SPATIALLY REGULARIZING HIGH ANGULAR RESOLUTION DIFFUSION IMAGING

Shangbang Rao

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
2014

Approved by:

Joseph G. Ibrahim

Hongtu Zhu

Donglin Zeng

Wei Sun

Pew-Thian Yap

ABSTRACT

Shangbang Rao: Spatially Regularizing High Angular Resolution Diffusion Imaging
(Under the direction of Joseph G. Ibrahim and Hongtu Zhu)

Many recent high angular resolution diffusion imaging (HARDI) reconstruction techniques have been introduced to infer ensemble average propagator (EAP),describing the three-dimensional (3D) average diffusion process of water molecules or the angular structure information contained in EAP, orientation distribution function (ODF). Most of these methods perform reconstruction independently at each voxel, which essentially ignoring the functional nature of the HARDI data at different voxels in space. The aim of my thesis is to develop methods which can spatially and adaptively infer the EAP, or ODF of water diffusion in regions with complex fiber configurations.

In Chapter 3, we propose a penalized multi-scale adaptive regression model (PMARM) framework to spatially and adaptively infer the ODF of water diffusion in regions with complex fiber configurations. We first represent DW-MRI signals using Spherical Harmonic (SH) basis, then apply PMARM on advanced statistical methods to calculate the coefficients of SH representation, from which ODF representation is calculated using Funk-Radon transformation. PMARM reconstructs the ODF at each voxel by adaptively borrowing the spatial information from the neighboring voxels. We show in the real and simulated data sets that PMARM can substantially reduce the noise level, while improving the ODF reconstruction.

In Chapter 4, we propose a robust multi-scale adaptive and sequential smoothing (MASS) method framework to robustly, spatially and adaptively infer the EAP of water diffusion in regions with complex fiber configurations. We first calculate spherical polar Fourier basis representation of the DW-MRI signals, and then apply MASS adaptively and sequentially updating SPF representation by borrowing the spatial information from the neighboring

voxels. We show in the real and simulated data sets that MASS can reduce the angle detection errors on fiber crossing area and provides more accurate reconstructions than standard voxel-wise methods and robust MASS performs very well with the presence of outliers.

In Chapter 5, we extend multi-scale adaptive method framework to dictionary learning methods, and show that by adding smoothing technique, we can significantly improve the accuracy of EAP reconstruction and reduce the angle detection errors on fiber crossing, even in very low signal-to-noise ratio situation.

ACKNOWLEDGMENTS

First and foremost, I am extremely grateful to both of my advisors, Prof. Joseph G. Ibrahim and Prof. Hongtu Zhu for the continuous support of my Ph.D study and research, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having better mentors for my Ph.D study.

Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Donglin Zeng, Prof. Wei Sun, and Prof. Pew-Thian Yap, for their encouragement, insightful comments, and hard questions.

My special thanks to Dr. Jian Cheng from the Multimodal Imaging in Neuro Disorders (MIND) lab at UNC for all the discussions and insights on coding.

I would like to thank all the members in UNC Biostatistics and Imaging Analysis lab. Without all the study groups and paper discussions, my thesis would not have come to a successful completion.

I also want to thank my fellow classmates for their friendship and help through these many years at UNC. In particular, I would like to thank Ruoqing Zhu, Ryan May, Amy Shi, Emil Cornea, Guanhua Chen, Zhengzheng Tang, Yi Zhang, Yingqi Zhao, Joanne Lin, Chih-Da Wu, Xin Zhou, Christopher Bryant.

Many thanks to my parents, Yanhe Rao and Aifeng Wu, for their unwavering support.

And finally, I would like to thank my beautiful wife, Anran, for the care and love that makes every new day even better than the last, and my lovely daughter, Yeye, a source of unending joy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

How can we get information about the human brain anatomy and in particular,about cerebral white matter? Cerebral dissection used to be the only way to access the neural architecture (Dejerine 1895, Gray 1918). Then, anatomists started using chemical markers to do neuronography (MacLean and Pribram 1953, Selden et al. 1998). More recently, neural fiber tractography based on local injection of chemical markers and subsequent observation of the induced propagation yielded high-quality connectivity mapping in the cat and monkey cerebral cortex (Selden et al. 1998). As of today, diffusion-weighted (DW) magnetic resonance imaging (MRI) is the unique noninvasive technique capable of quantifying the anisotropic diffusion of water molecules in biological tissues like the human brain white matter.

The great success of DW-MRI comes from its capability to accurately describe the geometry of the underlying microstructure. DW-MRI captures the average diffusion of water molecules, which probes the structure of the biological tissue at scales much smaller than the imaging resolution. The diffusion of water molecules is Brownian under normal unhindered conditions, but in fibrous structure such as white matter, water molecules tend to diffusion along fibers. Due to this physical phenomenon, DW-MRI is able to obtain information about the neural architecture in vivo.

Shortly after the first acquisitions of diffusionweighted images (DWI) in vivo (Moseley et al. 1990, Osment et al. 1990), Basser et al (Basser et al. 1994b;a) proposed the rigorous formalism of the diffusion tensor (DT) model. Diffusion tensor imaging (DTI) describes the three-dimensional(3D) nature of anisotropy in tissues by assuming that the average diffusion of water molecules follows a Gaussian distribution. DTI has now proved to be

extremely useful to study the normal and pathological human brain. However, the Gaussian assumption is over-simplifying the diffusion of water molecules and thus has some limitations. While the Gaussian assumption is adequate for voxels in which there is only a single fiber orientation (or none), it breaks down for voxels in which there is more complicated internal structure. In fact, it is currently thought that between one third to two thirds of imaging voxels in the human brain white matter contain multiple fiber bundle crossings (Behrens et al. 2007).

High Angular Resolution Diffusion Imaging (HARDI) is a category of reconstruction methods proposed to avoid the Gaussian EAP assumption and resolve the complex fiber configurations. In this thesis we are interested in the reconstruction and processing of the Ensemble Average Propagator (EAP), describing the diffusion process to obtain richer information on the complex microstructure of biological tissues, and its various features like Orientation Distribution Functions (ODF).HARDI methods can be separated into two classes, i.e. single shell HARDI (sHARDI) and multiple shell HARDI (mHARDI). sHARDI methods like the most famous one Q-Ball Imaging (QBI) (Tuch 2004, Anderson 2005, Hess et al. 2006, Descoteaux et al. 2007). mHARDI methods like Spherical Polar Fourier Imaging. Most of these methods perform reconstruction independently at each voxel, which essentially ignoring the functional nature of the HARDI data at different voxels in space. The aim of my thesis is to develop methods which can spatially and adaptively infer the ensemble average propagator (EAP), or EAP features such as orientation distribution function (ODF) of water diffusion in regions with complex fiber configurations.

Chapter 2:This chapter provides literature review for some HARDI methods in DW-MRI, especially analytical Q-Ball Imaging(QBI) and Spherical Polar Fourier Imaging(SPFI), and some important statistical estimation methods.

Chapter 3: This chapter proposes a penalized multi-scale adaptive regression model (PMARM) framework to spatially and adaptively infer the orientation distribution function (ODF) of water diffusion in regions with complex fiber configurations.

Chapter 4: This chapter proposes a robust multi-scale adaptive and sequential smoothing

(MASS) method framework to robustly, spatially and adaptively infer the EAP of water diffusion in regions with complex fiber configurations.

Chapter 5: This chapter extends multi-scale adaptive framework to SPFI via dictionary learning.

CHAPTER 2: LITERATURE REVIEW

2.1 Diffusion Magnetic Resonance Imaging (dMRI)

Diffusion Magnetic Resonance Imaging (dMRI) is a widely used in-vivo imaging technique to explore the information of neural micro-structure by probing the diffusion of water molecules. So far it is still the unique non-invasive method to reveal the micro-geometry of nervous tissues noninvasively and to explore the neural connectome in living human subjects. The diffusion of water molecules is constrained by the surrounding structures including nerves, cells and surrounding tissue Figure 2.1. For example, qualitatively water molecules diffuse fast along fibers and slowly cross fibers. Thus measuring the diffusion process quantitatively is crucial to understanding the neural micro-structure and fiber directions.

2.1.1 Magnetic Resonance Imaging (MRI)

The principles of Magnetic Resonance Imaging (MRI) are based on spin which is the rotation of a particle around some axis. Spin is a fundamental quantum characteristic of elementary particles like protons, electrons. Some nuclei have the property to align with a magnetic field $\mathbf{B_0}$ if their mass number, i.e. the summation number of protons and neutrons, is odd. Essentially their spin aligned along $\mathbf{B_0}$. Without the external stimulus by magnetic field $\mathbf{B_0}$, the macroscopic magnetization $\mathbf{M} = \mathbf{0}$. In MRI, the particles considered are hydrogen nucleus because human body is largely composed of water molecules. Each water molecule has two hydrogen nuclei or protons. When a person is inside the magnetic field $\mathbf{B_0}$ of the scanner, the average magnetic moment of many protons becomes aligned with the direction of the field $\mathbf{B_0}$ which is assumed as the z-axis by convection, in the meanwhile, the magnetization vector of spins precesses around $\mathbf{B_0}$ with an angular frequency known as the

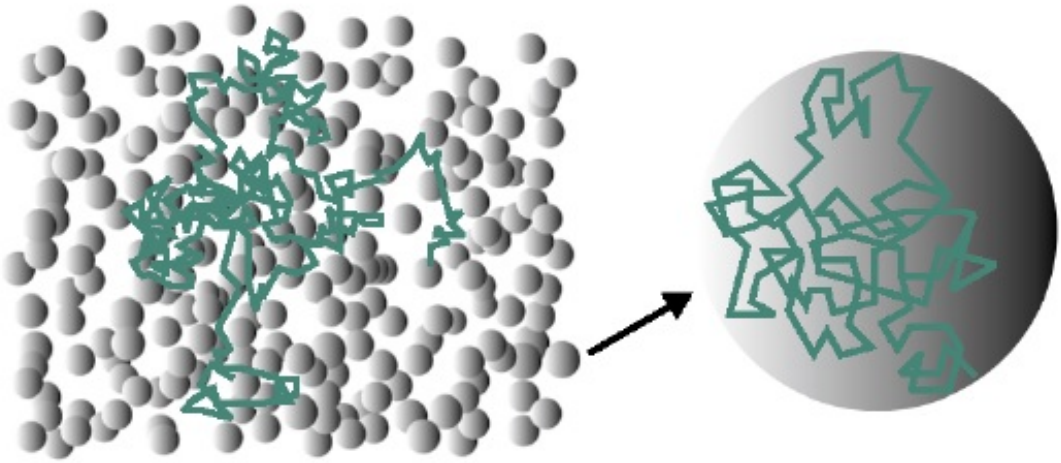


Figure 2.1: The water diffusion in biological tissues may be hindered by biological cells and other environment surrounding tissues. The figure is from (Johansen-Berg and Behrens 2009)

Larmor frequency. See Figure 2.2. When a Radio-Frequency (RF) is applied to the spins with the resonance frequency, the energy of RF is absorbed by the spins with low energy configuration and changes them into high energy configuration. Then the spins change their alignments. After RF is turned off, the spins begin to recover the alignment with $\mathbf{B_0}$, and finally return to the thermal equilibrium with low energy configuration. This is called as the relaxation phase. See Figure 2.3. Normally 90° and 180° RFs are used, which change the direction of spin with 90° and 180°. The Spin-lattice relaxation is the mechanism by which $\mathbf{M_z}$, the longitudinal component of the magnetization vector comes into thermodynamic equilibrium with its surroundings (the "lattice") in NMR and MRI. It is characterized by the spin-lattice relaxation time, a time constant known as $\mathbf{T_1}$. The spin-spin relaxation is the mechanism by which $\mathbf{M_{xy}}$, the transverse component of the magnetization vector, exponentially decays towards its equilibrium value in MRI, it is characterized by the spin-spin relaxation time, known as $\mathbf{T_2}$.

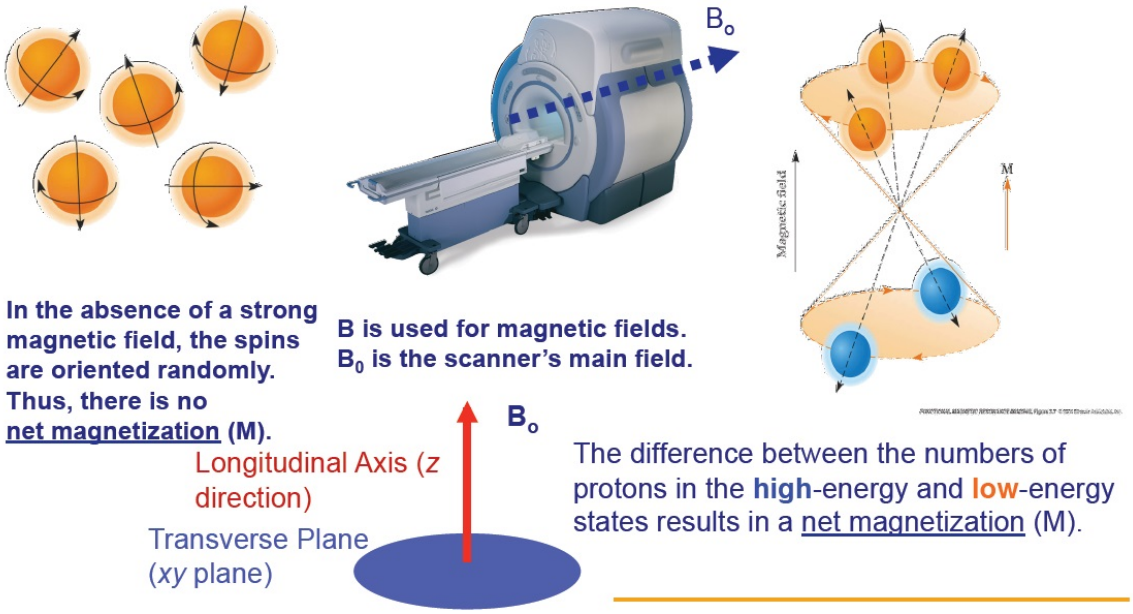


Figure 2.2: When inside the magnetic field $\mathbf{B_0}$ of the scanner, the average magnetic moment of many protons becomes aligned with the direction of the field $\mathbf{B_0}$ and the magnetization vector of spins precesses around $\mathbf{B_0}$. The figure is from the imaging course offered by Dr. Hongtu Zhu, UNC Chapel Hill

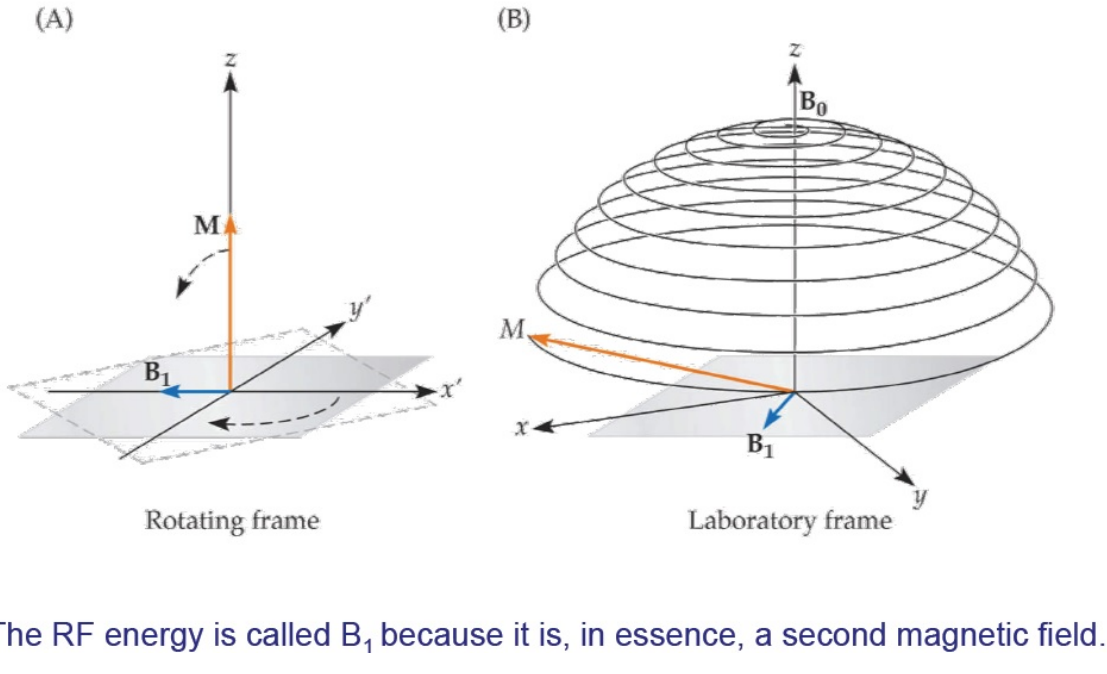


Figure 2.3: When a Radio-Frequency (RF) $\mathbf{B_1}$ is applied to the spins with the resonance frequency, the spins change their alignments. The figure is from the imaging course offered by Dr. Hongtu Zhu, UNC Chapel Hill

2.1.2 Diffusion Weighted Imaging (DWI)

Diffusion imaging is an MRI method that produces in vivo magnetic resonance images of biological tissues sensitized with the local characteristics of molecular diffusion, generally water (but other moieties can also be investigated using MR spectroscopic approaches). Regular MRI acquisition utilizes the behaviour of protons in water to generate contrast between clinically relevant features of a particular subject. The versatile nature of MRI is due to this capability of producing contrast related to the structure of tissues at microscopic level. In a typical $\mathbf{T_1}$-weighted image, water molecules in a sample are excited with the imposition of a strong magnetic field. This causes many of the protons in water molecules to precess simultaneously, producing signals in MRI. In $\mathbf{T_2}$-weighted images, contrast is produced by measuring the loss of coherence or synchrony between the water protons. When water is in an environment where it can freely tumble, relaxation tends to take longer. In certain clinical situations, this can generate contrast between an area of pathology and the surrounding healthy tissue.

To sensitize MRI images to diffusion, instead of a homogeneous magnetic field, the homogeneity is varied linearly by a pulsed field gradient. Since precession is proportional to the magnet strength, the protons begin to precess at different rates, resulting in dispersion of the phase and signal loss. Another gradient pulse is applied in the same magnitude but with opposite direction to refocus or rephase the spins. The refocusing will not be perfect for protons that have moved during the time interval between the pulses, and the signal measured by the MRI machine is reduced. The classical diffusion gradient sequence used in dMRI is the Pulsed Gradient Spin-Echo (PGSE) sequence proposed by Stejskal and Tanner (Stejskal and Tanner 1965). See Figure 2.4 for the sketch map of this sequence. This sequence uses two gradient pulses $\mathbf{G}(t)$ with duration time δ. The 90° RF pulse is applied to flip the magnetization in the transverse plane. Due to local magnetic field inhomogeneities, some spins slow down and some spins speed up. After a time Δ separating the two gradient pulses, the 180° pulse combined with the second gradient pulse is applied to refocuses the phase of spins so that slower spins lead ahead and the fast ones trail behind. The spin echo

process occurs when the spins recover their net magnetization.

Diffusion Weighted Imaging (DWI) signals, i.e. $S(b)$ with the diffusion weighting factor $b = \gamma^2\delta^2(\Delta - \delta/3)||\mathbf{G}||^2$ introduced by Dr. Lebihan in (LeBihan et al. 1986), and $S(0)$ with $b = 0$ is the baseline signal without any gradient, where γ is the proton gyromagnetic ratio, $\mathbf{G} = ||\mathbf{G}||\mathbf{u}$ is the diffusion sensitizing gradient pulse, $\tau = \Delta - \frac{1}{3}\delta$ is normally used to describe the effective diffusion time (LeBihan et al. 1986, Basser et al. 1994b). The signal intensity at each voxel in DWI is dependent on both surrounding structures and given weighted magnetic gradient (LeBihan et al. 1986). See Figure 2.5 for the DWI images $S(b)$ with different b values and different gradient directions $\mathbf{u}$. It can be seen that the DWI images are very noise, especially for large b values.

The diffusion weighted signal attenuation $E(b) = \frac{S(b)}{S(0)}$ is given by Stejskal-Tanner equation(Stejskal and Tanner 1965)

$$E(b) = \frac{S(b)}{S(0)} = \exp(-bD) \tag{2.1}$$

Where D is known as the Apparent Diffusion Coefficient (ADC) which reflects the property of surrounding tissues. In general case, ADC D is also dependent on $\mathbf{G}$ in a complex way, however free diffusion assumes D is only dependent on the direction of $\mathbf{G}$, i.e. $\mathbf{u} = \mathbf{G}/||\mathbf{G}||$.

Under narrow pulse condition, i.e. the duration time δ is much smaller than the separation time between two pluses Δ , $\mathbf{G}(t)$ is a constant $\mathbf{G}$ during δ. Then we introduce $\mathbf{q}$ vector as

$$\mathbf{q} = q\mathbf{u} = (2\pi)^{-1}\gamma\int_0^\delta \mathbf{G}(t)dt = (2\pi)^{-1}\gamma\delta\mathbf{G} \tag{2.2}$$

which can be seen as a vector in $\mathbf{q}$-space. Then the diffusion weighting factor can be represented by $\mathbf{q}$, i.e.

$$b = \gamma^2\delta^2(\Delta - \delta/3)||\mathbf{G}||^2 = 4\pi^2\tau q^2 \tag{2.3}$$

The early works in dMRI reported that the ADC D depends on gradient direction $\mathbf{u}$ and used two or three DWI images in different directions to detect the properties of tissues

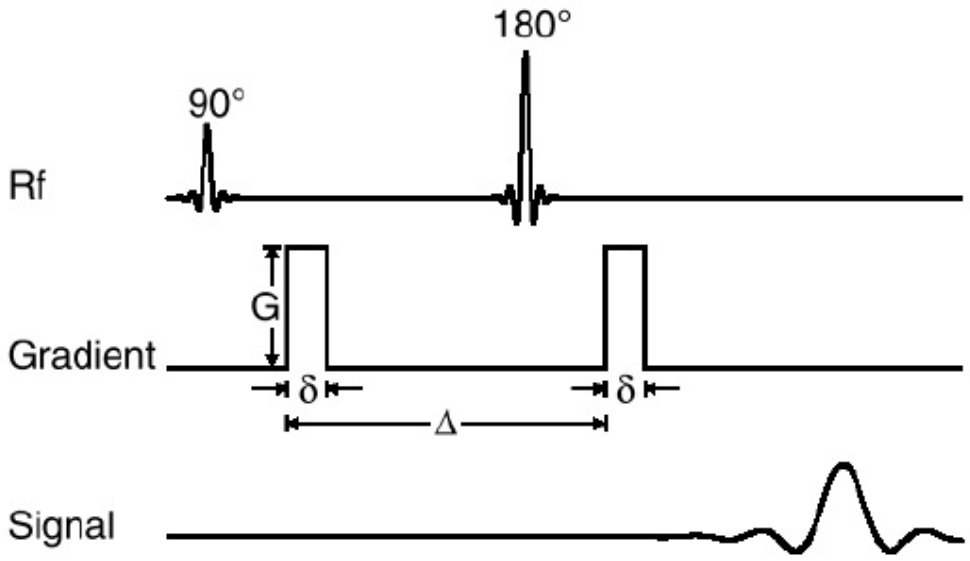


Figure 2.4: Pulsed Gradient Spin-Echo (PGSE) sequence introduced by Stejskal and Tanner(Stejskal and Tanner 1965). δ is the duration of the diffusion gradient pulses and Δ is the time between two diffusion gradient pulses.

(Moseley et al. 1990, Douek et al. 1991). Then Dr. Basser introduced diffusion tensor (Basser et al. 1994b) to represent ADC as

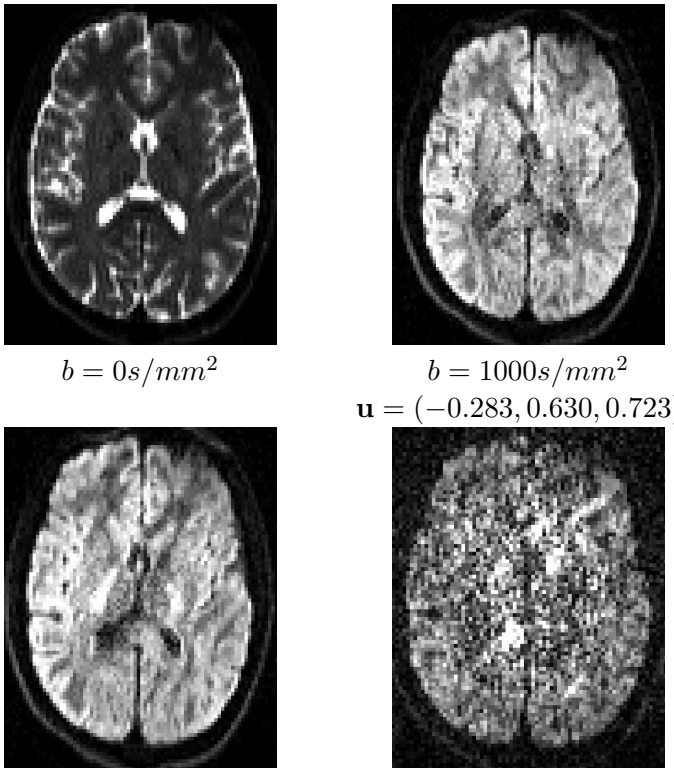
$$D(\mathbf{u}) = \mathbf{u}^T \mathbf{D} \mathbf{u} \tag{2.4}$$

$\mathbf{D}$ is called as the diffusion tensor, which is a 3×3 symmetric positive definite matrix independent of $\mathbf{u}$. This method is called as Diffusion Tensor Imaging (DTI). See Section 2.1.3 for more details.

2.1.3 Diffusion Tensor Imaging(DTI)

Dr. Basser proposed to model the ADC as a quadratic form parameterized by the diffusion tensor $\mathbf{D}$ in 2.4 (Basser et al. 1994b). Then the Stejskal-Tanner equation becomes

$$E(b) = \frac{S(b)}{S(0)} = \exp(-b\mathbf{u}^T \mathbf{D} \mathbf{u}) \tag{2.5}$$



$b = 0s/mm^2$

$b = 1000s/mm^2$
$\mathbf{u} = (-0.283, 0.630, 0.723)^T$

$b = 1000s/mm^2$
$\mathbf{u} = (-0.454, -0.861, 0.230)^T$

$b = 3000s/mm^2$
$\mathbf{u} = (-0.283, 0.630, 0.723)^T$

Figure 2.5: DWI images for different b-values and gradients.The data is from UNC BIAS lab.

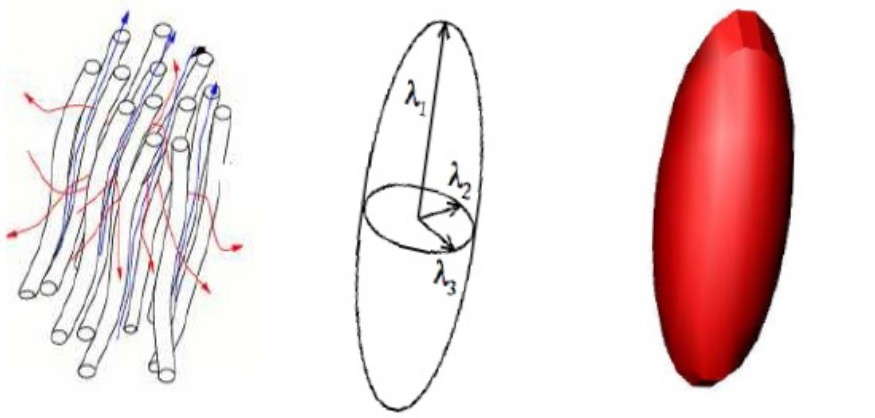


Figure 2.6: Diffusion tensor representation from (Descoteaux 2008).

The diffusion tensor $\mathbf{D} \in Sym_3^+$ is independent of b value and gradient direction $\mathbf{u}$, where Sym_3^+ is the space of 3×3 symmetric positive definite matrix. $\mathbf{D}$ can be eigendecomposed into three positive eigenvalues and corresponding eigenvectors, which is useful to define some scalar indices containing biological meaning.

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} = \lambda_1 \boldsymbol{v}_1 \boldsymbol{v}_1^T + \lambda_2 \boldsymbol{v}_2 \boldsymbol{v}_2^T + \lambda_3 \boldsymbol{v}_3 \boldsymbol{v}_3^T \tag{2.6}$$

Since $\mathbf{D}$ is symmetric it has six unknown coefficients that we need to estimate.Hence, DTI needs at least six DW images and one unweighted diffusion image ($b = 0s/mm^2$) to solve the system of equations. DTI estimation methods go from classical linear and non-linear least-squares (Basser et al. 1994b) or more complex methods which consider positive definite constraint or Rician noise(Tschumperlé and Deriche 2003, Chefd'hotel et al. 2004, Koay et al. 2006, Fillard et al. 2007). See Figure 2.6 for the sketch map of tensor representation and free diffusion along fibers.

Some useful scalar indices can be obtained from tensor D. The most important two

indices are Fractional Anisotropy (FA) and Mean Diffusivity (MD) Pierpaoli and Basser (1996) defined as

$$\mathbf{FA} = \frac{\sqrt{3}||\mathbf{D} - \frac{\mathbf{1}}{\mathbf{3}}\mathbf{Trace}(\mathbf{D})\mathbf{I}||}{\sqrt{2}||\mathbf{D}||} = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.7}$$

$$\mathbf{MD} = \frac{1}{3}\mathrm{Trace}(\mathbf{D}) = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{2.8}$$

Mixture of tensor model is a natural generalization of tensor model, where the signal is assumed to be a mixture of signals generated from tensors $\{\mathbf{D}_i\}_{i=1}^K$

$$E(b) = \sum_i^K \boldsymbol{w}_i \exp(-b\mathbf{u}^{\mathbf{T}}\mathbf{D}_i\mathbf{u}) \tag{2.9}$$

Based on some biological priors, the number of tensors is normally less than 3, typically K = 2. People normally use gradient descent method (typically the Levenberg-Marquardt minimization) Tuch (2002) to find a local minimum of the cost function in

$$\min_{\{\boldsymbol{w}_i, \mathbf{D}_i\}} \sum \sum_{j=1}^{N_s} \left(E_j - \sum_i^K \boldsymbol{w}_i \exp(-b_j \mathbf{u}_j^T \mathbf{D}_i \mathbf{u}_j) \right) \tag{2.10}$$

which is unstable and the result is sensitive to the initial point.

2.1.4 High Angular Resolution Diffusion Imaging(HARDI)

The term High Angular Resolution Diffusion Imaging (HARDI) was first proposed by Tuch (Tuch et al. 1999, Tuch 2002), where a finer angular resolution sampling scheme than conventional DTI sampling scheme was considered. The original HARDI term in (Tuch et al. 1999, Tuch 2002) means single shell sampling (only one b value).See Figure 2.7(c). However the mixture of tensor model in (Tuch et al. 1999, Tuch 2002) actually can be also used in Cartesian sampling in Figure 2.7(b) and sparse sampling (multiple b values) in Figure 2.7(d). With the development of MRI scanner, the acquisition time is reduced, which makes multiple shell data more practical and maybe available in clinical study in the

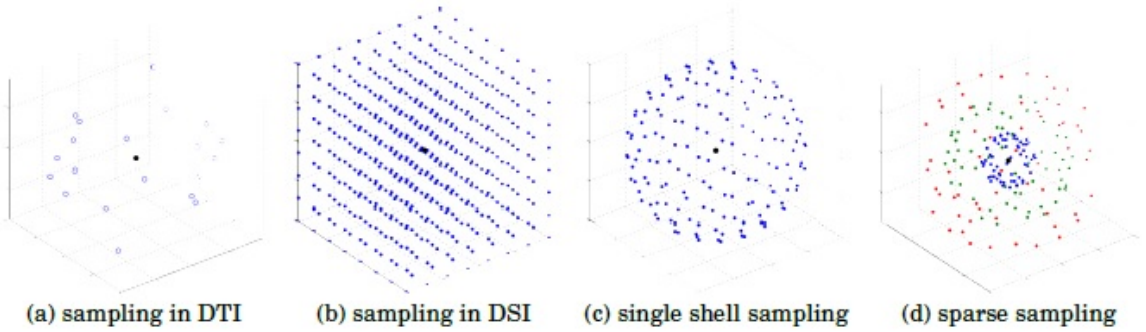


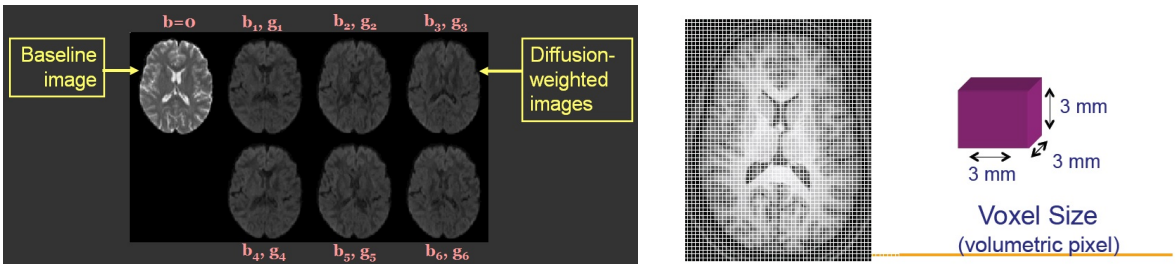
Figure 2.7: Several kinds of sampling in **q**-space.The black dot in $\mathbf{q} = (\mathbf{0}, \mathbf{0}, \mathbf{0})^{\mathbf{T}}$ the baseline image without diffusion gradient.(a) sampling used in DTI, normally less than 20 DWI images are used; (b) dense Cartesian sampling used in DSI. Note in practice the Cartesian samples inside a given Ball are used; (c) single shell sampling used in sHARDI methods, e.g. QBI, DOT etc; (d) sparse sampling used in mHARDI methods, e.g. DPI, SHORE,SPFI.

near future. Some research works proposed to estimate Orientation Distribution Functions or EAPs in multiple shell sampling (Liu et al. 2004, Assemlal et al. 2009, Özarslan et al. 2009, Descoteaux et al. 2010). Thus in this thesis, the term HARDI methods include all modeling methods beyond DTI. The HARDI methods which only can be used in single shell data are called as sHARDI methods. The HARDI methods which can be used in multiple shell data are called as mHARDI methods.

We usually acquire n normalized HARDI data with each image containing N voxels for each subject. Thus, we observe n normalized HARDI measurements $\{(E(b_i; \mathbf{v}), \mathbf{g}_i, b_i) : i = 1, \cdots, n\}$ at voxel $\mathbf{v} \in \mathcal{V}$, where $\mathbf{g}_i = (g_{i,1}, g_{i,2}, g_{i,3})^T$ is the gradient vector. These HARDI measurements can be also represented as **q**-space measurements $\{E(\mathbf{q}_i) : i = 1, \cdots, n\}$. We usually omit putting voxel $\mathbf{v}$, if no confusion in context. See Figure 2.8.

2.1.5 Diffusion Spectrum Imaging (DSI)

The EAP formalism provides a powerful framework to describe and predict the diffusion behavior in complex materials (Tuch 2002). Under the narrow pulse assumption(Stejskal and Tanner 1965), the relationship between the diffusion signal attenuation,$E(\mathbf{q})$, in q-space and the EAP, $p(\mathbf{R})$, in real space at each voxel in a common space $\mathcal{V}$, where $\mathbf{q} = q\mathbf{u} \in R^3$



DWI voxels in DWI

Figure 2.8: A set of diffusion-weighted (DW) images acquired with different gradient directions $\mathbf{g}_i$ and b_i.

and $\mathbf{R} = R\mathbf{r} \in R^3$, is given by a Fourier transform (FT) relationship (Callaghan 1991) such that

$$p(\mathbf{R}) = \int_{\mathbf{q} \in R^3} E(\mathbf{q}) e^{-2\pi i \mathbf{q} \cdot \mathbf{R}} d\mathbf{q} \tag{2.11}$$

A straightforward idea is to estimate $p(\mathbf{R})$ using fast Fourier transform from exhaustive signal samples (Callaghan 1991, Tuch 2002, Wedeen et al. 2000; 2005). This technique is called as Diffusion Spectrum Imaging (DSI). In practice, only limited number of samples are obtained and narrow pulse assumption is not always satisfied. (Wedeen et al. 2005) used 515 DWI images in a Cartesian sampling lattice in $\mathbf{q}$-space and the signal in $\mathbf{q}$-space was premultiplied by a Hanning window to obtain smooth attenuation of the signal at high $\mathbf{q}$ values. See Fig. 2.7(b) for the sketch map of the Cartesian sampling. In (Wedeen et al. 2005), the narrow pulse assumption is violated, the results are still exciting and show clearly some crossing fibers, which means that even though the narrow pulse assumption is violated, the Fourier transform can still obtain meaningful EAPs.

(Wedeen et al. 2005) visualized the EAP profile, or called iso-surface of EAP, which is the EAP with given radius R_0, i.e. $p(R_0\mathbf{r}) = p(R\mathbf{r})|_{R=R_0}$ The maxima of EAP profile were used to describe fiber directions later in many HARDI works (Assemlal et al. 2009, Özarslan et al. 2006; 2009, Descoteaux et al. 2010) See Figure 2.9 for the EAP profile with different radius R. The larger the radius R, the sharper the EAP profile is. However, EAP profile with large R has more estimation error. Thus normally $R = 15\mu m$is used in EAP profile

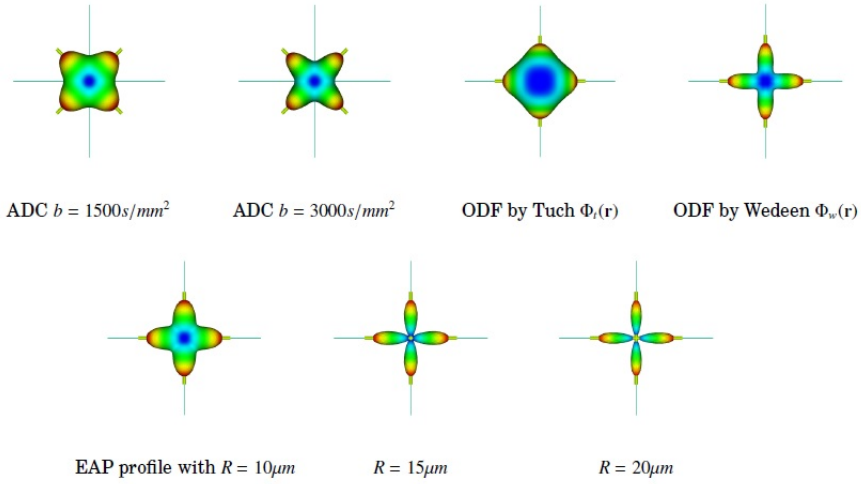


Figure 2.9: Fiber directions and ADC profiles with different b values, two kinds of ODFs, EAP profiles with different radius R.

to detect the fiber directions (Özarslan et al. 2006, Descoteaux et al. 2010). (Wedeen et al. 2005) also proposed another important feature of EAP, i.e. the Orientation Distribution Function (ODF), defined as

$$\Phi_w(\mathbf{r}) \stackrel{\text{def}}{=} \int_0^\infty p(\mathbf{R}) R^2 dR \tag{2.12}$$

It is called as ODF by Wedeen, denoted by $\Phi_w(\mathbf{r})$. $\Phi_w(\mathbf{r})$ is the marginal distribution of EAP $p(\mathbf{R})$, so the integration of $\Phi_w(\mathbf{r})$ over $\mathbb{S}^2$ i naturally 1. (Wedeen et al. 2005) proposed to first estimate EAP via numerical Fourier transform, then estimate the ODF in Eq.2.12 by numerical integration. Like the EAP profile, the maxima of ODFs are also normally assumed to be the directions of underlying fibers. Please see Figure 2.10 for EAP in 3D space and its two features, i.e. EAP profile and ODF.

2.1.6 Q-Ball Imaging (QBI)

Q-Ball Imaging (QBI) is the most widely used HARDI method. DSI needs a dense Cartesian sampling with a large range of b value, which makes it impractical (Wedeen et al. 2000; 2005). QBI was proposed to estimate the several kinds of ODFs, not EAP, from single shell sampling demonstrated in Figure 2.7(c), rather than Cartesian sampling inside a given

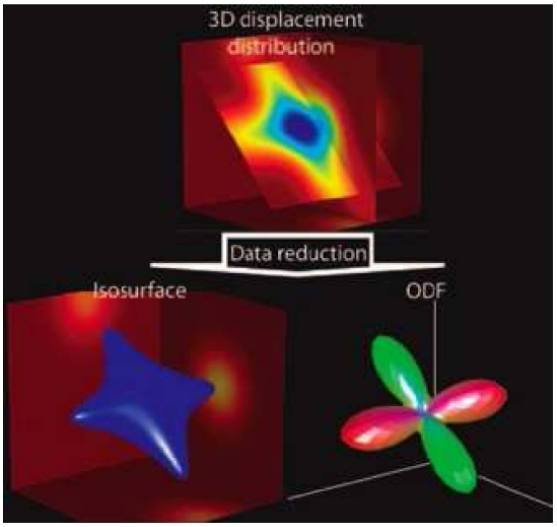


Figure 2.10: EAP in 3D **R**-space, and its two features, i.e. EAP profile (or called iso-surface of EAP) and ODF. The figure is from (Hagmann et al. 2006)

ball used in DSI in Figure 2.7(b).

Original Q-Ball Imaging

QBI was first proposed by Dr. Tuch in (Tuch 2002; 2004) in a numerical way and then was improved by an analytical way based on Spherical Harmonic basis in (Anderson 2005, Hess et al. 2006, Descoteaux et al. 2007). Instead of estimation of EAP, Dr. Tuch proposed to estimate a kind of ODF defined as

$$\Phi_t(\mathbf{r}) \stackrel{\text{def}}{=} \frac{1}{Z} \int_0^\infty p(R\mathbf{r}) dR \tag{2.13}$$

where Z is the normalization factor which makes $\int_{\mathbb{S}^2} \Phi_t(\mathbf{r}) d\mathbf{r} = 1$. This ODF is called as ODF by Tuch and denoted by $\Phi_t(\mathbf{r})$. Note $\Phi_t(\mathbf{r})$ is different from the ODF $\Phi_w(\mathbf{r})$ defined in Eq.2.12. $\Phi_w(\mathbf{r})$ is the marginal PDF of EAP which does not need artificial normalization factor, however, $\Phi_t(\mathbf{r})$ needs the normalization factor Z to make it as a PDF. Dr. Tuch proposed to estimate $\Phi_t(\mathbf{r})$ directly from samples of $E(\mathbf{q})$ in single shell data based on Funk-Radon Transform (FRT). The estimation of $\Phi_t(\mathbf{r})$ through FRT is calculated numerically.

The numerical QBI was later replaced by analytical QBI based on the representation of E(q) using SH basis. There were several groups which independently proposed the same analytical QBI using SHs. (Anderson 2005) obtained the analytical solution by considering the rotation property of SHs. (Hess et al. 2006) used addition theorem and rotation property of SHs, and considered Tikhonov regularization in least square estimation. (Descoteaux et al. 2007) applied 3D Funk-Hecke theorem to find the analytical solution of FRT. (Descoteaux et al. 2007) also proposed a simple and useful Laplace-Beltrami regularization scheme in least square estimation, which was shown to outperform the simple Tikhonov regularization and later became very popular for general least square estimation of spherical functions in HARDI domain.

Analytical QBI represents the signal $E(\mathbf{q})$ as SH basis, i.e.

$$E(q_0\mathbf{u}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}Y_l^m(\mathbf{u}) \tag{2.14}$$

where $Y_l^m(\mathbf{u})$ is the symmetric real spherical harmonic with order l and degree m. See (Descoteaux 2008) for more information on SHs. The coefficients $\{c_{lm}\}$ are normally estimated from signal samples by minimizing a least square cost function with Laplace- Beltrami regularization in (Descoteaux et al. 2007)

$$||B_M\mathbf{c} - \mathbf{E}||^2 + \mathbf{c}^T\Lambda\mathbf{c} \tag{2.15}$$

Where $\mathbf{c} = (c_{00}, \cdots, c_{LL})^T$ is the coefficient vector with $(L+1)(L+2)/2$ elements, $\mathbf{E} = (E_1, \cdots, E_{N_s})^T$ is the signal vector with N_s samples, $B_M$is the $N_s \times (L+1)(L+2)/2$ basis matrix generated by SHs, and Λ is the diagonal matrix with elements $Lambda_{lm} = \lambda l^2(l+1)^2$. $\mathbf{c}^T\Lambda\mathbf{c}$ is the Laplace-Beltrami regularization term. The least square problem has the closed form solution as

$$\mathbf{c} = (B_M^TB_M + \Lambda)^{-1}B_M^T\mathbf{E} \tag{2.16}$$

Based on Funk-Hecke theorem, the estimated ODF from FRT of $E(\mathbf{q})$ can be analytically obtained from the estimated $\{c_{lm}\}$ by:

$$\Phi_t(\mathbf{r}) = \frac{1}{Z}\mathrm{FRT}\{E(q_0\mathbf{u})\}(\mathbf{r}) = \frac{1}{Z}\sum_{l=0}^{L}\sum_{m=-l}^{l} 2\pi P_l(0) c_{lm} Y_l^m(\mathbf{u}) \tag{2.17}$$

where $P_l(0)$ is the Legendre polynomial of order l evaluated at 0.

Note in practice the ODF by Tuch $\Phi_t(\mathbf{r})$ is much smooth. The peaks of the ODF are only a little higher than the baseline values. Dr. Tuch proposed a minmax normalization method for visualization of $\Phi_t(\mathbf{r})$ to enhance the peaks of ODFs. Min-max normalization is a linear scaling to transform the ODF values into [0; 1], and it has been a commonly used way to visualize $\Phi_t(\mathbf{r})$ in literature. However, the minmax normalization also enhances the peaks of the ODFs in the area with isotropic diffusion.

Dr. Tuch also proposed a useful scalar index, named Generalized Fractional Anisotropy (FA), to describe the anisotropy of the ODFs, which can be seen as a generalization of previous FA in DTI model.

$$\mathrm{GFA}\{\Phi_t(\mathbf{r})\} \stackrel{\mathrm{def}}{=} \sqrt{\frac{N\sum_{i=1}^{N}(\Phi_t(\mathbf{r}_i) - \langle\Phi_t(\mathbf{r})\rangle)^2}{(n-1)\sum_{i=1}^{N}\Phi_t(\mathbf{r}_i)^2}} \tag{2.18}$$

where $\langle\Phi_t(\mathbf{r})\rangle$ is the mean of $\Phi_t(\mathbf{r})$. If the ODF is represented by SH basis with coefficients $\{c_{lm}\}$,the GFA can be represented by

$$\mathrm{GFA}\{\Phi_t(\mathbf{r})\} = \frac{||\Phi_t(\mathbf{r}) - \langle\Phi_t(\mathbf{r})\rangle||}{||\Phi_t(\mathbf{r})||} = \sqrt{1 - \frac{c_{00}^2}{\sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}^2}} \tag{2.19}$$

This is because of the orthogonality of SHs.

Exact Q-Ball Imaging

The ODF by Tuch is approximated by circle integration in original QBI, which has intrinsic limitations as we have discussed above. Exact QBI was proposed by several groups

independently (Wu et al. 2008, Canales-Rodrıguez et al. 2009, Aganj et al. 2010, Tristán-Vega et al. 2009a, Tristán-Vega et al. 2010) to estimate ODFs through a plane integration, not a circle integration.

Based on the famous projection-slice theorem in Fourier transform, the projection of $p(R\mathbf{r})$ along direction $\mathbf{r}$ r, i.e. the radial integration, equals to the integration of $E(\mathbf{q})$ in the orthogonal plane Π_r

In (Aganj et al. 2009; 2010), it is shown that if $E(\mathbf{q})$ follows the radial mon-exponential model, then we can use the single shell data to approximate $\Phi_w(\mathbf{r})$ by using

$$\Phi_w(\mathbf{r}) \approx \frac{1}{4\pi} + \frac{1}{16\pi^2}\mathbf{FRT}\{\Delta_b \ln(-\ln(E(\mathbf{u})))\}, \tag{2.20}$$

where Δ_b is the Laplace-Beltrami operator. In (Aganj et al. 2010), one may consider a model given by

$$\ln(-\ln(E(q\mathbf{u}))) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm} Y_l^m(\mathbf{u}) + \epsilon. \tag{2.21}$$

and considering $\Delta_b Y_l^m(\mathbf{u}) = -l(l+1)Y_l^m(\mathbf{u})$, we have

$$\Phi_w(\mathbf{r}) = \frac{1}{4\pi} - \frac{1}{8\pi} l(l+1) P_l(0) c_{lm} Y_l^m(\mathbf{u}) \tag{2.22}$$

The coefficients $\{c_{lm}\}$ can be estimated through a least square fitting from the samples of $\ln(-\ln(E(q_0\mathbf{u})))$ Note based on the above formula the first coefficient is $c_{00} = \frac{1}{\sqrt{4\pi}}$,then the integration of the estimated ODF is $\int_{\mathbb{S}^2} c_{00} Y_0^0(\mathbf{u}) d\mathbf{u} = 1$. Thus the estimated $\Phi_w(\mathbf{r})$ is naturally normalized.

2.1.7 Spherical Deconvolution (SD)

Spherical Deconvolution (SD) methods generalize the mixture model from discrete case to continuous case. In previous mixture of tensor model, $E(\mathbf{q})$ is assumed to be generated from K tensors. (Tournier et al. 2004; 2007) proposed to consider the continuous mixture

model as

$$E(q\mathbf{u}) = \int_{\mathbb{S}^2} \Phi_f(\mathbf{r}) R(\mathbf{r}^T\mathbf{u}) d\mathbf{r} \tag{2.23}$$

where $\Phi_f(\mathbf{r})$ is called as the fiber ODF (fODF) which needs to be estimated and $R(\mathbf{r}^T\mathbf{u})$ is the typical signal generated from one fiber. The spherical deconvolution is a modelbased method because it assumes the typical signal $R(\mathbf{r}^T\mathbf{u})$ and linear combination in the convolution. Mixture of tensor model is suffering from the model selection of the number of tensors and local minima of cost function. However, SD can be solved analytically by considering the Funk-Hecke theorem and representing the $E(q\mathbf{u}$ and $R(\mathbf{r}^T\mathbf{u})$ using SHs (Descoteaux et al. 2008a). The continuous weighting function $\Phi_f(\mathbf{r})$ avoids the limitation of mixture of tensor model in discrete case.

Note the SD method can be also used in some EAP features generated from signal. For example, consider Φ_t estimated from FRT, then based on the linearity of FRT we have

$$\Phi_t(\mathbf{r}) = \text{FRT}\{E(q\mathbf{u})\} = \int_{\mathbb{S}^2} \Phi_f(\mathbf{w}) \text{FRT}\{R(\mathbf{w}^T\mathbf{u})\} d\mathbf{w} = \int_{\mathbb{S}^2} \Phi_f(\mathbf{w}) \Phi_t^R(\mathbf{r}^T\mathbf{w}) d\mathbf{w} \tag{2.24}$$

Thus, if we use FRT to estimate $\Phi_t(\mathbf{r})$, the SD performed on $E(q\mathbf{u})$ is equivalent with SD performed on estimated $\Phi_t(\mathbf{r})$ (Descoteaux et al. 2008a). Since $\Phi_t(\mathbf{r})$ estimated from FRT is normally very smooth. SD becomes a good option to obtain the sharpened fiber ODF $\Phi_f(\mathbf{r})$.

2.1.8 Diffusion Propagator Imaging (DPI)

Diffusion Propagator Imaging (DPI) was proposed to model the signal $E(\mathbf{q})$ as the solution of Laplace's equation (Descoteaux et al. 2009; 2010). In DPI, the signal is assumed to be

$$E(q\mathbf{u}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} (\frac{c_{lm}}{q^{l+1}} + d_{lm} q^l) Y_l^m(\mathbf{u}) \tag{2.25}$$

Then the EAP is estimated from incomplete 3D integration inside the ball with a given radius q_{max}, because the complete integration in $\mathbb{R}^3$ does not converge.

$$\begin{aligned} p(R_0\mathbf{r}) &= \frac{1}{Z}\int_0^{q_{max}}\int_{\mathbb{S}^2} E(q\mathbf{u})q^2 e^{-2\pi i q R_0 \mathbf{u}^\mathbf{T}\mathbf{r}} dq d\mathbf{u} \\ &= \frac{1}{Z}\sum_{l=0}^{L}\sum_{m=-l}^{m=l}(p_{0l}(R_0)c_{lm} + p_{1l}(R_0)d_{lm})Y_l^m(\mathbf{r}) \end{aligned} \tag{2.26}$$

Where $p_{0l}(R_0)$ and $p_{1l}(R_0$ are given in (Descoteaux et al. 2010), Z is the normalization factor, q_{max} is the maximum q value used in DPI acquisition. The coefficents $\{c_{lm}\}$ and $\{d_{lm}\}$ can be calculated from DWI samples via a standard least square estimation, then we have the representation for EAP. However DPI is a model-based method, because it assumes $\Delta E(\mathbf{q};\mathbf{v}) = \mathbf{0}$

After obtaining the coefficients, DPI also proposed several EAP features analytically from incomplete radial integration, such as two ODFs.

$$\Phi_t(\mathbf{r}) = \frac{1}{Z}\int_0^{R_{max}} p(R\mathbf{r})dR = \frac{1}{Z}\sum_{l=0}^{L}\sum_{m=-l}^{l}(-1)^{l/2}(t_{0l}c_{lm} - t_{1l}c_{lm} + t_{2l}d_{lm})Y_l^m(\mathbf{r}) \tag{2.27}$$

$$t_{0l} = \frac{2^l\pi^{l-1}}{(2l-1)!!}\left(\frac{R_{max}^{l-1}}{l-1}\right), t_{1l} = \frac{\pi(l+1)!!}{2q_{max}^{l-1}(l/2+1)!}, t_{2l} = \frac{(-1)^{l/2}q_{max}^2(l-1)!!}{2^{l/2+2}(2\pi)^{3/2}(l/2+1)!}$$

$$\Phi_w(\mathbf{r}) = \int_0^{R_{max}} p(R\mathbf{r})R^2 dR = \sum_{l=0}^{L}\sum_{m=-l}^{l}(-1)^{l/2}(m_{0l}c_{lm} - m_{1l}c_{lm} + m_{2l}d_{lm})Y_l^m(\mathbf{r}) \tag{2.28}$$

$$m_{0l} = \frac{2^l\pi^{l-1}}{(2l-1)!!}\left(\frac{R_{max}^{l+1}}{l+1}\right), m_{1l} = \frac{(l-1)!!}{2\pi q_{max}^{l+1}(l/2-1)!}, m_{2l} = \frac{(l+1)!!}{2\pi 2^{l/2}(l/2)!}$$

2.1.9 Spherical Polar Fourier Imaging (SPFI)

Spherical Polar Fourier Imaging (SPFI) was first proposed by Dr. Assemlal in (Assemlal et al. 2008; 2009, Assemlal 2010). The diffusion signal $E(\mathbf{q})$ is represented by Spherical Polar Fourier (SPF) basis. SPF basis has SHs in spherical part and Gaussian-Laguerre functions in radial part and is a 3D orthonormal basis. Spherical polar Fourier imaging (SPFI) is a

model-free and fast HARDI method for multiple-shell data (Cheng et al. 2010).Let $G_k(q)$ be the Gaussian-Laguerre function and $B_{k,l,m}(\mathbf{q}) = G_k(q)Y_l^m(\mathbf{u})$ be spherical polar Fourier basis. The SPFI is to fit a model given by

$$E(\mathbf{q}) = \sum_{k=0}^{K}\sum_{l=0}^{L}\sum_{m=-l}^{l} a_{k,l,m}B_{k,l,m}(\mathbf{q}), \tag{2.29}$$

In (Assemlal et al. 2008),$\{a_{k,l,m}\}$ was proposed to be estimated by two methods, a least square fit and a nonlinear robust estimation which considers the Rician noise. After estimating $\{a_{k,l,m}\}$,a linear transformation was used to obtain the coefficeints $\{c_{k,l,m}\}$ of EAP profile $p(\mathbf{R}_0)$ represented by SH for a given R_0.

$$p(R_0\mathbf{r}) = \sum_{l=0}^{L}\sum_{m=-l}^{l}\left\{4(-1)^{l/2}\frac{\zeta^{0.5l+1.5}\pi^{l+1.5}R_0^l}{\Gamma(l+1.5)}\sum_{k=0}^{K} f_{k,l,m}(\zeta,R_0)a_{k,l,m}\right\}Y_l^m(\mathbf{u}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{l,m}Y_l^m(\mathbf{u}) \tag{2.30}$$

$$f_{k,l,m}(\zeta,R_0) = \kappa_k(\zeta)\sum_{i=0}^{k}(-1)^i\begin{pmatrix} k+0.5 \\ k-i \end{pmatrix}\frac{1}{i!}2^{0.5l+i-0.5}\Gamma(0.5l+i+1.5)1F1(\frac{2i+l+3}{2};l+\frac{3}{2};-2\pi^2R_0^2\zeta) \tag{2.31}$$

$$G_k(||\mathbf{q}_i||) = \kappa_k(\zeta)\exp(-\frac{||\mathbf{q}_i||^2}{2\zeta})L_k^{1/2}(\frac{||\mathbf{q}_i||^2}{\zeta}) \qquad \kappa_k(\zeta) = \left[\frac{2}{\zeta^{3/2}}\frac{n!}{\Gamma(n+3/2)}\right] \tag{2.32}$$

$$1F1(a;b;x) = \sum_{k=0}^{\infty}\frac{(a)_k x^k}{(b)_k k!}, \qquad (a)_k = (a(a+1)...(a+l-1)), \text{with}(a)_0 = 1 \tag{2.33}$$

The linear transform from $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ could be implemented as an matrix multiplication. This transformation is independent with the data, since $\{f_{k,l,m}(\zeta,R_0)\}$ only depends on ζ and R_0. Once a R_0 and the basis are given, trasformation matrix can be calculated. And since in SPFI, Only the value of $1F1$ at the fixed value $-2\pi^2R_0^2\zeta$ is needed, so the transformation matrix only needs to be calculated once.

2.2 Estimation methods

Most of the methods reviewed in previous sections assume that for each voxel $\mathbf{v}$,

$$f(E(\mathbf{q}_i;\mathbf{v})) = \mathbf{x}_i^T\beta(\mathbf{v}) + \epsilon_i(\mathbf{v}), \tag{2.34}$$

where $f(\cdot)$ is a given transformation function (e.g., $f(s) = s$ or $f(s) = \log(s)$), $\mathbf{x}_i$ is a $p \times 1$ vector of covariates, which depends on $\mathbf{q}_i$ (or $(\mathbf{b}_i, r_i)$), $\beta(\mathbf{v})$ is a $p \times 1$ vector of regression coefficients, and $\epsilon_i(\mathbf{v})$ is an error term with mean zero and variance $\sigma_i^2(\mathbf{v})$.

These methods focus on reconstructing $\beta(\mathbf{v})$ by solving a regularized linear least-squares optimization problem

$$\widehat{\beta}(\mathbf{v}) = \text{minimize}_{\beta(\mathbf{v})}||\mathbf{y}(\mathbf{v}) - \mathbf{X}\beta(\mathbf{v})||^2 + \rho(\beta(\mathbf{v});\lambda(\mathbf{v})), \tag{2.35}$$

where $\mathbf{y}(\mathbf{v}) = (f(E(\mathbf{q}_1;\mathbf{v})), \cdots, f(E(\mathbf{q}_n;\mathbf{v})))^T$, $\mathbf{X}$ is an $n \times p$ matrix with the $i-$th row being $\mathbf{x}_i$, and $\rho(\beta(\mathbf{v});\lambda(\mathbf{v}))$ is a penalty function with $\lambda(\mathbf{v})$ being a tuning parameter. Different penalty functions have been proposed in the literature. For instance, for analytical Q-ball imaging, the Laplacian-Beltrami regularization assumes

$$\rho(\beta(\mathbf{v});\lambda(\mathbf{v})) = \lambda(\mathbf{v})\sum_{l=0}^{L}\sum_{m=-l}^{l} l^2(l+1)^2\beta_{lm}^2 = \lambda(\mathbf{v})\beta(\mathbf{v})^T\Lambda\beta(\mathbf{v}), \tag{2.36}$$

where $\Lambda = \text{diag}(0, 4, 4, 4, \cdots, L^2(L+1)^2, \cdots, L^2(L+1)^2)$. Alternatively, we may consider other penalty functions, such as LASSO or generalized LASSO. Specifically, the penalty function of the generalized LASSO (Tibshirani and Taylor 2011) is given by

$$\rho(\beta(\mathbf{v});\lambda(\mathbf{v})) = \lambda(\mathbf{v})||D\beta(\mathbf{v})||_1, \tag{2.37}$$

where $||\cdot||_1$ is the L_1 norm and D is a specified $p \times p$ filter matrix.

Low-dimensional projections estimator

We review low-dimensional projections estimator (Zhang and Zhang 2014) as follows. Consider the linear model of $\mathbf{y} = \mathbf{X}\beta + \epsilon$, where $\epsilon \sim N(0, \sigma^2\mathbf{I})$. The low-dimensional projections estimator (LDPE) of β can be obtained as a one-step self bias correction from the initial estimator,

$$\hat{\beta}_j := \hat{\beta}_j^{(init)} + \frac{\mathbf{z}_j^T\{\mathbf{y} - \mathbf{X}\hat{\beta}_j^{(init)}\}}{\mathbf{z}_j^T\mathbf{x}_j}, \tag{2.38}$$

where $\mathbf{z}_j = \mathbf{x}_j^{\perp}$ is the residual of the least squares fit of $\mathbf{x}_j$ on $\mathbf{X}_{-j} = (\mathbf{x}_k, k \neq j)$. For QBI, $\mathbf{X}$ is the Spherical Harmonic representation of the signal directions, which are uniformly distributed on the sphere, and the SH basis is an orthonormal basis, so $\mathbf{z}_j = \mathbf{x}_j$ in this case. Moreover, $\hat{\beta}^{(init)}$ is obtained from the scaled LASSO (Sun and Zhang 2012) procedure as follow:

$$\{\hat{\beta}^{(init)}, \hat{\sigma}\} = \text{argmin}_{\beta,\sigma}\left\{\frac{|\mathbf{y} - \mathbf{X}\beta|_2^2}{2\sigma n} + \sigma/2 + \lambda_0|\beta|_1\right\} \tag{2.39}$$

where $\lambda_0 = \sqrt{(2/n)\log p}$, n is the number of the sampling directions, and p is the number of SH basis functions.

Robust Regression

The linear regression loss function, $\rho(\mathbf{e}) = ||\mathbf{e}||_2^2$ increases sharply with the size of the residual. Least squares estimates for regression models are highly sensitive to (not robust against) outliers. One alternative is to use $\rho(\mathbf{e}) = ||\mathbf{e}||_1$, the absolute value as a loss function instead of squaring the residual. This achieves robustness, but is hard to work with in practice because the absolute value function is not differentiable.

(Huber 1964) proposed a compromise between these two loss functions

$$\rho(e_i) = \begin{cases} e_i^2 & \text{if } |e_i| \leq c \\ c(2|e_i| - c) & \text{if } |e_i| > c \end{cases} \tag{2.40}$$

Where c is a parameter that controls the robustness level, and a smaller value of c usually leads to more robust estimation. Huber argued that $c = 1.345$ is a good choice, and showed that asymptotically, it is 95% as efficient as least square if the true distribution is normal and much more efficient in many other cases.(Huber and Ronchetti 1975)

Huber's estimators can be obtained by minimizing a loss function , or equivalently solving

$$\sum_i \rho'(e_i)\mathbf{x}_i = 0 \tag{2.41}$$

There are closed form solutions and fast algorithms for solving the least squares problem as well as the weighted least squares problem: Huber's estimators can be obtained by minimizing a loss function , or equivalently solving

$$\sum_i \omega_i e_i \mathbf{x}_i = 0 \tag{2.42}$$

Thus, a convenient way to solve for 2.41 is to use an iteratively reweighted least squares (IRLS) algorithm, in which we calculate $\omega_i = \rho'(e_i)/e_i$, solve the weighted least squares problem, re-calculate the weights, re-solve, and so on until convergence.

The preceding derivations are slightly oversimplied, in that the arguments for setting $c = 1.345$ are based on the assumption that the response variable has known variance 1. In reality, of course, this is not true, and we must apply the loss functions to the scaled residuals, i.e. replace every $\sum_i \rho(e_i)$ with $\sum_i \rho(e_i/s)$, and $\sum_i \rho'(e_i)$ with $\sum_i \rho'(e_i/s)$, where s is an estimated scale parameter. While a number of other estimators have been proposed,

the simplest is based on the median absolute deviation of the residuals:

$$\text{MAD} = \text{median}\{|e_i|\} \tag{2.43}$$

where $\hat{s} = \text{MAD}/0.6745$, based on the idea that, for the standard normal, $E(\text{MAD}) = 0.6745$.

2.3 Spatial Regularization

A key feature in HARDI is its spatial constraint. Specifically, the orientation and anisotropy of any single fiber bundles change smoothly from one voxel to the next, particularly along the dominant fiber orientation, whereas it may change dramatically at the boundaries between tracts and interfaces with gray matter structures and cerebrospinal fluid (CSF) spaces. Moreover, ODF or EAP is expected to change smoothly from one voxel to the next in the same fiber crossing region, whereas it may change dramatically at the boundaries of fiber crossing regions and surrounding fiber bundles. This is a very important and powerful constraint that can be exploited to improve the reconstruction in HARDI. However, most of current estimating methods are voxel-wise methods and do not make use of the spatial constraint of HARDI. To explicitly exploit such spatial constraint, we develop multiscale adaptive smoothing technique to spatially and adaptively update $\{\beta(\mathbf{v}) : \mathbf{v} \in \mathcal{V}\}$. (Li et al. 2011)

CHAPTER 3: SPATIALLY REGULARIZING HARDI VIA SPHERICAL HARMONICS

High angular resolution diffusion imaging (HARDI) has recently been of great interest in mapping the orientation of intra-voxel crossing fibers, and such orientation information allows one to infer the connectivity patterns prevalent among different brain regions and possible changes in such connectivity over time for various neurodegenerative and neuropsychiatric diseases. The aim of this chapter is to propose a penalized multi-scale adaptive regression model (PMARM) framework to spatially and adaptively infer the orientation distribution function (ODF) of water diffusion in regions with complex fiber configurations. In PMARM, we reformulate the HARDI imaging reconstruction as a weighted regularized least-squares regression (WRLSR) problem. Similarity and distance weights are introduced to account for spatial smoothness of HARDI, while preserving the unknown discontinuities (e.g., edges between white matter and grey matter) of HARDI. The L_1 penalty function is introduced to ensure the sparse solutions of ODFs, while a scaled L_1 weighted estimator is calculated to correct the bias introduced by the L_1 penalty at each voxel. In PMARM, we integrate the multiscale adaptive regression models (Li et al. 2011), the propagation-separation method (Polzehl and Spokoiny 2000), and Lasso (least absolute shrinkage and selection operator) (Tibshirani 1996) to adaptively estimate ODFs across voxels. Experimental results indicate that PMARM can reduce the angle detection errors on fiber crossing area and provide more accurate reconstruction than standard voxel-wise methods.

3.1 Introduction

Diffusion magnetic resonance imaging (dMRI) is a popular imaging technique for tracking the effective diffusion of water molecules, which is constrained by the surrounding structures,

such as nerves or cells, in the human brain *in vivo*. Because water molecules tend to diffuse fast along the pathways of white matter fibers and slow cross fibers, tracking its diffusion with dMRI allows one to map the microstructure and organization of those pathways (Basser and Pierpaoli 1996). Measuring the diffusion process quantitatively is critical for a quantitative assessment of the integrity of anatomical connectivity in white matter. A reconstruction step to dMRI is to estimate the normalized signal attenuation $E(\mathbf{q};\mathbf{v})$, the orientation distribution function (ODF) $O(\mathbf{r};\mathbf{v})$, and Ensemble Average Propagator (EAP) $p(\mathbf{R};\mathbf{v})$ at each voxel $\mathbf{v}$ in a common space $\mathcal{V}$, where $\mathbf{q} = q\mathbf{u} \in R^3$, $q = ||\mathbf{q}||_2$ and $\mathbf{R} = R\mathbf{r} \in R^3$, $R = ||\mathbf{R}||_2$, respectively, represent the effective gradient direction and displacement direction.

Raw HARDI images, as a result of elevated b-factor and decreased voxel size, suffer from depressed signal-to-noise ratio (SNR) levels, which make the problem of reconstructing HARDI data be of particular practical importance and challenging. Many existing methods perform reconstruction independently at each voxel, which essentially ignores the functional nature of the HARDI data at different voxels in space. Most of these methods model $E(\mathbf{q};\mathbf{v})$ or $O(\mathbf{r};\mathbf{v})$ as a linear combination of some known or unknown basis functions and then computes the model parameters by using a regularized linear least-squares optimization. Recently, there has been a great interest in incorporating spatial smoothness constraints into the HARDI reconstruction algorithm. The key assumption of this approach is that the orientation and anisotropy of any single fiber population are expected to vary smoothly along the dominant fiber orientation, except at the boundaries between tracts and interfaces with gray matter structures and cerebrospinal fluid spaces. Until recently, a few number of different approaches have been developed starting from smoothing raw HARDI images (Descoteaux et al. 2008b, Becker et al. 2012b;a), smoothing procedures in ODF space (Kim et al. 2009, Goh et al. 2011), spatial DTI (Tabelow et al. 2008, Yu et al. 2013, Yu and Li 2013, Liu et al. 2013), to spatial HARDI, which reconstructs and denoises all ODFs simultaneously (Raj et al. 2011).

The aim of this chapter is to develop a penalized multi-scale adaptive regression model (PMARM) framework to spatially and adaptively infer ODFs across all voxels. Similar to

(Raj et al. 2011), PMARM is also a simultaneous reconstruction and denoising procedure. However, PMARM differs significantly from the method in (Raj et al. 2011) in two major ways. First, we use similarity and distance weights to account for local spatial smoothness of HARDI, while preserving the unknown local discontinuities. In contrast, the spatial HARDI method uses a single smoothness regularization term to control global smoothness. Thus, PMARM should be more robust to heterogenous noise levels across different locations. Second, PMARM uses the L_1 regularization to ensure the sparse solutions of ODFs, wheres the spatial HARDI enforces Tikhonov regularization with several global tuning parameters in order to stabilize the estimated ODFs.

3.2 Methods

3.2.1 Model Formulation

We usually acquire n normalized HARDI data with each image containing N voxels for each subject. Thus, we observe n normalized HARDI measurements $\{(E(b_i;\mathbf{v}),\mathbf{g}_i,b_i) : i = 1,\cdots,n\}$ at voxel $\mathbf{v} \in \mathcal{V}$, where $\mathbf{g}_i = (g_{i,1}, g_{i,2}, g_{i,3})^T$ is the gradient vector. Based on (2.3), these HARDI measurements can be also represented as **q**-space measurements $\{E(\mathbf{q}_i) : i = 1,\cdots,n\}$ We usually omit putting voxel $\mathbf{v}$, if no confusion in context. See Figure 2.8.

Q-ball Imaging (QBI) is the most widely used HARDI method (Tuch et al. 1999; 2002), since it only needs single shell data to estimate the ODF, whose maxima agree with the fiber directions, and is very easy to be implemented. Q-Ball imaging, which is based on the Funk-Radon Transform (FRT) (Tuch 2004), was first proposed in a numerical way and then was improved by an analytical way based on spherical harmonics (Anderson 2005, Hess et al. 2006, Descoteaux et al. 2007). In Tuch (2004), a FRT approximation is proposed to estimate a kind of ODF defined as

$$\Phi_t(\mathbf{r};\mathbf{v}) \stackrel{\text{def}}{=} \frac{1}{Z}\int_0^\infty P(R\mathbf{r};\mathbf{v})dR \approx \text{FRT}\{E(q\mathbf{u};\mathbf{v})\}(\mathbf{r}), \tag{3.1}$$

where FRT is the Funk-Radon transform and Z is the normalization factor which makes $\int_{\mathbb{S}^2} \Phi_t(\mathbf{r};\mathbf{v})d\mathbf{r} = 1$. Recently, analytical QBI is proposed by representing E(q) as a linear combination of SH bases. Specifically, the analytical QBI is to fit a model given by

$$E(q\mathbf{u}_i;\mathbf{v}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}(\mathbf{v})H_l^m(\mathbf{u}_i) + \epsilon_i(\mathbf{v}), \tag{3.2}$$

where $H_l^m(\mathbf{u})$ is the symmetric real spherical harmonic with order l and degree m (Aganj et al. 2010, Descoteaux et al. 2007).

In Wedeen et al. (2005; 2000), another representation of ODF is given by

$$\Phi_w(\mathbf{r};\mathbf{v}) \stackrel{\text{def}}{=} \int_0^\infty P(R\mathbf{r};\mathbf{v})R^2 dR, \tag{3.3}$$

which is the marginal distribution of EAP such that it does not need the artificial normalization factor Z. One may estimate $\Phi_w(\mathbf{r};\mathbf{v})$ by calculating the numerical radial integration of a pre-estimated EAP estimated from diffusion spectrum imaging (DSI). In (Aganj et al. 2009) and (Aganj et al. 2010), it is shown that if $E(\mathbf{q};\mathbf{v})$ follows the radial multi-exponential model, then we can use the single shell data to approximate $\Phi_w(\mathbf{r};\mathbf{v})$ by using

$$\Phi_w(\mathbf{r};\mathbf{v}) \approx \frac{1}{4\pi} + \frac{1}{16\pi^2}\mathbf{FRT}\{\Delta_b \ln(-\ln(E(\mathbf{u};\mathbf{v})))\}, \tag{3.4}$$

where Δ_b is the Laplace-Beltrami operator. In (Aganj et al. 2010), one may consider a model given by

$$\ln(-\ln(E(q\mathbf{u}_i;\mathbf{v}))) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm}(\mathbf{v})H_l^m(\mathbf{u}_i) + \epsilon_i(\mathbf{v}). \tag{3.5}$$

This method is called constant solid angle QBI, denoted by cQBI in this paper, it can be further generalized to multiple $q-$shells based on the multi-exponential model (Aganj et al. 2010). $\Phi_w(\mathbf{r};\mathbf{v})$ has better angular resolution than $\Phi_t(\mathbf{r};\mathbf{v})$ normally used in original QBI, although normally $\Phi_t(\mathbf{r};\mathbf{v})$ is more robust to noise. $\Phi_w(\mathbf{r};\mathbf{v})$ normally does not need

the artificial normalization, sharpening technique like min-max normalization or spherical deconvolution. That is the main reason why we focus on reconstructing $\Phi_w(\mathbf{r};\mathbf{v})$ using cQBI for this chapter

3.2.2 Estimation Procedures

We reconstruct $\beta(\mathbf{v})$ by solving a regularized linear least-squares optimization problem

$$\widehat{\beta}(\mathbf{v}) = \text{minimize}_{\beta(\mathbf{v})}||\mathbf{y}(\mathbf{v}) - \mathbf{X}\beta(\mathbf{v})||^2 + \rho(\beta(\mathbf{v});\lambda(\mathbf{v})), \tag{3.6}$$

where $\mathbf{y}(\mathbf{v}) = (f(E(\mathbf{q}_1;\mathbf{v})),\cdots,f(E(\mathbf{q}_n;\mathbf{v})))^T$, $f(\cdot) = \ln(-\ln(\cdot))$, $\mathbf{X}$ is an $n\times p$ matrix with the $i-$th row being $\mathbf{x}_i$, and $\rho(\beta(\mathbf{v});\lambda(\mathbf{v}))$ is a penalty function with $\lambda(\mathbf{v})$ being a tuning parameter. Different penalty functions have been proposed in the literature. For instance, for analytical Q-ball imaging, the Laplacian-Beltrami regularization assumes

$$\rho(\beta(\mathbf{v});\lambda(\mathbf{v})) = \lambda(\mathbf{v})\sum_{l=0}^{L}\sum_{m=-l}^{l} l^2(l+1)^2\beta_{lm}^2 = \lambda(\mathbf{v})\beta(\mathbf{v})^T\Lambda\beta(\mathbf{v}), \tag{3.7}$$

where $\Lambda = \text{diag}(0,4,4,4,\cdots,L^2(L+1)^2,\cdots,L^2(L+1)^2)$. Alternatively, we may consider other penalty functions, such as LASSO, adaptive LASSO, generalized LASSO, or smoothly clipped absolute deviation (SCAD) (Fan and Li 2001, Tibshirani and Taylor 2011, Zou 2006b, Tibshirani 1996). Specifically, the penalty function of the generalized LASSO (Tibshirani and Taylor 2011) is given by

$$\rho(\beta(\mathbf{v});\lambda(\mathbf{v})) = \lambda(\mathbf{v})||D\beta(\mathbf{v})||_1, \tag{3.8}$$

where $||\cdot||_1$ is the L_1 norm and D is a specified $p\times p$ filter matrix. However, the standard LASSO assumes D to be an identity matrix. Following the Laplacian-Beltrami regularization, we may set $D = \Lambda^{1/2}$ in order to give different weights to the coefficients at different orders for Q-ball imaging. In this case, $\widehat{\beta}(\mathbf{v})$ is close to an adaptive LASSO estimator (Zou 2006b).

Many estimation methods have been developed to solve the regularized linear least-squares optimization (3.6). For instance, it is computationally easy to compute $\widehat{\beta}_{LB}(\mathbf{v}) = (\mathbf{X}^T\mathbf{X}+\lambda(\mathbf{v})\Lambda)^{-1}\mathbf{X}^T\mathbf{y}(\mathbf{v})$ for the Laplacian-Beltrami regularization (Descoteaux et al. 2007), but it is very sensitive to noise. As shown in the second row of Figure 3.1, there are many false maxima in ODF based on $\widehat{\beta}_{LB}(\mathbf{v})$. Although the LASSO method has been widely used and yields a sparse estimate of $\beta(\mathbf{v})$, denoted by $\widehat{\beta}_{LO}(\mathbf{v})$, LASSO can introduce substantial bias in the estimation of $\beta(\mathbf{v})$ Figure 3.1, even though the estimated ODF from LASSO seems to be slightly better than the one based on $\widehat{\beta}_{LB}(\mathbf{v})$. In (Zhang and Zhang 2014), a low-dimensional projections estimator (LDPE), denoted by $\widehat{\beta}_{LE}(\mathbf{v})$, is developed to address such bias issue Figure 3.1. The LDPE estimator is included section 2.2. From here on, we primarily consider the three estimation methods including the Laplace-Beltrami regularization, LASSO, and LDPE, even though extension to other cases is definitely feasible. After calculating $\widehat{\beta}(\mathbf{v})$, we can calculate ODFs and infer their maxima.

A key feature in HARDI is its spatial constraint. Specifically, the orientation and anisotropy of any single fiber bundles change smoothly from one voxel to the next, particularly along the dominant fiber orientation, whereas it may change dramatically at the boundaries between tracts and interfaces with gray matter structures and cerebrospinal fluid (CSF) spaces. Moreover, the ODF is expected to change smoothly from one voxel to the next in the same fiber crossing region, whereas it may change dramatically at the boundaries of fiber crossing regions and surrounding fiber bundles. This is a very important and powerful constraint that can be exploited to improve the reconstruction in HARDI. However, the methods in (3.6) of estimating ODF are voxel-wise methods and do not make use of the spatial constraint of HARDI.

To explicitly exploit such a spatial constraint, we develop a penalized multiscale adaptive regression model (PMARM) to spatially and adaptively update $\{\beta(\mathbf{v}) : \mathbf{v} \in \mathcal{V}\}$ by integrating various penalization methods (Fan and Li 2001, Tibshirani and Taylor 2011, Zou 2006b, Tibshirani 1996, Zhang and Zhang 2014), multiscale adaptive regression models (Li et al. 2011), and the propagation-separation method (Polzehl and Spokoiny 2000). The key idea of

PMARM is to combine HARDI signals in a neighboring sphere of voxel $\mathbf{v}$ to make inference on $\beta(\mathbf{v})$ at voxel $\mathbf{v}$. Specifically, let $B(\mathbf{v}, h)$ be a sphere with radius h centered at voxel $\mathbf{v}$ and $\omega(\mathbf{v}, \mathbf{v}'; h)$ be a weight function of triple $(\mathbf{v}, \mathbf{v}', h)$ such that

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

PMARM is based on a set of weighted penalization functions, denoted by $P_n(\beta(\mathbf{v}); \omega, h)$, which is defined as follows:

$$\begin{aligned} P_n(\beta(\mathbf{v}); \omega, h) &= \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega(\mathbf{v}, \mathbf{v}'; h) ||\mathbf{y}(\mathbf{v}') - \mathbf{X}\beta(\mathbf{v})||^2 + \rho(\beta(\mathbf{v}); \lambda(\mathbf{v})) \\ &= ||\mathbf{y}_w(\mathbf{v}; h) - \mathbf{X}\beta(\mathbf{v})||^2 + \rho(\beta(\mathbf{v}); \lambda(\mathbf{v})) + \text{constant}, \end{aligned} \tag{3.9}$$

where $\mathbf{y}_w(\mathbf{v}; h) = \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega(\mathbf{v}, \mathbf{v}'; h)\mathbf{y}(\mathbf{v}')$. Given the current weights $\{\omega(\mathbf{v}, \mathbf{v}'; h) : \mathbf{v}, \mathbf{v}' \in \mathcal{V}\}$, we consider the weighted GEE estimator of $\beta(\mathbf{v})$, denoted by $\hat{\beta}(\mathbf{v}, h)$, which satisfies

$$\hat{\beta}(\mathbf{v}, h) = \text{argmin}_{\beta(\mathbf{v})} P_n(\beta(\mathbf{v}); \omega, h). \tag{3.10}$$

It is critical to choose a good $\omega(\mathbf{v}, \mathbf{v}'; h)$ in preventing oversmoothing the estimates of $\beta(\mathbf{v})$ across voxels, while preserving the edges between different structures, such as fiber bundles, crossing fibers, or gray matter regions. A good $\omega(\mathbf{v}, \mathbf{v}'; h)$ should quantify the similarity between $\beta(\mathbf{v})$ and $\beta(\mathbf{v}')$ or their corresponding ODFs. Specifically, if $\beta(\mathbf{v})$ and $\beta(\mathbf{v}')$ substantially differ from each other, then the HARDI signals in voxel $\mathbf{v}'$ do not contain too much information on $\beta(\mathbf{v})$ and thus $\omega(\mathbf{v}, \mathbf{v}'; h)$ should be close to 0. However, if $\beta(\mathbf{v})$ and $\beta(\mathbf{v}')$ are close to each other indicating that the HARDI signals in voxel $\mathbf{v}'$ contain useful information on $\beta(\mathbf{v})$, then $\omega(\mathbf{v}, \mathbf{v}'; h)$ should be significantly larger than zero. See the explicit expression of $\omega(\mathbf{v}, \mathbf{v}'; h)$ in Section 3.2.3.

3.2.3 PMARM

We develop the PMARM procedure to adaptively determine $\mathbf{w}$ and estimate $\beta(\mathbf{v})$ across all voxels $\mathbf{v} \in \mathcal{V}$. Our multiscal adaptive strategy starts with building a sequence of nested spheres with increasing radii $h_0 = 0 < h_1 < \cdots < h_S = r$ ranging from the smallest scale $h_0 = 0$ to the largest scale $h_S = r$ at each voxel $\mathbf{v}$. At the scale $h_0 = 0$, we just calculate $\hat{\beta}(\mathbf{v}; h_0) = \widehat{\beta}(\mathbf{v})$ voxel-wisely without using any spatial information. This corresponds to setting $w(\mathbf{v}, \mathbf{v}'; h_0) = 1$ if $\mathbf{v} = \mathbf{v}'$ and 0 otherwise. Then, based on the signals contained in voxels d and d', we use methods as detailed below to calculate the weights $w(\mathbf{v}, \mathbf{v}'; h_1)$ at scale h_1 for all voxels $\mathbf{v}$. After getting the new weights $w(\mathbf{v}, \mathbf{v}'; h_1)$, we can update $\hat{\beta}(\mathbf{v}; h_1)$. Then we can sequentially determine $w(\mathbf{v}, \mathbf{v}'; h_s)$ and then adaptively estimate $\hat{\beta}(\mathbf{v}; h_s)$. From $h_0 = 0$ to $h_S = r$, a path diagram of the multiscale adaptive strategy is given below:

$$
\begin{array}{ccccccc}
w(\mathbf{v}, \mathbf{v}'; h_0) & & w(\mathbf{v}, \mathbf{v}'; h_1) & & \cdots & & w(\mathbf{v}, \mathbf{v}'; h_S = r) \\
\Downarrow & \nearrow & \Downarrow & \nearrow & \cdots & & \Downarrow \\
\hat{\beta}(\mathbf{v}; h_0) & & \hat{\beta}(\mathbf{v}; h_1) & & \cdots & & \hat{\beta}(\mathbf{v}; h_S)
\end{array}
$$

PMARM consists of three key steps: (I) an initialization step, (II) a weighted estimation step, and (III) a stop checking step. In the initialization step, we prefix a geometric series $\{h_s = c_h^s : s = 1, ..., S\}$ of radii with $h_0 = 0$, where $c_h \in (1, 2)$, say $c_h = 1.15$ and $S = 10$. We use a small c_h in order to prevent incorporating too many neighboring voxels at the beginning, and this improves the robustness of the procedure and the accuracy of the parameter estimation. At $h_0 = 0$, we solve the regularized linear least-squares optimization problem (3.6) for different penalty functions in order to calculate $\hat{\beta}(\mathbf{v}; h_0) = \widehat{\beta}(\mathbf{v})$ across all voxels $\mathbf{v}$. We then set $s = 1$ and $h_1 = c_h$.

In the weighted estimation step, we first compute $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ to characterize the similarity between the two estimated ODFs based on $\hat{\beta}(\mathbf{v}; h_s)$ and $\hat{\beta}(\mathbf{v}'; h_s)$ at voxels $\mathbf{v}$ and $\mathbf{v}'$ and the adaptive weights $\omega(\mathbf{v}, \mathbf{v}'; h_s)$, which are defined as

$$
\omega(\mathbf{v}, \mathbf{v}'; h_s) = \frac{K_{loc}(||\mathbf{v} - \mathbf{v}'||_2 / h_s) K_{st}(\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1}) / C_n)}{\sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} K_{loc}(||\mathbf{v} - \mathbf{v}'||_2 / h_s) K_{st}(\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1}) / C_n)}, \tag{3.11}
$$

where $K_{loc}(u)$ and $K_{st}(u)$ are two nonnegative kernel functions with compact support, C_n is a number associated with n, and $||\cdot||_2$ denotes the Euclidean norm of a vector (or a matrix).

We compute $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ as the similarity between the estimated ODFs in voxels $\mathbf{v}$ and $\mathbf{v}'$ for HARDI as follows. For instance, for QBI in Example 1, we transform $\hat{\beta}(\mathbf{v}; h_{s-1})$ to its corresponding ODF representation, denoted as $\widehat{ODF}(\mathbf{v}; h_{s-1})$, which is a linear combination of the SH basis functions. Since the SH basis is orthonormal, we may set $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1}) = ||\widehat{ODF}(\mathbf{v}; h_{s-1}) - \widehat{ODF}(\mathbf{v}'; h_{s-1})||_2$.

The weights $K_{loc}(||\mathbf{v} - \mathbf{v}'||_2/h_s)$ give less weight to the voxel $\mathbf{v}' \in B(\mathbf{v}, h_s)$, whose location is far from the voxel $\mathbf{v}$. The $K_{loc}(\cdot)$ is a regular kernel function for smoothing the smoothed curves or surfaces. Some common choices of $K_{loc}(\cdot)$ include the Gaussian kernel and Epanechnikov kernel (Tabelow et al. 2006; 2008, Polzehl and Spokoiny 2000). We use $K_{loc} = (1 - u^2)_+$ throughout this paper. The weights $K_{st}(\cdot)$ downweight the voxels that are dissimilar to voxel d. The $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ takes large values if the ODFs in voxel $\mathbf{v}$ differ significantly from those in voxel $\mathbf{v}'$. We set $K_{st} = \exp(-u^2/a)$, where a is a positive number.

After the calculation of $\omega(\mathbf{v}, \mathbf{v}'; h_s)$, we calculate the weighted HARDI signals of voxel $\mathbf{v}$, denoted by $\mathbf{y}_w(\mathbf{v}; h_s) = \sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} w(\mathbf{v}, \mathbf{v}'; h_s)\mathbf{y}(\mathbf{v}')$. Then, we use $\mathbf{y}_w(\mathbf{v}; h_s)$ to compute $\hat{\beta}(d; h_s)$ and $\widehat{ODF}(\mathbf{v}; h_s)$ at voxel $\mathbf{v}$. The computation of PMARM at each iteration is of the same order as that for the voxel-wise approach. Thus, this multiscale adaptive method provides an efficient method for adaptively exploring the neighboring voxels of each voxel. Since PMARM sequentially includes more data at each iteration, it will adaptively increase the statistical efficiency in estimating $\beta(\mathbf{v})$ in a homogenous region, while decreasing the variation of the weights $w(\mathbf{v}, \mathbf{v}'; h_s)$.

In the stop checking step, after the first iteration, we start to calculate a stopping criterion based on the L_2 distance between $\widehat{ODF}(\mathbf{v}; h_s)$ and $\widehat{ODF}(\mathbf{v}; h_{s-1})$, denoted by $\text{Dist}_s(\mathbf{v})$. We use $\text{Dist}_s(\mathbf{v})$ to determine whether 'bad' HARDI signals from neighboring voxels lead to a dramatic change in the estimated $\widehat{ODF}(\mathbf{v}; h_{s-1})$. If $\text{Dist}_s(\mathbf{v}) > C_s$, where C_s is a positive scalar, then we set $\widehat{ODF}(\mathbf{v}; h_s) = \widehat{ODF}(\mathbf{v}; h_{s-1})$ and $s = S$ for voxel $\mathbf{v}$. If $s = S$ for all voxels, we stop. If $\text{Dist}_s(\mathbf{v}) \leq C_s$, then we set $h_{s+1} = c_h h_s$, increase s by 1,

and continue with the weighted estimation step. In practice, different voxels may stop at different bandwidths, indicating that different degrees of smoothness are used to reconstruct HARDI.

We set $C_s = \chi^2(1)^{0.6/s}\bar{D}_{med}$ to prevent oversmoothing, where $\chi^2(1)^a$ is the upper $1-a$ percentile of the $\chi^2(1)$ distribution. As s increases, C_s decreases to zero. Moreover, $\bar{D}_{med}$ is chosen to be the median of $\{\text{Dist}(\mathbf{v}, \mathbf{v}'; h_0) : \mathbf{v} \neq \mathbf{v}'\}$, where $\mathbf{v}$ and $\mathbf{v}'$ are M preselected voxels $\{\mathbf{v}_i\}_{i=1}^M$ from HARDI. Specifically, we select these M voxels from regions with different Generalized Fractional Anisotropy (GFA) values. For instance, for QBI, if the ODF $\Phi(\mathbf{r})$ is represented by the SH basis with coefficients $\{c_{lm}\}$, it is shown in Özarslan et al. (2005) that the GFA can be represented by

$$\mathbf{GFA}\{\Phi_t(\mathbf{r})\} = \sqrt{1 - \frac{\beta_{00}^2}{\sum_{l=0}^{L}\sum_{m=-l}^{m=l}\beta_{lm}^2}}. \tag{3.12}$$

Finally, we summarize the PMARM algorithm 1 for the adaptive estimation of the ODF at voxel $\mathbf{v}$ below.

Algorithm 1: PMARM at voxel $\mathbf{v}$

Input: Signals $\mathbf{y}(\mathbf{v})$ and design matrix $\mathbf{X}$

Output: Estimated ODF $\widehat{ODF}(\mathbf{v}; h_S)$

Estimate $\hat{\beta}(\mathbf{v})$ from (3.6), $\lambda_{LB} = 0.006$, and $\lambda_{l_1} = 0.02$.

for $s \leftarrow 1$ ***to*** S **do**

- calculate the weights $w(\mathbf{v}, \mathbf{v}'; h_s)$ for $d' \in B(d, h_s)$ by (3.11);
- calculate the weighted signals of voxel $\mathbf{v}$ by using
- $\mathbf{y}_w(\mathbf{v}; h_s) = \sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} w(\mathbf{v}, \mathbf{v}'; h_s)\mathbf{y}(\mathbf{v}')$;
- calculate $\hat{\beta}(\mathbf{v}; h_s)$ based on (3.10);
- calculate $\widehat{ODF}(\mathbf{v}; h_s)$.
- If $\text{Dist}_s(\mathbf{v}) > C_s$, $\widehat{ODF}(\mathbf{v}; h_S) = \widehat{ODF}(\mathbf{v}; h_s)$, and $s = S$,
- **else**
  - $h_{s+1} = c_h h_s$;

return $\widehat{ODF}(\mathbf{v}; h_S)$.

3.2.4 Maxima Extraction

Based on $\widehat{ODF}(\mathbf{v}; h_S)$ at voxel $\mathbf{v}$, we need to extract its maxima in order to infer fiber directions. Although there are other more complicated methods for extracting ODF maxima, such as the method presented in (Hlawitschka and Scheuermann 2005), spherical Newton's method (Tuch 2004), and Powell's method (Jansons and Alexander 2003), we take a simple thresholding approach in this paper. Specifically, we project the estimated ODF onto the sphere tessellated with a triangle mesh, which has 2562 points on the unit sphere. If the estimated ODF value at a mesh point is greater than the corresponding value at all its neighboring mesh points and this estimated ODF value is greater than max(ODF)/2, then the direction at this mesh point is regarded as a maximum. This thresholding method avoids selecting small peaks that may appear due to noise.

3.3 Simulation Studies and Two Real Examples

In this section, we use Monte Carlo simulations and two real examples to evaluate the finite-sample performance of PMARM and compare PMARM with other estimation methods. All computations for these numerical examples were done in Matlab on an IBM ThinkCentre M50 workstation. The computation for PMARM is relatively efficient. The computational time for PMARM can be further reduced by using other computer languages, such as C++.

3.3.1 Simulation Studies

We examined the finite sample performance of our PMARM on decting crossing fibers by using synthetic HARDI data generated from the multi-tensor model (Alexander et al. 2002, Tuch 2004). We simulated the diffusion-weighted signals according to

$$E(\mathbf{q}_i; \mathbf{v}) = \sqrt{(\sum_{t=1}^{T} p_t e^{-b_i \mathbf{u}_i^T \mathbf{D}_t(\mathbf{v}) \mathbf{u}_i} + \sigma\epsilon_{i1})^2 + (\sigma\epsilon_{i2})^2} \tag{3.13}$$

for $i = 1, \cdots, n$, where $\mathbf{q}_i = q_i \mathbf{u}_i$ with $\mathbf{u}_i$ being a unit vector, T is the number of fibers, p_t is the weight for t-th fiber, b is the b-value and $\mathbf{D}_t(\mathbf{v})$ is the tensor matrix for the t-th fiber, SNR $= 1/\sigma$, and ϵ_{i1} and ϵ_{i2} are independently simulated from the standard normal distribution. We used the multi-tensor model (3.13) to generate different 10×10 phantoms with different regions of interest (ROIs) with 81 sampling directions on the hemisphere for the 3rd order tessellation of the icosahedron and $b = 2000\text{s}/\text{mm}^2$. Specifically, voxels with a single fiber were generated from a single tensor model using diffusion tensor profiles with eigenvalues $[1.7, 0.3, 0.3] \times 10^{-3}\text{mm}^2/\text{s}$, voxels with two fiber directions were generated by two-tensor model $E(\mathbf{q}_i; \mathbf{v}) = e^{-b_i \mathbf{u}_i^T \mathbf{D}_1(\mathbf{v}) \mathbf{u}_i}/2 + e^{-b_i \mathbf{u}_i^T \mathbf{D}_2(\mathbf{v}) \mathbf{u}_i}/2$, and voxels isotropic tensors were generated by the single tensor model using diffusion tensor profiles with eigenvalues $[1, 1, 1] \times 10^{-3}\text{mm}^2/\text{s}$. We estimated ODF at each voxel by using the three voxel-wise estimation methods including cQBI, LASSO and LDPE, and their corresponding three PMARMs including p-cQBI, p-LASSO and p-LDPE . For PMARM, we set $K_{st} = \exp(-u^2/4)$, $C_n = 1$ and $\lambda = 0.02$ in LASSO. Then we extracted the ODF maxima aligned with fiber directions.

Angle Detection in First Phantom with 90^o crossing fibers

In the first phantom, we included four different ROIs including isotropic ROIs, two single fiber ROIs with its direction going either along the $x-$axis (ROI1) or along the $y-$axis (ROI2), and the 90^o crossing fiber ROIs(ROI3); Figure 3.2 presents the estimated ODF images for this type of phantom. The left panel on the top row presents the recovered ODF from noise free data, whereas the other three panels on the top row present those from the data with SNR=10 by using cQBI, LASSO, and LDPE, respectively. Three panels on the bottom row present the recovered ODFs from the same dataset by using p-cQBI, p-LASSO and p-LDPE. Generally, PMARMs outperform the voxel-wise methods in terms of detecting the isotropic regions and consistently recovering the ODFs with fiber crossing.

To quantify the accuracy of detection angle, we generated 1,000 data sets for three different SNRs including 10 ,15, and 20. We estimated the ODFs by using voxel-wise cQBI,

LASSO and LDPE and their corresponding three PMARMs. Then we extracted the ODF maxima aligned with fiber directions. For voxels with a single fiber, we calculated angle detection errors by comparing recovered fiber directions with the ground truth. For voxels with two crossing fibers, we calculated angle detection errors by comparing recovered crossing angles with the ground truth. Mean angle detection errors at each voxel are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for each ROI are presented in Table 3.1.

Table 3.1 reveals that the mean angle errors are substantially reduced for the three PMARM methods. It may indicate that PMARM can efficiently exploit spatial smoothness for reconstructing ODFs, while reducing noise leading to better angle detection. Among the three methods, LASSO and LDPE outperform cQBI in terms of the mean angle error, since LASSO and LDPE force smaller ODF coefficients to be zero, leading to a more stable recovery of ODFs. Moreover, LDPE outperforms LASSO in terms of the mean angle error, since non-zero coefficients of LDPE are unbiased compared with LASSO estimators (Zhang and Zhang 2014).

Angle Detection in Second Phantom with Four Quadrants of ROIs

In the second phantom, we included four different types of regions of interest (ROIs) including isotropic ROIs, two single fiber ROIs with its direction going either along the $x-$axis (ROI1) or along the $y-$axis (ROI2), and the 90^o crossing fiber ROIs (ROI3). See Figure 3.3 for details. We used the same estimation methods and SNRs based on 1,000 simulated data sets. Table 3.2 presents the mean angle errors for the voxel-wise methods and their corresponding PMARMs.

Angle Detection in Third Phantom with 75^o Crossing Fibers

In the third phantom, we included four different types of regions of interest (ROIs) including isotropic ROIs, two single fiber ROIs with its direction from left to right going up (ROI1) or from left to right going down (ROI2), and the 75^o crossing fiber ROIs(ROI3);

We used the same setting as previous two phantom images simulations. Figure 3.4 presents the estimated ODF images for this phantom. Table 3.3 includes the average values of these detection errors for each ROI, indicating substantial reduction in mean angle errors by using PMARM.

Angle Detection in Fourth Phantom with Twisted Crossing

In the fourth phantom, we included a twisted crossing region in the middle, one fiber along $x-$axis (ROI1), and the other fiber with changing angles with $x-$axis from 30^o, 45^o , 60^o, 75^o, to 90^o and then from 90^o, 75^o, 60^o, 45^o to 30^o. We marked all region with single fiber, which is not along $x-$axis as ROI2, and all regions with crossing fibers as ROI3. We used the same setting as the previous phantom simulations. Figure 3.5 presents the estimated ODF images. Table 3.4 includes the average values of these detection errors for each ROI, indicating substantial reduction in mean angle errors by using PMARM.

3.3.2 Summary of Simulations

We have the following findings. First, the three PMARM methods outperform the three voxel-wise methods in all ROIs, especially in ROIs with crossing fibers. Second, the estimated ODFs in the single fiber ROIs are sharper and it is easier to detect the maxima. Third, the L_1 penalty based methods outperform the methods based on the Laplacian-Beltrami regularization in the isotropic ROIs, since in the isotropic regions, the ODF coefficients are small and tend to be suppressed to zero.

3.3.3 Pig Brain

The pig brain data set comes from a post-mortem porcine brain and was kindly provided by Tim Dyrby from The Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital, Hvidovre, Denmark (Dyrby et al. 2011). The acquisition uses a spherical acquisition scheme with 61 unique gradient directions and b-value of 3146 s/mm^2. Each diffusion weighted image has 10 slices with in-plane resolution 128×128 with voxel size

$0.5 \times 0.5 \times 0.5$ mm^3.

We used model 3.5 to estimate the ODF by using voxel-wise cQBI, LASSO and LDPE and their corresponding three PMARMs. We set $K_{st} = \exp(-u^2/4)$ and $\lambda = 0.02$ in LASSO. The results are shown in Figure 3.6. All PMARM methods lead to better ODF reconstruction results, in terms of smoother ODFs along the fiber tract. In the regions with fiber crossings, the main fibers are easier to detect than less small noisy fibers. Since in PMARM, we exploit the spatial constraints to estimate ODFs, PMARM can reduce the noise level in HARDI, while improving the ODF reconstruction, especially in some ROIs. Generally, PMARM can improve the recovery of crossing fibers.

3.3.4 Human Brain

We also tested our methods in a real human data set with b-value 3000 s/mm^2, 140 gradients, dimension $128 \times 96 \times 60$, and voxel size $2 \times 2 \times 2mm^3$. We used model 3.5 to estimate ODF by using voxel-wise cQBI, LASSO and LDPE and their corresponding three PMARMs. We set $K_{st} = \exp(-u^2/4)$ and $\lambda = 0.02$ in LASSO. The results are shown in Figure 3.7.

We can see from the figure that ODFs along the fiber tracts are smoother and regional seperations are clearer in PMARMs results. Because all PMARM methods can achieve noise reduction and obtain more accurate ODF reconstruction results in HARDI by exploiting the spatial constraints to estimate ODFs, especially in some gray matter ROIs, PMARM can improve the recovery of isotropic regions.

3.4 Conclusion

We have introduced a penalized multiscale adaptive model (PMARM) framework to adaptively reconstruct the ODF across all voxels from HARDI signals. PMARM reconstructs the ODF at each voxel by adaptively borrowing the spatial information from the neighboring voxels. We have shown in the real and simulated data sets that PMARM can substantially reduce the noise level, while improving the ODF reconstruction. We have shown that the

L_1 penalty function outperforms the Laplacian-Beltrami regularization and leads to sparse ODF solution, and can better detect the isotropic regions.

3.5 Appendix

We review the Spherical Harmonics (SH), normally indicated by H_l^m, where l denotes the order and m is the phase factor. The SH basis is a function basis for complex functions on the unit sphere. Specifically, the SH basis functions are given as follows:

$$H_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi},$$

where $\theta \in [0,\pi]$, $\phi \in [0,2\pi]$, and P_l^m is an associated Legendre polynomial. For $k = 0,2,4,...,l$ and $m = -k,...,0,...,k$, we define the new index $j : j(k,m) = (k^2+k+2)/2+m$ and define a modified SH basis as follows:

$$Y_j = \begin{cases} \sqrt{2}\cdot Re(H_k^{|m|}) & if \;\; m<0 \\ H_l^0 & if \;\; m=0 \\ \sqrt{2}(-1)^{m+1} Im(H_k^m) & if \;\; m>0 \end{cases}$$

where $Re(Y_k^m)$ and $Im(Y_k^m)$ represent the real and imaginary parts of Y_k^m respectively. The basis is designed to be symmetric, real, and orthonormal (Descoteaux et al. 2006).

Table 3.1: Mean angle errors of different ROIs in the first phantom with 90 degree crossing. 1,000 simulated data sets were used.

ROI1: fibers along $x-$axis								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.48	1.47	2.06	0.17	0.58	0.31
2000	4	15	2.46	0.68	0.9	0.02	0.42	0.24
2000	4	20	1.92	0.34	0.42	0.003	0.41	0.13
2000	6	10	4.94	2.55	5.44	1.24	1.53	1.04
2000	6	15	3.69	1.64	3.36	0.31	1.02	0.63
2000	6	20	2.9	1.01	2.06	0.07	0.89	0.51

ROI2: fibers along $y-$axis								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.48	1.28	2.05	0.11	0.62	0.26
2000	4	15	2.47	0.55	0.85	0.01	0.45	0.16
2000	4	20	1.93	0.26	0.38	0.001	0.39	0.11
2000	6	10	4.95	2.34	5.51	1.07	1.56	1.02
2000	6	15	3.7	1.43	3.39	0.22	1.09	0.60
2000	6	20	2.92	0.84	2.04	0.05	0.92	0.49

ROI3: 90^o crossing								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.61	1.63	2.52	0.1	2.14	0.19
2000	4	15	2.29	0.88	0.70	0.01	0.23	0.10
2000	4	20	1.78	0.41	0.21	0.001	0.12	0.07
2000	6	10	10.56	4.76	11.27	1.60	3.95	1.59
2000	6	15	7.12	2.99	5.57	0.16	1.66	0.66
2000	6	20	5.26	2.04	2.37	0.02	1.14	0.47

Table 3.2: Mean angle errors of different ROIs in the second phantom with four quadrants. 1,000 simulated data sets were used.

ROI1: fibers along $x-$axis								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.48	1.33	2.05	0.79	0.56	0.26
2000	4	15	2.46	0.57	0.89	0.07	0.40	0.16
2000	4	20	1.92	0.27	0.41	0.01	0.39	0.064
2000	6	10	4.97	2.41	5.50	5.00	1.51	1.01
2000	6	15	3.70	1.53	3.42	2.09	1.04	0.55
2000	6	20	2.93	0.91	2.09	0.61	0.89	0.46

ROI2: fibers along $y-$axis								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.47	1.35	2.06	0.81	0.64	0.24
2000	4	15	2.47	0.58	0.84	0.07	0.45	0.13
2000	4	20	1.92	0.28	0.39	0.01	0.39	0.092
2000	6	10	4.91	2.37	5.46	4.96	1.54	1.05
2000	6	15	3.67	1.51	3.37	2.06	1.08	0.57
2000	6	20	2.89	0.89	2.05	0.60	0.92	0.46

ROI3: 90^o crossing								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.63	1.95	2.56	0.84	2.05	0.19
2000	4	15	2.29	1.24	0.71	0.02	0.26	0.10
2000	4	20	1.79	0.71	0.22	0.001	0.13	0.08
2000	6	10	10.64	5.59	11.38	10.08	3.86	1.82
2000	6	15	7.13	3.64	5.67	2.50	1.69	0.75
2000	6	20	5.24	2.65	2.43	0.21	1.15	0.54

Table 3.3: Mean angle errors of different ROIs in the third phantom with 75 degree crossing. 1,000 simulated data sets were used.

ROI1: from left to right going up								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.60	1.94	3.68	1.85	5.37	3.05
2000	4	15	2.73	1.33	2.72	1.14	3.95	1.55
2000	4	20	2.30	1.06	2.22	0.94	3.04	1.03
2000	6	10	4.97	2.82	6.65	3.58	6.42	4.44
2000	6	15	3.80	2.08	5.04	2.41	4.91	2.51
2000	6	20	3.10	1.62	4.00	1.66	3.92	1.96

ROI2: from left to right going down								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.61	1.88	3.69	1.81	5.37	2.99
2000	4	15	2.74	1.29	2.74	1.12	3.98	1.53
2000	4	20	2.31	1.03	2.23	0.93	3.02	1.07
2000	6	10	4.99	2.78	6.69	3.55	6.37	4.33
2000	6	15	3.82	2.05	5.06	2.34	4.91	2.50
2000	6	20	3.12	1.58	4.02	1.60	3.91	1.97

ROI3: 75^o crossing								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	7.93	5.75	9.24	7.71	24.52	9.99
2000	4	15	6.01	5.61	7.64	6.94	10.08	8.44
2000	4	20	5.81	5.4	6.77	5.91	8.95	7.84
2000	6	10	9.19	4.77	10.95	6.52	25.27	11.37
2000	6	15	6.44	3.55	8.12	5.75	11.09	8.94
2000	6	20	5.14	3.08	6.73	5.80	9.43	8.17

Table 3.4: Mean angle errors of different ROIs in the fourth phantom with twisted crossing. 1,000 simulated data sets were used.

ROI1: fibers along $x-$axis								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.47	2.08	2.03	1.10	2.77	1.26
2000	4	15	2.46	1.02	0.85	0.24	2.01	1.06
2000	4	20	1.92	0.59	0.39	0.08	1.61	0.65
2000	6	10	4.95	3.67	5.49	5.15	6.40	3.72
2000	6	15	3.69	2.14	3.36	2.42	5.06	4.69
2000	6	20	2.91	1.38	2.03	0.96	4.18	3.48

ROI2: single fibers except the ones along $x-$axis								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	6.03	4.46	6.83	5.31	6.91	4.46
2000	4	15	4.29	3.54	4.67	4.03	4.87	4.18
2000	4	20	3.94	3.35	4.47	3.92	4.68	4.14
2000	6	10	8.91	7.11	9.84	9.31	10.39	7.89
2000	6	15	6.49	4.51	6.49	5.08	8.06	7.49
2000	6	20	5.13	3.78	4.87	3.84	6.70	5.57

ROI3: twisted crossing								
b value	SH order	SNR	cQBI	p-cQBI	LASSO	p-LASSO	LDPE	p-LDPE
2000	4	10	3.69	3.18	3.04	2.52	3.5	3.03
2000	4	15	2.85	2.61	2.00	1.9	2.63	2.5
2000	4	20	2.43	2.39	1.59	1.28	2.24	2.22
2000	6	10	5.05	4.31	5.87	5.55	6.67	4.82
2000	6	15	3.89	3.27	4.01	3.37	5.28	5.04
2000	6	20	3.22	2.83	2.85	2.28	4.41	4.02

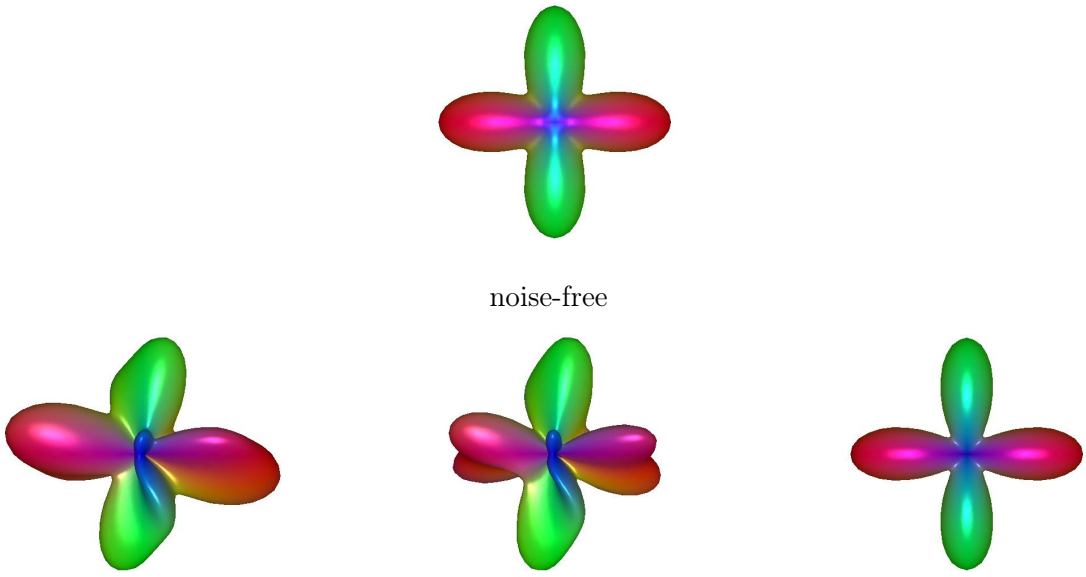


Figure 3.1: Simulation results: in the first row, ODF reconstruction of noise-free data; in the second row, ODFs of the data with noise(SNR=10) using cQBI, Lasso and LDPE, respectively.

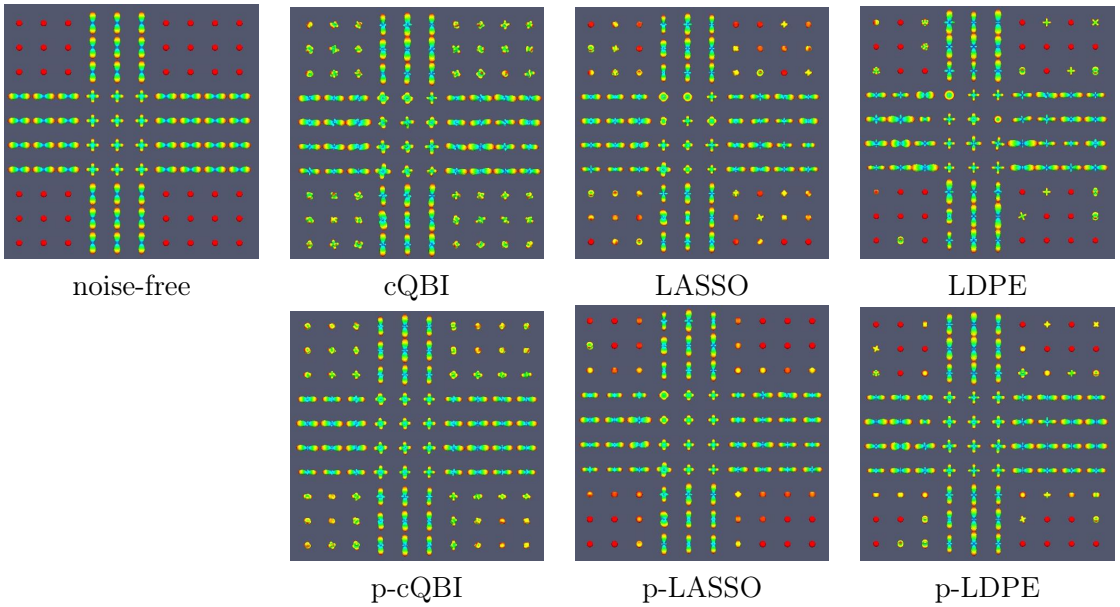


Figure 3.2: Simulation results for the first phantom: ODF reconstruction results of simulated data with 90 degree crossing.

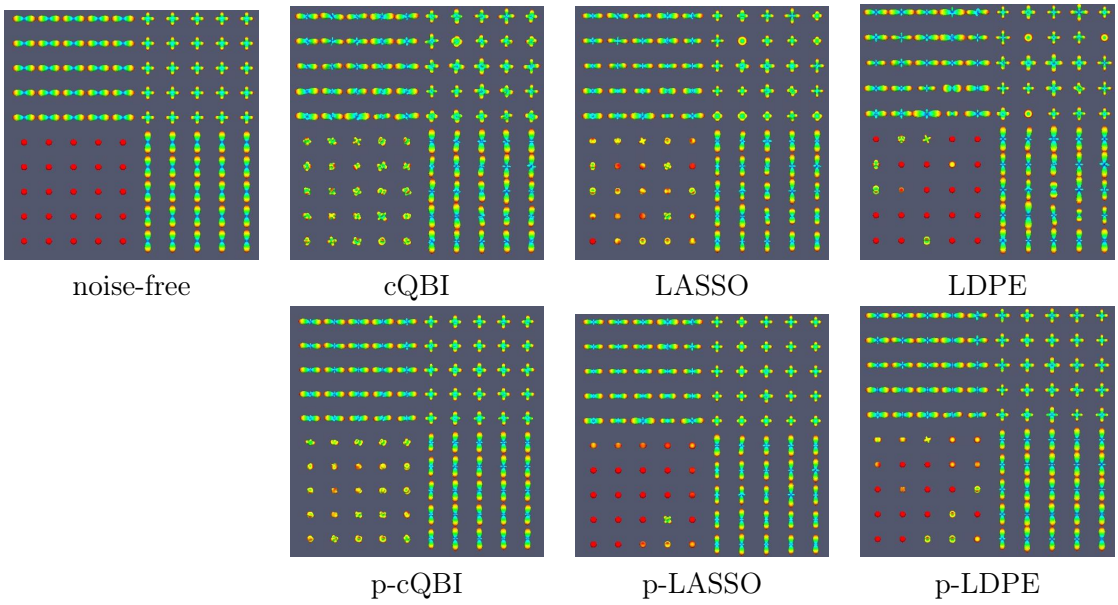


Figure 3.3: Simulation results for the second phantom: ODF reconstruction results on simulated data with four quadrants.

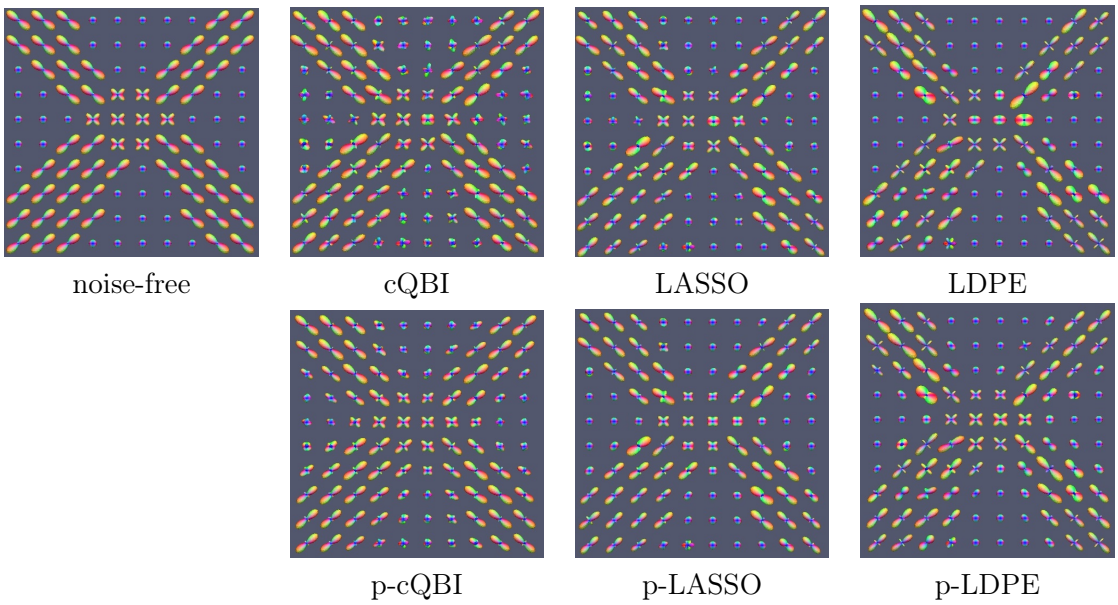


Figure 3.4: Simulation results for the third phantom: ODF reconstruction results on simulated data with 75 degree crossing.

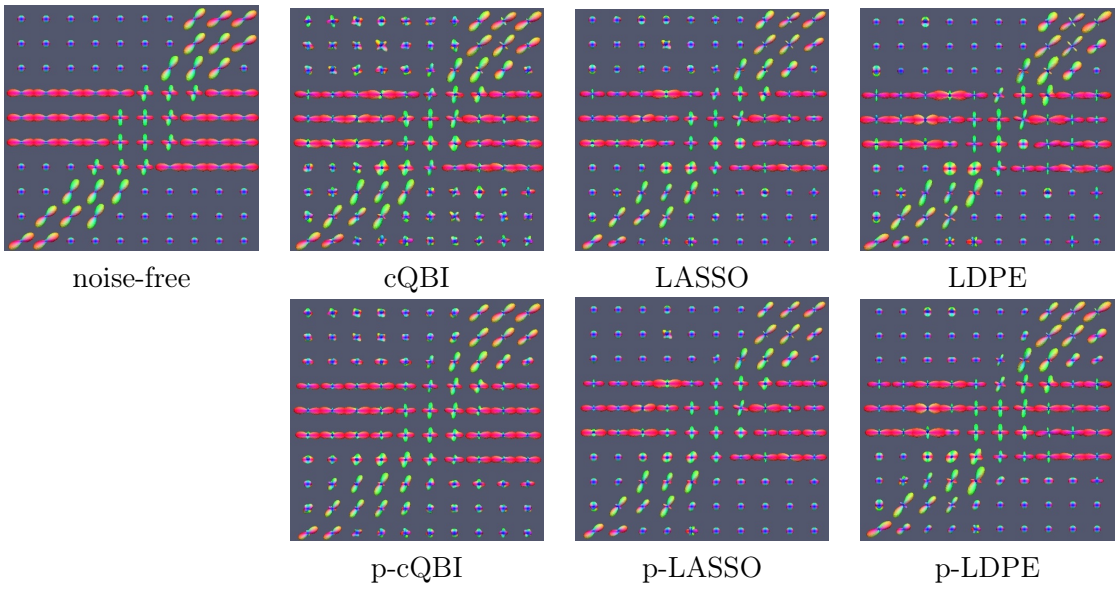


Figure 3.5: Simulation results for the fourth phantom with: ODF reconstruction results on simulated data with twisted crossing.

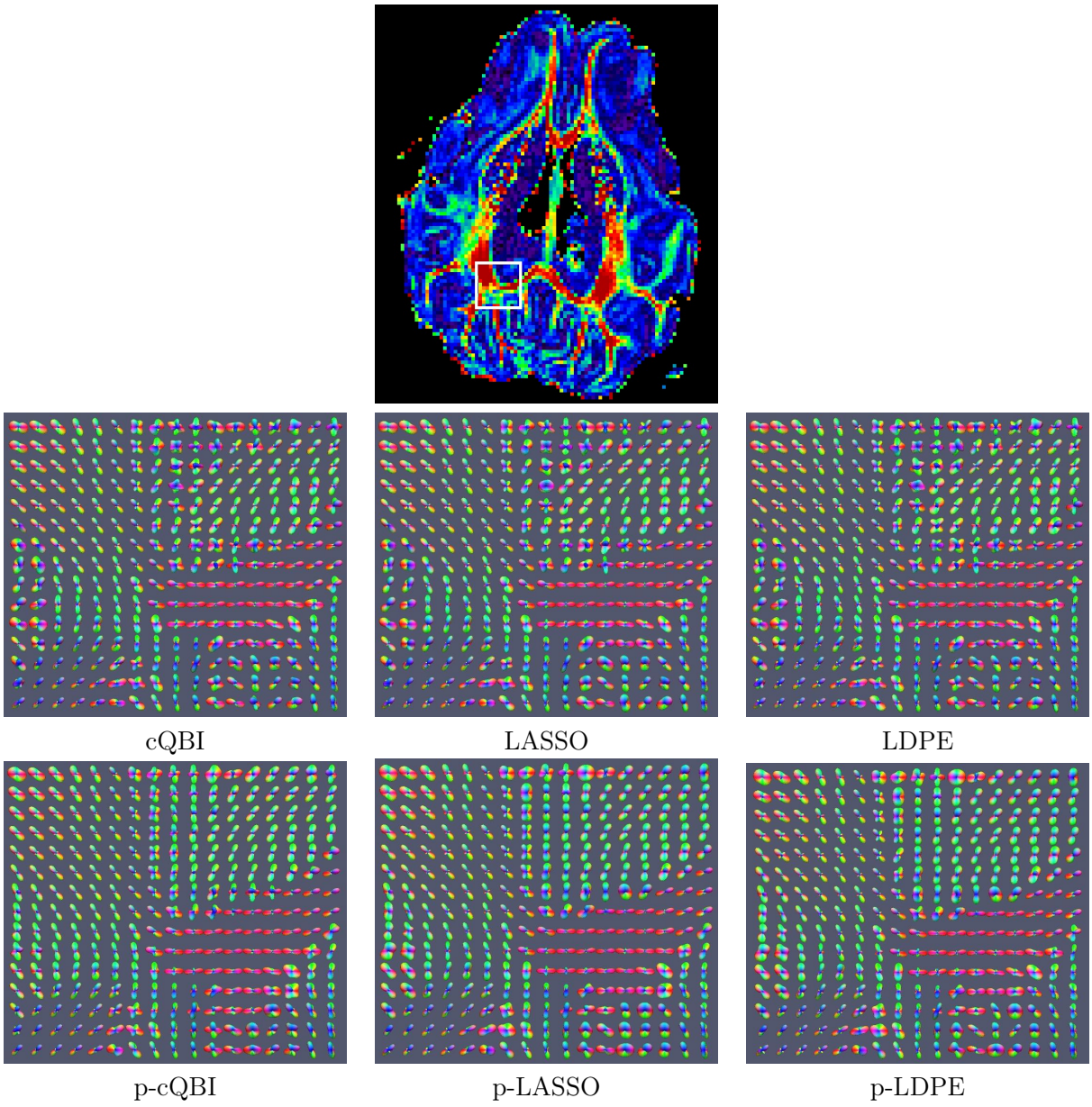


Figure 3.6: ODF reconstruction results for the pig dataset: the first row shows the region of interest on GFA map; the second row shows the ODF in the selected ROI without PMARM; and the third row shows the ODF with PMARM. All the ODFs are min-max normalized.

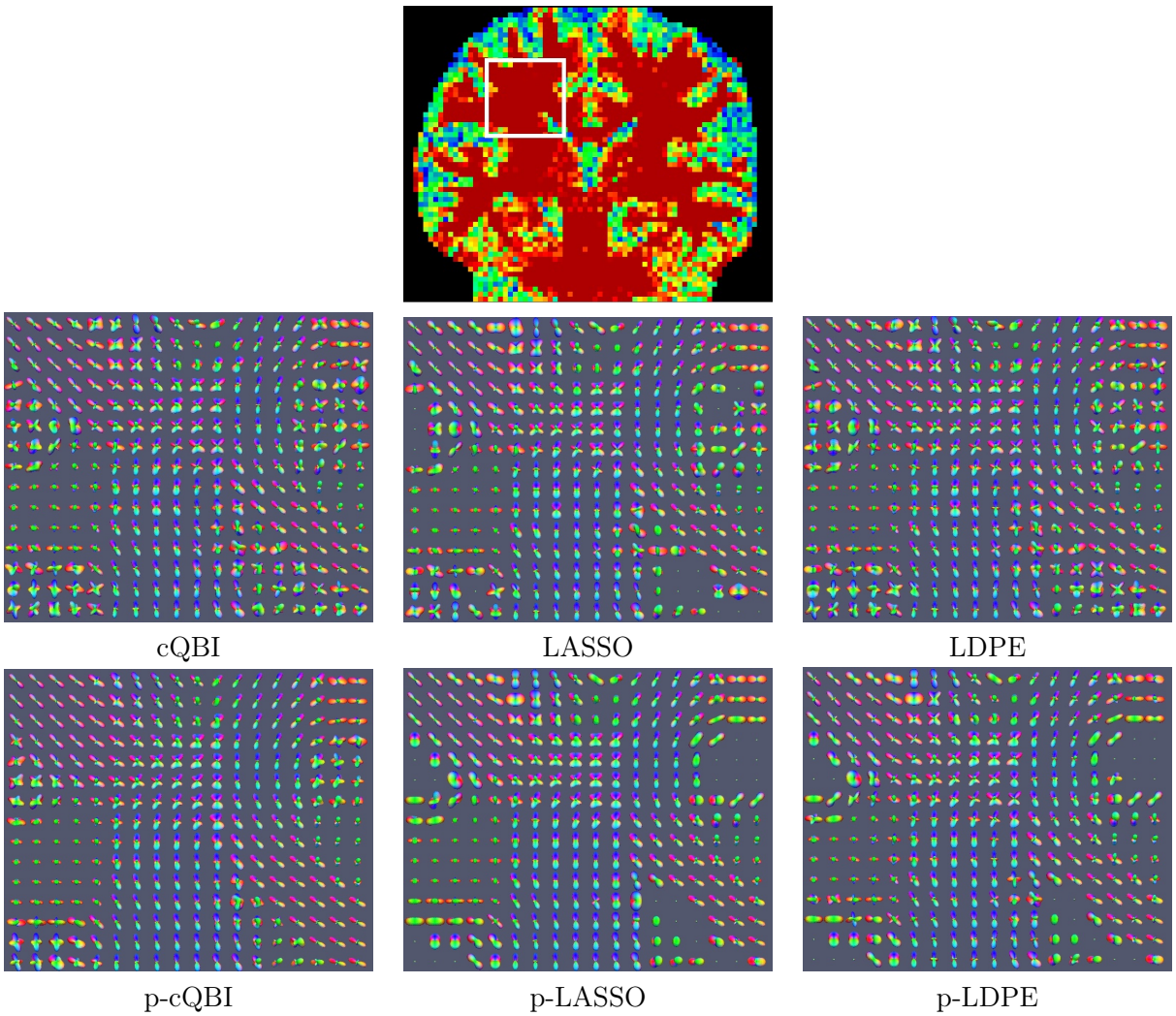


Figure 3.7: ODF reconstruction results for the human dataset: the first row shows the region of interest on GFA map; the second row shows the ODF in the selected ROI without PMARM; the third row shows the ODF with PMARM. All the ODFs are min-max normalized.

CHAPTER 4: ROBUST AND SPATIALLY ADAPTIVE EAP RECONSTRUCTION

Many recent high angular resolution diffusion imaging (HARDI) reconstruction techniques have been introduced to infer ensemble average propagator (EAP),describing the three-dimensional (3D) average diffusion process of water molecules. Recently, several analytical EAP reconstruction schemes have been proposed, for example,Diffusion Orientation Transform (DOT), Diffusion Specturm Imaging (DSI) and Spherical Polar Fourier Imaging (SPFI) and so on. Among which, DSI and SPFI are two important methods to estimate the EAP from the signal. Especially, SPFI is a model-free fast analytical EAP reconstruction method, which does not need any assumption of data and does not need too many samplings. However, current existing methods perform reconstruction independently at each voxel by minizing the square error loss function, which is very sensitive to noise and outliers. This essentially ignoring the functional nature of the HARDI data at different voxels in space.

The aim of this paper is to propose a robust multi-scale adaptive and sequential smoothing (MASS) method framework to robustly, spatially and adaptively infer the EAP of water diffusion in regions with complex fiber configurations. In robust MASS, we reformulate the HARDI imaging reconstruction as a robust regression problem using Huber's loss function. Similarity and distance weights are introduced to account for spatial smoothness of HARDI, while preserving the unknown discontinuities (e.g., edges between white matter and grey matter) of HARDI. We integrate robust regression estimation with the propagation-separation method (Polzehl and Spokoiny, 2000) to adaptively estimate EAPs across voxels. Experimental results indicate that MASS can reduce the angle detection errors on fiber crossing area and provides more accurate reconstructions than standard voxel-wise methods and robust MASS performs very well with the presence of outliers.

4.1 Introduction

Diffusion Magnetic Resonance Imaging (dMRI) is a widely used in-vivo imaging technique to explore the information of neural micro-structure by probing the diffusion of water molecules. So far it is still the unique non-invasive method to reveal the micro-geometry of nervous tissues noninvasively and to explore the neural connectome in living human subjects. The diffusion of water molecules is constrained by the surrounding structures including nerves, cells and surrounding tissue. For example, qualitatively water molecules diffuse fast along fibers and slowly cross fibers. Thus measuring the diffusion process quantitatively is crucial to understanding the neural micro-structure and fiber directions.

The quest of diffusion-weighted (DW) imaging is to non-invasively recover information about the diffusion of water molecules in biological tissues (Le Bihan et al. 2003). Many recent high angular resolution diffusion imaging (HARDI) techniques (Descoteaux 2008, Alexander 2005) have been proposed to recover the complex white matter geometry. However, these orientation functions derived from single-shell HARDI (one b-value) only capture the angular structure of the diffusion process and are therefore typically used for fiber tractography applications. Full threedimensional (3D) ensemble average propagator (EAP) describing the diffusion process to obtain richer information on the complex microstructure of biological tissues.

The EAP formalism provides a powerful framework to describe and predict the diffusion behavior in complex materials (Tuch 2002). Under the narrow pulse assumption(Stejskal and Tanner 1965), the relationship between the diffusion signal attenuation,$E(\mathbf{q};\mathbf{v})$, in q-space and the EAP, $p(\mathbf{R};\mathbf{v})$, in real space at each voxel $\mathbf{v}$ in a common space $\mathcal{V}$, where $\mathbf{q} = q\mathbf{u} \in R^3$ and $\mathbf{R} = R\mathbf{r} \in R^3$, is given by a Fourier transform (FT) relationship (Callaghan 1991) such that

$$p(\mathbf{R};\mathbf{v}) = \int_{\mathbf{q}\in\mathbf{R^3}} E(\mathbf{q};\mathbf{v}) e^{-2\pi i \mathbf{q}\cdot\mathbf{R}} d\mathbf{q} \tag{4.1}$$

Various methods already exist to reconstruct the EAP or estimate EAP features. The most common and famous model, is the diffusion tensor model (Basser et al. 1994b). Although

very successful in many neuroscience studies, diffusion tensor imaging (DTI) is limited by the Gaussian assumption (free diffusion model) and cannot account for complex fiber configurations. In free diffusion, diffusion behaves according to Fick's first law, often called Gaussian diffusion, and the diffusion tensor model is appropriate. However, such diffusion excludes observed in vivo phenomena such as restriction, heterogeneity, anomalous diffusion, and finite boundary permeability (Tuch 2002). Recent modeling techniques suggest the composite and hindered restricted model of diffusion (CHARMED) (Assaf et al. 2004), the "ball & multi-stick" model (Behrens et al. 2007), or other more sophisticated models (Alexander 2008). Note that CHARMED is designed to use multiple spherical shell diffusion data.

A large family of techniques, mostly based on different mathematical representation of the signal, use multiple q-shell acquisitions in order to reconstruct signal features or EAP features, such as generalized high order tensors (Liu et al. 2004) based on cumulant expansions; or the diffusion orientation transform (DOT) (Özarslan et al. 2006); or a fourth order Cartesian tensor representation of the probability profile (Barmpoutis et al. 2008);or better diffusion ODFs (Aganj et al. 2009, Tristán-Vega et al. 2009b) than obtained from q-ball imaging (QBI) (Tuch et al. 2002). Unfortunately, for most of these methods, many DW measurements are still needed. Moreover, most of these methods do not recover the full EAP but features of it, making several assumptions that remain to be validated

Diffusion Propagator Imaging (DPI) was proposed to model the signal $E(\mathbf{q};\mathbf{v})$ as the solution of Laplace's equation (Descoteaux et al. 2009; 2010). In DPI, the signal is assumed to be

$$E(q\mathbf{u};\mathbf{v}) = \sum_{l=0}^{L}\sum_{m=-l}^{l}(\frac{c_{lm}}{q^{l+1}} + d_{lm}q^l)Y_l^m(\mathbf{u}) \tag{4.2}$$

Then the EAP is estimated from incomplete 3D integration inside the ball with a given radius q_{max}, because the complete integration in $\mathbb{R}^3$ does not converge.

$$\begin{aligned} p(R_0\mathbf{r};\mathbf{v}) &= \frac{1}{Z}\int_0^{q_{max}}\int_{\mathbb{S}^2} E(q\mathbf{u};\mathbf{v})q^2 e^{-2\pi i q R_0 \mathbf{u}^\mathbf{T}\mathbf{r}} dq d\mathbf{u} \\ &= \frac{1}{Z}\sum_{l=0}^{L}\sum_{m=-l}^{m=l}(p_{0l}(R_0)c_{lm} + p_{1l}(R_0)d_{lm})Y_l^m(\mathbf{r}) \end{aligned} \tag{4.3}$$

Where $p_{0l}(R_0)$ and $p_{1l}(R_0)$ are given in (Descoteaux et al. 2010), Z is the normalization factor, q_{max} is the maximum q value used in DPI acquisition. The coefficents $\{c_{lm}\}$ and $\{d_{lm}\}$ can be calculated from DWI samples via a standard least square estimation, then we have the representation for EAP. However DPI is a model-based method, because it assumes $\Delta E(\mathbf{q};\mathbf{v}) = \mathbf{0}$

Then (Cheng et al. 2010) proposed a Model-free and Analytical EAP Reconstruction via Spherical Polar Fourier Diffusion MRI. Spherical Polar Fourier Imaging (SPFI) was first proposed by Dr. Assemlal in (Assemlal et al. 2008; 2009, Assemlal 2010). It represents the diffusion signal $E(q\mathbf{u};\mathbf{v})$ with Spherical Polar Fourier basis denoted by B_{klm}^{SPF}

The aim of this paper is to develop a robust multi-scale adaptive and sequential smoothing (MASS) framework to spatially and adaptively infer EAPs across all voxels. Similar to (Raj et al. 2011), robust MASS is also a simultaneous reconstruction and denoising procedure. However, MASS differs significantly from the method in (Raj et al. 2011). First, we use similarity and distance weights to account for spatial smoothness of HARDI, while preserving the unknown discontinuities. In contrast, the spatial HARDI method use a smoothness regularization term. Second, robust MASS uses robust method to estimate the EAPs, which is more stable again the outliers, wheres the spatial HARDI enforces Tikhonov regularization in order to stabilize the estimated EAPs. Third, robust MASS integrates WRLSR with the propagation-separation method to sequentially and adaptively estimate EAP at each voxel.

Section 4.2 of this paper presents MASS for HARDI reconstruction. In Section 4.3, we conduct simulation studies with the known ground truth to examine the finite sample performance of MASS on robust regression and least square estimators. Section 4.4 illustrates an application of the proposed methods in a real neuroimaging dataset. We present concluding remarks in Section 4.5.

4.2 Methods

4.2.1 Model Formulation

We usually acquire n normalized HARDI data with each image containing N voxels for each subject. Thus, we observe n normalized HARDI measurements $\{(E(b_i;\mathbf{v}),\mathbf{g}_i,b_i) : i=1,\cdots,n\}$ at voxel $\mathbf{v}\in\mathcal{V}$, where $\mathbf{g}_i=(g_{i,1},g_{i,2},g_{i,3})^T$ is the gradient vector. Based on (2.3), these HARDI measurements can be also represented as $\mathbf{q}$-space measurements $\{E(\mathbf{q}_i) : i=1,\cdots,n\}$. We usually omit putting voxel $\mathbf{v}$, if no confusion in context. See Figure 2.8. Most HARDIs assume that

$$f(E(\mathbf{q}_i;\mathbf{v}))=\mathbf{x}_i^T\beta(\mathbf{v})+\epsilon_i(\mathbf{v}), \tag{4.4}$$

where $f(\cdot)$ is a given transformation function (e.g., $f(s)=s$ or $f(s)=\log(s)$), $\mathbf{x}_i$ is a $p\times 1$ vector of covariates, which depends on $\mathbf{q}_i$ (or $(\mathbf{b}_i,r_i)$), $\beta(\mathbf{v})$ is a $p\times 1$ vector of regression coefficients, and $\epsilon_i(\mathbf{v})$ is an error term with mean zero and variance $\sigma_i^2(\mathbf{v})$. In practice, $E(\mathbf{q}_i;\mathbf{v})$ equals the ratio of magnetic resonance signal measured at $\mathbf{q}_i$, denoted by $S(\mathbf{q}_i;\mathbf{v})$, to the magnetic resonance signal measured at $\mathbf{0}$, denoted by $S(\mathbf{0};\mathbf{v})$. Since the signal-to-noise ratio in $S(\mathbf{0};\mathbf{v})$ is very high, we ignore the noise component of $S(\mathbf{0};\mathbf{v})$. Model 4.4 is general enough to cover many existing HARDIs. In the literature, for generalized DTI and high order tensor (HOT), it is common to set $f(E(\mathbf{q}_i;\mathbf{v}))=\log(E(\mathbf{q}_i;\mathbf{v}))$ and represent $\log(E(\mathbf{q}_i;\mathbf{v}))$ as a polynomial function of $\mathbf{q}_i$, whereas for most other HARDIs, such as Q-ball imaging (QBI) or diffusion orientation transform (DOT), it is common to set $f(E(\mathbf{q}_i;\mathbf{v}))=E(\mathbf{q}_i;\mathbf{v})$ and approximate $E(\mathbf{q}_i;\mathbf{v})$ by a linear combination of some basis functions.

Spherical polar Fourier imaging (SPFI) is a model-free and fast HARDI method for multiple-shell data (Cheng et al. 2010). Let $B_{k,l,m}(\mathbf{q})=G_k(q)Y_l^m(\mathbf{u})$ be spherical polar Fourier basis, where $Y_l^m(\mathbf{u})$ is the l order m degree Spherical Harmonic (SH) basis and $R_k(q)$ is the Gaussian-Laguerre polynomial basis. It was proposed to sparsely represent

$E(\mathbf{q};\mathbf{v})$ (Assemlal et al. 2009). The SPFI is to fit a model given by

$$E(\mathbf{q}_i;\mathbf{v}) = \sum_{k=0}^{K}\sum_{l=0}^{L}\sum_{m=-l}^{l} a_{k,l,m}(\mathbf{v})B_{k,l,m}(\mathbf{q}_i) + \epsilon_i(\mathbf{v}), \tag{4.5}$$

where K and L are large integers.

$$B_{k,l,m}(\mathbf{q}_i) = G_k(||\mathbf{q}_i||)Y_l^m(\mathbf{u}) \tag{4.6}$$

$$R_k(||\mathbf{q}_i||) = \kappa_k(\zeta)\exp(-\frac{||\mathbf{q}_i||^2}{2\zeta})L_k^{1/2}(\frac{||\mathbf{q}_i||^2}{\zeta}) \tag{4.7}$$

$$\kappa_k(\zeta) = \left[\frac{2}{\zeta^{3/2}}\frac{k!}{\Gamma(k+3/2)}\right]^{1/2} \tag{4.8}$$

In this case, if we set $\mathbf{x}_i = (B_{0,0,0}(\mathbf{q}_i),\cdots,B_{K,L,L}(\mathbf{q}_i))^T$ and $\beta(\mathbf{v}) = (a_{0,0,0}(\mathbf{v}),\cdots,a_{K,L,L}(\mathbf{v}))^T$, then SPFI can be regarded as a special case of model (4.4). Moreover, it can be shown that $p(R_0\mathbf{r};\mathbf{v})$ can be written as

$$p(R_0\mathbf{r};\mathbf{v}) = \sum_{l=0}^{L}\sum_{m=-l}^{l}\left\{4(-1)^{l/2}\frac{\zeta^{0.5l+1.5}\pi^{l+1.5}R_0^l}{\Gamma(l+1.5)}\sum_{k=0}^{K} f_{k,l,m}(\zeta,R_0)a_{k,l,m}\right\}Y_l^m(\mathbf{u}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_{l,m}Y_l^m(\mathbf{u}) \tag{4.9}$$

$$f_{k,l,m}(\zeta,R_0) = \kappa_k(\zeta)\sum_{i=0}^{k}(-1)^i\binom{k+0.5}{k-i}\frac{1}{i!}2^{0.5l+i-0.5}\Gamma(0.5l+i+1.5)1F1(\frac{2i+l+3}{2};l+\frac{3}{2};-2\pi^2R_0^2\zeta) \tag{4.10}$$

$$1F1(a;b;x) = \sum_{k=0}^{\infty}\frac{(a)_k x^k}{(b)_k k!}, \qquad (a)_k = (a(a+1)...(a+l-1)), \text{with}(a)_0 = 1 \tag{4.11}$$

The Implementation includes two steps. The first step is to estimate coefficients $\{a_{k,l,m}\}$ of the signal. The second step is the linear analytical transform $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ of EAP profile $p(\mathbf{R_0})$, and the second step is independent of the first step.

The linear transform from $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ could be implemented as an matrix multiplication. This transformation is independent with the data, since $\{f_{k,l,m}(\zeta,R_0)\}$ only depends on ζ and R_0. Once a R_0 and the basis are given, trasformation matrix can be calculated. And since in SPFI, Only the value of $1F1$ at the fixed value $-2\pi^2R_0^2\zeta$ is needed, so the transformation matrix only needs to be calculated once.

The basis matrix is calculated using $q = b^{1/2}$ and $\zeta = 700$ where b values are from image

acquisition. The reasoning for this choice is that considering $E(\mathbf{q}) = \exp(-4\pi^2\tau q^2 D)$, $b = 4\pi^2\tau q^2$, and a typical diffusion coefficient of $D = 0.7 \times 10^{-3} mm^2/s$, a typical b-value $b = 3000 s/mm^2$, we set $\zeta = \frac{1}{8\pi^2\tau \times 0.7 \times 10^{-3}}$. If $4\pi^2\tau = 1$, then ζ is about 700. The transformation matrxi from $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ will be calcuated by setting $R_0 = 15\mu m$.

4.2.2 Estimation Procedures

(Cheng et al. 2010) used lease square (LS) estimation with regularization terms for spherical and radial parts. For LS estimation, denote signal vector by $Y = [E(\mathbf{q}_i)]_{n\times 1}$, the basis matrix by $X = [B_{k,l,m}(\mathbf{q}_i)]_{n\times(L+1)(L+2)(K+1)/2}$, and the spherical and radial regularization diagonal matrices respectively by $\mathbf{L} = [l(l+1)]$ and $\mathbf{K} = [k(k+1)]$. Then the coefficient vector $\beta = [a_{k,l,m}] = (X^T X + \lambda_l \mathbf{L}^T\mathbf{L} + \lambda_k \mathbf{K}^T\mathbf{K})^{-1} X^T Y$, where λ_l and λ_k are the regularization terms for spherical and radial parts. The implementation is very fast, but the LS estimation is highly sensitive to (not robust against) outliers. Figure 4.1 shows that with the presence of outliers, LS estiamtion with regularization may not recover the underlying structure correctly.

In this paper, we are considering a robust estimation by minizing Huber's loss function $\sum_i \rho(\epsilon_i(\mathbf{v}))$, or equivalently solving $\sum_i \rho'(\epsilon_i(\mathbf{v}))\mathbf{x}_i = 0$, with

$$\rho(\epsilon_i(\mathbf{v})) = \begin{cases} \epsilon_i(\mathbf{v})^2 & \text{if } |\epsilon_i(\mathbf{v})| \le c \\ c(2|\epsilon_i(\mathbf{v})| - c) & \text{if } |\epsilon_i(\mathbf{v})| > c \end{cases} \tag{4.12}$$

Where c is a parameter that controls the robustness level, and a smaller value of c usually leads to more robust estimation. Huber argued that $c = 1.345$ is a good choice, and showed that asymptotically, it is 95% as efficient as least square if the true distribution is normal and much more efficient in many other cases.(Huber and Ronchetti 1975)

There are closed form solutions and fast algorithms for solving the least squares problem as well as the weighted least squares problem: Huber's estimators can be obtained by

minimizing a loss function , or equivalently solving

$$\sum_i \omega_i \epsilon_i(\mathbf{v}) \mathbf{x}_i = 0 \tag{4.13}$$

Thus, a convenient way to solve for 4.13 is to use an iteratively reweighted least squares (IRLS) algorithm, in which we calculate $\omega_i = \rho'(\epsilon_i(\mathbf{v}))/\epsilon_i(\mathbf{v})$, solve the weighted least squares problem, re-calculate the weights, re-solve, and so on until convergence.

The preceding derivations are slightly oversimplied, in that the arguments for setting $c = 1.345$ are based on the assumption that the response variable has known variance 1. In reality, of course, this is not true, and we must apply the loss functions to the scaled residuals, i.e. replace every $\sum_i \rho(\epsilon_i(\mathbf{v}))$ with $\sum_i \rho(\epsilon_i(\mathbf{v})/s)$, and $\sum_i \rho'(\epsilon_i(\mathbf{v}))$ with $\sum_i \rho'(\epsilon_i(\mathbf{v})/s)$, where s is an estimated scale parameter. While a number of other estimators have been proposed, the simplest is based on the median absolute deviation of the residuals:

$$\text{MAD} = \text{median}\{|\epsilon_i(\mathbf{v})|\} \tag{4.14}$$

where $\hat{s} = \text{MAD}/0.6745$, based on the idea that, for the standard normal, $E(\text{MAD}) = 0.6745$.

A key feature in HARDI is its spatial constraint. Specifically, the orientation and anisotropy of any single fiber bundles change smoothly from one voxel to the next, particularly along the dominant fiber orientation, whereas it may change dramatically at the boundaries between tracts and interfaces with gray matter structures and cerebrospinal fluid (CSF) spaces. Moreover, the EAP profile or ODF is expected to change smoothly from one voxel to the next in the same fiber crossing region, whereas it may change dramatically at the boundaries of fiber crossing regions and surrounding fiber bundles. This is a very important and powerful constraint that can be exploited to improve the reconstruction in HARDI. However, the methods mentioned in previous sections are voxel-wise methods and do not make use of the spatial constraint of HARDI.

To explicitly exploit such spatial constraint, we develop a robust multi-scale adaptive

and sequential smoothing (MASS) to spatially and adaptively update $\{\beta(\mathbf{v}) : \mathbf{v} \in \mathcal{V}\}$. The key idea of MASS is to combine HARDI signals in a neighboring sphere of voxel $\mathbf{v}$ to make inference on $\beta(\mathbf{v})$ at the voxel $\mathbf{v}$. Specifically, let $B(\mathbf{v}, h)$ be a sphere with radius h centered at voxel $\mathbf{v}$ and $\omega(\mathbf{v}, \mathbf{v}'; h)$ be a weight function of triple $(\mathbf{v}, \mathbf{v}', h)$ such that

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

MASS is based on a set of weighted quadratic function, denoted by $l_n(\beta_j(\mathbf{v}); \omega, h)$ for j-th component of $\beta(\mathbf{v})$ as follows:

$$l_n(\beta_j(\mathbf{v}); \omega_j, h) = \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega_j(\mathbf{v}, \mathbf{v}'; h)(\hat{\beta}_j(\mathbf{v}') - \beta_j(\mathbf{v}))^2 \tag{4.15}$$

$$\hat{\beta}_j(\mathbf{v}) = \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega_j(\mathbf{v}, \mathbf{v}'; h)\hat{\beta}_j(\mathbf{v}') \tag{4.16}$$

It is critical to choose a good $\omega(\mathbf{v}, \mathbf{v}'; h)$ in preventing oversmoothing the estimates of $\beta(\mathbf{v})$ across voxels, while preserving the edges between different structures, such as fiber bundles, crossing fibers, or gray matter regions. A good $\omega(\mathbf{v}, \mathbf{v}'; h)$ should quantify the similarity between $\beta(\mathbf{v})$ and $\beta(\mathbf{v}')$ or their corresponding EAPs. Specifically, if $\beta(\mathbf{v})$ and $\beta(\mathbf{v}')$ substantially differ from each other, then the HARDI signals in voxel $\mathbf{v}'$ do not contain too much information on $\beta(\mathbf{v})$ and thus $\omega(\mathbf{v}, \mathbf{v}'; h)$ should be close to 0. However, if $\beta(\mathbf{v})$ and $\beta(\mathbf{v}')$ are close to each other indicating that the HARDI signals in voxel $\mathbf{v}'$ contain useful information on $\beta(\mathbf{v})$, then $\omega(\mathbf{v}, \mathbf{v}'; h)$ should be significantly bigger than zero. See the explicit expression of $\omega(\mathbf{v}, \mathbf{v}'; h)$ in Section 4.2.3.

4.2.3 Multi-scale Adaptive and Sequential Smoothing (MASS)

We develop the MASS procedure to adaptively determine $\mathbf{w}$ and estimate $\beta(\mathbf{v})$ across all voxels $\mathbf{v} \in \mathcal{V}$. Our multiscal adaptive strategy starts with building a sequence of nested spheres with increasing radiues $h_0 = 0 < h_1 < \cdots < h_S = r$ ranging from the smallest scale $h_0 = 0$ to a large scale $h_S = r$ at each voxel $\mathbf{v}$. At the scale $h_0 = 0$, we just calculate

$\hat{\beta}(\mathbf{v}; h_0) = \widehat{\beta}(\mathbf{v})$ voxel-wisely without using any spatial information. It corresponds to setting $w(\mathbf{v}, \mathbf{v}'; h_0) = 1$ if $\mathbf{v} = \mathbf{v}'$ and 0 otherwise. Then, based on the signals contained in voxels d and d', we use methods as detailed below to calculate weights $w(\mathbf{v}, \mathbf{v}'; h_1)$ at scale h_1 for all voxels $\mathbf{v}$. After getting the new weights $w(\mathbf{v}, \mathbf{v}'; h_1)$, we can update $\hat{\beta}(\mathbf{v}; h_1)$. Then we can sequentially determine $w(\mathbf{v}, \mathbf{v}'; h_s)$ and adaptively update $\hat{\beta}(\mathbf{v}; h_s)$. From $h_0 = 0$ to $h_S = r$, a path diagram of the multiscale adaptive strategy is given below:

$$\begin{array}{ccccccc} w(\mathbf{v}, \mathbf{v}'; h_0) & & w(\mathbf{v}, \mathbf{v}'; h_1) & & \cdots & w(\mathbf{v}, \mathbf{v}'; h_S = r) \\ \Downarrow & \nearrow & \Downarrow & \nearrow & \cdots & \Downarrow \\ \hat{\beta}(\mathbf{v}; h_0) & & \hat{\beta}(\mathbf{v}; h_1) & & \cdots & \hat{\beta}(\mathbf{v}; h_S) \end{array}$$

MASS consists of three key steps: (I) an initialization step, (II) a weighted estimation step, and (III) a stop checking step. In the initialization step, we prefix a geometric series $\{h_s = c_h^s : s = 1, ..., S\}$ of radii with $h_0 = 0$, where $c_h \in (1, 2)$, say $c_h = 1.15$ and $S = 10$. We use small c_h in order to prevent incorporating too many neighboring voxels at the beginning, and thus it improves the robustness of the procedure and the accuracy of parameter estimation. At $h_0 = 0$, we obtain $\hat{\beta}(\mathbf{v}; h_0) = \widehat{\beta}(\mathbf{v})$ across all voxels by minimizing 4.13. We then set $s = 1$ and $h_1 = c_h$.

In the weighted estimation step, we first compute $D_{\beta_j}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ to characterize the similarity between $\hat{\beta}_j(\mathbf{v}; h_{s-1})$ and $\hat{\beta}_j(\mathbf{v}'; h_{s-1})$ at voxels $\mathbf{v}$ and $\mathbf{v}'$ and the adaptive weights $\omega_j(\mathbf{v}, \mathbf{v}'; h_s)$, which are defined as

$$D_{\beta_j}(\mathbf{v}, \mathbf{v}'; h_{s-1}) = (\hat{\beta}_j(\mathbf{v}'; h_{s-1}) - \hat{\beta}_j(\mathbf{v}; h_{s-1}))^2 / \Sigma(\hat{\beta}_j(\mathbf{v}; h_{s-1})) \tag{4.17}$$

$$\Sigma(\hat{\beta}_j(\mathbf{v}; h_{s-1})) = \sum_{\mathbf{v}' \in B(\mathbf{v}, h_{s-1})} \omega_j(\mathbf{v}, \mathbf{v}'; h_{s-1})^2 \Sigma(\hat{\beta}_j(\mathbf{v}'; h_0)) \tag{4.18}$$

$$\omega_j(\mathbf{v}, \mathbf{v}'; h_s) = \frac{K_{loc}(||\mathbf{v} - \mathbf{v}'||_2 / h_s) K_{st}(D_{\beta_j}(\mathbf{v}, \mathbf{v}'; h_{s-1}) / C_n)}{\sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} K_{loc}(||\mathbf{v} - \mathbf{v}'||_2 / h_s) K_{st}(D_{\beta_j}(\mathbf{v}, \mathbf{v}'; h_{s-1}) / C_n)}, \tag{4.19}$$

where $K_{loc}(u)$ and $K_{st}(u)$ are two nonnegative kernel functions with compact support, C_n is a number associated with n, and $||\cdot||_2$ denotes the Euclidean norm of a vector (or a matrix).

We then calculate

$$\hat{\beta}_j(\mathbf{v}; h_s) = \text{argmin}_{\beta_j(\mathbf{v})} l_n(\beta_j(\mathbf{v}); \omega_j, h_{s-1}) = \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega_j(\mathbf{v}, \mathbf{v}'; h_{s-1}) \hat{\beta}_j(\mathbf{v}'; h_{s-1}) \quad (4.20)$$

The weights $K_{loc}(||\mathbf{v} - \mathbf{v}'||_2/h_s)$ give less weight to the voxel $\mathbf{v}' \in B(\mathbf{v}, h_s)$, whose location is far from the voxel $\mathbf{v}$. The K_{loc} is a regular kernel function for smoothing the smoothed curves or surfaces. Some common choices of $K_{loc}(\cdot)$ include the Gaussian kernel and Epanechnikov kernel (Tabelow et al. 2006; 2008, Polzehl and Spokoiny 2000). We use $K_{loc} = (1 - u^2)_+$ throughout this paper. The weights $K_{st}(\cdot)$ downweight voxel $\mathbf{v}'$ with large $D_{\beta_j}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ which indicates a large difference between $\hat{\beta}_j(\mathbf{v}; h_{s-1})$ and $\hat{\beta}_j(\mathbf{v}'; h_{s-1})$. Although differnet choice of $K_{st}(\cdot)$ have been suggested in the propagation-separation method (Polzehl and Spokoiny 2000; 2006, Polzehl et al. 2010), we have tested these kernel functions and found that $K_{st} = \exp(-u)$ performs reasonably well.

The scale C_n is used to penalize the similarity between any two voxels $\mathbf{v}$ and $\mathbf{v}'$ in a similar manner to bandwidth, and an appropriate choice of C_n is crucial for the behavior of the propagation-separation method. As discussed in (Polzehl and Spokoiny 2000; 2006), a propagation condition independent of the observations at hand can be used to specify C_n, The basic idea of the propagation condition is that the impact of the statistical penalty in $K_{st}(D_{\beta_j}(\mathbf{v}, \mathbf{v}'; h_{s-1})/C_n)$ should be negligible under a homogeneous model $\beta_j(\mathbf{v}) \equiv$ constant yielding almost free smoothing within homogeneous regions. However, we take an alternative approach to choose C_n here. Specifically, a good choice of C_n should balance between the sensitivity and specificity of MASS. We choose $C_n = n^{0.4}\chi_1^2(0.8)$, where $\chi_1^2(a)$ is the upper a-percentile of the χ_1^2 distribution.

In the stop checking step, after the first iteration, we start to calculate a stopping criterion based on a normalized distance between $\hat{\beta}_j(\mathbf{v}; h_s)$ and $\hat{\beta}_j(\mathbf{v}; h_0)$ given by

$$D_{\beta_j}(\mathbf{v}; h_s, h_0) = (\hat{\beta}_j(\mathbf{v}; h_s) - \hat{\beta}_j(\mathbf{v}; h_0))^2 / \Sigma(\hat{\beta}_j(\mathbf{v}; h_0)) \quad (4.21)$$

Then, if $D_{\beta_j}(\mathbf{v}; h_s, h_0)$ is greater than C_s, where $C_s = \chi^2_1(0.8/s)$ in our implementation, then we set $\hat{\beta}_j(\mathbf{v}; h_S) = \hat{\beta}_j(\mathbf{v}; h_{s-1})$ and $s = S$ for the j-th component and voxel $\mathbf{v}$. If $s = S$ for all voxels, we stop. If $D_{\beta_j}(\mathbf{v}; h_s, h_0) \leq C_s$, then we set $h_{s+1} = c_h h_s$, increase s by 1, and continue with the weighted estimation step. It should be noted that different components of $\hat{\beta}(\mathbf{v}; h_S)$ may stop at different bandwidths, indicating that different degrees of smoothness are used to reconstruct HARDI.

We usually set the maximal stop S to be relatively small, say between 10 and 20, and thus each $B(\mathbf{v}; h_S)$ only contains a relatively small number of voxels. As S increases,the number of neighboring voxels in $B(\mathbf{v}; h_S)$ increases exponentially. It increase the chances of oversmoothing $\beta_j(\mathbf{v})$ when $\mathbf{v}$ is near the edge of distinct regions. Moreover, in order to prevent oversmoothing $\beta_j(\mathbf{v})$, we gradually decrease C_s with the number of iteration.

Finally, we summarize the MASS algorithm 2 for the adaptive estimation of ODF at voxel $\mathbf{v}$ below.

Algorithm 2: MASS at voxel $\mathbf{v}$

Input: Signals $\mathbf{Y}(\mathbf{v})$ and design matrix $\mathbf{X}$

Output: Estimated EAP profile at Radius $R_0 = 15\mu m$, $\widehat{p}(\mathbf{v}; h_S)|_{R_0}$

```
Estimate β̂(v) from (4.13).
for j-th component in β̂(v) do
    for s ← 1 to S do
        calculate the weights w_j(v, v'; h_s) for d' ∈ B(d, h_s) by (4.19);
        calculate β̂_j(v; h_s) by minizing (4.15).
        if D_{β_j}(v; h_s, h_0) > C_s, then
            β̂_j(v; h_S) = β̂_j(v; h_{s-1}), and s = S,
        else
            h_{s+1} = c_h h_s;
return p̂(v; h_S)|_{R_0}.
```

4.2.4 Maxima Extraction

Based on $\widehat{p}(\mathbf{v}; h_S)|_{R_0}$ at voxel $\mathbf{v}$, we need to extract its maxima in order to infer fiber directions. Although there are other more complicated methods for extracting maximal directions, such as the method presented in (Hlawitschka and Scheuermann 2005), spherical

Newton's method (Tuch 2004), and Powell's method (Jansons and Alexander 2003), we take a simple thresholding approach in this paper. Specifically, we project the estimated EAP onto the sphere tessellated with a triangle mesh, which has 2562 points on the unit sphere. If the estimated EAP value at a mesh point is greater than the corresponding value at all its neighboring mesh points and this estimated EAP value is greater than max(EAP)/2, then the direction at this mesh point is regarded as a maximum. This thresholding method avoids selecting small peaks that may appear due to noise.

4.3 Simulation Study

4.3.1 Data Generating

We examined the finite sample performance of our MASS on decting crossing fibers by using synthetic HARDI data generated from the multi-tensor model (Alexander et al. 2002, Tuch 2004). We simulated the diffusion-weighted signals according to

$$E(\mathbf{q}_i;\mathbf{v}) = \sqrt{(\sum_{k=t}^{T} p_t e^{-b_i \mathbf{u}_i^T \mathbf{D}_t(\mathbf{v})\mathbf{u}_i} + \sigma\epsilon_{i1})^2 + (\sigma\epsilon_{i2})^2} \text{ for } i = 1, \cdots, n \tag{4.22}$$

for $i = 1, \cdots, n$, where $\mathbf{q}_i = q_i\mathbf{u}_i$ with $\mathbf{u}_i$ being a unit vector, T is the number of fibers, p_t is the weight for t-th fiber, b is the b-value and $\mathbf{D}_t(\mathbf{v})$ is the tensor matrix for the t-th fiber, SNR $= 1/\sigma$, and ϵ_{i1} and ϵ_{i2} are independently simulated from the standard normal distribution. We used the multi-tensor model 4.22 to generate different phantoms with different regions of interest (ROIs) with 81 sampling directions on the hemisphere for the 3rd order tessellation of the icosahedron and $b = 500, 1000, 2000, 3000\text{s/mm}^2$. Specifically, voxels with a single fiber were generated from a single tensor model using diffusion tensor profiles with eigenvalues $[1.7, 0.3, 0.3] \times 10^{-3}\text{mm}^2/\text{s}$, voxels with two fiber directions were generated by two-tensor model $E(\mathbf{q}_i;\mathbf{v}) = e^{-b_i\mathbf{u}_i^T\mathbf{D}_1(\mathbf{v})\mathbf{u}_i}/2 + e^{-b_i\mathbf{u}_i^T\mathbf{D}_2(\mathbf{v})\mathbf{u}_i}/2$, and voxels isotropic tensors were generated by the single tensor model using diffusion tensor profiles with eigenvalues $[1, 1, 1] \times 10^{-3}\text{mm}^2/\text{s}$. We estimated EAP at each voxel by using the two

voxel-wise estimation methods including least square estiamtion (LS), Robust regression (RR), and their corresponding MASS including LS-MASS, RR-MASS. For our smoothing procedure, we set $K_{st} = \exp(-u)$, $S = 10$. For SPFI, we use all data in 4 shells and chose $SNR = 15$, and $K = 2$,$L = 4$, $\lambda_l = 1e-7$, $\lambda_k = 5e-8$, $\zeta = 700$ for all experiments. Then we extracted the EAP maxima aligned with fiber directions.

4.3.2 Angle Detection in First Phantom with 90^o crossing fibers

In the first phantom, we included four different ROIs including isotropic ROIs, two single fiber ROIs with its direction going either along the $x-$axis (ROI1) or along the $y-$axis (ROI2), and the 90^o crossing fiber ROIs(ROI3); To add outliers to the data, we randomly select 16 out of 81 directions, the data for these 16 directions are generated using different underground structure, which is 45^o degree rotation from that used for other directions.

Figure 4.2 presents the estimated EAP images for this type of phantom from different settings. The left panel gives the EAP reconstruction from noise free and outlier free data; The middle panel gives the LS result from noise free but outlier added data; the right panel gives the RR result from noise free but outlier added data. We can see that LS estiamtion is sensitive to outliers and RR estimation gives better EAP estimation for outlier added data.

Figure 4.4 gives the comparison of EAP estiamtions with and without MASS on data with SNR=10, but NO outliers: The first column gives the EAP reconstruction results using LS and LS-MASS. We can see that LS-MASS gives smoother result than LS. First two panels in the second column gives the EAP reconstruction results using robust estimation without (RR) and with MASS (RR-MASS); the third panel gives the ground truth. We can see that RR-MASS is the closest to ground truth. Because MASS can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

Figure 4.5 gives the comparison of EAP estiamtions with and without MASS on outliers added data with SNR=15. The first column gives the EAP reconstruction results using LS and with LS-MASS. We can see that LS-MASS gives smoother result than LS, but it still can not deal with the outlier. First two panels in the second column gives the EAP

reconstruction results using RR and RR-MASS; the third panel gives the ground truth. We can see that RR-MASS is the closest to ground truth. Because MASS can not only reduce the noise by incorporation useful neighborhood information into the estimation procedure, but also use robust estimation which is more stable agains the outlier.

To quantify the accuracy of detection angle, we generated 1,000 data sets without outliers for three different SNRs including 10 ,15, and 20. We estimated the EAPs by using voxel-wise LS, RR, LS-MASS and RR-MASS . Then we extracted the EAP maxima aligned with fiber directions. For voxels with a single fiber, we calculated angle detection errors by comparing recovered fiber directions with the ground truth. For voxels with two crossing fibers, we calculated angle detection errors by comparing recovered crossing angles with the ground truth. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for voxels with one fiber and two fibers are presented in Table 4.1.

Similarly, we generated 1,000 data sets with outliers for three different SNRs including 10 ,15, and 20. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for voxels with one fiber and two fibers are presented in Table 4.2.

Table 4.1 and Table 4.2 reveal that the mean of the angular errors are substantially reduced and the percentage of detecting correct number of fibers are increased by adding our adaptive smoothing techique. It may indicate that the proposed adaptive smoothing techique can efficiently exploit spatial smoothness for reconstructing EAPs, while reducing noise leading to better angle detection. LS type estimations outperform RR esitmations in terms of the mean angle error on data without added outliers, while RR esitmations outperform LS estimations on data with added outliers. This is because LS type of estimations are sensitive to outliers and RR downweights the outliers in the data, leading to more stable recovery of EAPs. The percentage of detecting correct number of fibers are similar from LS

and RR estimations.

4.3.3 Angle Detection in Second Phantom with Twisted Crossing

In the second phantom, we included a twisted crossing region in the middle, one fiber along $x-$axis (ROI1), and the other fiber with changing angles with $x-$axis from 30^o, 45^o , 60^o, 75^o, to 90^o and then from 90^o, 75^o, 60^o, 45^o to 30^o. We marked all region with single fiber, which is not along $x-$axis as ROI2, and all regions with crossing fibers as ROI3. To add outliers to the data, we randomly select 13 out of 81 directions, the data for these 13 direcions are generated using different underground structure, which is 45^o degree rotation from that used for other directions.

Figure 4.3 presents the estimated EAP images for this type of phantom from different settings. The left panel gives the EAP reconstruction from noise free and outlier free data; The middle panel gives the LS result from noise free but outlier added data; the right panel gives the RR result from noise free but outlier added data. We can see that LS estiamtion is sensitive to outliers and RR estimation gives better EAP estimation for outlier added data.

Figure 4.6 gives the comparison of EAP estiamtions with and without MASS on data with SNR=10, but NO outliers: The first column gives the EAP reconstruction results using LS estimation without and with MASS. We can see that LS-MASS gives smoother result than LS. First two panels in the second column gives the EAP reconstruction results using robust estimation without and with MASS; the third panel gives the ground truth. We can see that RR-MASS result is the closest to ground truth. Because MASS can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

Figure 4.7 gives the comparison of EAP estiamtions with and without MASS on outliers added data with SNR=15: we rotated the underlying crossing 45^o counterclockwisely in randomly selected 13 out of 81 directions to mimic subject movement. The first column gives the EAP reconstruction results using LS and LS-MASS. We can see that LS-MASS gives smoother result than LS, but they still can not deal with the outlier. First two panels in the second column gives the EAP reconstruction results using RR and RR-MASS; the

third panel gives the ground truth. We can see that RR-MASS result is the closest to ground truth. Because MASS can not only reduce the noise by incorporation useful neighborhood information into the estimation procedure, but also use robust estimation which is more stable agains the outlier.

To quantify the accuracy of detection angle, we generated 1,000 data sets without outliers for three different SNRs including 10 ,15, and 20. We estimated the EAPs by using voxel-wise LS, RR, LS-MASS and RR-MASS . Then we extracted the EAP maxima aligned with fiber directions. For voxels with a single fiber, we calculated angle detection errors by comparing recovered fiber directions with the ground truth. For voxels with two crossing fibers, we calculated angle detection errors by comparing recovered crossing angles with the ground truth. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for voxels with one fiber and two fibers are presented in Table 4.3.

Similarly, we generated 1,000 data sets with outliers for three different SNRs including 10 ,15, and 20. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for voxels with one fiber and two fibers are presented in Table 4.4.

Table 4.3 and Table 4.4 reveal that the mean of the angular errors are substantially reduced and the percentage of detecting correct number of fibers are increased by adding our adaptive smoothing techique. It may indicate that the proposed adaptive smoothing techique can efficiently exploit spatial smoothness for reconstructing EAPs, while reducing noise leading to better angle detection. LS type estimations outperform RR esitmations in terms of the mean angle error on data without added outliers, while RR esitmations outperform LS estimations on data with added outliers. This is because LS type of estimations are sensitive to outliers and RR downweights the outliers in the data, leading to more stable recovery of EAPs.

4.4 Real Data Analysis

We tested our proposed method on public data from the NIH Human Connectome Project (HCP). 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. Diffusion weighting consisted of 3 shells of b=1000, 2000, and 3000 s/mm^2 interspersed with an approximately equal number of acquisitions on each shell within each run. Figure 4.8 gives EAP recovery result using LS estimation. One slice is presented and two ROIs are selected for furthur analysis.

In order to test our multi-scale adaptive and sequential smoothing (MASS) method, we added racian noise with $SNR = 15$ to the HCP data, then recovered the EAPs using LS, LS-MASS, RR and RR-MASS, where MASS is setup the same as simulation study, except here $\lambda_l = 5e-9$ and $\lambda_k = 1e-9$. (Figure 4.9) shows the EAP recovery of ROI1 from Figure 4.8. Panel (a) and (c) give the EAP results from noise added data using LS and LS-MASS; Panel (b) and (d) give the EAP results from noise added data using RR and RR-MASS; Panel (e) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from LS-MASS and RR-MASS are closer to (e) when compared to LS and RR respectively. In this data, LS-MASS and RR-MASS performly equally well. Similar results are shown in Figure 4.10, the EAP recovery of ROI2 from Figure 4.8.

4.5 Conclusion

We have introduced a robust multi-scale adaptive and sequential smoothing (MASS) framework to adaptively and sequentially reconstruct the EAPs across all voxels from HARDI signals. In simulation data analysis, we have shown that adding MASS to the regular LS estimation or robust estimation can substantially reduce the angle detection error and increase the accuracy of detecting the correct number of fibers in each voxels.

Because MASS reconstructs the EAPs at each voxel by adaptively borrowing the spatial information from the neighbouring voxels, then can substantially reduce the noise level, while improving the EAP reconstruction. This is also shown in the real data study. The other important contribution of this paper is that we also show with the presence of the outliers, like subject movement in the scanner, robust estimation works much better than the regular LS estimation, as it downweights the abnormal signals, leading to more stable reconstruction of the EAP.

4.6 Acknowledgment

Data used in Real data analysis were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Table 4.1: The mean of anglular errors under given EAP configuration and the percentage of correct number of detected EAP maximum in the first phantom. 1,000 simulated data sets were used and NO outliers were added.

All voxels with one fiber

SNR	LS	LS-MASS	RR	RR-MASS
10	2.09 (100 %)	0.48 (100 %)	3.02 (100 %)	1.12 (100 %)
15	0.78 (100 %)	0.03 (100 %)	1.55 (100 %)	0.20 (100 %)
20	0.22 (100 %)	0.01 (100 %)	0.70 (100 %)	0.03 (100 %)

All voxels with two fibers

SNR	LS	LS-MASS	RR	RR-MASS
10	5.89 (99.08 %)	2.77 (99.87 %)	6.31 (95.25 %)	2.99 (99.52 %)
15	3.02 (99.99 %)	1.29 (100 %)	2.82 (99.93 %)	1.44 (99.99 %)
20	2.17 (100 %)	0.57 (100 %)	2.06 (100 %)	0.83 (100 %)

Table 4.2: The mean of anglular errors under given EAP configuration and the percentage of correct number of detected EAP maximum in the first phantom. 1,000 simulated data sets were used and outliers were added.

All voxels with one fiber

SNR	LS	LS-MASS	RR	RR-MASS
10	7.21 (100 %)	6.40 (100 %)	6.53 (99.61 %)	5.20 (100 %)
15	6.78 (100 %)	5.83 (100 %)	5.07 (100 %)	4.26 (100 %)
20	6.59 (100 %)	5.30 (100 %)	4.47 (100 %)	3.94 (100 %)

All voxels with two fibers

SNR	LS	LS-MASS	RR	RR-MASS
10	22.77 (83.67 %)	13.60 (92.82 %)	16.99 (82.39 %)	11.82 (92.98 %)
15	12.62 (94.17 %)	5.92 (99.13 %)	10.25 (94.57 %)	5.34 (98.58 %)
20	7.23 (98.74 %)	3.89 (99.88 %)	5.83 (98.93 %)	3.16 (99.67 %)

Table 4.3: The mean of anglular errors under given EAP configuration and the percentage of correct number of detected EAP maximum in the second phantom. 1,000 simulated data sets were used and NO outliers were added.

All voxels with one fiber

SNR	LS	LS-MASS	RR	RR-MASS
10	2.42 (100 %)	1.36 (100 %)	3.22 (100 %)	1.99 (100 %)
15	1.43 (100 %)	0.84 (100 %)	2.03 (100 %)	1.10 (100 %)
20	1.00 (100 %)	0.74 (100 %)	1.39 (100 %)	0.85 (100 %)

All voxels with two fibers

SNR	LS	LS-MASS	RR	RR-MASS
10	6.26 (93.03 %)	4.74 (97.26 %)	5.72 (91.53 %)	4.20 (97.49 %)
15	3.98 (99.56 %)	2.62 (99.92 %)	3.73 (99.57 %)	3.06 (99.89 %)
20	2.74 (99.99 %)	1.56 (100 %)	3.18 (99.99 %)	2.85 (100 %)

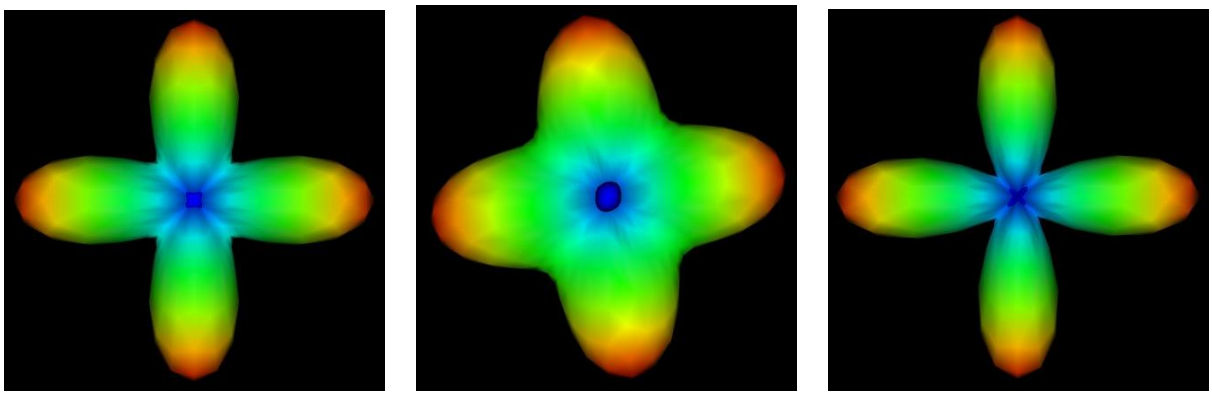
Table 4.4: The mean of anglular errors under given EAP configuration and the percentage of correct number of detected EAP maximum in the second phantom. 1,000 simulated data sets were used and outliers were added.

All voxels with one fiber

SNR	LS	LS-MASS	RR	RR-MASS
10	4.64 (100 %)	4.07 (100 %)	4.85 (99.61 %)	3.53 (100 %)
15	3.70 (100 %)	3.63 (100 %)	3.09 (100 %)	2.12 (100 %)
20	1.83 (100 %)	1.07 (100 %)	1.73 (100 %)	0.80 (100 %)

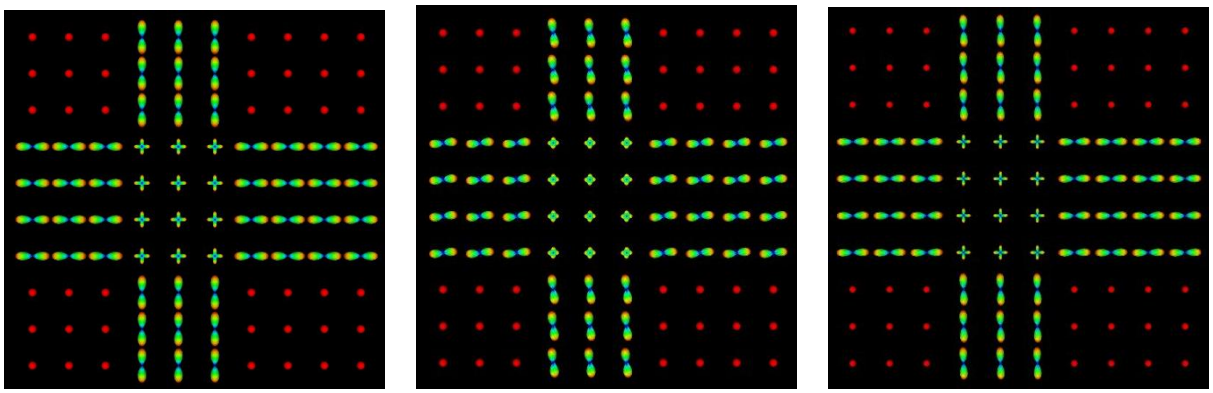
All voxels with two fibers

SNR	LS	LS-MASS	RR	RR-MASS
10	9.60 (71.37 %)	9.52 (75.05 %)	9.47 (78.04 %)	7.37 (82.68 %)
15	8.72 (81.85 %)	8.29 (87.84 %)	6.52 (89.42 %)	5.14 (93.69 %)
20	7.45 (92.21 %)	6.15 (97.57 %)	4.59 (97.89 %)	3.23 (99.36 %)



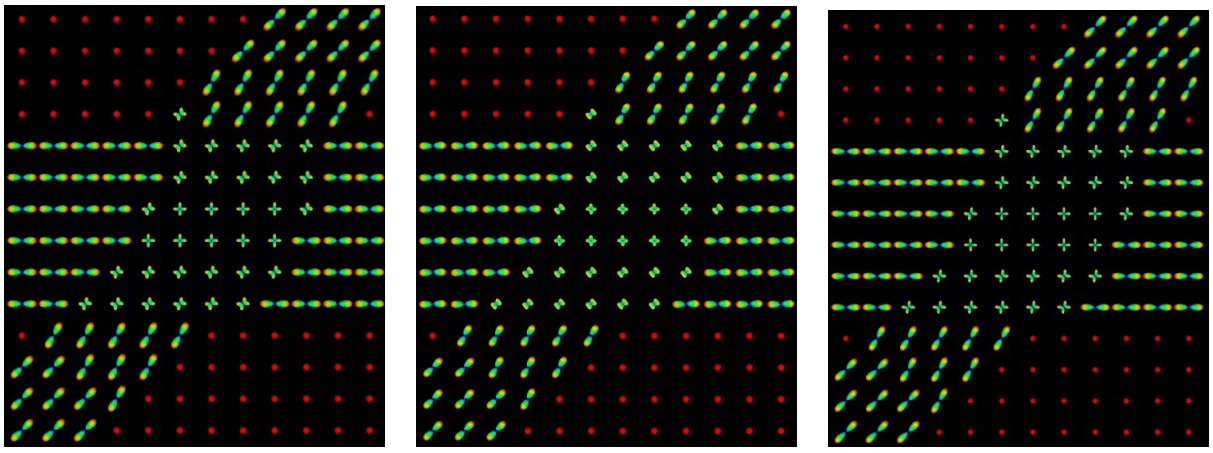
LS LS result with outliers RR result with outliers

Figure 4.1: Comparison of LS estiamtion with robust estimation: we rotated the underlying crossing 45^o counterclockwisely in randomly selected 16 out of 81 directions to mimic subject movement. The left panel gives the EAP reconstruction from noise free data; The middle panel gives the LS result from outlier added data; the right panel gives the Robust regression result from outlier added data. We can see that LS estiamtion is sensitive to outliers and robust regression gives better EAP estimation for outlier added data.



LS LS result with outliers RR result with outliers

Figure 4.2: Comparison of LS estiamtion with robust estimation on outliers added data: we rotated the underlying crossing 45^o counterclockwisely in randomly selected 16 out of 81 directions to mimic subject movement. The left panel gives the EAP reconstruction from noise free and outlier free data; The middle panel gives the LS result from noise free but outlier added data; the right panel gives the RR result from noise free but outlier added data. We can see that LS estiamtion is sensitive to outliers and RR estimation gives better EAP estimation for outlier added data.



LS LS result with outliers RR result with outliers

Figure 4.3: Comparison of LS estiamtion with robust estimation on outliers added data: we rotated the underlying crossing 45^o counterclockwisely in randomly selected 13 directions to mimic subject movement. The left panel gives the EAP reconstruction from noise free and outlier free data; The middle panel gives the LS result from noise free but outlier added data; the right panel gives the RR result from noise free but outlier added data. We can see that LS estiamtion is sensitive to outliers and RR estimation gives better EAP estimation for outlier added data.

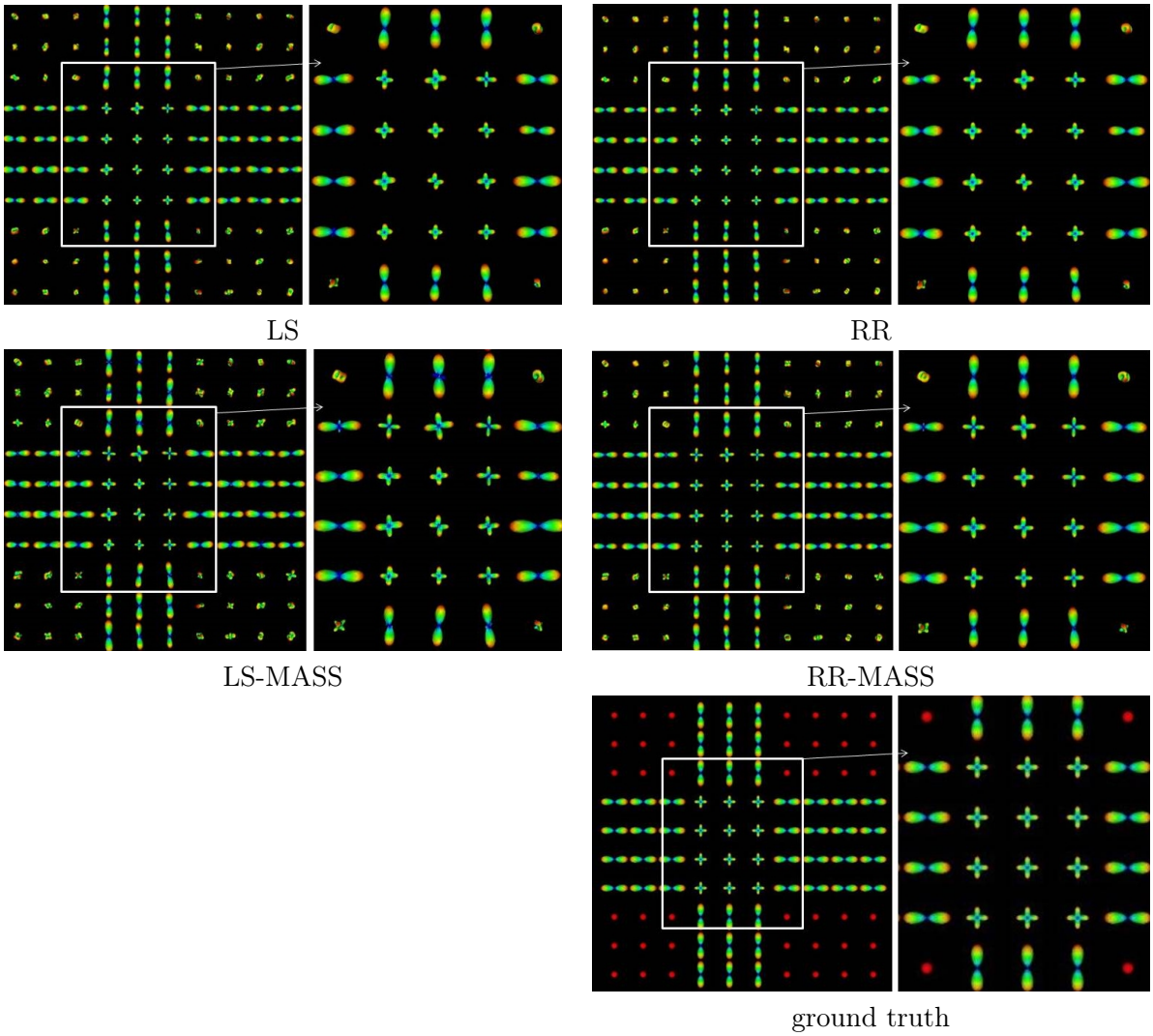


Figure 4.4: Comparison of EAP estiamtions with and without MASS on data with SNR=10, but NO outliers: The first column gives the EAP reconstruction results using LS estimation without and with MASS. We can see that LS with MASS gives smoother result than without. First two panels in the second column gives the EAP reconstruction results using robust estimation without and with MASS; the third panel gives the ground truth. We can see that RR resutl with MASS is the closest to ground truth. Because MASS can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

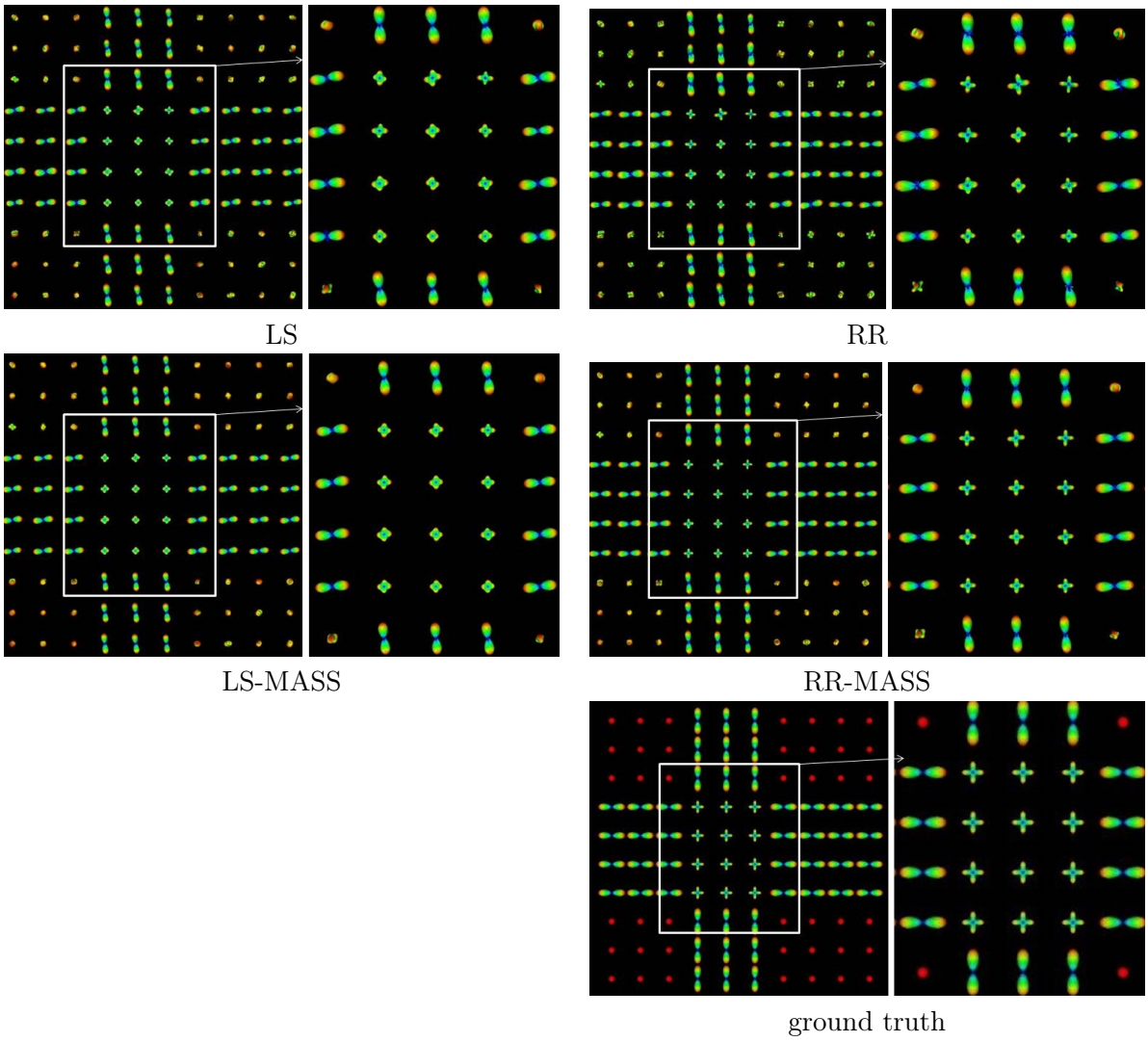


Figure 4.5: Comparison of EAP estiamtions with and without MASS on outliers added data with SNR=15: we rotated the underlying crossing 45^o counterclockwisely in randomly selected 16 out of 81 directions to mimic subject movement. The first column gives the EAP reconstruction results using LS estimation without and with MASS. We can see that LS with MASS gives smoother result than without MASS, but it still can not deal with the outlier. First two panels in the second column gives the EAP reconstruction results using robust estimation without and with MASS; the third panel gives the ground truth. We can see that RR resutl with MASS is the closest to ground truth. Because MASS can not only reduce the noise by incorporation useful neighborhood information into the estimation procedure, but also use robust estimation which is more stable agains the outlier.

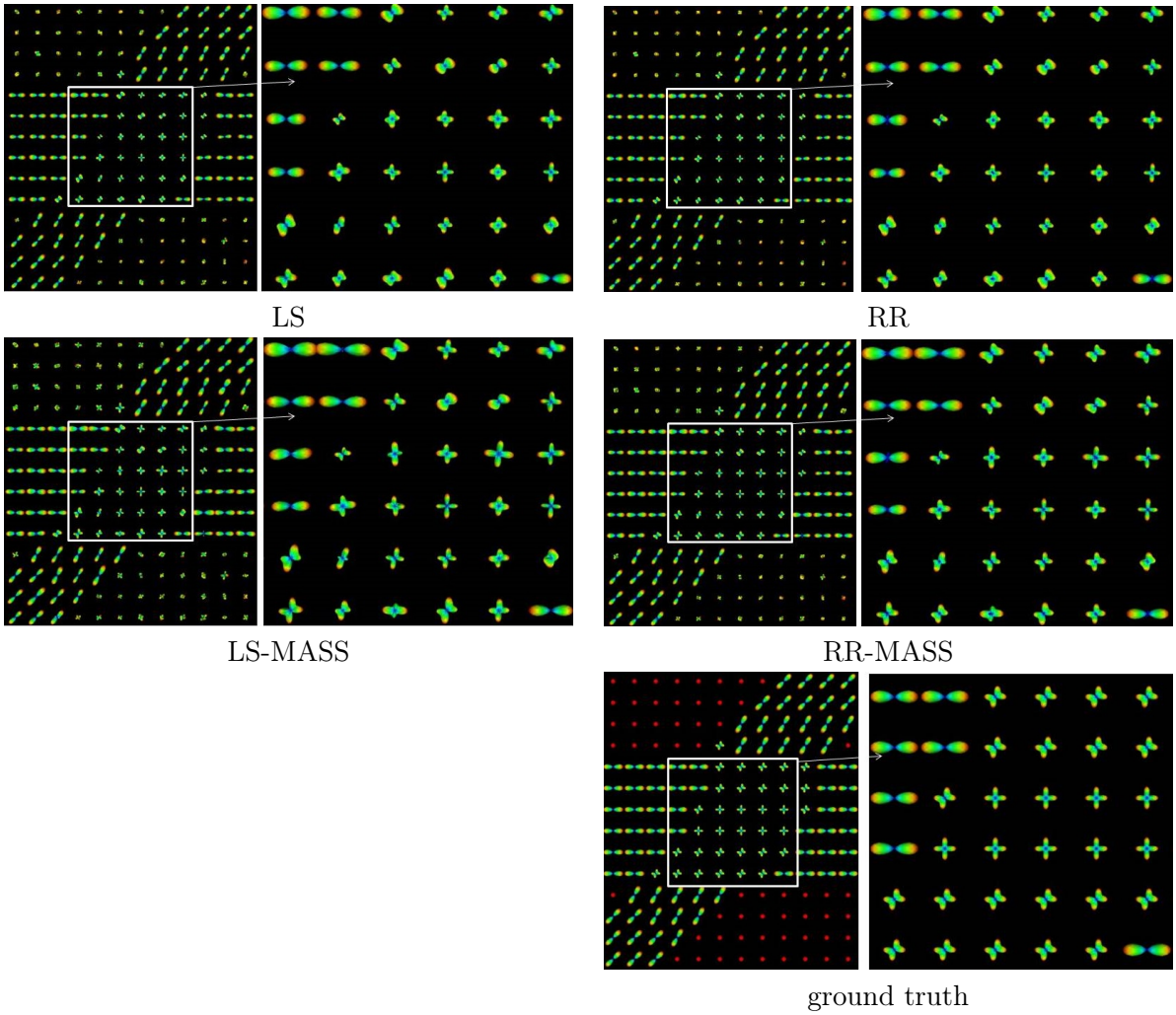


Figure 4.6: Comparison of EAP estiamtions with and without MASS on data with SNR=10, but NO outliers: The first column gives the EAP reconstruction results using LS estimation without and with MASS. We can see that LS with MASS gives smoother result than without. First two panels in the second column gives the EAP reconstruction results using robust estimation without and with MASS; the third panel gives the ground truth. We can see that RR resutl with MASS is the closest to ground truth. Because MASS can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

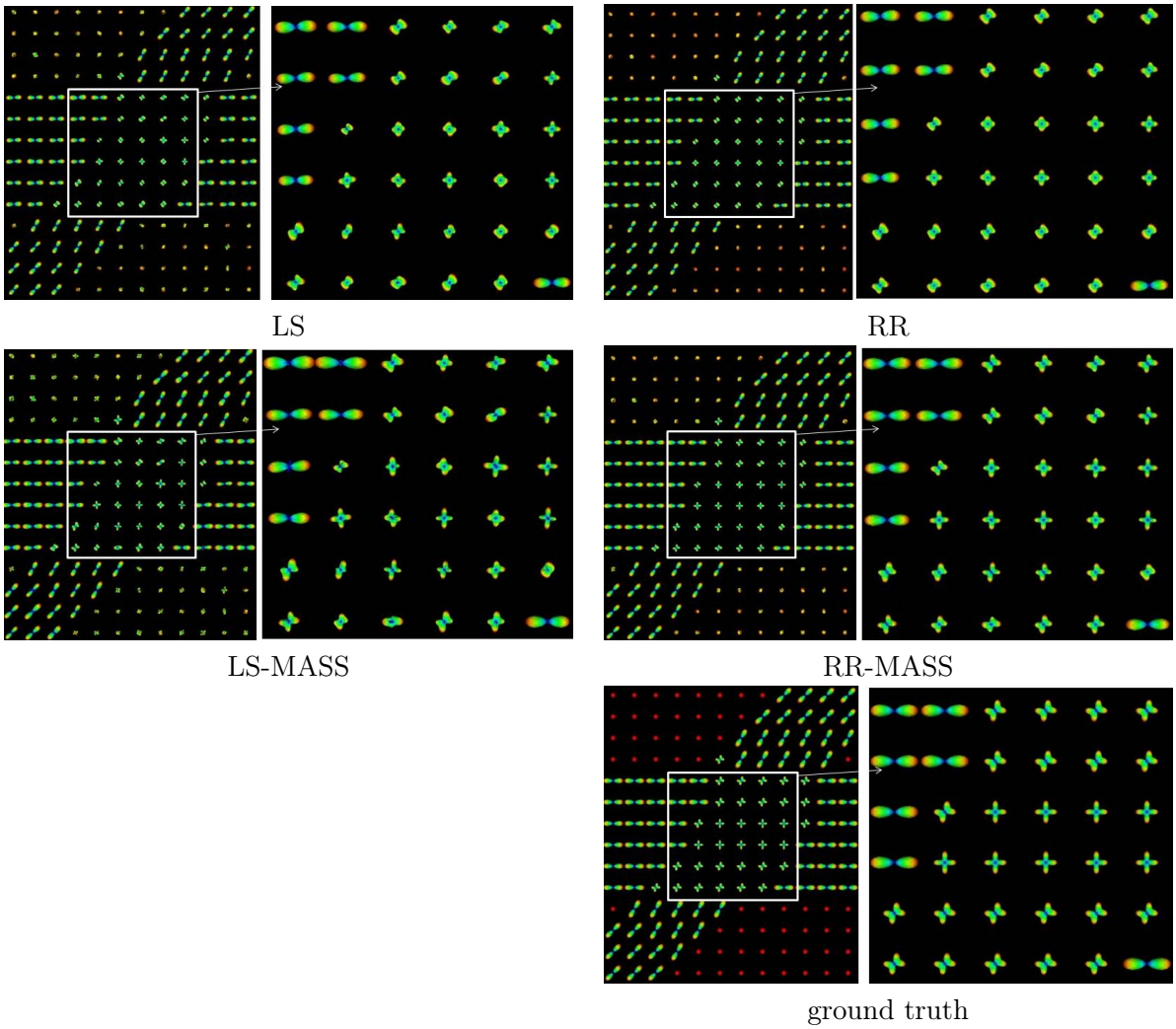


Figure 4.7: Comparison of EAP estiamtions with and without MASS on outliers added data with SNR=15: we rotated the underlying crossing 45^o counterclockwisely in randomly selected 13 out of 81 directions to mimic subject movement. The first column gives the EAP reconstruction results using LS estimation without and with MASS. We can see that LS with MASS gives smoother result than without MASS, but it still can not deal with the outlier. First two panels in the second column gives the EAP reconstruction results using robust estimation without and with MASS; the third panel gives the ground truth. We can see that RR resutl with MASS is the closest to ground truth. Because MASS can not only reduce the noise by incorporation useful neighborhood information into the estimation procedure, but also use robust estimation which is more stable agains the outlier.

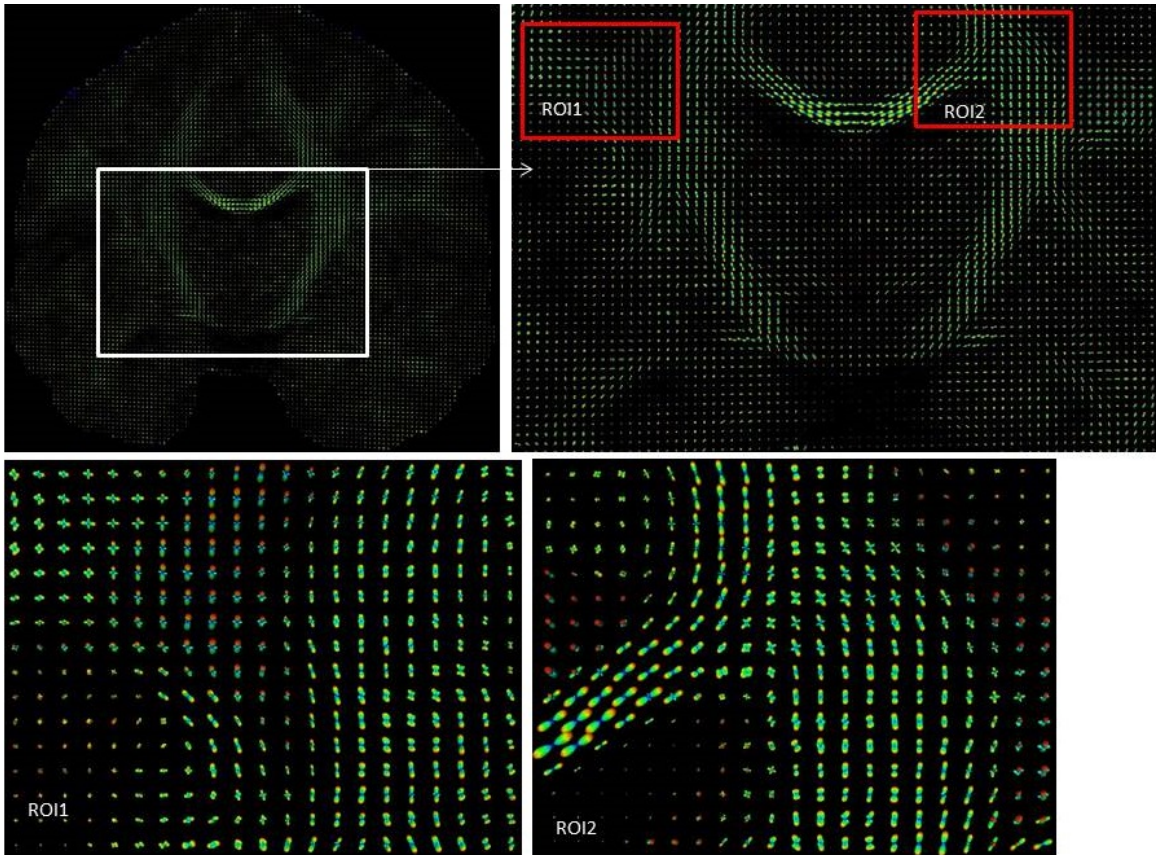
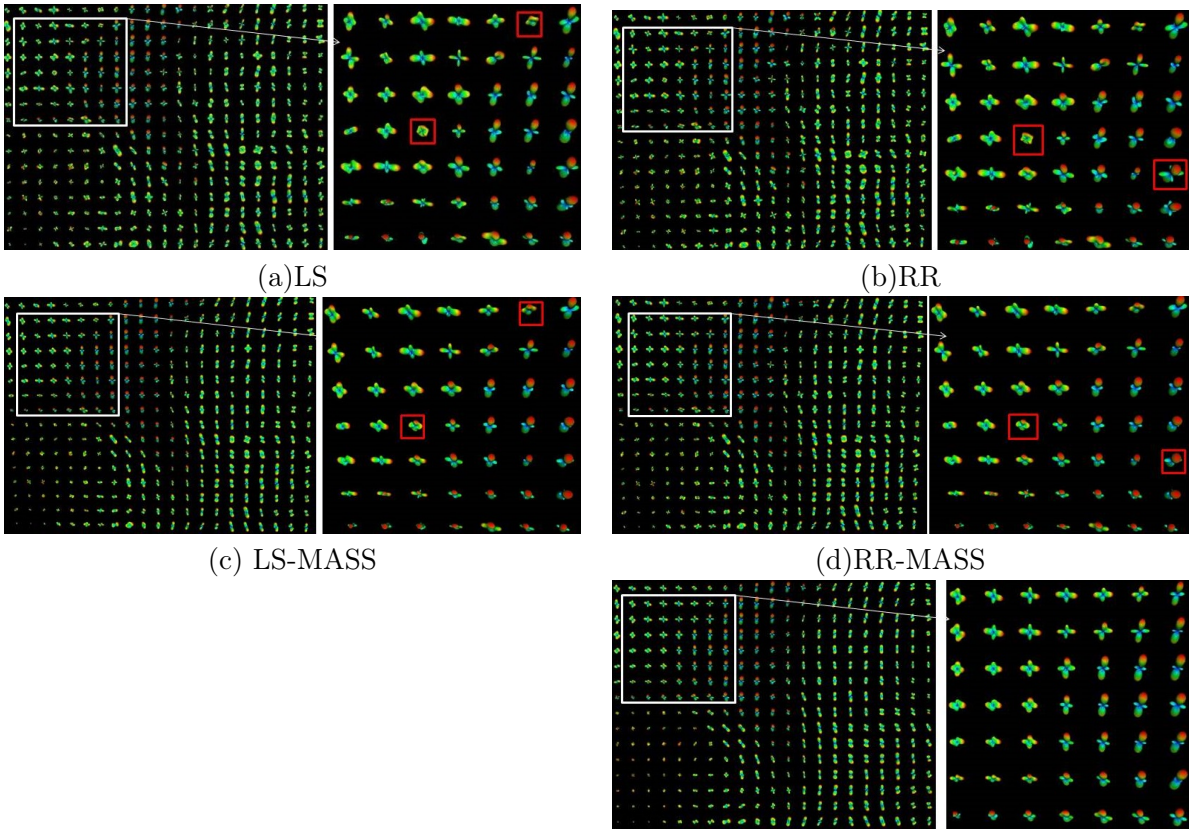


Figure 4.8: EAP recovery result of data from the NIH Human Connectome Project using LS estimation.

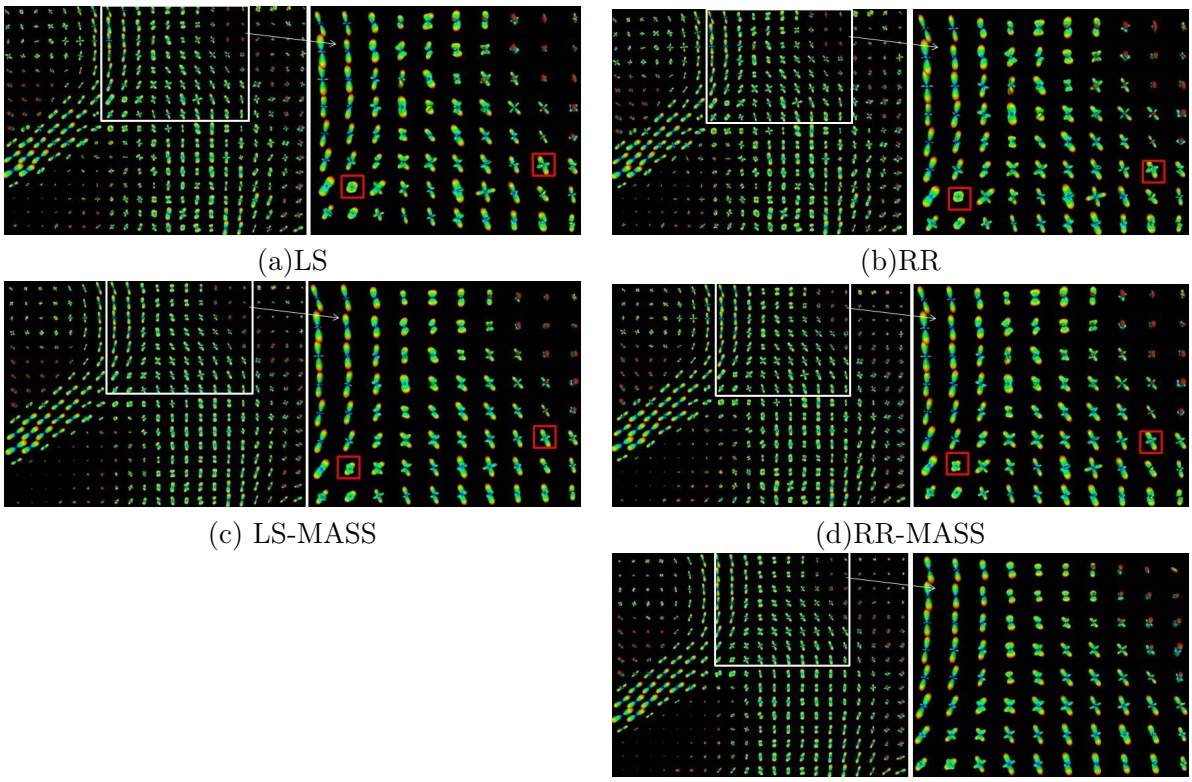


(a)LS (b)RR

(c) LS-MASS (d)RR-MASS

(e)LS from original HCP

Figure 4.9: EAP recovery of ROI1 from Figure 4.8. Panel (a) and (c) give the EAP results from noise added data using LS and LS-MASS; Panel (b) and (d) give the EAP results from noise added data using RR and RR-MASS; Panel (e) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from LS-MASS and RR-MASS are closer to (e) when compared to LS and RR respectively. In this data, LS-MASS and RR-MASS performly equally well.



(a)LS (b)RR

(c) LS-MASS (d)RR-MASS

(e)LS from original HCP

Figure 4.10: EAP recovery of ROI2 from Figure 4.8. Panel (a) and (c) give the EAP results from noise added data using LS and LS-MASS; Panel (b) and (d) give the EAP results from noise added data using RR and RR-MASS; Panel (e) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from LS-MASS and RR-MASS are closer to (e) when compared to LS and RR respectively. In this data, LS-MASS and RR-MASS performly equally well.

CHAPTER 5: SPARSE MULTI-SCALE ADAPTIVE MODEL (SMAM)

Diffusion magnetic resonance imaging (dMRI) is an important tool that allows non-invasive investigation of neural architecture of the brain. In Chapter 3 and Chapter 4, we have proposed multi-scale adaptive framework to spatially and adaptively infer ODFs and EAPs across all voxels, However the reconstruction results from signal with low SNR (≤ 10) are still not satisfactory. In this chapter, we develop a sparse multi-scale adaptive model (SMAM) to spatially and adaptively infer the EAP of water diffusion in regions with complex fiber configurations using voxel-adaptive dictionary. We show SMAM can provide great EAP reconstruction for signals with low SNR.

5.1 Introduction

Diffusion Magnetic Resonance Imaging (dMRI) is a widely used in-vivo imaging technique to explore the information of neural micro-structure by probing the diffusion of water molecules. So far it is still the unique non-invasive method to reveal the micro-geometry of nervous tissues noninvasively and to explore the neural connectome in living human subjects. The diffusion of water molecules is constrained by the surrounding structures including nerves, cells and surrounding tissue. For example, qualitatively water molecules diffuse fast along fibers and slowly cross fibers. Thus measuring the diffusion process quantitatively is crucial to understanding the neural micro-structure and fiber directions.

A central problem in dMRI is to estimate the Ensemble Average Propagator (EAP) $p(\mathbf{R};\mathbf{v})$, which describes fully the probability distribution of water molecule displacement $\mathbf{R}$ from a limited number of measurements of the signal attenuation $E(\mathbf{q};\mathbf{v})$. Under the narrow pulse assumption(Stejskal and Tanner 1965), the relationship between the $E(\mathbf{q};\mathbf{v})$

and $p(\mathbf{R};\mathbf{v})$, at each voxel $\mathbf{v}$ in a common space $\mathcal{V}$, where $\mathbf{q} = q\mathbf{u} \in R^3$ and $\mathbf{R} = R\mathbf{r} \in R^3$, is given by a Fourier transform (FT) relationship (Callaghan 1991) such that

$$p(\mathbf{R};\mathbf{v}) = \int_{\mathbf{q}\in\mathbf{R^3}} E(\mathbf{q};\mathbf{v}) e^{-2\pi i \mathbf{q}\cdot\mathbf{R}} d\mathbf{q} \tag{5.1}$$

Various methods already exist to reconstruct the EAP or estimate EAP features. The most common and famous model, is the diffusion tensor model (Basser et al. 1994b). Although very successful in many neuroscience studies, diffusion tensor imaging (DTI) is limited by the Gaussian assumption (free diffusion model) and cannot account for complex fiber configurations. A large family of techniques, mostly based on different mathematical representation of the signal, use multiple q-shell acquisitions in order to reconstruct signal features or EAP features, such as generalized high order tensors (Liu et al. 2004) based on cumulant expansions; or the diffusion orientation transform (DOT) (Özarslan et al. 2006); or a fourth order Cartesian tensor representation of the probability profile (Barmpoutis et al. 2008);or better diffusion ODFs (Aganj et al. 2009, Tristán-Vega et al. 2009b) than obtained from q-ball imaging (QBI) (Tuch et al. 2002). Unfortunately, for most of these methods, many DW measurements are still needed. Moreover, most of these methods do not recover the full EAP but features of it, making several assumptions that remain to be validated. Some attempts to sparsely represent the diffusion signal have already been performed. Spherical Polar Fourier Expression (SPFE) was proposed to sparsely represent to signal (Assemlal et al. 2008; 2009). Based on SPFE, Spherical Polar Fourier Imaging(SPFI), a novel technique for model-free analytical reconstruction of the EAP profile from the signals was proposed (Cheng et al. 2010). This continuous representation is based on the Spherical Polar Fourier (SPF) basis and provides closed-form expressions for EAP and ODF computation.

Recovering a latent function from a small number of samples in Fourier domain is a classic problem in Compressed Sensing (CS) theory (Donoho 2006). A good basis that allows sparse representation is crucial for the reconstruction. Although some analytic bases, including discrete basis like wavelets (Menzel et al. 2011), and continuous basis like the SPF basis,

have been proposed as sparse bases for EAP estimation, a sparser basis can be learned from well chosen exemplars via Dictionary Learning (DL) techniques based on CS theory (Aharon et al. 2006, Mairal et al. 2010). (Bilgic et al. 2012) learns a discrete dictionary via the K-SVD (Aharon et al. 2006) approach and uses it in FOCal Underdetermined System Solver algorithm for EAP estimation. This strategy dramatically reduces the number of samples and scanning time required by DSI. However, because their dictionary is composed of a set of discrete basis vectors, Bilgic et al.'s approach suffers from numerical errors similar to DSI. (Merlet et al. 2012) learns a continuous dictionary, parametrized as a linear combination of some atoms adopted from SPF basis, from synthetic Gaussian signals, where the learned basis has the closed forms for ODF and EAP estimation due to the results of SPF basis (Cheng et al. 2010). However, there are some inherent limitations in both theoretical analysis and practical usage in (Merlet et al. 2012). For example, they learned the scale parameter ζ associated with the SPF basis from the training data, instead of the testing data. In addition, they have also neglected isotropic exemplars in the training data, causing over-fitting problems in less anisotropic areas such as the grey matter.

(Cheng et al. 2013) Dictionary Learning - Spherical Polar Fourier Imaging (DL-SPFI), for effective compressed-sensing reconstruction of the diffusion signal and the EAP. This approach offers a number of advantages over (Merlet et al. 2012). First, DL-SPFI dramatically reduces the dimensionality of the optimization problem by working in a small subspace of the SPF coefficients, instead of **q**-space. Second, the dictionary learned using DL-SPFI can be applied optimally and adaptively to each voxel by voxel-dependent determination of the optimal scale parameter. Third, DL-SPFI considers the constraint $E(0) = 1$ during both learning and estimation processes. However, all these DL methods perform reconstruction independently at each voxel by minimizing the square error loss function, which is very sensitive to noise and outliers. This essentially ignoring the functional nature of the HARDI data at different voxels in space.

The aim of this chapter is to develop a sparse multi-scale adaptive model (SMAM) to

spatially and adaptively infer the EAP of water diffusion in regions with complex fiber configurations using voxel-adaptive dictionary. In SMAM, we reformulate the HARDI imaging reconstruction as a regression problem using DL-SPFI. Similarity and distance weights are introduced to account for spatial smoothness of HARDI, while preserving the unknown discontinuities (e.g., edges between white matter and grey matter) of HARDI. We integrate DL estimation with the propagation-separation method (Polzehl and Spokoiny, 2000) to adaptively estimate EAPs across voxels. Experimental results indicate that SMAM can reduce the angle detection errors on fiber crossing area and provides more accurate reconstructions than the original DL-SPFI method and SMAM performs very well when the signals have low signal to noise ratio.

5.2 Methods

5.2.1 Model Formulation

We usually acquire n normalized HARDI data with each image containing N voxels for each subject. Thus, we observe n normalized HARDI measurements $\{(E(b_i;\mathbf{v}),\mathbf{g}_i,b_i) : i=1,\cdots,n\}$ at voxel $\mathbf{v}\in\mathcal{V}$, where $\mathbf{g}_i=(g_{i,1},g_{i,2},g_{i,3})^T$ is the gradient vector. Based on (2.3), these HARDI measurements can be also represented as $\mathbf{q}$-space measurements $\{E(\mathbf{q}_i) : i=1,\cdots,n\}$. We usually omit putting voxel $\mathbf{v}$, if no confusion in context. See Figure 2.8.

Most HARDIs assume that

$$f(E(\mathbf{q}_i;\mathbf{v}))=\mathbf{x}_i^T\beta(\mathbf{v})+\epsilon_i(\mathbf{v}), \tag{5.2}$$

where $f(\cdot)$ is a given transformation function (e.g., $f(s)=s$ or $f(s)=\log(s)$), $\mathbf{x}_i$ is a $p\times 1$ vector of covariates, which depends on $\mathbf{q}_i$ (or $(\mathbf{b}_i,r_i)$), $\beta(\mathbf{v})$ is a $p\times 1$ vector of regression coefficients, and $\epsilon_i(\mathbf{v})$ is an error term with mean zero and variance $\sigma_i^2(\mathbf{v})$. In practice, $E(\mathbf{q}_i;\mathbf{v})$ equals the ratio of magnetic resonance signal measured at $\mathbf{q}_i$, denoted by $S(\mathbf{q}_i;\mathbf{v})$,

to the magnetic resonance signal measured at $\mathbf{0}$, denoted by $S(\mathbf{0};\mathbf{v})$. Since the signal-to-noise ratio in $S(\mathbf{0};\mathbf{v})$ is very high, we ignore the noise component of $S(\mathbf{0};\mathbf{v})$. Model 5.2 is general enough to cover many existing HARDIs. In the literature, for generalized DTI and high order tensor (HOT), it is common to set $f(E(\mathbf{q}_i;\mathbf{v})) = \log(E(\mathbf{q}_i;\mathbf{v}))$ and represent $\log(E(\mathbf{q}_i;\mathbf{v}))$ as a polynomial function of $\mathbf{q}_i$, whereas for most other HARDIs, such as Q-ball imaging (QBI) or diffusion orientation transform (DOT), it is common to set $f(E(\mathbf{q}_i;\mathbf{v})) = E(\mathbf{q}_i;\mathbf{v})$ and approximate $E(\mathbf{q}_i;\mathbf{v})$ by a linear combination of some basis functions.

Spherical polar Fourier imaging (SPFI) is a model-free and fast HARDI method for multiple-shell data (Cheng et al. 2010). Let $B_{k,l,m}(\mathbf{q}) = G_k(q)Y_l^m(\mathbf{u})$ be spherical polar Fourier basis, where $Y_l^m(\mathbf{u})$ is the l order m degree Spherical Harmonic (SH) basis and $R_k(q)$ is the Gaussian-Laguerre polynomial basis. It was proposed to sparsely represent $E(\mathbf{q};\mathbf{v})$ (Assemlal et al. 2009). The SPFI is to fit a model given by

$$E(\mathbf{q}_i;\mathbf{v}) = \sum_{k=0}^{K}\sum_{l=0}^{L}\sum_{m=-l}^{l} a_{k,l,m}(\mathbf{v})B_{k,l,m}(\mathbf{q}_i) + \epsilon_i(\mathbf{v}), \tag{5.3}$$

where K and L are large integers.

$$B_{k,l,m}(\mathbf{q}_i) = G_k(||\mathbf{q}_i||)Y_l^m(\mathbf{u}) \tag{5.4}$$

$$R_k(||\mathbf{q}_i||) = \kappa_k(\zeta)\exp(-\frac{||\mathbf{q}_i||^2}{2\zeta})L_k^{1/2}(\frac{||\mathbf{q}_i||^2}{\zeta}) \tag{5.5}$$

$$\kappa_k(\zeta) = \left[\frac{2}{\zeta^{3/2}}\frac{k!}{\Gamma(k+3/2)}\right]^{1/2} \tag{5.6}$$

In this case, if we set $\mathbf{x}_i = (B_{0,0,0}(\mathbf{q}_i),\cdots,B_{K,L,L}(\mathbf{q}_i))^T$ and $\beta(\mathbf{v}) = (a_{0,0,0}(\mathbf{v}),\cdots,a_{K,L,L}(\mathbf{v}))^T$, then SPFI can be regarded as a special case of model (5.2). For each voxel $\mathbf{v}$, the SPF coefficients $\beta = (a_{0,0,0},\cdots,a_{K,L,L})^T$ can be estimated from the signal attenuation measurements $\{E(\mathbf{q}_i)\}$ via least square fitting with l_2 or l_1 regularization, where the constraint $E(0) = 1$ can be imposed by adding artificial samples at $||\mathbf{q}|| = 0$ (Cheng et al. 2010; 2011). It also

can be imposed by the equality obtained from $E(0) = 1$, $\sum_{k=0}^{K} a_{k,l,m} G_k(0) = \sqrt{4\pi}\delta_l^0, 0 \leq l \leq L, -l \leq m \leq l$ (Cheng et al. 2013). Based on this, we can separate the coefficient vector β into $\beta = (\beta_0^T, \beta'^T)^T$, where $\beta_0 = (a_{0,0,0}, \cdots, a_{0,L,L})^T$, $\beta' = (a_{1,0,0}, \cdots, a_{K,L,L})^T$, then β_0^T can be represented by β'^T

$$a_{0,l,m} = \frac{1}{G_0(0)} \left(\sqrt{4\pi}\delta_l^0 - \sum_{k=1}^{K} a_{k,l,m} G_k(0) \right), 0 \leq l \leq L, -l \leq m \leq l \tag{5.7}$$

Let $\mathbf{q}_i = q_i \mathbf{u}_i$, then the 5.3 can be written as

$$E(\mathbf{q}_i) - \frac{G_0(q_i|\zeta)}{G_0(0|\zeta)} = \sum_{k=1}^{K} \sum_{l=0}^{L} \sum_{m=-l}^{l} a_{k,l,m} \left(G_k(q_i|\zeta) - \frac{G_k(0|\zeta)}{G_0(0|\zeta)} G_0(q_i|\zeta) \right) Y_l^m(\mathbf{u}_i) + \epsilon_i, \tag{5.8}$$

$$X'\beta' = E' \tag{5.9}$$

where

$$X' = \begin{bmatrix} \left(G_1(q_1|\zeta) - \frac{G_1(0|\zeta)}{G_0(0|\zeta)} G_0(q_1|\zeta) \right) Y_0^0(\mathbf{u}_1) & \cdots & \left(G_K(q_1|\zeta) - \frac{G_K(0|\zeta)}{G_0(0|\zeta)} G_0(q_1|\zeta) \right) Y_L^L(\mathbf{u}_1) \\ \vdots & \ddots & \vdots \\ \left(G_1(q_n|\zeta) - \frac{G_1(0|\zeta)}{G_0(0|\zeta)} G_0(q_n|\zeta) \right) Y_0^0(\mathbf{u}_n) & \cdots & \left(G_K(q_n|\zeta) - \frac{G_K(0|\zeta)}{G_0(0|\zeta)} G_0(q_n|\zeta) \right) Y_L^L(\mathbf{u}_n) \end{bmatrix} \tag{5.10}$$

$$E' = \left[E(\mathbf{q}_1) - \frac{G_0(q_1|\zeta)}{G_0(0|\zeta)}, \cdots, E(\mathbf{q}_n) - \frac{G_0(q_n|\zeta)}{G_0(0|\zeta)} \right]^T$$

$$\beta' = (a_{1,0,0}, \cdots, a_{K,L,L})^T$$

β' can be calculated by the least square fitting with l_2 or l_1 regularization or some robust methods. Then β_0 can be obtained using 5.7, and the estimated β satisfies $E(0) = 1$. Moreover, it can be shown that the EAP at radius R_0, $p(R_0\mathbf{u}; \mathbf{v})$ can be written as

$$p(R_0\mathbf{r}; \mathbf{v}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} \left\{ 4(-1)^{l/2} \frac{\zeta^{0.5l+1.5} \pi^{l+1.5} R_0^l}{\Gamma(l+1.5)} \sum_{k=0}^{K} f_{k,l,m}(\zeta, R_0) a_{k,l,m} \right\} Y_l^m(\mathbf{u}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} c_{l,m} Y_l^m(\mathbf{u}) \tag{5.11}$$

$$f_{k,l,m}(\zeta, R_0) = \kappa_k(\zeta) \sum_{i=0}^{k} (-1)^i \binom{k+0.5}{k-i} \frac{1}{i!} 2^{0.5l+i-0.5} \Gamma(0.5l+i+1.5) 1F1(\frac{2i+l+3}{2}; l+\frac{3}{2}; -2\pi^2 R_0^2 \zeta) \tag{5.12}$$

$$1F1(a;b;x) = \sum_{k=0}^{\infty} \frac{(a)_k x^k}{(b)_k k!}, \qquad (a)_k = (a(a+1)...(a+l-1)), \text{with} (a)_0 = 1 \tag{5.13}$$

The Implementation includes two steps. The first step is to estimate coefficients β' of the signal, then calculate $\{a_{k,l,m}\}$ using 5.7. The second step is the linear analytical transform $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ of EAP profile $p(\mathbf{R_0})$, and the second step is independent of the first step. The linear transform from $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ could be implemented as an matrix multiplication. This transformation is independent with the data, since $\{f_{k,l,m}(\zeta, R_0)\}$ only depends on ζ and R_0. Once a R_0 and the basis are given, transformation matrix can be calculated. And since in SPFI, Only the value of $1F1$ at the fixed value $-2\pi^2 R_0^2 \zeta$ is needed, so the transformation matrix only needs to be calculated once.

The basis matrix is calculated using $q = b^{1/2}$ and $\zeta = 700$ where b values are from image acquisition. The reasoning for this choice is that considering $E(\mathbf{q}) = \exp(-4\pi^2 \tau q^2 D)$, $b = 4\pi^2 \tau q^2$, and a typical diffusion coefficient of $D = 0.7 \times 10^{-3} mm^2/s$, a typical b-value $b = 3000 s/mm^2$, we set $\zeta = \frac{1}{8\pi^2 \tau \times 0.7 \times 10^{-3}}$. If $4\pi^2\tau = 1$, then ζ is about 700. The transformation matrix from $\{a_{k,l,m}\}$ to $\{c_{l,m}\}$ will be calculated by setting $R_0 = 15\mu m$.

5.2.2 Dictionary Learning

Now let's consider the l_1 regularization problem

$$\min_{B,\Lambda} \sum_i ||\Lambda \beta_i'||_1 \;\; s.t. \;\; ||X'\beta_j' - E_j'||_2^2 \leq \epsilon_{DL} \;\; \forall j \tag{5.14}$$

$$\Leftrightarrow \min_{C,D} \sum_i ||c_i||_1 \;\; s.t. \;\; ||X'Dc_j - E_j'||_2^2 \leq \epsilon_{DL} \;\; \forall j \tag{5.15}$$

Where $B = (\beta_1', \cdots, \beta_Q')$ is the SPF coefficient matrix. The transform matrix D will result in a transformed SPF basis $X'D$ that can be used for even sparser representation of the signal. $C = (c_1', \cdots, c_Q')$ is the new coefficient matrix in association with the transformed basis.

(Bilgic et al. 2012) proposed to learn a dictionary from real data, as done in DL-FOCUSS, but the learned dictionary may be significantly affected by noise and the samll sample size. (Cheng et al. 2013) proposed an alternative solution to perform DL using some synthetic

data that approximate well the real signal. They proposed to learn a continuous basis using mixtures of Gaussian signals by solving

$$\min_{C,D} \sum_i ||c_i||_1 \;\; s.t. \;\; ||Dc_j - \beta'_j||_2^2 \leq \epsilon_{DL} \;\; \forall j \tag{5.16}$$

which is equivalent to 5.15, due to the orthogonality of the SPF basis for large K and L. Threshold ϵ_{DL} can be chosen simply as 0.01 for unit-norm normalized $\{\beta'_j\}$. (Cheng et al. 2013) proved that (1) the single tensor model is sufficient to learn a dictionary which sparsifies the multi-Gaussian signals; (2)The parameter ζ should be determined adaptively from testing signals. For signals generated from single tensor model with fixed mean diffusivity (MD) d_0, then for large enough K, fixed L, and small enough ϵ_{DL}, the optimal scale ζ for the DL problem 5.16 is $\zeta_* = (8\pi^2\tau d_0)^{-1}$. Signals were generated using the single tensor model with $d_0 = 0.7 \times 10^{-3} mm^2$,$\zeta_0 = (8\pi^2\tau d_0)^{-1}$, MD in range $[0.5, 0.9] \times 10^{-3}$, FA in range [0, 0.9], 321 directions equally distributed on unit sphere. The corresponding SPF coefficients $\{\beta'_j\}$ in 5.16 were computed with $K = 4, L = 8$ via LS estimation. The dictionary D was learned using the online method in (Mairal et al. 2010), with identity matrix as the initial value. By solving 5.16, 254 atoms were learned in D, including the isotropic atoms $\{B_{k,0,0}(\mathbf{q}_i)\}_{k=1}^K$. Note that the isotropic atoms are important so that grey matter and the CSF can be sparsely represented.

5.2.3 Estimation Procedures

After learning the dictionary D, we compute the estimation for voxel $\mathbf{v}$ using l_1 regularization. Let E denote the signal vector for voxel $\mathbf{v}$, we can compute the scale parameter $\zeta = (8\pi^2\tau d)^{-1}$ based on the estimated MD value d for the signal vector. Then the new basis matrix X' and new signal vector E' can be computed base on ζ. The coefficients c by computed by

$$\min_c ||X'Dc - E'||_2^2 + \lambda \sum_{j=1}^{p} \frac{c_j}{|\hat{c}_j|^\gamma} \tag{5.17}$$

where $\hat{c}$ is the least square estimator using LB regularization. (λ, γ) can be selected by two-dimensional cross-validation. Note that this is one type of adaptive lasso method (Zou 2006a), which has been proved to have oracle property. After estimating c, the SPF coefficients $\beta = (\beta_0^T, \beta'^T)^T$ can be computed by $\beta' = Dc$ and 5.7

A key feature in HARDI is its spatial constraint. Specifically, the orientation and anisotropy of any single fiber bundles change smoothly from one voxel to the next, particularly along the dominant fiber orientation, whereas it may change dramatically at the boundaries between tracts and interfaces with gray matter structures and cerebrospinal fluid (CSF) spaces. Moreover, the EAP profile or ODF is expected to change smoothly from one voxel to the next in the same fiber crossing region, whereas it may change dramatically at the boundaries of fiber crossing regions and surrounding fiber bundles. This is a very important and powerful constraint that can be exploited to improve the reconstruction in HARDI. However, the methods mentioned in previous sections are voxel-wise methods and do not make use of the spatial constraint of HARDI.

To explicitly exploit such spatial constraint, we develop a sparse multi-scale adaptive model (SMAM) to spatially and adaptively update $\{c(\mathbf{v}) : \mathbf{v} \in \mathcal{V}\}$. The key idea of SMAM is to combine HARDI signals in a neighboring sphere of voxel $\mathbf{v}$ to make inference on $c(\mathbf{v})$ at the voxel $\mathbf{v}$. Specifically, let $B(\mathbf{v}, h)$ be a sphere with radius h centered at voxel $\mathbf{v}$ and $\omega(\mathbf{v}, \mathbf{v}'; h)$ be a weight function of triple $(\mathbf{v}, \mathbf{v}', h)$ such that

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

SMAM is based on a set of weighted penalization functions, denoted by $P_n(c(\mathbf{v}); \omega, h)$, which is defined as follows:

$$P_n(c(\mathbf{v}); \omega, h) \quad = \quad ||E'_w(\mathbf{v}; h) - X'Dc(\mathbf{v})||_2^2 + \rho(c(\mathbf{v}); \lambda(\mathbf{v})) \tag{5.18}$$

where $E'_w(\mathbf{v}; h) = \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega(\mathbf{v}, \mathbf{v}'; h) E'(\mathbf{v}')$. Given the current weights $\{\omega(\mathbf{v}, \mathbf{v}'; h)$:

$\mathbf{v}, \mathbf{v}' \in \mathcal{V}\}$, we consider the weighted GEE estimator of $c(\mathbf{v})$, denoted by $\hat{c}(\mathbf{v}, h)$, which satisfies

$$\hat{c}(\mathbf{v}, h) = \text{argmin}_{c(\mathbf{v})} P_n(c(\mathbf{v}); \omega, h). \tag{5.19}$$

It is critical to choose a good $\omega(\mathbf{v}, \mathbf{v}'; h)$ in preventing oversmoothing the estimates of $c(\mathbf{v})$ across voxels, while preserving the edges between different structures, such as fiber bundles, crossing fibers, or gray matter regions. A good $\omega(\mathbf{v}, \mathbf{v}'; h)$ should quantify the similarity between $c(\mathbf{v})$ and $c(\mathbf{v}')$ or their corresponding EAPs. Specifically, if $c(\mathbf{v})$ and $c(\mathbf{v}')$ substantially differ from each other, then the HARDI signals in voxel $\mathbf{v}'$ do not contain too much information on $c(\mathbf{v})$ and thus $\omega(\mathbf{v}, \mathbf{v}'; h)$ should be close to 0. However, if $c(\mathbf{v})$ and $c(\mathbf{v}')$ are close to each other indicating that the HARDI signals in voxel $\mathbf{v}'$ contain useful information on $c(\mathbf{v})$, then $\omega(\mathbf{v}, \mathbf{v}'; h)$ should be significantly bigger than zero. See the explicit expression of $\omega(\mathbf{v}, \mathbf{v}'; h)$ in Section 5.2.4.

5.2.4 SMAM

We develop the SMAM procedure to adaptively determine $\mathbf{w}$ and estimate $c(\mathbf{v})$ across all voxels $\mathbf{v} \in \mathcal{V}$. Our multiscal adaptive strategy starts with building a sequence of nested spheres with increasing radiues $h_0 = 0 < h_1 < \cdots < h_S = r$ ranging from the smallest scale $h_0 = 0$ to a large scale $h_S = r$ at each voxel $\mathbf{v}$. At the scale $h_0 = 0$, we just calculate $c(\mathbf{v}; h_0) = \widehat{c}(\mathbf{v})$ voxel-wisely without using any spatial information. It corresponds to setting $w(\mathbf{v}, \mathbf{v}'; h_0) = 1$ if $\mathbf{v} = \mathbf{v}'$ and 0 otherwise. Then, based on the signals contained in voxels d and d', we use methods as detailed below to calculate weights $w(\mathbf{v}, \mathbf{v}'; h_1)$ at scale h_1 for all voxels $\mathbf{v}$. After getting the new weights $w(\mathbf{v}, \mathbf{v}'; h_1)$, we can update $c(\mathbf{v}; h_1)$. Then we can sequentially determine $w(\mathbf{v}, \mathbf{v}'; h_s)$ and adaptively update $c(\mathbf{v}; h_s)$. From $h_0 = 0$ to $h_S = r$, a path diagram of the multiscale adaptive strategy is given below:

$$\begin{array}{ccccccc} w(\mathbf{v}, \mathbf{v}'; h_0) & & w(\mathbf{v}, \mathbf{v}'; h_1) & & \cdots & w(\mathbf{v}, \mathbf{v}'; h_S = r) \\ \Downarrow & \nearrow & \Downarrow & \nearrow & \cdots & \Downarrow \\ c(\mathbf{v}; h_0) & & c(\mathbf{v}; h_1) & & \cdots & c(\mathbf{v}; h_S) \end{array}$$

SMAM consists of three key steps: (I) an initialization step, (II) a weighted estimation step, and (III) a stop checking step. In the initialization step, we prefix a geometric series $\{h_s = c_h^s : s = 1, ..., S\}$ of radii with $h_0 = 0$, where $c_h \in (1, 2)$, say $c_h = 1.15$ and $S = 10$. We use small c_h in order to prevent incorporating too many neighboring voxels at the beginning, and thus it improves the robustness of the procedure and the accuracy of parameter estimation. At $h_0 = 0$, we obtain $c(\mathbf{v}; h_0) = \widehat{c}(\mathbf{v})$ across all voxels by minimizing (5.18). We then set $s = 1$ and $h_1 = c_h$.

In the weighted estimation step, we first compute $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ to characterize the similarity between the two estimated Signals based on $X'D\hat{c}(\mathbf{v}; h_{s-1})$ and $X'D\hat{c}(\mathbf{v}'; h_{s-1})$ at voxels $\mathbf{v}$ and $\mathbf{v}'$ and the adaptive weights $\omega(\mathbf{v}, \mathbf{v}'; h_s)$, which are defined as

$$\omega(\mathbf{v}, \mathbf{v}'; h_s) = \frac{K_{loc}(||\mathbf{v} - \mathbf{v}'||_2/h_s)K_{st}(\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})/C_n)}{\sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} K_{loc}(||\mathbf{v} - \mathbf{v}'||_2/h_s)K_{st}(\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})/C_n)}, \tag{5.20}$$

where $K_{loc}(u)$ and $K_{st}(u)$ are two nonnegative kernel functions with compact support, C_n is a number associated with n, and $||\cdot||_2$ denotes the Euclidean norm of a vector (or a matrix).

We compute $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ as the similarity between the estimated signals in voxels $\mathbf{v}$ and $\mathbf{v}'$ for HARDI as follows, we may set $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1}) = ||X'D\hat{c}(\mathbf{v}; h_{s-1}) - X'D\hat{c}(\mathbf{v}'; h_{s-1})||_2/||X'D\hat{c}(\mathbf{v}; h_{s-1})||_2$.

The weights $K_{loc}(||\mathbf{v} - \mathbf{v}'||_2/h_s)$ give less weight to the voxel $\mathbf{v}' \in B(\mathbf{v}, h_s)$, whose location is far from the voxel $\mathbf{v}$. The $K_{loc}(\cdot)$ is a regular kernel function for smoothing the smoothed curves or surfaces. Some common choices of $K_{loc}(\cdot)$ include the Gaussian kernel and Epanechnikov kernel (Tabelow et al. 2006; 2008, Polzehl and Spokoiny 2000). We use $K_{loc} = (1 - u^2)_+$ throughout this paper. The weights $K_{st}(\cdot)$ downweight the voxels that are dissimilar to voxel d. The $\text{Dist}(\mathbf{v}, \mathbf{v}'; h_{s-1})$ takes large values if the estimated signals in voxel $\mathbf{v}$ differ significantly from those in voxel $\mathbf{v}'$. We set $K_{st} = \exp(-u^2/a)$, where a is a positive number.

After the calculation of $\omega(\mathbf{v}, \mathbf{v}'; h_s)$, we calculate the weighted HARDI signals of voxel $\mathbf{v}$, denoted by $E'_w(\mathbf{v}; h_s) = \sum_{\mathbf{v}' \in B(\mathbf{v}, h_s)} w(\mathbf{v}, \mathbf{v}'; h_s)E'(\mathbf{v}')$. Then, we use $E'_w(\mathbf{v}; h_s)$ to compute

$\hat{c}(d; h_s)$ at voxel $\mathbf{v}$. The computation of SMAM at each iteration is of the same order as that for the voxel-wise approach. Thus, this multiscale adaptive method provides an efficient method for adaptively exploring the neighboring voxels of each voxel. Since SMAM sequentially includes more data at each iteration, it will adaptively increase the statistical efficiency in estimating $c(\mathbf{v})$ in a homogenous region, while decreasing the variation of the weights $w(\mathbf{v}, \mathbf{v}'; h_s)$.

In the stop checking step, after the first iteration, we start to calculate a stopping criterion based on the normalized L_2 distance between $\widehat{c}(\mathbf{v}; h_s)$ and $\widehat{c}(\mathbf{v}; h_{s-1})$, denoted by $\text{Dist}_s(\mathbf{v}) = ||X'D\hat{c}(\mathbf{v}; h_s) - X'D\hat{c}(\mathbf{v}'; h_{s-1})||_2 / ||X'D\hat{c}(\mathbf{v}; h_{s-1})||_2$. We use $\text{Dist}_s(\mathbf{v})$ to determine whether 'bad' HARDI signals from neighboring voxels lead to a dramatic change in the estimated $\widehat{c}(\mathbf{v}; h_{s-1})$. If $\text{Dist}_s(\mathbf{v}) > C_s$, where C_s is a positive scalar, then we set $\widehat{c}(\mathbf{v}; h_s) = \widehat{c}(\mathbf{v}; h_{s-1})$ and $s = S$ for voxel $\mathbf{v}$. If $s = S$ for all voxels, we stop. If $\text{Dist}_s(\mathbf{v}) \leq C_s$, then we set $h_{s+1} = c_h h_s$, increase s by 1, and continue with the weighted estimation step. In practice, different voxels may stop at different bandwidths, indicating that different degrees of smoothness are used to reconstruct HARDI. We set $C_s = \chi^2(1)^{0.6/s} \bar{D}_{med}$ to prevent oversmoothing, where $\chi^2(1)^a$ is the upper $1 - a$ percentile of the $\chi^2(1)$ distribution. As s increases, C_s decreases to zero. Moreover, $\bar{D}_{med}$ is chosen to be the median of $\{\text{Dist}(\mathbf{v}, \mathbf{v}'; h_0) : \mathbf{v} \neq \mathbf{v}'\}$, where $\mathbf{v}$ and $\mathbf{v}'$ are M preselected voxels $\{\mathbf{v}_i\}_{i=1}^M$ from HARDI.

Finally, we summarize the SMAM algorithm 3 for the adaptive estimation of the ODF at voxel $\mathbf{v}$ below.

5.2.5 Maxima Extraction

Based on $\widehat{p}(\mathbf{v}; h_S)|_{R_0}$ at voxel $\mathbf{v}$, we need to extract its maxima in order to infer fiber directions. Although there are other more complicated methods for extracting maximal directions, such as the method presented in (Hlawitschka and Scheuermann 2005), spherical Newton's method (Tuch 2004), and Powell's method (Jansons and Alexander 2003), we take a simple thresholding approach in this paper. Specifically, we project the estimated EAP onto the sphere tessellated with a triangle mesh, which has 2562 points on the unit sphere.

Algorithm 3: PMARM at voxel $\mathbf{v}$

Input: Signals $E(\mathbf{v})$ and SPF matrix $\mathbf{X}$

Output: Estimated EAP profile at Radius $R_0 = 15\mu m$, $\widehat{p}(\mathbf{v}; h_S)|_{R_0}$

Calculate the MD value d, and $\zeta = (8\pi^2\tau d)^{-1}$.
Calculate X' and E' from 5.10.
Estimate $\hat{\mathbf{c}}(\mathbf{v})$ from (5.19).
for $s \leftarrow 1$ ***to*** S **do**
    calculate the weights $w(\mathbf{v}, \mathbf{v}'; h_s)$ for $d' \in B(d, h_s)$ by (5.20);
    calculate the weighted signals of voxel $\mathbf{v}$ by using
     $E'_w(\mathbf{v}; h) = \sum_{\mathbf{v}' \in B(\mathbf{v},h)} \omega(\mathbf{v}, \mathbf{v}'; h) E'(\mathbf{v}')$;
    calculate $\hat{\mathbf{c}}(\mathbf{v}; h_s)$ based on (5.19);
    calculate $\text{Dist}_s(\mathbf{v})$.
    If $\text{Dist}_s(\mathbf{v}) > C_s$, $\hat{\mathbf{c}}(\mathbf{v}; h_S) = \hat{\mathbf{c}}(\mathbf{v}; h_{s-1})$, and $s = S$,
    else
        $h_{s+1} = c_h h_s$;
return $\widehat{p}(\mathbf{v}; h_S)|_{R_0}$.

If the estimated EAP value at a mesh point is greater than the corresponding value at all its neighboring mesh points and this estimated EAP value is greater than max(EAP)/2, then the direction at this mesh point is regarded as a maximum. This thresholding method avoids selecting small peaks that may appear due to noise.

5.3 Simulation Study

5.3.1 Data Generating

We examined the finite sample performance of our SMAM on decting crossing fibers by using synthetic HARDI data generated from the multi-tensor model (Alexander et al. 2002, Tuch 2004). We simulated the diffusion-weighted signals according to

$$E(\mathbf{q}_i; \mathbf{v}) = \sqrt{(\sum_{t=1}^{T} p_t e^{-b_i \mathbf{u}_i^T \mathbf{D}_t(\mathbf{v}) \mathbf{u}_i} + \sigma\epsilon_{i1})^2 + (\sigma\epsilon_{i2})^2} \text{ for } i = 1, \cdots, n \tag{5.21}$$

for $i = 1, \cdots, n$, where $\mathbf{q}_i = q_i \mathbf{u}_i$ with $\mathbf{u}_i$ being a unit vector, T is the number of fibers, p_t is the weight for t-th fiber, b is the b-value and $\mathbf{D}_t(\mathbf{v})$ is the tensor matrix for the t-th fiber, SNR $= 1/\sigma$, and ϵ_{i1} and ϵ_{i2} are independently simulated from the standard normal

distribution. We used the multi-tensor model (5.21) to generate different phantoms with different regions of interest (ROIs) with 81 sampling directions on the hemisphere for the 3rd order tessellation of the icosahedron and $b = 500, 1000, 2000, 3000\text{s/mm}^2$. Specifically, voxels with a single fiber were generated from a single tensor model using diffusion tensor profiles with eigenvalues $[1.7, 0.3, 0.3] \times 10^{-3}\text{mm}^2/\text{s}$, voxels with two fiber directions were generated by two-tensor model $E(\mathbf{q}_i; \mathbf{v}) = e^{-b_i \mathbf{u}_i^T \mathbf{D}_1(\mathbf{v}) \mathbf{u}_i}/2 + e^{-b_i \mathbf{u}_i^T \mathbf{D}_2(\mathbf{v}) \mathbf{u}_i}/2$, and voxels isotropic tensors were generated by the single tensor model using diffusion tensor profiles with eigenvalues $[1, 1, 1] \times 10^{-3}\text{mm}^2/\text{s}$. We estimated EAP at each voxel by using the weighted l_1 and its corresponding SMAM, weighted l_1-SMAM. For our smoothing procedure, we set $K_{st} = \exp(-u^2)$, $C_n = 1$, $S = 10$. For SPFI, we use all data in 4 shells and chose $SNR = 5, 7, 10, 12$, and $K = 4$,$L = 8$, $\lambda = 1e - 8$, $\gamma = 0.5$, $\zeta = 700$ for all experiments. Then we extracted the EAP maxima aligned with fiber directions.

5.3.2 Angle Detection in First Phantom with 90^o crossing fibers

In the first phantom, we included four different ROIs including isotropic ROIs, two single fiber ROIs with its direction going either along the $x-$axis (ROI1) or along the $y-$axis (ROI2), and the 90^o crossing fiber ROIs(ROI3).

Figure 5.1 presents the estimated EAP images for this type of phantom from different settings. The first row gives the EAP reconstruction results using Weighted l_1 estimation without and with SMAM. We can see that Weighted l_1 with SMAM gives smoother result than without. The panel in the second row shows the EAP reconstruction from noise free data. We can see that the resutl with SMAM is the closer to noise-free result. Because SMAM can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

To quantify the accuracy of detection angle, we generated 1,000 data sets for three different SNRs including 5 ,7, 10 and 12. We estimated the EAPs by using voxel-wise weighted l_1 and weighted l_1-SMAM . Then we extracted the EAP maxima aligned with fiber directions. For voxels with a single fiber, we calculated angle detection errors by

comparing recovered fiber directions with the ground truth. For voxels with two crossing fibers, we calculated angle detection errors by comparing recovered crossing angles with the ground truth. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for voxels with one fiber and two fibers are presented in Table 5.1.

Table 5.1 reveals that the mean of the angular errors are substantially reduced and the percentage of detecting correct number of fibers are increased by adding our adaptive smoothing techique. It may indicate that the sparse multi-scale adaptive model can efficiently exploit spatial smoothness for reconstructing EAPs, while reducing noise leading to better angle detection. For the voxels with multiple fibers, adding SMAM can significantly improve the The percentage of detecting correct number of fibers, especially in the data with low SNR.

5.3.3 Angle Detection in Second Phantom with Twisted Crossing

In the second phantom, we included a twisted crossing region in the middle, one fiber along $x-$axis (ROI1), and the other fiber with changing angles with $x-$axis from 30^o, 45^o , 60^o, 75^o, to 90^o and then from 90^o, 75^o, 60^o, 45^o to 30^o. We marked all region with single fiber, which is not along $x-$axis as ROI2, and all regions with crossing fibers as ROI3.

Figure 5.2 presents the estimated EAP images for this type of phantom from different settings. The first row gives the EAP reconstruction results using Weighted l_1 estimation without and with SMAM. We can see that Weighted l_1 with SMAM gives smoother result than without. The panel in the second row shows the EAP reconstruction from noise free data. We can see that the resutl with SMAM is the closer to noise-free result. Because SMAM can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

To quantify the accuracy of detection angle, we generated 1,000 data sets for three different SNRs including 5 ,7, 10 and 12. We estimated the EAPs by using voxel-wise

weighted l_1 and weighted l_1-SMAM . Then we extracted the EAP maxima aligned with fiber directions. For voxels with a single fiber, we calculated angle detection errors by comparing recovered fiber directions with the ground truth. For voxels with two crossing fibers, we calculated angle detection errors by comparing recovered crossing angles with the ground truth. The mean of the angular errors at each voxel and the percentage of detecting correct number of fibers are calculated based on the 1000 simulations using each estimation method. The average values of these detection errors for voxels with one fiber and two fibers are presented in Table 5.2.

Table 5.2 reveals that the mean of the angular errors are substantially reduced and the percentage of detecting correct number of fibers are increased by adding our adaptive smoothing techique. It may indicate that the sparse multi-scale adaptive model can efficiently exploit spatial smoothness for reconstructing EAPs, while reducing noise leading to better angle detection. For the voxels with multiple fibers, adding SMAM can significantly improve the The percentage of detecting correct number of fibers, especially in the data with low SNR.

5.4 Real Data Analysis

We tested our proposed method on public data from the NIH Human Connectome Project (HCP). 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. Diffusion weighting consisted of 3 shells of b=1000, 2000, and 3000 s/mm^2 interspersed with an approximately equal number of acquisitions on each shell within each run. Figure 5.3 gives EAP recovery result using weighted l_1 estimation. One slice is presented and two ROIs are selected for furthur analysis.

In order to test our sparse multi-scale adaptive model(SMAM) method, we added racian noise with $SNR = 5$ to the HCP data, then recovered the EAPs using weighted l_1 and

weighted l_1-SMAM, where SMAM is setup the same as simulation study. Figure 5.4 shows the EAP recovery of ROI1 from Figure 5.3. Panel (a) gives the EAP results from noise added data using Weighted l_1; Panel (b) gives the EAP results from noise added data using Weighted l_1-MASS; Panel (c) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from Weighted l_1-MASS is much closer to (c) when compared to Weighted l_1. Similar results are shown in Figure 5.5, the EAP recovery of ROI2 from Figure 5.3.

Panel (a) gives the EAP results from noise added data using Weighted l_1; Panel (b) gives the EAP results from noise added data using Weighted l_1-MASS; Panel (c) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from Weighted l_1-MASS is much closer to (c) when compared to Weighted l_1.

5.5 Conclusion

We have introduced a sparse multi-scale adaptive model (SMAM) framework to adaptively and sequentially reconstruct the EAPs across all voxels from HARDI signals using SPFI dictionary learning. In simulation data analysis, we have shown that adding SMAM to the weighted l_1 estimation can substantially reduce the angle dectection error and increase the accuracy of dectecting the correct number of fibers in each voxels. Because SMAM reconstructs the EAPs at each voxel by adaptively borrowing the spatial information from the neighboring voxels, then can substantially reduce the noise level, while improving the EAP reconstruction. This is also shown in the real data study.

Table 5.1: The mean of anglular errors under given EAP configuration and the percentage of correct number of detected EAP maximum in the first phantom. 1,000 simulated data sets were used .

All voxels with one fiber

SNR	Weighted l_1	Weighted l_1-SMAM
5	5.41 (94.35 %)	2.050 (99.97 %)
7	3.25 (99.66 %)	1.25 (100 %)
10	1.93 (100 %)	0.70 (100 %)
12	1.47 (100 %)	0.49 (100 %)

All voxels with two fibers

SNR	Weighted l_1	Weighted l_1-SMAM
5	16.18 (50.99 %)	6.95 (91.77 %)
7	9.00 (76.98 %)	3.56 (98.52 %)
10	4.87 (95.35 %)	1.93 (99.62 %)
12	3.85 (98.17 %)	1.52 (99.48 %)

Table 5.2: The mean of anglular errors under given EAP configuration and the percentage of correct number of detected EAP maximum in the second phantom. 1,000 simulated data sets were used .

All voxels with one fiber

SNR	Weighted l_1	Weighted l_1-SMAM
5	5.49 (94.77 %)	3.10 (99.99 %)
7	3.78 (99.73 %)	2.69 (100.00 %)
10	2.66 (100.00 %)	2.23 (100.00 %)
12	2.26 (100.00 %)	2.13 (100.00 %)

All voxels with two fibers

SNR	Weighted l_1	Weighted l_1-SMAM
5	11.57 (50.87 %)	7.99 (90.95 %)
7	8.05 (76.83 %)	5.75 (98.55 %)
10	5.32 (94.81 %)	3.53 (99.46 %)
12	4.19 (98.04 %)	2.82 (99.48 %)

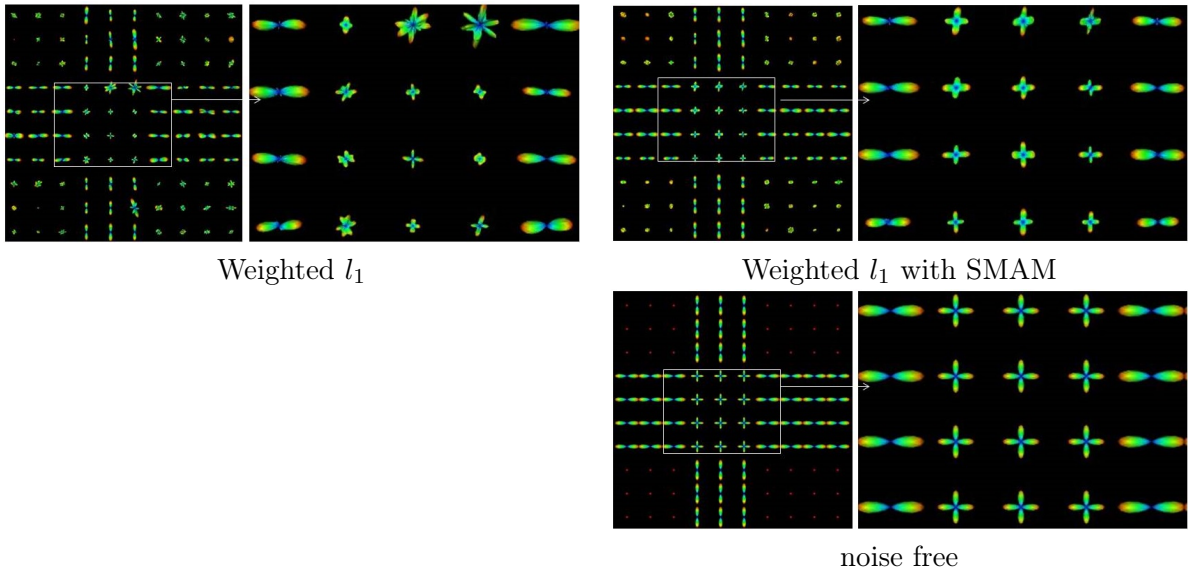


Figure 5.1: Comparison of EAP estiamtions with and without SMAM on data with 90 degree crossing and SNR=5: The first row gives the EAP reconstruction results using Weighted l_1 estimation without and with SMAM. We can see that Weighted l_1 with SMAM gives smoother result than without. The panel in the second row shows the EAP reconstruction from noise free data. We can see that the resutl with SMAM is the closer to noise-free result. Because SMAM can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

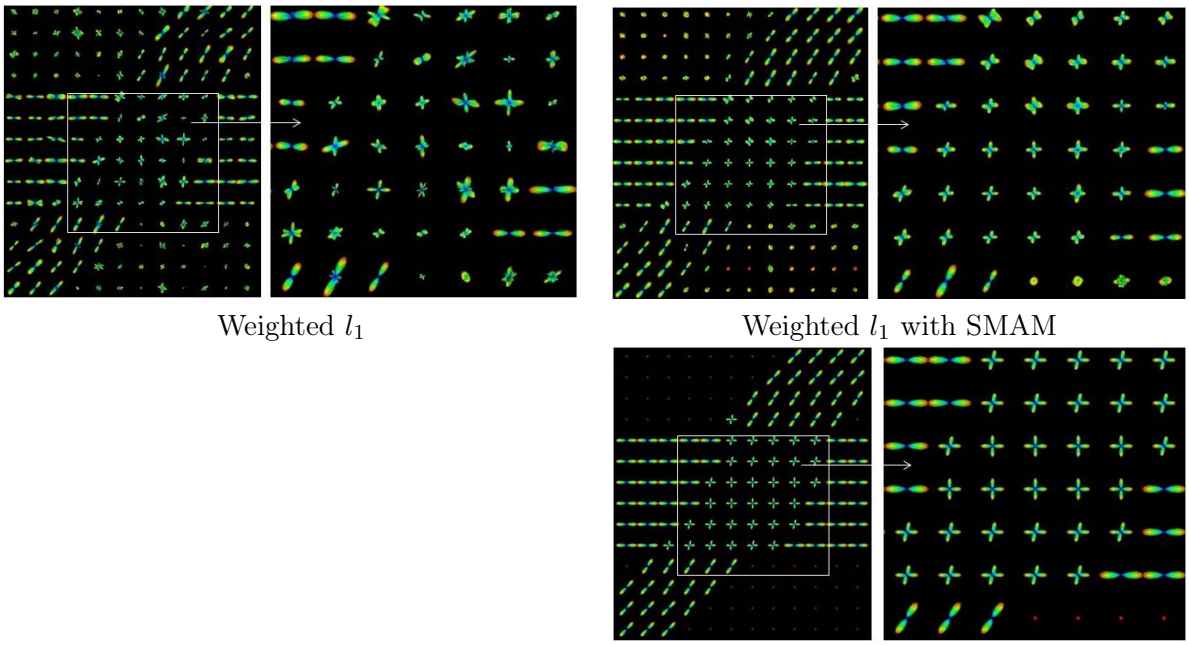


Figure 5.2: Comparison of EAP estiamtions with and without SMAM on data with twisted crossing and SNR=5: The first row gives the EAP reconstruction results using Weighted l_1 estimation without and with SMAM. We can see that Weighted l_1 with SMAM gives smoother result than without. The panel in the second row shows the EAP reconstruction from noise free data. We can see that the resutl with SMAM is the closer to noise-free result. Because SMAM can reduce the noise by incorporation useful neighborhood information into the estimation procedure.

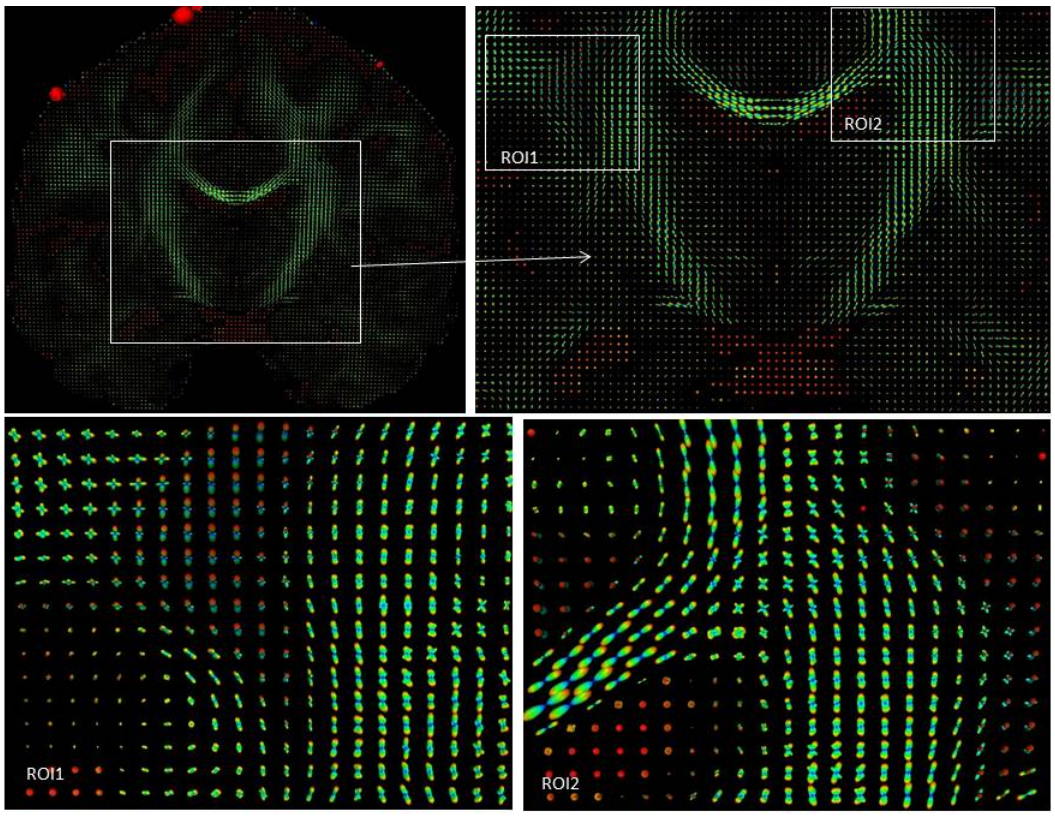
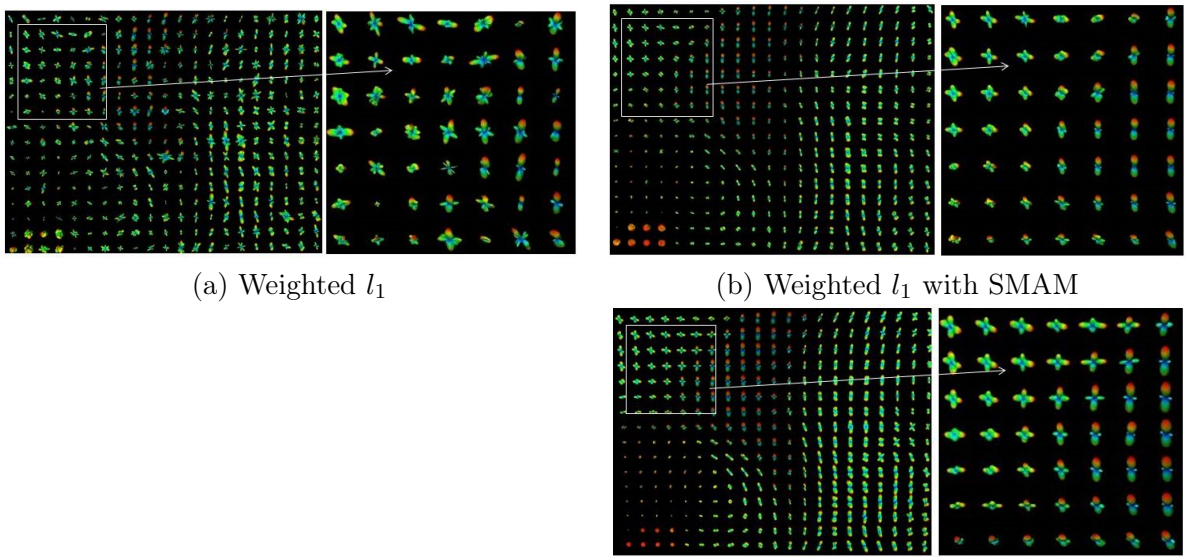


Figure 5.3: EAP recovery result of data from the NIH Human Connectome Project using weighted l_1 estimation.

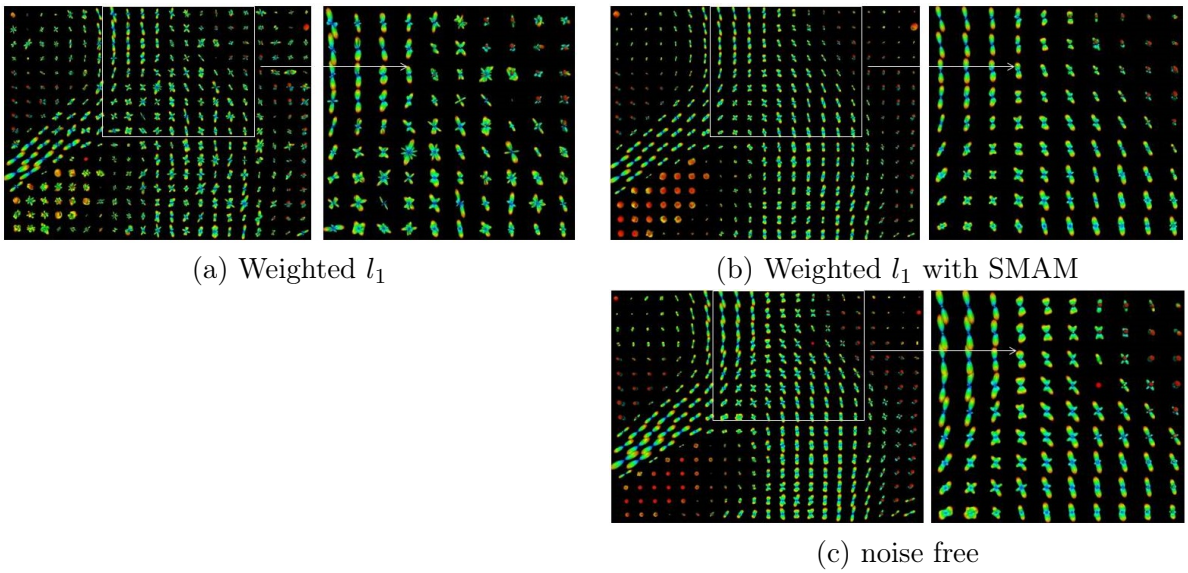


(a) Weighted l_1

(b) Weighted l_1 with SMAM

(c) noise free

Figure 5.4: EAP recovery of ROI1 from Figure 5.3. Panel (a) gives the EAP results from noise added data using Weighted l_1; Panel (b) gives the EAP results from noise added data using Weighted l_1-MASS; Panel (c) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from Weighted l_1-MASS is much closer to (c) when compared to Weighted l_1.



(a) Weighted l_1

(b) Weighted l_1 with SMAM

(c) noise free

Figure 5.5: EAP recovery of ROI2 from (Figure 5.3). Panel (a) gives the EAP results from noise added data using Weighted l_1; Panel (b) gives the EAP results from noise added data using Weighted l_1-MASS; Panel (c) give EAP recovery from original HCP data, meaning without added racian noise. We can see that results from Weighted l_1-MASS is much closer to (c) when compared to Weighted l_1.

CHAPTER 6: DISCUSSION

In summary, we have introduced three multi-scale adaptive smoothing and denoising procedures, PMARM, MASS and SMAM for improving ODF or EAP reconstructions from DWI images. We have shown that adding these procedures to regular statistical estimating methods can substantially reduce the angle detection error and increase the accuracy of detecting the correct number of fibers in each voxels. Because these procedures reconstructs the ODFs or EAPs at each voxel by adaptively borrowing the spatial information from the neighboring voxels, then can substantially reduce the noise level, while improving the EAP reconstruction. This is shown in both the simulation studies and real data studies.

However, these adaptive smoothing procedures are not perfect either. The most important issue with these procedures is how to avoid over-smoothing, especially on the boundary of anisotropic and isotropic regions, i.e. the boundary of white matter and grey matter. PMARM and SMAM are denoising signals directly while MASS is smoothing the parameter estimations. Based on our experiments, PMARM and SMAM are relatively easier over-smooth the ODFs and EAPs compared to MASS. In order to avoid over-smoothing, a good stopping criterion is required. In our future projects, we are planning to combine denoising the signal and smoothing the parameter estimates together to possibly develope an more intelligent stopping rule which can balance over-smoothing and under-smoothing.

In three procedures we proposed, **q**-space signals are represented by continuous bases, such as Spherical Harmonics basis and Spherical Polar Fourier basis, which allow the closed forms of EAP and the ODF estimation. Among these procedures, reconstructing 3D ensemble average propagator (EAP) describing the diffusion process to obtain richer information on the complex microstructure of biological tissues is more attractive than reconstructing

the ODF, which only captures the angular structure of the diffusion process.

For two procedures using SPFI, dictionary learning method provide sparser representation for EAPs and handle signals with low signal-to-noise ratio. In SMAM, we denoise the signal after removing the approximated isotropic Gaussian part calculated using voxel specific parameter ζ, therefore the effectiveness of the denoising is more obvious than using MASS. However, the learned dictionary used in SMAM is from synthetic single fiber data, it may not be applied to every image data. In the future project, we want to combine this learning technique and our smoothing technique to learn the dictionary from the real data.

REFERENCES

Aganj, I., Lenglet, C., and Sapiro, G. (2009), "ODF reconstruction in q-ball imaging with solid angle consideration," in *ISBI*.

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.

Aharon, M., Elad, M., and Bruckstein, A. (2006), "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *Signal Processing, IEEE Transactions on*, 54, 4311–4322.

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, 331–340.

Alexander, D. C. (2005), "Multiple-Fiber Reconstruction Algorithms for Diffusion MRI," *Annals of the New York Academy of Sciences*, 1064, 113–133.

— (2008), "A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features," *Magnetic Resonance in Medicine*, 60, 439–448.

Anderson, A. (2005), "Measurement of fiber orientation distributions using high angular resolution diffusion imaging," *Magnetic Resonance in Medicine*, 54, 1194–1206.

Assaf, Y., Freidlin, R., Rohde, G., and Basser, P. (2004), "New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter," *Magnetic Resonance in Medicine*, 52, 965–978.

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

Assemlal, H.-E. (2010), "Diffusion MR image analysis for the estimation of tissues local architecture," Ph.D. thesis, the University of Caen.

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.

Barmpoutis, A., Vemuri, B. C., and Forder, J. R. (2008), "Fast displacement probability profile approximation from hardi using 4th-order tensors," in *ISBI*.

Basser, P. J., Mattiello, J., and LeBihan, D. (1994a), "Estimation of the effective self-diffusion tensor from the NMR spin echo," *Journal of Magnetic Resonance Series B*, 103, 247–254.

— (1994b), "MR Diffusion Tensor Spectroscropy and Imaging," *Biophysical Journal*, 66, 259–267.

Basser, P. J. and Pierpaoli, C. (1996), "Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI," *Journal Of Magnetic Resonance*, 111, 209–219.

Becker, S., Tabelow, K., Mohammadi, S., Weiskopf, N., and Polzehl, J. (2012a), "Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by msPOAS," Tech. rep., Weierstrass-Institute.

Becker, S., Tabelow, K., Voss, H., Anwander, A., Heidemann, R., and Polzehl, J. (2012b), "Position-orientation adaptive smoothing of diffusion weighted magnetic resonance data (POAS)," *Med. Image Anal.,*, 16, 1142–1155.

Behrens, T., Berg, H., Jbabdi, S., Rushworth, M., and Woolrich, M. (2007), "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?" *Neuroimage*, 34, 144–155.

Bilgic, B., Setsompop, K., Cohen-Adad, J., Yendiki, A., Wald, L. L., and Adalsteinsson, E. (2012), "Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries," *Magnetic Resonance in Medicine*.

Callaghan, P. T. (1991), *Principles of nuclear magnetic resonance microscopy*, Oxford University Press.

Canales-Rodrıguez, E. J., Melie-Garcıa, L., and Iturria-Medina, Y. (2009), "Mathematical description of q-space in spherical coordinates: Exact q-ball imaging," *Magnetic Resonance In Medicine*, 61, 1350–1367.

Chefd'hotel, C., Tschumperlé, D., Deriche, R., and Faugeras, O. (2004), "Regularizing Flows for Constrained Matrix-Valued Images," *Journal of Mathematical Imaging and Vision*, 20, 147–162.

Cheng, J., Ghosh, A., Jiang, T., and Deriche, R. (2010), "Model-free and analytical EAP reconstruction via spherical polar Fourier diffusion MRI," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, Springer, pp. 590–597.

Cheng, J., Jiang, T., Deriche, R., Shen, D., and Yap, P.-T. (2013), "Regularized Spherical Polar Fourier Diffusion MRI with Optimal Dictionary Learning," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, Springer, pp. 639–646.

Cheng, J., Jiang, T., Deriche, R. D., et al. (2011), "Theoretical analysis and practical insights on EAP estimation via a unified HARDI framework," in *MICCAI workshop on Computational Diffusion MRI*.

Dejerine, J. J. (1895), *Anatomie des centres nerveux*, vol. 1, Rueff.

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

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2006), "Apparent Diffusion Coefficients from High Angular Resolution Diffusion Images: Estimation and Application," *Magnetic Resonance in Medicine*, 56, 395–410.

— (2007), "Regularized, Fast and Robust Analytical Q-ball Imaging," *Magnetic Resonance in Medicine*, 58, 497–510.

Descoteaux, M., Deriche, R., Bihan, D., Mangin, J., and Poupon, C. (2010), "Multiple q-Shell Diffusion Propagator Imaging," *Medical Image Analysis*.

Descoteaux, M., Deriche, R., Bihan, D. L., Mangin, J.-F., and Poupon, C. (2009), "Diffusion Propagator Imaging: Using Laplace's Equation and Multiple Shell Acquisitions to Reconstruct the Diffusion Propagator," in *IPMI*.

Descoteaux, M., Deriche, R., Knösche, T. R., and Anwander, A. (2008a), "Deterministic and Probabilistic Tractography Based on Complex Fiber Orientation Distributions," *IEEE Transactions in Medical Imaging*, 28, 269–286.

Descoteaux, M., Wiest-Daesslé, N., Prima, S., Barillot, C., and Deriche, R. (2008b), "Impact of rician adapted non-local means filtering on HARDI," *Medical Image Computing and Computer-Assisted Intervention*, 5242, 122–130.

Donoho, D. (2006), "Compressed sensing," *Information Theory, IEEE Transactions on*, 52, 1289–1306.

Douek, P., Turner, R., Pekar, J., Patronas, N., and LeBihan, D. (1991), "MR color mapping of myelin fiber orientation," *Journal of Computer Assisted Tomography*, 15, 923–929.

Dyrby, T. B., Baaré, W. F., Alexander, D. C., Jelsing, J., Garde, E., and Søgaard, L. V. (2011), "An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets," *Human brain mapping*, 32, 544–563.

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.

Fillard, P., Pennec, X., Arsigny, V., and Ayache, N. (2007), "Clinical DT-MRI Estimation, Smoothing, and Fiber Tracking With Log-Euclidean Metrics," *IEEE Transactions On Medical Imaging*, 26, 1472–1482.

Goh, I., Lenglet, C., Thompson, P. M., and Vidald, R. (2011), "A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry," *Neuroimage*, 56, 1181–1201.

Gray, H. (1918), *Anatomy of the human body*, Lea & Febiger.

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, S205–S223.

Hess, C. P., Mukherjee, P., Han, E. T., Xu, D., and Vigneron, D. B. (2006), "Q-Ball Reconstruction of Multimodal Fiber Orientations Using The Spherical Harmonic Basis," *Magnetic Resonance In Medicine*, 56, 104–117.

Hlawitschka, M. and Scheuermann, G. (2005), "HOT-lines: Tracking lines in higher order tensor fields," in *Visualization, 2005. VIS 05. IEEE*, IEEE, pp. 27–34.

Huber, P. J. (1964), "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, 35, 73–101.

Huber, P. J. and Ronchetti, E. M. (1975), "Robustness of Design," *Robust Statistics, Second Edition*, 239–248.

Jansons, K. and Alexander, D. C. (2003), "Persistent angular structure: new insights fom diffusion magnetic resonance imaging data," *Inverse Problems*, 19, 1031–1046.

Johansen-Berg, H. and Behrens, T. E. (2009), *Diffusion MRI: From quantitative measurement to In vivo neuroanatomy*, Elsevier.

Kim, Y., Thompson, P., Toga, A., Vese, L., and Zhan, L. (2009), "HARDI denoising: variational regularization of the spherical apparent diffusion coefficient sADC," *Information Processing in Medical Imaging*, 515–527.

Koay, C. G., Chang, L.-C., Carew, J. D., Pierpaoli, C., and Basser, P. J. (2006), "a unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging," *Journal of Magnetic Resonance*, 182, 115–125.

Le Bihan, D. et al. (2003), "Looking into the functional architecture of the brain with diffusion MRI," *Nature Reviews Neuroscience*, 4, 469–480.

LeBihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., and Jeantet, M. L. (1986), "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, 61, 401–407.

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.

Liu, C., Bammer, R., Acar, B., and Moseley, M. E. (2004), "Characterizing Non-Gaussian Diffusion by Using Generalized Diffusion Tensors," *Magnetic Resonance In Medicine*, 51, 925–937.

Liu, M., Vemuri, B., and Deriche, R. (2013), "A robust variational approach for simultaneous smoothing and estimation of DTI," *NeuroImage*, 67, 33–41.

MacLean, P. D. and Pribram, K. H. (1953), "Neuronographic analysis of medial and basal cerebral cortex. I. Cat," *J. Neurophysiol*, 16, 312–323.

Mairal, J., Bach, F., Ponce, J., and Sapiro, G. (2010), "Online learning for matrix factorization and sparse coding," *The Journal of Machine Learning Research*, 11, 19–60.

Menzel, M. I., Tan, E. T., Khare, K., Sperl, J. I., King, K. F., Tao, X., Hardy, C. J., and Marinelli, L. (2011), "Accelerated diffusion spectrum imaging in the human brain using compressed sensing," *Magnetic Resonance in Medicine*, 66, 1226–1233.

Merlet, S., Caruyer, E., and Deriche, R. (2012), "Parametric dictionary learning for modeling EAP and ODF in diffusion MRI," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, Springer, pp. 10–17.

Moseley, M., Cohen, Y., Mintorovitch, J., Kucharczyk, J., Tsuruda, J., Weinstein, P., and Norman, D. (1990), "Evidence of Anisotropic Self-Diffusion," *Radiology*, 176, 439–445.

Osment, P., Packer, K., Taylor, M., Attard, J., Carpenter, T., Hall, L., Herrod, N., Doran, S., Gordon, R., Hahn, E., et al. (1990), "NMR Imaging of Fluids in Porous Solids [and Discussion]," *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 333, 441–452.

Özarslan, E., Koay, C., Shepherd, T., Blackband, S., and Basser, P. (2009), "Simple harmonic oscillator based reconstruction and estimation for three-dimensional q-space mri," in *International Society for Magnetic Resonance in Medicine (ISMRM)*, p. 1396.

Özarslan, E., Shepherd, T. M., Vemuri, B. C., Blackband, S. J., and Mareci, T. H. (2006), "Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT)," *NeuroImage*, 31, 1086–1103.

Özarslan, E., Vemuri, B., and Mareci, T. (2005), "Generalized scalar measures for diffusion MRI using trace, variance, and entropy," *Magnetic Resonance in Medicine*, 53, 866–876.

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. G. (2000), "Adaptive weights smoothing with applications to image restoration," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62, 335–354.

— (2006), "Propagation-separation approach for local likelihood estimation," *Probab. Theory Relat. Fields*, 135, 335–362.

Polzehl, J., Voss, H. U., and Tabelow, K. (2010), "Structural adaptive segmentation for statistical parametric mapping," *NeuroImage*, 52, 515–523.

Raj, A., Hess, C., and Mukherjee, P. (2011), "Spatial HARDI: Improved visualization of complex white matter architecture with Bayesian spatial regularization," *Neuroimage*, 54, 396–409.

Selden, N. R., Gitelman, D. R., Salamon-Murayama, N., Parrish, T. B., and Mesulam, M.-M. (1998), "Trajectories of cholinergic pathways within the cerebral hemispheres of the human brain." *Brain*, 121, 2249–2257.

Stejskal, E. and Tanner, J. (1965), "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient," *The Journal of Chemical Physics*, 42, 288–292.

Sun, T. and Zhang, C.-H. (2012), "Scaled sparse linear regression," *Biometrika*, 99, 879–898.

Tabelow, K., Polzehl, J., Spokoiny, V., and Voss, H. U. (2008), "Diffusion tensor imaging: Structural adaptive smoothing," *NeuroImage*, 39, 1763–1773.

Tabelow, K., Polzehl, J., Voss, H. U., and Spokoiny, V. (2006), "Analyzing fMRI experiments with structural adaptive smoothing procedures," *NeuroImage*, 33, 55–62.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society (Series B)*, 58, 267–288.

Tibshirani, R. J. and Taylor, J. (2011), "The solution path of the generalized lasso," *Ann. Statist.*, 39, 1335–1826.

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., Gadian, D., and Connelly, A. (2004), "direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution," *NeuroImage*, 23, 1176–1185.

Tristán-Vega, A., Westin, C., and Aja-Fernández, S. (2009a), "Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging," *NeuroImage*, 47, 638–650.

Tristán-Vega, A., Westin, C.-F., and Aja-Fernández, S. (2009b), "Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging," *NeuroImage*, 47, 638–650.

Tristán-Vega, A., Westin, C.-F., and Aja-Fernández, S. (2010), "a new methodology for the estimation of fiber populations in the white matter of the brain with the Funk-Radon transform," *NeuroImage*, 49, 1301–1315.

Tschumperlé, D. and Deriche, R. (2003), "Variational Frameworks for DT-MRI Estimation, Regularization and Visualization," in *ICCV*.

Tuch, D., Weisskoff, R., Belliveau, J., and Wedeen, V. (1999), "High angular resolution diffusion imaging of the human brain," in *Proceedings of the 7th Annual Meeting of ISMRM*.

Tuch, D. S. (2002), "Diffusion MRI of Complex Tissue Structure," Ph.D. thesis, MIT.

— (2004), "Q-ball imaging," *Magnetic Resonance in Medicine*, 52, 1358–1372.

Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., and Wedeen, V. J. (2002), "High Angular Resolution Diffusion Imaging Reveals Intravoxel White Matter Fiber Heterogeneity," *Magnetic Resonance in Medicine*, 48, 577–582.

Wedeen, V., Reese, T., Tuch, D., Weigel, M., Dou, J.-G., Weiskoff, R., and Chessler, D. (2000), "Mapping fiber orientation spectra in cerebral white matter with Fourier transform diffusion MRI," in *Proceedings of the International Society of Magnetic Resonance in Medicine*, p. 82.

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.

Wu, Y.-C., Field, A. S., and Alexander, A. L. (2008), "Computation of Diffusion Function Measures in q -Space Using Magnetic Resonance Hybrid Diffusion Imaging," *IEEE Transactions On Medical Imaging*, 27, 858–865.

Yu, T. and Li, P. (2013), "Spatial shrinkage estimation of diffusion tensors on diffusion weighted imaging data," *Journal of the American Statistical Association*, 108, 864–875.

Yu, T., Zhang, C., Alexander, A. L., and Davidson, R. J. (2013), "Local tests for identifying anisotropic diffusion areas in human brain with DTI," *Annals of Applied Statistics*, 7, 201–225.

Zhang, C. H. and Zhang, S. (2014), "Confidence intervals for low-dimensional parameters with high-dimensional data," *Journal of Royal Statistical Society B*, 76, 217–242.

Zou, H. (2006a), "The adaptive lasso and its oracle properties," *Journal of the American statistical association*, 101, 1418–1429.

— (2006b), "The adaptive lasso and its oracle properties," *Journal of the American Statistical Association*, 101, 1418–1429.