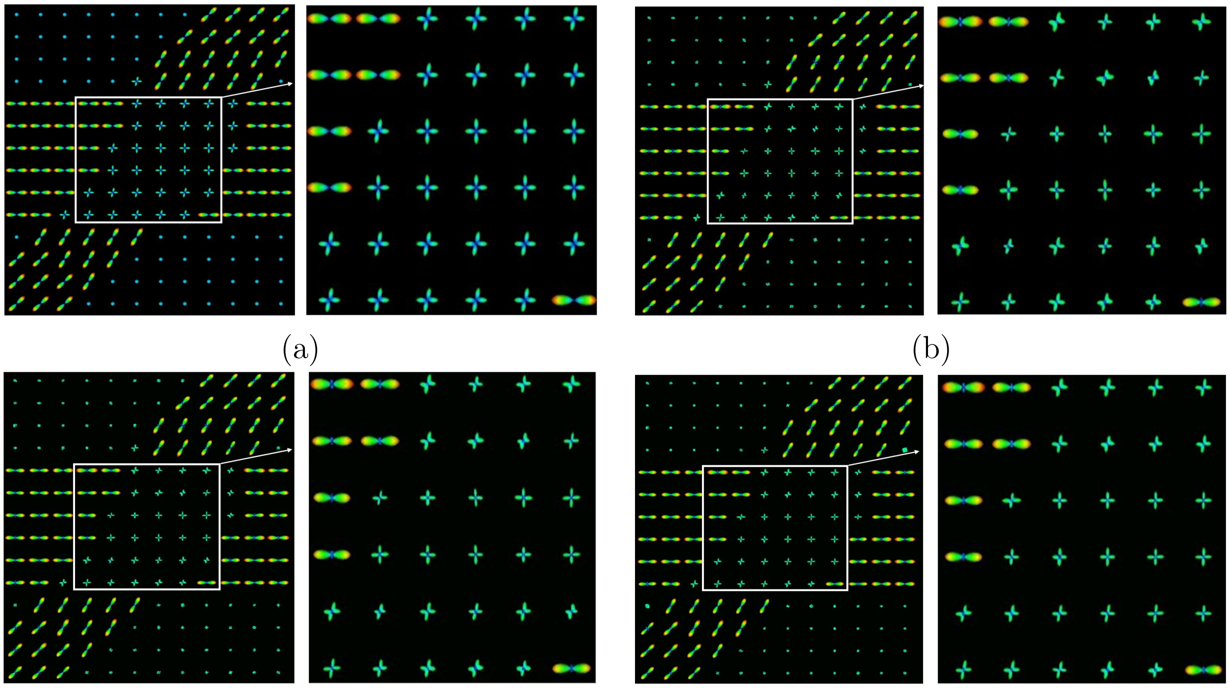
Furthermore, we examined the performance of the two-stage framework quantitatively. We simulated 500 data sets for SNR = 10, 15, and 20, respectively. We first calculated the overall MSE and the result was summarized in Table 3.1. We further extracted the EAP maxima aligned with fiber directions in the anisotropic(with fiber) region and calculated the angle detection errors, which is defined as the difference between the recovered fiber direction with the ground truth for the voxels with a single fiber and the difference between the recovered crossing angles with the ground truth for the voxels with two crossing fibers. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers were calculated and the result was summarized in the Table 3.2. We can see that



(a) (b)

(c) (d)

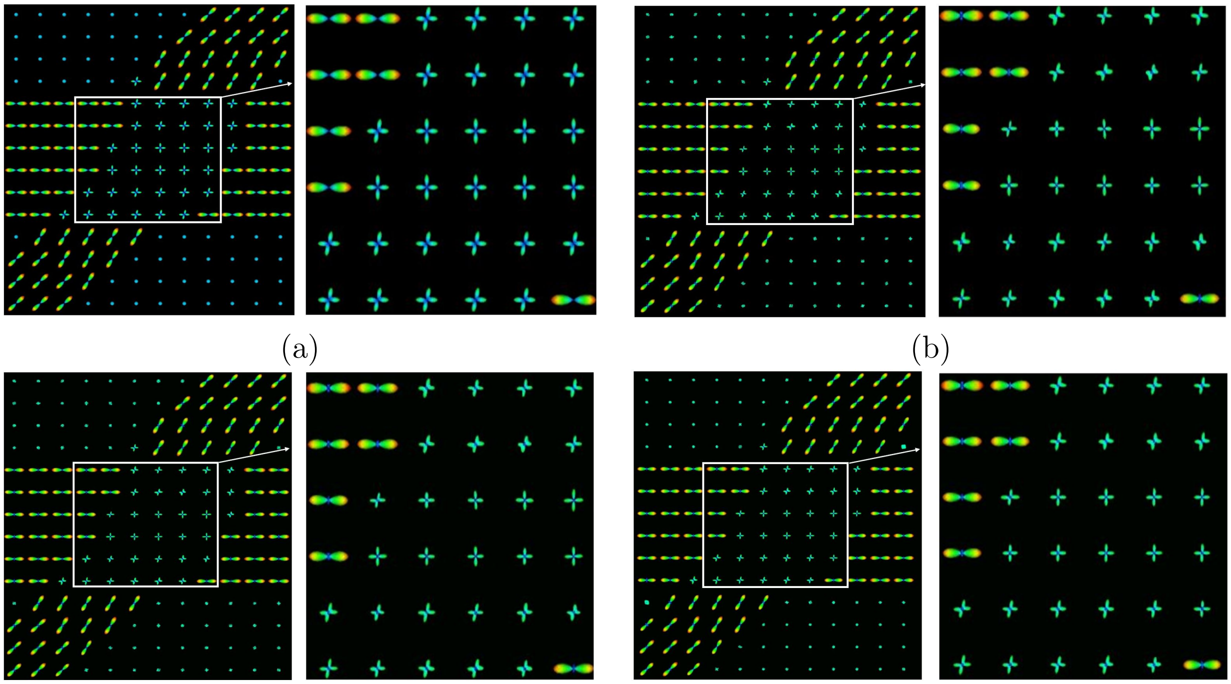
Figure 3.1: Comparison of three estimation methods based on simulated data sets corresponding to SNR=10: (a) the EAP image of ground truth; (b) the EAP image based on the SCAD estimation; (c) the EAP image based on the YF estimation; and (d) the EAP image based on the NLM estimation;



(a) (b)

(c) (d)

Figure 3.2: Comparison of three estimation methods based on simulated data sets corresponding to SNR=15: (a) the EAP image of ground truth; (b) the EAP image based on the SCAD estimation; (c) the EAP image based on the YF estimation; and (d) the EAP image based on the NLM estimation;



(a) (b)

(c) (d)

Figure 3.3: Comparison of three estimation methods based on simulated data sets corresponding to SNR=20: (a) the EAP image of ground truth; (b) the EAP image based on the SCAD estimation; (c) the EAP image based on the YF estimation; and (d) the EAP image based on the NLM estimation;

SNR	Noise	SCAD	YF	NLM
10	547.16 (3.06)	222.19 (2.32)	154.10 (2.19)	129.10 (0.38)
15	240.02 (1.35)	84.70 (0.95)	56.69 (0.87)	54.40 (1.37)
20	134.32 (0.76)	44.15 (0.52)	29.97 (0.47)	33.05 (1.32)

Table 3.1: The MSE of EAP image at three SNR levels.

both YF and NLM can reduce the MSE level and mean of the angular errors and increase the percentage of detecting correct number of fibers at all the noise levels. Under low noise level(SNR = 20), YF performs better than NLM while NLM performs better than YF under strong noise level(SNR = 10 and 15).

All voxels with one fiber

SNR	SCAD	YF	NLM
10	2.75±0.48 (99.97 %)	2.48±0.58 (99.96 %)	1.64±0.85 (99.98 %)
15	1.23±0.78(100 %)	1.04±0.84 (100 %)	0.73±0.89 (100 %)
20	0.78±0.90 (100 %)	0.69±0.89 (100 %)	0.62±0.86 (100 %)

All voxels with two fibers

SNR	SCAD	YF	NLM
10	7.66±1.08 (88.45 %)	7.22±1.17 (89.98 %)	6.04±1.10 (93.19 %)
15	4.14±0.69(99.87 %)	3.76±0.68 (99.88 %)	3.39±0.84 (99.92 %)
20	2.76±0.63(100 %)	2.37±0.61 (100 %)	2.40±1.11 (1000 %)

Table 3.2: The means of angular errors and percentages of correct number (in parentheses) of detected EAP maximum at three SNR levels.

3.5 Real Data Applications

We applied our proposed method to reconstruct EAP image via SPFI on a raw diffusion Magnetic Resonance Imaging(dMRI) of a healthy adult. The dMRI data contains 24 shells with 313 volumes corresponding to 13 b_0s and 300 b values ($n = 300$). The b values range from 125 to 3000 s/mm^2 with a step size 125 and the number of volume in each shell ranges from 1 to 24. The imaging protocol is as follows: 140×140 imaging matrix, 1.5mm $\times$ 1.5mm $\times$ 1.5mm resolution, TE=89 ms, TR=2513 ms and multi band factor 5. The image data was normalized into $[0, 1]$ by the average of b_0 images.

We first fitted SCAD to get the estimates of SPF coefficients through (3.4.2) and then smoothed the coefficients using YF and NLM respectively. For the spatial bandwidth h of YF and NLM, since we did have the ground truth image, we minimized the overall generalized cross-validation (GCV) (Fan and Gijbels, 1996) score given by

$$\sum_{i=1}^{p} \mathrm{GCV}_i(h) = \sum_{i=1}^{p} \frac{\widetilde{\beta}_i^{\mathrm{T}}(I_D - W_i)^{\mathrm{T}}(I_D - W_i)\widetilde{\beta}_i}{[1 - {N_D}^{-1}\mathrm{tr}(W_i)]^2}, \tag{3.5.1}$$

where I_D is an $N_D \times N_D$ identity matrix and W_i is from equation (3.2.5). We chose tuning parameter λ as $\lambda_K = 1e^{-7}$ and $\lambda_L = 5e^{-6}$. All other parameter settings were same with the simulation studies. Figure 3.4 presents one slice and two selected ROIs of the EAP image based on the SCAD estimation. Figures 3.5 and 3.6 show the estimated EAP images for the selected ROI1 and ROI2 of Figure 3.4, respectively. We observe that YF and NLM lead to better EAP reconstruction results than SCAD in terms of smoother EAPs along fiber tracts. This is because SCAD fits the regression model voxel by voxel and does not consider the spatial structure of imaging data.

We further used the reconstructed EAP to perform whole-brain tractography as described in (Wu et al., 2020, 2019). We first calculated the EAP samples of each voxel from the reconstructed EAP on 362 directions and normalized them into $[0, 1]$ by the maximum of all the directions. Tractography then followed the fiber orientation derived from EAP samples and made use of generalized fractional anisotropy (GFA) image (Özarslan et al., 2005) as stopping criteria and the prior information from previous step to generate reliable results. Whole-brain tractography was conducted using deterministic fiber tracking method with 2 seeds per voxel within brain parenchyma. The tractography dataset came up with approximately 1M fiber streamlines and and each streamline has a step size of 0.2 mm. All the tractography results went through a quality check manually. Taking advantage of streamline geometry rather than the anatomical landmarks, the streamlines were then clustered into fiber bundles automatically (Wu et al., 2020). Specifically, each streamline was mapped from the native point space to a Hilbert space and then represented by the linear combination of a series of the cosine basis functions(Chung et al., 2010). This representation only relied on the degrees of the cosine basis functions and robust to the number of sampling points along the fiber. Tract Dictionary Learning (Kumar et al., 2019b) was further performed based on the cosine coefficients to learn the dictionary for each bundle. The $\mathbb{L}_1$ penalty was added to enforce each dictionary have sparse coefficients. Finally, each fiber was classified to a fiber bundle by minimizing the distance to each dictionary. We chose the cluster size as 1000 and the results of Cortico and Thalamic Pathway Parietal and Cortico-Fugal Parietal bundles were summarized in Figure 3.7. The axial and sagital views were displayed and we can see that NLM and YF lead to better tractography results than SCAD in terms of smoother fiber tracts.

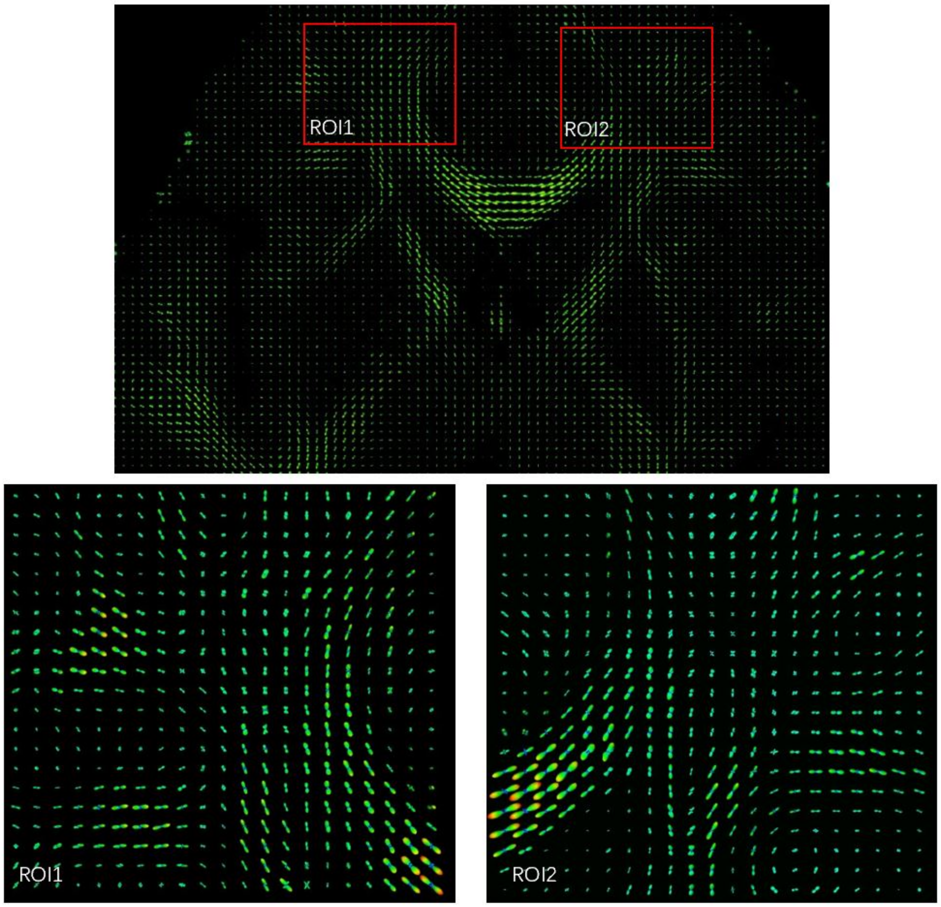
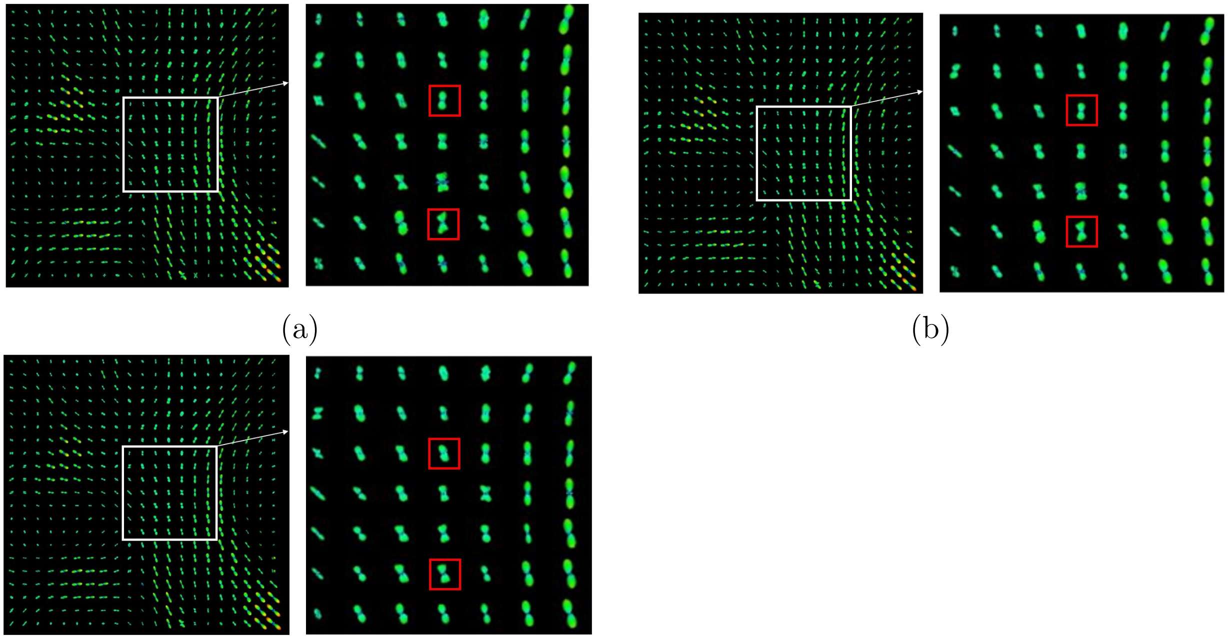


Figure 3.4: The EAP image based on the SCAD estimation on a selected slice and two ROIs(ROI1 and ROI2) from the raw 24-shell data of the healthy subject.

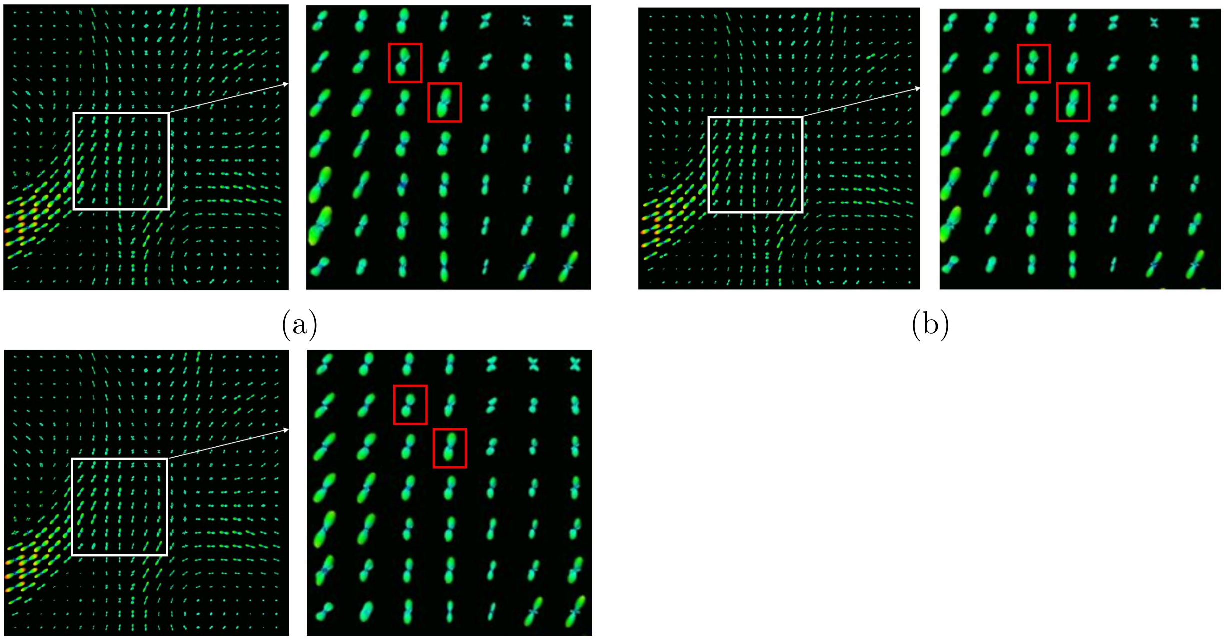


(a)

(b)

(c)

Figure 3.5: Comparisons of the three estimation methods for ROI1 in Figure 3.4: (a) the EAP image based on the SCAD estimation; (b) the EAP image based on the YF estimation; (c) the EAP image based on the NLM estimation.

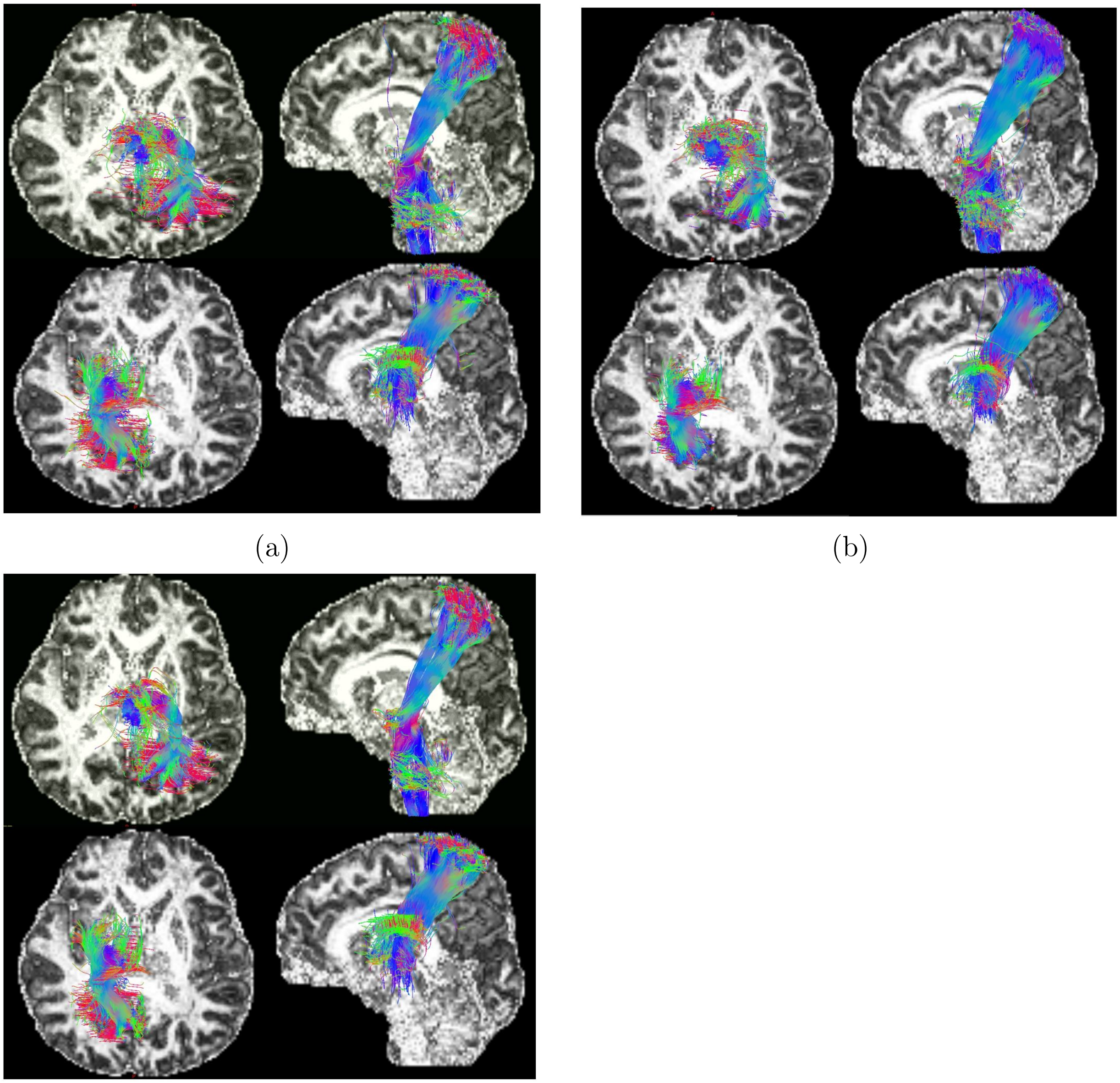


(a)

(b)

(c)

Figure 3.6: Comparisons of the four estimation methods for ROI2 in Figure 3.4: (a) the EAP image based on the SCAD estimation; (b) the EAP image based on the YF estimation; (c) the EAP image based on the NLM estimation.



(a)

(b)

(c)

Figure 3.7: Comparisons of tractography results obtained from tracking the estimated EAPs by using the three estimation methods based on the Cortico and Thalamic Pathway Parietal (top) and Cortico-Fugal Parietal (bottom) bundles: (a) the SCAD estimation; (b) the YF estimation; and (c) the NLM estimation.

3.6 Discussion

We have introduced a two-stage method to denoise and reconstruct imaging data. Our method first fits the regularized regression model utilizing covariates information among subjects at each voxel. Then the kernel-based smoothing methods YF and NLM are applied by borrowing the spatial information from the neighboring voxels adaptively to keep the feature of piecewise smooth regions with jump discontinuities. Moreover, we derive the oracle inequalities and minimax rates on MSE of YF and NLM under different noise levels and calculate the optimal choice of bandwidth of kernel functions.

Some important issues need to be addressed in future research. First, our initial penalized $M-$estimator relies on the condition that the number of covariates that related to the imaging data should be much smaller than the sample size. Therefore, when this condition does not hold, some other type of regularized regression model, such as Ridge Regression can be considered to use. Second, we assume the measurement errors in (3.2.1) are independent and identical copies of normal distribution. We can further extend it to be other complex distribution with correlation between different voxels (Zhu et al., 2014).

CHAPTER 4: LIFESPAN POPULATION-BASED STRUCTURAL CONNECTOME MAPPING

4.1 Introduction

With recent development in imaging technologies, many large-scale biomedical studies, such as the UK Biobank (Miller et al., 2016) and Human Connectome Project (HCP) (Sotiropoulos et al., 2013; Van Essen et al., 2013), have collected/are collecting massive multi-modal imaging data (e.g., structural magnetic resonance imaging (MRI), diffusion MRI (dMRI), and functional MRI (fMRI)) with high spatial and temporal resolution as well as other complex information (e.g., genomics). The HCP represents the first large-scale attempt to map the human brain's connections in high resolution from around 1,200 healthy adults aged 22-35 years old to study white matter (WM) structure (Glasser et al., 2016). The Lifespan Human Connectome Studies (LHCS) projects are extending the HCP protocols to accurately map brain connections in much larger cohorts of developing, adult, and aging subjects compared with the HCP (Howell et al., 2019; Bookheimer et al., 2019; Harms et al., 2018; Somerville et al., 2018). Furthermore, fourteen Disease Human Connectome Studies (DHCS) studies collected disease connectome data from subjects with certain clinical diagnoses (e.g., Alzheimer). These rich data will allow us to quantify the dynamic process of WM across the lifespan and address many fundamental questions on the connectional organization of the human brain, which may eventually establish a baseline that would help identify connectivity abnormalities across different spatial and/or temporal scales in brain disorders, such as schizophrenia and Alzheimer (Durston, 2010; Satterthwaite et al., 2014; Knickmeyer et al., 2014; Van Essen et al., 2013). Identifying such abnormalities could transform our understanding of the origins of these disorders and inspire new, urgently needed approaches to urgently needed prevention, diagnosis, and treatments.

To map the structural connectome across the lifespan at the population level, the main challenge is the large intra-subject heterogeneity of WM fibers and the complex inter-subject spatio-temporal patterns of fibers. Particularly, the quality of WM tractographic data varies strongly by age, anatomical tract, tractography algorithm, and acquisition parameters (e.g., directions), among others, so it is critically important to accurately segment WM tractographic data (St-Jean et al., 2019; Guevara et al., 2011a; Buchanan et al., 2014; Wakana et al., 2007; Guevara et al., 2020; Schilling et al., 2019).

To address this challenge, existing analytical methods include three major types (i) tract-based spatial statistics (TBSS)(Smith et al., 2006) and other voxel-based analysis methods (Smith et al., 2006; Schwarz et al., 2014; Snook et al., 2007) (ii) tract-based fiber parcellation methods (Fornito et al., 2013; Zhu et al., 2011; Yeatman et al., 2012; Cousineau et al., 2017; Jin et al., 2014; Heiervang et al., 2006; Ciccarelli et al., 2003; Wang et al., 2016; Wassermann et al., 2010; Garyfallidis et al., 2017; Olivetti et al., 2017; Sharmin et al., 2016) and (iii) hybrid method(O'Donnell et al., 2013; Guevara et al., 2017; Zhang et al., 2018). TBSS is a robust method that projects WM diffusion properties onto a whole brain WM skeleton. It is independent of tratographic data but relies on the existing registration methods, which therefore does not have individual fiber tract specificity and can not model WM fiber architecture in the registration process (Zalesky et al., 2010; Yeatman et al., 2012). Tract-based methods include various fiber parcellation methods, consisting of fiber clustering strategies that directly cluster WM fiber bundles (O'Donnell et al., 2013; Guevara et al., 2017; Jin et al., 2014; Guevara et al., 2011b; Garyfallidis et al., 2017) and cortical-parcellation-based strategies that focus on the SC among different brain regions of interest (ROIs) (Siless et al., 2020; Zhang et al., 2018; O'Donnell et al., 2013; de Reus and van den Heuvel, 2013; Zalesky et al., 2010). Compared to voxel-based methods, tract-based methods can group the fibers into anatomically meaningful fiber bundles and visualize them. However, the drawback of this approach is the reliance either on the choice of clustering method and that of the similarity metric for comparing streamlines (Zhang et al., 2014) or the choice of parcellation atlas. The hybrid method such as population-based structural connectome (PSC) (Zhang et al., 2018) can utilize the geometric information of fibers, including shape and location, to increase the robustness of extracted WM bundles between two ROIs and extract discriminative and reproducible geometric features for parcellation-based connectome analysis.

In this paper, we propose a new hybrid method lifespan population-based structural connectome (LPSC). The goal of LPSC is to create a parcellation-based tractographic skeleton (PBTS) atlas for accurately mapping dMRI and structural MRI data acquired from n LHCS subjects with multiple scans measured at different times on the atlas across the lifespan. The reliable construction of such atlas allows us to quantify the development of WM within individual subjects across time and understand its variations across groups. Figure 4.1 shows the overview of the LPSC framework, which consists of three key stages. The Innovations of LPSC are summarized as follows. **(i)** The LPSC can be represented as a major improvement of TBSS and all other tract-based approaches, including our prior PSC, while inheriting the robustness of TBSS. Specifically, we use LHCP-Y to train PBTS, so PBTS has much higher signal-to-noise ratio than TBSS. **(ii)** The LPSC as a tract-based atlas is directly applicable to dMRI obtained from various neuroimaging studies even without invoking fiber tracking. Thus, it is extremely valuable for many existing non-HCP studies, which may not have high-quality tractographic data. **(iii)** The LPSC is to build the weighted graph structure to sparsely represent high-dimensional tractographic data. Since the graph structure is built through Cole-Anticevic Brain-wide Network Partition (CAB-NP) (Ji et al., 2019), our LPSC enjoys stronger functional and structural organization information than existing tract-based approaches.

4.2 Materials and Methods

4.2.1 Overview

The proposed LPSC framework has three main steps, as shown in Figure 4.1. These are (i) whole-brain structural connectome construction; (ii) creation of fiber skeleton; and (iii) Sparse representation of tractographic data. In Sections 4.2.2-4.2.4, we introduce each of

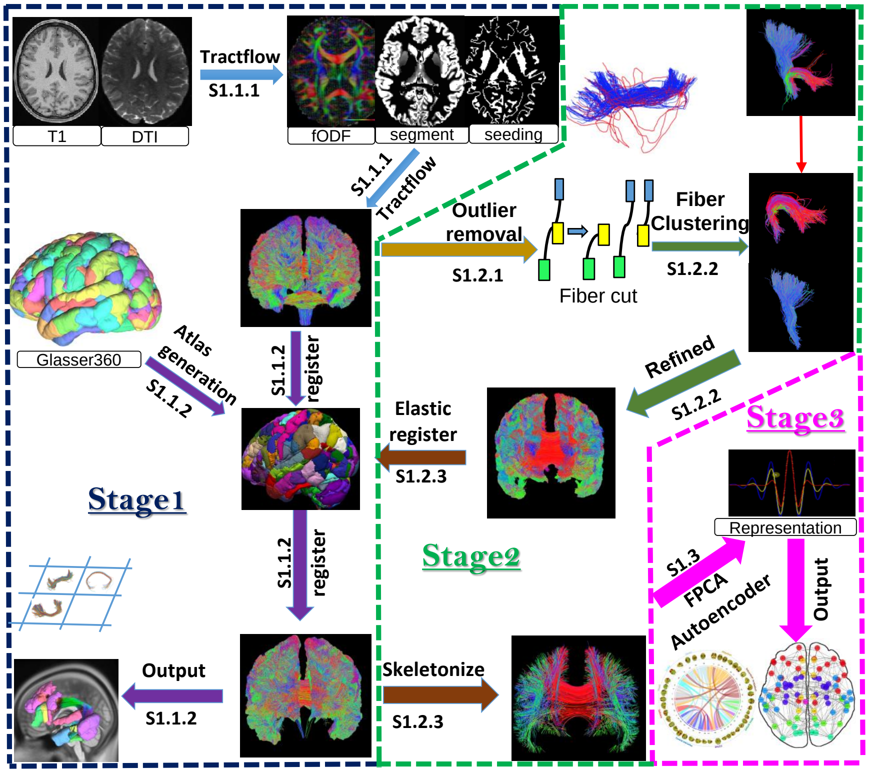


Figure 4.1: The workflow of LPSC.

these modules in detail. Section 4.2.5 describes HCP and its test-retest dataset.

4.2.2 Whole-brain structural connectome construction

The first stage is to construct the whole-brain trajectory data set and map an initial structural connectome based on two key steps, including (S1.1.1) the construction of high-quality whole-brain tractographic data; and (S1.1.2) the robust brain parcellation and connectome extraction..

Whole-brain tractography (S1.1.1) is the first key step of our LPSC framework, which is to reliably reconstruct the whole-brain tractographic data for all scans of all subjects through state-of-the-art tractography algorithms. We use the diffusion MRI tractography processing pipeline, called Tractoflow (Theaud et al., 2020), and then further improve the precision of tractography under the cortex by incorporating brain cortical surface boundary information (St-Onge et al., 2018). Specifically, the Tractoflow pipeline first processes the DWI, which includes raw DWI preprocessing (denoising, eddy correction, brain extraction, N4 bias correction, cropping, intensity normalization and resampling), diffusion tensor imaging (DTI) metrics and fiber orientation distribution function(Descoteaux et al., 2007b; Tournier et al., 2007) (fODF) metrics computation. Then it processes the T1 image, which includes raw T1 preprocessing (denoising, N4 bias correction, resampling, brain extraction, and cropping), registration to the DWI space, the tissue segmentation and the computation of tracking masks and seeding maps. Finally, Tractoflow utilizes the the tractography algorithm presented by Girard et al. (2014) to generate a set of fibers that represents the underlying WM architecture. This tracking algorithm generates the tractogram using the probabilistic method based on fODF image and seeding mask. To avoid the fiber streamlines stop in the middle of the WM without reaching the gray matter, we utilize the surface-enhanced tractography (SET) method (St-Onge et al., 2018) to ensure the fiber curves connect cortical regions intersect with cortical surfaces, such that any cortical parcellation can be applied to get a connectivity adjacency matrix and we can further improve such cortical parcellation. We discard the fibers which have less than three points or only one of the endpoints intersect with cortical surface and we also filter the fiber streamlines such that the fiber length is between 10 and 250 mm. The output of (**S1**.1.**1**) is a whole-brain tractographic data set, denoted as $\mathcal{F}_{it}$, and the associated diffusion profiles (e.g., fractional anistropy), denoted as $\mathcal{P}_{it}$ for the i -th subject at t -th time. The set $\mathcal{F}_{it}$ consists of N_{it} three-dimensional curves such that $\mathcal{F}_{it} = \{F_{it,1}(\cdot), \ldots, F_{it,N_{it}}(\cdot)\}$ and $F_{it,k}$ is represented as a sequence of 3D points $p_{(it,k),j} = \left(x_{(it,k),j}, y_{(it,k),j}, z_{(it,k),j}\right) \in R^3$ for $j = 1, \ldots, m_{(it,k)}$ and $k = 1, \ldots, N_{it}$. Moreover, $\mathcal{P}_{it} = \{P_{it,1}(\cdot), \ldots, P_{it,N_i}(\cdot)\}$ also consists of N_{it} 1D functions and $P_{it,k}(s)$ is observed along the fiber tract $F_{it,k}(s)$.

In our analysis of real data, the tractography dataset came up with approximately 1M fiber streamlines and and each streamline has a step size of 0.2 mm for each individual in the HCP dataset (with isotropic voxel size of 1.25 mm).

Brain parcellation and connectome extraction. In (**S1**.1.**2**), we use a state-of-the-art brain atlas with known parcellation to extract SC across subjects in three steps. (i) We use the Cole-Anticevic Brain-wide Network Partition (CAB-NP) (Ji et al., 2019) consisting of 360 cortical parcels (Glasser et al., 2016) to define the nodes of the SC network, enabling treating brain as a single network. These 360 parcels belong to 12 brain functional networks, such as visual, language, and default. The CAB-NP parcellation improves neuroanatomical precision for studying the structural and functional organization of human brain. (ii) We register all individual brains and CAB-NP to the common MNI space using Freesurfer (Fischl, 2012) and apply the transformation to tractographic data for each subject. We then map the ROI labels of CAB-NP to the points on individual cortical surface. Since all the generated fiber curves connect cortical regions intersect with cortical surfaces, we extract the endpoints of the fibers and snap the endpoints of each fiber to the nearest points on the individual cortical surface by minimizing the $\mathbb{L}_2$ distance such that we can assign the ROI labels of the points on cortical surface to the fiber endpoints. (iii) We group fibers based on the connections of their endpoints with different ROIs after parcellation and we calculate the fiber counts between different ROI pairs for each subject. The output of Stage 1 is a parcellation of $\{\mathcal{F}_{it}\}$ into tractographs that connect each ROI pair.

4.2.3 Creation of fiber skeleton

Our second stage is to compute a parcellation-based fiber skeleton. Its three key steps include (S1.2.1) outlier filtering; (S1.2.2) refined brain connectome; and (S1.2.3) elastic fiber registration.

Outlier filtering In (**S1.2.1**), we identify and remove fibers that do not follow major WM pathways as outliers in order to estimate the fiber bundles and establish a reliable network among all ROI pairs, since almost all tractographic algorithms produce false fiber tracts due to the intrinsic limitations of the existing technology, such as low resolution of dMRI and inaccurate choices of tracking parameters. (Schilling et al., 2019). Similar with Zhang et al. (2018), we use the Quickbundle method based on the direct-flip (MDF) distance (Garyfallidis et al., 2010, 2012) to detect outlier fibers. Specifically, for each fiber $s = [s_1, s_2, \cdots, s_K]$ and its flipped version $s^F = [s_K, s_{K-1}, \cdots, s_1]$, the MDF distance is defined as follows,

$$d_{\text{direct}}(s,t) = d(s,t) = \frac{1}{K}\sum_{i=1}^{K} \|s_i - t_i\|,$$

$$d_{\text{flipped}}(s,t) = d\left(s, t^F\right) = d\left(s^F, t\right),$$

$$\mathrm{MDF}(s,t) = \min\left(d_{\text{direct}}(s,t), d_{\text{flipped}}(s,t)\right),$$

where $\|\cdot\|$ is the $\mathbb{L}_2$ distance and we downsample the number of points $K = 100$ on each streamline. MDF distance is very fast and easy to compute and also considers the fiber orientation issues. After calculating the MDF distance of each pair of the fibers, the clustering process, similar to k-means (MacQueen et al., 1967), is performed. A fiber is assigned to a cluster if its MDF distance to the cluster centroid is smaller than a pre-selected parameter θ_t and then the cluster centroid is also updated correspondingly. Finally, Quickbundle method groups all the fibers into several fiber clusters and the outliers are then defined as the clusters with very few fibers inside. The output of (S1.2.1) is a set of outliers-removed fiber bundles across all the possible ROI pairs for each scan.

Refined brain connectome In (**S1.2.2**), we propose to refine the brain connectome. Basically, there are three key steps: (i) merge ROIs with similar connectivity pattern; We first extract the fiber count vector of each ROI and calculate the mean correlation of the fiber count vector between each ROI pair across all the subjects. We merge fiber tracks in $(R_v, R_{v'})$ together, where R_v and R_{v_1} are the ROIs which have large mean correlation, belong to the same functional network and are spatially neighboring. In this way, we can also preserve the functional network structure. These fiber curves and their endpoints are used to refine the initial parcellation. (ii) remove ROI pair with low reproducibility; With the merged ROIs, we quantify the at the reproducibility on test-retest dataset and discard the fiber curves in the ROI pair with low reproducibility. (iii) cluster fiber curves within ROI pair; We further refine the brain connectivity matrix across subjects by cluster fiber curves within ROI pair. That is, we utilize the shape features of fiber curves to refine our initial brain atlas to increase the discriminative power of the extracted structural connectome. As an illustration, we choose a specific ROI pair, denoted as $(R_v, R_{v'})$. If the number of fiber tracks connecting R_v and $R_{v'}$ is large, then it is important to identify these pathways across subjects and scans and refine the parcellation such that each ROI pair only contains one major fiber pathway. We use the Tract Dictionary Learning (TractDL) as described in Wu et al. (2020) to cluster the fibers. Specifically, each fiber is mapped from the native point space to a Hilbert space and then represented by the linear combination of a series of the cosine basis functions(Chung et al., 2010). This representation only depends on the degrees of the cosine basis functions and robust to the number of sampling points along the fiber. Tract Dictionary Learning(Kumar et al., 2019b) is further performed based on the cosine coefficients to learn the dictionary for each bundle. The $\mathbb{L}_1$ penalty is added to enforce each dictionary has sparse coefficients. Finally, each fiber is classified to a fiber bundle by minimizing the distance to each bundle dictionary.

Elastic fiber registration In (**S1.2.3**), we will use a group-wise elastic registration method to register fiber bundles within each ROI pair across subjects. Although some registration tools are available for fibers Durrleman et al. (2011); Zvitia et al. (2010); Garyfallidis et al. (2015); Jin et al. (2014); Wassermann et al. (2011); Zhang et al. (2018); O'Donnell et al. (2012), they have major limitations (e.g., pinching effects). Please see discussions in pages 84-88 of Srivastava and Klassen (2016). Instead, we propose to use a square-root velocity (SRV) representation for analyzing fiber shapes in the Euclidean space under an elastic metric Srivastava and Klassen (2016). The elastic metric is developed specifically for the analysis of curve shapes. Specifically, due to this SRV representation, a complicated elastic metric simplifies to the $\mathbb{L}_2$ metric, the re-parameterization group acts by isometries, and the space of unit length curves becomes the unit sphere. The use of the elastic metric seeks the best correspondence between points to become correspondence-free. Finally, we use the elastic metric to calculate the mean fiber and establish the correspondence between fibers for each ROI pair across all subjects. Without loss of generality, for each specific pair $(R_v, R_{v'})$, we will use its mean fiber, denoted as $S_{(v,v')}$, as its skeleton. **The final output of Stage 2 is a parcellation-based tractographic skeleton, abbreviated as PBTS, which is given by PBTS** $= \cup_{v,v'=1}^{V} \left\{(R_v, R_{v'}) \oplus S_{(v,v')}\right\}$, **where** V **is the total number of ROIs.**

4.2.4 Sparse representation of tractographic data

Our third stage is to sparsely represent $\mathcal{F}_{it}$. This step consists of both basis learning and efficient representation based on a matrix of basis functions across all pairs of ROIs across time. After alignment, all fiber tracts connecting a specific ROI pair across subjects and time share many shape similarities. Without loss of generality, we focus on a ROI pair, $(R_v, R_{v'})$. We perform functional principal component analysis (fPCA) on each of the x, y and z components of all aligned fiber tracts, separately, in order to learn a set of basis functions denoted as $\mathcal{L}_{(v,v')} = \left\{\phi_{(x,y,z,t),m} : m = 1, \ldots, M\right\}$, where M is the number of fPCA basis and

$\phi_{(x,y,z,t),m}$ are the fPCA basis functions. Each fiber curve connecting $(R_v, R_{v'})$ at time t can be sparsely represented as a linear combination of the basis functions in $\mathcal{L}_{(v,v')}$. **The output of this stage is the graph of fPCA basis functions across all ROI pairs, denoted as GBF**$= \cup_{v,v'=1}^{V} \left\{ (R_v, R_{v'}) \oplus \mathcal{L}_{(v,v')} \right\}$**.**

4.2.5 Real Datasets

The dMRI data of HCP dataset is accquired using HARDI. A full dMRI session includes 6 runs (each approximately 9 minutes and 50 seconds), representing 3 different gradient tables, with each table acquired once with right-to-left and left-to-right phase encoding polarities, respectively. Each gradient table includes approximately 90 diffusion weighting directions plus 6 b=0 acquisitions interspersed throughout each run. There are three shells ($b = 1000, 2000$, and 3000 s/mm^2) with an approximately equal number of acquisitions on each shell for each run. More details regarding data acquisition and preprocessing can be found at Van Essen et al. (2012) and Sotiropoulos et al. (2013). In our stage (S1.1.1), all the three shells are used for fODF computation and only the $b = 1000s/mm^2$ is used for DTI metrics calculation. We extracted 856 subjects with both raw dMRI and anatomical T1-weighted MRI data from the 1206-subject release of the HCP dataset. Among the extracted 856 subjects, 43 subjects have one repeated acquisitions, with an approximate two-week interval between two acquisitions. Therefore, the test-retest dataset of HCP is composed of 86 acquisitions.

4.3 Experimental Results

In this section, we evaluate the following three aspects of our LPSC framework.

(I) Choice of optimal parameters in LPSC: There are several parameters in LPSC that play an important role for generating reproducible and reliable results. We quantify reproducibility on test-retest dataset to select these parameters.

(II) Validation of reproducibility: We validate the robustness and reproducibility of the connectome results generated from LPSC.

(III) Demonstration of groupwise analysis: We extract the FA maps of the HCP data and

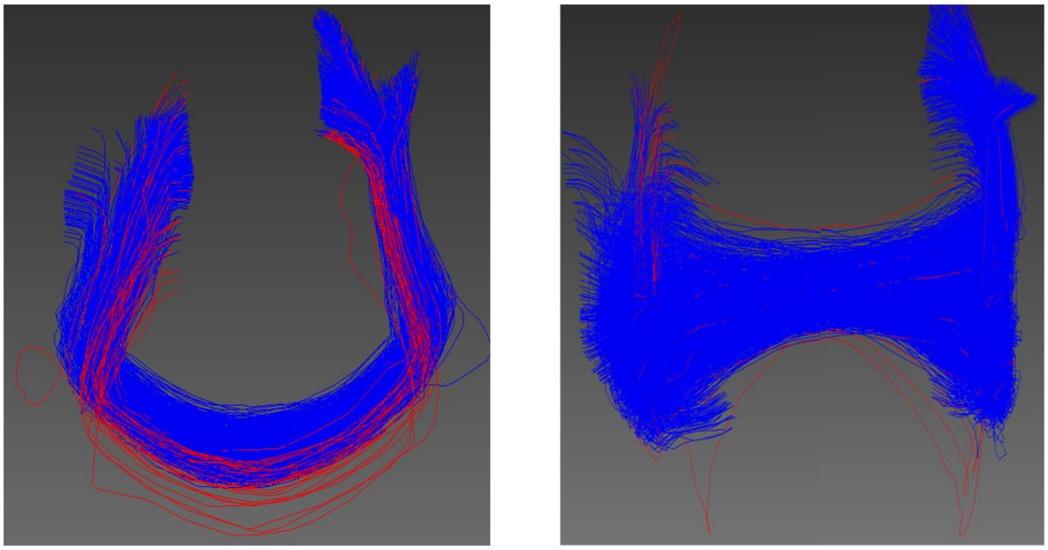


Figure 4.2: Outlying streamlines (red) and major WM pathways (blue) for two selected ROI pairs

project them on our PBTS. We use these data to demonstrate the application of our PBTS in delineating the association between WM diffusion properties with a set of covariates of interest.

4.3.1 Choice of optimal parameters in LPSC

There are several parameters in LPSC that play an important role for generating reproducible and reliable results. We quantify reproducibility on test-retest dataset to select these parameters.

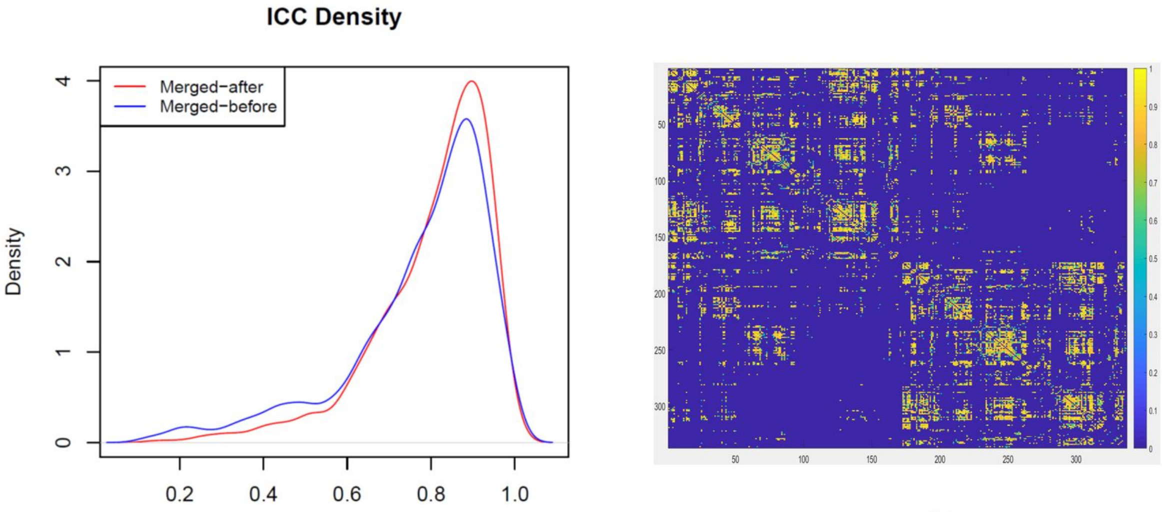
Outlier threshold The clustering threshold θ_t in QuickBundle affects the number of the outliers detected. Based on our previous findings (Zhang et al., 2018), QuickBundle can hardly detect any outliers for $\theta_t > 10$ mm, while QuickBundle may detect too many outliers for $\theta_t < 5$. Previously, we set $\theta_t = 8$ mm to remove the apparent outlying streamlines that do not follow any major WM pathways. At this time, we more conservatively set $\theta_t = 5$ mm to make the skeleton cleaner. Figure 4.2 displays some apparent outlying streamlines in red for two selected ROI pairs.

Parameters in refined brain connectome In the first step of (**S1.2.2**), we calculate the mean correlation of the fiber count vector between each ROI pair across all the subjects in test-retest dataset and then merge 23 ROI pairs that have mean correlation larger than 0.7, belong to the same functional network and are spatially neighboring. In this case, dimension of the connectivity matrix is reduced from $360 * 360$ to $337 * 337$. To justify that merging ROIs can improve the reproducibility, we calculate intraclass correlation coefficient (ICC) (Prckovska et al., 2016; Welton et al., 2015) at the test-retest dataset, which is defined as

$$\mathrm{ICC} = \frac{1}{Ns^2} \sum_{n=1}^{N} \left(x_{n,1} - \bar{x}\right) \left(x_{n,2} - \bar{x}\right)$$

where $x_{n,1}$ and $x_{n,2}$ are the fiber counts for the nth subject in test and retest dataset respectively, $\bar{x}$ and s^2 are the mean and stand deviation of fiber counts in the whole test-retest dataset. Figure 4.3 (a) shows the density of ICCs calculated by the merged-before and merged-after ROIs and we can see the pattern that ICCs are improved after merging the ROIs. We further perform two-sample t-test and mann-whitney test and conclude that mean ICC is improved significantly by both of the two tests ($p < 10^{-4}$). In the second step of (**S1.2.2**), we recalculate ICCs on the merged ROIs. Since ICC< 0.4 represents poor reproducibility(Cicchetti, 1994), we discard these ROI pairs. We also remove the ROIs with the percentage of the missingness at the test-retest dataset larger than 50% to ensure the connectivity between any two ROIs exist for the majority of the subjects. Figure 4.3 (b) shows displays the ICC matrix calculated by the merged-after ROIs. In the third step of (**S1.2.2**), we cluster fiber curves within the ROI pair that has mean fiber count larger than 500 in test-retest dataset. The number of cluster K is determined by maximizing Calinski-Harabasz index, which is defined as

$$\frac{n-K}{K-1} \frac{\sum_{k=1}^{K} \sum_{i \in I_k} \left\| s_i^{\{k\}} - c^{\{k\}} \right\|^2}{\sum_{k=1}^{K} n_k \left\| c^{\{k\}} - c \right\|^2},$$



(a) (b)

Figure 4.3: ICC of fiber counts.(a)density of ICCs calculated by the merged-before and merged-after ROIs; (b)ICC matrix calculated by the merged-after ROIs.

where $s_i^{\{k\}}$ is the i-th streamline in the k-th cluster I_k, $c^{\{k\}}$ is the mean streamline in I_k, c is the the mean streamline in the whole ROI pair, n_k is the number of streamlines in I_k and $n = \sum_{k=1}^{k} n_k$. Clustering results for the ROI pairs left Area 6mp and right Primary Motor Cortex, and left Primary Motor Cortex and left Area 1/3a are summarized in Figure 4.4 (a) and (b) respectively. Each fiber cluster is remarked with a distinct color.

4.3.2 Reproducibility of connectomes generated by LPSC

A good LPSC framework should generate very similar connectome results based on the scans acquired within a short time period from the same subject. In this section, we evaluate and validate the reproducibility of structural connectomes at both network and streamline levels on the test-retest dataset. All the parameters are set exactly the same as discussed in section 4.3.1 to process the test-retest dataset.

Reproducibility of connectomes at network level We first consider the reproducibility at the binary network level. We extract the $337 * 337$ fiber count matrix $A = (a_{i,j})$ and generate a binary network matrix $B = (b_{i,j})$ for each 86 scans in test-retest dataset, where

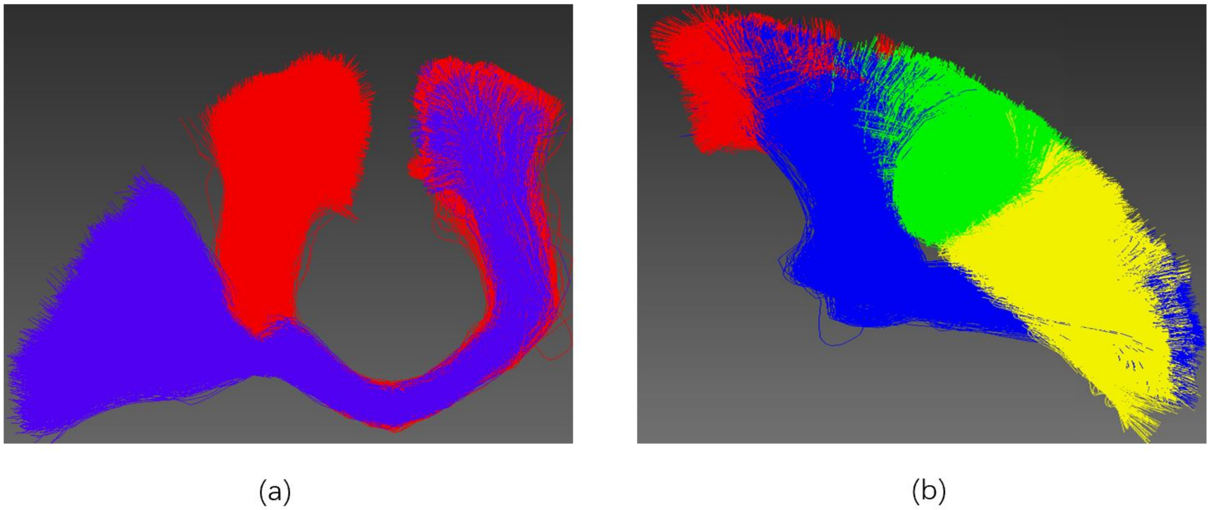

(a) (b)

Figure 4.4: Clustering results for two selected ROI pairs: (a)Left Area 6mp and right Primary Motor Cortex; (b)Left Primary Motor Cortex and left Area 1/3a.

$b_{i,j} = 1$ if $b_{i,j} \geq \theta_{bin}$ and $b_{i,j} = 0$ otherwise. Here we simply choose $\theta_{bin} = 20$ as an example and more details regarding how to choose the optimal θ_{bin} can be found in Zhang et al. (2018). From Figure 4.5 (a), we can see that there is only small amount of the non-zero element in the difference of the binary network matrix of the two scans from the same subject, whereas the amount is larger in the the difference of the binary network matrix of two different subjects. Figure 4.5 (b) displays the pairwise distance matrices between the 86 binary network matrices. The distance is calculated as the sum of the $\mathbb{L}_2$ distance of the corresponding element between each two matrices. To quantify the reproducibility at the network (matrix) level, we further calculate the distance-based ICC (dICC)(Zhang et al., 2018). dICC can be viewed as an extension of ICC to the multi-variate case and is defined as

$$\mathrm{dICC} = (\bar{d}^2_{bs} - \bar{d}^2_{ws})/\bar{d}^2_{bs},$$

where $\bar{d}^2_{bs}$ and $\bar{d}^2_{ws}$ respectively represent the average squared distance between subjects and within multiple scans of a subject. The subject-related block diagonal pattern in Figure 4.5 (b) and the dICC around 0.64 indicate the strong reproducibility of our connectome results.

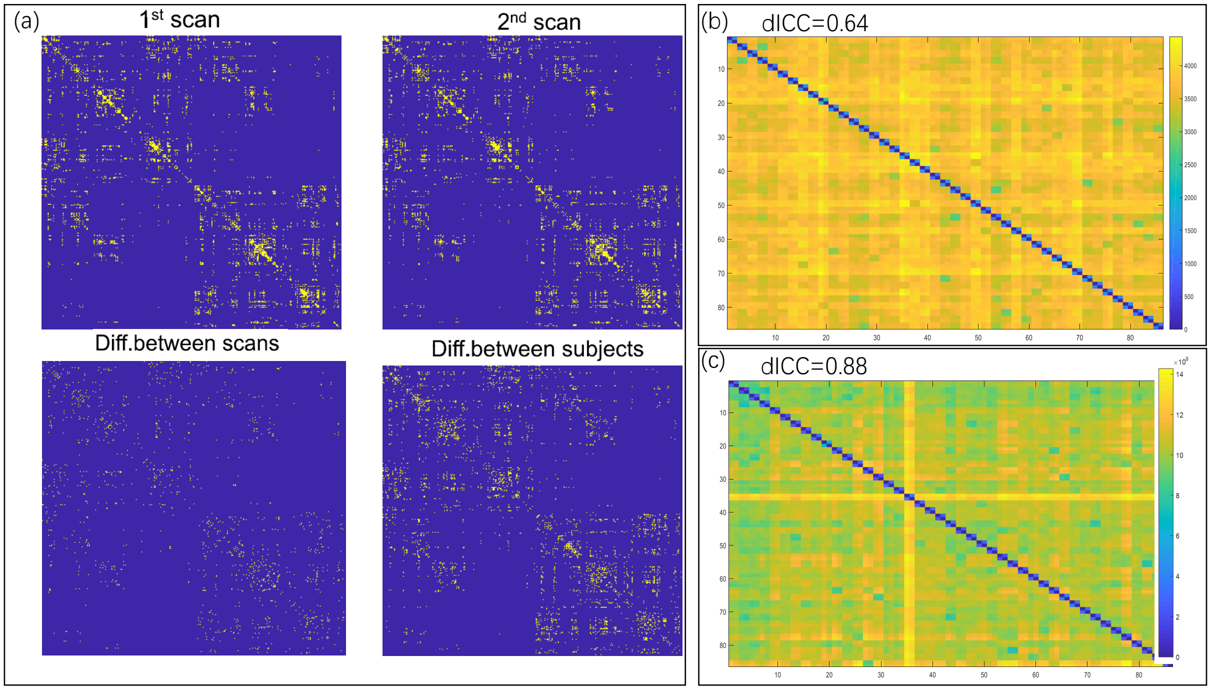


Figure 4.5: Reproducibility analysis at network level. (a) Binary network matrices difference from two different scans of the same subject and the difference from two different subjects; (b) Pairwise distance matrix between 86 binary network extracted from the test-retest dataset; (c) Pairwise distance matrix between 86 weighted network extracted from the test-retest dataset

Furthermore, we can observe the similar pattern for the distance matrices of weighted network (i.e. regular fiber count matrix) in Figure 4.5 (c).

Reproducibility of connectomes at streamline level At the streamline level, we extract the FA values of each streamline for each scan and calculate the mean FA curves for each ROI pair. We use the $\mathbb{L}_2$ distance between the mean FA curves to further calculate the dICC score at test-retest dataset. We summarized the results in Figure 4.6.

Figure 4.6 (a) displays the streamlines that connect left and right Supplementary and Cingulate Eye Field (SCEF) regions and the corresponding FA values and mean FA curve along them from the two scans of two randomly selected subjects in test-retest dataset. These streamlines belong to part of the corpus callosum bundle. We find that the pattern of

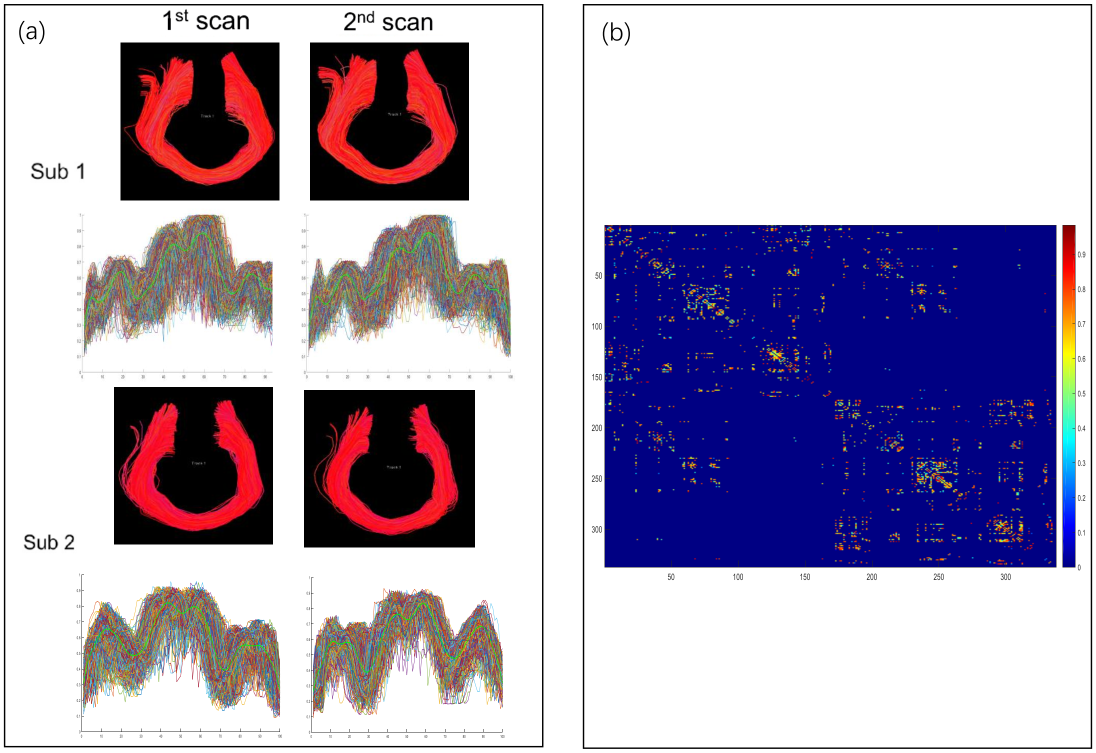


Figure 4.6: Reproducibility analysis at streamline level. (a) Streamlines that connect left and right SCEF regions extracted from two subjects in test-retest dataset. FA values and mean curve (green line) are also plotted; (b) dICC score matrix based on mean FA curves.

FA values and streamlines is different across subjects but is consistent within the different scans from the same subjects. Figure 4.6 (b) presents the dICC calculated on the mean FA curves. We can observe most of the ROI pairs have dICC> 0.6, which indicates the strong reproducibility at streamline level.

4.3.3 Groupwise analysis

In this section, we illustrate the application of our LPSC for groupwise analysis in HCP dataset. We extract the FA maps for each subjects and study the association between FA values along the PBTS and several covariates of interest.

We first project the individual FA map on our PBTS. Specifically, we merge the streamlines of each ROI pair across subjects and randomly select 1000 streamlines from each ROI pair to establish a parcellation-based tractographic coordinate system (PBTCS). For illustration,

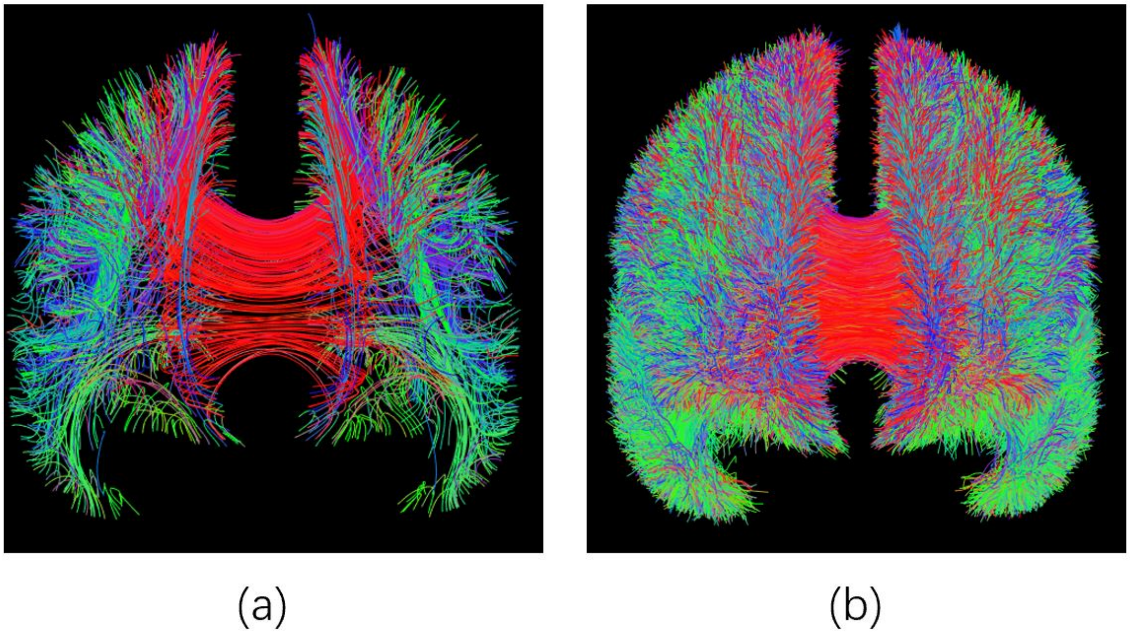


Figure 4.7: The skeleton and coordinate system used for projection of individual FA map: (a) parcellation-based tractographic skeleton (PBTS); (b) parcellation-based tractographic coordinate system (PBTSC) .

we choose a specific ROI pair $(R_v, R_{v'})$ as an example. We denote the i th streamline $t_i = [t_i^1, t_i^2, \cdots, t_i^{100}]$ selected from the ROI pair $(R_v, R_{v'})$, $i = 1, \cdots, 1000$ and the skeleton of this ROI pair as $S_{(v,v')} = [s^1_{(v,v')}, s^2_{(v,v')}, \cdots, s^{100}_{(v,v')}]$. Then the projected individual FA map $F_{(v,v')} = [f^1_{(v,v')}, f^2_{(v,v')}, \cdots, f^{100}_{(v,v')}]$ is defined as

$$f^k_{(v,v')} = \sum_{i=1}^{1000} w_{i,k} f_i^k,$$

where $k = 1, \cdots, 100$, f_i^k is the FA value of the individual map at point t_i^k. Moreover, $w_{i,k}$ is given by

$$w_{i,k} = \frac{g_h(\|s^k_{(v,v')} - t_i^k\|))}{\sum_{k=1}^{1000} g_h(\|s^k_{(v,v')} - t_i^k\|)},$$

where g_h is the gaussian kernel function with bandwidth h. Figure 4.7 displays PBTS and PBTCS.

FA values along the streamline can be modeled as functions of the coordinates on the streamline. Therefore, we use the Functional analysis of diffusion tensor tract statistics (FADTTS)(Zhu et al., 2011) pipeline to characterize the structure of the variability along the streamline. Specifically, FADTTS fits a multivariate varying coefficient model as follows

$$y_i(s) = \mathbf{x}_i^T \beta(s) + \eta_i(s) + \varepsilon_i(s),$$

where $y_i(s)$ is the FA value within a given ROI pair at position/arc-length s for the i-th subject. Age, gender, intelligence score (IS), BMI, max drink (Max drinks in a single day in past 12 months), and intercept term are added into $x_i \in \mathbb{R}^6$ and $\beta(s) = [\beta_1(s), \cdots, \beta_6(s)]^T \in \mathbb{R}^6$ is a vector of six functional covariate coefficients. $\eta_i(s)$ and $\varepsilon_i(s)$ are the stochastic processes with zero mean function to characterize the inter-subject variability and error structure respectively. In our analysis, sample size is $n = 977$ (57 subjects are removed from the total 1034 subjects due to missing covariates) and the FA values are chosen from two ROI pairs: left and right SCEF, and left and right Superio Frontal Language Area (SFL). These two ROI pairs both belong to part of the corpus callosum bundle.

Figure 4.8 displays the FA curves and the estimated coefficient function of each covariate for the two ROI pairs. We can see that the coefficient function of gender have potential positive effect and the coefficient function of BMI have potential negative effect. To test the significance of these effects, we perform the following global test

$$H_0 : \beta_i(s) = 0 \;\text{ for all } s \text{ v.s. } \quad H_1 : \beta_i(s) \neq 0, \tag{4.3.1}$$

for $i = 2, \cdots 6$. The test statistic has been shown has asymptotic weighted χ^2 distribution and the pvalue p is calculated by wild bootstrap method(Zhu et al., 2012). The coefficient function of gender has $p < 10^{-3}$ and the coefficient function of BMI has $p = 0.02$. Other coefficient functions do not have significant pvalues.

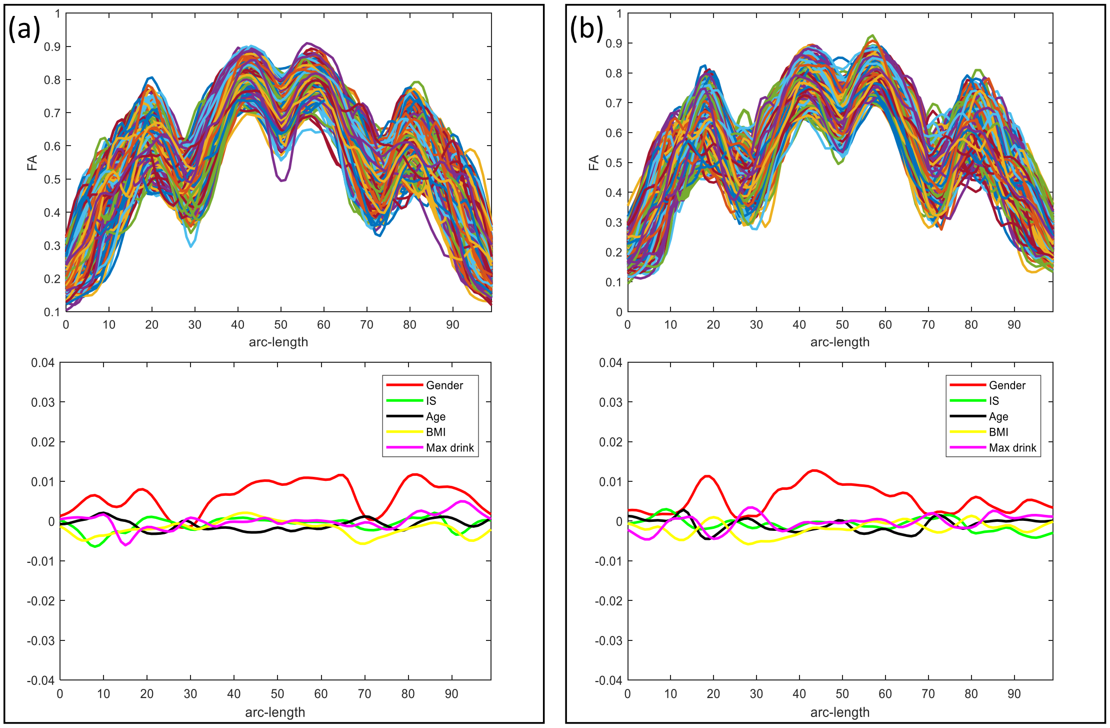


Figure 4.8: FA curves and the estimated coefficient function for the two ROI pairs: (a)Left and right SCEF; (b)Left and right SFL.

4.4 Discussion

We have developed a LPSC framework and created a PBTS atlas for mapping population structural connectivity networks from large scale neuroimaging dataset with multiple scans measured at different times across the lifespan. LPSC can be represented as a major improvement of TBSS and all other tract-based approaches with higher signal-to-noise ratio and stronger functional and structural organization information, while inheriting the robustness of TBSS. LPSC also enables us to study brain structural connectivity at network level and explore the shape information of white matter at streamline level for the high-resolution data such as HCP. Furthermore, PBTS as a tract-based atlas is directly applicable to dMRI obtained from various neuroimaging studies even without invoking fiber tracking. Thus, it is extremely valuable for many existing non-HCP studies, which may not have high-quality tractographic data. We have also applied LPSC on the HCP test-retest to illustrate its reproducibility.

Although we have demonstrated the use of LPSC, one limitation needs to be addressed in the future research. The CAB-NP atlas we used for brain parcellation does not consider the subcortical regions. To explore the connectivity in subcortical region, the subcortical parcellation such as woGSR(Ji et al., 2019) can be utilized. Besides, since the streamlines in subcortical regions do not intersect with surfaces, other volume-based method maybe required to cut the subcortical streamlines.

CHAPTER 5: DECONFOUNDING BY EIGEN-SHRINKAGE IN HIGH-DIMENSIONS

5.1 Introduction

5.1.1 Background

Large observational dataset emerge rapidly in many biomedical research fields. Associated large-scale statistical inference often focuses on the mean profile and variance-covariance structure of data matrix, especially on testing and estimating the mean difference of numerous collected markers (e.g., image voxels, genetic variants) for one variable of interest (e.g., disease status, age, treatment group). In many research fields, the practical tools to do such inference are massive univariate models, where regressions are separately fitted on thousands or even millions of markers, with the goal to prioritize and filter out important ones. For example, a genome-wide association study (GWAS) is often performed to detect the association between a phenotype (e.g., disease/trait) and a large number of single nucleotide polymorphisms (SNPs) collected across an entire genome. The common approach in GWAS is to assess the marginal association between phenotype and single SNP each at a time, while adjusting for the same set of covariates. Similar single-feature univariate procedures are widely applied in epigenome-wide association studies (EWAS) on DNA methylation data (Teschendorff and Relton, 2018), transcriptome-wide association studies (TWAS) on gene expression data (Pasaniuc and Price, 2017), and various voxel-wise analysis on neuroimaging (Smith and Nichols, 2018). In what follows, we refer to the dependent variable in a univariate model as *feature* and the independent variable of interest as *primary variable.*

Unmeasured hidden confounders, however, are well known challenges in such observational studies. Here hidden confounders are defined as the unmeasured factors that are correlated with both primary variable and features. For example, in high-throughout biomedical experiments, a large number of markers are produced by a series operation involving complicated reagents, hardware, software and well-trained lab members (Leek et al., 2010). The variability of these factors within or among studies can often cause unwanted variation of the features. When these factors happen to be correlated with the primary variable, they will become hidden confounders. Famous examples include ancestral history in GWAS (Price et al., 2006), batch effects in microarray experiments caused by non-biological factors such as lab members and collection times (Benito et al., 2004), and the head motion, breathing rate/depth change, blood pressure, and scanner in neuroimaging (Smith and Nichols, 2018).

5.1.2 Low-rank models

Low-rank models are powerful tools in large-scale inference. In this paper, we consider the following high-dimensional low-rank model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{Z}\boldsymbol{D} + \boldsymbol{E}, \tag{5.1.1}$$

where $\boldsymbol{Y} \in \mathbb{R}^{n\times p}$ is a matrix of p collected features from n samples, $\boldsymbol{X} \in \mathbb{R}^{n\times q}$ contains the primary variable and potentially other observed covariates, $\boldsymbol{Z} \in \mathbb{R}^{n\times k}$ stands for k hidden confounding factors, $\boldsymbol{E} \in \mathbb{R}^{n\times p}$ is a noise matrix, and $\boldsymbol{B} \in \mathbb{R}^{q\times p}$ and $\boldsymbol{D} \in \mathbb{R}^{r\times p}$ are regression coefficients. Testing and estimating $\boldsymbol{B}$ is the primary goal in statistical inference. Here we ignore the difference between primary variable and other covariates, and assume both $\boldsymbol{X}$ and $\boldsymbol{Z}$ are full rank with $\max(k, q) < n$. In practice, $\boldsymbol{Z}$ is unmeasured and thus omitted from model (5.1.1), leading to the following misspecified regression model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B}^{*} + \boldsymbol{E}^{*}. \tag{5.1.2}$$

When $\boldsymbol{X}$ and $\boldsymbol{Z}$ are independent (i.e., $\boldsymbol{Z}$ is not confounder but some unobserved factors), ordinary least squares (OLS) estimator of $\boldsymbol{B}^{*}$ remains as an unbiased estimator of $\boldsymbol{B}$. $\boldsymbol{Z}\boldsymbol{D}$ is absorbed into $\boldsymbol{E}^{*}$ and thus results in unwanted variance-covariance structure in test statistics of $\boldsymbol{B}^{*}$. Addressing the unwanted dependency of test statistics has been an active research area for large-scale inference. For example, the false discovery rate (FDR) control procedures designed for independent tests (Benjamini and Hochberg, 1995; Storey, 2002) have been extended to correlated tests with special dependency structure (weak or positive) (Benjamini and Yekutieli, 2001; Storey et al., 2003; Clarke et al., 2009), and to more generally correlated tests (Efron, 2010; Fan and Han, 2016; Friguet et al., 2009; Sun et al., 2015; Fan et al., 2017). When $\boldsymbol{X}$ and $\boldsymbol{Z}$ are dependent, the scenario might become more complex. Since $\boldsymbol{Z}$ confounds the relationship of $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\boldsymbol{B}^*$ is a mixing of effects in $\boldsymbol{B}$ and $\boldsymbol{D}$. The OLS estimator of $\boldsymbol{B}^*$ is a biased estimator of $\boldsymbol{B}$ and the test statistics of $\boldsymbol{B}^*$ are not only correlated but also confounded (Wang et al., 2017), resulting in inflated type I error and failure of FDR control. In this paper, $\boldsymbol{X}$ and $\boldsymbol{Z}$ are allowed to be highly correlated and we consider the situation in which $\boldsymbol{X}$ and $\boldsymbol{Z}$ are independent as a special case.

5.1.3 Existing methods

The research challenge to control for hidden confounders in gene expression studies is termed *Surrogate Variable Analysis* (SVA) by Leek and Storey (2007). The goal of SVA is to estimate a surrogate variable $\widehat{\boldsymbol{U}}$ of $\boldsymbol{Z}$ from the observed data $(\boldsymbol{Y}, \boldsymbol{X})$, then the large-scale inference on $\boldsymbol{B}$ can be performed with the OLS estimator from

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \widehat{\boldsymbol{U}}\widehat{\boldsymbol{V}} + \boldsymbol{E}.$$

To achieve validated inference of $\boldsymbol{B}$, the surrogate $\widehat{\boldsymbol{U}}$ needs to represent the column space of $\boldsymbol{Z}$, donated as $\boldsymbol{C}(\boldsymbol{Z})$. That is, SVA aims to recover $\boldsymbol{C}(\boldsymbol{Z})$ from $(\boldsymbol{Y}, \boldsymbol{X})$, where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are allowed to be highly correlated. Along this line, existing methods to estimate $\widehat{\boldsymbol{U}}$ include two-step SVA (2-SVA) (Leek and Storey, 2007), iteratively re-weighted SVA (IRW-SVA) (Leek and Storey, 2008), independent SVA (ISVA) (Teschendorff et al., 2011), remove unwanted variation (RUV) (Gagnon-Bartsch and Speed, 2012; Gagnon-Bartsch et al., 2017; Gerard and Stephens, 2017), latent effect adjustment after primary projection (LEAPP) (Sun et al., 2012; Wang et al., 2017; McKennan and Nicolae, 2018a,b), direct surrogate variable analysis (dSVA) (Lee et al., 2017), and latent factor mixed models (LFMM) (Caye and Francois, 2018; Frichot et al., 2013). These approaches are mainly proposed for large-scale multiple testing problem in genetics and omics data, but now are increasingly recognized in other fields, such as neuroimaging (Fortin et al., 2016; Guillaume et al., 2018), data integration (Fortin et al., 2018, 2017), and predictive modeling (Fromer et al., 2016; Parker et al., 2014). A detailed review and comparison of these methods is given in Section **??**.

5.1.4 Motivations

To recover $\boldsymbol{C}(\boldsymbol{Z})$ from feature matrix $\boldsymbol{Y}$, one need to carefully address the effects of $\boldsymbol{XB}$. It is notable that $\boldsymbol{C}(\boldsymbol{Z})$ can not be efficiently estimated from the residualized feature matrix $\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}^*}$ of model (5.1.2), because the OLS estimator $\widehat{\boldsymbol{B}^*}$ is biased when $\boldsymbol{C}(\boldsymbol{X})$ (the column space of $\boldsymbol{X}$) and $\boldsymbol{C}(\boldsymbol{Z})$ are not orthogonal. Since $\boldsymbol{X}$ is observed, feature projection/rotation constructed by $\boldsymbol{X}$ has been shown to be useful. For example, in LEAPP (Sun et al., 2012; Wang et al., 2017; McKennan and Nicolae, 2018a,b), Householder transformation is the key technique, which transforms feature matrix $\boldsymbol{Y}$ and separate it into two parts. The signals associated with $\boldsymbol{C}(\boldsymbol{X})$ (including those from $\boldsymbol{ZD}$) are concentrated to the first q rows of $\boldsymbol{Y}$, therefore $\boldsymbol{D}$ can be easily estimated from the left $n-q$ rows of $\boldsymbol{Y}$.

In this paper, we propose an eigen-shrinkage projection (ESP) method, which utilizes a new and simpler projection on feature matrix $\boldsymbol{Y}$. Consider the projection $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T$ on $\boldsymbol{Y}$ given by

$$\begin{aligned} \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y} &= \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \left(\boldsymbol{XB} + \boldsymbol{ZD} + \boldsymbol{E}\right) \\ &= \boldsymbol{D}_\gamma \boldsymbol{\Sigma}_{XB} \boldsymbol{R}_B^T + \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD} + \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{E} \\ &= \boldsymbol{Y}_1 + \boldsymbol{Y}_2 + \boldsymbol{Y}_3, \end{aligned}$$

where $\boldsymbol{XB} = \boldsymbol{Q}_X \begin{bmatrix} \boldsymbol{\Sigma}_{XB} \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{R}_B^T$ is the singular value decomposition (SVD) of $\boldsymbol{XB}$ with $\boldsymbol{\Sigma}_{XB} = \mathrm{Diag}(\sigma_1, \cdots, \sigma_q)$ are non-zero singular values, $\boldsymbol{Q}_X \in \mathbb{O}_{n,n}$ and $\boldsymbol{R}_B \in \mathbb{O}_{q,p}$ are left and right singular vectors, respectively, and

$$\boldsymbol{D}_\gamma = \mathrm{Diag}\big(\gamma_1, \gamma_2, \cdots, \gamma_q, 1, \cdots, 1\big) \in R^{n\times n}, \quad \gamma_i \in (0,1), \quad i = 1, \cdots, q.$$

We will show that under mind conditions, the projected data satisfies $\boldsymbol{Y}_2 >> \boldsymbol{Y}_1$ and $\boldsymbol{Y}_2 >> \boldsymbol{Y}_3$. Therefore,

$$\mathrm{svd}_k(\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}) \approx \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Z}\boldsymbol{D}$$

where $\mathrm{svd}_k(\boldsymbol{A}) = \boldsymbol{Q}_A \boldsymbol{\Sigma}_A^* \boldsymbol{R}_A^T$ is the rank k approximation of $\boldsymbol{A} = \boldsymbol{Q}_A \boldsymbol{\Sigma}_A \boldsymbol{R}_A^T$, in which $\boldsymbol{\Sigma}^*$ is generated by top k hard-thresholding of singular values in $\boldsymbol{\Sigma}_A$. Project the data back, we have the estimator for $\boldsymbol{Z}\boldsymbol{D}$

$$(\widehat{\boldsymbol{U}}\widehat{\boldsymbol{V}})_\gamma = \boldsymbol{Q}_X \boldsymbol{D}_\gamma^{-1} \mathrm{svd}_k(\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}). \tag{5.1.3}$$

In equation (5.1.3), $\boldsymbol{Y}$ is left-projected onto $\boldsymbol{Q}_X^T$ and then is shrunk by $\boldsymbol{D}_\gamma$. The top k left singular vector from the $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}$ is used to approximate the column space of $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Z}\boldsymbol{D}$. After left-projecting by $\boldsymbol{Q}_X \boldsymbol{D}_\gamma^{-1}$, $\boldsymbol{C}(\boldsymbol{Z})$ can be estimated by $\widehat{\boldsymbol{U}}$. Intuitively, shrinkage projection is not efficient because penalty may also eliminate the effects from $\boldsymbol{Z}$ since $\boldsymbol{C}(\boldsymbol{Z})$ and $\boldsymbol{C}(\boldsymbol{X})$ are not orthogonal. However, our theoretical analysis in later sections shows that as long as $\boldsymbol{C}(\boldsymbol{Z})$ and $\boldsymbol{C}(\boldsymbol{X})$ are not exactly the same, such loss is ignorable and the eigenvalue gap (eigen-gap) between $\boldsymbol{Z}\boldsymbol{D}$ and $\boldsymbol{E}$ is preserved.

Our ESP has natural connection with the regularized regression method recently proposed in LFMM (Caye and Francois, 2018). For example, When ridge penalty is posted on $\boldsymbol{B}$, LFMM-Ridge solves the following loss function

$$L_\lambda(\boldsymbol{U}\boldsymbol{V}, \boldsymbol{B}) = ||\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}^T - \boldsymbol{U}\boldsymbol{V}||_F^2 + \lambda||\boldsymbol{B}||_F^2, \quad \lambda \in (0, \infty),$$

where $||\boldsymbol{B}||_F = \mathrm{Tr}(\boldsymbol{B}\boldsymbol{B}^T)$ is the Frobenius norm of $\boldsymbol{B}$, and λ is a regularization parameter. The closed-form solution for $\widehat{\boldsymbol{U}}\widehat{\boldsymbol{V}}$ is

$$\begin{aligned}&(\widehat{\boldsymbol{U}}\widehat{\boldsymbol{V}})_\lambda = \boldsymbol{Q}_X \boldsymbol{D}_\lambda^{-1} \mathrm{svd}_k(\boldsymbol{D}_\lambda \boldsymbol{Q}_X^T \boldsymbol{Y}), \quad \text{and} \\ &\boldsymbol{D}_\lambda = \mathrm{Diag}\left(\sqrt{\frac{\lambda}{\lambda+\sigma_1^2}}, \cdots, \sqrt{\frac{\lambda}{\lambda+\sigma_q^2}}, 1, \cdots, 1\right).\end{aligned}$$

Thus, ridge penalized regression is a special case of shrinkage projection, where $\gamma_i \equiv \sqrt{\lambda/(\lambda+\sigma_i^2)}$, This close connection reveals that traditional ridge regression methods could be used for hidden confounder problem where $\boldsymbol{C}(\boldsymbol{X})$ and $\boldsymbol{C}(\boldsymbol{Z})$ are not orthogonal, because the signals from $\boldsymbol{Z}\boldsymbol{D}$ will not vanish after penalization. Particularly, though the OLS estimator $\widehat{\boldsymbol{B}^*}$ of model (5.1.2) does not work to recover $\boldsymbol{C}(\boldsymbol{Z})$, ridge estimator can solve this problem by adding L_2 penalty onto $\boldsymbol{B}^*$.

5.1.5 Structure and notations

We provide theoretical guarantees of ridge projection for hidden confounder problem under high-dimensional low-rank data settings. In Section 5.2, we show that penalized projection can eliminate the signals from primary variable while preserve the eigenvalue-gap (eigen-gap) between unmeasured factors and noises, which enables hidden confounder estimation from the projected data. Numerical experiments in Section 5.3 are used to illustrate our theoretical arguments and the finite sample performance of ridge-projected projection.

We make use of the following notations frequently. $\mathrm{tr}(\boldsymbol{A})$ is the trace of matrix $\boldsymbol{A}$, $\mathrm{Diag}(a_1, \cdots, a_n)$ is the diagonal matrix with elements $(a_1, \cdots, a_n)$ on the diagonal, $\to$ donates the convergence of a series of real numbers, $\to_p$ represents the in probability convergence of a series of random variables, and $\to_{a.s.}$ is the almost surely convergence of a series of random variables. $\mathrm{I}(\cdot)$ is the indicator function, and $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x} = \sum_{i=1}^p x_i^2$ is the squared l_2 norm of $p \times 1$ vector $\boldsymbol{x}$, and $\|\boldsymbol{x}\|_{\boldsymbol{\Sigma}}^2 = \boldsymbol{x}^T\boldsymbol{\Sigma}\boldsymbol{x}$ is the norm induced by $\boldsymbol{\Sigma}$. In addition, $o(1)$ and $O(1)$ define the small o and big O, $o_p(1)$ and $O_p(1)$ define the small o and big O in probability, and c, C are some generic constant numbers, $\otimes$ is the Kronecker product, $\mathrm{vec}(\boldsymbol{A})$ is vectorized a matrix $\boldsymbol{A}$. Further, $\varphi_v(\cdot)$ is a function that returns the v_{th} largest singular value of an input matrix. Let $\mathbb{O}_{p,r} = \left\{V \in \mathbb{R}^{p\times r} : V^\top V = \mathbf{I_r}\right\}$ be the set of all matrices with orthonormal columns. We define $\mathbf{I_n}$ as the $n \times n$ identity matrix.

5.2 Methods

5.2.1 Statistical models

Using the same notations in 5.1.2, let $\boldsymbol{Y} = [\boldsymbol{y}_1, \cdots \boldsymbol{y}_n]^T \in \mathbb{R}^{n\times p}$, $\boldsymbol{X} \in \mathbb{R}^{n\times q}$, and $\boldsymbol{Z} \in \mathbb{R}^{n\times k}$, we have

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{Z}\boldsymbol{D} + \boldsymbol{E}, \tag{5.2.1}$$

where $\boldsymbol{E} \in \mathbb{R}^{n\times p}$ is the error matrix, $\boldsymbol{B} \in \mathbb{R}^{q\times p}, \boldsymbol{D} \in \mathbb{R}^{k\times p}$ are fixed effects.

Given an $n \times 1$ tuning parameter vector $\boldsymbol{\Gamma} = (\gamma, \cdots, \gamma, 1, \cdots, 1)$ with the first q elements to be γ and the rest to be 1. We let $\boldsymbol{D}_\gamma = \mathrm{Diag}\big(\boldsymbol{\Gamma}\big)$. The proposed method to estimate the column space of $\boldsymbol{Z}$ has the following three steps.

(1). Project the data matrix $\boldsymbol{Y}$ onto $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T$, obtain the projected data $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}$.

(2). Obtain the first approximation of $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}$ using the rank k singular value decomposition (SVD) of the matrix $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}$, donated as $\mathrm{svd}_k(\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y})$.

(3). Project the data back, then the estimator for $\boldsymbol{W} = \boldsymbol{Z}\boldsymbol{D}$ is given by

$$\widehat{\boldsymbol{W}} = \boldsymbol{Q}_X \boldsymbol{D}_\gamma^{-1} \mathrm{svd}_k(\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}).$$

The estimator of $\boldsymbol{Z}$ can be achieved by performing extra SVD on $\widehat{\boldsymbol{W}}$.

In the first step, we use $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T$ to make a rotation and then a shrinkage on $\boldsymbol{Y}$. Compared with the confounding factor, the primary variable is shrinked much more. Therefore, the factors of the SVD in the second step mainly come from the projected confounding factor space. Finally, we use $\boldsymbol{Q}_X \boldsymbol{D}_\gamma^{-1}$ to project and rotate the confounding factor space back.

5.2.2 Statistical properties

To study how well we recover the column (singular) space of $\boldsymbol{Z}$, we need the additional notations. We decompose $\boldsymbol{ZD}$ as $\boldsymbol{ZD} = \boldsymbol{Q}_Z \begin{bmatrix} \boldsymbol{\Sigma}_{ZD} \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{R}_D^T$, where $\boldsymbol{\Sigma}_{ZD} \in \mathbb{R}^{k\times k}$ is the diagonal matrix of non-zero singular values, $\boldsymbol{Q}_X \in \mathbb{O}_{n,n}$ and $\boldsymbol{R}_B \in \mathbb{O}_{q,p}$ are left and right singular vectors, respectively. We further define

$$\boldsymbol{Q}_X^T \boldsymbol{Q}_Z = \begin{bmatrix} \boldsymbol{\alpha}_1^T & \boldsymbol{\alpha}_2^T \\ \boldsymbol{\beta}_1^T & \boldsymbol{\beta}_2^T \end{bmatrix}$$

$$\boldsymbol{\tau} = \boldsymbol{R}_D^T \boldsymbol{R}_B$$

where $\boldsymbol{\alpha}_1^T \in \mathbb{R}^{q\times k}$, $\boldsymbol{\alpha}_2^T \in \mathbb{R}^{q\times(n-k)}$, $\boldsymbol{\beta}_1^T \in \mathbb{R}^{(n-q)\times k}$, $\boldsymbol{\beta}_2^T \in \mathbb{R}^{(n-q)\times(n-k)}$, and $\boldsymbol{\tau} \in \mathbb{R}^{q\times k}$. Decompose $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD}$ as

$$\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD} = \begin{bmatrix} \boldsymbol{U}_1 & \boldsymbol{U}_{1\perp} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{V}_1^\top \\ \boldsymbol{V}_{1\perp}^\top \end{bmatrix}$$

where

$$\boldsymbol{\Sigma}_1 = \mathrm{Diag}\left\{\varphi_1\left(\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD}\right), \cdots, \varphi_k\left(\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD}\right)\right\} \in \mathbb{R}^{k\times k},$$

$$\boldsymbol{\Sigma}_2 = \mathrm{Diag}\left\{0, \cdots, 0\right\} \in \mathbb{R}^{(n-k)\times(p-k)},$$

are singular values, and $\boldsymbol{U}_1 \in \mathbb{O}_{n,q}$, $\boldsymbol{U}_{1\perp} \in \mathbb{O}_{n,n-k}$, $\boldsymbol{V}_1 \in \mathbb{O}_{p,k}$ and $\boldsymbol{V}_{1\perp} \in \mathbb{O}_{p,p-k}$ are corresponding singular vectors. We further partition $\boldsymbol{Y}$ and $\boldsymbol{XB} + \boldsymbol{ZD}$ in the same way as for $\boldsymbol{ZD}$,

$$\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T (\boldsymbol{XB} + \boldsymbol{ZD}) = \begin{bmatrix} \widetilde{\boldsymbol{U}}_1 & \widetilde{\boldsymbol{U}}_{1\perp} \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\boldsymbol{\Sigma}}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \widetilde{\boldsymbol{\Sigma}}_2 \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\boldsymbol{V}}_1^\top \\ \widetilde{\boldsymbol{V}}_{1\perp}^\top \end{bmatrix},$$

and

$$\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y} = \begin{bmatrix} \widehat{\boldsymbol{U}}_1 & \widehat{\boldsymbol{U}}_{1\perp} \end{bmatrix} \cdot \begin{bmatrix} \widehat{\boldsymbol{\Sigma}}_1 & 0 \\ 0 & \widehat{\boldsymbol{\Sigma}}_2 \end{bmatrix} \cdot \begin{bmatrix} \widehat{\boldsymbol{V}}_1^\top \\ \widehat{\boldsymbol{V}}_{1\perp}^\top \end{bmatrix},$$

where $\widetilde{\boldsymbol{U}}_1$, $\widetilde{\boldsymbol{U}}_{1\perp}$, $\widetilde{\boldsymbol{V}}_1$, $\widetilde{\boldsymbol{V}}_{1\perp}$, $\widetilde{\boldsymbol{\Sigma}}_1$ and $\widetilde{\boldsymbol{\Sigma}}_2$ and $\widehat{\boldsymbol{U}}_1$, $\widehat{\boldsymbol{U}}_{1\perp}$, $\widehat{\boldsymbol{V}}_1$, $\widehat{\boldsymbol{V}}_{1\perp}$, $\widehat{\boldsymbol{\Sigma}}_1$ and $\widehat{\boldsymbol{\Sigma}}_2$ have the same structures as $\boldsymbol{U}_1$, $\boldsymbol{U}_{1\perp}$, $\boldsymbol{V}_1$, $\boldsymbol{V}_{1\perp}$, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$.

We further assume the following assumptions.

Assumption 5.1. *Both p and n increase to ∞.*

Assumption 5.2. $p^{-1/2}\varphi_k(\boldsymbol{ZD}) \to \infty$ *and* $\frac{p}{\varphi_k(\boldsymbol{ZD})} = o(n)$.

Assumption 5.3. *The distribution of the entries of $\boldsymbol{E}$, E_{ij} is assumed to satisfy*

$$E_{ij} \overset{iid}{\sim} \mathcal{G}_\tau, \quad 1 \le i \le n, 1 \le j \le m,$$

where the the class of distributions $\mathcal{G}_\tau$ for some $\tau > 0$ is defined as

$$\textit{If } Z \sim \mathcal{G}_\tau, \textit{ then } \mathbb{E}Z = 0, \mathrm{Var}(Z) = 1, \mathbb{E}\exp(tZ) \le \exp(\tau t), \forall t \in \mathbb{R}.$$

Assumption 5.4. $0 \le \|\boldsymbol{\alpha}_1\| < 1$.

Assumption 5.2 represents the signal to noise ratio in the data. Assumption 5.3 describes the sub-gaussian distribution of error matrix $\boldsymbol{E}$. Assumption 5.4 implies that the singular space of $\boldsymbol{XB}$ and $\boldsymbol{ZD}$ are not fully correlated.

We first investigate the rate of estimating $\boldsymbol{U}_1$ with $\widetilde{\boldsymbol{U}}_1$ and estimating $\widetilde{\boldsymbol{U}}_1$ with $\widehat{\boldsymbol{U}}_1$. The results are summarized in Lemma 5.2.1 and 5.2.2 respectively.

Lemma 5.2.1. Under assumptions 5.1-5.4, we have $\|\sin\Theta(\boldsymbol{U}_1, \widetilde{\boldsymbol{U}}_1)\| = O(r_1)$, where

$$r_1 = \frac{\gamma^3 \varphi_1^2(\boldsymbol{XB})}{\varphi_q^2(\boldsymbol{ZD})} + \frac{\gamma\|\boldsymbol{\tau}\|\varphi_1(\boldsymbol{XB})}{\varphi_q(\boldsymbol{ZD})}.$$

Lemma 5.2.2. Under assumptions 5.1-5.4, we have

$$\|\sin\Theta(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1)\| = O_p(r_2),$$

where $r_2 = \sqrt{\frac{n}{\varphi_k^2(\boldsymbol{ZD})}}$.

Theorem 5.2.3. Under assumptions 5.1-5.4, we have

$$\|\sin\Theta(\boldsymbol{U}_1, \widehat{\boldsymbol{U}}_1)\| = O_p(r_3),$$

where $r_3 = r_1 + r_2$.

Lemma 5.2.1 illustrates that as we give a shrinkage γ on $\boldsymbol{XB}$, we can estimate the singular space of $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD}$ with the singular space of $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T(\boldsymbol{XB} + \boldsymbol{ZD})$ well. This is because the shrinkage γ directly affects on the singular values of $\boldsymbol{XB}$ while $\boldsymbol{ZD}$ can preserve its singular values since the singular space of $\boldsymbol{XB}$ and $\boldsymbol{ZD}$ are not fully correlated. Lemma 5.2.2 mainly addresses the perturbation bounds for the low-rank matrix $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T(\boldsymbol{XB} + \boldsymbol{ZD})$ due to the noise matrix $\boldsymbol{E}$. Theorem 5.2.3 combines the rates from Lemmas 5.2.1 and 5.2.2 together and describes the rate of $\sin\Theta$ when estimating the singular space of $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZD}$ using $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{ZY}$.

To further investigate how we recover the column space of $\boldsymbol{ZD}$ after we apply $\boldsymbol{Q}_X^T \boldsymbol{D}_\gamma^{-1}$ to project the space back, we need the following additional assumptions.

Assumption 5.5. *There exits* $\boldsymbol{W}_2 \in \mathbb{O}_{k,k}$ *such that* $\|(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1 \boldsymbol{W}_2)_{[1:k,:]}\| = O_p(\gamma r_2)$ *and* $\|(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1 \boldsymbol{W}_2)_{[(k+1):n,:]}\| = O_p(r_2)$.

Assumption 5.5 implies that when we estimate $\widetilde{\boldsymbol{U}}_1$ by $\widehat{\boldsymbol{U}}_1$ the rate for the first k rows is γ times smaller than the rest of the rows. To see the intuition behind this, we first have $\|(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1 \boldsymbol{W})\| \asymp \|\sin\Theta(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1)\| = O_p(r_2)$ by the Lemma 1 in Tony Cai and Zhang (2017).

Then consider

$$\begin{aligned}
\widehat{\boldsymbol{U}}_1 = \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y} \widehat{\boldsymbol{V}}_1 \widehat{\boldsymbol{\Sigma}}_1^{-1} &= \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T (\boldsymbol{X}\boldsymbol{B} + \boldsymbol{Z}\boldsymbol{D}) \widehat{\boldsymbol{V}}_1 \widehat{\boldsymbol{\Sigma}}_1^{-1} + \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{E} \widehat{\boldsymbol{V}}_1 \widehat{\boldsymbol{\Sigma}}_1^{-1} \\
&\approx D_\gamma \boldsymbol{Q}_X^T (\boldsymbol{X}\boldsymbol{B} + \boldsymbol{Z}\boldsymbol{D}) \widetilde{\boldsymbol{V}}_1 \widetilde{\boldsymbol{\Sigma}}_1^{-1} + \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{E} \widetilde{\boldsymbol{V}}_1 \widetilde{\boldsymbol{\Sigma}}_1^{-1} \\
&= \widetilde{\boldsymbol{U}}_1 + \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{E} \widetilde{\boldsymbol{V}}_1 \widetilde{\boldsymbol{\Sigma}}_1^{-1}.
\end{aligned}$$

We can see that the difference of $\widetilde{\boldsymbol{U}}_1$ and $\widehat{\boldsymbol{U}}_1$ is led by the first order approximation $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{E} \widetilde{\boldsymbol{V}}_1 \widetilde{\boldsymbol{\Sigma}}_1^{-1}$, the first k rows of which have the stand deviation γ times smaller than the rest rows.

Then we can introduce our Theorem 5.2.4.

Theorem 5.2.4. Under assumptions 5.1-5.5, there exits $\boldsymbol{W} \in \mathbb{O}_{k,k}$ such that

$$\|\boldsymbol{Q}_X \boldsymbol{D}_\gamma^{-1} (\widehat{\boldsymbol{U}}_1 - \boldsymbol{U}_1 \boldsymbol{W})\| = O_p(r_4),$$

where

$$r_4 = r_1/\gamma + r_2 = \sqrt{\frac{n}{\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})}} + \frac{\gamma^2 \|\boldsymbol{\alpha}_1\| \varphi_1^2(\boldsymbol{X}\boldsymbol{B})}{\sqrt{1 - \|\boldsymbol{\alpha}_1\|^2} \varphi_q^2(\boldsymbol{Z}\boldsymbol{D})} + \frac{\|\boldsymbol{\tau}\| \varphi_1(\boldsymbol{X}\boldsymbol{B})}{\varphi_q(\boldsymbol{Z}\boldsymbol{D})}.$$

If we assume $\|\boldsymbol{\tau}\| = o(\frac{n}{\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})})$, we have

$$r_4 = \sqrt{\frac{n}{\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})}} + \frac{\gamma^2 \|\boldsymbol{\alpha}_1\| \varphi_1^2(\boldsymbol{X}\boldsymbol{B})}{\sqrt{1 - \|\boldsymbol{\alpha}_1\|^2} \varphi_q^2(\boldsymbol{Z}\boldsymbol{D})}$$

Theorem 5.2.4 describes that we can recover the column space of $\boldsymbol{Z}$ by $\boldsymbol{Q}_X \boldsymbol{D}_\gamma^{-1} \widehat{\boldsymbol{U}}_1$ and it also gives us the insight of how choose the parameter γ. As long as we choose γ such that $\gamma = o(\frac{n^{1/4}\sqrt{\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})}}{\varphi_1(\boldsymbol{X}\boldsymbol{B})})$, we have $r_4 = O_p(\sqrt{\frac{n}{\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})}}$ if $\|\boldsymbol{\tau}\|$ is small enough. Some examples regarding the scale of $\|\boldsymbol{\tau}\|$ can be seen in the Condition 5 of Lee et al. (2017) and the scenario where $\boldsymbol{B}$ is spare in Wang et al. (2017).

5.3 Numeric Experiments

In this section,we perform three sets of Monte Carlo simulations to examine the performances of the proposed and several existing methods under various simulation setups. In particular, we are interested in investigating how correlation structures among the primary and hidden confounders, as well as the correlation among the associated effects of these variables affect the performances of these methods. We also explore the impact of the sparsity of the effects and the number of hidden confounders.

5.3.1 Methods to compare

The comparison methods are described below, with the name by which each will be referred to in parentheses:

1. the hidden confounders are assumed to be known and are included as covariates in the regression model (5.1.1) (Known);
2. the hidden confounders are completely ignored and analysis is done with the reduced model (5.1.2) (Not.Adj);
3. two-step surrogate variable analysis from Leek and Storey (2007) (2-SVA);
4. iteratively reweighted surrogate variable analysis from Leek and Storey (2008) (IRW-SVA);
5. PCA on original feature matrix (PCA);
6. PCA on the residualized matrix where effects of the observed variables are removed from each feature (rPCA);
7. four-step RUV method of Gagnon-Bartsch et al. (2017) with q being estimated by PA procedure of Buja and Eyuboglu (1992) (RUV4-k1);
8. four-step RUV method of Gagnon-Bartsch et al. (2017) with r being estimated by its own built-in procedure (RUV4-k2)
9. LEAPP with robust regression from Wang et al. (2017) (LEAPP);

5.3.2 Simulation Setups

In the first set of simulations, we consider the case where the effect of primary variables and hidden confounders are independent. We generate $p = 5000$ features with sample size $n = 100$. In each simulated dataset, the i_{th} feature of the j_{th} sample, y_{ij}, is generated from the following model:

$$y_{ij} = \beta_{1i}x_{1j} + \delta_{1i}z_{1j} + \epsilon_{ij} \quad (i = 1, \cdots, p; j = 1, \cdots, n), \tag{5.3.1}$$

where x_{1j} is the primary variable, z_{1j} is a continuous hidden confounder, with corresponding β_{1i} and δ_{1i}, respectively. The random error ϵ_{ij} is generated from $N(0, 1)$. The primary variable x_{1j} and the hidden confounder z_{1j} are jointly generated from the bivariate normal distribution

$$\begin{pmatrix} x_{1j} \\ z_{1j} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right),$$

where ρ is used to control the degree of correlation between x_{1j} and z_{1j}. Four different values of ρ are considered to allow for increasing correlation: 0.1, 0.3, 0.6 and 0.9.

One set of features of size $m = 500$, 1000 and 2000 is selected to have non-zero β_{1i} coefficient respectively, which is generated from $N(0, 1)$. To mimic real microarray data where the hidden confounders affect a subset of features relatively strongly, 2000 features are set to be associated with z_{1j}, with associated coefficients generated from $N(0, 1)$ and correlated with the non-zero β_{1i}s with correlation coefficient $\rho_1 = 0$. The overlap of the non-zero index between β_{1i} and δ_{1i} are set as $s = 0.1m$, $0.5m$ and $0.9m$. Therefore, we have in total 36 simulation setups: $\rho = 0.1$, 0.3, 0.6 and 0.9; $m = 500$, 1000 and 2000; and $s = 0.1m$, $0.5m$ and $0.9m$. For each simulation setup, a total of 200 replications were conducted.

In the second set of simulations, we consider the case where the effect of primary variables and hidden confounders have correlation. Specifically, we set $\rho = 0.1$ such that β_{1i} and δ_{1i} have the correlation 0.1 on the common nonzero index. All the other setting are kept exactly the same as in the first set of simulations.

The number of surrogate variables r is estimated by the PA procedure of Buja and Eyuboglu (1992) for all of the algorithms, unless otherwise stated. In RUV4-k1 and RUV4-k2, we randomly select 300 or 400 features ($6-8\%$) as the negative controls. This percentage mimics the proportion of housekeeping genes in real microarray datasets, and is similar to those used in Gagnon-Bartsch et al. (2017) (10%) and Lee et al. (2017) (6%).

5.3.3 Simulation Results

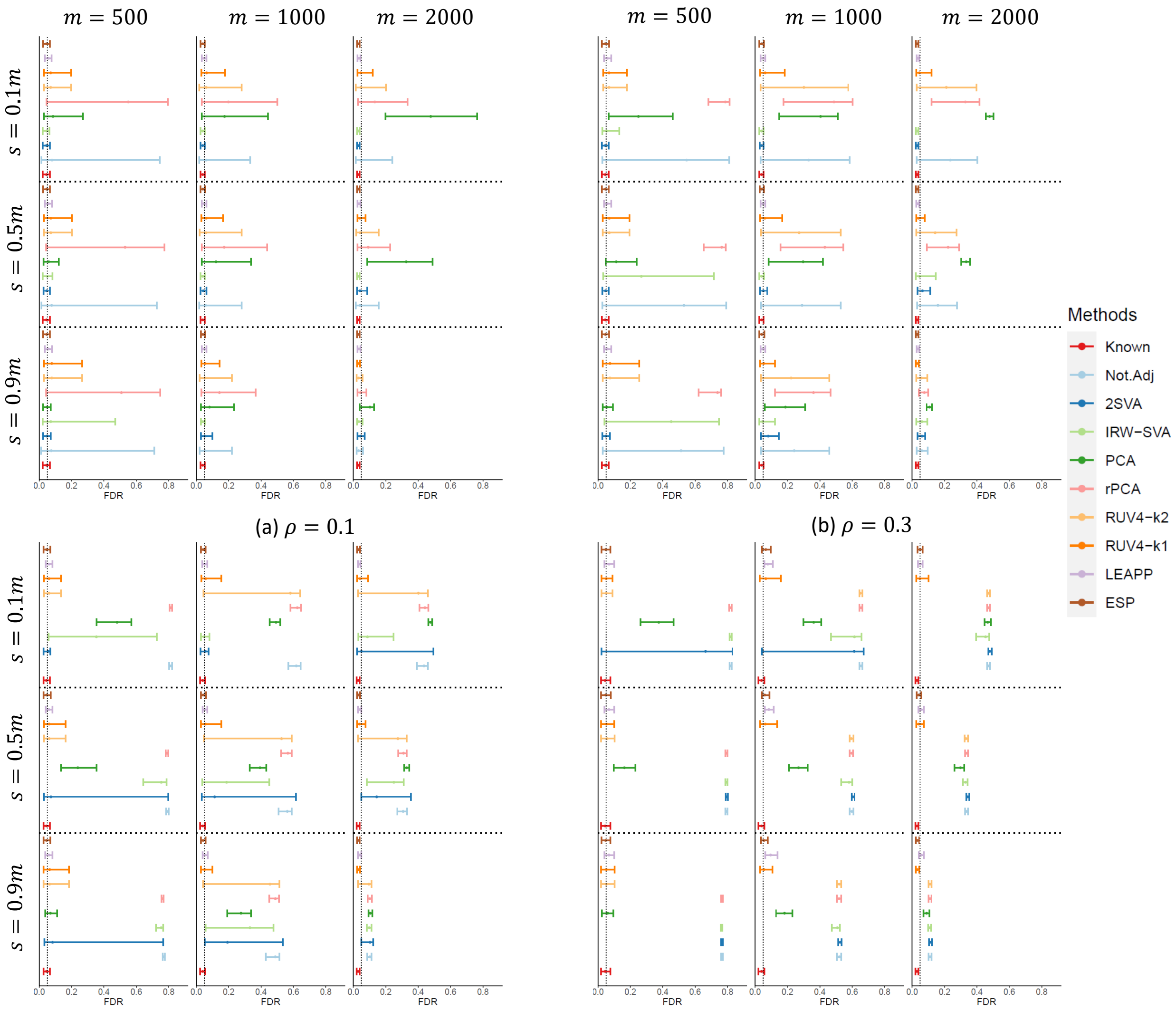
The FDR results are displayed in Figures 5.1 and 5.4. As expected, Not.Adj has inflated FDR and the inflation increases with ρ in all cases. Performance of PCA is acceptable when the hidden confounders have much larger effects than the primary variable (Figures 5.1 (a), $m = 500$). In this case, the hidden confounders clearly stand out and the top PCs from PCA on $\boldsymbol{Y}$ can sufficiently identify them. However, when the effects of the hidden confounders are similar to those of the observed variables (Figures 5.1, $m = 2000$), the FDR of PCA becomes inflated since the top r PCs can no longer recover the hidden confounders efficiently. For all simulated cases, rPCA has noticeably inflated FDR regardless of ρ. In all cases, both 2-SVA and IRW-SVA have noticeably inflated FDR when ρ is large and the inflation increases with ρ. In addition, 2-SVA has noticeably inflated FDR when the effects of the primary and hidden confounders are correlated.

The performance of RUV4-k1 varies. It suffers from inflated FDR for some cases (Figures 5.1 $m = 1000$), possibly due to the lack of a perfect set of negative controls. For RUV4-k2, the built-in estimation procedure for r (Gagnon-Bartsch et al., 2017) does not always perform well and thus has noticeably inflated FDR. Both RUV4-k1 and RUV4-k2 also have inflated FDR when the effects of the primary and hidden confounders are correlated (Figure 5.4). In contrast, LEAPP performs well in most cases, though it may have slightly inflated FDR which increases with ρ. In all cases, ESP performs closest to Known and has well-controlled FDR.

Mean squared errors (MSE) results are summarized in Figures 5.2 and 5.5. The MSE of

LEAPP, RUV-k1 and ESP are comparable and are close to those of Known. Other methods have inflated MSE when ρ is large. In many neuroimaging studies, a key step is to select and regress out a list of confounders from the imaging features for downstream follow-up analysis. Therefore, a method's ability to recover the hidden confounders is also examined. Specifically, a linear regression with the first hidden confounder $z_1 = (z_{11}, \cdots, z_{1n})$ as response variable and estimated hidden confounders as covariates is fitted for each of the method, except for the Known and Not.Adj. R-squared (R^2) results are presented in Figures 5.3 and 5.6. Clearly, the hidden confounders estimated from 2-SVA, IRW-SVA, PCA, rPCA and RUV4-k2 all depart noticeably from z_1. The remaining methods, RUV4-k1, LEAPP and ESP perform better, but ESP is the clear winner.

In summary, ESP provides the most stable FDR control for data with complicated dependence structures and can serve as a powerful alternative method for detecting and recover hidden confounding variables.



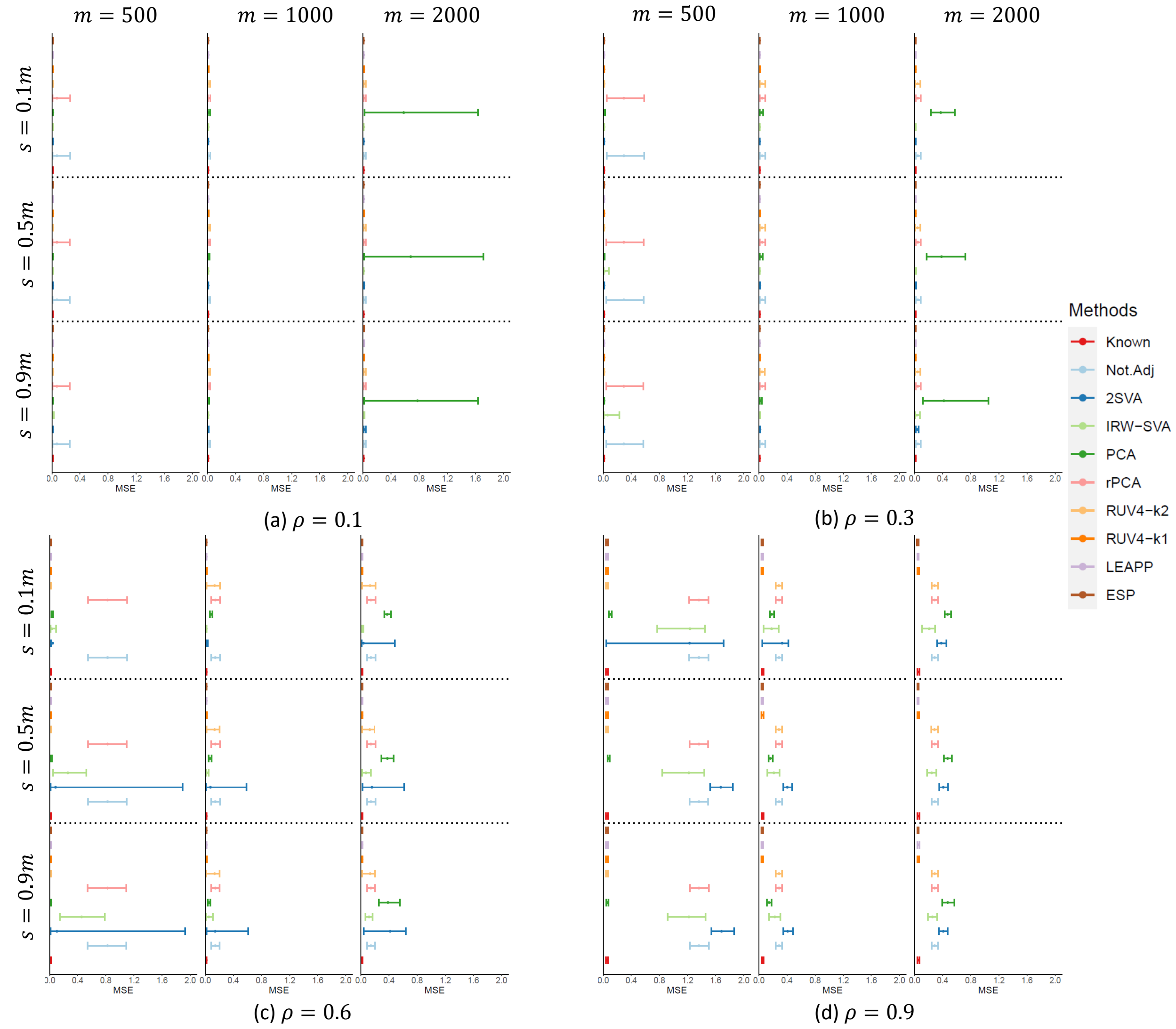
(a) ρ = 0.1

(b) ρ = 0.3

(c) ρ = 0.6

(d) ρ = 0.9

Figure 5.1: FDR results of the proposed and competing method in first set of simulations. The error bars are one empirical 95% confidence interval over 200 repeated simulations. The dashed vertical line is the 5% FDR level.



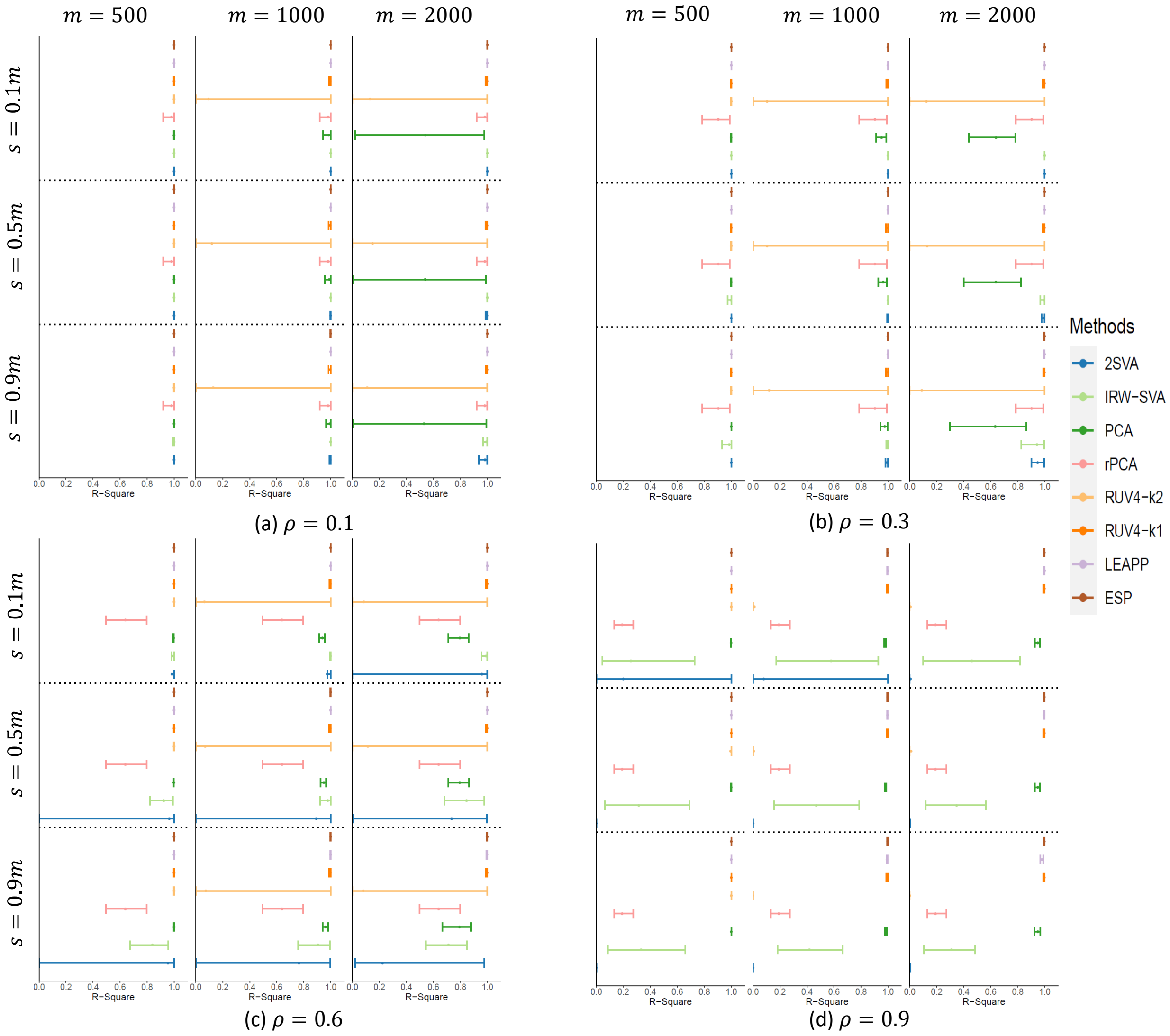
(a) $\rho = 0.1$

(b) $\rho = 0.3$

(c) $\rho = 0.6$

(d) $\rho = 0.9$

Figure 5.2: MSE results of the proposed and competing method in first set of simulations. The error bars are one empirical 95% confidence interval over 200 repeated simulations.



(a) $\rho = 0.1$

(b) $\rho = 0.3$

(c) $\rho = 0.6$

(d) $\rho = 0.9$

Figure 5.3: R^2 results of the proposed and competing method in first set of simulations. The error bars are one empirical 95% confidence interval over 200 repeated simulations.

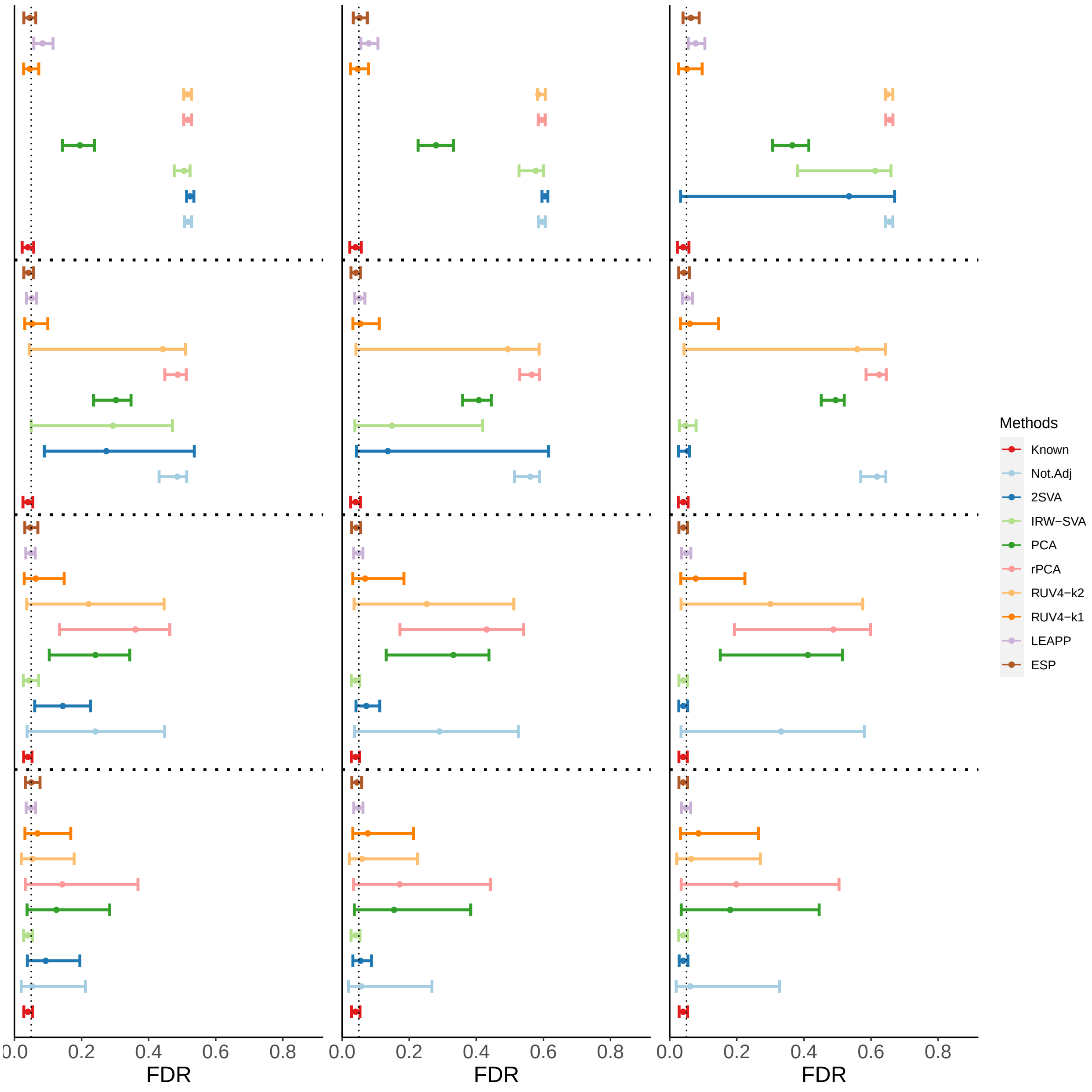


Figure 5.4: FDR results of the proposed and competing method in second set of simulations. The error bars are one empirical 95% confidence interval over 200 repeated simulations. The dashed vertical line is the 5% FDR level.

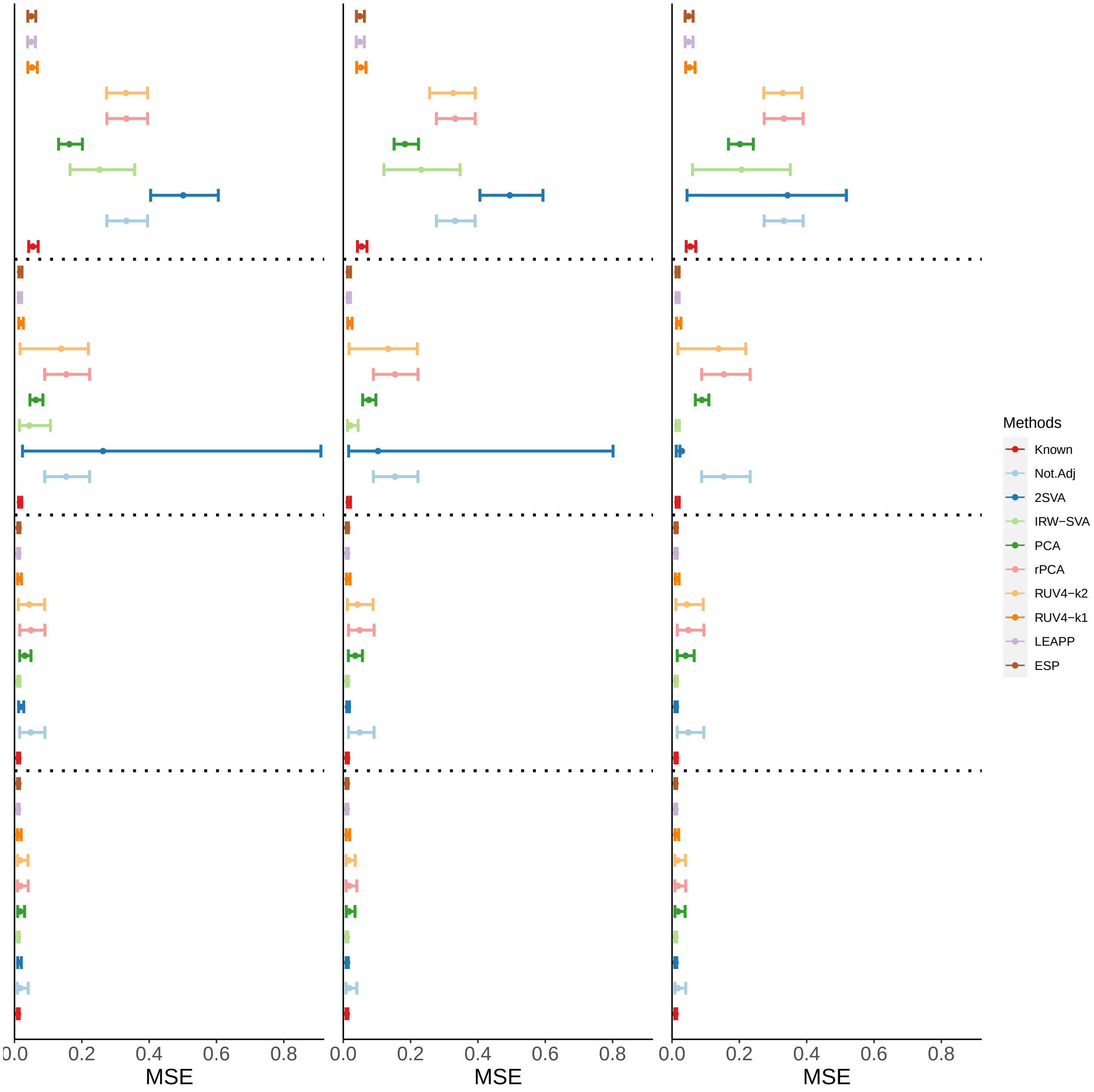


Figure 5.5: MSE results of the proposed and competing method in second set of simulations. The error bars are one empirical 95% confidence interval over 200 repeated simulations.

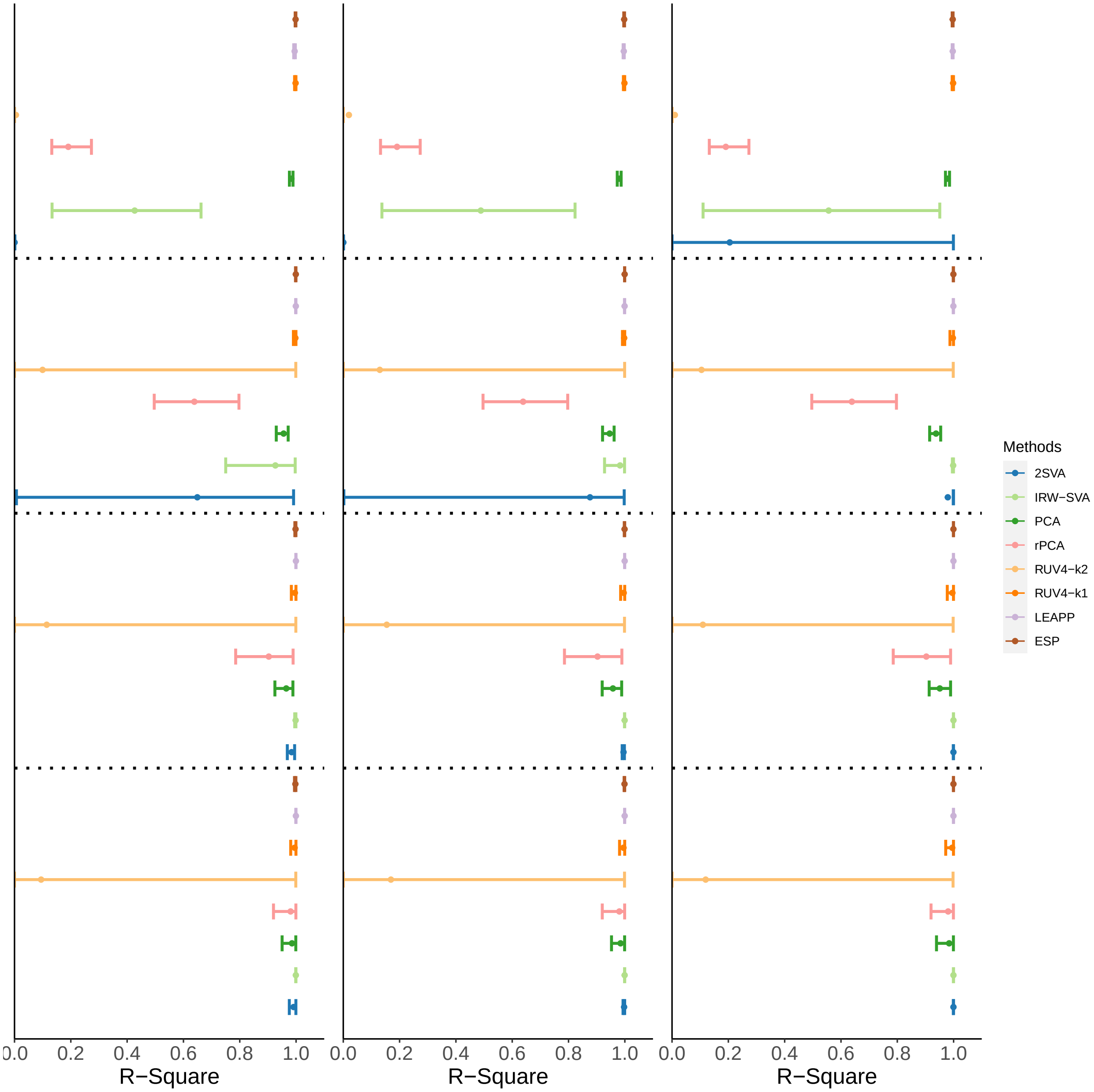


Figure 5.6: R^2 results of the proposed and competing method in second set of simulations. The error bars are one empirical 95% confidence interval over 200 repeated simulations.

APPENDIX A: TECHNICAL DETAILS OF CHAPTER 3

The appendix collects the proofs of Theorems 3.3.1, 3.3.2, 3.3.3, 3.3.4, 3.3.5.

A.1 Proof of Theorem 3.3.1

Proof of Theorem 3.3.1. For simplicity, we omit $\mathbf{d}_0$ in $\mathcal{L}_{\mathbf{d}_0}(\boldsymbol{\beta}^*)$. We prove a stronger result that the estimated coefficient vector $\widehat{\boldsymbol{\beta}}(\mathbf{d}_0)$ is exactly equal to the oracle estimator $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)$. For notational simplicity, we write $\boldsymbol{\beta}^* = \boldsymbol{\beta}^*(\mathbf{d}_0)$ and $\widehat{\boldsymbol{\beta}}^\lambda = \widehat{\beta}(\mathbf{d}_0, \lambda)$. Let $F(\boldsymbol{\beta}) = \mathcal{L}(\boldsymbol{\beta}) + p_\lambda(\boldsymbol{\beta})$, where with slight abuse of notation we write $p_\lambda(\boldsymbol{\beta}) = \sum_{j=1}^p p_\lambda(\boldsymbol{\beta}_j)$.

To show $\widehat{\boldsymbol{\beta}}^\lambda = \widehat{\boldsymbol{\beta}}^0$, we first establish some basic properties of $\widehat{\boldsymbol{\beta}}^\lambda$ and $\widehat{\boldsymbol{\beta}}^0$. We show that $\widehat{\boldsymbol{\beta}}^0$ is a stationary point of (3.2.2). Suppose $||\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^0)||_\infty \le \lambda/2$. Since $\lim_{x\to 0} p'_\lambda(x) = \lambda$, there exists a subgradient $\nabla p_\lambda(\widehat{\boldsymbol{\beta}}^0)$ such that

$$(\nabla F(\widehat{\boldsymbol{\beta}}^0))_{\mathcal{S}^c} = (\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^0))_{\mathcal{S}^c} + (\nabla p_\lambda(\widehat{\boldsymbol{\beta}}^0))_{\mathcal{S}^c} = \mathbf{0}.$$

Moreover by the assumptions of the theorem we have $\min_{j\in\mathcal{S}} |\widehat{\beta}_j^0| \ge \gamma\lambda$, and thus

$$(\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^0))_{\mathcal{S}} + (\nabla p_\lambda(\widehat{\boldsymbol{\beta}}^0))_{\mathcal{S}} = \mathbf{0}.$$

Adding the above two equalities gives $\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^0) + \nabla p_\lambda(\widehat{\boldsymbol{\beta}}^0) = \mathbf{0}$ and thus $\widehat{\boldsymbol{\beta}}^0$ is a stationary point of (3.2.2).

Apparently the global minimum $\widehat{\boldsymbol{\beta}}^\lambda$ is also a stationary point of the program (3.2.2). We collect some lemmas which characterize the basic property of $\widehat{\boldsymbol{\beta}}^\lambda$.

Lemma A.1.1. Assume Assumptions 3.1 and 3.2. Suppose $\rho > 3\mu/4$, $r = 2R$ and $2\|\nabla\mathcal{L}(\boldsymbol{\beta}^*)\|_\infty \le \lambda$. Then

$$\|(\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*)_{\mathcal{A}^c}\|_1 \le \|(\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*)_{\mathcal{A}}\|_1,\ \|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_2 \le 1.5(\rho - 3\mu/4)^{-1}\lambda\sqrt{s},$$

where $\mathcal{A}$ is the subset indexing the largest s coordinates of $\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*$ in magnitude.

Proof of Lemma A.1.1. Since $\widehat{\boldsymbol{\beta}}^\lambda \in \mathbb{B}_2(R) := \{x : \|x\|_1 \le R\}$, we have $\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_1 \le 2R$.

Using Assumption 3.2 with $r = 2R$ acquires

$$\rho\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_2^2 - \tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_1^2 \le \langle \nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^\lambda) - \nabla\mathcal{L}(\boldsymbol{\beta}^*), \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle.$$

Using the first order optimality condition that

$$-\langle \nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^\lambda) + \nabla p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda), \boldsymbol{\beta}^* - \widehat{\boldsymbol{\beta}}^\lambda\rangle \le 0$$

for some $\nabla p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) \in \partial p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda)$, we obtain

$$\langle \nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^\lambda) - \nabla\mathcal{L}(\boldsymbol{\beta}^*), \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle \le \langle -\nabla p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) - \nabla\mathcal{L}(\boldsymbol{\beta}^*), \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle. \tag{A.1.1}$$

Since $g(x) := p_\lambda(x) + \mu x^2/2$ is convex, we have

$$g(\boldsymbol{\beta}^*) - g(\widehat{\boldsymbol{\beta}}^\lambda) + \langle g'(\widehat{\boldsymbol{\beta}}^\lambda), \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle \ge 0,$$

or equivalently

$$-\langle \nabla p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda), \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle \le g(\boldsymbol{\beta}^*) - g(\widehat{\boldsymbol{\beta}}^\lambda) + \langle \mu\widehat{\boldsymbol{\beta}}^\lambda, \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle \le p_\lambda(\boldsymbol{\beta}^*) - p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) + \frac{\mu}{2}\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_2^2.$$

Plugging the above inequality into (A.1.1) and after some algebra, we obtain

$$\langle \nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^\lambda) - \nabla\mathcal{L}(\boldsymbol{\beta}^*), \widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\rangle \le \|\nabla\mathcal{L}(\boldsymbol{\beta}^*)\|_\infty\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_1 + p_\lambda(\boldsymbol{\beta}^*) - p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) + \frac{\mu}{2}\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_2^2,$$

and thus

$$\begin{aligned}(\rho - \mu/2)\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_2^2 &\le p_\lambda(\boldsymbol{\beta}^*) - p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) + \left(\|\nabla\mathcal{L}(\boldsymbol{\beta}^*)\|_\infty + 2R\tau\frac{\log p}{n}\right)\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_1 \\ &\le p_\lambda(\boldsymbol{\beta}^*) - p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) + \frac{\lambda}{2}\|\widehat{\boldsymbol{\beta}}^\lambda - \boldsymbol{\beta}^*\|_1,\end{aligned} \tag{A.1.2}$$

provided $\|\nabla\mathcal{L}(\boldsymbol{\beta}^*)\|_\infty + 2R\tau(\log p)/n \le \lambda/2$. Now using the fact that $\lambda|x| \le p_\lambda(x) + \mu x^2/2$

implied by the convexity of $p_\lambda(x)+\mu x^2/2-\lambda|x|$ and the subadditivity of p_λ which is then implied by Assumption 3.1 (iii), we have

$$(\rho-3\mu/4)\|\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*\|_2^2 \le p_\lambda(\boldsymbol{\beta}^*)-p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda)+\frac{p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda)+p_\lambda(\boldsymbol{\beta}^*)}{2},$$

or equivalently

$$0\le 2(\rho-3\mu/4)\|\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*\|_2^2 \le 3p_\lambda(\boldsymbol{\beta}^*)-p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda).$$

For the RHS of the inequality above, we have

$$\begin{aligned}
3p_\lambda(\boldsymbol{\beta}^*)-p_\lambda(\widehat{\boldsymbol{\beta}}^\lambda) &= 3p_\lambda\big(\boldsymbol{\beta}^*_{\mathcal{S}}\big)-p_\lambda\big(\widehat{\boldsymbol{\beta}}^\lambda_{\mathcal{S}}\big)-p_\lambda\big(\widehat{\boldsymbol{\beta}}^\lambda_{\mathcal{S}^c}\big)\\
&\le 3p_\lambda\big((\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{S}}\big)-p_\lambda\big(\widehat{\boldsymbol{\beta}}^\lambda_{\mathcal{S}^c}\big) && \text{(subadditivity)}\\
&= 3p_\lambda\big((\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{S}}\big)-p_\lambda\big((\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{S}^c}\big)\\
&\le 3p_\lambda\big((\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}}\big)-p_\lambda\big((\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}^c}\big)\\
&\le \lambda\big(3\|(\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}}\|_1-\|(\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}^c}\|_1\big), && \text{(Assumption 3.1 (iii))}
\end{aligned}$$

where $\mathcal{A}$ is the subset indexing the largest s coordinates of $\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*$ in magnitude. Therefore, $\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*$ falls in the ℓ_1 cone of $\{\Delta:\|\Delta_{\mathcal{A}^c}\|_1\le 3\|\Delta_{\mathcal{A}}\|_1\}$, that is

$$\|(\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}^c}\|_1 \le \|(\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}}\|_1.$$

Thus we must have

$$2(\rho-3\mu/4)\|\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*\|_2^2 \le 3\lambda\|(\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}}\|_1 \le 3\lambda\sqrt{s}\|(\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*)_{\mathcal{A}}\|_2,$$

which implies $\|\widehat{\boldsymbol{\beta}}^\lambda-\boldsymbol{\beta}^*\|_2\le 1.5(\rho-3\mu/4)^{-1}\lambda\sqrt{s}$.

□

Now we proceed to show that $\widehat{\boldsymbol{\beta}}^{\lambda} = \widehat{\boldsymbol{\beta}}^{0}$. We first show that $\widehat{\boldsymbol{\beta}}^{\lambda}$ is supported on $\mathcal{S}$. By Lemma A.1.1, $\widehat{\boldsymbol{\beta}}^{\lambda}$ is an interior local minimum and thus $\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^{\lambda}) + \nabla p_{\lambda}(\widehat{\boldsymbol{\beta}}^{\lambda}) = 0$, or equivalently $\nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{\lambda}) + \lambda\boldsymbol{\zeta} = \mathbf{0}$, for some $\boldsymbol{\zeta} \in \partial\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1$. Applying the LRSC condition, i.e., Assumption 3.2, we have

$$-\tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\|_1^2 + \rho\|\widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\|_2^2 \le \langle \nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^{0}) - \nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^{\lambda}), \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\rangle,$$

which implies

$$\langle \nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{0}) - \nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{\lambda}), \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\rangle \ge (\rho - \mu)\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_2^2 - \tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_1^2.$$

Also, $\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^{0}) + \partial p_{\lambda}(\widehat{\boldsymbol{\beta}}^{0}) = \mathbf{0}$ together with $\nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{\lambda}) + \lambda\boldsymbol{\zeta} = \mathbf{0}$ yields

$$\begin{aligned}
0 &= \langle \nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{\lambda}), \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\rangle + \lambda\langle \boldsymbol{\zeta}, \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\rangle \\
&= \langle \nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{\lambda}) - \nabla\bar{\mathcal{L}}(\widehat{\boldsymbol{\beta}}^{0}), \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\rangle + \lambda\langle \boldsymbol{\zeta}, \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}\rangle - \lambda\langle \boldsymbol{\zeta}_0, \widehat{\boldsymbol{\beta}}^{0} - \widehat{\boldsymbol{\beta}}^{\lambda}\rangle \\
&\le \tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_1^2 - (\rho - \mu)\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_2^2 + \lambda\langle \boldsymbol{\zeta}, \widehat{\boldsymbol{\beta}}^{0}\rangle + \lambda\langle \boldsymbol{\zeta}_0, \widehat{\boldsymbol{\beta}}^{\lambda}\rangle - \lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1 - \lambda\|\widehat{\boldsymbol{\beta}}^{0}\|_1.
\end{aligned} \tag{A.1.3}$$

Rearranging the above inequality, we obtain

$$\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1 - \lambda\langle \boldsymbol{\zeta}_0, \widehat{\boldsymbol{\beta}}^{\lambda}\rangle \le \tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_1^2 - (\rho - \mu)\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_2^2.$$

We need the following result.

Lemma A.1.2. If $\|\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^{0})\|_\infty \le \lambda/2$ and $\tau s(\log p/n)^{3/2} \le c\lambda$ for some small enough constant c, then

$$\|\widehat{\boldsymbol{\Delta}}\|_1 \le \frac{5}{2}\sqrt{s}\|\widehat{\boldsymbol{\Delta}}\|_2.$$

Proof of Lemma A.1.2. Rearranging (A.1.3) acquires

$$(\rho-\mu)\|\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0}\|_2^2-\tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0}\|_1^2\leq\lambda\langle\boldsymbol{\zeta},\widehat{\boldsymbol{\beta}}^{0}\rangle+\lambda\langle\boldsymbol{\zeta}_0,\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0}\rangle-\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1.$$

Now since $\mathrm{supp}(\widehat{\boldsymbol{\beta}}^{0})\subseteq\mathcal{S}$, we have

$$\lambda\langle\boldsymbol{\zeta},\widehat{\boldsymbol{\beta}}^{0}\rangle-\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1\leq\lambda\|\widehat{\boldsymbol{\beta}}^{0}\|_1-\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1\leq\lambda(\|(\widehat{\boldsymbol{\beta}}^{0}-\widehat{\boldsymbol{\beta}}^{\lambda})_{\mathcal{S}}\|_1-\|(\widehat{\boldsymbol{\beta}}^{0}-\widehat{\boldsymbol{\beta}}^{\lambda})_{\mathcal{S}^c}\|_1).$$

For $\lambda\langle\boldsymbol{\zeta}_0,\widehat{\boldsymbol{\beta}}^{\lambda}\rangle$, since $\|\boldsymbol{\zeta}_{0\mathcal{S}^c}\|_\infty\leq\lambda/2$, we have

$$\lambda\langle\boldsymbol{\zeta}_0,\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0}\rangle\leq\lambda\|(\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0})_{\mathcal{S}}\|_1+\lambda/2\|(\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0})_{\mathcal{S}^c}\|_1.$$

For simplicity, let $\widehat{\boldsymbol{\Delta}}=\widehat{\boldsymbol{\beta}}^{\lambda}-\widehat{\boldsymbol{\beta}}^{0}$. Plugging the above bounds into the first inequality acquires

$$(\rho-\mu)\|\widehat{\boldsymbol{\Delta}}\|_2^2-\tau\frac{\log p}{n}\|\widehat{\boldsymbol{\Delta}}\|_1^2\leq 2\lambda\|\widehat{\boldsymbol{\Delta}}_{\mathcal{S}}\|_1-\lambda/2\|\widehat{\boldsymbol{\Delta}}_{\mathcal{S}^c}\|_1,$$

which, under the assumption that

$$\tau\log p\|\widetilde{\boldsymbol{\Delta}}\|_1\leq C\tau s(\log p/n)^{3/2}\leq\lambda/2 \qquad \text{(by Assumption)}$$

for some large enough constant C, then yields

$$\frac{\lambda}{2}\|\widehat{\boldsymbol{\Delta}}\|_1\leq 2\lambda\|\widehat{\boldsymbol{\Delta}}_{\mathcal{S}}\|_1-\frac{\lambda}{2}\|\widehat{\boldsymbol{\Delta}}_{\mathcal{S}^c}\|_1,$$

or equivalently

$$2\|\widehat{\boldsymbol{\Delta}}_{\mathcal{S}^c}\|_1\leq 3\|\widehat{\boldsymbol{\Delta}}_{\mathcal{S}}\|_1.$$

The statement follows. □

Applying Lemma A.1.2, we obtain

$$\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1 - \lambda\langle \boldsymbol{\zeta}_0, \widehat{\boldsymbol{\beta}}^{\lambda}\rangle \le \tau\frac{\log p}{n}\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_1^2 - (\rho-\mu)\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_2^2 \le -\frac{\rho-\mu}{2}\|\widehat{\boldsymbol{\beta}}^{\lambda} - \widehat{\boldsymbol{\beta}}^{0}\|_2^2,$$

if $13\tau s(\log p)/n \le \rho - \mu$. Thus $\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1 - \lambda\langle \boldsymbol{\zeta}_0, \widehat{\boldsymbol{\beta}}^{\lambda}\rangle \le 0$ and we must have $\lambda\|\widehat{\boldsymbol{\beta}}^{\lambda}\|_1 = \lambda\langle \boldsymbol{\zeta}_0, \widehat{\boldsymbol{\beta}}^{\lambda}\rangle$, which holds only if $\mathrm{supp}(\widehat{\boldsymbol{\beta}}^{\lambda}) \subseteq \mathcal{S}$.

After showing $\mathrm{supp}(\widehat{\boldsymbol{\beta}}^{\lambda}) \subseteq \mathcal{S}$, in order to prove $\widehat{\boldsymbol{\beta}}^{\lambda} = \widehat{\boldsymbol{\beta}}^{0}$, it suffices to show the strict convexity of F over the sparse cone. This can be reduced to show that $\bar{\mathcal{L}} = \mathcal{L} - h$ is strictly convex since $\lambda\|\boldsymbol{\beta}\|_1$ is convex. This is obvious by the LRSC condition as long as $\rho > \mu + \tau(\log p)/n$. Thus the result follows on the event that $\{\|\nabla\mathcal{L}(\boldsymbol{\beta}^*)\|_\infty \le \lambda/2\} \cap \{\|\nabla\mathcal{L}(\widehat{\boldsymbol{\beta}}^{0})\|_\infty \le \lambda/2\}$.

□

A.2 Validation of Assumption 3.8

In this section, we use the examples of L_2 loss and Huber loss functions to validate assumption 3.8.

A.2.1 Example 1: Least squares estimator

Divide the matrix X as $X = (X_{\mathcal{S}(\mathbf{d}_0)}, X_{\mathcal{S}(\mathbf{d}_0)^C})$. For least square estimator, we have $\mathcal{L}(\boldsymbol{\beta}; Y(\mathbf{d}_0), X) = \frac{1}{2n}\|Y(\mathbf{d}_0) - X\boldsymbol{\beta}\|_2^2$ and $\widehat{\boldsymbol{\beta}}^{0}(\mathbf{d}_0) = (X_{\mathcal{S}(\mathbf{d}_0)}^T X_{\mathcal{S}(\mathbf{d}_0)})^{-1} X_{\mathcal{S}(\mathbf{d}_0)}^T Y(\mathbf{d}_0)$.

Then by Assumption 3.3, we have

$$\|\mathrm{bias}(\widehat{\boldsymbol{\beta}}^{0}(\mathbf{d}_0))\|_2 := \|E\widehat{\boldsymbol{\beta}}^{0}(\mathbf{d}_0) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_2 = 0$$

and for any unit vector $\boldsymbol{u} \in \mathcal{R}^{p\times 1}$ we have

$$\begin{aligned}
&\mathrm{var}(\boldsymbol{u}^T\widehat{\boldsymbol{\beta}}^{0}(\mathbf{d}_0)) = \mathrm{var}(\boldsymbol{u}^T(X_{\mathcal{S}(\mathbf{d}_0)}^T X_{\mathcal{S}(\mathbf{d}_0)})^{-1} X_{\mathcal{S}(\mathbf{d}_0)}^T \epsilon(\mathbf{d}_0)) \\
&= \|X_{\mathcal{S}(\mathbf{d}_0)}(X_{\mathcal{S}(\mathbf{d}_0)}^T X_{\mathcal{S}(\mathbf{d}_0)})^{-1}\boldsymbol{u}\|^2 \mathrm{var}\Big(\frac{\boldsymbol{u}^T(X_{\mathcal{S}(\mathbf{d}_0)}^T X_{\mathcal{S}(\mathbf{d}_0)})^{-1} X_{\mathcal{S}(\mathbf{d}_0)}^T}{\|X_{\mathcal{S}(\mathbf{d}_0)}(X_{\mathcal{S}(\mathbf{d}_0)}^T X_{\mathcal{S}(\mathbf{d}_0)})^{-1}\boldsymbol{u}\|}\epsilon(\mathbf{d}_0)\Big) \\
&\le \boldsymbol{u}^T(X_{\mathcal{S}(\mathbf{d}_0)}^T X_{\mathcal{S}(\mathbf{d}_0)})^{-1}\boldsymbol{u}\sigma^2 \asymp \frac{1}{n}\sigma^2.
\end{aligned}$$

A.2.2 Example 2: Robust regression estimator

For robust regression estimator, we have $\boldsymbol{\mathcal{L}}_{\mathbf{d}_0}(\boldsymbol{\beta}^*) = \frac{1}{n}\sum_{i=1}^n \ell_\tau\left(\epsilon_i(\mathbf{d}_0)\right)$ and $\nabla\boldsymbol{\mathcal{L}}_{\mathbf{d}_0}(\boldsymbol{\beta}^*) = \frac{1}{n}X^T_{\mathcal{S}(\mathbf{d}_0)}\ell'_\tau(\epsilon(\mathbf{d}_0))$ where

$$\ell_\tau(x) = \begin{cases} x^2/2 & \text{if } |x| \le \tau \\ \tau|x| - \tau^2/2 & \text{if } |x| > \tau \end{cases},$$

$\ell'_\tau(\epsilon(\mathbf{d}_0)) = (\ell'_\tau(\epsilon_1(\mathbf{d}_0)), \cdots, \ell'_\tau(\epsilon_n(\mathbf{d}_0)))^T$ and $\ell'_\tau(x) = xI(|x| \le \tau) + \tau\,\mathrm{sign}(x)I(|x| > \tau)$. We further have

$$|E\ell'_\tau(\epsilon_i(\mathbf{d}_0))| \le \frac{\sigma^2 - E\left[\ell'_\tau(\epsilon_i)^2\right]}{\tau} \le \frac{\sigma^2}{\tau},$$

and $\Sigma_{\ell'_\tau(\epsilon(\mathbf{d}_0))} := \mathrm{cov}(\ell'_\tau(\epsilon(\mathbf{d}_0)))$ is a diagonal matrix with identical diagonal element $\mathrm{var}\left[\ell'_\tau(\epsilon_1)^2\right]$.

Refer to the equation (C.25) in Sun et al. (2020) and take expectation with $\eta = 1$. Then for some $a_0 > 0$ and $\tau \asymp \sqrt{n}$, we have

$$\begin{aligned}
&E\|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_2^2 \le Ea_0^{-2}\|\nabla\boldsymbol{\mathcal{L}}_{\mathbf{d}_0}(\boldsymbol{\beta}^*)\|_2^2 \lesssim E\|\nabla\boldsymbol{\mathcal{L}}_{\mathbf{d}_0}(\boldsymbol{\beta}^*)\|_2^2 \\
&= \frac{1}{n^2}E\ell'_\tau(\epsilon(\mathbf{d}_0))^T X_{\mathcal{S}(\mathbf{d}_0)}X^T_{\mathcal{S}(\mathbf{d}_0)}\ell'_\tau(\epsilon(\mathbf{d}_0)) \\
&= \frac{1}{n^2}E\ell'_\tau(\epsilon(\mathbf{d}_0))^T X_{\mathcal{S}(\mathbf{d}_0)}X^T_{\mathcal{S}(\mathbf{d}_0)}E\ell'_\tau(\epsilon(\mathbf{d}_0)) + \frac{1}{n^2}\mathrm{tr}(X_{\mathcal{S}(\mathbf{d}_0)}X^T_{\mathcal{S}(\mathbf{d}_0)}\Sigma_{\ell'_\tau(\epsilon(\mathbf{d}_0))}) \\
&\le \frac{\sigma^4}{\tau^2} + \frac{E\left[\ell'_\tau(\epsilon_1)^2\right]}{n^2}\mathrm{tr}((X_{\mathcal{S}(\mathbf{d}_0)}X^T_{\mathcal{S}(\mathbf{d}_0)}) \lesssim \frac{\sigma^4}{\tau^2} + \frac{s}{n}\sigma^2 \lesssim \frac{s}{n}\sigma^2
\end{aligned}$$

Finally we can have

$$\|\mathrm{bias}(\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0))\|_2 := \|E\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_2 \le \sqrt{\frac{s}{n}\sigma^2},$$

and for any unit vector $\boldsymbol{u} \in \mathcal{R}^{p\times 1}$

$$\mathrm{var}(\boldsymbol{u}^T\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)) \le \frac{s}{n}\sigma^2.$$

A.3 Proof of Theorem 3.3.2

We prove both upper bound and lower bound. We start with the upper bound.

Upper Bound: For every $\mathbf{d}_0 \in \mathcal{D}_0 \cap \Omega$, consider the oracle estimator $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)$ in (3.3.1) and define the local linear estimator of $\boldsymbol{f}^*_\Omega(\mathbf{d}_0)$ based on $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)$

$$\widehat{\boldsymbol{f}^0_\Omega}(\mathbf{d}_0) = (\widehat{f}^0_1(\mathbf{d}_0), \cdots, \widehat{f}^0_n(\mathbf{d}_0))^T = X \sum_{\mathbf{d}_m \in \mathcal{D}_0 \cap \Omega} w_h(\mathbf{d}_m, \mathbf{d}) \widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0) = \sum_{\mathbf{d}_m \in \mathcal{D}_0 \cap \Omega} w_h(\mathbf{d}_m, \mathbf{d}) \breve{\boldsymbol{f}}(\mathbf{d}_m),$$

where $w_h(\mathbf{d}_m, \mathbf{d})$ is defined in (3.2.5) and $\breve{\boldsymbol{f}}(\mathbf{d}_m) = X\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m) = (\breve{f}_1(\mathbf{d}_0), \cdots, \breve{f}_n(\mathbf{d}_0))^T$.

Define $\mathcal{N}(\mathbf{d}_0, h) = \{\mathbf{d} \in \mathcal{D}_0 \cap \Omega, K_h(\mathbf{d} - \mathbf{d}_0) > 0\}$, then we have $|\mathcal{N}(\mathbf{d}_0, h)| \asymp N_D h^3$. Further define $Z \in \mathcal{R}^{|\mathcal{N}(\mathbf{d}_0,h)|\times 4}$, $\breve{F}_0 \in \mathcal{R}^{|\mathcal{N}(\mathbf{d}_0,h)|\times n}$ and $\breve{B}_0 \in \mathcal{R}^{|\mathcal{N}(\mathbf{d}_0,h)|\times p}$ as concatenating $\mathbf{z}_h(\mathbf{d}_m - \mathbf{d})$, $\breve{\boldsymbol{f}}(\mathbf{d}_m)$ and $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m)$ row-wisely for each $\mathbf{d}_m = (d_{m,1}, d_{m,2}, d_{m,3}) \in \mathcal{N}(\mathbf{d}_0, h)$. Then (3.2.3) can be rewritten as

$$G(B) =: \sum_{\mathbf{d}_m \in \mathcal{N}(\mathbf{d}_0,h)} \|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m) - B^{\mathrm{T}}\mathbf{z}_h(\mathbf{d}_m - \mathbf{d}_0)\|^2.$$

Minimizing $G(B)$ with respect to B, we can obtain that $\widehat{B} = (Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}\breve{B}_0 \in \mathcal{R}^{4\times p}$ and further have

$$\widehat{\boldsymbol{f}^0_\Omega}(\mathbf{d}_0) = X\widehat{B}^{\mathrm{T}}e = \left((Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}\breve{F}_0\right)^{\mathrm{T}} e,$$

where $e = (1, 0, 0, 0)^{\mathrm{T}}$.

We rewrite $E\|\widehat{\boldsymbol{f}^0_\Omega}(\mathbf{d}_0)(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2$ as

$$E\|\widehat{\boldsymbol{f}^0_\Omega}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 = \sum_{i=1}^{n} E(\widehat{f}^o_i(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 = \sum_{i=1}^{n}\{\mathrm{var}(\widehat{f}^o_i(\mathbf{d}_0)) + (E(\widehat{f}^o_i(\mathbf{d}_0)) - f^*_i(\mathbf{d}_0))^2\}.$$

By the assumption 3.8, we have $\mathrm{var}(\widehat{f}^o_i(\mathbf{d}_0)) \asymp \sigma^2 s^2 n^{-1} e^{\mathrm{T}}(Z^{\mathrm{T}}Z)^{-1}e$. As $N_D h^3 \to \infty$, we have

$|\mathcal{N}(\mathbf{d}_0,h)|^{-1}Z^{\mathrm{T}}Z \to M_0$, where

$$M_0 = \begin{pmatrix} 1 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/3 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/3 & 1/4 \\ 1/2 & 1/4 & 1/4 & 1/3 \end{pmatrix}$$

is a positive-definite matrix. Hence as long as $N_D h^3$ is large enough, there is a constant $C_1 > 0$ such that $C_1^{-1}|\mathcal{N}(\mathbf{d}_0,h)| \le \lambda_{\min}(Z^{\mathrm{T}}Z) \le \lambda_{\max}(Z^{\mathrm{T}}Z) \le C_1|\mathcal{N}(\mathbf{d}_0,h)|$. Consequently, we have $\mathrm{var}(\widehat{f}_i^o(\mathbf{d}_0)) \asymp \sigma^2 s^2 n^{-1}|\mathcal{N}(\mathbf{d}_0,h)|^{-1} \asymp \sigma^2 s^2 n^{-1} N_D^{-1} h^{-3}$.

Next we will consider the bias term $(E(\widehat{f}_i^o(\mathbf{d}_0)) - f_i^*(\mathbf{d}_0))^2$. By the assumption 3.8, we have

$$\begin{aligned} |E(\widehat{f}_i^o(\mathbf{d}_0)) - f_i^*(\mathbf{d}_0)| &= |\sum_{\mathbf{d}_m \in \mathcal{N}(\mathbf{d}_0,h)} \omega(\mathbf{d}_m,\mathbf{d}_0) E\breve{f}_i(\mathbf{d}_m) - f_i^*(\mathbf{d}_0)| \\ &\le \sum_{\mathbf{d}_m \in \mathcal{N}(\mathbf{d}_0,h)} \omega(\mathbf{d}_m,\mathbf{d}_0)(|f_i^*(\mathbf{d}_m) - f_i^*(\mathbf{d}_0)| + |E\breve{f}_i(\mathbf{d}_m) - f_i^*(\mathbf{d}_m)|) \\ &= C_0 \sum_{\mathbf{d}_m \in \mathcal{N}(\mathbf{d}_0,h)} \omega(\mathbf{d}_m,\mathbf{d}_0)(\|\mathbf{d}_m - \mathbf{d}_0\|_\infty + |E\breve{f}_i(\mathbf{d}_m) - f_i^*(\mathbf{d}_m)|) \\ &\lesssim h + \sqrt{\sigma^2 s^2 n^{-1}}. \end{aligned}$$

Consequently, we have

$$E\|\widehat{\boldsymbol{f}}^o_\Omega(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 \lesssim n(\sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 + \sigma^2 s^2 n^{-1}).$$

Using a similar argument, for every $\mathbf{d}_0 \in \mathcal{D}_0 \cap \Omega^C$ we also have

$$E\|\widehat{\boldsymbol{f}}^o_{\Omega^C}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 \lesssim n(\sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 + \sigma^2 s^2 n^{-1}).$$

Therefore, define $\widehat{\boldsymbol{f}}^o(\mathbf{d}_0) = 1_{\{\mathbf{d}_0\in\Omega\}}\widehat{\boldsymbol{f}}^o_\Omega(\mathbf{d}_0) + 1_{\{\mathbf{d}_0\in\Omega^c\}}\widehat{\boldsymbol{f}}^o_{\Omega^C}(\mathbf{d}_0)$ and we have

$$\mathrm{MSE}_f(\widehat{f}^o) = \frac{\sum_{\mathbf{d}_0\in\mathcal{D}_0} E\|\widehat{\boldsymbol{f}}^o(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2}{nN_D} \lesssim (\sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 + \sigma^2 s^2 n^{-1}).$$

Let $\mathcal{E}$ be the event $\{\widehat{\mathcal{S}}(\mathbf{d}_0,\lambda) = \mathcal{S}(\mathbf{d}_0)$ for all $\mathbf{d}_0 \in \mathcal{D}_0\}$. On the event $\mathcal{E}$, we have $\widehat{\boldsymbol{\beta}}(\mathbf{d}_0,\lambda) = \widehat{\boldsymbol{\beta}}^0(\mathbf{d}_0)$. Consequently, we have

$$\begin{aligned}
&\mathrm{MSE}_f(\widehat{f}) = \frac{\sum_{\mathbf{d}_0\in\mathcal{D}_0}\sum_{i=1}^n E(\widehat{f}_i(\mathbf{d}_0) - f_i^*(\mathbf{d}_0))^2}{nN_D} \\
=\ & \frac{\sum_{\mathbf{d}_0\in\mathcal{D}_0}\{E\|\widehat{\boldsymbol{f}}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 I(\mathcal{E}) + E\|\widehat{\boldsymbol{f}}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 I(\mathcal{E}^C)\}}{nN_D} \\
=\ & \frac{\sum_{\mathbf{d}_0\in\mathcal{D}_0}\{E\|\widehat{\boldsymbol{f}}^o(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 I(\mathcal{E}) + E\|\widehat{\boldsymbol{f}}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 I(\mathcal{E}^C)\}}{nN_D} \\
\lesssim\ & \sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 + \sigma^2 s^2 n^{-1} + \frac{\sum_{\mathbf{d}_0\in\mathcal{D}_0}\{E\|\widehat{\boldsymbol{f}}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 I(\mathcal{E}^C)\}}{nN_D}.
\end{aligned}$$

And since $|\widehat{f}_i(\mathbf{d}_0) - f^*{}_i(\mathbf{d}_0)| \le 1$ by clipping, we have $n^{-1}E\|\widehat{\boldsymbol{f}}(\mathbf{d}_0) - \boldsymbol{f}^*(\mathbf{d}_0)\|^2 I(\mathcal{E}^C) \lesssim P(\mathcal{E}^C)$, which implies that

$$\mathrm{MSE}_f(\widehat{f}) \lesssim \sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 + \sigma^2 s^2 n^{-1} + P(\mathcal{E}^C)$$

and

$$\inf_h \mathcal{R}_n(\widehat{f}) \lesssim (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} + \sigma^2 s^2 n^{-1} + P(\mathcal{E}^C)$$

when $h \asymp (\sigma^2 s^2 n^{-1} N_D^{-1})^{1/5}$. By assumptions 3.3 and 3.6, we take $\delta = (1/(pN_D))^{C_1}$ and we have

$$P(\mathcal{E}^C) = \delta = o((\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5}).$$

We can further get the upper bound $\inf_h \mathcal{R}_n(\widehat{f}) \lesssim (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \vee \sigma^2 s^2 n^{-1}$.

Lower Bound: Next let us consider the lower bound of the risk $\mathcal{R}_n(\widehat{f})$ and we need to assume the $\epsilon_i(\mathbf{d})$'s are i.i.d. Gaussian random variables with mean 0 and variance σ^2. For arbitrary $\mathbf{d}_0 \in \mathcal{D} \cup \Omega$, we first consider $E(\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0))^2$. Define $g_0(\mathbf{d}) \equiv 0 : [0,1]^3 \to \mathcal{R}^1$, and $g_1(\mathbf{d}) = Lh\prod_{j=1}^3 H((d_j - d_{j,0})/h)$, where $h = c_0(\sigma^2 s^2 n^{-1} N_D^{-1})^{1/5}$ for some $c_0 > 0$, and $H(u) = aH_0(2u), H_0(u) = e^{-1/(1-u^2)} I(|u| \leq 1)$.

Step 1. $g_0(\cdot), g_1(\cdot) \in \mathcal{H}_3(1, C_0)$.

Step 2. Define $A = 1/2Lc_0\prod_{j=1}^3 H(0)$, $\psi = (\sigma^2 s^2 n^{-1} N_D^{-1})^{1/5}$, and $s_0 = \psi A$, then we have $E\{\psi^{-2}(\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0))^2\} \geq A^2 P(|\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0)| \geq s_0)$, and $|g_1(\mathbf{d}_0) - g_0(\mathbf{d}_0)| = Lh\prod_{j=1}^3 H(0) = 2A\psi = 2s_0$. Hence

$$\begin{aligned} & \inf_{\widehat{f}} \sup_{f \in \mathcal{F}} E\{\psi^{-2} \|\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0)\|^2\} \\ \geq \;& A^2 \inf_{\widehat{f}} \sup_{f \in \mathcal{F}} P(|\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0)| \geq s_0) \\ \geq \;& A^2 \inf_{\widehat{f}} \max_{f \in \{g_0, g_1\}} P(|\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0)| \geq s_0). \end{aligned}$$

And we further define $p_e := \inf_{\widehat{f}} \max_{f \in \{g_0, g_1\}} P(|\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0)| \geq s_0)$

Step 3. Define $K(P_{g_0}, P_{g_1}) = E_{P_{g_0}}(\log(dP_{g_0}/dP_{g_1}))$, where P_{g_j} is the distribute function of $y_i(\mathbf{d}_0)$ with $f_i(\mathbf{d}_0) = g_j(\mathbf{d}_0)$ for $j = 0, 1$. Then we have

$$\begin{aligned} K(P_{g_0}, P_{g_1}) &= \sum_{\mathbf{d}_0 \in \mathcal{D}_0} \int \log \frac{\phi(x; 0, \sigma^2(\mathbf{d}_0))}{\phi(x; g_1(\mathbf{d}_0), \sigma^2(\mathbf{d}_0))} \phi(x; 0, \sigma^2(\mathbf{d}_0)) dx \\ &= \sum_{\mathbf{d}_0 \in \mathcal{D}_0} \frac{g_1(\mathbf{d}_0)^2}{2\sigma^2(\mathbf{d}_0)} \\ &= \frac{1}{2\sigma^2} L^2 h^2 \sum_{\mathbf{d}_0 \in \mathcal{D}_0} \prod_{j=1}^3 H^2\left(\frac{d_j - d_{j,0}}{h}\right) \\ &\leq \frac{1}{2\sigma^2} L^2 h^2 e^{-6} \sum_{\mathbf{d}_0 \in \mathcal{D}_0} \prod_{j=1}^3 I(|d_j - d_{j,0}| \leq h/2) \\ &\leq \frac{1}{2\sigma^2} L^2 h^2 e^{-6} N_D h^3 \\ &\leq \frac{1}{2\sigma^2} L^2 e^{-6} N_D h^5. \end{aligned}$$

Recall that $h = c_0(\sigma^2 s^2 n^{-1} N_D^{-1})^{1/5}$, then one can see that $K(P_{g_0}, P_{g_1}) \leq 2L^2 e^{-6} s^2 n^{-1} c_0^5 := \alpha$.

Consequently, by lemma A.8.1, we have $p_e \geq \max\{e^{-\alpha}/4, (1-\sqrt{\alpha/2})/2\}$. Hence,

$$\liminf_{n\to\infty} \inf_{\widehat{f}} \sup_{f\in\mathcal{F}} E\{\psi^{-2}\|\widehat{f}_i(\mathbf{d}_0) - f_i(\mathbf{d}_0)\|^2\} \geq \delta_1,$$

where $\delta_1 = A^2 \max\{e^{-\alpha}/4, (1-\sqrt{\alpha/2})/2\} = 8^{-1}L^2 \prod_{j=1}^{3} H(0)[\alpha/(L^2e^{-6})]^{2/5} \max\{e^{-\alpha}/2, 1-\sqrt{\alpha/2}\}$. Because c_1 is free of i and $\mathbf{d}_0$, we have

$$\inf_h \mathcal{R}_n(\widehat{f}_h) \geq \mathcal{R}_n^* \geq \delta_1\psi^2 \asymp (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5}$$

Combing the lower and the upper bound, we have

$$(\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \lesssim \inf_h \mathcal{R}_n(\widehat{f}_h) \lesssim (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \vee \sigma^2 s^2 n^{-1}$$

If $N_D \lesssim (\sigma^{-2}s^{-2}n)^{\frac{3}{2}}$, we have $\inf_h \mathcal{R}_n(\widehat{f}_h) \asymp \mathcal{R}_n^* \asymp (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5}$. This completes the proof of Theorem 3.3.2.

A.4 Proof of Theorem 3.3.3

Without loss of generality, assume that $\mathbf{d}_0 \in \Omega$. Define the event

$$E_0 = \{\max_{\mathbf{d}_m \in B(\mathbf{d}_0,h)} \|\widehat{\boldsymbol{\beta}}(\mathbf{d}_m,\lambda) - \boldsymbol{\beta}^*(\mathbf{d}_m)\|_2 \leq t\}$$

. When the event $\mathcal{E} \cap E_0$ is true, for $\mathbf{d}_m \in \Omega \bigcap B(\mathbf{d}_0,h)$ and $j = 1, \cdots p$, we have

$$\begin{aligned}
&|\widehat{\beta}_j(\mathbf{d}_m,\lambda) - \widehat{\beta}_j(\mathbf{d}_0,\lambda)| \\
&\leq |\beta_j^*(\mathbf{d}_m) - \beta_j^*(\mathbf{d}_0)| + |\widehat{\beta}_j(\mathbf{d}_m,\lambda) - \beta_j^*(\mathbf{d}_m)| + |\widehat{\beta}_j(\mathbf{d}_0,\lambda) - \beta_j^*(\mathbf{d}_0)| \\
&\leq C_0 h + 2\|\widehat{\boldsymbol{\beta}}(\mathbf{d}_0,\lambda) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_2 \\
&\leq C_0 h + 2t,
\end{aligned}$$

and for $\mathbf{d}_m \in \Omega^C \bigcap B(\mathbf{d}_0, h)$, there exists $\mathbf{d} \in \partial\Omega$ such that

$$\begin{aligned}
&|\widehat{\beta}_j(\mathbf{d}_m, \lambda) - \widehat{\beta}_j(\mathbf{d}_0, \lambda)| \\
&\geq |\beta_j^*(\mathbf{d}_m) - \beta_j^*(\mathbf{d}_0)| - |\widehat{\beta}_j(\mathbf{d}_m, \lambda) - \beta_j^*(\mathbf{d}_m)| - |\widehat{\beta}_j(\mathbf{d}_0, \lambda) - \beta_j^*(\mathbf{d}_0)| \\
&\geq |\beta_{j,\Omega}^*(\mathbf{d}) - \beta_{j,\Omega^C}^*(\mathbf{d})| - |\beta_{j,\Omega^C}^*(\mathbf{d}_m) - \beta_{j,\Omega^C}^*(\mathbf{d})| - |\beta_{j,\Omega}^*(\mathbf{d}_0) - \beta_{j,\Omega}^*(\mathbf{d})| \\
&- 2\|\widehat{\boldsymbol{\beta}}(\mathbf{d}_0, \lambda) - \boldsymbol{\beta}^*(\mathbf{d}_0)\|_2 \\
&\geq 1/C_0 - 2C_0 h - 2t.
\end{aligned}$$

Since $h = o(1)$, we can pick, for example, $h_y = \frac{1}{2C_0}$ and $t = \frac{1}{6C_0}$ such that

$$C_0 h + 2t < h_y < 1/C_0 - 2C_0 h - 2t,$$

then the YF estimator equals the estimator under the oracle case. Consequently,

$$\begin{aligned}
& E(\widehat{f}_{i,h,h_y}^{\text{YF}}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 \\
= \;& E\{(\widehat{f}_{i,h,h_y}^{\text{YF}}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 1_{\mathcal{E}}\} + E\{(\widehat{f}_{i,h,h_y}^{\text{YF}}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 1_{\mathcal{E}^C}\} \\
= \;& E\{(\widehat{f}_{i,h,h_y}^{\text{YF}}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 1_{E_0 \cap \mathcal{E}}\} + E\{(\widehat{f}_{i,h,h_y}^{\text{YF}}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 1_{E_0^C \cap \mathcal{E}}\} \\
+ \;& E\{(\widehat{f}_{i,h,h_y}^{\text{YF}}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 1_{\mathcal{E}^C}\} \\
\leq \;& E(\widehat{f}_i(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 + P(E_0^C \cap \mathcal{E}) + P(\mathcal{E}^C)
\end{aligned}$$

For $P(E_0^C \cap \mathcal{E})$, by lemma A.1.1,we have

$$P(\|\widehat{\boldsymbol{\beta}}(\mathbf{d}_m, \lambda) - \boldsymbol{\beta}^*(\mathbf{d}_m)\|_2 \geq t) \leq P(\|\nabla \mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m))\|_\infty \geq Ct/\sqrt{s})$$

for some $C > 0$, which implies

$$P(\max_{\mathbf{d}_m \in B(\mathbf{d}_0, h)} \|\widehat{\boldsymbol{\beta}}(\mathbf{d}_m, \lambda) - \boldsymbol{\beta}^*(\mathbf{d}_m)\|_2 \geq t) \leq P(\max_{\mathbf{d}_m \in B(\mathbf{d}_0, h)} \|\nabla \mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m))\|_\infty \geq C_0 t/\sqrt{s}).$$

Then by assumption 3.3, for $t = \frac{1}{6C_0} = O(1)$, we take $\sigma \lesssim \sqrt{n/s \log(pN_D)}$, $\delta = h^3$ and we have $P(E_0^C \cap \mathcal{E}) \leq \delta = o((\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5})$. We also have $P(\mathcal{E}^C) = o((\sigma^2 s n^{-1} N_D^{-1})^{2/5})$ in the proofs of Theorem 3.3.2. Therefore, we have

$$E(\widehat{f}_{i,h,h_y}^{\mathrm{YF}}(\mathbf{d}_0) - f_i^*(\mathbf{d}_0))^2 \lesssim E(\widehat{f}_{i,h}(\mathbf{d}_0) - f_i^*(\mathbf{d}_0))^2.$$

Further we have $\inf_{h,h_y} \mathcal{R}(\widehat{f}_{h,h_y}^{\mathrm{YF}}) \asymp \inf_h \mathcal{R}_n(\widehat{f}_h)$. This completes the proof of Theorem 3.3.3.

A.5 Proof of Theorem 3.3.4

In fact, when $h_{\mathcal{P}} = 1/N_D^{1/3}$, the NLM method degenerates into the YF method. Hence, as long as the conditions in Theorem 3.3.3 is assumed, the claims of Theorem 3.3.4 is true.

A.6 Validation of Assumption 3.9

In this section, we use the examples of L_2 loss and Huber loss functions to validate assumption 3.9.

A.6.1 Example 1: Least Square estimator

For any $\mathbf{d}_m \in B(\mathbf{d}_0, h)$, we have $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m) - \boldsymbol{\beta}^*(\mathbf{d}_m) = \nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m))$. Therefore,

$$\begin{aligned}\max_{\mathbf{d}_m \in B(\mathbf{d}_0,h)} \|\bar{\beta}_{\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{\mathcal{P}_{\mathbf{d}_m}}\|_\infty &= \max_{\mathbf{d}_m \in B(\mathbf{d}_0,h)} \|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k \in \mathcal{P}_0} \widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m + \mathbf{d}_k) - \boldsymbol{\beta}^*(\mathbf{d}_m + \mathbf{d}_k))\|_\infty \\ &= \max_{\mathbf{d}_m \in B(\mathbf{d}_0,h)} \|m_{\mathcal{P}}^{-1} \sum_{\mathbf{d}_k \in \mathcal{P}_0} \nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m + \mathbf{d}_k))\|_\infty.\end{aligned} \tag{A.6.1}$$

Denote the j th element of $\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m + \mathbf{d}_k))$ as $\nabla\mathcal{L}^{(j)}(\boldsymbol{\beta}^*(\mathbf{d}_m + \mathbf{d}_k))$, $a = (m_{\mathcal{P}}^{-1}, \cdots, m_{\mathcal{P}}^{-1})^T \in \mathcal{R}^{m_{\mathcal{P}}}$ and $b \in \mathcal{R}^{m_{\mathcal{P}}}$ with each element to be $\nabla\mathcal{L}^{(j)}(\boldsymbol{\beta}^*(\mathbf{d}_m + \mathbf{d}_k))$ for some $\mathbf{d}_k \in \mathcal{P}_0$. Then for any $s > 0$ we have

$$\begin{aligned}&\mathbb{P}\left\{|m_{\mathcal{P}}^{-1} \sum_{\mathbf{d}_k \in \mathcal{P}_0} \nabla\mathcal{L}^{(j)}(\boldsymbol{\beta}^*(\mathbf{d}_m + \mathbf{d}_k))| > t\right\} \leq \frac{Ee^{sa^t b}}{e^{st}} \\ &= e^{-st} \prod_{i=1}^{m_{\mathcal{P}}} Ee^{sb_i/m_{\mathcal{P}}}.\end{aligned}$$

By Assumption 3.3 and lemma 1.5 in Rigollet and Hütter (2015),

$$Ee^{sb_i/m_{\mathcal{P}}} \le e^{2\sigma^2 s^2/(nm_{\mathcal{P}}^2)}.$$

Therefore,

$$\mathbb{P}\left\{|m_{\mathcal{P}}^{-1}\sum_{\mathbf{d}_k\in\mathcal{P}_0}\nabla\mathcal{L}^{(j)}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k))|>t\right\} \le e^{-st+2\sigma^2 s^2/(nm_{\mathcal{P}})}.$$

And by minimizing s, we have

$$\mathbb{P}\left\{|m_{\mathcal{P}}^{-1}\sum_{\mathbf{d}_k\in\mathcal{P}_0}\nabla\mathcal{L}^{(j)}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k))|>t\right\} \le e^{-nm_{\mathcal{P}}t^2/(8\sigma^2)}.$$

Taking the union bound over $j=1,\cdots,p$ and $\mathbf{d}_m\in B(\mathbf{d}_0,h)$, we have

$$\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}\sum_{\mathbf{d}_k\in\mathcal{P}_0}\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k))\|_\infty \lesssim \sigma\sqrt{\log(pN_Dh^3/\delta)/(nm_{\mathcal{P}})} \tag{A.6.2}$$

holds with probability at least $1-\delta$. We take $\delta=h^3$ in (A.6.2), and we have

$$\begin{aligned}
&\mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k\in\mathcal{P}_0}\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)\|_\infty\right.\\
&\left.\gtrsim \sigma\frac{s+\log(m_{\mathcal{P}}N_D)}{n}+\sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})}\right\}\\
&\le \mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k\in\mathcal{P}_0}\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)\|_\infty\gtrsim\sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})}\right\}\\
&\le \delta=o(h^2)=o((\sigma^2s^2n^{-1}N_D^{-1})^{2/5}).
\end{aligned}$$

A.6.2 Example 2: Robust regression estimator

By Lemma A.8.3, we have for any $\mathbf{d}_m$

$$\mathbb{P}\left\{\|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m)-\boldsymbol{\beta}^*(\mathbf{d}_m)-\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m))\|_2\ge C_3\sigma\frac{s+t}{n}\right\}\le 3e^{-t}.$$

Therefore,

$$\begin{aligned}
&\mathbb{P}\left\{\|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k\in\mathcal{P}_0}\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)-\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)))\|_2\geq C_3\sigma\frac{s+t}{n}\right\}\\
&\leq 1-\mathbb{P}\left\{\|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)-\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k))\|_2\leq C_3\sigma\frac{s+t}{n}\right\}^{m_{\mathcal{P}}}\\
&=1-(1-3e^{-t})^{m_{\mathcal{P}}}\\
&\lesssim m_{\mathcal{P}}e^{-t}.
\end{aligned}$$

Taking the union bound over $\mathbf{d}_m\in B(\mathbf{d}_0,h)$, we have

$$\begin{aligned}
&\mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k\in\mathcal{P}_0}\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)-\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)))\|_2\geq C_3\sigma\frac{s+t}{n}\right\}\\
&\lesssim m_{\mathcal{P}}N_Dh^3e^{-t}.
\end{aligned}\tag{A.6.3}$$

We take $t=\log(m_{\mathcal{P}}N_D)$ in (A.6.3) and $\delta=h^3$ in (A.6.2). Combine the two equations together and we have

$$\begin{aligned}
&\mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|\bar{\beta}_{\mathcal{P}_{\mathbf{d}_m}}-\bar{\beta}^*_{\mathcal{P}_{\mathbf{d}_m}}\|_\infty\gtrsim\sigma\frac{s+\log(m_{\mathcal{P}}N_D)}{n}+\sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})}\right\}\\
&=\mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k\in\mathcal{P}_0}\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)\|_\infty\right.\\
&\left.\gtrsim\sigma\frac{s+\log(m_{\mathcal{P}}N_D)}{n}+\sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})}\right\}\\
&\leq\mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}(\sum_{\mathbf{d}_k\in\mathcal{P}_0}\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m+\mathbf{d}_k)-\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)-\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k)))\|_2\right.\\
&\left.\gtrsim\sigma\frac{s+\log(m_{\mathcal{P}}N_D)}{n}\right\}\\
&+\mathbb{P}\left\{\max_{\mathbf{d}_m\in B(\mathbf{d}_0,h)}\|m_{\mathcal{P}}^{-1}\sum_{\mathbf{d}_k\in\mathcal{P}_0}\nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m+\mathbf{d}_k))\|_\infty\gtrsim\sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})}\right\}\\
&\leq\delta+h^3\\
&=o(h^2)=o((\sigma^2s^2n^{-1}N_D^{-1})^{2/5}).
\end{aligned}\tag{A.6.4}$$

A.7 Proof of Theorem 3.3.5

For arbitrary $\mathbf{d}_0 \in \mathcal{D}_0$, if $\mathcal{P}_{\mathbf{d}_0} \cap \partial\Omega \neq \emptyset$, we have $E(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f_{*,i}(\mathbf{d}_0))^2 \leq 1$. If $\mathcal{P}_{\mathbf{d}_0} \cap \partial\Omega = \emptyset$, without loss of generality we assume that $\mathcal{P}_{\mathbf{d}_0} \subset \Omega$. For any $\mathbf{d}_m \in B(\mathbf{d}_0, h)$ and $j = 1, \cdots p$, we have

$$\bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_0}} = \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}} + \bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} + \bar{\beta}_{j,\mathcal{P}_{\mathbf{d}_0}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}},$$

where $\bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}}$ is the average of $\beta^*_j(\mathbf{d}_m)$ within patch $\mathcal{P}_{\mathbf{d}_m}$ and we assume that all patches are of same size with $m_{\mathcal{P}} \asymp N_D h^3_{\mathcal{P}}$. Define $\mathcal{P}_0$ the generic patch centered at 0 and

$$E_{\mathcal{P}_{\mathbf{d}_m}} = \{\max_{\mathbf{d}_m \in B(\mathbf{d}_0,h)} \|\bar{\beta}_{\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{\mathcal{P}_{\mathbf{d}_m}}\|_\infty \leq \zeta := \sigma \frac{s + \log(m_{\mathcal{P}} N_D)}{n} + \sigma\sqrt{\log(pN_D)/(nm_{\mathcal{P}})}\}.$$

By assumption 3.9, We have $P(E^C_{\mathcal{P}_{\mathbf{d}_0}}) = o(\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5})$. Moreover, we can rewrite $\bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}}$ as

$$\bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}} = m_{\mathcal{P}}^{-1} \sum_{\mathbf{d}_k \in \mathcal{P}_0} (\beta^*_j(\mathbf{d}_0 + \mathbf{d}_k) - \beta^*_j(\mathbf{d}_m + \mathbf{d}_k)).$$

If $\mathbf{d}_m \in \Omega$ with $\mathcal{P}_{\mathbf{d}_m} \subset \Omega$, then by the C_0-Lipschitz assumption, we have

$$|\bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}}| \leq C_0 \|\mathbf{d}_m - \mathbf{d}_0\|_\infty \leq C_0 h.$$

If $\mathbf{d}_m \in \Omega^C$, then there exists a point $\mathbf{d} \in B(\mathbf{d}_0) \cap \partial\Omega$, such that for $\mathbf{d}_k \in \Omega$, $\beta^*_j(\mathbf{d}_k) = \beta^*_{j,\Omega}(\mathbf{d}) + (\beta^*_{j,\Omega}(\mathbf{d}_k) - \beta^*_{j,\Omega}(\mathbf{d}))$ with $|\beta^*_{j,\Omega}(\mathbf{d}_k) - \beta^*_{j,\Omega}(\mathbf{d})| \leq C_0 h$, and for $\mathbf{d}_k \in \Omega^C$, $\beta^*_j(\mathbf{d}_k) =$

$\beta^*_{j,\Omega^C}(\mathbf{d}) + (\beta^*_{j,\Omega^C}(\mathbf{d}_k) - \beta^*_{j,\Omega^C}(\mathbf{d}))$ with $|\beta^*_{j,\Omega^C}(\mathbf{d}_k) - \beta^*_{j,\Omega^C}(\mathbf{d})| \le C_0 h$. Hence

$$\begin{aligned}
&\bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}} \\
&= \beta^*_{j,\Omega}(\mathbf{d}) + m_{\mathcal{P}}^{-1} \sum_{\mathbf{d}_k \in \mathcal{P}_{\mathbf{d}_m}} [\beta^*_{j,\Omega}(\mathbf{d}_k) - \beta^*_{j,\Omega}(\mathbf{d})] \\
&- \beta^*_{j,\Omega}(\mathbf{d}) \frac{|\mathcal{P}_{\mathbf{d}_0} \cap \Omega|}{|\mathcal{P}_{\mathbf{d}_0}|} - m_{\mathcal{P}}^{-1} \sum_{\mathbf{d}_k \in \mathcal{P}_{\mathbf{d}_0} \cap \Omega} [\beta^*_{j,\Omega}(\mathbf{d}_k) - \beta^*_{j,\Omega}(\mathbf{d})] \\
&- \beta^*_{j,\Omega^C}(\mathbf{d}) \frac{|\mathcal{P}_{\mathbf{d}_0} \cap \Omega^C|}{|\mathcal{P}_{\mathbf{d}_0}|} - m_{\mathcal{P}}^{-1} \sum_{\mathbf{d}_k \in \mathcal{P}_{\mathbf{d}_0} \cap \Omega^C} [\beta^*_{j,\Omega^C}(\mathbf{d}_k) - \beta^*_{j,\Omega^C}(\mathbf{d})] \\
&= (\beta^*_{j,\Omega}(\mathbf{d}) - \beta^*_{j,\Omega^C}(\mathbf{d})) \frac{|\mathcal{P}_{\mathbf{d}_0} \cap \Omega^C|}{|\mathcal{P}_{\mathbf{d}_0}|} + R
\end{aligned}$$

where $|R| \le 2C_0 h$. Further by lemma A.8.2, we know that $|\mathcal{P}_{\mathbf{d}_m} \cap \Omega^C|/|\mathcal{P}_{\mathbf{d}_m}| \ge (2C_0)^{-3}$. Then we have $|\bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_m}} - \bar{\beta}^*_{j,\mathcal{P}_{\mathbf{d}_0}}| \ge 2^{-3}C_0^{-4} - 2C_0 h$. Since $h \to 0$, $C_0 \asymp 1$ and $\zeta \to 0$ by assumption 3.6, we can select $h_y \asymp C_0^{-4} \asymp 1$ such that

$$C_0 h + 2\zeta + \le h_y \le 2^{-3} C_0^{-4} - 2C_0 h - 2\zeta. \tag{A.7.1}$$

This can ensure that all voxels $\mathbf{d}_m \in B(\mathbf{d}_0, h)$ such that $\mathcal{P}_{\mathbf{d}_m} \subset \Omega$ are included in the neighborhood of $\mathbf{d}_0$, while no voxels in Ω^C are included under the event $E_{\mathcal{P}_{\mathbf{d}_0}}$. Define $B_{\mathbf{d}_0} = \{\mathbf{d} \in \mathcal{D}_0 : \mathbf{d} \in B(\mathbf{d}_0, h)\}$, $B^0_{\mathbf{d}_0} = \{\mathbf{d} \in \mathcal{D}_0 : \mathbf{d} \in B(\mathbf{d}_0, h), \mathcal{P}_{\mathbf{d}} \subset \Omega\}$, and $A_{\mathbf{d}_0} = \cup_{j=1,\cdots,p}\{\mathbf{d} \in \mathcal{D}_0 : K^{\text{adj}}_{j,h,h_y}(\mathbf{d}_0 - \mathbf{d}) > 0\}$. Since under $E_{\mathcal{P}_{\mathbf{d}_0}}$, we have $B^0_{\mathbf{d}_0} \subset A_{\mathbf{d}_0}$, which implies $E_{\mathcal{P}_{\mathbf{d}_0}} \subset \{B^0_{\mathbf{d}_0} \subset A_{\mathbf{d}_0}\} \subset \cup_{B^0_{\mathbf{d}_0} \subset A \subset B_{\mathbf{d}_0}} \{A_{\mathbf{d}_0} = A\}$ and $1_{E_{\mathcal{P}_{\mathbf{d}_0}}} \le \sum_{B^0_{\mathbf{d}_0} \subset A \subset B_{\mathbf{d}_0}} 1_{A_{\mathbf{d}_0} = A}$, and further

$$\begin{aligned}
&E(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 \\
&= E\{(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 1_{\mathcal{E}}\} + E\{(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 1_{\mathcal{E}^C}\} \\
&= E\{(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 1_{\{E_{\mathcal{P}_{\mathbf{d}_0}}\} \cap \mathcal{E}}\} + E\{(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 1_{\{E^C_{\mathcal{P}_{\mathbf{d}_0}}\} \cap \mathcal{E}}\} \\
&+ E\{(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 1_{\mathcal{E}^C}\} \\
&\le \sum_{B^0_{\mathbf{d}_0} \subset A \subset B_{\mathbf{d}_0}} P(A_{\mathbf{d}_0} = A) E(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 + P(E^C_{\mathcal{P}_{\mathbf{d}_0}}) + P(\mathcal{E}^C).
\end{aligned}$$

By lemma A.8.2, $B^0_{\mathbf{d}_0}$ contains a ball of radius $\asymp h$. Therefore, $|B^0_{\mathbf{d}_0}|/|B_{\mathbf{d}_0}| \asymp 1$ and we have

$$\begin{aligned}
&\sum_{B^0_{\mathbf{d}_0} \subset A \subset B_{\mathbf{d}_0}} P(A_{\mathbf{d}_0} = A) E(\widehat{f}^{\text{NLM}}_{i,h,h_y}(\mathbf{d}_0) - f^*_i(\mathbf{d}_0))^2 \\
&\leq \sum_{B^0_{\mathbf{d}_0} \subset A \subset B_{\mathbf{d}_0}} P(A_{\mathbf{d}_0} = A) O(h^2 + \sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 \vee \sigma^2 s^2 n^{-1}) \\
&\leq O(\sigma^2 s^2 n^{-1} N_D^{-1} h^{-3} + h^2 \vee \sigma^2 s^2 n^{-1}).
\end{aligned}$$

We also have $P(E^C_{\mathcal{P}_{\mathbf{d}_0}}) = o(h^2)$ and $P(\mathcal{E}^C) = o(h^2)$. Hence we get that

$$\text{MSE}_f(\widehat{f}^{\text{NLM}}_{h,h_y}) \leq \frac{|Q|}{N_D} + O(\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \vee \sigma^2 s^2 n^{-1},$$

where $Q = \{\mathbf{d} : \mathcal{P}_{\mathbf{d}} \cap \partial\Omega \neq \emptyset\}$ is a subset of $\{\mathbf{d} : dist(\mathbf{d}, \partial\Omega) < h_{\mathcal{P}}\}$, hence we have $|Q| \leq C_2 N_D h_{\mathcal{P}}$. Consequently,

$$\text{MSE}_f(\widehat{f}^{\text{NLM}}_{h,h_y}) \leq O(h_{\mathcal{P}} + (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} \vee \sigma^2 s^2 n^{-1}).$$

Optimizing the right term of above formula over $h_{\mathcal{P}}$ in ζ subject to (A.7.1), one can obtain that

$$\begin{aligned}
\inf_{h,h_y} \mathcal{R}(\breve{f}^{\text{NLM}}_{h,h_y}) &\leq O(\frac{(\sigma^2 n^{-1} \log p N_D)^{1/3}}{N_D^{1/3}} \vee (\sigma^2 s^2 n^{-1} N_D^{-1})^{2/5} + \sigma^2 s^2 n^{-1}) \\
&= O(\frac{(\sigma^2 n^{-1} \log p N_D)^{1/3}}{N_D^{1/3}} \vee \sigma^2 s^2 n^{-1}).
\end{aligned}$$

This completes the proof of Theorem 3.3.5.

A.8 Auxiliary Results

This section collects some auxiliary results.

Lemma A.8.1. (Theorem 2.2 (iii) of Tsybakov (2008).) Let P_0 and P_1 be two probability measures. If the KL divergence $K(P_0, P_1) \leq \alpha < \infty$, then

$$p_{e,1} := \frac{1}{2}\left(P_0\left(\psi^* \neq 0\right) + P_1\left(\psi^* \neq 1\right)\right) \geq \max\left(\frac{1}{4}\exp(-\alpha), \frac{1 - \sqrt{\alpha/2}}{2}\right),$$

where

$$\psi^* = \begin{cases} 0, & \text{if } p_0 \geq p_1 \\ 1, & \text{otherwise} \end{cases}$$

and p_0 and p_1 are the densities of P_0 and P_1.

Lemma A.8.2. (Lemma 8.2 of Arias-Castro et al. (2012).) Let $\phi : \mathcal{R}^3 \to \mathcal{R}^3$ is injective with both ϕ, ϕ^{-1} C-Lipschitz. The for $\Omega = \phi(B(0,1))$ and

$$\rho(A) := \inf_{h \in (0,1)} \inf_{x \in A} \sup \left\{ \frac{\operatorname{Vol}(B(y,s))}{\operatorname{Vol}(B(x,h))} : B(y,s) \subset B(x,h) \cap A \right\},$$

where Vol denotes the volume. We have $\min\left(\rho(\Omega), \rho\left(\Omega^c\right)\right) \geq (2C)^{-3}$.

Lemma A.8.3. (Theorem 7 (II) of Sun et al. (2020).) For $t > 0$ and $\tau_0 \geq \sigma$, the estimator $\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m)$ with $\tau = \tau_0\sqrt{n/(s+t)}$ satisfies

$$\mathbb{P}\left\{ \|\widehat{\boldsymbol{\beta}}^0(\mathbf{d}_m) - \boldsymbol{\beta}^*(\mathbf{d}_m) - \nabla\mathcal{L}(\boldsymbol{\beta}^*(\mathbf{d}_m))\|_2 \geq C_3\sigma\frac{s+t}{n} \right\} \leq 3e^{-t}$$

provided $n \geq C_2(s+t)$.

APPENDIX B: TECHNICAL DETAILS OF CHAPTER 5

The appendix collects the proofs of Lemmas 5.2.1, 5.2.2 and Theorems 5.2.3 and 5.2.4.

B.1 Proof of Lemma 5.2.1

Proof of Lemma 5.2.1. Define

$$\alpha := \varphi_{\min}(\boldsymbol{U}_1^T \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T(\boldsymbol{X}\boldsymbol{B} + \boldsymbol{Z}\boldsymbol{D})\boldsymbol{V}_1) = \varphi_{\min}(\boldsymbol{\Sigma}_1 + \boldsymbol{U}_1^T \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}\boldsymbol{V}_1) \asymp \varphi_k(\boldsymbol{Z}\boldsymbol{D}),$$

$$\beta := \|\boldsymbol{U}_\perp^T \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T(\boldsymbol{X}\boldsymbol{B} + \boldsymbol{Z}\boldsymbol{D})\boldsymbol{V}_\perp\| = \|\boldsymbol{U}_\perp^T \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}\boldsymbol{V}_\perp\| \le \gamma\varphi_1(\boldsymbol{X}\boldsymbol{B}),$$

$$z_{21} := \|\mathbb{P}_{\boldsymbol{U}_{1\perp}} \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}\mathbb{P}_{\boldsymbol{V}_1}\| = \|\boldsymbol{U}_{1\perp}\boldsymbol{U}_{1\perp}^T \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}\boldsymbol{V}_1\boldsymbol{V}_1^T\|,$$

and

$$z_{12} := \|\mathbb{P}_{\boldsymbol{U}_1} \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}\mathbb{P}_{\boldsymbol{V}_{1\perp}=}\| = \|\boldsymbol{U}_1\boldsymbol{U}_1^T \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}\boldsymbol{V}_{1\perp}\boldsymbol{V}_{1\perp}^T\|.$$

We rewrite $\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B}$ as

$$\boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{X}\boldsymbol{B} = \gamma \begin{bmatrix} \mathbf{I_q} \\ \mathbf{0} \end{bmatrix} \boldsymbol{\Sigma}_{XB}\boldsymbol{R}_B^T.$$

We further observe that there exist $\boldsymbol{W}_0 \in \mathbb{R}^{k\times k}$ and $\boldsymbol{W}_0' \in \mathbb{O}_{k,k}$ such that

$$\boldsymbol{U}_1 = \boldsymbol{D}_\gamma \begin{bmatrix} \boldsymbol{\alpha}_1^T \\ \boldsymbol{\beta}_1^T \end{bmatrix} \boldsymbol{W}_0,$$

and

$$\boldsymbol{V}_1 = \boldsymbol{R}_D\boldsymbol{W}_0',$$

with $\|\boldsymbol{W}_0\| \in [1, 1/\sqrt{1-\|\boldsymbol{\alpha}_1\|^2}]$. Then we have

$$\begin{aligned} z_{12} &\le \gamma^2\|\boldsymbol{W}_0^T\boldsymbol{\alpha}_1\|\varphi_1(\boldsymbol{X}\boldsymbol{B}) \\ &= O(\gamma^2\frac{\|\boldsymbol{\alpha}_1\|}{\sqrt{1-\|\boldsymbol{\alpha}_1\|^2}}\varphi_1(\boldsymbol{X}\boldsymbol{B})), \end{aligned}$$

and

$$\begin{aligned} z_{21} &\leq \gamma \varphi_1(\boldsymbol{XB}) \|\boldsymbol{\tau}\| \\ &= O(\gamma \varphi_1(\boldsymbol{XB}) \|\boldsymbol{\tau}\|). \end{aligned}$$

Therefore, following the Theorem 1 in Tony Cai and Zhang (2017) and we have

$$\begin{aligned} \|\sin \Theta(\boldsymbol{U}_1, \widetilde{\boldsymbol{U}}_1)\| &\leq \frac{\alpha z_{21} + \beta z_{12}}{\alpha^2 - \beta^2 - z_{21}^2 \wedge z_{12}^2} \wedge 1 \\ &= O(\frac{\gamma^3 \|\boldsymbol{\alpha}_1\| \varphi_1^2(\boldsymbol{XB})}{\sqrt{1 - \|\boldsymbol{\alpha}_1\|^2 \varphi_q^2(\boldsymbol{ZD})}} + \frac{\gamma \|\boldsymbol{\tau}\| \varphi_1(\boldsymbol{XB})}{\varphi_q(\boldsymbol{ZD})}). \end{aligned}$$

□

B.2 Proof of Lemma 5.2.2

Proof of Lemma 5.2.2. For convenience, denote $\boldsymbol{Y}_0 = \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{Y}$, $\boldsymbol{X}_0 = \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T (\boldsymbol{XB} + \boldsymbol{ZD})$, $\boldsymbol{M}_3 = \boldsymbol{D}_\gamma \boldsymbol{Q}_X^T \boldsymbol{E}$ and we have $\mathbb{E} \boldsymbol{Y}_0 \boldsymbol{Y}_0^T = \boldsymbol{X}_0 \boldsymbol{X}_0^T + p \boldsymbol{D}_\gamma^2$. Since

$$\boldsymbol{X}_0 \boldsymbol{X}_0^T = \begin{bmatrix} \widetilde{\boldsymbol{U}}_1 & \widetilde{\boldsymbol{U}}_{1\perp} \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\boldsymbol{\Sigma}}_1^2 & \boldsymbol{0} \\ \boldsymbol{0} & \widetilde{\boldsymbol{\Sigma}}_2^2 \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\boldsymbol{U}}_1^\top \\ \widetilde{\boldsymbol{U}}_{1\perp}^\top \end{bmatrix},$$

it yields $\mathbb{E} \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 = \widetilde{\boldsymbol{\Sigma}}_1^2 + \boldsymbol{N}_1$ and $\varphi_{\min}\left(\widetilde{\boldsymbol{\Sigma}}_1^2 + \boldsymbol{N}_1\right) \leq \varphi_k^2(\boldsymbol{X}_0) + p$, where $\boldsymbol{N}_1 = p \widetilde{\boldsymbol{U}}_1^T \boldsymbol{D}_\gamma^2 \widetilde{\boldsymbol{U}}_1$. Let $\boldsymbol{N} = (\widetilde{\boldsymbol{\Sigma}}_1^2 + \boldsymbol{N}_1)^{-\frac{1}{2}}$, then we have $\mathbb{E} \boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} = \boldsymbol{I}_k$. It follows that

$$\begin{aligned} \varphi_k^2(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1) &\geq \varphi_k^2(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N}) \left\{\varphi_k^2(\boldsymbol{X}_0) + p\right\} \\ &= \left\{\varphi_k^2(\boldsymbol{X}_0) + p\right\} \\ &\cdot \varphi_k \left(\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} - \mathbb{E} \boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} + \mathbb{E} \boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N}\right) \\ &\geq \left\{\varphi_k^2(\boldsymbol{X}_0) + p\right\} \left\{1 - \varphi_1 \left(\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} - \boldsymbol{I}_k\right)\right\} \\ &= \left\{\varphi_k^2(\boldsymbol{X}_0) + p\right\} \left(1 - \|\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} - \boldsymbol{I}_k\|_2\right) \end{aligned} \tag{B.2.1}$$

Step 1. we prove that

$$\mathbb{P}\left[\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1) \geq (1-x)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}\right] \geq 1 - c\cdot\exp\left[Ck - c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x \wedge x^2\right],$$

for some constants C and c.

For any unit vector $\boldsymbol{\mu}\in\mathbb{R}^k$, we have

$$\begin{aligned}
&\boldsymbol{\mu}^\top\boldsymbol{N}^\top\widetilde{\boldsymbol{U}}_1^\top\boldsymbol{Y}_0\boldsymbol{Y}_0^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{I}_k\boldsymbol{\mu} = \boldsymbol{\mu}^\top\boldsymbol{N}^\top\widetilde{\boldsymbol{U}}_1^\top\boldsymbol{Y}_0\boldsymbol{Y}_0^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu} - \mathbb{E}\boldsymbol{\mu}^T\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{Y}_0\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu}\\
&= \boldsymbol{\mu}^\top\boldsymbol{N}^\top\widetilde{\boldsymbol{U}}_1^\top\boldsymbol{X}_0\boldsymbol{X}_0^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu} - \mathbb{E}\boldsymbol{\mu}^T\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{X}_0\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu}\\
&+ 2\boldsymbol{\mu}^\top\boldsymbol{N}^\top\widetilde{\boldsymbol{U}}_1^\top\boldsymbol{X}_0\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu} - 2\mathbb{E}\boldsymbol{\mu}^\top\boldsymbol{N}^\top\widetilde{\boldsymbol{U}}_1^\top\boldsymbol{X}_0\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu}\\
&+ \boldsymbol{\mu}^\top\boldsymbol{N}^\top\widetilde{\boldsymbol{U}}_1^\top\boldsymbol{X}_0\boldsymbol{X}_0^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu} - \mathbb{E}\boldsymbol{\mu}^T\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu}\\
&= 2(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu} + (\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T\left[\boldsymbol{M}_3\boldsymbol{M}_3^T - \mathbb{E}(\boldsymbol{M}_3\boldsymbol{M}_3^T)\right](\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu}).
\end{aligned}$$

Then for fixed unit vector $\boldsymbol{\mu}\in\mathbb{R}^k$, we vectorized $\boldsymbol{M}_3$ into $\overrightarrow{\boldsymbol{e}}\in\mathbb{R}^{np}$ as follows,

$$\overrightarrow{\boldsymbol{e}} = \left(m_{11}, m_{21}, \ldots m_{n1}, m_{12}, m_{22}, \ldots m_{n2} \ldots m_{1p} \ldots m_{np}\right)^T.$$

We also repeat $(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T$ block for p times and introduce

$$\overrightarrow{\boldsymbol{D}} = \begin{bmatrix} (\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T & & \\ & \ddots & \\ & & (\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T \end{bmatrix} \in \mathbb{R}^{(np)\times(np)}$$

and we can see that

$$\|\overrightarrow{\boldsymbol{D}}\| = \|(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T\| = \|\boldsymbol{N}\boldsymbol{\mu}\|_2^2 \leq \|\boldsymbol{N}\|^2 = \left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}^{-1},$$

and

$$\|\overrightarrow{\boldsymbol{D}}\|_F^2 = p\|(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\boldsymbol{\mu})^T\|_F^4 = p\|\boldsymbol{N}u\|_2^2 \leq p\|\boldsymbol{N}\|^4 = p\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}^{-2}.$$

By Hanson-Wright Inequality (Theorem 1 in Rudelson et al. (2013)), we have

$$\begin{aligned}&\mathbb{P}\{|(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\left[\boldsymbol{M}_3\boldsymbol{M}_3^T-\mathbb{E}(\boldsymbol{M}_3\boldsymbol{M}_3^T)\right](\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)|>x\} = \mathbb{P}\{|\overrightarrow{\boldsymbol{e}}^T\overrightarrow{\mathrm{D}}\overrightarrow{\boldsymbol{e}}-\mathbb{E}\overrightarrow{\boldsymbol{e}}^T\overrightarrow{\mathrm{D}}\overrightarrow{\boldsymbol{e}}|>x\}\\&\leq 2\exp\left(-c\min\left(\frac{x^2\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)^2}{pa^4},\frac{x\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)}{a^2}\right)\right),\end{aligned} \tag{B.2.2}$$

where a is the upper bound of the sub-gaussian norm of m_{ij}. Next, we bound

$$(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u = \mathrm{tr}\left\{\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\right\} = \overrightarrow{\boldsymbol{e}}^T\mathrm{Vec}\left\{\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\right\}.$$

We know that $\varphi_1\left(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}\right)\leq 1$ and

$$\begin{aligned}\|\mathrm{Vec}\{\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\}\|_2^2 &= \|\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\|_F^2 = \|\boldsymbol{N}u\|_2^2\cdot\|(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\|_2^2\\&\leq\|\boldsymbol{N}\|^2\leq\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}^{-1}.\end{aligned}$$

Since $\overrightarrow{\boldsymbol{e}}^T\mathrm{Vec}\left\{\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\right\}$ can be viewed as the summation of the independent sub-gaussian random variables, we have

$$\mathbb{P}\left(|(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u|>x\right)\leq C\exp\left[-cx^2\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}\right]. \tag{B.2.3}$$

Combining equations (B.2.2) and (B.2.3) implies

$$\begin{aligned}
&\mathbb{P}(|u^T\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\mathbf{Y}_0\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u-u^T\boldsymbol{I}_qu|>x)\\
&=\mathbb{P}\Big(\{|(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\left[\boldsymbol{M}_3\boldsymbol{M}_3^T-\mathbb{E}(\boldsymbol{M}_3\boldsymbol{M}_3^T)\right](\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)|>x/2\}\\
&\cup\{2|(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u|>x/2\}\Big)\\
&\leq\mathbb{P}\left(\{|(\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\left[\boldsymbol{M}_3\boldsymbol{M}_3^T-\mathbb{E}(\boldsymbol{M}_3\boldsymbol{M}_3^T)\right](\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)|>x/2\}\right)\\
&+\mathbb{P}\left(\{2|(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u)^T\boldsymbol{M}_3^\top\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u|>x/2\}\right)\\
&\leq c\cdot\exp\left[-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\right].
\end{aligned}$$

.

Next, the ϵ-net argument (Lemma 5 in Tony Cai and Zhang (2017)) leads to

$$\mathbb{P}(|\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\mathbf{Y}_0\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}-\boldsymbol{I}_k|>x)\leq c\cdot\exp\left[Ck-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\right].$$

With equation (B.2.1), we finally have

$$\mathbb{P}\left[\varphi_k^2(\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_1)\geq(1-x)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}\right]\geq 1-c\cdot\exp\left[Ck-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\right] \tag{B.2.4}$$

Step 2. We prove that

$$\mathbb{P}\left[\varphi_{k+1}\left(\mathbf{Y}_0\right)>\left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)+p\right\}(1+x)\right]\leq c\cdot\exp\left[C(n-k)-c\left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\right]$$

for some constants C and c. We first use the fact that

$$\varphi_{k+1}\left(\mathbf{Y}_0\right)=\varphi_{k+1}\left(\mathbf{Y}_0^T\right)=\min_{\operatorname{rank}(B)=k}\|\mathbf{Y}_0^T-B\|\leq\|\mathbf{Y}_0^T-\mathbf{Y}_0^T[\widetilde{\boldsymbol{U}}_1,\mathbf{0}]\|=\varphi_1(\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}),$$

$$\varphi_1^2(\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp})=\varphi_1(\widetilde{\boldsymbol{U}}_{1\perp}^T\mathbf{Y}_0\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp})=\varphi_1(\widetilde{\boldsymbol{U}}_{1\perp}^T\mathbf{Y}_0\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}-\mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T\mathbf{Y}_0\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}+\mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T\mathbf{Y}_0\mathbf{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}),$$

and

$$\mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp} = \widetilde{\boldsymbol{\Sigma}}_2^2 + p\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{D}_\gamma^2 \widetilde{\boldsymbol{U}}_{1\perp}.$$

Therefore, we have

$$\begin{aligned}\varphi_1^2(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}) &\leq \varphi_1(\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}) + \varphi_1(\mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}) \\ &= \varphi_{k+1}^2\left(\boldsymbol{X}_0\right) + p + \|\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\|.\end{aligned} \tag{B.2.5}$$

Let $\boldsymbol{N}' = \{\widetilde{\boldsymbol{\Sigma}}_2^2 + p\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{D}_\gamma^2 \widetilde{\boldsymbol{U}}_{1\perp}\}^{-\frac{1}{2}}$, then

$$\|\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\| \leq \left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right) + p\right\} \|(\boldsymbol{N}')^T \widetilde{\boldsymbol{U}}_{1\perp}^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp} \boldsymbol{N}' - \boldsymbol{I}_{n-k}\|$$

Finally, following the same procedure in Step 1 and replacing $\boldsymbol{N}$ by $\boldsymbol{N}'$, we have

$$\mathbb{P}\left[\varphi_{k+1}\left(\boldsymbol{Y}_0\right) > \left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right) + p\right\}(1+x)\right] \leq c \cdot \exp\left[C(n-k) - c\left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right) + p\right\} x \wedge x^2\right] \tag{B.2.6}$$

Step3. In this step, we study $\mathbb{P}_{\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1} \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}$. Note that

$$\begin{aligned}\|\mathbb{P}_{\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1} \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\| &= \|\mathbb{P}_{\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N}} \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\| \\ &= \|(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N})\left\{(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N})^T (\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N})\right\}^{-1} (\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N})^T \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\| \\ &\leq \varphi_{\min}^{-1}(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N}) \|\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\|.\end{aligned}$$

We analyze $\varphi_{\min}(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N})$ and $\|\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\|$ separately. For $\varphi_{\min}(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N})$, we have

$$\varphi_{\min}^2(\boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N}) = \varphi_{\min}(\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N}) \geq 1 - \|\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{Y}_0 \boldsymbol{Y}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} - \boldsymbol{I}_k\|.$$

Recall that

$$\mathbb{P}(|\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{Y}_0\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}-\boldsymbol{I}_k|>x)\leq c\cdot\exp\left[Ck-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\right],$$

it follows that

$$\mathbb{P}\{\varphi_{\min}^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N})\geq 1-x\}\geq 1-c\cdot\exp\left[Ck-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\right].$$

Set $x=1/2$, we could choose C_{gap} large enough such that whenever $\varphi_k^2\left(\boldsymbol{X}_0\right)\geq C_{gap}k$, we can have

$$Ck-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}x\wedge x^2\leq-\frac{c}{8}\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}.$$

Under this setting,

$$\mathbb{P}\{\varphi_{\min}^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N})\geq\frac{1}{2}\}\geq 1-C\cdot\exp\left[-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}\right]. \tag{B.2.7}$$

For $\|\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{Y}_0\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|$, since $\widetilde{\boldsymbol{U}}_1^T\boldsymbol{X}_0\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_{1\perp}=0$, we have the following decomposition

$$\begin{aligned}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{Y}_0\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}&=\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T(\boldsymbol{X}_0+\boldsymbol{M}_3)(\boldsymbol{X}_0+\boldsymbol{M}_3)^T\widetilde{\boldsymbol{U}}_{1\perp}\\&=\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{X}_0\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}+\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}+\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\end{aligned} \tag{B.2.8}$$

For the first term in (B.2.8),

$$u_1^T\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{X}_0\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}u_2=\overrightarrow{\boldsymbol{e}}^T\mathrm{Vec}\left\{\widetilde{\boldsymbol{U}}_{1\perp}u_2(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u_1)^T\right\},$$

for any unit vectors $u_1\in\mathbb{R}^k$ and $u_2\in\mathbb{R}^{n-k}$. Since

$$\|\mathrm{Vec}\left\{\widetilde{\boldsymbol{U}}_{1\perp}u_2(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u_1)^T\right\}\|_2^2=\|\widetilde{\boldsymbol{U}}_{1\perp}u_2(\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u_1)^T\|_F^2=\|\widetilde{\boldsymbol{U}}_{1\perp}u_2\|_2^2\cdot\|\boldsymbol{X}_0^T\widetilde{\boldsymbol{U}}_1\boldsymbol{N}u_1\|_2^2\leq 1,$$

it implies

$$\mathbb{P}(|u_1^T \boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{X}_0 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp} u_2| \geq x) \leq C \exp(-cx^2 / \|\widetilde{\boldsymbol{U}}_{1\perp} u_2 (\boldsymbol{X}_0^T \widetilde{\boldsymbol{U}}_1 \boldsymbol{N} u_1)^T\|_F^2) \leq C \exp(-cx^2).$$

Similarly, for the third term in (B.2.8),

$$\mathbb{P}(|u_1^T \boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{X}_0^T \widetilde{\boldsymbol{U}}_{1\perp} u_2| \geq x) \leq C \exp(-cx^2).$$

Therefore, by ϵ-net argument, we have

$$\mathbb{P}(\|\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{X}_0 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp}\| \geq x) \leq C \exp(Cn - cx^2), \tag{B.2.9}$$

and

$$\mathbb{P}(\|\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{X}_0^T \widetilde{\boldsymbol{U}}_{1\perp}\| \geq x) \leq C \exp(Cn - cx^2). \tag{B.2.10}$$

Now we study the second term $\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp}$. Note that $\mathbb{E}\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp} = p\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{D}_\gamma^2 \widetilde{\boldsymbol{U}}_{1\perp}$. Then we can similarly have

$$\begin{aligned}
&\mathbb{P}(|u_1^T (\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp}) u_2| \geq x) \\
&\leq C \exp\left[-c \min\left\{\frac{x^2 \{\varphi_k^2(\boldsymbol{X}_0) + p\}}{pa^4}, \frac{x\sqrt{\varphi_k^2(\boldsymbol{X}_0) + p}}{a^2}\right\}\right]
\end{aligned}$$

and ϵ-net argument leads to

$$\begin{aligned}
&\mathbb{P}(\|(\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\boldsymbol{N}^T \widetilde{\boldsymbol{U}}_1^T \boldsymbol{M}_3 \boldsymbol{M}_3^T \widetilde{\boldsymbol{U}}_{1\perp})\| \geq x) \\
&\leq C \exp\left[Cn - c \min\left\{\frac{x^2 \{\varphi_k^2(\boldsymbol{X}_0) + p\}}{pa^4}, \frac{x\sqrt{\varphi_k^2(\boldsymbol{X}_0) + p}}{a^2}\right\}\right]
\end{aligned} \tag{B.2.11}$$

To bound $\|\mathbb{E}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}\|$, we define

$$\begin{aligned}t := \|\mathbb{E}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}\| &= \|p\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{D}_\gamma^2\widetilde{\boldsymbol{U}}_{1\perp}\| \\ &\leq p\|\boldsymbol{N}\| = \frac{p}{\sqrt{\varphi_k^2\left(\boldsymbol{X}_0\right)+p}}.\end{aligned}$$

Therefore, for $x > t$

$$\begin{aligned}&\mathbb{P}(\|(\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}\| \geq x) \\ &\leq \mathbb{P}(\|\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}\| \geq x) \\ &\leq \mathbb{P}(\|\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}\| + \|\mathbb{E}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp})\| \geq x) \\ &\leq \mathbb{P}(\|\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp} - \mathbb{E}\boldsymbol{N}^T\widetilde{\boldsymbol{U}}_1^T\boldsymbol{M}_3\boldsymbol{M}_3^T\widetilde{\boldsymbol{U}}_{1\perp}\| \geq x - t).\end{aligned} \tag{B.2.12}$$

By combining equations (B.2.8)-(B.2.12), we have

$$\begin{aligned}&\mathbb{P}(\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\| \leq x) \\ &\geq 1 - C\cdot\exp\left[-c\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}\right] - C\exp\left\{Cn - c\min\left(t_1, t_2, t_3\right)\right\}.\end{aligned} \tag{B.2.13}$$

for $x > t$, where $t_1 = x^2$, $t_2 = p^{-1}(x-t)^2\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\}$, and $t_3 = (x-t)\sqrt{\varphi_k^2\left(\boldsymbol{X}_0\right)+p}$. Through Steps 1-3, we have proved three important inequalities in equations (B.2.4), (B.2.6), and (B.2.13).

Step 4.1: We first focus on the scenario that $\varphi_k^2\left(\boldsymbol{X}_0\right) \geq C_{gap}(\sqrt{np}+n)$ for some large constant C_{gap}. For equation (B.2.4), we take

$$x = \frac{\varphi_k^2\left(\boldsymbol{X}_0\right)}{3\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}},$$

and rewrite it as

$$\begin{aligned}&\mathbb{P}\left\{\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1) \le \varphi_k^2\left(\boldsymbol{X}_0\right) + p - \frac{1}{3}\varphi_k^2\left(\boldsymbol{X}_0\right)\right\} \\ &\le C \cdot \exp\left[Ck - c\min\left\{\varphi_k^2\left(\boldsymbol{X}_0\right), \frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{3\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}\right\}\right].\end{aligned} \tag{B.2.14}$$

and for equation (B.2.6), we take

$$x = \frac{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)}{3\left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)+p\right\}},$$

and it becomes

$$\begin{aligned}&\mathbb{P}\left\{\varphi_{k+1}\left(\boldsymbol{Y}_0\right) > \varphi_k^2\left(\boldsymbol{X}_0\right) + p + \frac{1}{3}\varphi_k^2\left(\boldsymbol{X}_0\right)\right\} \\ &\le c \cdot \exp\left[C(n-k) - c\min\left\{\varphi_k^2\left(\boldsymbol{X}_0\right), \frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{3\left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)+p\right\}}\right\}\right].\end{aligned} \tag{B.2.15}$$

When C_{gap} is large enough, it holds for some C that

$$\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+p} \ge \frac{C_{gap}^2(\sqrt{np}+n)^2}{C_{gap}(\sqrt{np}+n)+p} \ge Cn.$$

Then we have

$$\begin{aligned}&c\min\left[\varphi_k^2\left(\boldsymbol{X}_0\right), \frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{3\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}\right] - Ck = c\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+p} - Ck \\ &\ge c\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+p} - Cn \ge \frac{c}{2}\cdot\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+p}\end{aligned}$$

and

$$\begin{aligned}&c\min\left[\varphi_k^2\left(\boldsymbol{X}_0\right), \frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{3\left\{\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)+p\right\}}\right] - C(n-k) \\ &\ge c\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)+p} - Cn \\ &\ge c\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+p} - Cn \ge \frac{c}{2}\cdot\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right)+p}.\end{aligned}$$

For equation (B.2.13), we take

$$x = -2t + \sqrt{\varphi_k^2\left(\boldsymbol{X}_0\right) + p},$$

then $\min(t_1, t_2, t_3) = t_1 = \left\{-2t + \sqrt{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}\right\}^2$ and we have

$$\begin{aligned} c\min\left(t_1, t_2, t_3\right) - Cn &= c\left\{-2t + \sqrt{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}\right\}^2 - Cn \\ &\geq \frac{c}{2}\cdot\left\{\varphi_k^2\left(\boldsymbol{X}_0\right) + p\right\} \\ &\geq \frac{c}{2}\cdot\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}. \end{aligned}$$

To sum up, we denote the event $\boldsymbol{Q}$ as

$$\begin{aligned} \boldsymbol{Q} = \Big\{&\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1) \geq \varphi_k^2\left(\boldsymbol{X}_0\right) + p - \frac{1}{3}\varphi_k^2\left(\boldsymbol{X}_0\right), \\ &\varphi_{k+1}^2\left(\boldsymbol{Y}_0\right) \leq \varphi_{k+1}^2\left(\boldsymbol{X}_0\right) + p + \frac{1}{3}\varphi_k^2\left(\boldsymbol{X}_0\right), \\ &\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\| \leq -2t + \sqrt{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}\Big\}. \end{aligned}$$

When $\varphi_k^2\left(\boldsymbol{X}_0\right) \geq C_{gap}(\sqrt{np} + n)$ for some large constant C_{gap}, we have

$$\mathbb{P}\left(\boldsymbol{Q}^c\right) \leq C\exp\left\{-c\frac{\varphi_k^4\left(\boldsymbol{X}_0\right)}{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}\right\}$$

By the basic properties of exponential functions,

$$\mathbb{P}\left(\boldsymbol{Q}^c\right) \leq C\frac{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}{\varphi_k^4\left(\boldsymbol{X}_0\right)} \leq C\frac{(n + t^2)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right) + p\right\}}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right) - \varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2}. \tag{B.2.16}$$

Under event $\boldsymbol{Q}$, we can use the Proposition 1 in Tony Cai and Zhang (2017) and obtain

$$\begin{aligned}
\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1\right)\|^2 &\leq \frac{\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1)\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|^2}{\left\{\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1) - \varphi_{k+1}^2(\boldsymbol{Y}_0)\right\}^2} \leq \frac{\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1)\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|^2}{\left\{\varphi_k^2(\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1) - \varphi_{k+1}^2(\boldsymbol{Y}_0)\right\}^2} \\
&\leq C\frac{\left\{\frac{2}{3}\varphi_k^2\left(\boldsymbol{X}_0\right) + p\right\}\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|^2}{\left\{\frac{1}{3}\varphi_k^2\left(\boldsymbol{X}_0\right) - \varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2} \\
&= C\frac{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right) + p\right\}\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|^2}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right) - \varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2}.
\end{aligned}$$

Here we use the fact that $x^2/\left(x^2 - y^2\right)^2$ is a decreasing function of x and increasing function of y for $x > y \geq 0$. Next, note that $\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1\right)\| \leq 1$, it follows that

$$\begin{aligned}
&\mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1\right)\|^2 \\
&= \mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1\right)\|^2\mathbf{1}_{\boldsymbol{Q}} + \mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1\right)\|^2\mathbf{1}_{\boldsymbol{Q}^c} \\
&\leq C\frac{\varphi_k^2\left(\boldsymbol{X}_0\right) + p}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right) - \varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2}\mathbb{E}\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|^2\mathbf{1}_{\boldsymbol{Q}} + \mathbb{P}(\boldsymbol{Q}^c).
\end{aligned} \tag{B.2.17}$$

It remains to study $\mathbb{E}\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|^2\mathbf{1}_{\boldsymbol{Q}}$. Denote $\boldsymbol{T}=\|\mathbb{P}_{\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_1}\boldsymbol{Y}_0^T\widetilde{\boldsymbol{U}}_{1\perp}\|$ and we have

$$
\begin{aligned}
&\mathbb{E}\boldsymbol{T}^2\mathbf{1}_{\boldsymbol{Q}}\leq\mathbb{E}\boldsymbol{T}^2\mathbf{1}_{\left\{\boldsymbol{T}^2\leq\left(\sqrt{\varphi_k^2(\boldsymbol{X}_0)+p}-2t\right)^2\right\}}\\
&=\int_0^\infty\mathbb{P}\left(\boldsymbol{T}^2\mathbf{1}_{\left\{\boldsymbol{T}^2\leq\left(\sqrt{\left(\varphi_k^2(\boldsymbol{X}_0)+p\right)}-2t\right)^2\right\}}\geq u\right)du\\
&\leq\left(\sqrt{C_xn}+t\right)^2+\int_{\left(\sqrt{C_xn}+t\right)^2}^{\left(\sqrt{\left(\varphi_k^2(\boldsymbol{X}_0)+p\right)}-2t\right)^2}\mathbb{P}\left(\boldsymbol{T}^2\mathbf{1}_{\left\{\boldsymbol{T}^2\leq\left(\sqrt{\left(\varphi_k^2(\boldsymbol{X}_0)+p\right)}-2t\right)^2\right\}}\geq u\right)du\\
&\leq\left(\sqrt{C_xn}+t\right)^2+\int_{\left(\sqrt{C_xn}+t\right)^2}^{\left(\sqrt{\left(\varphi_k^2(\boldsymbol{X}_0)+p\right)}-2t\right)^2}C\cdot\exp\left(-c\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)\right)du\\
&+C\exp\left(Cn\right)\cdot\int_{\left(\sqrt{C_xn}+t\right)^2}^{\left(\sqrt{\left(\varphi_k^2(\boldsymbol{X}_0)+p\right)}-2t\right)^2}\exp\left(-cu\right)du\\
&\leq\left(\sqrt{C_xn}+t\right)^2+C\left(\sqrt{\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)}\right)^2\cdot\exp\left(-c\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)\right)\qquad\text{(B.2.18)}\\
&+C\exp\left(Cn\right)\cdot\int_{\left(\sqrt{C_xn}+t\right)^2}^{\left(\sqrt{\left(\varphi_k^2(\boldsymbol{X}_0)+p\right)}\right)^2}\exp\left(-cu\right)du\\
&\leq\left(\sqrt{C_xn}+t\right)^2+C\left(\sqrt{\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)}\right)^2\cdot\exp\left(-c\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)\right)\\
&+\frac{C}{c}\exp\left(Cn\right)\cdot\exp\left(-c\left(\sqrt{C_xn}+t\right)^2\right)\\
&\leq\left(\sqrt{C_xn}+t\right)^2+C\left(\sqrt{\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)}\right)^2\cdot\exp\left(-c\left(\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right)\right)\\
&+\frac{C}{c}\exp\left(Cn\right)\cdot\exp\left(-cC_xn\right)\\
&\leq\left(\sqrt{C_xn}+t\right)^2+C+\frac{C}{c}\exp\left(Cn\right)\cdot\exp\left(-cC_xn\right).
\end{aligned}
$$

As can be seen that we can choose C_x large enough, but only relying on other constants C,c in the inequalities above, to ensure

$$\mathbb{E}\boldsymbol{T}^2\mathbf{1}_{\boldsymbol{Q}}\leq C(n+t^2).\tag{B.2.19}$$

Now combine equations (B.2.16), (B.2.17), (B.2.19) and the trivial upper bound 1, we obtain

$$\mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1,\widehat{\boldsymbol{U}}_1\right)\|^2 \le C\frac{(n+t^2)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)-\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2}\wedge 1.$$

Step 4.2. We next study the case that $\varphi_k^2\left(\boldsymbol{X}_0\right) \le C_{gap}(\sqrt{np}+n)$. Note that

$$\begin{aligned}
&\frac{(n+t^2)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)-\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2} \ge \frac{(n+t^2)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}{\varphi_k^4\left(\boldsymbol{X}_0\right)}\\
&\ge \frac{(n+t^2)\left\{C_{gap}(\sqrt{np}+n)+p\right\}}{C_{gap}^2(2n\sqrt{np}+n^2+np)}\\
&\ge \frac{C_{gap}(\sqrt{np}+n)+p}{C_{gap}^2(2\sqrt{np}+n+p)}\\
&\ge \frac{C_{gap}(\sqrt{np}+n)+\frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma})}{C_{gap}^2(2\sqrt{np}+n+p)} \ge \min(1,\frac{c}{C_{gap}^2}).
\end{aligned}$$

Then

$$\mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1,\widehat{\boldsymbol{U}}_1\right)\|^2 \le 1 \le C\frac{(n+t^2)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)-\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2}\wedge 1.$$

In summary, we have

$$\mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1,\widehat{\boldsymbol{U}}_1\right)\|^2 \le C\frac{(n+t^2)\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)+p\right\}}{\left\{\varphi_k^2\left(\boldsymbol{X}_0\right)-\varphi_{k+1}^2\left(\boldsymbol{X}_0\right)\right\}^2}\wedge 1$$

regardless of $\varphi_k^2\left(\boldsymbol{X}_0\right)$. Finally, by assumption 5.2, we have

$$\mathbb{E}\|\sin\Theta\left(\widetilde{\boldsymbol{U}}_1,\widehat{\boldsymbol{U}}_1\right)\|^2 \le C\frac{n}{\varphi_k^2\left(\boldsymbol{X}_0\right)} \asymp \frac{n}{\varphi_k^2(\boldsymbol{Z}\boldsymbol{D})} = r_2^2.$$

□

B.3 Proof of Theorem 5.2.3

Proof of Theorem 5.2.3. Followed by the triangle inequality of $\sin\Theta$. □

B.4 Proof of Theorem 5.2.4

Proof of Theorem 5.2.4. By the Lemma 1 in Tony Cai and Zhang (2017), there exist $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ such that

$$\|(\widetilde{\boldsymbol{U}}_1 - \boldsymbol{U}_1\boldsymbol{W}_1)\| \asymp \|\sin\Theta(\boldsymbol{U}_1, \widetilde{\boldsymbol{U}}_1)\| = O_p(r_1)$$

and

$$\|(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1\boldsymbol{W}_2)\| \asymp \|\sin\Theta(\widetilde{\boldsymbol{U}}_1, \widehat{\boldsymbol{U}}_1)\| = O_p(r_2).$$

Take $\boldsymbol{W} = \boldsymbol{W}_1\boldsymbol{W}_2$ and we have

$$\begin{aligned}
&\|\boldsymbol{Q}_X\boldsymbol{D}_\gamma^{-1}(\widehat{\boldsymbol{U}}_1 - \boldsymbol{U}_1\boldsymbol{W})\| = \|\boldsymbol{Q}_X\boldsymbol{D}_\gamma^{-1}(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1\boldsymbol{W}_2)\| + \|\boldsymbol{Q}_X\boldsymbol{D}_\gamma^{-1}(\widetilde{\boldsymbol{U}}_1 - \boldsymbol{U}_1\boldsymbol{W}_1)\boldsymbol{W}_2\| \\
&\le 1/\gamma\|(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1\boldsymbol{W}_2)_{[1:k,:]}\| + \|(\widehat{\boldsymbol{U}}_1 - \widetilde{\boldsymbol{U}}_1\boldsymbol{W}_2)_{[(k+1):n,:]}\| + 1/\gamma\|(\widetilde{\boldsymbol{U}}_1 - \boldsymbol{U}_1\boldsymbol{W}_1)\| \\
&= O_p(r_1/\gamma + r_2) = O_p(\sqrt{\frac{n}{\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})}} + \frac{\gamma^2\|\boldsymbol{\alpha}_1\|\varphi_1^2(\boldsymbol{X}\boldsymbol{B})}{\sqrt{1-\|\boldsymbol{\alpha}_1\|^2}\varphi_q^2(\boldsymbol{Z}\boldsymbol{D})} + \frac{\|\boldsymbol{\tau}\|\varphi_1(\boldsymbol{X}\boldsymbol{B})}{\varphi_q(\boldsymbol{Z}\boldsymbol{D})}).
\end{aligned}$$

□

REFERENCES

Aganj, I., Lenglet, C., and Sapiro, G. (2009). ODF reconstruction in q-ball imaging with solid angle consideration. In *Proceedings of the 6th International Symposium on Biomedical Imaging*, pages 84–91.

Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., and Harel, N. (2010). Reconstruction of the orientation distribution function in single and multiple shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine* **2,** 554–566.

Alexander, D., Barker, G., and Arridge, S. (2002). Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data,. *Magnetic Resonance in Medicine* **48,** 3319340,.

Alexander, D. C. (2005). Multiple-fiber reconstruction algorithms for diffusion mri. *Annals of the New York Academy of Sciences* **1064,** 113–133.

Arias-Castro, E., Salmon, J., and Willett, R. (2012). Oracle inequalities and minimax rates for non-local means and related adaptive kernel-based methods. *SIAM Journal on Imaging Sciences* **5,** 944–992.

Assemlal, H., Tschumperle, D., and Brun, L. (2008). Efficient computation of pdf-based characteristics from diffusion mr signal. *Medical Image Computing and Computer-Assisted Intervention MICCAI,* pages 70–78.

Assemlal, H., Tschumperle, D., and Brun, L. (2009). Efficient and robust computation of pdf features from diffusion mr signal. *Medical ImMedical Image Analysis,* **13,** 715–729.

Assemlal, H.-E., Tschumperlé, D., and Brun, L. (2009). Efficient and robust computation of PDF features from diffusion MR signal. *Medical Image Analysis* **13,** 715–729.

Awate, S. P. and Whitaker, R. T. (2006). Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28,** 364–376.

Azzabou, N., Paragios, N., and Guichard, F. (2007). Image denoising based on adapted dictionary computation. *2007 IEEE International Conference on Image Processing* pages 109–112.

Basser, P. J., Mattiello, J., and LeBihan, D. (1994). Mr diffusion tensor spectroscropy and imaging. *Biophysical Journal* **66,** 259–267.

Becker, S., Tabelow, K., Mohammadi, S., Weiskopf, N., and Polzehl, J. (2014). Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by mspoas. *NeuroImage* **95,** 90–105.

Benito, M., Parker, J., Du, Q., Wu, J., Xiang, D., Perou, C. M., and Marron, J. S. (2004). Adjustment of systematic microarray data biases. *Bioinformatics* **20,** 105–114.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* **57,** 289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of statistics* pages 1165–1188.

Bijsterbosch, J., Harrison, S. J., Jbabdi, S., Woolrich, M., Beckmann, C., Smith, S., and Duff, E. P. (2020). Challenges and future directions for representations of functional brain organization. *Nature neuroscience* **23,** 1484–1495.

Bookheimer, S. Y., Salat, D. H., Terpstra, M., Ances, B. M., Barch, D. M., Buckner, R. L., Burgess, G. C., Curtiss, S. W., Diaz-Santos, M., Elam, J. S., et al. (2019). The lifespan human connectome project in aging: an overview. *NeuroImage* **185,** 335–348.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning,* **3,** 1122.

Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., Thompson, M. R., and Venkatesan, R. (2014). *Magnetic Resonance Imaging: Physical Principles and Sequence Design.* Wiley-Blackwell.

Buades, A., Coll, B., and Morel, J. M. (2005). A review of image denoising algorithms, with a new one. *Multiscale Model Simul* **4,** 490–530.

Buchanan, C. R., Pernet, C. R., Gorgolewski, K. J., Storkey, A. J., and Bastin, M. E. (2014). Test-retest reliability of structural brain networks from diffusion MRI. *NeuroImage* **86,** 231–243.

Buja, A. and Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate behavioral research* **27,** 509–540.

Caye, K. and Francois, O. (2018). Lfmm 2.0: Latent factor models for confounder adjustment in genome and epigenome-wide association studies. *Biorxiv* page 255893.

Chang, J., Meng, G., Wang, L., Xiang, S., and Pan, C. (2018). Deep self-evolution clustering,. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pages 5879–5887,.

Chang, J., Wang, L., Meng, G., Xiang, S., and Pan, C. (2017). Deep adaptive image clustering,. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pages 5879–5887,.

Cheng, J., Deriche, R., Jiang, T., Shen, D., and Yap, P.-T. (2014). Non-Negative Spherical Deconvolution (NNSD) for estimation of fiber Orientation Distribution Function in single-/multi-shell diffusion MRI. *NeuroImage* **101,** 750–764.

Cheng, J., Ghosh, A., Deriche, R., and Jiang, T. (2010). Model-Free, Regularized, Fast, and Robust Analytical Orientation Distribution Function Estimation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'10)*, volume 6361, pages 648–656.

Cheng, J., Ghosh, A., Jiang, T., and Deriche, R. (2010a). Model-free and analytical eap reconstruction via spherical polar fourier diffusion mri. *Medical Image Computing and Computer-Assisted Intervention MICCAI,* pages 70–78.

Cheng, J., Ghosh, A., Jiang, T., and Deriche, R. (2010b). Model-free and analytical eap reconstruction via spherical polar fourier diffusion mri. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 590–597. Springer.

Cheng, J. and Zhu, H. (2016). Handbook of neuroimaging data analysis,. *PLoS Biology,* .

Cheng, J.and Merlet, S., Ghosh, A., Caruyer, E., Jiang, T., and Deriche, R. (2011). Compressive sensing ensemble average propagator estimation via l1 spherical polar fourier imaging. *Medical Image Computing and Computer-Assisted Intervention MICCAI,* .

Chung, M. K., Adluru, N., Lee, J. E., Lazar, M., Lainhart, J. E., and Alexander, A. L. (2010). Cosine series representation of 3d curves and its application to white matter fiber bundles in diffusion tensor imaging. *Statistics and its interface* **3,** 69.

Ciccarelli, O., Parker, G., Toosy, A., Wheeler-Kingshott, C., Barker, G., Boulby, P., Miller, D., and Thompson, A. (2003). From diffusion tractography to quantitative white matter tract measures: a reproducibility study. *Neuroimage* **18,** 348–359.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment* **6,** 284.

Clarke, S., Hall, P., et al. (2009). Robustness of multiple testing procedures against dependence. *The Annals of Statistics* **37,** 332–358.

Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., and Barillot, C. (2008). An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. *IEEE transactions on medical imaging* **27,** 425–441.

Cousineau, M., Jodoin, P.-M., Garyfallidis, E., Côté, M.-A., Morency, F. C., Rozanski, V., Grand'Maison, M., Bedell, B. J., and Descoteaux, M. (2017). A test-retest study on Parkinson's PPMI dataset yields statistically significant white matter fascicles. *NeuroImage: Clinical* **16,** 222–233.

Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing* **16,** 2080–2095.

de Reus, M. A. and van den Heuvel, M. P. (2013). The parcellation-based connectome: limitations and extensions. *Neuroimage* **80,** 397–404.

Descoteaux, M. (2008). *High Angular Resolution Diffusion MRI: from Local Estimation to Segmentation and Tractograpghy.* PhD thesis, INRIA Sophia Antipolis.

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2007a). Regularized, Fast and Robust Analytical Q-ball Imaging. *Magnetic Resonance in Medicine* **58,** 497–510.

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2007b). Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **58,** 497–510.

Dong, W., Zhang, L., Shi, G., and Li, X. (2012). Nonlocally centralized sparse representation for image restoration. *IEEE transactions on Image Processing* **22,** 1620–1630.

Donnell, L. and Westin, C.-F. (2007). Automatic tractography segmentation using a high--dimensional white matter atlas,. *IEEE Transactions in Medical Imaging,* **26,** 32–45,.

Durrleman, S., Fillard, P., Pennec, X., Trouvé, A., and Ayache, N. (2011). Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage* **55,** 1073–1090.

Durston, S. (2010). Imaging genetics in adhd. *NeuroImage* **53,** 832–838.

Efron, B. (2010). Correlated z-values and the accuracy of large-scale statistical estimates. *Journal of the American Statistical Association* **105,** 1042–1055.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *Annals of Statistics,* **32,** 407–499.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications*. Chapman and Hall/CRC.

Fan, J. and Han, X. (2016). Estimation of the false discovery proportion with unknown dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* .

Fan, J., Ke, Y., Sun, Q., and Zhou, W.-X. (2017). Farm-test: Factor-adjusted robust multiple testing with false discovery control. *arXiv preprint arXiv:1711.05386* .

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association* **96,** 1348–1360.

Fan, L., Zhang, F., Fan, H., and Zhang, C. (2019). Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art* **2,** 1–12.

Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., Reisert, M., Sakaie, K., Tensaouti, F., Yo, T., Mangin, J., and Poupon, C. (2011). Quantitative evaluation of 10 tractography algorithms on a realistic diffusion mr phantom,. *NeuroImage,,* **56,** 2209234,.

Fischl, B. (2012). Freesurfer. *Neuroimage* **62,** 774–781.

Fornito, A., Zalesky, A., and Breakspear, M. (2013). Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage* **80,** 426–444.

Fortin, J.-P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P. J., et al. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* **167,** 104–120.

Fortin, J.-P., Parker, D., Tunc, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., et al. (2017). Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* **161,** 149–170.

Fortin, J.-P., Sweeney, E. M., Muschelli, J., Crainiceanu, C. M., Shinohara, R. T., and ADNI (2016). Removing inter-subject technical variability in magnetic resonance imaging studies. *NeuroImage* **132,** 198–212.

Frichot, E., Schoville, S. D., Bouchard, G., and François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular biology and evolution* **30,** 1687–1699.

Friguet, C., Kloareg, M., and Causeur, D. (2009). A factor model approach to multiple testing under dependence. *Journal of the American Statistical Association* **104,** 1406–1415.

Fromer, M., Roussos, P., Sieberts, S. K., Johnson, J. S., Kavanagh, D. H., Perumal, T. M., Ruderfer, D. M., Oh, E. C., Topol, A., Shah, H. R., et al. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature neuroscience* **19,** 1442.

Gagnon-Bartsch, J. A., Jacob, L., and Speed, T. P. (2017). Removing unwanted variation: Exploiting negative controls for high dimensional data analysis. *IMS Monographs. Cambridge: Cambridge University Press* page in press.

Gagnon-Bartsch, J. A. and Speed, T. P. (2012). Using control genes to correct for unwanted variation in microarray data. *Biostatistics* **13,** 539–552.

Garyfallidis, E., Brett, M., Correia, M. M., Williams, G. B., and Nimmo-Smith, I. (2012). Quickbundles, a method for tractography simplification. *Frontiers in neuroscience* **6,**.

Garyfallidis, E., Brett, M., and Nimmo-Smith, I. (2010). Fast dimensionality reduction for brain tractography clustering. In *16th Annual Meeting of the Organization for Human*

Brain Mapping.

Garyfallidis, E., Côté, M.-A., Rheault, F., Sidhu, J., Hau, J., Petit, L., Fortin, D., Cunanne, S., and Descoteaux, M. (2017). Recognition of white matter bundles using local and global streamline-based registration and clustering. *NeuroImage* .

Garyfallidis, E., Ocegueda, O., Wassermann, D., and Descoteaux, M. (2015). Robust and efficient linear registration of white-matter fascicles in the space of streamlines. *NeuroImage* **117,** 124 – 140.

Gerard, D. and Stephens, M. (2017). Unifying and generalizing methods for removing unwanted variation based on negative controls. *arXiv preprint arXiv:1705.08393* .

Girard, G., Whittingstall, K., Deriche, R., and Descoteaux, M. (2014). Towards quantitative connectivity analysis: reducing tractography biases. *NeuroImage* **98,** 266 – 278.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. *Nature* **536,** 171–178.

Glasser, M. F., Smith, S. M., and et al (2016). The human connectome project's neuroimaging approach. *Nature Neuroscience* **19,** 1175–1187.

Guevara, M., Guevara, P., Roman, C., and Mangin, J. (2020). Superficial white matter: A review on the dmri analysis methods and applications. *NeuroImage* **214,** 116673.

Guevara, M., Román, C., Houenou, J., Duclap, D., Poupon, C., Mangin, J. F., and Guevara, P. (2017). Reproducibility of superficial white matter tracts using diffusion-weighted imaging tractography. *NeuroImage* **147,** 703–725.

Guevara, P., Duclap, D., Poupon, C., Marrakchi-Kacem, L., Fillard, P., Le Bihan, D., Leboyer, M., Houenou, J., and Mangin, J.-F. (2012). Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas,. *NeuroImage,* **61,** 1083–1099,.

Guevara, P., Poupon, C., Rivière, D., Cointepas, Y., Descoteaux, M., Thirion, B., and Mangin, J. (2011a). Robust clustering of massive tractography datasets. *NeuroImage* **54,** 1975–1993.

Guevara, P., Poupon, C., Rivière, D., Cointepas, Y., Descoteaux, M., Thirion, B., and Mangin, J. (2011b). Robust clustering of massive tractography datasets. *NeuroImage* **54,** 1975–1993.

Guillaume, B., Wang, C., Poh, J., Shen, M. J., Ong, M. L., Tan, P. F., Karnani, N., Meaney, M., and Qiu, A. (2018). Improving mass-univariate analysis of neuroimaging data by modelling important unknown covariates: Application to epigenome-wide association studies. *NeuroImage* **173,** 57–71.

Hagmann, P., Jonasson, L., Maeder, P., Thiran, J.-P., Wedeen, V. J., and Meuli, R. (2006). Understanding diffusion mr imaging techniques: From scalar diffusion weighted imaging to diffusion tensor imaging and beyond. *RadioGraphics* **26,** 205–223.

Harms, M. P., Somerville, L. H., Ances, B. M., Andersson, J., Barch, D. M., Bastiani, M., Bookheimer, S. Y., Brown, T. B., Buckner, R. L., Burgess, G. C., et al. (2018). Extending the human connectome project across ages: Imaging protocols for the lifespan development and aging projects. *NeuroImage* **183,** 972–984.

Heiervang, E., Behrens, T., Mackay, C., Robson, M., and Johansen-Berg, H. (2006). Between session reproducibility and between subject variability of diffusion MR and tractography measures. *Neuroimage* **33,** 867–877.

Howell, B. R., Styner, M. A., Gao, W., Yap, P.-T., Wang, L., Baluyot, K., Yacoub, E., Chen, G., Potts, T., Salzwedel, A., et al. (2019). The unc/umn baby connectome project (bcp): an overview of the study design and protocol development. *NeuroImage* **185,** 891–905.

Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics,* **35,** 73–101.

Jbabdi, S. and Behrens, T. E. (2013). Long-range connectomics,. *Annals of the New York Academy of Sciences,* **23,** 83–93,.

Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., and Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage* **103,** 411–426.

Ji, J. L., Spronk, M., Kulkarni, K., Repovs., G., Anticevic, A., and Cole, M. W. (2019). Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage* **185,** 35–57.

Jin, Y., Shi, Y., Zhan, L., Gutman, B. A., de Zubicaray, G. I., McMahon, K. L., Wright, M. J., Toga, A. W., and Thompson, P. M. (2014). Automatic clustering of white matter fibers in brain diffusion MRI with an application to genetics. *NeuroImage* **100,** 75–90.

Johansen-Berg, H. and Behrens, T. E. (2009). *Diffusion MRI: From quantitative measurement to In vivo neuroanatomy.* Elsevier.

Jones, D. K. (2010). *Diffusion MRI.* Oxford University Press.

Katkovnik, V., Foi, A., Egiazarian, K., and Astola, J. (2010). From local kernel to nonlocal multiple-model image denoising. *International Journal of Computer Vision* **86,** 1–32.

Knickmeyer, R. C., Wang, J. P., Zhu, H. T., Geng, X., Woolson, S., Hamer, R. M., Konneker, T., Lin, W. L., Styner, M., and Gilmore, J. H. (2014). Common variants in psychiatric risk genes predict brain structure at birth. *Cerebra Cortex* **24,** 1230–1246.

Korostelev, A. P. and Tsybakov, A. B. (2012). *Minimax theory of image reconstruction*, volume 82. Springer Science & Business Media.

Kumar, K., Siddiqi, K., and Desrosiers, C. (2019a). White matter fiber analysis using kernel dictionary learning and sparsity priors,. *Pattern Recognition,* **95,** 83–95,.

Kumar, K., Siddiqi, K., and Desrosiers, C. (2019b). White matter fiber analysis using kernel dictionary learning and sparsity priors. *Pattern Recognition* **95,** 83–95.

Lazar, M. (2010). Mapping brain anatomical connectivity using white matter tractography,. *NMR in Biomedicine,* **23,** 8219835,.

Lazar, N. (2008a). *The statistical analysis of functional MRI data.* Springer Science & Business Media.

Lazar, N. A. (2008b). *The Statistical Analysis of Functional MRI Data.* Springer.

Le Bihan, D. and Iima, M. (2015). Diffusion magnetic resonance imaging: what water tells us about biological tissues. *PLoS Biology* **13,**.

Lee, J. (1983). Digital image smoothing and the sigma filter. *Computer Vision, Graphics,* *and Image Processing* **24,** 255–269.

Lee, S., Sun, W., Wright, F. A., and Zou, F. (2017). An improved and explicit surrogate variable analysis procedure by coefficient adjustment. *Biometrika* **104,** 303–316.

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., Geman, D., Baggerly, K., and Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics* **11,** 733–739.

Leek, J. T. and Storey, J. D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS genetics* **3,** e161.

Leek, J. T. and Storey, J. D. (2008). A general framework for multiple testing dependence. *Proceedings of the National Academy of Sciences* **105,** 18718–18723.

Li, H., Xue, Z., Guo, L., Liu, T., Hunter, J., and Wong, S. (2010). A hybrid approach to automatic clustering of white matter fibers,. *NeuroImage,* **49,** 1249–1258,.

Li, Y., Zhu, H., Shen, D., Lin, W., Gilmore, J. H., and Ibrahim, J. G. (2011). Multiscale adaptive regression models for neuroimaging data. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **73,** 559–578.

Lindquist, M. A., Loh, J. M., and Yue, Y. R. (2010). Adaptive spatial smoothing of fmri images. *Statistics and its Interface* **3,** 3–13.

Loh, P.-L. and Wainwright, M. J. (2015). Regularized m-estimators with nonconvexity: Statistical and algorithmic theory for local optima. *The Journal of Machine Learning Research* **16,** 559–616.

MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.

Mahmoudi, M. and Sapiro, G. (2005). Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE Signal Processing Letters* **12,** 839–842.

McKennan, C. and Nicolae, D. (2018a). Accounting for unobserved covariates with varying degrees of estimability in high dimensional biological data. *arXiv preprint arXiv:1801.00865* .

McKennan, C. and Nicolae, D. (2018b). Estimating and accounting for unobserved covariates in high dimensional correlated data. *arXiv preprint arXiv:1808.05895* .

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., Bartsch, A. J., Jbabdi, S., Sotiropoulos, S. N., Andersson, J. L., et al. (2016). Multimodal population brain imaging in the uk biobank prospective epidemiological study. *Nature neuroscience* **19,** 1523.

Mori, S. and van Zijl, P. C. M. (2002). Fiber tracking: principles and strategies a technical review,. *NMR in Biomedicine,* **15,** 468–480,.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications* **9,** 141–142.

Nguyen, H., Patel, V., Nasrabadi, N., and Chellappa, R. (2012). Kernel dictionary learning,. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pages 2021–2024,.

O'Donnell, L., Wells III, W. M., Golby, A., and Westin, C. (2012). Unbiased groupwise registration of white matter tractography. *International Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI)* pages 123–130.

O'Donnell, L. J., Golby, A. J., and Westin, C. (2013). Fiber clustering versus the parcellation-based connectome. *NeuroImage* **80,** 283–289.

Olivetti, E., Berto, G., Gori, P., Sharmin, N., and Avesani, P. (2017). Comparison of distances for supervised segmentation of white matter tractography. In *Pattern Recognition in Neuroimaging (PRNI), 2017 International Workshop on*, pages 1–4. IEEE.

Özarslan, E., Vemuri, B. C., and Mareci, T. H. (2005). Generalized scalar measures for diffusion mri using trace, variance, and entropy. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **53,** 866–876.

Parker, H. S., Bravo, H. C., and Leek, J. T. (2014). Removing batch effects for prediction problems with frozen surrogate variable analysis. *PeerJ* **2,** e561.

Pasaniuc, B. and Price, A. L. (2017). Dissecting the genetics of complex traits using summary association statistics. *Nature Reviews Genetics* **18,** 117.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Pierpaoli, C. and Basser, P. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance In Medicine,* **36,** 893–906.

Polzehl, J. and Spokoiny, V. (2000a). Adaptive weights smoothing with applications to image restoration. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **62,** 335–354.

Polzehl, J. and Spokoiny, V. G. (2000b). Adaptive weights smoothing with applications to image restoration. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **62,** 335–354.

Polzehl, J., Voss, H. U., and Tabelow, K. (2010). Structural adaptive segmentation for statistical parametric mapping. *NeuroImage* **52,** 515–523.

Prckovska, V., Rodrigues, P., Puigdellivol Sanchez, A., Ramos, M., Andorra, M., Martinez Heras, E., Falcon, C., Prats Galino, A., and Villoslada, P. (2016). Reproducibility of the structural connectome reconstruction across diffusion methods. *Journal of Neuroimaging* **26,** 46–57.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38,** 904–909.

Reisert, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., and Kiselev, V. (2011). Global fiber reconstruction becomes practical,. *NeuroImage,* **54,** 9559962,.

Rigollet, P. and Hütter, J.-C. (2015). High dimensional statistics. *Lecture notes for course 18S997* **813,** 814.

Rudelson, M., Vershynin, R., et al. (2013). Hanson-wright inequality and sub-gaussian concentration. *Electronic Communications in Probability* **18,**.

Satterthwaite, T. D., Elliott, M. A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M. E., Hopson, R., Jackson, C., Keefe, J., Riley, M., et al. (2014). Neuroimaging of the philadelphia neurodevelopmental cohort. *Neuroimage* **86,** 544–553.

Schilling, K. G., Nath, V., Hansen, C., Parvathaneni, P., Blaber, J., Gao, Y., Neher, P., Aydogan, D. B., Shi, Y., Ocampo-Pineda, M., et al. (2019). Limits to anatomical accuracy of diffusion tractography using modern approaches. *NeuroImage* **185,** 1–11.

Schwarz, C. G., Reid, R. I., Gunter, J. L., Senjem, M. L., Przybelski, S. A., Zuk, S. M., Whitwell, J. L., Vemuri, P., Josephs, K. A., Kantarci, K., et al. (2014). Improved DTI registration allows voxel-based analysis that outperforms tract-based spatial statistics. *Neuroimage* **94,** 65–78.

Sharmin, N., Olivetti, E., and Avesani, P. (2016). Alignment of tractograms as linear assignment problem. In *Computational Diffusion MRI*, pages 109–120. Springer.

Siless, V., Chang, K., Fischl, B., and Yendiki, A. (2018). Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas,. *NeuroImage,* **166,** 32–45,.

Siless, V., Davidow, J. Y., Nielsen, J., Fan, Q., Hedden, T., Hollinshead, M., Beam, E., Bustamante, C. M. V., Garrad, M. C., Santillana, R., Smith, E. E., Hamadeh, A., Snyder, J., Drews, M. K., Van Dijk, K., Sheridan, M., Somerville, L. H., and Yendiki, A. (2020). Registration-free analysis of diffusion mri tractography data across subjects through the human lifespan. *NeuroImage* **214,** 116703.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* **31,** 1487–1505.

Smith, S. M., M., J., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M., Matthews, P., and Behrens, T. E. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* **31,** 1487–1505.

Smith, S. M. and Nichols, T. E. (2018). Statistical challenges in “big data” human neuroimaging. *Neuron* **97,** 263–268.

Snook, L., Plewes, C., and Beaulieu, C. (2007). Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. *Neuroimage* **34,** 243–252.

Somerville, L. H., Bookheimer, S. Y., Buckner, R. L., Burgess, G. C., Curtiss, S. W., Dapretto, M., Elam, J. S., Gaffrey, M. S., Harms, M. P., Hodge, C., et al. (2018). The lifespan human connectome project in development: A large-scale study of brain connectivity development in 5–21 year olds. *NeuroImage* **183,** 456–468.

Sotiropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D. A., Yacoub, E., Lenglet, C., Essen, D. C. V., Ugurbil, K., and Behrens, T. E. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage* **80,** 125 – 143.

Springenberg, J. T., Dosovitskiy, A., Brox, T., and Riedmiller, M. (2014). Striving for simplicity: The all convolutional net,. *arXiv preprint arXiv:1412.6806* .

Srivastava, A. and Klassen, E. (2016). *Functional and Shape Data Analysis*. Springer, New York.

St-Jean, S., Chamberland, M., Viergever, M. A., and Leemans, A. (2019). Reducing variability in along-tract analysis with diffusion profile realignment. *NeuroImage* **199,** 663–679.

St-Onge, E., Daducci, A., Girard, G., and Descoteaux, M. (2018). Surface-enhanced tractography (set). *NeuroImage* **169,** 524–539.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64,** 479–498.

Storey, J. D. et al. (2003). The positive false discovery rate: a bayesian interpretation and the q-value. *The Annals of Statistics* **31,** 2013–2035.

Sun, Q., Zhou, W.-X., and Fan, J. (2020). Adaptive huber regression. *Journal of the American Statistical Association* **115,** 254–265.

Sun, W., Reich, B. J., Tony Cai, T., Guindani, M., and Schwartzman, A. (2015). False discovery control in large-scale spatial multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **77,** 59–83.

Sun, Y., Zhang, N. R., and Owen, A. B. (2012). Multiple hypothesis testing adjusted for latent variables, with an application to the agemap gene expression data. *The Annals of Applied Statistics* pages 1664–1688.

Sykov, E. and Nicholson, C. (2008). Diffusion in brain extracellular space,. *Physiological Reviews,* **88,** 1277–1340,.

Tabelow, K., Polzehl, J., Spokoiny, V., and Voss, H. U. (2008a). Diffusion tensor imaging: Structural adaptive smoothing. *NeuroImage,* **39,** 176391773.

Tabelow, K., Polzehl, J., Spokoiny, V., and Voss, H. U. (2008b). Diffusion tensor imaging: structural adaptive smoothing. *NeuroImage* **39,** 1763–1773.

Teschendorff, A. E. and Relton, C. L. (2018). Statistical and integrative system-level analysis of dna methylation data. *Nature Reviews Genetics* **19,** 129.

Teschendorff, A. E., Zhuang, J., and Widschwendter, M. (2011). Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. *Bioinformatics* **27,** 1496–1505.

Theaud, G., Houde, J. C., Bore, A., Rheault, F., Morency, F., and Descoteaux, M. (2020). Tractoflow: A robust, efficient and reproducible diffusion mri pipeline leveraging nextflow & singularity. *NeuroImage* **218,** 116889.

Thiebaut de Schotten, M., ffytche, D., Bizzi, A., Dell'Acqua, F., Allin, M., Walsh, M., Murray, R., Williams, S., Murphy, D., and Catani, M. (2011). Atlasing location,asymmetry and inter-subject variability of white matter tracts in the human brain with mr diffusion tractography,. *NeuroImage,* **54,** 49–59,.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **58,** 267–288.

Tibshirani, R. J. and Taylor, J. (2011). The solution path of the generalized lasso. *The Annals of Statistics,* **39,** 1335–1371.

Tolga, T. (2009). Principal neighborhood dictionaries for nonlocal means image denoising. *IEEE Transactions on Image Processing* **18,** 2649–2660.

Tony Cai, T. and Zhang, a. (2017). Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics,. *Annals of statistics* pages 60–89.

Tournier, J., Calamante, F., Connelly, A., et al. (2007). Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35,** 1459–1472.

Tournier, J.-D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* **35,** 1459–1472.

Tristán-Vega, A., Westin, C.-F., and Aja-Fernández, S. (2009). Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. *NeuroImage* **47,** 638–650.

Tsybakov, A. B. (2008). *Introduction to Nonparametric Estimation.* Springer.

Tuch, D. S. (2002a). *Diffusion MRI of Complex Tissue Structure.* PhD thesis, MIT.

Tuch, D. S. (2002b). Diffusion mri of complex tissue structure. *PhD thesis, MIT* .

Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine,* **52,** 1358–1372.

Tuch, D. S., Weisskoff, R. M., Belliveau, J. W., and J., W. V. (1999). High angular resolution diffusion imaging of the human brain. *Proceedings of the 7th Annual Meeting of ISMRM,* .

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., Consortium, W.-M. H., et al. (2013). The wu-minn human connectome project: an overview. *Neuroimage* **80,** 62–79.

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W., et al. (2012). The Human Connectome Project: a data acquisition perspective. *Neuroimage* **62,** 2222–2231.

Wakana, S., Caprihan, A., Panzenboeck, M., Fallon, J., Perry, M., Gollub, R., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., and Mori, S. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimaging* **36,** 630–644.

Wang, D., Luo, Y., Mok, V. C., Chu, W. C., and Shi, L. (2016). Tractography atlas-based spatial statistics: Statistical analysis of diffusion tensor image along fiber pathways. *NeuroImage* **125,** 301–310.

Wang, J., Zhao, Q., Hastie, T., Owen, A. B., et al. (2017). Confounder adjustment in multiple hypothesis testing. *Annals of Statistics* **45,** 1863–1894.

Wassermann, D., Bloy, L., Kanterakis, E., Verma, R., and Deriche, R. (2010). Unsupervised white matter fiber clustering and tract probability map generation: Applications of a gaussian process framework for white matter fibers. *NeuroImage* **51,** 228–241.

Wassermann, D., Rathi, Y., Bouix, S., Kubicki, M., Kikinis, R., Shenton, M., and Westin, C. (2011). White matter bundle registration and population analysis based on Gaussian processes. In *Information Processing in Medical Imaging*, pages 320–332. Springer.

Watson, G. S. (1964). Smooth regression analysis. *Sankhya: The Indian Journal of Statistics, Series A* **26,** 359–372.

Wedeen, V. J., Hagmann, P., Tseng, W.-Y. I., Reese, T. G., and Weisskoff, R. M. (2005). Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance In Medicine,* **54,** 1377–1386.

Welton, T., Kent, D. A., Auer, D. P., and Dineen, R. A. (2015). Reproducibility of graph-theoretic brain network metrics: a systematic review. *Brain Connect* **5,** 193–202.

Wu, T. and Lange, K. (2008). Coordinate descent algorithms for lasso penalized regression. *Annals of Applied Statistics,* **2,** 224–244.

Wu, Y., Hong, Y., Ahmad, S., Lin, W., Shen, D., Yap, P.-T., Consortium, U. B. C. P., et al. (2020). Tract dictionary learning for fast and robust recognition of fiber bundles. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 251–259. Springer.

Wu, Y., Hong, Y., Feng, Y., Shen, D., and Yap, P.-T. (2020). Mitigating gyral bias in cortical tractography via asymmetric fiber orientation distributions. *Medical Image Analysis* **59,** 101543.

Wu, Y., Lin, W., Shen, D., Yap, P.-T., Consortium, U. B. C. P., et al. (2019). Asymmetry spectrum imaging for baby diffusion tractography. In *Information Processing in Medical Imaging*, pages 319–331.

Wu, Y., Lin, W., Shen, D., Yap, P.-T., Consortium, U. B. C. P., et al. (2020). Tract dictionary learning for fast and robust recognition of fiber bundles. In *Proceedings of the 28th annual meeting of the International Society of Magnetic Resonance in Medicine*.

Yan, H., Carmichael, O., Paul, D., and Peng, J. (2018). Estimating fiber orientation distribution from diffusion mri with spherical needlets. *Medical Image Analysis* **46,** 57–72.

Yaroslavsky, L. P. (1985). *Digital picture processing. An introduction.* Springer.

Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., and Feldman, H. M. (2012). Tract profiles of white matter properties: Automating fiber-tract quantification. *PLOS ONE* **7,** 1–15.

Yuan, Y., Zhu, H., Lin, W., and Marron, J. (2012). Local polynomial regression for symmetric positive definite matrices. *Journal of Royal Statistical Society B,* **74,** 104–117.

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yucel, M., Pantelis, C., and Bullmore, E. T. (2010). Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* **50,** 970–983.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics* **38,** 894–942.

Zhang, F., Wu, Y., Norton, I., Rigolo, L., Rathi, Y., Makris, N., and O'Donnell, L. J. (2018). An anatomically curated fiber clustering white matter atlas for consistent white matter tract parcellation across the lifespan. *Neuroimage* **179,** 429–447.

Zhang, T., Chen, H., Guo, L., Li, K., Li, L., Zhang, S., Shen, D., Hu, X., and Liu, T. (2014). Characterization of u-shape streamline fibers: Methods and applications. *Medical image analysis* **18,** 795–807.

Zhang, T., Li, F., Beckes, L., and Coan, J. A. (2013). A semi-parametric model of the hemodynamic response for multi-subject fmri data. *NeuroImage* **75,** 136–145.

Zhang, T., Shen, H., and Li, F. (2016). Linear and non-linear models for fmri time series analysis. *Handbook of Neuroimaging Data Analysis* page 309.

Zhang, Z., Descoteaux, M., Zhang, J., Girard, G., Chamberland, M., Dunson, D., Srivastava, A., and Zhu, H. (2018). Mapping population based structural connectomes. *NeuroImage* **172,** 130–145.

Zhu, H., Fan, J., and Kong, L. (2014). Spatially varying coefficient model for neuroimaging data with jump discontinuities. *Journal of the American Statistical Association* **109,** 1084–1098.

Zhu, H., Kong, L., Li, R., Styner, M., Gerig, G., Lin, W., and Gilmore, J. H. (2011). Fadtts: functional analysis of diffusion tensor tract statistics. *NeuroImage* **56,** 1412–1425.

Zhu, H., Li, R., and Kong, L. (2012). Multivariate varying coefficient model for functional responses. *Annals of statistics* **40,** 2634.

Zhu, H., Li, Y., Ibrahim, J. G., Shi, X., An, H., Chen, Y., Gao, W., Lin, W., Rowe, D. B., and Peterson, B. S. (2009). Regression models for identifying noise sources in magnetic resonance images. *Journal of the American Statistical Association* **104,** 623–637.

Zhu, T., Hu, R., Qiu, X., Taylor, M., Tso, Y., Yiannoutsos, C., Navia, B., Mori, S., Ekholm, S., Schifitto, G., et al. (2011). Quantification of accuracy and precision of multi-center dti measurements: a diffusion phantom and human brain study. *Neuroimage* **56,** 1398–1411.

Zhu, T., Liu, X., Gaugh, M. D., Connelly, P. R., Ni, H., Ekholm, S., Schifitto, G., and Zhong, J. (2009). Evaluation of measurement uncertainties in human diffusion tensor imaging (dti)-derived parameters and optimization of clinical dti protocols with a wild bootstrap analysis. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* **29,** 422–435.

Zvitia, O., Mayer, A., Shadmi, R., Miron, S., and Greenspan, H. K. (2010). Co-registration of white matter tractographies by adaptive-mean-shift and Gaussian mixture modeling. *IEEE Transactions on Medical Imaging* **29,** 132–145.