

Spatially Regularized Spherical Reconstruction: A Cross-Domain Filtering Approach for HARDI Signals

by

Iván Camilo Salgado Patarroyo

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2013

© Iván Camilo Salgado Patarroyo 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Despite the immense advances of science and medicine in recent years, several aspects regarding the physiology and the anatomy of the human brain are yet to be discovered and understood. A particularly challenging area in the study of human brain anatomy is that of brain connectivity, which describes the intricate means by which different regions of the brain interact with each other. The study of brain connectivity is deeply dependent on understanding the organization of white matter. The latter is predominantly comprised of bundles of myelinated axons, which serve as connecting pathways between approximately 10^{11} neurons in the brain. Consequently, the delineation of fine anatomical details of white matter represents a highly challenging objective, and it is still an active area of research in the fields of neuroimaging and neuroscience, in general.

Recent advances in medical imaging have resulted in a quantum leap in our understanding of brain anatomy and functionality. In particular, the advent of *diffusion magnetic resonance imaging* (dMRI) has provided researchers with a non-invasive means to infer information about the connectivity of the human brain. In a nutshell, dMRI is a set of imaging tools which aim at quantifying the process of water diffusion within the human brain to delineate the complex structural configurations of the white matter. Among the existing tools of dMRI, *high angular resolution diffusion imaging* (HARDI) offers a desirable trade-off between its reconstruction accuracy and practical feasibility. In particular, HARDI excels in its ability to delineate complex directional patterns of the neural pathways throughout the brain, while remaining feasible for many clinical applications.

Unfortunately, HARDI presents a fundamental trade-off between its ability to discriminate crossings of neural fiber tracts (i.e., its angular resolution) and the signal-to-noise ratio (SNR) of its associated images. Consequently, given that the angular resolution is of fundamental importance in the context of dMRI reconstruction, there is a need for effective algorithms for denoising HARDI data. In this regard, the most effective denoising approaches have been observed to be those which exploit both the angular and the spatial-domain regularity of HARDI signals. Accordingly, in this thesis, we propose a formulation of the problem of reconstruction of HARDI signals which incorporates regularization assumptions on both their angular and their spatial domains, while leading to a particularly simple numerical implementation. Experimental evidence suggests that the resulting cross-domain regularization procedure outperforms many other state of the art HARDI denoising methods. Moreover, the proposed implementation of the algorithm supersedes the original reconstruction problem by a sequence of efficient filters which can be executed in parallel, suggesting its computational advantages over alternative implementations.

Acknowledgements

I would like to express my gratitude towards Dr. Oleg Michailovich for his unconditional support, admirable patience, and invaluable guidance over the past three years. I also want to extend my gratitude to my colleagues and friends: Sudipto Dolui, Daniel Otero, Alan Kuurstra, Eric Lum, Quan Zhou, João Battistella, and Jade Deng. Aside from the very insightful discussions we have had in the last few years, they have made my graduate experience extremely enjoyable. Additionally, I want to thank Dr. Edward Vrscay for the priceless advice and extremely kind disposition to collaborate and assist at all times. I also want to express my gratitude towards Dr. Bill Bishop and Dr. Zhou Wang for their feedback and insightful suggestions regarding the contents and the format of this thesis. Finally, I would like to thank my entire family and in particular to Gloria Patarroyo Torres, Maria Torres de Patarroyo, Luis Fernando Salgado, Elizabeth Saldarriaga, Luis Armando Patarroyo, Martha Patarroyo, Jorge Patarroyo, and Fernando Salgado Suárez. Their unconditional support, immense patience, and priceless advice have been indispensable for the completion of my graduate degree as a whole.

Table of Contents

| | |
|--|-----------|
| List of Tables | v |
| List of Figures | vi |
| 1 Introduction | 1 |
| 1.1 Medical Imaging | 2 |
| 1.2 Brain Imaging | 2 |
| 1.2.1 Non-MRI Techniques | 3 |
| 1.2.2 MRI-based Techniques | 4 |
| 1.3 Diffusion Imaging | 4 |
| 1.3.1 dMRI Notations | 6 |
| 1.3.2 DW-MRI | 8 |
| 1.3.3 DTI | 8 |
| 1.3.4 MSDI and DSI | 9 |
| 1.3.5 HARDI | 10 |
| 1.4 Main Contributions | 10 |
| 1.5 Outline | 11 |
| 2 Background | 13 |
| 2.1 Principles of MRI | 13 |
| 2.1.1 The Spin and Alignment of Spin Systems | 14 |

| | | |
|----------|---|-----------|
| 2.1.2 | RF Excitation | 15 |
| 2.1.3 | Precession around a Static Magnetic Field | 15 |
| 2.1.4 | Relaxed Precession around a Static Magnetic Field | 16 |
| 2.1.5 | Magnetic Field Gradients | 17 |
| 2.1.6 | MRI Measurements and Noise | 18 |
| 2.2 | Principles of Diffusion MRI | 20 |
| 2.2.1 | The Bloch-Torrey Equations and Related Expressions | 23 |
| 2.2.2 | Isotropic vs. Anisotropic Diffusion and the Diffusion Signal | 24 |
| 2.2.3 | The Q-space Formalism | 25 |
| 2.2.4 | The Fourier Relationship between E and P | 26 |
| 2.3 | Diffusion Tensor Imaging (DTI) and Its Limitations | 26 |
| 2.4 | HARDI and Q-ball Imaging | 28 |
| 2.4.1 | Q-Ball Imaging and HARDI Contrasts | 30 |
| 2.5 | Filtering Methods for HARDI Data | 30 |
| 2.5.1 | MRI De-noising Methodologies | 31 |
| 2.5.2 | Angular Smoothing | 32 |
| 2.5.3 | Non-Local Means | 33 |
| 2.5.4 | ℓ_1 -Based HARDI Reconstruction | 33 |
| 2.5.5 | Denoising of HARDI Signals using FEM and TV | 35 |
| 2.5.6 | Other Filtering Methods for HARDI | 36 |
| 2.5.7 | Conclusions | 37 |
| 3 | Main Contributions | 38 |
| 3.1 | Notations | 38 |
| 3.2 | Spatial Regularization for HARDI Data | 39 |
| 3.2.1 | TV for Scalar Functions over \mathbb{R}^n and the ROF Model | 41 |
| 3.2.2 | Scalar TV over \mathcal{M} | 42 |
| 3.2.3 | TV Definitions for Vector Valued Functions | 42 |

| | | |
|----------|--|-----------|
| 3.3 | Angular Regularization | 44 |
| 3.4 | Separable Regularization | 47 |
| 3.5 | Numerical Implementation | 48 |
| 4 | Results | 52 |
| 4.1 | In Silico Experiments | 52 |
| 4.2 | Real-Life Experiments | 56 |
| 4.2.1 | The FiberCup Phantom Experiments | 56 |
| 4.2.2 | In Vivo Experiments | 59 |
| 5 | Summary and Conclusions | 67 |
| 6 | Future Work | 69 |
| 6.1 | Spherical Deconvolution | 69 |
| 6.2 | Fiber Continuity | 70 |
| | References | 71 |

List of Tables

| | | |
|-----|---|----|
| 3.1 | Pseudocode for the proposed algorithm (SR^2). | 50 |
| 4.1 | Acronyms, references, and input signal types corresponding to the various methods under comparison. | 54 |
| 4.2 | NMSE produced by the various methods under comparison. | 56 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | dMRI methodologies: trade-off between complexity and practical feasibility. | 7 |
| 1.2 | Sampling schemes used in dMRI. | 7 |
| 2.1 | The effect of magnetic field gradients. | 18 |
| 2.2 | Spatial encoding via magnetic field gradients. | 19 |
| 2.3 | Sample MRI scan and associated k -space measurements. | 21 |
| 2.4 | A sample dephase-rephase diffusion encoding sequence. | 22 |
| 2.5 | Sample pulsed gradient spin echo sequence. | 23 |
| 2.6 | Various DTI contrasts. | 27 |
| 2.7 | The limitations of DTI. | 28 |
| 2.8 | Sample fODFs and their dODFs counterparts. | 29 |
| 3.1 | Three sample diffusion-encoding slices. | 40 |
| 3.2 | A set of sample spherical HARDI signals. | 45 |
| 4.1 | HARDI phantom and its associated ADC and noisy images. | 54 |
| 4.2 | In silico phantom construction. | 55 |
| 4.3 | Sample reconstructions using various algorithms (SNR = 8). | 57 |
| 4.4 | Sample reconstructions using various algorithms (SNR = 16). | 58 |
| 4.5 | The FiberCup phantom | 59 |
| 4.6 | Real-life dODF reconstructions by means of SH2 and of SR ² | 60 |
| 4.7 | In vivo dMRI data set and reconstruction results. | 62 |

| | | |
|------|--|----|
| 4.8 | In vivo reconstruction results (region 1). | 63 |
| 4.9 | In vivo reconstruction results (region 2). | 64 |
| 4.10 | In vivo reconstruction results (region 3). | 65 |
| 4.11 | In vivo reconstruction results (region 4). | 66 |

Chapter 1

Introduction

Diffusion magnetic resonance imaging (dMRI) provides clinicians and researchers with an effective means to study the anatomy of the human brain. dMRI methods infer anatomical information with high levels of precision by quantifying the process of water diffusion within the human brain. Prevalent examples of dMRI clinical tools used in practice include fiber tractography [60, 81], a modelling technique used to recreate the neural tracts within the brain, and diffusion contrasts such as average mean diffusivity and fractional anisotropy maps, both of which reveal relevant information for the accurate diagnosis of several mental disorders [79, 101, 67].

In order to accurately depict the highly complex anatomical features present within the human brain, dMRI techniques have progressively adopted more demanding image acquisition schemes, elaborated modelling assumptions and estimation procedures over time. Presently, among the existing dMRI-based techniques, high angular resolution diffusion imaging (HARDI) offers a very desirable balance between its level of accuracy and its practical feasibility [9]. Unfortunately, just like any other dMRI methodology, HARDI is known to exhibit a trade-off between the directional resolution and the signal-to-noise ratio (SNR) of its related contrasts. As such, high angular resolution in HARDI data is often achieved only at the expense of poorer SNR levels, which suggests the need for effective HARDI de-noising/reconstruction algorithms.

The purpose of this thesis is to introduce a particularly effective de-noising algorithm for HARDI data. The detailed design of the proposed algorithm is motivated by two regularity assumptions about HARDI signals, namely that HARDI signals showcase smooth behaviour across their angular domain, and piece-wise smooth behaviour across their spatial domain. Before formally presenting the design and implementation details of the

proposed algorithm in subsequent sections, this chapter gives some relevant background information which motivates the use of this de-noising method in practice. In particular, an overview of some currently widespread brain imaging modalities is provided, including dMRI, immediately followed by an outline of the main contributions of the thesis in this context. Finally, an outline of the remaining contents of the thesis is presented at the end.

1.1 Medical Imaging

Medical imaging techniques offer the ability to capture information regarding the internal properties of the human body without the need for invasive interventions such as surgery or endoscopy [92]. It is customary to distinguish different imaging techniques by the intrinsic properties they aim to capture. Based on this criterion, imaging techniques are commonly categorized into different *imaging modalities*. For instance, functional magnetic resonance imaging (fMRI) provides information about blood perfusion or blood flow, while positron emission tomography (PET) may provide information about metabolism or receptor binding, making these two imaging techniques belong to different image modalities. On the other hand, diffusion tensor imaging (DTI) and high angular resolution diffusion imaging (HARDI) both belong to the dMRI imaging modality, since they only differ in terms of the signal models they employ and on their requirements for image acquisition, but they aim to describe the same physical properties of water diffusion.

1.2 Brain Imaging

There are several imaging modalities available for the purpose of medical diagnosis. Central to this thesis is the use of these techniques for the purpose of brain imaging or neuroimaging, that is, for imaging certain characteristics of the human brain. A rich set of imaging methodologies has been developed for capturing either functional information (e.g., cerebral perfusion) or structural information (e.g., proton density in tissue, geometrical features of tissue, etc.) about the human brain, which we refer hereafter to as functional and structural brain imaging techniques, respectively. Functional brain imaging differs from its structural counterpart since it allows for the study of processes and characteristics related not only to the structural features of the brain, but also to the subject's specific brain activity during the execution of particular activities or exposure to stimuli. These imaging techniques are often carried out on awake human subjects who are engaged in sensory, motor or cognitive tasks. The use of these techniques has opened a range of possibilities to study, for example,

human cognition [56]. Structural and functional imaging studies can be performed both in normal and in psychiatric patients, thereby offering the ability to study the anatomical and functional abnormalities associated with several mental disorders [56, 61].

1.2.1 Non-MRI Techniques

Among the existing structural brain imaging techniques which do not employ MR imagery, X-ray computed tomography (X-ray CT) is possibly the most prevalent in practice. The advent of X-ray CT has provided clinicians with invaluable scanning and diagnostic tools due to its associated high spatial resolution and contrast levels as well short acquisition times. The applications of CT are not limited to the study of the brain, but also to various clinical applications including pulmonary, abdominal, and cardiac imaging [57].

X-ray CT allows one to obtain 3-D images of internal organs of the human body from 2-D image slices of the same, and it relies on the principle of projective radiography to do so. Even though the machinery and processing tools used in X-ray CT have been improved significantly over the past years, X-ray exposure remains a drawback of this imaging modality, due to the adverse health effects that this procedure may incur in human subjects [75].

Non-MRI functional brain imaging techniques, on the other hand, include positron emission tomography (PET), single-photon emission computed tomography (SPECT), electroencephalography (EEG), and magnetoencephalography (MEG). In PET and SPECT, images are generated by the emission of gamma rays from the human body, which takes place after introducing a positron-emitting radionuclide, called the radiotracer, into the subject's body. The radiotracer reacts selectively based on brain function and thus allows radiologists to study functional characteristics of this particular organ [92]. Electric-magnetic neuroimaging techniques such as EEG and MEG, on the other hand, aim at measuring the electromagnetic fields associated with neural activity [56]. As opposed to PET and SPECT, these methods do not require the use of radioactive substances. Moreover, they offer superior temporal resolution (in the range of milliseconds) than PET, SPECT, and functional MRI (fMRI) (in the range of one or more seconds), suggesting their superiority for detection of neuronal activity. Despite the superior temporal resolution of these methods, however, they exhibit a fundamental drawback arising from the inherent difficulty to identify the locations of sources responsible for the sensed electromagnetic data within the human brain [56]. As a result, the practical value of these techniques is heavily dependent on the size of the region of interest; allowing for the accurate study of localized regions of the human brain, while preventing one to study complex processing tasks involving networks of interacting brain regions.

1.2.2 MRI-based Techniques

MRI techniques offer high spatial resolution levels, a high degree of image contrast, and are non-invasive in nature [92]. Altogether, these desirable characteristics of MRI have positioned it as the prominent brain imaging technique in both the functional and structural domains [56, 92]. MRI relies on the principle of nuclear magnetic resonance (NMR), whereby the imaging mechanism can be made sensitive to the chemical composition of the tissue under examination. In particular, MRI exploits the abundance of water molecules in human tissue to provide images with relevant diagnostic information and fine scale visualizations of anatomical details.

There are three main standard MRI contrasts which are named according to the underlying physical quantity they aim to describe [92]. Thus, for example, the intensities on T_1 -weighted and T_2 -weighted MRI contrasts correspond to the so-called T_1 and T_2 relaxation times, respectively. These quantities are measures describing the *precession* of the magnetic moments associated with hydrogen atoms under the presence of an external magnetic field. On the other hand, as its name suggests, the intensities from proton density (PD) contrasts correspond to hydrogen atom density values. Based on these contrasts, it is possible to study the anatomical and the functional aspects of the human brain, which represent the main focus of the structural MRI (sMRI) and functional MRI (fMRI) imaging modalities, respectively

1.3 Diffusion Imaging

A particular characteristic of standard MRI contrasts is that they are solely sensitive to the chemical composition of the examined tissue. Thus, regions comprised of chemically homogeneous regions of human tissue are difficult to discern based on standard MRI contrasts, even when their anatomy features are entirely different [77]. This is the fundamental limitation of standard MRI contrasts which makes them inadequate for the study of the white matter within the human brain, which is chemically homogeneous, but whose underlying anatomy is highly complex and not very well understood. White matter in the human brain is inherently complex, being comprised of several *neural fiber tracts*, which are bundles of coherently oriented axons serving as the connection pathways between the approximately 10^{11} neurons within the human brain.

The key to sensitize MRI imagery to the white matter anatomical features lies in the physical process known as *diffusion*. Diffusion is a mass transport mechanism whereby particles displace over time within a medium according to the change in their concentration

across the medium. At a microscopic level, diffusion is described by a partial differential equation (PDE) known as Fick’s second law [33]. In a sufficiently general case for the purpose of our discussion, letting $\phi(\mathbf{x}, t) \in \mathbb{R}^+$ denote the concentration of the particles at spatial location $\mathbf{x} \in \mathbb{R}^3$ and time $t \in \mathbb{R}^+$, a general version of Fick’s second law can be expressed as

$$\frac{\partial \phi}{\partial t}(\mathbf{x}, t) = \operatorname{div}(D(\phi, \mathbf{x}) \nabla \phi(\mathbf{x}, t)), \quad (1.1)$$

where $D(\phi, \mathbf{x})$ is called the diffusion coefficient for ϕ at \mathbf{x} , while $\operatorname{div}(\cdot)$ and $\nabla(\cdot)$ are the standard divergence and the gradient operators over the \mathbf{x} coordinate, respectively. Conceptually, equation (1.1) dictates that particles tend to move from regions of low concentration to those of high concentration at a rate depending on D and on the degree of variation of the concentration around that location. It is known that D depends on the physical properties of the particles and of the medium, to the extent that knowledge of this function allows one to probe the physical properties of the medium in which the particles are present, including its geometry.

A different quantity which is conceptually related to D is the so-called *ensemble average propagator* (EAP) function, P . The EAP, commonly denoted as $P : \Omega_{\mathbf{x}} \times \Omega_{\mathbf{r}} \subset \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^+$ is a function which assigns to each spatial location \mathbf{x}_0 and each displacement vector \mathbf{r} , both in three dimensional space, the probability $P(\mathbf{r} \mid \mathbf{x}_0; \tau) d\mathbf{r}$ (often denoted simply as $P(\mathbf{r} \mid \mathbf{x}_0) d\mathbf{r}$) that a water molecule initially located at \mathbf{x}_0 moves to a different predefined position $\mathbf{x}_0 + d\mathbf{r}$ in a fixed amount of time τ . Not surprisingly, like D , the EAP contains essential information for the reconstruction of underlying neural fiber anatomy [61]. Diffusion MRI (dMRI) techniques exploit the tight connection between the anatomical features of white matter and quantities such as D or P . Particularly, these techniques infer the complex anatomical details of white matter via estimates of related quantities to D and P obtained under a wide range of modelling assumptions.

The origins of diffusion imaging date back to 1950 when the spin echoes discovered by Hahn were sought to sense diffusion processes [54]. Shortly after, in 1954, Carr and Purcell developed a method for quantifying water diffusion using NMR measurements [24]. These achievements were followed by the discoveries of Stejskal and Tanner in 1965, who were among the first to estimate the diffusion coefficient of water particles using NMR and a pulsed magnetic field gradient system [103]. These discoveries, combined with the fact that many regions of the white matter are comprised of fiber bundles that are larger than the image resolution of MRI (typically about 2 mm along the three spatial coordinates), suggested that anatomical details of these fibers could be captured using MRI techniques [77].

1.3.1 dMRI Notations

In dMRI, MR scans are sensitized to water diffusion through the use of diffusion-encoding gradients, which are spatially varying magnetic fields whose intensities change linearly as a function of spatial location. These sensitized MRI scans are commonly referred to as diffusion-encoded images, and are parametrized by their acquisition settings, including the parameters of the diffusion-encoding gradients used (e.g., their strength and direction). As a result, it is convenient to distinguish diffusion-encoded images based on their associated \mathbf{q} -value, where $\mathbf{q} \in \mathbb{R}^3$ is a quantity which encapsulates all the determining acquisition parameters at once [23]. Based on this concept, one can introduce a coordinate system known as the \mathbf{q} -space to characterize dMRI images. Mathematically, the \mathbf{q} -space formalism models a diffusion-encoded image E as

$$E : \Omega_{\mathbf{x}} \times \Omega_{\mathbf{q}} \rightarrow \mathbb{R}^+, \quad \Omega_{\mathbf{x}}, \Omega_{\mathbf{q}} \subset \mathbb{R}^3, \quad (1.2)$$

where $\mathbb{R}^+ := \{r \in \mathbb{R} \mid r \geq 0\}$, $\Omega_{\mathbf{x}}$ represents a region of interest in the 3 dimensional space (aka the spatial domain of E), and $\Omega_{\mathbf{q}}$ is the \mathbf{q} -space. Under this notation, each diffusion-encoded image is of the form $E(\cdot, \mathbf{q}) : \Omega_{\mathbf{x}} \rightarrow \mathbb{R}^+$, and one refers to $E(\cdot, \mathbf{v})$ as the diffusion-encoded image corresponding to \mathbf{q} -value of $\mathbf{v} \in \Omega_{\mathbf{q}}$. A fundamental result in dMRI is that the \mathbf{q} -space data is conveniently related to the EAP via the three dimensional Fourier transform [23], so that $P(\mathbf{r}|\mathbf{x}) = \mathcal{F}_{3D,\mathbf{q}}\{E\}(\mathbf{x}, \mathbf{r})$, where $\mathcal{F}_{3D,\mathbf{q}}$ denotes the 3 dimensional Fourier transform with respect to the \mathbf{q} coordinate. As a result of the Fourier relationship between P and E , the variables $\mathbf{r} \in \Omega_{\mathbf{r}}$ and $\mathbf{q} \in \Omega_{\mathbf{q}}$ are commonly referred to as the displacement vector and the wave vector, respectively.

Various methods of dMRI differ in terms of their underlying diffusion models (e.g., assumptions imposed on the EAP or on D), and in terms of the sampling scheme employed to sample the \mathbf{q} -space (see Figure 1.2). The principal dMRI techniques are diffusion weighted imaging (DW-MRI), diffusion tensor imaging (DTI), high angular resolution diffusion imaging (HARDI), multi shell diffusion imaging (MSDI), and diffusion spectrum imaging (DSI). A fundamental trade-off showcased by all these techniques, shown in Figure 1.1, is between their reconstruction accuracy and their acquisition requirements, the latter of which dictates their feasibility in practical settings. We now provide a brief review of the most prevalent dMRI techniques found in practice. For a more comprehensive review of dMRI techniques, the reader is referred to [9].

1.3.2 DW-MRI

The term diffusion weighted MRI (DW-MRI) is used to refer to the earliest approaches for generating contrasts using dMRI information. DW-MRI consists of measuring the degree of water diffusion along three mutually orthogonal directions, that is, sampling the \mathbf{q} – space at only three mutually orthogonal \mathbf{q} –values $\{\mathbf{q}_x, \mathbf{q}_y, \mathbf{q}_z\}$ [101]. The images thus obtained, along with a standard T_2 -weighted image, are combined together to create a single contrast used to infer information about the water diffusion process within the brain.

There are two common ways to combine the acquired images. On the one hand, one may normalize all the diffusion-encoded images by the standard T_2 scan (thereby isolating the effects of water diffusion present in the signals), and subsequently calculate the geometric average of the normalized diffusion-encoded images [101]. Alternatively, one can estimate the trace of the so-called apparent diffusion coefficient (ADC) matrix at each spatial location $\mathbf{x} \in \mathbb{R}^3$, $ADC(\mathbf{x}) \in \mathbb{R}^{3 \times 3}$, the entries of which are closely related to the diffusion coefficient D . The contrast thus obtained provides a measure of the degree of anisotropy of water diffusion in the human brain at each spatial coordinate in the image [101].

The use of DW-MRI contrasts has allowed clinicians to obtain relevant information about several cerebral diseases including acute ischemic stroke, neoplasms, intracranial infections, and traumatic brain injury, among others [101]. Additionally, DW-MRI contrasts have been successfully used to discern pathologies which are simply indistinguishable by means of standard MRI scans, such as acute stroke and other processes that manifest with sudden neurologic deficits [101]. Unfortunately, in several cases, limiting water diffusion measurements to three directions results in insufficient information for the accurate reconstruction of the anatomy of white matter, motivating the need for alternative acquisition schemes or simplifying models of water diffusion.

1.3.3 DTI

Diffusion tensor imaging (DTI) differs from DW-MRI in that it allows one to infer information about water diffusion in any possible direction, as opposed to only three. To attain this level of informativeness, DTI uses a parametric approach, whereby the EAP assumed to be a Gaussian probability density function in three dimensions. This assumption in turn entails that the EAP is fully characterized by a 3×3 symmetric positive definite matrix, known as the diffusion tensor, which plays the role of the covariance matrix associated with the EAP. Hence, DTI only requires a minimum of seven MRI scans to estimate the EAP [14, 61]: six diffusion-encoded scans obtained using different diffusion-encoding gradients, and a standard T_2 -weighted image for normalization. It is important to note that,

in practice, 20 to 30 diffusion-encoded images are usually obtained in DTI, as opposed to the minimum of 7, in order to improve the robustness and the accuracy of DTI estimation methods. Moreover, the diffusion-encoded images are commonly obtained by sampling the \mathbf{q} – space over a spherical shell, as illustrated in Figure 1.2.

DTI is currently the most widespread diffusion MRI technique used in practice due to the usefulness of its associated contrasts. These contrasts are based on scalar features such as *relative anisotropy* (RA), *volume ratio* (VR), and *fractional anisotropy* (FA) [79], or on vectorial features based on the geometrical characteristics of the ellipsoid corresponding to the diffusion tensor (see Section 2.3 for more information). DTI contrasts have played a determining role in the study and diagnosis of numerous brain-related diseases and abnormalities including Alzheimer’s disease [28], ischemic stroke [30], epilepsy [98], and schizophrenia [47], to name a few (see [105] for a more comprehensive review).

Despite the success of DTI in clinical settings, however, the underlying model of DTI showcases fundamental limitations which render its contrasts inaccurate in the context of diffusion imaging of the brain. In particular, DTI reconstructions are unable to depict accurately complex water diffusion profiles present in white matter, including those corresponding to neural fiber-crossings, fiber branching, and disk-shape anisotropy [77].

1.3.4 MSDI and DSI

To overcome the accuracy limitations of DTI, the dMRI methodologies known as multi-shell diffusion imaging (MSDI) [8] and diffusion spectrum imaging (DSI) [120] abandon the Gaussian parametric model of the EAP which is unable describe multi-modal diffusion profiles [61]. In the case of MSDI and DSI, the acquisition of samples in the \mathbf{q} – space is not restricted to a single spherical shell, as it is in the case of DTI and HARDI. On the one hand, MSDI samples the \mathbf{q} – space over multiple spherical shells, as illustrated in Figure 1.2. The result of these modifications to the sampling scheme is that both angular and *radial* diffusion information is captured using this methodology. On the other hand, the sampling scheme in DSI usually consists of sampling the \mathbf{q} – space over a rectangular grid as opposed to as a spherical shell . The number of \mathbf{q} – space samples acquired in DSI (typically around 500), is determined by the desired estimation accuracy of the EAP via direct application of the fast Fourier transform to the acquired diffusion signal E . Not surprisingly, DSI reconstructions offer a higher degree of accuracy as compared to those of the alternative dMRI methodologies, and it provides a model-free EAP estimation procedure which makes it invulnerable to model errors. Unfortunately, both MSDI and DSI utilize sampling schemes which are often impractical in clinical settings (e.g., [120]), and they are therefore rarely used in practice.

1.3.5 HARDI

High angular resolution diffusion imaging (HARDI), first introduced in [113], is a dMRI technique which differs from DTI in terms of the number of samples which are obtained in the \mathbf{q} – space. HARDI-based methods typically require about 60 to 80 samples in the \mathbf{q} – space, enough information to abandon the oversimplifying assumptions/models of DW-MRI and DTI, while being more feasible in practice than MSDI and DSI [9]. As suggested by Figure 1.1, HARDI offers a very desirable trade-off between practical feasibility and accuracy.

As a result of the increased flexibility of HARDI, a variety of parametric and non-parametric estimation procedures have been proposed based on this dMRI technique, including mixture models [112], spherical deconvolution [110], and Q-ball imaging (QBI) [114]. Among the numerous HARDI estimation procedures, possibly the most widespread is that of QBI. In QBI, the \mathbf{q} – space data is used to estimate ψ , the diffusion orientation distribution function (dODF) [114, 61], which is related to the EAP via

$$\psi(\mathbf{u}) = \int_0^\infty P(\alpha\mathbf{u}) d\alpha, \quad \forall \mathbf{u} \in \mathbb{S}^2, \quad (1.3)$$

so that the dODF evaluated at a particular direction $\mathbf{u} \in \mathbb{S}^2$ is equivalent to a radial projection of EAP in the direction of \mathbf{u} . As a result, one may interpret the dODF as a quantity which captures solely the angular dependencies of the EAP. In this way, the dODF describes the likelihood of a spin to displace in any possible direction.

As opposed to MSDI and DSI based approaches which are able to describe both the angular and the radial dependency of the EAP, most HARDI estimation techniques can only estimate the angular behaviour of this quantity, thereby achieving less accurate and complete reconstructions of the diffusion signal. Nevertheless, the increased image acquisition times associated with the former methods commonly outweigh their benefits over HARDI in practical settings. This fact, along with the improved accuracy of HARDI over DW-MRI and DTI, position HARDI as the best compromise in terms of usefulness and feasibility among the alternative dMRI methods.

1.4 Main Contributions

The scope of this thesis is restricted to reconstruction problems associated with HARDI. In particular, the problem of HARDI de-noising is addressed herein. It is known that

HARDI, just like in any other diffusion MRI technique, exhibits a particular trade-off between angular resolution and signal-to-noise ratio (SNR) which arises from the physical phenomena governing diffusion imaging acquisition [62]. Consequently, provided that high angular resolution is an essential requirement for HARDI reconstructions, it is necessary to deal with distortions in the acquired data which arise as a result of the presence of noise. Such distortions often obscure features which are highly relevant for the purpose of clinical diagnosis and accurate anatomical descriptions of the tissue under examination [62], and therefore jeopardize the overall usefulness of HARDI in practice. As a result, there is a need for effective, robust, and efficient tools for the purpose of HARDI data de-noising.

Existing HARDI de-noising techniques can be broadly categorized into three main classes based on the type(s) of regularity they impose on the underlying signals. In the context of HARDI data, regularity may be imposed over the spatial and/or the angular domain of the signals, and therefore methods can be divided into those which take into account 1) spatial regularity only, 2) angular regularity only, and 3) both spatial and angular regularity. It is expected that the methods belonging to the third class are the most effective at eliminating noise in general, and one of the main objectives of this thesis is to examine the validity of this hypothesis. The second main objective of this thesis is to study different types of regularization which can be employed for the purpose of HARDI de-noising. In doing so, a particular de-noising strategy of the third class is introduced and its performance is compared with that of other methods using both simulated and real-life data. The proposed algorithm advocates the use of the regularized least-squares spherical harmonic fit of [36] for the purpose of angular regularization, in conjunction with a total variation (TV) regularization model for the purpose of spatial regularization. The algorithm is compared to various alternatives based on spherical smoothing or on different TV regularization priors, including an adaptation of the standard TV semi-norm to functions defined over the direct product space $\mathbb{R}^3 \times \mathbb{S}^2$. Finally, the mathematical foundations of the proposed algorithm are presented along with relevant practical considerations. Although not rigorously shown, it is also suggested that the newly introduced method offers computational advantages based on the fact that it can be broken down into independent subroutines which may be executed in a parallel fashion.

1.5 Outline

The remaining portion of this thesis is organized as follows: Chapter 2 outlines the existing relevant literature and other background information which is pertinent for subsequent discussions. Chapter 3 presents in detail the main contribution of this thesis, that is,

the proposed HARDI de-noising algorithm known as spatially regularized spherical reconstruction (SR^2). The motivation behind this method, its mathematical formulation, and its implementation details are presented in this chapter. Additionally, some theoretical and practical aspects related to the Rudin-Osher-Fatemi (ROF) model for functions defined over the direct product space $\mathbb{R}^3 \times \mathbb{S}^2$ is presented in this chapter. Chapter 4 presents both simulation and in vivo experimental studies of the performance of the proposed and reference algorithms. Finally, Chapter 5 presents a short summary of the main contributions from the previous chapters, while Chapter 6 presents a closing discussion of relevant directions for future work in the context of HARDI signal de-noising and other related estimation problems.

Chapter 2

Background

Water accounts for roughly 60% of the overall chemical composition of the human body [52]. Whereas conventional MRI contrasts make use of this fact to produce informative scans of the chemical composition of human tissue, they still face the limitation of being insensitive to the micro-structure of white matter [61, 67, 77, 115]. One of the main attributes of dMRI which differentiates it from basic MRI contrasts is its ability to quantify the diffusion of water in the human brain, which in turn allows one to infer the geometrical characteristics of neural fiber tracts present in white matter. The practical relevance of dMRI is exemplified by its applications to the study and diagnosis of several brain-related disorders [61, 107, 85, 70, 77]. Additionally, dMRI has revealed new frontiers in the study of brain connectivity as a whole - a current active area of research which is believed to be crucial for the proper understanding of the function of the human brain [81, 60, 77].

This section presents some principles of MRI and dMRI relevant to future discussions, including an overview of the dMRI methods which are prevalent in current practical and research settings. Following this discussion, a brief description of several existing filtering methods for HARDI data is provided.

2.1 Principles of MRI

There are three conventional MRI contrasts which are used for different medical applications and studies, namely T_1 -weighted, T_2 -weighted, and proton density (PD) contrasts. The understanding of these contrasts is fundamental for the development of diffusion MRI theory.

The basic principles behind MRI rely on the concept of the spin of an atom nucleus. In particular, all conventional MRI contrasts can be understood by studying the following three fundamental mechanisms: 1) the alignment of the spins within a static magnetic field in steady state, 2) the change in orientation of the spins using radio frequency (RF) signals (RF excitation), and 3) the relaxed (attenuated) precession of the spins around a static magnetic field in transient state. Essentially, conventional MRI contrasts are visual representations of the parameters associated with these mechanisms.

2.1.1 The Spin and Alignment of Spin Systems

The spin is a quantum mechanics characteristic of fundamental particles which shares many similarities with the concept of angular momentum in classical mechanics [23, 92]. In the context of nuclear magnetic resonance (NMR), the spin of a hydrogen nucleus can be conceived as the rotation of a proton around an axis. It is known that any nucleus possessing a spin has an associated microscopic magnetic field - a phenomenon that can be rigorously derived from the principles of quantum physics, but which is often explained by means of classical mechanics arguments for simplicity. Specifically, since protons are rotating positively charged particles, it is expected that their rotation induces a microscopic magnetic field by Ampere's law [92]. As a result, each proton has an associated microscopic magnetic field which in turn has a magnetic moment vector $\boldsymbol{\mu} \in \mathbb{R}^3$ given by:

$$\boldsymbol{\mu} = \gamma \boldsymbol{\Phi}, \tag{2.1}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^3$ is the angular momentum of the proton, and $\gamma = 2.675 \times 10^8$ rad/T/s is the *gyromagnetic ratio* of hydrogen nuclei.

When the hydrogen atoms in a sample are not exposed to an extraneous magnetic field, there is no preferred orientation for their magnetic moment vectors. Therefore, at the macroscopic level, the individual contributions of the many $\boldsymbol{\mu}$'s to the total magnetic field cancel each other out and the resulting spin system has no apparent macroscopic magnetic field. Under the presence of an external magnetic field, however, the principle of *nuclear magnetism* implies that individual spins in a spin system will tend to align with the applied magnetic field, resulting in a non-zero macroscopic magnetic field.

Consider a spin system comprised of N hydrogen nuclei with associated $\{\boldsymbol{\mu}_n\}_{n=1}^N \subset \mathbb{R}^3$ magnetization vectors, and subjected to an external magnetic field $\mathbf{B}_0 = B_0 \hat{z}$, where $\hat{z} = (0, 0, 1)$ in Cartesian coordinates. Because of \mathbf{B}_0 , the spin system will showcase a non-zero macroscopic magnetic magnetization vector $\mathbf{M} = \sum_{n=1}^N \boldsymbol{\mu}_n$. If the spin system is subjected to the static magnetic field \mathbf{B}_0 indefinitely without any other disturbances applied

to it, \mathbf{M} will eventually *align* with \mathbf{B}_0 and reach the equilibrium value $\mathbf{M} = \mathbf{M}_0 = M_0\hat{z}$, where

$$M_0 = \frac{B_0\gamma^2\hbar^2P_D}{4kT}, \quad (2.2)$$

where \hbar is the reduced Planck's constant, k is the Boltzmann's constant, T is the temperature in degrees Kelvin, and P_D is the number of targeted nuclei per unit volume. P_D is commonly referred to as the proton density of the sample, and it constitutes the underlying quantity of interest in P_D MRI contrasts [92].

2.1.2 RF Excitation

As its name suggests, a radio frequency (RF) pulse is an oscillating magnetic field of finite duration which operates at radio frequencies. Consider the scenario presented in the previous subsection, whereby \mathbf{M} arises after a static magnetic field is applied in the \hat{z} direction until \mathbf{M} becomes $\mathbf{M}_0 = M_0\hat{z}$. At any point in time, it is possible to manipulate the orientation of \mathbf{M} temporarily by applying an RF pulse with frequency $\omega_0 = \gamma|\mathbf{B}_0|$. This process, known as RF excitation, is of fundamental importance for obtaining MRI contrasts, as will soon be seen.

RF pulses are named according to the change in spin orientation they cause, quantified using the so-called flip angle, which is the angle difference between the orientation of the spins prior and after the RF pulse is applied. As such, an RF pulse which changes the main orientation of a spin system by α radians has a flip angle of α rad and is called an α RF pulse.

RF excitation is indispensable for the generation of MRI scans. If \mathbf{B} is present after the application of an RF pulse, the spins with modified orientation will undergo relaxed *precession* around the static magnetic field \mathbf{B} , a mechanism which will be explained in the next subsection. The parameters of this precession, among which are the so-called T_1 and T_2 relaxation times, are the basis for T_1 and T_2 weighted MRI contrasts, respectively.

2.1.3 Precession around a Static Magnetic Field

Suppose that a static magnetic field $\mathbf{B}_0 = B_0\hat{z}$ is applied to a spin system of interest until $\mathbf{M} = M_0\hat{z}$. Additionally, suppose that α RF pulse is applied at time $t = 0$, so that $\mathbf{M}(0^+)$ is oriented at an angle α away from the \hat{z} plane. Under these circumstances, \mathbf{M} is known to undergo precession around \mathbf{B}_0 . To see why this is the case, we first think of \mathbf{M} as a function of time and space, i.e., $\mathbf{M} = \mathbf{M}(\mathbf{x}, t)$, where $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ is the

spatial coordinate, and $t \in \mathbb{R}^+$ is the time coordinate. We then note that, analogous to the torque exerted in a loop of current by an appropriately oriented external magnetic field, \mathbf{M} experiences a torque $\boldsymbol{\tau}$ under the presence of \mathbf{B}_0 . This torque is formally given by

$$\boldsymbol{\tau}(\mathbf{x}, t) = \frac{\partial \Phi(\mathbf{x}, t)}{\partial t} = \mathbf{M}(\mathbf{x}, t) \times \mathbf{B}_0, \quad (2.3)$$

where $\Phi = \frac{1}{\gamma} \sum_{n=1}^N \boldsymbol{\mu}_n = \frac{\mathbf{M}}{\gamma}$, so that

$$\frac{\partial \mathbf{M}(\mathbf{x}, t)}{\partial t} = \gamma \mathbf{M}(\mathbf{x}, t) \times \mathbf{B}_0, \quad (2.4)$$

whose solution is [92]

$$M_x(t) = M_0 \sin(\alpha) \cos(\gamma B_0 t + \phi), \quad (2.5)$$

$$M_y(t) = M_0 \sin(\alpha) \sin(\gamma B_0 t + \phi), \quad (2.6)$$

$$M_z(t) = M_0 \cos(\alpha), \quad (2.7)$$

where $M_0 = |\mathbf{M}(0)|$, $\phi \in \mathbb{R}$ is an arbitrary angle, and $\mathbf{M}(t) = (M_x(t), M_y(t), M_z(t))$ for all values of t .

Equations (2.5)-(2.7) describe the phenomenon of *precession* of \mathbf{M} about the applied magnetic field \mathbf{B}_0 at frequency ω_0 , known as the *Larmor frequency* and given by

$$\omega_0 = \gamma |\mathbf{B}_0|. \quad (2.8)$$

This process is analogous to the precession of a toy top due to gravity.

2.1.4 Relaxed Precession around a Static Magnetic Field

It is critical at this point to evaluate some of the implicit assumptions made in deriving equations (2.5)-(2.7), as they suggest that the precession of \mathbf{M} takes place for an arbitrary amount of time, which is not physically possible. In reality, precession is hindered by different natural phenomena caused by spin-spin and spin-lattice interactions [23], a process generally known as *spin relaxation*. As a result, the precession mechanism is temporary in reality, and as expected, the vector \mathbf{M} aligns with \mathbf{B}_0 in steady state.

Spin relaxation is mainly attributed to the so-called T_1 and T_2 relaxation mechanisms. These mechanisms represent the underlying quantities in T_1 -weighted and T_2 -weighted conventional MRI contrasts. In order to understand the origin of these mechanisms, it

is useful to decompose \mathbf{M} into its projections onto the $x - y$ plane and onto the z axis, giving rise to the quantities $\mathbf{M}_\perp := M_x\hat{x} + M_y\hat{y}$, and $\mathbf{M}_z = M_z\hat{z}$, often referred to as the *transverse* and the *longitudinal* components of \mathbf{M} , respectively.

On the one hand, T_1 relaxation refers to the mechanism by which the magnitude of the longitudinal component of \mathbf{M} rises from its initial value $M_z(0)$ to its final value $M(0)$ according to

$$M_z(t) = M_z(0)e^{-t/T_1} + M(0)(1 - e^{-t/T_1}), \quad (2.9)$$

where $M = \|\mathbf{M}\|_2 = \sqrt{M_x^2 + M_y^2 + M_z^2}$. This type of relaxation is attributed principally to the so-called spin-lattice relaxation, which involves an exchange of thermal energy between the thermal surrounding reservoir or “lattice” and the spin system [23]. At room temperature, T_1 attains values of about 0.1 s to 10 s in protons of dielectric materials.

Even though T_1 relaxation does have an impact in the evolution of \mathbf{M}_\perp also, it is mainly the interactions between different spins within the same spin system which describe the changes of this quantity over time. The so-called T_2 , transverse, or spin-spin relaxation phenomenon resulting from de-phasing effects among spins is the main mechanism responsible for this process. This type of relaxation is commonly represented mathematically by

$$M_x(t) = e^{-t/T_2}[M_x(0)\cos(\omega_0 t) + M_y(0)\sin(\omega_0 t)] \quad (2.10)$$

$$M_y(t) = e^{-t/T_2}[M_x(0)\sin(\omega_0 t) - M_y(0)\cos(\omega_0 t)]. \quad (2.11)$$

T_2 relaxation times are usually much smaller than T_1 , and typically fall between 10 μs and 10 s. Due to the faster decay rates associated with T_2 relaxation, and therefore faster acquisition times, mainly T_2 -weighted contrasts are employed in dMRI.

Equations (2.9)-(2.11) fully describe the relaxed precession of spins around a static magnetic field, and can be shown to be the solution to the so-called Bloch equations, which can be compactly expressed as:

$$\frac{\partial \mathbf{M}}{\partial t} = \gamma(\mathbf{M} \times \mathbf{B}_0) + \frac{1}{T_1}(M_0 - M_z)\hat{z} - \frac{1}{T_2}\mathbf{M}_\perp. \quad (2.12)$$

2.1.5 Magnetic Field Gradients

As mentioned earlier, the application of a constant magnetic field \mathbf{B}_0 (aka the main magnetic field), causes the spins of water nuclei to precess at the Larmor frequency given by (2.8). This result suggests the possibility to encode the spatial location of different

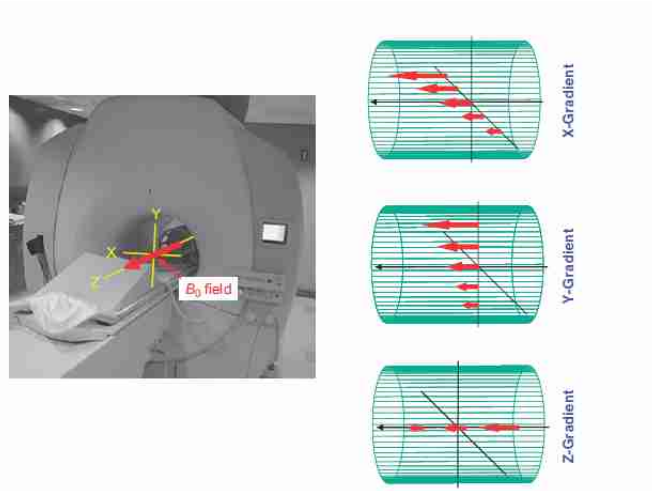


Figure 2.1: The main magnet in an MRI scanner and the effects of magnetic field gradients. The unaltered main magnetic field, \mathbf{B}_0 , is shown in the photograph in red. Also, depictions of sample linear inhomogeneities applied using magnetic field gradients are shown in the subfigures on the right. Credit: [77].

spins based on their Larmor frequencies under the presence of a spatially-varying magnetic field. In practice, such spatially-varying magnetic fields are generated by means of pulsed magnetic field gradients or simply gradients, which introduce different degrees of linear magnetic field inhomogeneities in the main magnetic field \mathbf{B}_0 . This fundamental principle lies at the heart of dMRI methodologies, as will be discussed in Section 2.2. Sample magnetic gradients are illustrated in Figure 2.1, and the use of gradients for the purpose of spatial encoding of water molecules is illustrated in Figure 2.2.

2.1.6 MRI Measurements and Noise

The acquisition of MRI scans entails several principles from physics and signal processing which are out of the scope of this discussion (see e.g., [92] for a detailed discussion of these principles). However, it is pertinent to address specific aspects of MRI acquisition to understand the statistical nature of diffusion-encoded images.

According to Faraday’s law, the time-changing magnetization vector \mathbf{M} should induce a voltage in a coil in its proximity, provided that the coil is not lying on the same plane as \mathbf{M} . In MRI, two receiving coils are employed for this purpose, both of which are sensitive only to changes occurring in \mathbf{M}_\perp .

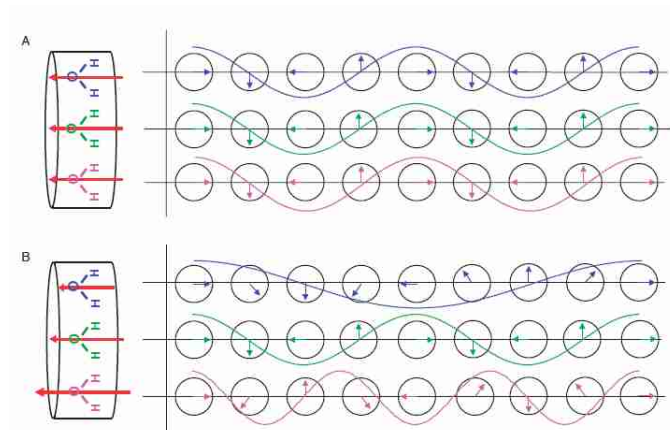


Figure 2.2: Encoding of spins using magnetic field gradients. A. A spatially-invariant magnetic field is applied to the spin system, in which case all the spins possess the same Larmor frequency and no spatial encoding is present. B. A spatially varying magnetic field is introduced by means of a magnetic field gradient so that spins at different spatial locations possess different Larmor frequencies. After removing the magnetic field gradient, the spins at different locations possess the same Larmor frequency but may possess different phases. Credit: [77].

After taking multiple physical factors into account and making multiple simplifications (see [53] for more details), it turns out that the induced voltage in the coils, \mathbf{V} , is related to \mathbf{M}_\perp via the 2D Fourier transform. In complex notation,

$$\mathbf{V}(k_x, k_y, t) \propto \int |\mathbf{M}_\perp(x, y, t)| e^{-i2\pi(xk_x + yk_y)} dx dy \quad (2.13)$$

$$= \mathcal{F}_{2D}\{|\mathbf{M}_\perp|\}(k_x, k_y), \quad (2.14)$$

which suggests that $|\mathbf{M}_\perp|$ may be estimated from samples of the induced voltage via the inverse fast Fourier transform. The domain of definition of \mathbf{V} is commonly referred in the literature as the k -space, and to obtain the samples of \mathbf{V} (i.e., to traverse the k -space), MRI acquisition protocols employ field gradients along the x and y directions [92].

In view of the above discussion, one expects the inverse Fourier transform of \mathbf{V} to be a real-valued image. In practice, however, due to additive white Gaussian noise (AWGN) contamination in the k -space and imperfections in the signal demodulation process, one has

$$\mathcal{F}_{2D}^{-1}\{\mathbf{V}\} = |\mathbf{M}_\perp| e^{i\alpha} + N, \quad \text{with } \alpha \in [0, 2\pi), N = N_r + iN_i, \quad (2.15)$$

where $N_r, N_i \sim \mathcal{N}(0, I\sigma^2)$ are two identically distributed, independent random images of the same size as $|\mathbf{M}_\perp|$, and α is an arbitrary phase shift. As a result, the j^{th} pixel of the output image $S := |\mathcal{F}_{2D}^{-1}(\mathbf{V})|$ is given by

$$S(j) = \sqrt{[N_r(j) + |\mathbf{M}_\perp(j)| \cos(\alpha)]^2 + [N_i(j) + |\mathbf{M}_\perp(j)| \sin(\alpha)]^2}, \quad (2.16)$$

As such, every pixel $S(j)$ of the output image S can be considered as a random variable obeying a Rician distribution with parameters $A = |\mathbf{M}_\perp(j)|$ and σ , with associated probability density function (PDF) given by

$$p_{S(j)|A}(s|a) = \begin{cases} \frac{s}{\sigma^2} \exp\left\{-\frac{a^2 + s^2}{2\sigma^2}\right\} I_0\left(\frac{as}{\sigma^2}\right), & s \geq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (2.17)$$

where $I_0(\cdot)$ denotes the zeroth-order modified Bessel function of the first kind.

A sample MRI contrast along with its associated \mathbf{V} signal (i.e., the k -space signal) are shown in Figure 2.3.

2.2 Principles of Diffusion MRI

In the previous section, it was shown that the orientation of the spins associated with water molecules in a given tissue can be manipulated using MRI techniques. This ability allows

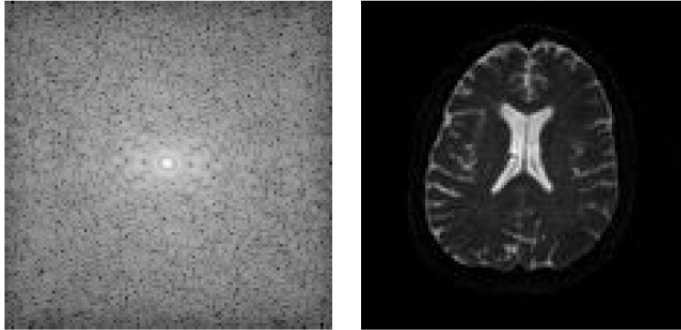


Figure 2.3: Left: Sample k -space representation of the induced coil voltage values corresponding to the MRI output image on the right. Credit: [9].

one to sense the diffusion of water molecules, and is therefore fundamentally relevant in the context of dMRI.

Diffusion, also termed random motion or intravoxel incoherent motion (IVIM), is one of the fundamental forms of particle displacement [77]. In addition to the details presented in Section 1.3, it is known that individual particles undergoing diffusion move randomly according to their probability to displace in each possible direction at each point in time. This probability is affected by the particular physical characteristics of the surrounding medium in which the particles are present. It should be noted that diffusion is different from bulk motion and flow, which are two additional forms of displacement relevant to dMRI, in which the motion of multiple particles is coherent in nature [77].

Conventional MRI scans can be utilized to capture diffusion information. To accomplish this goal, magnetic field gradient systems are commonly employed in practice. As mentioned in Section 2.1.5, the application of a magnetic field gradient allows one to encode spatial information of water molecules. Based on this principle, one can measure water diffusion information using a pair of gradients of opposite directions and equal magnitudes, as illustrated in Figure 2.4. As illustrated in this figure, a radio frequency (RF) excitation pulse is first applied over the region of interest causing all the spins to precess at the same Larmor frequency ω_0 and with equal phase. Following this step, a gradient field is applied over the t_2 time interval, causing spins in different spatial locations to rotate at different Larmor frequencies. After the system recovers the homogeneous \mathbf{B}_0 field during the time interval t_3 , all the spins regain their original Larmor frequency ω_0 ; however, their phases are not necessarily equal, giving rise to the term *de-phasing gradient* used to denote the first gradient. The fact that spins are out of phase originate a decay of the signal, as illustrated in Figure 2.4. Following the time interval t_3 , a second gradient is applied during t_4 . If

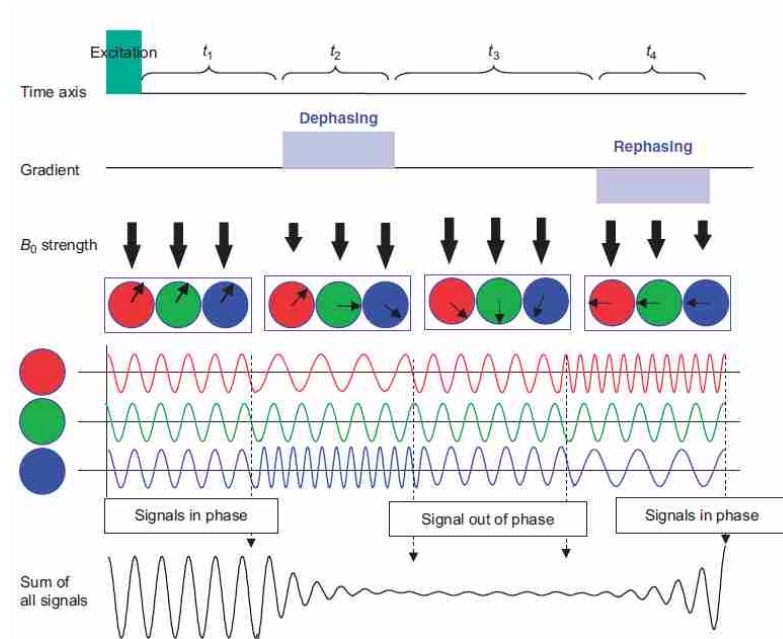


Figure 2.4: A sample dephase-rephase experiment for diffusion encoding. The colors of the circles are used to represent the different spatial locations of water molecules. The black arrows inside each circle are used to represent the spin direction associated with each molecule. Credit: [77].

the magnitude and duration of the second gradient are identical to those of the de-phasing gradient and if no displacement of particles takes place throughout the entire procedure, the spins should regain their original, coherent, phase upon termination of the t_4 interval. Therefore, the second gradient is commonly referred to as the *re-phasing gradient*.

Suppose now that water molecules undergo diffusion over the time interval t_3 . In such case, the displacement of particles occurring during this time interval would prevent the perfect re-phasing of the spins, therefore causing a loss of signal strength. The loss of signal intensity allows one to detect imperfect re-phasing of the spins, therefore providing a means for quantifying the degree of diffusion of water molecules in a sample. The output MRI scans which result from this procedure are referred to as *diffusion-encoded MRI* scans or simply diffusion-encoded images, due to their sensitivity to water diffusion. Note that the diffusion measurements which are attained using the above-mentioned procedure correspond to a fixed diffusion direction which depends on the orientation of the de-phasing and the re-phasing gradients. As a result, by controlling the orientation of the gradients, one can tune the direction in which diffusion is measured. This fact motivates the use

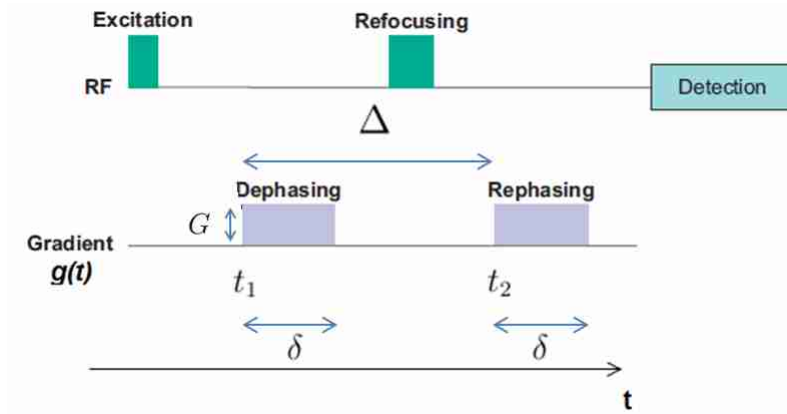


Figure 2.5: A sample pulsed gradient spin echo sequence. The refocusing RF pulse is a π RF pulse which is employed to attain an echo effect which in turn induces the slower T_2 decay rate, as opposed to T_2^* associated with pulse sequences without refocusing procedures. Adapted from [77].

of the term *diffusion-encoding gradients* to refer to these field gradients in the context of dMRI.

In practice, the pulse sequence presented in Figure 2.4 is often replaced by the more sophisticated pulsed gradient spin echo (PGSE) sequence introduced by Stejskal and Tanner in 1965 [103]. Many variants of the original PGSE sequence exist [77, 102], but the main ingredients present in most of them can be encapsulated by the particular sequence presented in Figure 2.5. The parameters that define the PGSE sequence are: the *mixing time* (Δ), the duration of the gradient pulses (δ), the strength of the gradients (G), and the starting times of the RF and gradient pulses. The main advantage of this type of sequence over the one in Figure 2.4 is its ability to attain a slower decay rate ($T_2 < T_2^*$) of the diffusion signals, which in turn result in higher values of SNR.

2.2.1 The Bloch-Torrey Equations and Related Expressions

The effects of water diffusion on MRI measurements were originally noted by Hahn in 1950 [54], and subsequently by Carr and Purcell in 1954 [24]. A few years later, in 1954, Torrey [109] adapted directly the Bloch equations presented in (2.12) to incorporate the effects of free isotropic diffusion. Torrey’s modifications were followed by those of Stejskal and Tanner [103], who considered the case of free anisotropic Gaussian water diffusion in

the same context in 1965. These modifications resulted in the general set of equations

$$\frac{\partial \mathbf{M}}{\partial t} = \gamma(\mathbf{M} \times \mathbf{B}_0) + \frac{1}{T_1}(M_0 - M_z)\hat{z} - \frac{1}{T_2}\mathbf{M}_\perp + \text{div}(D \nabla \mathbf{M}), \quad (2.18)$$

where div and ∇ represent the divergence and the gradient operators in the spatial coordinate only, respectively (e.g., $\nabla M = (\partial_x M, \partial_y M, \partial_z M)$), and D is the diffusion coefficient appearing in (1.1). In this particular context, D is referred to as the *apparent diffusion coefficient* (ADC), a name which suggests that the value of this quantity does not necessarily describe a Gaussian diffusion process, which is implicitly assumed by (2.18). Such a Gaussian assumption is often violated when the diffusion is hindered, as is the case in numerous regions of white matter in the brain [9, 5]. Equations (2.18) are of pivotal importance in the context of dMRI, as they define the signal formation model employed by some dMRI approaches.

2.2.2 Isotropic vs. Anisotropic Diffusion and the Diffusion Signal

In 1956, Torrey studied the particular case when $D(\mathbf{x}) \in \mathbb{R}^+, \forall \mathbf{x} \in \mathbb{R}^3$, in which case the diffusion process is called *free isotropic diffusion*, and equations (2.18) are known as the *Bloch-Torrey equations* [109]. In this case, the acquired MRI signal at spatial location \mathbf{x} and time t can be expressed as

$$S(\mathbf{x}, t) = |\mathbf{M}_\perp(\mathbf{x}, t)|e^{-bD(\mathbf{x})}, \quad (2.19)$$

where $b \in \mathbb{R}^+$, known as the *b-value*, is a function of the acquisition parameters, and its explicit expression depends on the type of pulse sequence employed. In the particular case of the PGSE sequence proposed by Stejskal and Tanner in [103] (see Figure 2.5), this parameter is given by $b = (\gamma\delta G)^2(\Delta - \delta/3)$ [68].

Despite the fact that isotropic diffusion describes accurately the diffusion profiles in some regions of the brain (e.g., regions composed predominantly of gray matter), it is not adequate to describe water diffusion in white matter in general. White matter is composed of numerous neural fiber tracts whose presence favour water diffusion along their direction, in which case equation (2.19) is clearly violated. In 1965, Stejskal and Tanner extended Torrey's work to the case when, for each $\mathbf{x} \in \mathbb{R}^3$, $D(\mathbf{x})$ in (2.18) is a 3×3 symmetric positive definite matrix. In this case, the underlying diffusion process corresponds to *free anisotropic Gaussian diffusion*, and the acquired MRI signal at spatial location \mathbf{x} and time t has the form

$$S(\mathbf{x}, \mathbf{u}, t) = |\mathbf{M}_\perp(\mathbf{x}, t)|e^{-b\mathbf{u}^T D(\mathbf{x})\mathbf{u}}, \quad (2.20)$$

where b is the b -value appearing in (2.19), and the orientation of $\mathbf{u} \in \mathbb{S}^2$ is parallel to that of the diffusion-encoding gradients.

From (2.19) and (2.20), it is evident that only the exponential term in S encapsulates information regarding the diffusion. Therefore, it is customary to isolate this term in S by dividing the acquired data by $S_0 = |\mathbf{M}_\perp(\mathbf{x}, t)|$, which is a conventional T_2 -weighted contrast acquired with no diffusion gradients in place, commonly referred to as the “ b_0 image.” Upon division by S_0 , the acquired signal at spatial location \mathbf{x} , diffusion direction \mathbf{u} and time t becomes

$$E(\mathbf{x}, \mathbf{u}) = S(\mathbf{x}, \mathbf{u}, t)/S_0(\mathbf{x}, t) = e^{-b\mathbf{u}^T D(\mathbf{x})\mathbf{u}}, \quad (2.21)$$

and the signal E is referred to as the *diffusion signal*.

2.2.3 The Q-space Formalism

As suggested by (2.20) and (2.21), a diffusion signal E depends on the parameters of the pulse gradient sequence employed, which are completely determined by the b -value and the diffusion direction \mathbf{u} . Under the \mathbf{q} -space formalism, the parameters of the diffusion-encoding gradient are encapsulated on the \mathbf{q} -value or wave-vector $\mathbf{q} \in \mathbb{R}^3$, which is given by

$$\mathbf{q} = \frac{\gamma \delta G \mathbf{u}}{2\pi}, \quad (2.22)$$

where G is the magnitude of the applied gradients (see Figure 2.5). In this case, an arbitrary diffusion signal may be represented as

$$E : \Omega_{\mathbf{x}} \times \Omega_{\mathbf{q}} \rightarrow \mathbb{R}_+ : (\mathbf{x}, \mathbf{q}) \mapsto E(\mathbf{x}, \mathbf{q}), \quad (2.23)$$

where $\Omega_{\mathbf{x}} \subset \mathbb{R}^3$ is a region of interest in space and $\Omega_{\mathbf{q}} \subset \mathbb{R}^3$ is the \mathbf{q} -space.

Given that the signal E is sampled over a spherical shell in the \mathbf{q} -space in the case of DTI and HARDI (implying that b is constant), the dependency of the signal E on the orientation of \mathbf{q} is often emphasized by replacing the argument \mathbf{q} by \mathbf{u} (as in equations (2.19) and (2.20)), in which case the signal E is modelled as a positively-valued function over $\mathcal{M} := \mathbb{R}^3 \times \mathbb{S}^2$,

$$E : (\Omega_{\mathbf{x}} \times \Omega_{\mathbf{u}}) \subset \mathcal{M} \rightarrow \mathbb{R}^+ : (\mathbf{x}, \mathbf{u}) \mapsto E(\mathbf{x}, \mathbf{u}), \quad (2.24)$$

where $\mathbb{S}^2 := \{\mathbf{u} \in \mathbb{R}^3 \mid \|\mathbf{u}\|_2 = 1\}$ is the unit sphere in \mathbb{R}^3 . For ease of notation, one occasionally drops the explicit dependency of the signals on the spatial coordinate (e.g., we write $E(\mathbf{q})$ as opposed to $E(\mathbf{x}, \mathbf{q})$), in which case it is implied that the analysis is carried solely over a fixed spatial location.

2.2.4 The Fourier Relationship between E and P

One important result in dMRI is that the diffusion signal E and the ensemble average propagator (EAP) are related via the 3D Fourier transform. Recall that the EAP is commonly denoted as $P : \Omega_{\mathbf{x}} \times \Omega_{\mathbf{r}} \subset \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^+$, so that $P(\mathbf{r} \mid \mathbf{x}_0; \tau)$ is the probability of a spin at initial position $\mathbf{x}_0 \in \Omega_{\mathbf{x}}$ to be at $\mathbf{x}_0 + \mathbf{r}$ after τ time units have elapsed. Using this formalism and notations, along with the so-called narrow pulse approximation (NPA), which assumes that $\delta \rightarrow 0$, one can express the diffusion signal E in terms of the EAP as [23, 32, 106]:

$$E(\mathbf{q}) = \int_{\mathbb{R}^3} P(\mathbf{r}) e^{i2\pi(\mathbf{q}\cdot\mathbf{r})} d\mathbf{x} = \mathcal{F}_{3D}^{-1}\{P\}(\mathbf{q}), \quad (2.25)$$

where \mathcal{F}_{3D} is the three dimensional Fourier transform. The equation above provides a useful relationship between the EAP and the diffusion signal E , and is therefore a central result for numerous dMRI techniques including HARDI.

2.3 Diffusion Tensor Imaging (DTI) and Its Limitations

As suggested by several studies [104, 116, 14, 13, 12, 88], a first order tensor, $D \in \mathbb{R}^{3 \times 3}$, is adequate for modelling the ADC of anisotropic diffusion at each spatial location, precisely as in (2.20). Following this assumption, and in relation to (2.20), DTI models the diffusion signal at spatial location $\mathbf{x} \in \mathbb{R}^3$ and diffusion direction $\mathbf{u} \in \mathbb{S}^2$ as

$$E(\mathbf{x}, \mathbf{u}) = e^{-\mathbf{b}\mathbf{u}^T D(\mathbf{x})\mathbf{u}}, \quad (2.26)$$

where the *diffusion tensor*

$$D(\mathbf{x}) = \begin{pmatrix} D_{xx}(\mathbf{x}) & D_{xy}(\mathbf{x}) & D_{xz}(\mathbf{x}) \\ D_{xy}(\mathbf{x}) & D_{yy}(\mathbf{x}) & D_{yz}(\mathbf{x}) \\ D_{xz}(\mathbf{x}) & D_{yz}(\mathbf{x}) & D_{zz}(\mathbf{x}) \end{pmatrix} \quad (2.27)$$

is a symmetric positive definite 3×3 matrix for every spatial location \mathbf{x} .

DTI normally employs a spherical sampling scheme on the \mathbf{q} – space as shown in Figure 1.2, and requires as few as 7 measurements for full characterization of the EAP. Six of these measurements are diffusion-encoded images used to estimate the entries of each $D(\mathbf{x})$, while the seventh measurement is a conventional T_2 -weighted image used to normalize

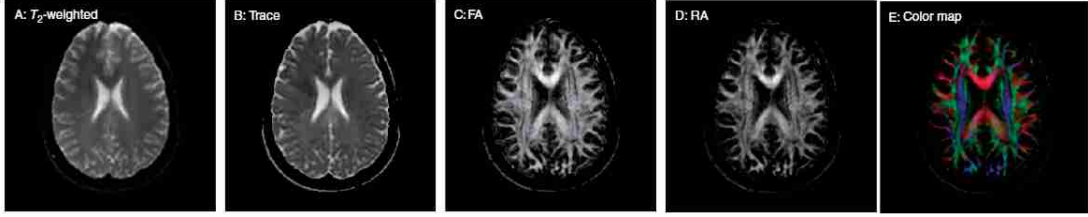


Figure 2.6: Different contrasts of an axial slice of the human brain: A) T_2 -weighted MRI contrast; B) $\langle \lambda \rangle$, C) FA map, D) RA map, E) orientation color map. For the FA and RA maps, high (bright) intensity values correspond to high anisotropy indices, indicating the predominance of a single diffusion direction. Credit: [77].

the diffusion signals as in (2.21). In practice, however, the robustness of DTI estimation methods is improved by obtaining more than the minimum number of measurements (often around 20 – 30) and performing a linear regression procedure to estimate D [61]. In either case, the overall acquisition times associated with DTI are relatively low as compared to those of other dMRI methods, which represents a major advantage of DTI over alternative methods of dMRI.

The usefulness of DTI as a diagnostic tool lies on its associated contrasts. These contrasts, which aim at discriminating regions based on their level of anisotropy, are computed based on the estimated diffusion tensor at each spatial location. The most widely used contrast measures are the so-called *relative anisotropy* and *fractional anisotropy* indices [79, 61], given by:

$$RA = \frac{1}{\sqrt{3}} \sqrt{\frac{(\lambda_1 - \langle \lambda \rangle)^2 + (\lambda_2 - \langle \lambda \rangle)^2 + (\lambda_3 - \langle \lambda \rangle)^2}{\langle \lambda \rangle^2}} \quad (2.28)$$

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \langle \lambda \rangle)^2 + (\lambda_2 - \langle \lambda \rangle)^2 + (\lambda_3 - \langle \lambda \rangle)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}},$$

where $\{\lambda_i\}_{i=1}^3$ are the eigenvalues of the diffusion tensor and $\langle \lambda \rangle = \sum_{i=1}^3 \lambda_i / 3$. Other relevant DTI contrasts include those based on the so-called Westin metrics [121] and color contrasts built from the tensor orientations [84, 40, 63]. Sample DTI contrasts are illustrated in Figure 2.6

Despite the indisputable success of DTI and its associated contrasts in numerous clinical applications [63, 84, 13, 28, 83], using the diffusion tensor model of the ADC along with the implicit assumption of anisotropic Gaussian diffusion leads to fundamental limitations

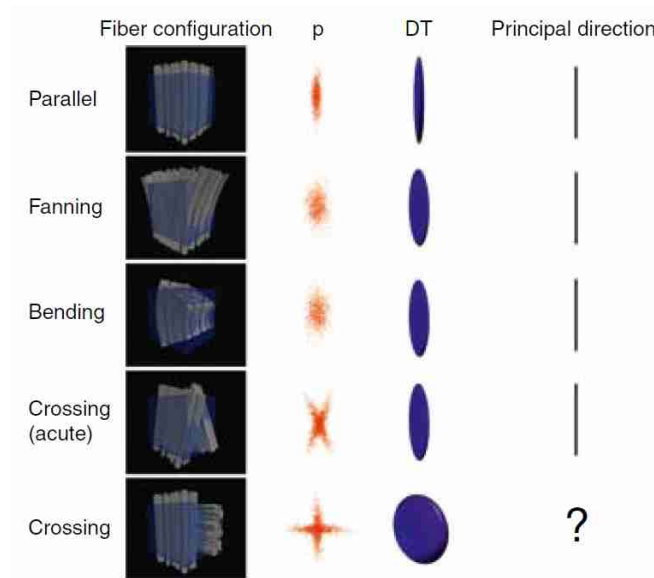


Figure 2.7: The limitations of DTI. From left to right: different fiber configurations of axon fibers on the first column. The second column shows the expected scatter pattern representing the EAP arising from each configuration on the left. The third column shows the best fit diffusion tensor for each configuration. The fourth column shows the principal direction corresponding to the best fit diffusion tensors. It can be seen that the principal directions of diffusion do not encapsulate the diffusion process accurately when diffusion is not entirely predominant in a single direction. Credit: [61].

of this imaging tool. Such limitations render the task of identifying complex diffusion profiles, including those shown in Figure 2.7 of fiber crossings, fiber bending, and fiber fanning, extremely challenging by means of DTI [61]. This limitation motivates the need for more complex models of water diffusion for the purpose of accurate delineation of white matter.

2.4 HARDI and Q-ball Imaging

High angular resolution diffusion imaging (HARDI), originally proposed in [113], is a dMRI technique which achieves a higher angular resolution to that of DTI at the expense of longer, yet clinically admissible acquisition times [9, 61]. The acquisition of HARDI signals often requires 60 to 80 samples in practice, all of which are restricted to a single spherical shell

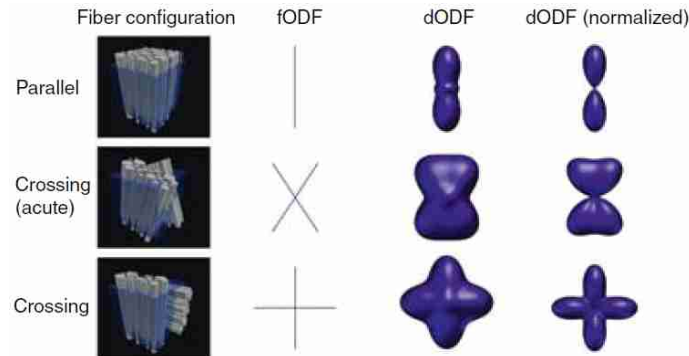


Figure 2.8: Sample fiber orientation distribution functions (fODFs) and diffusion distribution functions (dODFs) corresponding to different fiber configurations. Credit: [61].

in the \mathbf{q} – space as illustrated in Figure 1.2. As a result, a HARDI signal is a spherical function and so is its associated ADC (often assumed to be regular in some sense). Thus, as a direct generalization of (2.20), the diffusion signal at spatial location \mathbf{x} and diffusion direction \mathbf{u} is often represented in HARDI as

$$E(\mathbf{x}, \mathbf{u}) = e^{-bD(\mathbf{x}, \mathbf{u})} \quad (2.29)$$

where $\mathbf{u} \in \mathbb{S}^2$ is the direction of interest (i.e., the gradient direction), $b \in \mathbb{R}^+$ is the b –value, and $D : \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow \mathbb{R}^+$ is the (spherical) ADC.

The high angular resolution attained in HARDI allows one to work with more accurate models of diffusion and more sophisticated contrast types than in DTI. In particular, to overcome the limitations of DTI, HARDI often abandons the unimodal diffusion model of DTI and utilizes the concepts of fiber orientation distribution function (fODFs) and diffusion orientation distribution function (dODFs) instead. The fODF is a probability density function over the unit sphere which quantifies the fraction of fiber segments with a given orientation at each voxel [61]. In other words, the fODF at a given voxel is a function $f : \mathbb{S}^2 \rightarrow \mathbb{R}^+$ such that $f(\mathbf{u})$ is the fraction of fiber segments within that voxel with orientation parallel to \mathbf{u} . The dODF, on the other hand, is a scaled PDF defined over the unit sphere. At each direction $\mathbf{u} \in \mathbb{S}^2$, this function quantifies the relative likelihood of a water molecule to have undergone displacement due to diffusion in that particular direction over the experimental time. Due to the finite outer diameter of axon fibers, one expects to encounter a non-zero displacement probability in every possible direction. This means that fODFs are “sharper” in nature than dODFs [61]. These discrepancies between fODFs and dODFs are shown in Figure 2.8.

2.4.1 Q-Ball Imaging and HARDI Contrasts

For the purposes of the experimental studies in Chapter 4, we employ a particular method of dODF reconstruction called Q-ball imaging (QBI) [114]. In QBI, the dODF $\psi : \mathbb{S}^2 \rightarrow \mathbb{R}^+$ is given by a non-canonical radial projection

$$\psi(\mathbf{u}) = \int_0^\infty P(\alpha\mathbf{u}) d\alpha \quad (2.30)$$

of the ensemble average propagator, and it is estimated from the HARDI signal E via the Funk-Radon transform (FRT) [114]. In particular, in [114] it is shown that $\tilde{\psi} = \mathcal{FRT}\{E\}$, where \mathcal{FRT} is the FRT operator, is an approximation of ψ . The quality of this approximation depends on the b -value employed, and it improves as the value of this parameter is increased [114]. It has been shown that complex diffusion profiles which cannot be identified using DTI can be identified using QBI [61], including regions exhibiting fiber crossings and fiber branching. These advantages in turn have had an impact in the field of fiber tractography and in the study of brain regions showcasing white matter heterogeneity [2, 61, 6, 115, 16].

As in the case of DTI, there are contrasts and 3D models derived from HARDI signals which reveal important anatomical information of the human brain, some of which are believed to contain relevant information for the effective diagnosis of mental illnesses [94, 67, 38, 2, 16]. For instance, the informativeness of HARDI signals was exhaustively explored in [67], where the coefficients of spherical harmonics (SHs) associated to the acquired data were used to develop novel contrasts which characterize the local diffusion profiles of water molecules. Similarly, in [38], the total energy associated to each SH degree, which can be shown to be a rotationally invariant quantity, was used to generate image contrasts for the purpose of clinical diagnosis of brain diseases.

2.5 Filtering Methods for HARDI Data

Existing de-noising algorithms for HARDI data differ in terms of the quantity they aim to recover, the noise contamination model employed, the type of regularity priors imposed on the underlying signals, and on the mathematical tools used to formulate and solve the associated reconstruction problem(s). We now provide an overview of various HARDI de-noising approaches found in the literature.

2.5.1 MRI De-noising Methodologies

De-noising methods specifically designed for conventional MRI scans can be applied to each 3D image volume resulting from the restriction of the HARDI signal E to each angular coordinate $\mathbf{u} \in \mathbb{S}^2$ at a time.

A class of methods of this type are those which make use of the sparsifying properties of wavelets [119, 55, 80, 124, 123, 89, 11, 90] (please refer to [90] for a comprehensive review of these methods). For instance, the method of [89] proposes an adaptive wavelet domain estimation procedure based on a Bayesian framework, in which the wavelet coefficients of the original signal are estimated from a preliminary coefficient classification based on the correlation of significant image features across the resolution scales. A single parameter, $K > 0$, is used to tune the level of filtering to be performed by this method, and it does not require prior knowledge of the noise distribution. On the other hand, the method of [80] employs wavelet thresholding of noisy coefficients of the square-magnitude MR images, followed by a de-biasing procedure on the estimated coefficients to account for the non-central chi-square statistics of the noise. Finally, the approach of [4] proposes a wavelet-domain bilateral filter for enhancing MR images. This filter is applied to the squared-magnitude images and uses the undecimated wavelet transform.

Other methods of this type employ TV regularization for the purpose of MR imagery enhancement. For example, the method of [10] imposes a TV prior on the MRI scans along with a sparsity constraint using a dictionary comprised of discrete cosine functions. Another related algorithm is presented in [93], in which the authors propose to enforce sparsity constraints on MR signals using overcomplete dictionaries containing a combination of representation functions arising from multiple sparsifying transforms. Finally, the method of [117] introduces an adaptive approach to MRI de-noising using a TV minimization model whose associated regularization parameter is estimated using a local noise estimation technique.

Additional methods for MRI enhancement include those of [45, 48, 69, 42, 15, 65]. For instance, in the variance stabilization approach introduced in [45], the signal-dependent noise variance of MR signals is modified to appear roughly uniform via an invertible transformation applied to the data. The uniformity of the noise variance of the transformed image allows one to obtain performance improvements on conventional filters designed for additive Gaussian noise contamination. The anisotropic filter approach of [48], on the other hand, is a variant of the anisotropic filter of [87], and is a spatially adaptive filter specifically designed to blur homogeneous regions of the MRI image while preserving its region boundaries. Finally, the variational approach method from [69] proposes two smoothing filters for MRI data, each of which involves the minimization of a cost functional used to

measure the oscillations in the noisy data, and solved via a discrete version of a fourth-order PDE.

2.5.2 Angular Smoothing

The methods of [22] and [36] propose a smoothing procedure for signals defined over the unit-sphere. Given that HARDI signals are assumed to be scalar-valued spherical functions when restricted to any (fixed) spatial location, it is possible to employ these methods in a *voxel by voxel* fashion for the de-noising of this type of data sets. Since both methods impose smoothness regularity constraints over the unit sphere, they are said to perform *angular smoothing* with no spatial regularization constraints.

In particular, the diffusion-based smoothing method of [22] seeks to smooth a square integrable scalar valued signal $f : \mathbb{S}^2 \rightarrow \mathbb{R}$ by finding the solution at time $t \in \mathbb{R}^+$ to the PDE

$$\frac{\partial g}{\partial t}(\mathbf{u}, t) = \Delta_{\mathbb{S}^2} g(\mathbf{u}, t), \quad \forall \mathbf{u} \in \mathbb{S}^2, \quad (2.31)$$

with initial condition $g(\mathbf{u}, 0) = f(\mathbf{u}), \forall \mathbf{u} \in \mathbb{S}^2$, and where $\Delta_{\mathbb{S}^2}$ is the Laplace-Beltrami operator.

It turns out that the solution to (2.31) at time t is identical to the result of spherical convolution of f with the Gauss-Weierstrass kernel $g_\tau : \mathbb{S}^2 \rightarrow \mathbb{R}^+$, whose effective support is higher for higher values of τ . The effective support of the kernel, in turn, determines the degree of smoothing that the signal f undergoes, with more smoothing achieved with a greater effective support of the kernel. Thus, the ideal value of τ depends on the desired degree of smoothing of f , making lower (resp. higher) values of τ optimal when f is believed to have predominantly high (resp. low) frequency content. As such, given the variety of spherical signal types which arise in HARDI data, a major disadvantage of methods of this kind lies in the impossibility to find a single value of τ which is optimal for the smoothing of all of these functions.

A different approach to angular smoothing is that of [36], where the dODF is estimated by computing a regularized projection of the HARDI signal, E , into the set of real and symmetric spherical harmonics (SHs). The regularization is imposed using the Laplace-Beltrami operator, $\Delta_{\mathbb{S}^2}$, which provides a measure of smoothness of functions defined over the unit sphere. The formulation and implementation details of this method are presented in Chapter 3.

2.5.3 Non-Local Means

In [20, 21], the authors presented a novel de-noising method for scalar valued images which exploits the self-similarity properties present in natural images. This method, known as *non-local means* (NLM), obtains an estimate \hat{X} of the noise-free original image X with domain I , $\{X_s\}_{s \in I} \subset \mathbb{R}$, from the intensities of its noisy counterpart, $\{Y_s\}_{s \in I}$, according to:

$$\hat{X}_s = \frac{1}{C_s} \sum_{t \in V_s} w_{s,t} Y_t, \quad \text{with } C_s = \sum_{t \in V_s} w_{s,t}, \quad (2.32)$$

where $V_s \subset I$ is a set of indices defining a spatial neighbourhood of X_s , called the search window, and $w_{s,t} \geq 0$ aims at quantifying the degree of similarity between X_s and X_t . Each $w_{s,t}$ is computed based on the similarity between two local neighbourhoods of X_s and X_t . For instance, the original NLM filter computes $w_{s,t}$ as

$$w_{s,t} = \exp \left\{ -\frac{1}{h} \sum_{k \in P} \beta_k |Y_{s-k} - Y_{t-k}| \right\}, \quad (2.33)$$

where $P \subset I$ is the index set of a spatial neighbourhood around the center of the image coordinates. For instance, if $I \subset \mathbb{Z}^2$, this indexing set is commonly set to $P = \{(x_1, x_2) \in I \mid |x_1| \leq L_{x_1}, |x_2| \leq L_{x_2}\}$ for some $L_{x_1}, L_{x_2} \in \mathbb{N}$.

Different NLM algorithms differ in terms of their definitions of P , I , and $w_{s,t}$. Thus, for instance, in [34, 35, 108, 37], $w_{s,t}$ which are optimal under a Bayesian framework are developed for different noise statistics and noise contamination models. On the other hand, in [122, 66], the original NLM filter is altered for the purpose of dMRI de-noising. In particular, in [66], HARDI signals are modelled as functions over the $\mathbb{R}^3 \times \mathbb{S}^2$ space which undergo Rician noise contamination, and local neighbourhoods over $\mathbb{R}^3 \times \mathbb{S}^2$ are carefully defined to exploit the self-similarity of the images. In [122], on the other hand, the HARDI signals are modelled as vector-valued images over \mathbb{R}^3 , so that the averaging in (2.32) is performed using vectors, which correspond to sampled versions of spherical functions, rather than scalar voxel intensities as in the original NLM filter. Despite the excellent performance of NLM-based filters, their computational requirements represent their main drawback. As such, these filters are rarely used when dealing with large volumes of data, as is often the case in dMRI.

2.5.4 ℓ_1 -Based HARDI Reconstruction

A different class of methods exploit the robustness offered by the ℓ_1 -norm in the context of statistical estimation [96]. For instance, the method of [64] proposes to find an estimate

of the ADC as a minimizer of the cost functional

$$J(D) = \|D\|_{TV} + \sum_{i=1}^n \left[\lambda \int_{\Omega} |E(\mathbf{x}, \mathbf{u}_i) - S_0(\mathbf{x}) \exp\{-bD(\mathbf{x}, \mathbf{u}_i)\}| d\mathbf{x} \right] \quad (2.34)$$

$$- \sum_{i=1}^n \left[\mu \int_{\Omega} H(E(\mathbf{x}, \mathbf{u}_i), S_0(\mathbf{x})) \cdot \frac{\log(D(\mathbf{x}, \mathbf{u}_i))}{D(\mathbf{x}, \mathbf{u}_i)} d\mathbf{x} \right], \quad (2.35)$$

where $\Omega \in \mathbb{R}^3$ and $\{\mathbf{u}_i\}_{i=1}^n \subset \mathbb{S}^2$ are the sets of spatial locations and diffusion directions, respectively; λ, μ are tuning parameters; H is a function such that $H(a, b) = 0$ if $a \leq b$, $H(a, b) = 1$ if $a > b$; and

$$\|D\|_{TV} = \int_{\Omega} \sqrt{\sum_{i=1}^n \|\nabla D(\mathbf{x}, \mathbf{u}_i)\|_2^2} d\mathbf{x} \quad (2.36)$$

is a total variation (TV) semi-norm of D , where $\nabla D(\mathbf{x}, \mathbf{u}_i)$ is the gradient of $D_i(\mathbf{x}) := D(\mathbf{x}, \mathbf{u}_i)$ with respect to the \mathbf{x} coordinate. In (2.34), the first term accounts for spatial regularity of each diffusion-encoded image. The second term, which penalizes the level of the noise of the image, makes use of the L_1 -norm (ℓ_1 -norm for the discretized case) to improve the overall degree of robustness of the algorithm. Finally, the third term incorporates the constraint that $E(\mathbf{x}, \mathbf{u}) \leq S_0(\mathbf{x}, \mathbf{u}), \forall (\mathbf{x}, \mathbf{u}) \in \Omega \times \{\mathbf{u}_i\}_{i=1}^n$, as it is expected in view of (2.20) and (2.21). This method will be revisited in the experimental study presented in Chapter 4.

Other methods of HARDI de-noising employ ℓ_1 -norm regularization to exploit its sparsifying properties in the context of compressed sensing (CS) [39]. In [73], for example, the authors introduce a set of functions defined over \mathbb{S}^2 , called spherical ridgelets (SRs), whose L_2 -energy is concentrated along the great circles of \mathbb{S}^2 . It is argued that this particular characteristic of SRs is responsible for their sparsifying advantages over SHs in the context of HARDI signal approximations. In particular, it is shown that a SR least-squares fitting procedure regularized by the ℓ_1 -norm of the SR coefficients results in higher accuracy of reconstruction as compared to its SH-based counterpart in the context of CS.

A direct extension to this method is presented in [74], where it is proposed to minimize an alternative functional of the form

$$J(c) = \frac{1}{2} \|\mathcal{A}\{c\} - E\|_{\mathfrak{F}}^2 + \lambda \|c\|_1 + \mu \|\mathcal{A}\{c\}\|_{TV},$$

where $\mathcal{A}\{\cdot\}$ is the relation that maps the SR coefficients c to their corresponding linear combination of SRs, $\|\cdot\|_{\mathfrak{F}}$ is the Frobenius norm, and the TV semi-norm of a discrete

HARDI signal $f \in \mathbb{R}^{|\Omega| \times K}$ is given by

$$\|f\|_{TV} = \sum_{k=1}^K \sum_{i=1}^{|\Omega|} \|\nabla f(\mathbf{x}_i, \mathbf{u}_k)\|_2, \quad (2.37)$$

where $|\Omega|$ is the number of voxels in the image, K is the number of diffusion-encoded images, and ∇ represents a discrete gradient operator.

The inclusion of the additional TV regularization term in this modified functional introduces regularization in the spatial domain, in addition to the already present angular regularization attained by the SRs fitting described in [73].

2.5.5 Denoising of HARDI Signals using FEM and TV

In [72], de-noising of HARDI signals is achieved through the minimization of an energy functional comprised of a data fidelity term and two smooth regularization priors. McGraw et al. use this approach to apply regularization priors in both the spatial and the angular domain. The spatial regularization across a volumetric lattice is obtained using a TV semi-norm based prior, while the angular regularization is achieved through the finite element method (FEM) using first and second order smoothness constraints.

In particular, [72] employs the signal model $\hat{S} = S_0 \exp(-bD) + \eta$, where $\hat{S} : U \subset \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow \mathbb{R}^+$ is the acquired HARDI signal, $S_0 : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}^+$ is the b_0 -image, $D : U \subset \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow \mathbb{R}^+$ is the ADC, and $\eta : U \subset \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow \mathbb{R}^+$ is Rician noise. Under such a model, McGraw et al. propose to find the de-noised HARDI reconstruction $S = S_0 \exp(-bD)$ from its noisy counterpart \hat{S} as a minimizer of the cost functional

$$J(S) = \frac{\mu}{2} \int_U \int_{\mathbb{S}^2} |S(\mathbf{x}, \mathbf{u}) - \hat{S}(\mathbf{x}, \mathbf{u})|^2 d\mathbf{u} d\mathbf{x} + \int_{\mathbb{S}^2} \|\nabla_{\mathbb{S}^2} S\|^2 d\mathbf{u} + \int_U g(\mathbf{x}) \|\nabla_{\mathbf{x}} S\| d\mathbf{x}. \quad (2.38)$$

The first term in the cost functional is the standard data fidelity term; it ensures faithfulness of the reconstruction to the noisy data in the L_2 -sense. The second term enforces regularization over the spherical domain at each voxel, and incorporates both first and second smoothness constraints. Finally, the third term imposes spatial regularization using a modification to the color TV semi-norm proposed in [17]. In particular, this regularization term treats the discretized HARDI signal at each voxel as a vector-valued function

$S(\mathbf{x}) = [S_1(\mathbf{x}) S_2(\mathbf{x}) \dots S_K(\mathbf{x})]^T$, with each entry corresponding to a diffusion encoding direction, so that $\|\nabla_{\mathbf{x}} S\| = \sum_{i=1}^K \|\nabla S_i\|$, where ∇ is a discrete gradient map. Moreover, the coupling factor $g(\mathbf{x}) = 1/(a + \|GA(\mathbf{x})\|^2)$, where GA is the generalized anisotropy index introduced in [82], aims at preserving details at highly anisotropic regions while smoothing the rest of the signal. The parameter $\mu > 0$ controls the degree of smoothing present in the reconstruction. Low values of μ , result in smooth reconstructions that deviate significantly from the data acquired (i.e., appropriate for low frequency signals), while high values of μ encourage reconstructions which fit the data points closely, even if they are not smooth (appropriate for high frequency signals).

2.5.6 Other Filtering Methods for HARDI

Among the additional filtering methods for HARDI encountered in the literature are the ones in [29, 59, 6, 7, 100, 95]. For instance, in [29] Chen et al. propose a method for reconstructing the ADC from noisy HARDI data as the solution to a variational problem with positivity constraints on the ADC along with spatial regularization constraints on both the S_0 signal and on the SH coefficients of the ADC. They limit their model to SH coefficients up to degree four based on the observations in [46, 3]. Limiting the solution to SH expansions of degree four imposes smoothing constraints that might result in the oversight of important high frequency details which might be present in anisotropic regions of interest. Jansons et al. [59] propose a new statistic called persistent angular structure (PAS) for the characterization of diffusion profiles from HARDI data. The PAS is a spherical function which, when embedded in three dimensional space on a two dimensional sphere, has the Fourier transform that best fits the normalized HARDI measurements. On the other hand, Assemlal et al. proposed variational reconstruction methods for estimation of dODFs [6] and of PDFs [7] from HARDI data. For instance, the method of [6] takes into account the Rician statistics of the noise present in MRI data, and finds the dODF reconstruction as the minimizer of a cost functional. Spatial regularization constraints are imposed by introducing the gradient norm of the SH coefficients of the reconstruction. Furthermore, this method employs a “robust estimation function” and a “discontinuity-preserving regularization function” to improve the robustness to Rician noise and to better resolve inter-regional portions of the data, respectively. Finally, in [100] Savadjiev et al. propose a differential geometric approach for the regularization of HARDI data. Neuronal bundles are modelled as 3D curves, which are inferred from an initial set of dODFs based on the notion of co-helicity between vectors. Co-helicity serves as a measure of compatibility of local fiber orientations with respect to their immediate neighbours. Based on this quantity, the method is able to estimate the geometric properties of 3D curves (which represent white

matter tracts) which pass through every spatial location, thereby producing a regularized estimate of dODFs.

2.5.7 Conclusions

Altogether, the filtering methods described in this section suggest that HARDI signals exhibit both spatial *and* angular regularity. In other words, the effective reconstruction of HARDI signals through the use of model priors seem to require regularity constraints imposed on the restricted versions of the signals to the spatial domain (i.e., regularity of the signal across \mathbb{R}^3) as well as to the angular domain (i.e., regularity of the signal across \mathbb{S}^2). Ignoring either one of these factors would result in poorer reconstruction results, as one expects that the inclusion of any correct prior information about the underlying quantity of interest should result in higher quality of the estimates obtained.

Another important factor to HARDI reconstruction is the type of regularization chosen to incorporate prior information about the underlying signal(s). In particular, from the principles of physics governing the water diffusion process, one expects to observe smooth-like behaviour of HARDI signals over \mathbb{S}^2 at every fixed spatial location; thereby suggesting the appropriateness of strict angular smoothness constraints for the reconstruction of these signals. On the other hand, the intricate anatomy of the human brain suggests that diffusion MRI data is inherently non-smooth across space (3D images corresponding to HARDI data restricted to fixed arbitrary diffusion-directions). Instead, as in conventional MRI scans, one expects to encounter edges and textures which are routinely modelled as non-smooth behaviour in the underlying signals, thereby suggesting the sub-optimality of strict smoothness priors on these data. These two premises constitute the main motivation for the design of the filtering method proposed in Chapter 4, which incorporates prior knowledge about the regularity of HARDI signals in both their spatial and angular domains, and which offers superior performance over numerous state of the art de-noising methodologies for HARDI data.

Chapter 3

Main Contributions

The main contribution of this thesis is to develop a novel algorithm, known as spatially regularized spherical reconstruction (SR²), which exploits both the spatial and the angular regularity of the underlying signals in HARDI¹. The proposed method employs a variational framework which relies on two fundamental assumptions about the nature of the underlying signal of interest. In particular, HARDI signals are assumed to be smooth over their angular domain and piece-wise smooth over their spatial domain. In addition to being effective at de-noising HARDI data sets, the SR² algorithm is a numerical implementation which can be executed in a parallel fashion, suggesting its computational advantages over some alternative implementations.

In this section, the underlying assumptions behind the SR² method as well as the variational problem which it attempts to solve are introduced. Additionally, some details about the regularity constraints imposed on the HARDI data are discussed, and a detailed description of the numerical implementation associated with the proposed method is presented.

3.1 Notations

Consider a continuously defined, normalized, HARDI signal $E_c : \mathcal{M} := \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow \mathbb{R}^+$. In practice, the acquired HARDI signal E is a sampled version of E_c , so that $E : \Omega_{\mathbf{x}} \times \Omega_{\mathbf{u}} \rightarrow \mathbb{R}^+$, where $\Omega_{\mathbf{x}} \subset \mathbb{R}^3$ is a discrete lattice of spatial locations, and $\Omega_{\mathbf{u}} = \{\mathbf{u}_k\}_{k=1}^K \subset \mathbb{S}^2$ is the set of samples over \mathbb{S}^2 , so that K is the number of diffusion-encoded images acquired.

¹A summarized version of the main contribution presented herein can be found in [86].

In practice, it is often the case that $\Omega_{\mathbf{x}}$ corresponds to a rectangular lattice, so we may assume from now on that $\Omega_{\mathbf{x}} = \{(n, m, l) \in \mathbb{Z}^3 \mid 0 \leq n \leq N, 0 \leq m \leq M, 0 \leq l \leq L\}$ for some $N, M, L \in \mathbb{N}$, in which case $|\Omega_{\mathbf{x}}| = NML$ is the number of elements in $\Omega_{\mathbf{x}}$. For notational purposes, it is useful to consider three different representations of the HARDI signal E :

1. **As a 5D matrix:**

We may simply consider E as a matrix indexed by 5-tuples in which case

$$E = (E_{i,j,k,l,m,n})_{(i,j,k,m,n) \in \Omega_{\mathbf{x}} \times \Omega_{\mathbb{S}^2}},$$

where $\Omega_{\mathbb{S}^2} = \{(\phi_k, \theta_k), k = 0, 1, \dots, K\}$ is the set of angles in spherical coordinates corresponding to the locations of the samples of E_c over the unit sphere.

2. **As a sequence of 3-dimensional diffusion-encoded images:**

The restriction of E_c to any fixed angular coordinate corresponds to a continuously defined diffusion-encoded image. Thus one may consider E_c as a set of functions defined over \mathbb{R}^3 and indexed by the angular coordinate $\mathbf{u} \in \mathbb{S}^2$. Likewise, E can be considered as a sequence of K different three dimensional discrete diffusion-encoded images, that is:

$$E = \{E_k\}_{k=1}^K, \quad \text{where } E_k \in \mathbb{R}_+^{N \times M \times L}, k = 1, 2, \dots, K \quad (3.1)$$

3. **As a sequence of K -dimensional vectors (sampled spherical functions):**

the restriction of E_c to any fixed spatial coordinate corresponds to a spherical function, so that one may consider E_c as a set of functions defined over \mathbb{S}^2 and indexed by the spatial coordinate $\mathbf{x} \in \mathbb{R}^3$. Likewise, E can be considered as a sequence of K -dimensional vectors which result from sampling $|\Omega_{\mathbf{x}}|$ spherical functions at K different directions, that is:

$$E = \{E^i\}_{i=1}^{|\Omega_{\mathbf{x}}|}, \quad \text{where } E^i \in \mathbb{R}_+^K, i = 1, 2, \dots, |\Omega_{\mathbf{x}}| \quad (3.2)$$

3.2 Spatial Regularization for HARDI Data

The proposed algorithm employs the following fundamental assumption on the spatial behaviour of HARDI signals:

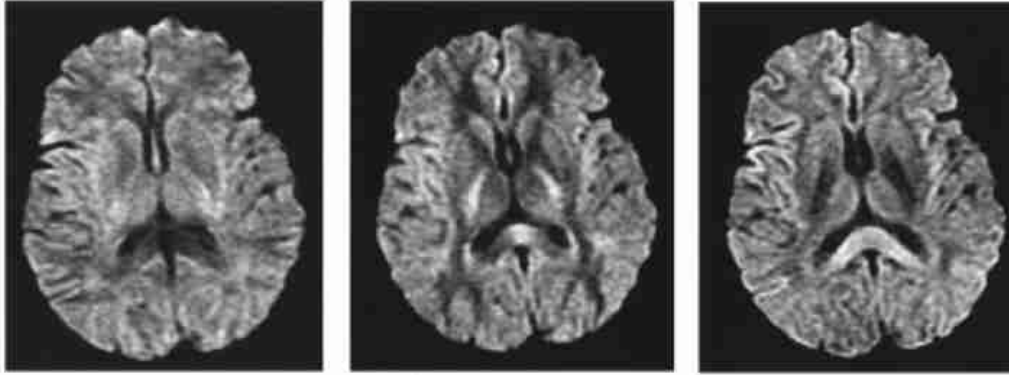


Figure 3.1: Three sample diffusion-encoded MRI “slices” or images. A piecewise-smooth model appears appropriate from the images. Credit: [101].

Spatial Regularity Assumption: The function $E_c(\cdot, \mathbf{u}) : \mathbb{R}^3 \rightarrow \mathbb{R}^+$ is a piecewise-smooth function over \mathbb{R}^3 for all $\mathbf{u} \in \mathbb{S}^2$, so that each E_k in (3.1) is a sampled version of a piece-wise smooth function over \mathbb{R}^3 .

Based on this regularity assumption, the proposed algorithm promotes the use of TV regularization over the spatial domain of HARDI signals. The appropriateness of this methodology is justified by the apparent piecewise-smooth nature of diffusion-encoded images, as illustrated in Figure 3.1.

TV regularization employed in the context of image processing was originally introduced in [97]. Ever since its introduction, numerous TV-based algorithms have emerged for different purposes, e.g., including image inpainting, super-resolution, and deconvolution [27, 25, 19] (see [26] for a thorough review of total variation regularization for image processing applications). The basic idea behind numerous TV-based algorithms is to exploit the fact that signals of bounded total variation may exhibit a *piecewise smooth* behaviour. It is this principle which makes TV regularization successful for the purpose of natural image reconstruction and enhancing, given that these images are often projected representations of 3D objects, and therefore comprised of a finite number of edges/discontinuities, textures, and large smooth regions.

The original definition of the TV semi-norm applies to scalar-valued functions defined over an open subset of an Euclidean space, that is, to functions of the form $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ for some $n \in \mathbb{N}$ and an open set U . Even though such a definition of the TV semi-norm can be applied in the context of HARDI de-noising with little modification, there is no guarantee that this approach is optimal, and it is worth considering different TV semi-

norm definitions for this type of signal. Numerous attempts have been made to extend the definition of the TV semi-norm for non-scalar valued functions or for functions with more complex domains [19, 51, 58]. In this section, multiple definitions of TV semi-norms will be presented, and the appropriateness of the channel-by-channel TV semi-norm for the purpose of HARDI de-noising will be highlighted.

3.2.1 TV for Scalar Functions over \mathbb{R}^n and the ROF Model

Let $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in L^1_{loc}(U)$, where $L^1_{loc}(U)$ is the set of locally integrable functions over U . The dual definition of the TV semi-norm of f is [26]

$$\|f\|_{TV(U)} := \sup_{\xi \in \mathcal{C}_c^1(U, \mathbb{E}^n)} \left\{ \int_U f(\mathbf{x}) \operatorname{div} \xi(\mathbf{x}) d\mathbf{x} \right\}, \quad (3.3)$$

where \mathbb{E}^n denotes the closed unit ball of \mathbb{R}^n , $\mathcal{C}_c^1(U, \mathbb{E}^n)$ is the set of continuously differentiable vector fields from U to \mathbb{E}^n with compact support, and $d\mathbf{x}$ is the Euclidean measure in \mathbb{R}^n .

In the case when $f \in \mathcal{C}^1(U)$, where $\mathcal{C}^1(U)$ is the set of scalar valued continuously differentiable functions over U , (3.3) can be reduced to [26]

$$\|f\|_{TV(U)} := \int_U \|\nabla f(\mathbf{x})\|_2 d\mathbf{x}, \quad (3.4)$$

where ∇ represents the standard gradient map. This definition naturally allows one to define a set of scalar functions over U , called the functions of bounded variation over U , denoted by $BV(U)$, and with the property that $f \in BV(U)$ if and only if $\|f\|_{TV(U)} < \infty$.

In this original context, the TV image de-noising problem, commonly referred to as the Rudin-Osher-Fatemi (ROF) problem [97], involving a noisy image g and original image f aimed at finding the solution to

$$\operatorname{argmin}_{f \in BV(U)} \left\{ \int_U |f(\mathbf{x}) - g(\mathbf{x})|^2 d\mathbf{x} + \|f\|_{TV} \right\}. \quad (3.5)$$

The problem in (3.5) is known to be convex [26], and its solution can be found directly using optimization methods applied directly to the *primal* problem in (3.5), for example, by gradient descent-based or majorization-minimization methods [97, 43, 26]. It is important to note, however, that significant performance advantages can be achieved by applying *dual-based* convex optimization techniques, in which case one employs the dual definition of (3.3) (see e.g., [25, 26, 126]).

3.2.2 Scalar TV over \mathcal{M}

Motivated by the original ROF model for scalar valued signals, an alternative definition of the TV semi-norm applicable to HARDI signals arises by setting $U = \Omega_{\mathbf{x}} \times \Omega_{\mathbb{S}^2} \subset \mathcal{M}$ as the domain of definition of the image, in which case it is customary to utilize the coordinate chart (x, y, z, θ, ϕ) , where $(x, y, z) \in \Omega_{\mathbf{x}} \subset \mathbb{R}^3$ and $(\theta, \phi) \in \Omega_{\mathbb{S}^2} \subset [0, \pi] \times [0, 2\pi)$. In this case, based on the Riemannian structure of \mathcal{M} , a natural extension of the TV semi-norm for a smooth function $f : \Omega_{\mathbf{x}} \times \Omega_{\mathbb{S}^2} \rightarrow \mathbb{R}$ is

$$\|f\|_{TV_{\mathcal{M}}(U)} := \int_U \|\nabla_{\mathcal{M}} f(x, y, z, \theta, \phi)\|_2 d\eta(U), \quad (3.6)$$

where $d\eta(U) = \sin(\theta) dx dy dz d\theta d\phi$, and

$$\nabla_{\mathcal{M}} f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial \theta}, \frac{1}{\sin \theta} \frac{\partial f}{\partial \phi} \right], \quad (3.7)$$

is the gradient of f . We refer to the TV definition in (3.6) as the $TV_{\mathcal{M}}$ semi-norm.

The solution to the ROF model associated with the $TV_{\mathcal{M}}$ semi-norm can be attained in an analogous fashion to that of (3.5), e.g., by means of [25]. In the case of $TV_{\mathcal{M}}$ regularization, a possible approach to estimate the derivatives along the θ and ϕ coordinates is by computing finite differences over equiangular sampling schemes of the unit sphere, see e.g., [58]. In this case, the implementation of the ROF minimizer is adapted by simply adding the two extra dimensions to the data corresponding to the two additional coordinates, while paying particular attention to the weights of the finite differences to ensure the stability of the iterative algorithm [58]. To bypass the requirement of equiangular sampling, we have also investigated the use of an analytical model similar to that of [76] whereby closed-form expressions of the derivatives are estimated. Unfortunately, this approach presents undesirable convergence properties which are still under investigation.

It should be noted that, as indicated by the experimental results in Chapter 4, the $TV_{\mathcal{M}}$ semi-norm is not appropriate in the context of HARDI signal reconstruction. This is due to the fact that HARDI signals rarely present discontinuities across their angular domain, and are therefore better estimated using stricter smoothness priors over \mathbb{S}^2 . It is possible, however, that other applications within dMRI signal estimation might benefit from the ROF model corresponding to the $TV_{\mathcal{M}}$ definition, as discussed in Chapter 5.

3.2.3 TV Definitions for Vector Valued Functions

The definition of the TV semi-norm in (3.4) can be extended in diverse ways for the case when f is a vector valued function of the form $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}^m$, with $m > 1$. These

extensions are commonly referred to as *vectorial TV* definitions, and in the discrete setting, they can be readily applied to HARDI data in view of the interpretation given by (3.2).

A family of vectorial TV definitions relies on concepts from differential geometry presented in [125]. In this work, di Zenzo addresses the problem of edge detection in vector-valued images by interpreting image $f : \mathbb{R}^2 \rightarrow \mathbb{R}^m : \mathbf{x} \mapsto (f^1(\mathbf{x}), f^2(\mathbf{x}), \dots, f^m(\mathbf{x}))$ as a 2-dimensional manifold embedded in \mathbb{R}^m . Based on this consideration, a metric tensor on this manifold can be defined as

$$g_{h,k}(\mathbf{x}) = \langle \partial_{x_h} f(\mathbf{x}), \partial_{x_k} f(\mathbf{x}) \rangle, \quad \forall \mathbf{x} \in \mathbb{R}^3 \quad \text{with } h, k = 1, 2, \quad (3.8)$$

where $\partial_{x_h} f := \left[\frac{\partial f^1}{\partial x_h}, \frac{\partial f^2}{\partial x_h}, \dots, \frac{\partial f^m}{\partial x_h} \right]$. The definition of this metric allows one to determine the direction $\nu(\mathbf{x})$ of the greatest rate of change of f at \mathbf{x} (i.e., the direction perpendicular to an edge), as well as the rate of change corresponding to this direction (i.e., the “strength” of an edge). In [99] Sapiro suggests that the strength of an edge is a function of how the two eigenvalues of g , denoted as λ_- and λ_+ , compare to each other. As a result, [99] introduced a vectorial TV definition of a vector-valued function f which replaces the term $\|\nabla f(\mathbf{x})\|_2$ (in the scalar case) at \mathbf{x} by

$$\|f\|_{TV_{SR}} = \int_U s(\lambda_-(\mathbf{x}), \lambda_+(\mathbf{x})) d\mathbf{x} \quad (3.9)$$

for an arbitrary scalar function s . This general definition of the vectorial TV semi-norm motivated the study of the properties offered by (3.9) under different definitions of s . For example, if one were to set $s(\lambda_-, \lambda_+) = \sqrt{\lambda_- + \lambda_+}$, one would obtain the so-called *pointwise Frobenius TV* definition [51],

$$\|f\|_{TV_F} := \int_{\Omega} \|Jf(\mathbf{x})\|_F d\mathbf{x}, \quad (3.10)$$

where $\|\cdot\|_F$ stands for the Frobenius norm and $Jf(\mathbf{x})$ is the Jacobian matrix of f at \mathbf{x} , so that $(Jf(\mathbf{x}))_{j,k} = \partial f^j(\mathbf{x})/\partial x_k$. Similarly, the authors of [51] use $s(\lambda_-, \lambda_+) = \sqrt{\lambda_+}$, in which case one has the alternative vectorial TV definition

$$\|f\|_{TV_J} := \int_{\Omega} \sqrt{\lambda_-(\mathbf{x})} d\mathbf{x} = \int_{\Omega} \sigma_1(Df(\mathbf{x})) d\mathbf{x}, \quad (3.11)$$

where $\sigma_1(Df(\mathbf{x}))$ is the largest singular value Df at \mathbf{x} .

It is worth noting that definitions (3.10) and (3.11) admit a dual definition, just like in the case of the standard definition of the TV for scalar functions. This property makes

these definitions attractive from an implementation point of view, as it allows for the application of efficient minimization techniques in the context of image de-noising [51].

A different family of vectorial TV definitions are those which treat the image $f : \mathbb{R}^2 \rightarrow \mathbb{R}^m : \mathbf{x} \mapsto (f^1(\mathbf{x}), f^2(\mathbf{x}), \dots, f^m(\mathbf{x}))$ as a vector-valued function with m channels, and compute the TV semi-norm of the image as a combination of the scalar TV semi-norms of each channel. In particular, these TV semi-norms are given by [17]

$$\|f\|_{TV_\alpha} = \sqrt[\alpha]{\sum_{i=1}^m \|f^i\|_{TV}^\alpha}, \quad (3.12)$$

with $\alpha = 1$ or 2 . $\|f\|_{TV_2}$ is known as the *color TV* of f , and it was introduced in [17]. This definition offers desirable properties in the context of color image de-noising, whereby channels are coupled in right proportions to result in a balanced smoothing effect in each channel, while allowing discontinuities in each of them [17]. On the other hand, the semi-norm $\|f\|_{TV_1}$ is commonly referred to as the *channel-by-channel TV*. This definition offers computational advantages in the context of optimization methods for TV regularized problems due to its lack of coupling between the channels. Unfortunately, the absence of coupling between the channels also results in undesired properties of this definition in the context of color image de-noising [17]. Despite these pitfalls of the $\|f\|_{TV_1}$ semi-norm, several reasons support its appropriateness in the context of HARDI image de-noising. On the one hand, the channel-by-channel TV definition offers computational advantages due to its separability, as will be seen in Section 3.5. Moreover, despite the fact that this TV definition disregards any type of coupling between the image channels, further angular regularization constraints specifically designed for HARDI signals can be used to overcome this flaw, as shown in Section 3.5.

3.3 Angular Regularization

The spatial regularization imposed on HARDI signals by means of channel-by-channel ℓ_1 TV regularization disregards any type of regularity/coupling across different channels (i.e., diffusion encoded images). However, according to the fundamental physical models of water diffusion [61, 23], one expects an arbitrary HARDI signal to be a smooth spherical function at any arbitrary spatial coordinate. This premise is further supported by the fact that the resolution in MRI, which is typically around $2 \text{ mm} \times 2 \text{ mm} \times 2 \text{ mm}$, allows one to obtain only estimates of the average signal contribution from diffusion taking place over a group of axonal fibers, as opposed to the single contributions arising from each

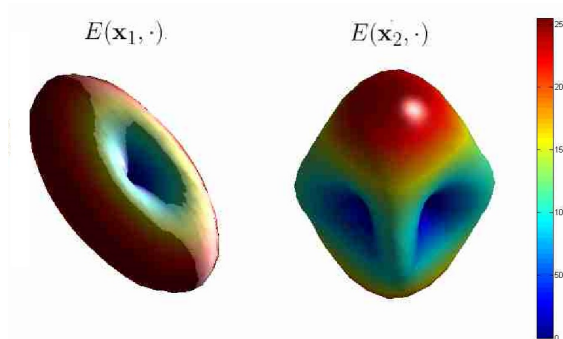


Figure 3.2: Two sample spherical signals arising from the restriction of the diffusion signal E at spatial locations $\mathbf{x}_1, \mathbf{x}_2 \in \Omega_{\mathbf{x}} \subset \mathbb{R}^3$. The value of $E(\mathbf{x}, \mathbf{u})$ is given by the distance from the origin to the surface in the direction $\mathbf{u} \in \mathbb{S}^2$. The figure illustrates that the HARDI signals showcase smooth behaviour over their angular domain.

fiber separately. As a result, as noted in [72], even though a HARDI signal may appear to be nearly discontinuous over \mathbb{S}^2 around a single predominating fiber, it is unlikely for the acquired signal to showcase this behaviour. Thus, HARDI signal channels are highly correlated in principle. This fundamental fact, which can be appreciated in Figure 3.2, motivates the angular regularity assumption for HARDI signals, which we state as follows:

Angular Regularity Assumption: The function $E_c(\mathbf{x}, \cdot) : \mathbb{S}^2 \rightarrow \mathbb{R}^+$ is a smooth function over \mathbb{S}^2 for all $\mathbf{x} \in \mathbb{R}^3$, so that each E^i in (3.2) is a sampled version of a smooth function over \mathbb{S}^2 .

In order to incorporate this angular regularity prior in the proposed methodology, we follow the approach of [36]. To this end, let $y_{n,l}$ be the real valued spherical harmonic (SH)² of degree n and order l and $Y_{n,l} \in [y_{n,l}(\mathbf{u}_1)y_{n,l}(\mathbf{u}_2) \cdots y_{n,l}(\mathbf{u}_K)]^T \in \mathbb{R}^K$. Furthermore, let \mathcal{Y} be the matrix whose columns are the vectors $Y_{n,l}$ and \mathbf{c}^i be the vector of SH coefficients corresponding to a truncated real SH expansion of E^i up to the degree $n = n_{max}$ inclusive,

²See, e.g., [78, 46] for detailed discussions on spherical harmonics in general and in the context of HARDI.

so that

$$\mathcal{Y} = \begin{pmatrix} y_{0,0}(\mathbf{u}_1) & y_{1,-1}(\mathbf{u}_1) & y_{1,0}(\mathbf{u}_1) & \cdots & y_{n_{max},n_{max}}(\mathbf{u}_1) \\ y_{0,0}(\mathbf{u}_2) & y_{1,-1}(\mathbf{u}_2) & y_{1,0}(\mathbf{u}_2) & \cdots & y_{n_{max},n_{max}}(\mathbf{u}_2) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y_{0,0}(\mathbf{u}_K) & y_{1,-1}(\mathbf{u}_K) & y_{1,0}(\mathbf{u}_K) & \cdots & y_{n_{max},n_{max}}(\mathbf{u}_K) \end{pmatrix}, c^i = \begin{pmatrix} c_{0,0}^i \\ c_{1,-1}^i \\ c_{1,0}^i \\ \vdots \\ c_{n_{max},n_{max}}^i \end{pmatrix}. \quad (3.13)$$

Under the angular regularity assumption, and using the notation in (3.2), one obtains that for each $i = 0, 1, \dots, |\Omega|$ and for large enough n_{max}

$$E^i \approx \mathcal{Y}c^i = \sum_{n=0}^{n_{max}} \sum_{|l| \leq n} c_{n,l}^i Y_{n,l}, \quad (3.14)$$

where $c_{n,l}^i$ is the SH coefficient of E^i corresponding to the real SH of order l and degree n . It is important to note that only even degree SHs are involved in the truncated expansion in (3.14), given that E is an even function with respect to its angular coordinate (i.e., $E(\mathbf{x}, -\mathbf{u}) = E(\mathbf{x}, \mathbf{u}), \forall (\mathbf{x}, \mathbf{u}) \in \Omega_{\mathbf{x}} \times \Omega_{\mathbf{u}}$). Thus, since there are $(n+1)/2$ possible SH orders associated with each SH degree n , and only even degrees are involved in the expansion, it follows that $\mathcal{Y} \in \mathbb{R}^{K \times P}$, where $P = (n_{max} + 1)(n_{max} + 2)/2$. Note that (3.14) suggests that signals arising from the restriction of E to a fixed spatial location are well approximated by bandlimited, and therefore, smooth spherical functions.

In [36], the HARDI reconstruction is obtained voxel-wise, in which case each E^i is reconstructed separately. In particular, Descoteaux et al. obtain the SH coefficients of the HARDI reconstruction by solving the regularized least-squares problem

$$\operatorname{argmin}_{c^i \in \mathbb{R}^K} \{ \|\mathcal{Y}c^i - E^i\|_F^2 + \lambda \|\Lambda c^i\|_F^2 \}, \quad (3.15)$$

where $\Lambda \in \mathbb{R}^{P \times P}$ is the diagonal matrix whose k^{th} entry is $n^2(n+1)^2$, where n is the degree of the SH corresponding to the k^{th} coefficient from c^i . The closed form solution to (3.15) is simply given by $c_{opt}^i = (\mathcal{Y}^T \mathcal{Y} + \lambda \Lambda)^{-1} \mathcal{Y}^T E^i$.

The method above offers great computational advantages and has been shown to perform remarkably well in the context of HARDI and ODF reconstruction, as shown in [36] and also in Chapter 4. It is worth mentioning that this approach corresponds to a least-squares SH fit, along with a low pass filter applied to the SH coefficients c^i , a fact which is easily perceived by noticing that higher degree coefficients are weighted more heavily than their lower degree counterparts in the cost function appearing in (3.15).

3.4 Separable Regularization

To combine the spatial and the angular regularization priors outlined in Sections 3.2 and 3.3, the spatial regularized spherical reconstruction (SR²) method minimizes the cost function

$$J(c) = \left\{ \frac{1}{2} \|\Upsilon(c) - E\|_2^2 + \overbrace{\frac{\lambda}{2} \|\Lambda(c)\|_2^2}^{\text{Angular regularization}} + \overbrace{\mu \|\Upsilon(c)\|_{TV_1}}^{\text{Spatial regularization}} \right\}, \quad (3.16)$$

over $\mathbb{R}^{P \times |\Omega_{\mathbf{x}}|}$, where

- $c \in \mathbb{R}^{P \times |\Omega_{\mathbf{x}}|}$ is the matrix of SH coefficients of the entire signal E , so that the i^{th} column of c corresponds to c^i ;
- $\Upsilon : \mathbb{R}^{P \times |\Omega_{\mathbf{x}}|} \rightarrow \mathbb{R}^{N \times M \times L \times K}$ maps the SH coefficients c of a HARDI signal E to its corresponding SH linear combination. This implies, e.g., that if $E^i = \mathcal{Y}c^i$ for each $i = 1, 2, \dots, |\Omega_{\mathbf{x}}|$, then $E = \Upsilon(c)$;
- $\Lambda : \mathbb{R}^{P \times |\Omega_{\mathbf{x}}|} \rightarrow \mathbb{R}^{P \times |\Omega_{\mathbf{x}}|}$ is as defined in Section 3.3;
- $\lambda, \mu \geq 0$ are tuning parameters which control the desired degrees of angular and spatial regularization, respectively;
- $\|\Upsilon(c) - E\|_2^2 = \sum_{i=1}^{|\Omega_{\mathbf{x}}|} \|\mathcal{Y}c^i - E^i\|_2^2$;
- $\|\Lambda(c)\|_2^2 = \sum_{i=1}^{|\Omega_{\mathbf{x}}|} \|\Lambda c^i\|_2^2$;
- $\|\cdot\|_{TV_1}$ is given by (3.12); and
- the $\|\cdot\|_2$ norms are computed by arranging the input argument into a one-dimensional vector and computing its ℓ^2 -norm.

Under the assumption of the measurement noise being additive white Gaussian noise (AWGN), the minimizer of J is optimal in a maximum-a-posteriori (MAP) sense upon interpreting the two regularity conditions as priors on the unknown signal. As mentioned in Section 2.1.6, the noise present in MRI signals follows a Rician distribution, and as such, the estimate above ceases to be MAP-optimal.

Even though Rician noise closely resembles Gaussian additive noise behaviour at high values of SNR [62], some studies suggest the need for effective algorithms derived under

a Rician contamination model. Multiple methods of this kind have been proposed for in the context of MRI reconstruction [37, 71, 31]. Unfortunately, these methods often involve significant computational burdens and/or ignore either the spatial or the angular regularity of the signals. For this reason, we will still advocate the use of the above functional, thereby implicitly assuming an AWGN contamination model, while acknowledging the potential improvements in performance that may arise by modifying the above functional according to a Rician noise contamination model, as was done in [38], for instance.

3.5 Numerical Implementation

The minimization problem in (3.16) involves *cross-domain* constraints (i.e., constraints both in the spatial and angular domains of the HARDI signal), making it difficult to solve directly from the analytical and computational point of view. However, if considered one at a time, both constraints are applied in a separable fashion (i.e., either in a voxel-by-voxel or in a channel-by-channel fashion), and as a result it is possible to employ the variable splitting techniques of classical optimization theory (see, e.g., [38, 50, 41, 1]) to simplify the complexity of this particular problem. For the problem at hand, we employ the alternating direction method of multipliers (ADMM) (the reader is referred to [18] for further details about ADMM and related minimization techniques).

As a preliminary step to the derivation of the algorithm, the unconstrained problem in (3.16) can be replaced by an equivalent constrained problem of the form

$$\begin{aligned} \min_{c,u} \left\{ \frac{1}{2} \|\mathbf{Y}(c) - E\|_2^2 + \frac{\lambda}{2} \|\Lambda(c)\|_2^2 + \mu \|u\|_{TV_1} \right\} \\ \text{s.t. } \mathbf{Y}(c) - u = 0, \end{aligned} \quad (3.17)$$

where u is an auxiliary optimization variable, and the optimization is carried over $c \in \mathbb{R}^{P \times |\Omega_{\mathbf{x}}|}$ and $u \in \mathbb{R}^{N \times M \times L \times K}$. The above problem can be readily solved using the method of augmented Lagrangian multipliers, which suggests an iterative update of c and u according to

$$\begin{aligned} (c^{t+1}, u^{t+1}) = \arg \min_{c,u} \left\{ \frac{1}{2} \|\mathbf{Y}(c) - E\|_2^2 + \frac{\lambda}{2} \|\Lambda(c)\|_2^2 + \right. \\ \left. + \mu \|u\|_{TV_1} + \frac{\delta}{2} \|\mathbf{Y}(c) - u + p^t\|_2^2 \right\} \\ p^{t+1} = p^t + (\mathbf{Y}(c^{t+1}) - u^{t+1}), \end{aligned} \quad (3.18)$$

where t stands for the iteration index, $\delta > 0$ is a predefined constant (e.g., $\delta = 0.5$) and $p^t \in \mathbb{R}^{N \times M \times L \times K}$ denotes the vector of (augmented) Lagrange multipliers at iteration t .

Finally, the optimization over c and u can be superseded by minimization with respect to c and u sequentially, in which case one can “split” the above problem into two simpler sub-problems, namely

$$\min_c \left\{ \frac{1}{2} \|\mathbf{Y}(c) - E\|_2^2 + \frac{\lambda}{2} \|\Lambda(c)\|_2^2 + \frac{\delta}{2} \|\mathbf{Y}(c) - (u^t - p^t)\|_2^2 \right\} \quad (3.19)$$

and

$$\min_u \left\{ \frac{\delta}{2} \|u - (\mathbf{Y}(c^{t+1}) + p^t)\|_2^2 + \mu \|u\|_{TV_1} \right\}, \quad (3.20)$$

giving rise to two important observations:

- The optimization in (3.19) is *separable in the spatial domain*, as it can be solved on a voxel-by-voxel basis using

$$[c^{t+1}]^i = (\mathcal{Y}^T \mathcal{Y} + \lambda \Lambda^T \Lambda)^{-1} \mathcal{Y}^T \left(\frac{E^i + \delta([u^t - p^t]^i)}{1 + \delta} \right), \quad (3.21)$$

for each $i = 1, 2, \dots, |\Omega_{\mathbf{x}}|$.

- The optimization in (3.20) is *separable in the diffusion domain*, as it can be solved on a channel-by-channel basis using

$$[u^{t+1}]_k = \arg \min_{u_k \in \mathbb{R}^{N \times M \times L \times K}} \left\{ \frac{1}{2} \|u_k - (\mathbf{Y}(c^{t+1}) + p^t)_k\|_2^2 + \frac{\mu}{\delta} \|u_k\|_{TV_1} \right\}, \quad (3.22)$$

for each $k = 1, 2, \dots, K$. Note that solving the above problem amounts to solving a sequence of scalar TV de-noising problems, which, as mentioned in Section 3.2, can be efficiently solved using methods such as the one in [25].

Thus, the proposed algorithm alternates the (separable) minimizations in (3.21) and (3.22), followed by updating the Lagrange multipliers according to (3.18). The iterations are terminated once the relative difference between subsequent solutions drops below a predefined tolerance (e.g., $\leq 0.1\%$) or until a fixed number of iterations is reached. In our experiments, an average number of required iterations to reach such a tolerance (with $p^0 = u^0 = 0$ and $\delta = 0.5$) has been observed to be 20. The pseudocode for the proposed method is shown in Table 3.1

Proposed Algorithm: SR²

Input: E (noisy data), $\delta, \lambda, \mu, NIT$ (number of iterations), $tol, tolTV$
Output: \hat{E} (de-noised signal)
foreach $i \in \{1, 2, \dots, |\Omega_{\mathbf{x}}|\}$ **and** $k \in \{1, 2, \dots, K\}$ **do**
 | $u_k \leftarrow 0_{N \times M \times L}$
 | $c^i \leftarrow 0_P$
end
 $cntr \leftarrow 0$
 $\hat{E}_{old} \leftarrow E$
 $p \leftarrow 0_{N \times M \times L \times K}$
while $err > tol$ **or** $(cntr < NIT)$ **do**
 | **foreach** $i \in \{1, 2, \dots, |\Omega_{\mathbf{x}}|\}$ **do**
 | i^{th} column of $c \leftarrow c^i$
 | **end**
 | **foreach** $k \in \{1, 2, \dots, K\}$ **do**
 | k^{th} $N \times M \times L$ volume of $u \leftarrow u_k$
 | **end**
 | **foreach** $i \in \{1, 2, \dots, |\Omega_{\mathbf{x}}|\}$ **do**
 | $c^i \leftarrow (\mathcal{Y}^T \mathcal{Y} + \lambda \Lambda)^{-1} \mathcal{Y}^T \left(\frac{E^i + \delta([u-p]^i)}{1+\delta} \right)$
 | **end**
 | **foreach** $k \in \{1, 2, \dots, K\}$ **do**
 | $u_k \leftarrow tv((\mathcal{Y}(c) + p)_k, tolTV, \lambda)^*$
 | **end**
 | $p \leftarrow p + (\mathcal{Y}(c) - u)$
 | $\hat{E} \leftarrow \mathcal{Y}(c)$
 | $err \leftarrow \text{norm}(\hat{E}_{old} - \hat{E})$
 | $\hat{E}_{old} \leftarrow \hat{E}$
 | $cntr++$
end

*: $tv(\cdot, \cdot, \cdot)$ stands for the implementation of [25].

Table 3.1: Pseudocode for solving the problem in (3.17).

It is also important to note that the degree of angular regularity imposed in the signal obtained is a function of n_{max} , the highest SH order included in the truncated SH expansion of the reconstruction. Thus, for instance, the higher the value of n_{max} , the “milder” the smoothness regularity constraints which are imposed on the reconstructions, as its corresponding truncated SH signal would allow for the presence of higher frequency signal components. It is therefore important to choose n_{max} not to induce oversmoothing effects in the reconstruction obtained. As an immediate implication of this fact, since $K \geq P = (n_{max} + 2)(n_{max} + 1)/2$ must be satisfied for the computation of the solution of each step (3.21), it is imperative to obtain enough diffusion-encoded directions in order to avoid oversmoothing effects. In the experiments presented in Chapter 4, setting $n_{max} = 8$ has been shown to produce high quality reconstructions, in which case $K \geq 45$.

Finally, it is worth noting that the SR^2 method and that introduced in [74] share multiple characteristics, but their relative effectiveness differ, being dependent on a specific application at hand. Thus, the method of [74] was intended to be used for reconstructing the diffusion signals from their sub-critical measurements. In such a case, the sparseness of signal representation (as provided through the use of spherical ridgelets) is essential for attaining useful reconstruction results. The method proposed herein, on the other hand, requires fully sampled data, while allowing one to bypass the computationally extensive procedure of sparse approximation, thereby substantially improving the efficiency of the reconstruction procedure.

Chapter 4

Results

In Chapter 3, two de-noising schemes for HARDI data were introduced, viz. the $TV_{\mathcal{M}}$ -based ROF model from Section 3.2.2 and the SR^2 method from Section 3.4. In this chapter, the superiority of the SR^2 algorithm over several alternative denoising methods is demonstrated by means of quantitative performance measures and visual means using both simulated and real-life HARDI data sets. It is concluded that SR^2 outperforms all the other alternative methods across different levels of noise and over different types of complex diffusion profiles. Additionally, it is observed that the performance of the $TV_{\mathcal{M}}$ -based ROF model from Section 3.2.2 is inferior as compared to that of most alternative HARDI de-noising methods, supporting the hypothesis presented in Section 3.2.2.

4.1 In Silico Experiments

In this section, a comparative study between the methods from Sections 3.2.2 and 3.4 and several other state-of-the art HARDI de-noising methodologies is presented. In particular, the study is based on quantitative measures and employs simulated HARDI data sets. The alternative methods used for this study are: 1) the constrained TV-denoising approach of [64], 2) the unconstrained vectorized TV-denoising approach of [19], and 3) the Tikhonov-type SH fitting procedure of [36]. It should be noted that all the algorithms can be applied to both the HARDI data E or to the ADC data $ADC = (-1/b) \log(E)$, as a result of the fact that these two signals possess the same dimensionality. Moreover, according to (2.29), there is a one to one relationship between these two signals, and therefore they can be obtained from one another.

The reference methods of [64] and [19] have been observed to produce more accurate reconstructions when applied to the ADC signals, as opposed to the original HARDI data E . For convenience, the above algorithms will be referred to as TV1 and TV2, respectively. We note that, although using different assumptions regarding the nature of measurement noises, both TV1 and TV2 would be conceptually equivalent to the SR^2 method had we decided to ignore the signal regularity over the \mathbf{u} coordinate by setting $\lambda = 0$ in (3.16). On the other hand, the reference method of [36] has been applied to both E and the ADC signal, with the resulting reconstruction procedures referred below to as SH1 and SH2, respectively. In this case also, it deserves mentioning that both SH1 and SH2 would have been equivalent to the proposed method, had we decided to ignore the spatial regularity of the diffusion data by setting $\mu = 0$ in (3.16). Finally, the $TV_{\mathcal{M}}$ -based method of Section 3.2.2 will be referred to as TV3. For ease of reference, Table 4.1 presents the names of all the algorithms used for this study, along with their acronyms and input data types.

2D In Silico Phantom A 2D set of simulated HARDI data was generated in the form of a 16×16 array of spherical HARDI signals obeying a Gaussian mixture model (GMM) with a spatially dependent number of “crossing fibers” of various orientations [112]. In particular, the model of the HARDI signal adopted assumes that the normalized HARDI signal is of the form

$$E(\mathbf{x}, \mathbf{u}) = \frac{1}{N(\mathbf{x})} \sum_{i=1}^{N(\mathbf{x})} e^{-b \mathbf{u}^T D_i(\mathbf{x}) \mathbf{u}}, \quad (4.1)$$

where $N(\mathbf{x}) \in \{1, 2, 3\}$ and each $D_i(\mathbf{x}) \in \mathbf{S}_{++}^3$ for each spatial coordinate \mathbf{x} . In this experiment, each $D_i(\mathbf{x})$ was designed to have eigenvalues $\{\lambda_i(\mathbf{x})\}_{i=1}^3$ with $\lambda_1(\mathbf{x}) > \lambda_2(\mathbf{x}) = \lambda_3(\mathbf{x})$. In this case, each spherical signal in the sum of (4.1) is said to have *cylindrical symmetry*¹. Geometrically, each term in the sum of (4.1) corresponds to a neuronal fiber whose direction is parallel to the main axis of $D_i(\mathbf{x})$. The latter is the direction of the eigenvector corresponding to the largest eigenvalue of $D_i(\mathbf{x})$.

Under this GM model, one may parametrize the HARDI signal in (4.1) in terms of: 1) its b -value (one for the entire signal), 2) the set of fiber orientations determined by all $D_i(\mathbf{x})$'s, and 3) the fractional anisotropy (FA) and mean diffusivity (MD) ($MD(D_i) = tr(D_i)/3$) indices corresponding to each $D_i(\mathbf{x})$. For this particular experiment, the parameters of the GMM were chosen so as to mimic a diffusion signal corresponding to a 16×16 region of interest supporting two “fiber bundles” crossing at a right angle with an additional

¹ The sets of eigenvalues from two different $D_i(\mathbf{x})$'s can be different, as long as the cylindrical symmetry condition is met in both cases

| Acronym | Input Signal | Reference | Equivalent to SR^2 when |
|---------|--------------|------------|------------------------------|
| TV1 | ADC | [64] | $\mu \neq 0, \lambda = 0$ |
| TV2 | ADC | [19] | $\mu \neq 0, \lambda = 0$ |
| TV3 | E | (proposed) | N/A |
| SH1 | E | [36] | $\mu = 0, \lambda \neq 0$ |
| SH2 | ADC | [36] | $\mu = 0, \lambda \neq 0$ |
| SR^2 | ADC | (proposed) | $\lambda \neq 0, \mu \neq 0$ |

Table 4.1: Acronyms, references, and input signal types corresponding to the various methods under comparison.

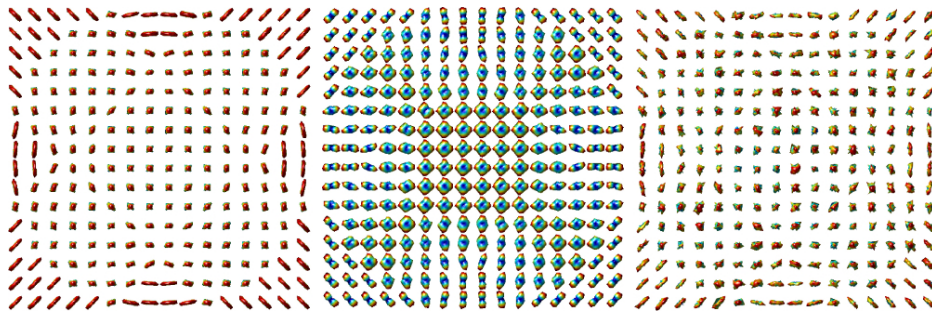


Figure 4.1: (left) Phantom for in silico experiments (E), (center) ADC profile of the phantom on the left (ADC), (right) Noisy (SNR=8) version of phantom on the left.

“bundle” traversing the plane in a circular pattern. The b -value was set to 2500 s/mm^2 , while the MD and the FA indices of individual fibers comprising the Gaussian mixtures were set at random from the ranges $[0.6, 0.9] \cdot 10^{-3}$ and $[0.65, 0.85]$, respectively. All the simulated signals were corrupted by different levels of Rician noise, giving rise to signal-to-noise (SNR) values in the range $[4, 20]$, where the SNR is given by the ratio between the maximum value of the normalized HARDI signal, E , and the scale (commonly denoted by σ) parameter of the Rician noise (see (2.17)). The number of sampling points (diffusion-encoding directions) was set to $K = 64$ (a practical amount which suggests $n_{max} = 8$). The sample points were distributed over the sphere using an equiangular scheme similar to that employed in [58]. The original and sample noisy versions of the phantom, along with the ADC profile of the noise-free phantom and are shown in Figure 4.1. Additionally, Figure 4.2 illustrates the composition of the phantom. For the purpose of quantitative comparison, the normalized mean-squared error (NMSE) metric has been employed. The

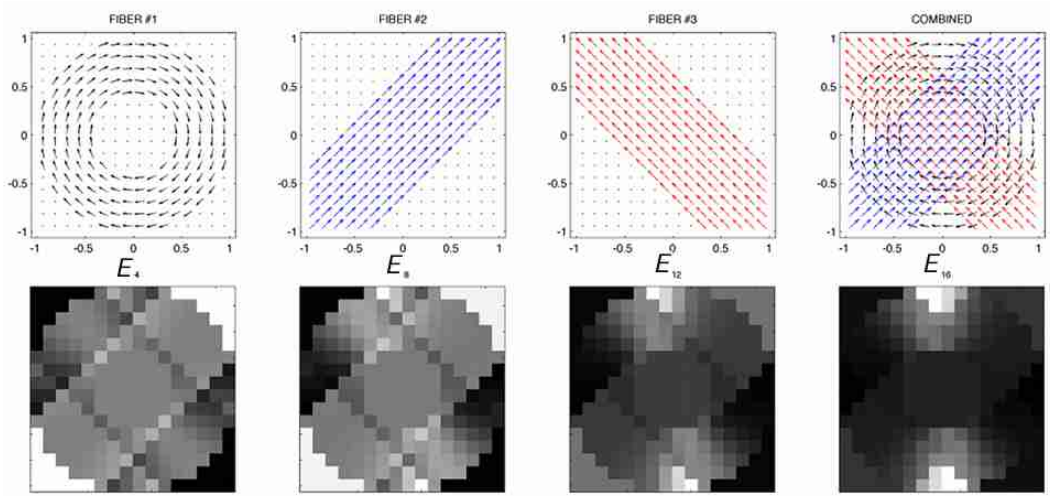


Figure 4.2: In silico phantom construction. (a-d) show vector fields whose orientations at each spatial location are parallel to the axonal fibers in the phantom at that same location. (a) circular axonal fiber; (b-c) two $\frac{\pi}{4}$ rad axonal fibers crossing at a right angle; (d) superposition of the three axonal fibers in (a-c); (d-g) sample 2D diffusion encoded images.

NMSE between the underlying original signal, f , and its reconstruction, \hat{f} , is given by

$$\text{NMSE}(f, \hat{f}) := \sqrt{\frac{\sum_{(n,m,l,k)} [f(n, m, l, k) - \hat{f}(n, m, l, k)]^2}{\sum_{(n,m,l,k)} f^2(n, m, l, k)}}. \quad (4.2)$$

The regularization parameters of the algorithms have been optimized so as to produce the smallest possible NMSE at every given level of SNR for a specific realization of the noise.

It should be noted that despite the rather intuitive interpretation offered by the NMSE metric in (4.2), this metric may not always adequately quantify the effects or the size of estimation artifacts or noises [118]. As a result, a different performance metric called the *structural similarity* (SSIM) index, has been proposed in [118]. Unfortunately, to the best of my knowledge, there is not a unique, clear direct extension of the standard SSIM index definition to HARDI signals. As a result, this metric has not been employed for the purpose of the quantitative performance study presented herein.

Sample reconstruction results for the various algorithms for two different levels of SNR are shown in Figure 4.3 and Figure 4.4. It can be seen from these images that the reconstructions by means of SR^2 resemble the original data more closely than those obtained

| SNR | Raw | TV1 | TV2 | TV3 | SH1 | SH2 | SR ² |
|-----------|--------|--------|--------|--------|--------|--------|-----------------|
| 4 | 0.4619 | 0.2802 | 0.2345 | 0.3036 | 0.2729 | 0.2305 | 0.2285 |
| 8 | 0.2326 | 0.1657 | 0.1704 | 0.1685 | 0.1261 | 0.1179 | 0.1063 |
| 12 | 0.1555 | 0.1292 | 0.1525 | 0.1229 | 0.0845 | 0.0815 | 0.0767 |
| 16 | 0.1169 | 0.1075 | 0.0900 | 0.0984 | 0.0650 | 0.0631 | 0.0610 |
| 20 | 0.0937 | 0.0908 | 0.0768 | 0.0821 | 0.0532 | 0.0518 | 0.0507 |

Table 4.2: NMSE produced by various compared methods. Each entry represents an ensemble average over 20 independent trials.

using the alternative methods. Additionally, Table 4.2 summarizes the NMSE values attained by the different algorithms at all the different levels of SNR considered. The results confirm the superior performance of the SR² method over the rest of the algorithms. It is also clear from this table that the performance of TV3 is significantly poorer than that of most other methods. The relatively poor performance of TV3 may be attributed to the sub-optimality of the piece-wise smooth model for HARDI signals when restricted to their angular domain, as mentioned in Section 3.2.2.

4.2 Real-Life Experiments

In addition to the experiments based on computer simulations, the performance of the algorithms have been assessed qualitatively by using real life data. In particular, two real-life diffusion data sets have been used for this purpose: 1) the data set from the FiberCup diffusion phantom reported in [44](available at www.lnao.fr/spip.php?article112) with the isotropic spatial resolution of 3 mm³, $b = 2000$ s/mm² and $K = 64$, and 2) an in vivo dMRI data set of a human brain from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database.

4.2.1 The FiberCup Phantom Experiments

The FiberCup phantom consists of hydrophobic acrylic fibers whose diameters are in the order of magnitude of the diameter of myelinated axons. These fibers were bundled together tightly, and the resulting bundles were oriented and stacked upon each other so as to mimic different types of fiber crossing configurations, resulting in an approximate fiber density of 1900 fibers/mm². In order to acquire the diffusion data, a cylindrical container was used

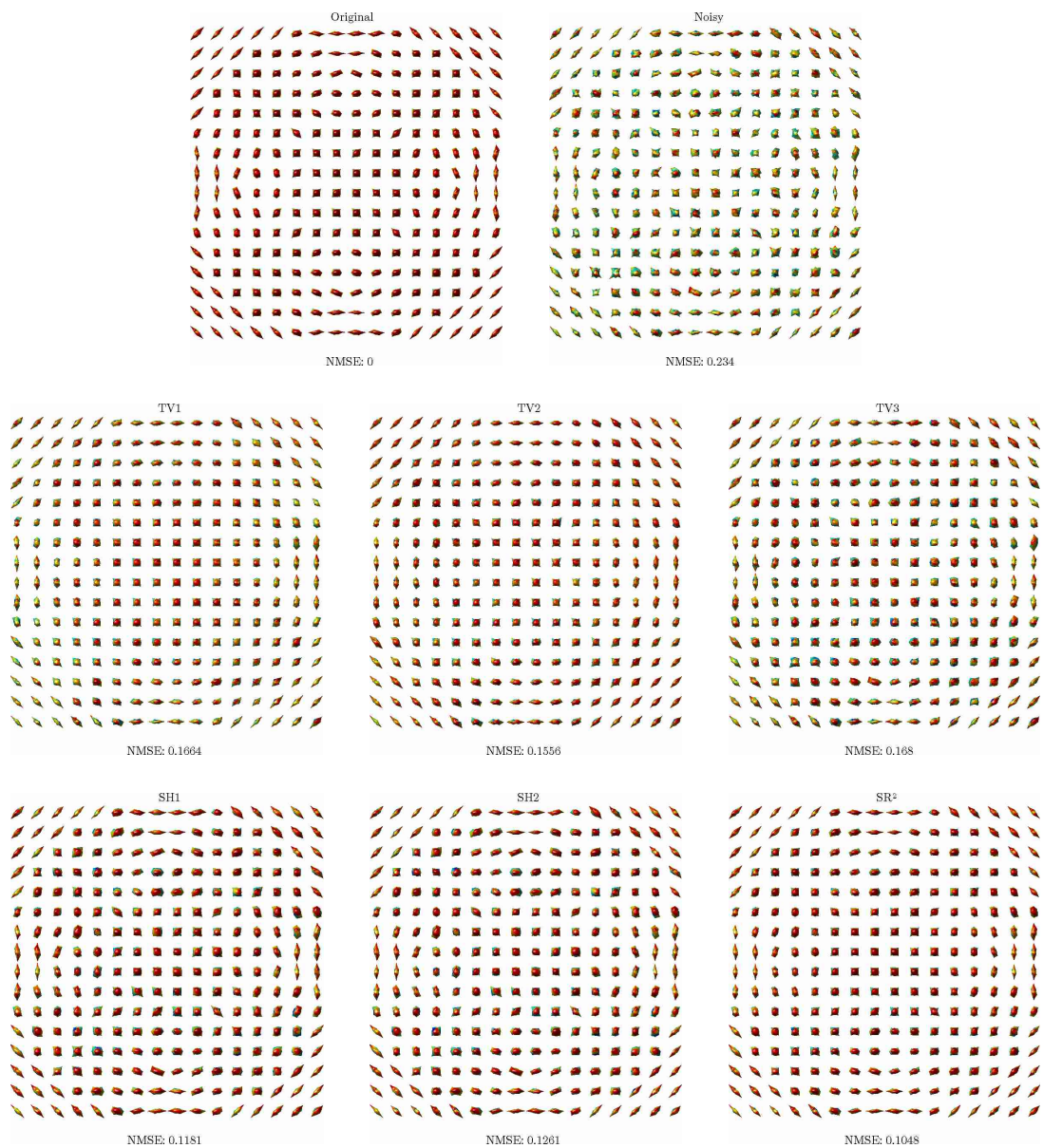


Figure 4.3: Sample reconstructions using various algorithms for the case $\text{SNR}= 8$.

to hold the phantom inside an MRI scanner. An echo-planar imaging (EPI) acquisition scheme was employed for obtaining the data using a single-shot diffusion-weighted, twice refocused pulse sequence. Some additional design details as well as detailed parameters for acquisition can be found in [91, 44]. The b_0 -image of the FiberCup phantom, as well as its

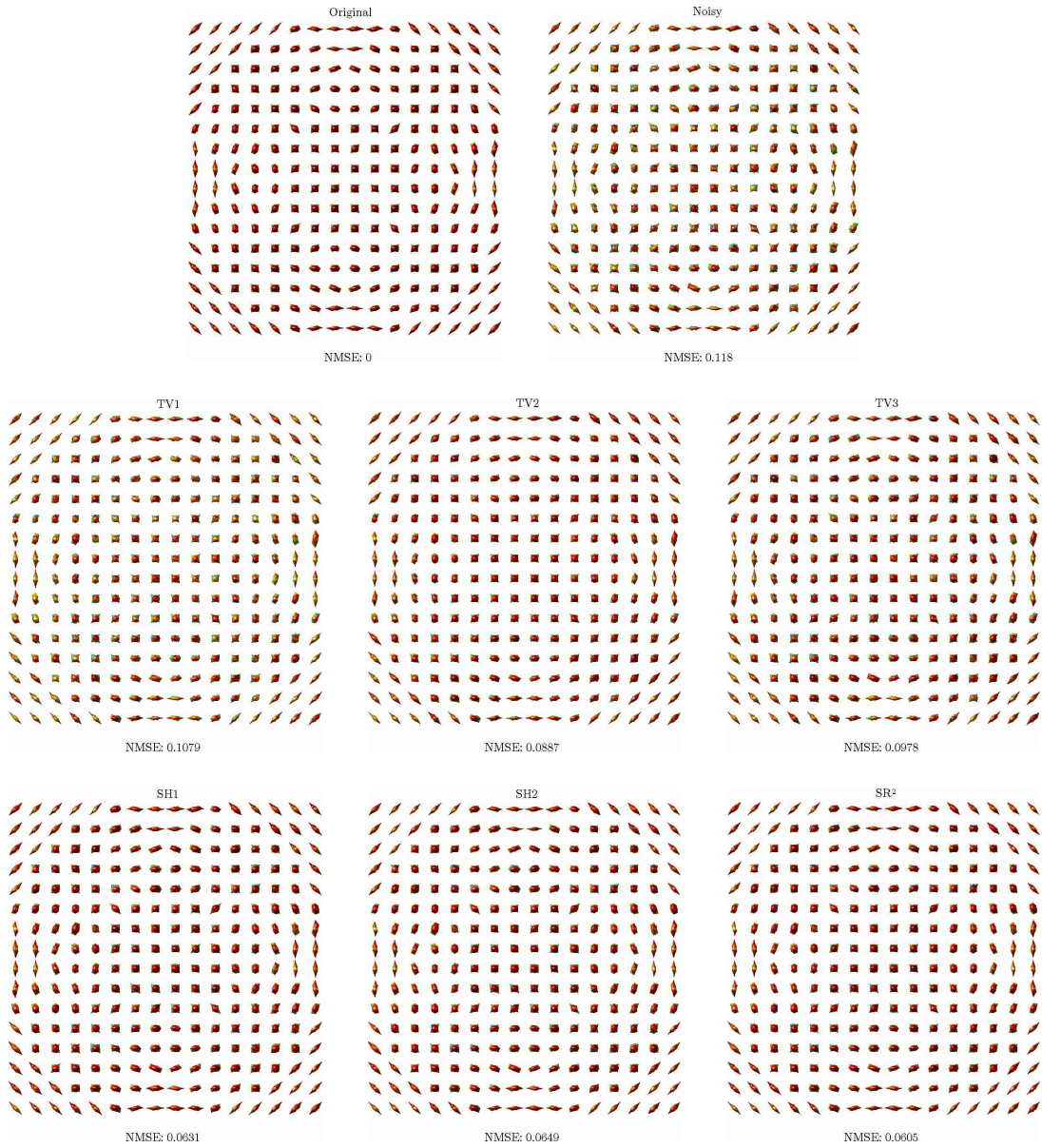


Figure 4.4: Sample reconstruction using various algorithms for the case $\text{SNR}= 16$.

diffusion data are illustrated in Figure 4.4. The three regions of the phantom indicated by the red boxes in Figure 4.5 correspond to sections showcasing a region of “fiber branching”, a “sharp turn”, and a “fiber crossing.”

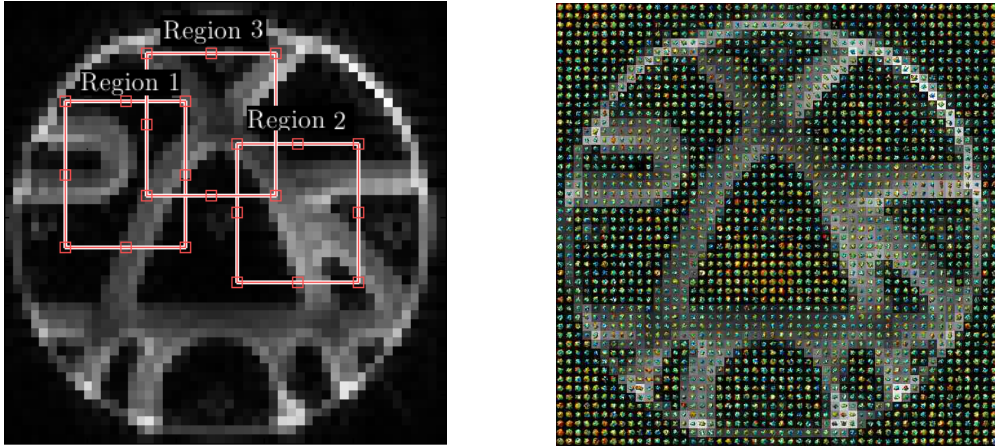


Figure 4.5: b_0 -image of the FiberCup phantom and its associated diffusion data. The three regions shown correspond to different diffusion profiles, region 1: sharp turn, region 2: fiber branching, and region 3: fiber crossing.

Based on the performance results from Section 4.1, the SH2 method has been chosen for the purpose of comparison against the SR^2 method, as it has shown to be the “second best” according to the simulation results. Figure 4.6 compares the performance of SH2 with that of SR^2 based on their associated dODFs [114] (see Section 2.4 for more details) computed based on the reconstructions from each of these algorithms, and subsequently applying the method presented in [36]. Notice that the background in these images corresponds to the b_0 -image of the phantom. In both cases, the regularization parameters of the algorithms were tuned to result in the most consistent reconstructions with respect to the known geometry of the phantom. As can be seen in Figure 4.6, the underlying diffusion profiles are much more accurately resolved via the dODFs estimated using SR^2 as compared to the case of SH2.

4.2.2 In Vivo Experiments

Concluding the experimental study, a real life dMRI data set of a human brain obtained from the Alzheimer’s disease neuroimaging initiative (ADNI) has been used to evaluate the performance of the proposed algorithm. The data set consists of a 41 diffusion-encoded 3D volumes (i.e., $K = 41$), each of which contains 59 axial slices acquired at $b = 1000$ s/mm² with a spatial resolution of 1.3672 mm \times 1.3672 mm \times 2.7 mm (for illustration purposes, only the results corresponding to a single axial slice will be presented). In particular,

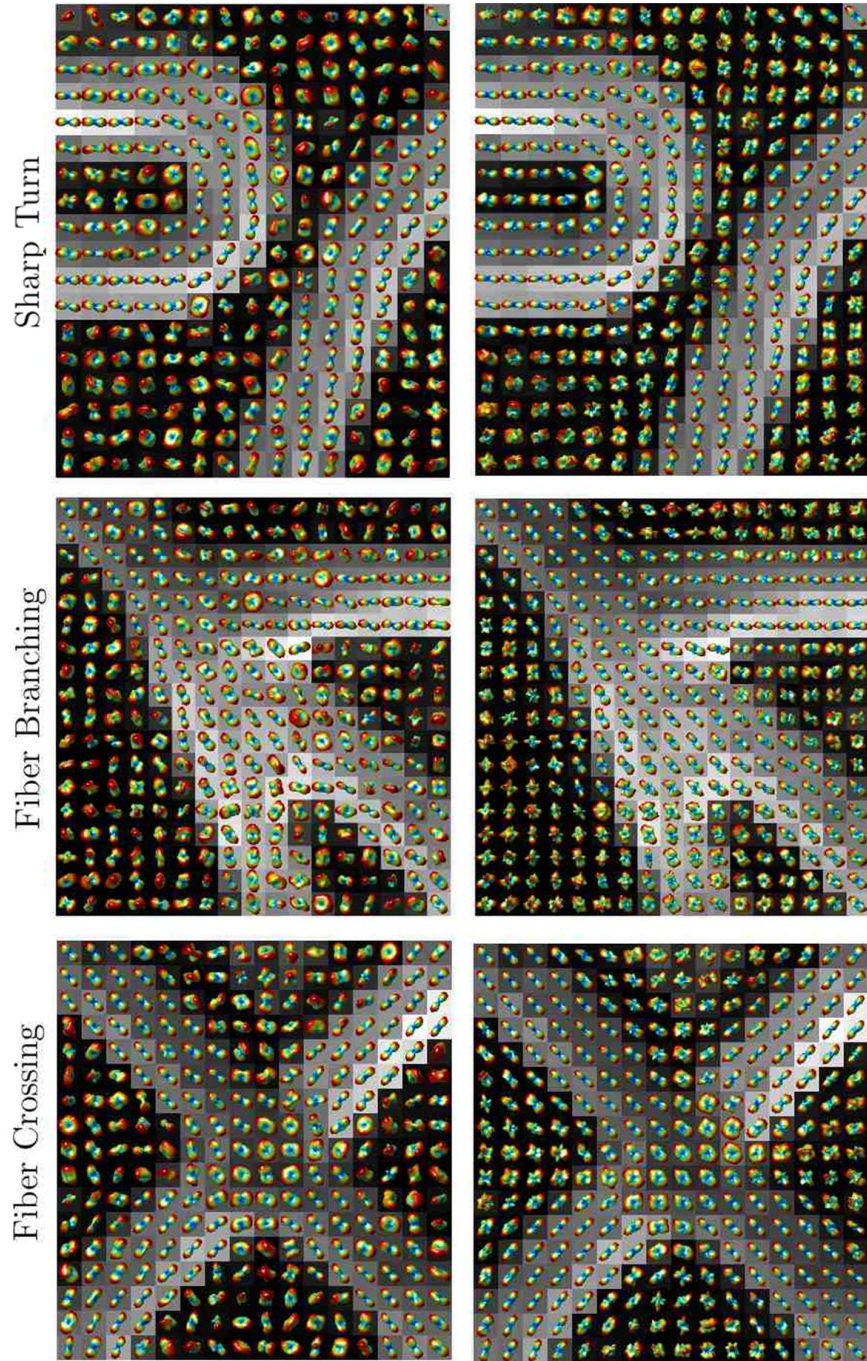


Figure 4.6: Real-life dODF reconstructions by means of SH2 (left) and of SR2 (right).

Figure 4.7 illustrates the b_0 -image, the estimated FA map (see Section 2.3), the noisy diffusion signal, and the reconstruction by means of SR^2 corresponding to the chosen axial slice of the human brain. The marked regions in the b_0 -image correspond to the close-up views presented in Figures 4.8 to 4.11 of the noisy and the reconstructed data. In the latter figures, the estimated FA map is presented in the background, and the reconstructed dODFs (see Section 2.4) are presented to assist in the identification of the diffusion profiles inferred from the method's reconstruction. The dODF estimates from the noisy data and from the reconstructed data have been obtained by direct application of the method of [36] with no regularization applied, i.e., by finding the best-fit SH coefficients to the HARDI data at each voxel in the least-squares sense.

The regularization effects of the SR^2 method can be observed from the sample spherical signals shown in Figure 4.7. Additionally, as can be seen in Figures 4.8 to 4.11, the obtained reconstruction by SR^2 clearly outlines the unimodal diffusion profiles associated with regions with high anisotropy indices, as indicated by the FA map estimate shown in the background.

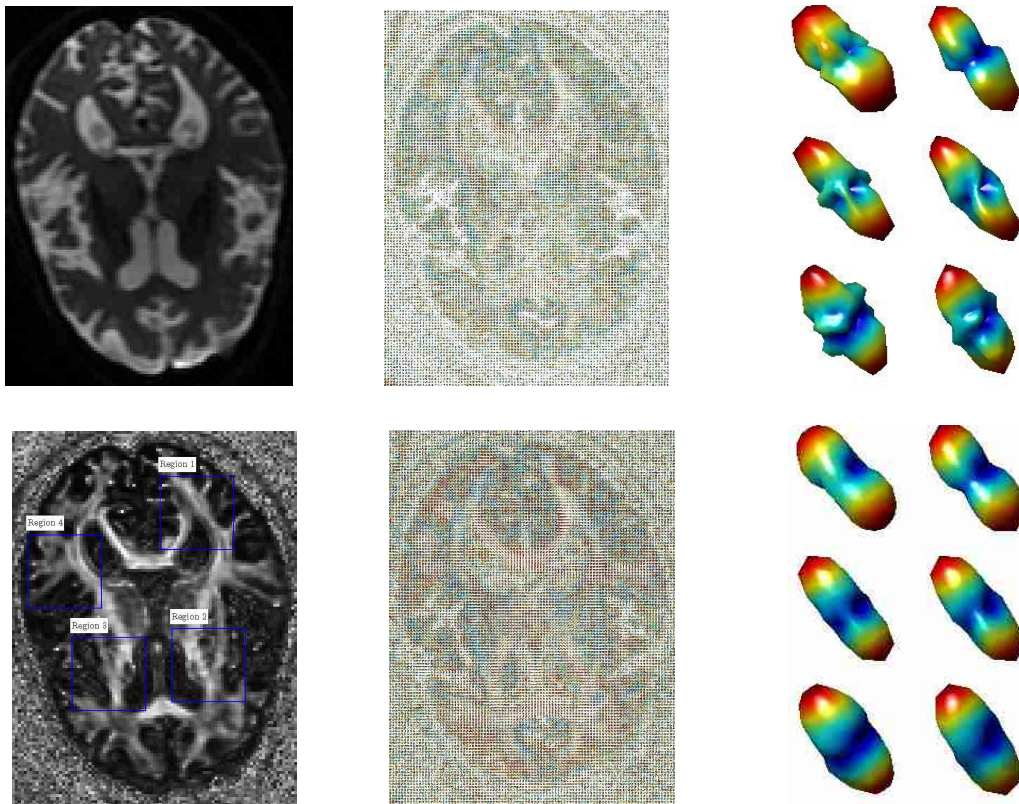
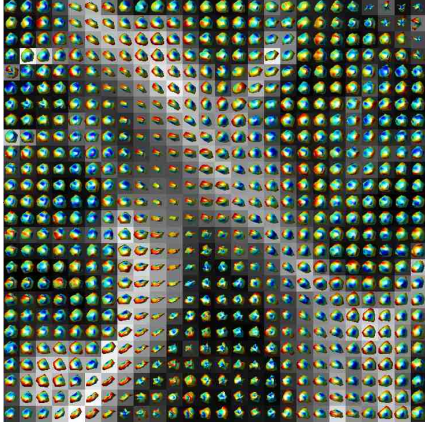
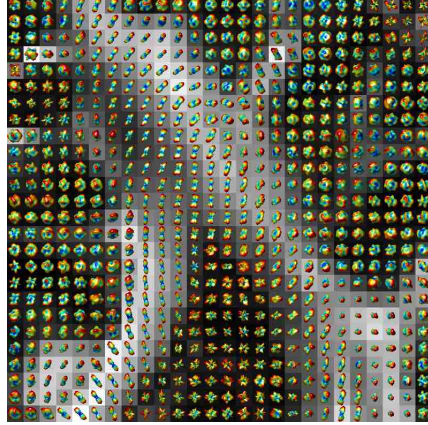


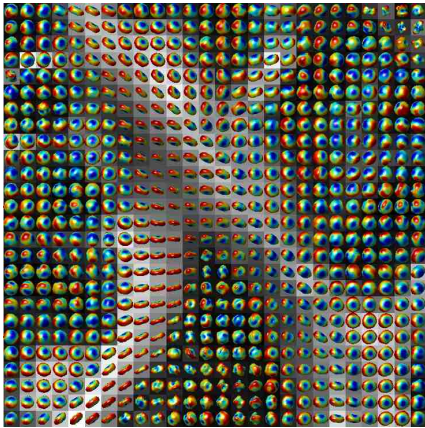
Figure 4.7: In vivo real-life dMRI data set and reconstruction by means of SR^2 . Left column: b_0 -image corresponding to an axial slice of the human brain (top) and FA map corresponding to the same slice (bottom). Center column: original (noisy) dMRI data (top) and reconstruction by means of SR^2 (bottom) corresponding to the axial slice on the left. Right column: sample dODF estimates from the noisy data (top), and their counterpart estimates achieved by means of SR^2 (bottom).



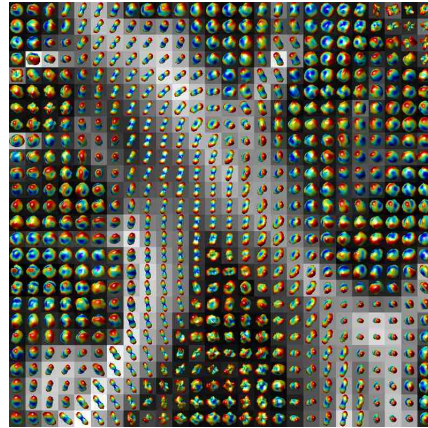
(a) Noisy HARDI data



(b) Noisy dODF estimate

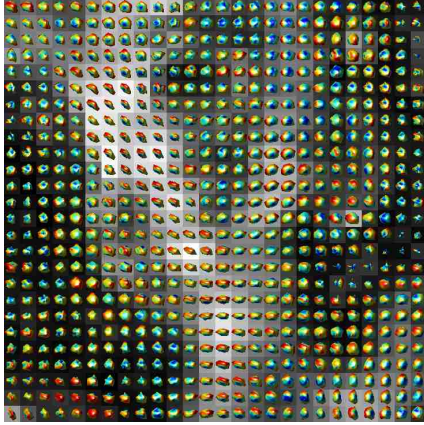


(c) SR² HARDI data estimate

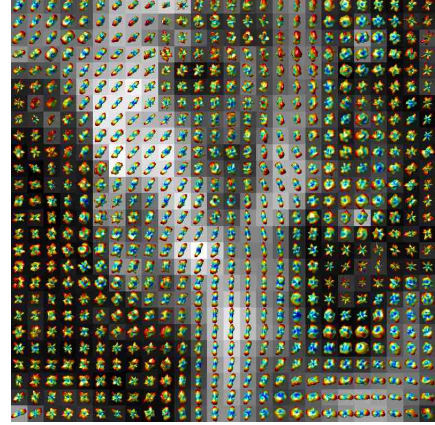


(d) SR² dODF estimate

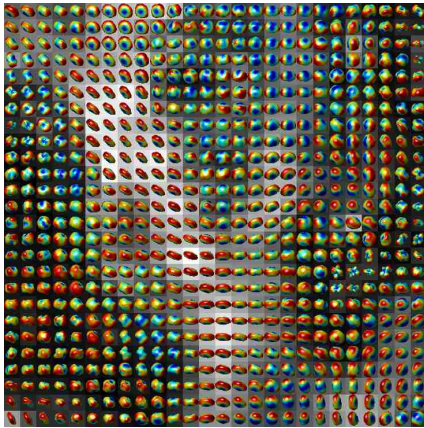
Figure 4.8: In vivo reconstruction results (region 1).



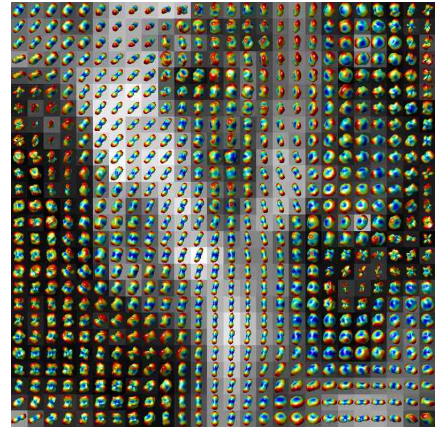
(a) Noisy HARDI data



(b) Noisy dODF estimate

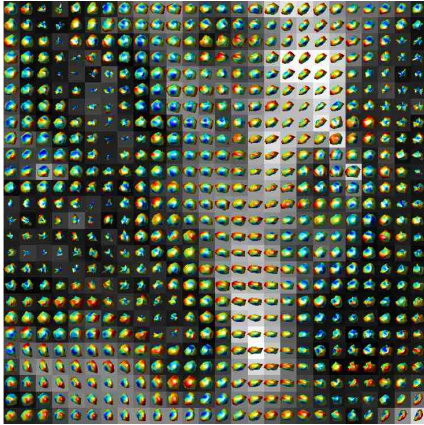


(c) SR^2 HARDI data estimate

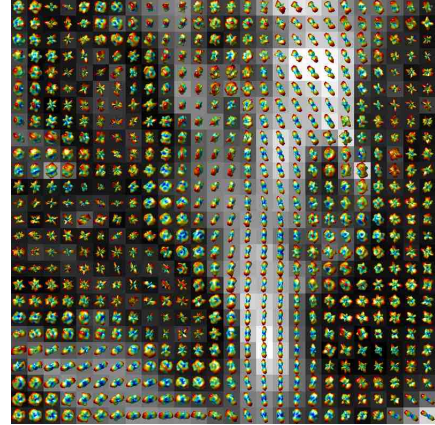


(d) SR^2 dODF estimate

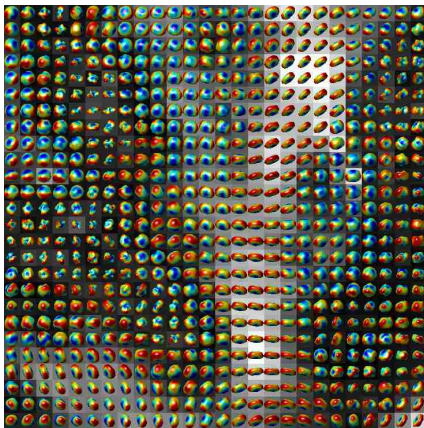
Figure 4.9: In vivo reconstruction results (region 2).



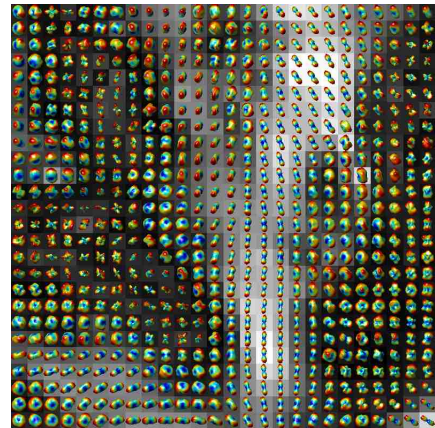
(a) Noisy HARDI data



(b) Noisy dODF estimate

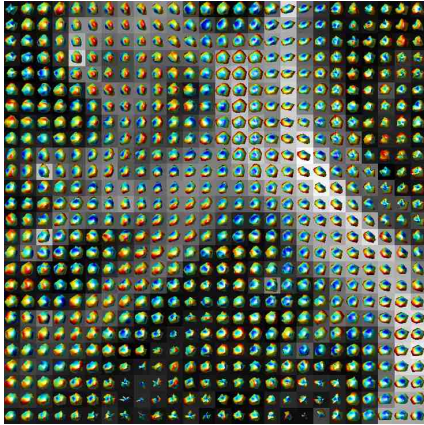


(c) SR² HARDI data estimate

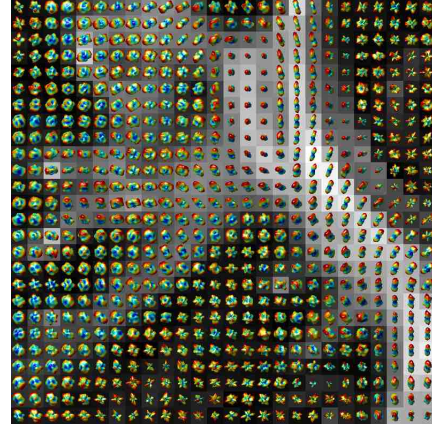


(d) SR² dODF estimate

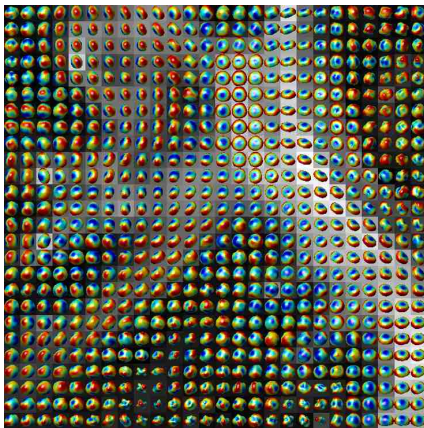
Figure 4.10: In vivo reconstruction results (region 3).



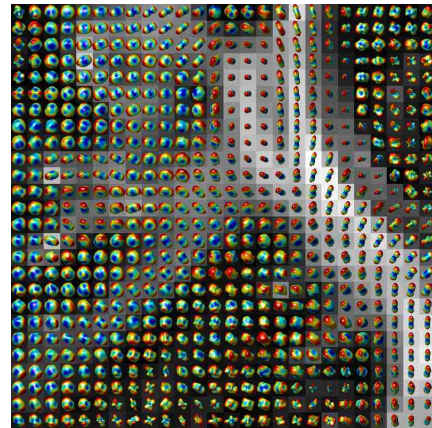
(a) Noisy HARDI data



(b) Noisy dODF estimate



(c) SR^2 HARDI data estimate



(d) SR^2 dODF estimate

Figure 4.11: In vivo reconstruction results (region 4).

Chapter 5

Summary and Conclusions

The purpose of this thesis was to propose a novel HARDI de-noising methodology known as spatially regularized spherical reconstruction (SR^2). As such, the theoretical underpinnings and the implementation details associated with this method have been presented in this work. Additionally, the performance of the algorithm has been tested against that of alternative HARDI de-noising methods using both simulated and real life HARDI data sets.

The apparent success of the proposed algorithm lies in its ability to exploit both the spatial and the angular regularity of HARDI signals, whose nature can be inferred from basic physics principles of water diffusion and through our knowledge of the anatomy of the human brain. Based on this information, the proposed algorithm uses a combination of: 1) the angular regularization scheme of [36] consisting of a regularized least squares SH fit procedure which amounts to a low pass filtering method in the domain of SH coefficients, and 2) a spatial regularization approach based on a TV minimization scheme applied to the diffusion-encoded images of the HARDI signals. The angular regularization procedure accepts a closed form solution separable with respect to the spatial coordinate of the HARDI signal (i.e., angular regularization is imposed to each voxel in the HARDI signal independently), therefore offering computational advantages over alternative iterative methods which can be used for the same purpose. On the other hand, the spatial regularization employs the channel-by-channel TV semi-norm definition, which renders the solution to the spatial regularization problem separable with respect to the angular coordinate of the HARDI signal. Thus, when imposed one at a time, both angular and spatial regularization schemes are separable in nature, while imposing both types of regularization on the HARDI signal simultaneously results in a cross-domain optimization problem. Nonetheless, employing the concept of variable splitting from convex optimiza-

tion, the cross-domain problem can be broken into smaller separable sub-problems which amount to one of the two regularization problems above, each accepting a parallelizable computational scheme. Despite the fact that no thorough performance study has been carried out to demonstrate the efficiency of the proposed algorithm, the separable nature of the computational scheme proposed suggests its potential computational advantages over alternative, non-separable implementations.

In addition to the apparent computational advantages offered by the proposed algorithm, the quality of its reconstructions has been shown to outperform that of alternative state-of-the-art methods both qualitatively and quantitatively using simulated and real life data sets. As a result, it has been concluded that the type of regularization scheme chosen appears adequate in the context of HARDI signal de-noising. In particular, it has been shown that combining the two regularization constraints simultaneously offers performance improvements over considering either one of them alone. Moreover, the use of TV regularization on HARDI signals has been further investigated by studying the performance of a TV regularization algorithm which employs a scalar TV semi-norm over the $\mathbb{S}^2 \times \mathbb{R}^3$ domain. This algorithm has been observed to be inferior to most alternative methods in terms of the quality of their reconstructions, suggesting the inadequacy of TV-based schemes for the purpose of angular regularization in the context of HARDI de-noising.

As a final remark, it has also been noted that the smoothness properties of the SR^2 method depend on the number of diffusion-encoded images (i.e., K) acquired. Therefore, it is important to choose K carefully as to avoid overly smoothing the reconstructions (with $K \geq 50$ being a reasonable minimum for $b \geq 2000$ s/mm²).

Chapter 6

Future Work

Direct extensions of the proposed methodology include modifications of (3.16) by the inclusion of positivity constraints on the HARDI signal and/or incorporation of compensation terms for the Rician statistics of the noise present in HARDI images (e.g., as in [38]). Additionally, here we discuss two potential variational dMRI estimation frameworks motivated by the understanding of the different types of HARDI signal regularization presented in Chapter 3. The first one applies TV regularization to fODFs over $\mathcal{M} := \mathbb{R}^3 \times \mathbb{S}^2$, and may be used for the purpose of spherical deconvolution. The second one incorporates a novel type of regularization for dMRI signals known as fiber continuity, and may be used for both de-noising of HARDI data as well as for spherical deconvolution of dODFs.

6.1 Spherical Deconvolution

A common problem of interest in HARDI is that of *spherical deconvolution*, e.g., [111]. In particular, one may model the fODF signal $f : \mathcal{M} \rightarrow \mathbb{R}^+$ as a blurred and noisy version of the b_0 -normalized HARDI data E at each voxel, in which case, the blurring effect can be modelled by means of a convolution operator \mathcal{H} . In particular, assuming an additive noise contamination model, $f(\mathbf{x}, \cdot) = \mathcal{H}\{E(\mathbf{x}, \cdot)\} + \eta = h * f(\mathbf{x}, \cdot) + \eta(\mathbf{x}, \cdot)$, $\forall \mathbf{x} \in \mathbb{R}^3$, where $h : \mathbb{S}^2 \rightarrow \mathbb{R}$ is a point spread function (PSF) (aka single fiber response) associated with \mathcal{H} and $\eta(\mathbf{x}, \cdot) : \mathbb{S}^2 \rightarrow \mathbb{R}$ is the noise present in the HARDI signal at spatial location \mathbf{x} .

In practice, an estimate of f from E can be obtained by finding the minimizer of the functional

$$\int_{\mathbb{S}^2} \int_U |\mathcal{H}\{f\}(\mathbf{x}, \mathbf{u}) - E(\mathbf{x}, \mathbf{u})|^2 d\eta(U) + J_r(f), \quad (6.1)$$

where $d\eta(U) = \sin(\theta) dx dy dz d\theta d\phi$, and $J_r(f)$ is a functional whose purpose is to incorporate some known regularity of the fODF. For instance, in [95], the authors employ a fiber continuity prior (see Section 6.2 for more details), while in [49] the authors use a second order regularization prior on the gradient of f . In view of the fact that the underlying fODF signals can, in principle, present an indefinite number of discontinuities over $\mathbb{R}^3 \times \mathbb{S}^2$, and based on the observations about the $\|\cdot\|_{TV_{\mathcal{M}}}$ presented in Chapters 2 and 3, it seems appropriate to use $J_r(f) = \|f\|_{TV_{\mathcal{M}}}$ in (6.1), in which case the minimizer of the associated cost functional can be found using standard TV-based deconvolution methods such as [27], but with the modified definition of the total variation semi-norm.

6.2 Fiber Continuity

An alternative spatial regularity constraint to the TV-based ones discussed in Section 2 is that of fiber continuity. The fiber continuity prior, introduced in [95], is expressed in terms of its associated cost functional

$$J_{FC}(f) = \int_{\mathbb{S}^2} \int_U |\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{u}) \cdot \mathbf{u}| d\eta(U), \quad (6.2)$$

where $\nabla_{\mathbf{x}}$ is the gradient operator in the spatial coordinate, i.e., $\nabla_{\mathbf{x}} f = (\partial_{x_1} f, \partial_{x_2} f, \partial_{x_3} f)$. Although fiber continuity regularization may be effective in the context of dODF estimation, if used on its own, any angular regularity of the underlying signals is ignored. In view of the conclusions drawn in Chapter 3, a possible alternative to the SR² method in the context of HARDI estimation is to replace its TV semi-norm penalizing term appearing in (3.16) by $J_{FC}(f)$, resulting in the associated cost functional

$$\int_{\mathbb{S}^2} \int_U (|\mathcal{H}\{f\}(\mathbf{x}, \mathbf{u}) - E(\mathbf{x}, \mathbf{u})|^2 + \mu |\Delta_{\mathbb{S}^2} \mathcal{H}\{f\}(\mathbf{x}, \mathbf{u})|^2) d\eta(U) + \lambda J_{FC}(f). \quad (6.3)$$

It goes without saying that additional constraints such as those regarding the positivity or the sparsity of the signals may improve the performance of the regularization schemes which arise from the minimization of the functionals in (6.1) and (6.3).

The applications above exemplify the potential of the regularization schemes presented in this dissertation in the context of dMRI reconstruction. Naturally, the same ideas can be applied to the reconstruction and enhancement of signals with similar domains to that of HARDI data, and can therefore be easily extended to such applications.

References

- [1] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo. Fast image recovery using variable splitting and constrained optimization. *IEEE Transactions on Image Processing*, 19(9):2345–2356, September 2010.
- [2] I. Aganj, C. Lenglet, N. Jahanshad, E. Yacoub, N. Harel, P. Thompson, and G. Sapiro. A Hough transform global probabilistic approach to multiple subject diffusion MRI tractography, 2010.
- [3] D. Alexander, G. Barker, and S. Arridge. Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.
- [4] C. S. Anand and J. S. Sahambi. Wavelet domain non-linear filtering for MRI denoising. *Magnetic Resonance Imaging*, 28(6):842–861, 2010.
- [5] Yaniv Assaf and Peter J. Basser. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain, 2005.
- [6] H. Assemlal, D. Tschumperle, and L. Brun. Fiber tracking on HARDI data using robust ODF fields. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 3, pages III – 133–III – 136, 2007.
- [7] H. Assemlal, D. Tschumperlé, and L. Brun. Robust variational estimation of PDF functions from diffusion MR signal. In *Computational Diffusion MRI Workshop held at the Medical Image Computing and Computer-Assisted Intervention*, pages 1–8, 2008.
- [8] H. Assemlal, D. Tschumperlé, and L. Brun. Evaluation of q-space sampling strategies for the diffusion magnetic resonance imaging. In *Proceedings of MICCAI '09*, pages 406–414, 2009.

- [9] Haz-Edine Assemlal, David Tschumperlé, Luc Brun, and Kaleem Siddiqi. Recent advances in diffusion MRI modeling: Angular and radial reconstruction. *Medical Image Analysis*, 15(4):369–396, 2011.
- [10] L. Bao, W. Liu, Z. Pu, L. Fanton, S. Rapacchi, P. Croisille, Y. Zhu, and I. Magnin. Sparse representation based MRI denoising with total variation. In *IEEE ICSP'08*, page in press, 2009.
- [11] P. Bao and L. Zhang. Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. *IEEE Transactions on Medical Imaging*, 22(9):1089–1099, September 2003.
- [12] P. Basser. Relationships between diffusion tensor and q-space MRI. *Magnetic Resonance in Medicine*, 47:392–397, 2002.
- [13] P. Basser and D. Jones. Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR Biomed.*, 15:456–467, 2002.
- [14] P. Basser and C. Pierpaoli. A simplified method to measure the diffusion tensor from seven MR images. *Magnetic Resonance in Medicine*, 39:928–934, 1998.
- [15] S. Basu, T. Fletcher, and R. Whitaker. Rician noise removal in Diffusion Tensor MRI. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages 117–125, 2006.
- [16] J. Berman, M. Lanza, L. Blaskey, J. Edgar, and T. Roberts. High angular resolution diffusion imaging probabilistic tractography of the auditory radiation. *American Journal of Neuroradiology*, 34(6), 2003.
- [17] P. Blomgren and T. Chan. Color tv: total variation methods for restoration of vector-valued images. *IEEE Trans. Image Proc.*, 7(3):304–309, 1998.
- [18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, 2011.
- [19] X. Bresson and T.F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Problems and Imaging*, 2(4):455–484, 2008.

- [20] A. Buades, B. Coll, and J. M. Morel. A non local algorithm for image denoising. In *IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2005*, volume 2, pages 60–65, 2005.
- [21] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation (SIAM interdisciplinary journal)*, 4(2):490–530, 2005.
- [22] T. Bulow. Spherical diffusion for 3D surface smoothing. In *3DPVT*, pages 449–459, 2002.
- [23] T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford University Press on Demand, 1993.
- [24] H. Y. Carr and E. M. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.*, 94:630–638, 1954.
- [25] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imag. Vis.*, 20:89–97, 2004.
- [26] A. Chambolle, V. Caselles, M. Novaga, D. Cremers, and T. Pock. An introduction to Total Variation for Image Analysis. 2009.
- [27] T. Chan and C. Wong. Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3):370–375, 1998.
- [28] S. Chen, Z. Kang, X. Hu, B. Hu, and Y. Zou. Diffusion tensor imaging of the brain in patients with Alzheimer’s disease and cerebrovascular lesions. *J Zhejiang Univ Sci B*, 8(4):242–7, 2007.
- [29] Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B. Vemuri, and Y. Liu. Estimation, smoothing, and characterization of apparent diffusion coefficient profiles from high angular resolution DWI. In *CVPR*, 2004.
- [30] Z. Chen, P. Ni, J. Zhang, Y. Ye, H. Xiao, G. Qian, S. Xu, J. Wang, X. Yang, J. Chen, B. Zhang, and Y. Zeng. Evaluating ischemic stroke with diffusion tensor imaging. *Neurol Res*, 30(7):720–6, 2008.
- [31] R.A. Clarke, P. Scifo, G. Rizzo, F. Dell’Acqua, G. Scotti, and F. Fazio. Noise correction on Rician distributed data for fibre orientation estimators. *IEEE Transactions on Medical Imaging*, 27(9):1242–1251, 2008.

- [32] D. G. Cory. Measurement of translational displacement probabilities by NMR: An indicator of compartmentation. *Magnetic Resonance in Medicine*, 14(3):435–444, 1990.
- [33] J. Crank. *The Mathematics of Diffusion*. Oxford Science Publications. [Eng] Clarendon Press, 1979.
- [34] C. A. Deledalle, F. Tupin, and L. Denis. Patch similarity under non Gaussian noise. In *Proceedings of International Conference on Image Processing*, 2011.
- [35] Charles-Alban Deledalle, Loïc Denis, and Florence Tupin. How to compare noisy patches? Patch similarity beyond Gaussian noise. *International Journal of Computer Vision*, 99(1):86–102, Aug. 2012.
- [36] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical Q-ball imaging. *Magn. Res. Med.*, 58(3):497–510, 2007.
- [37] S. Dolui, I. Patarroyo, and O. Michailovich. A new similarity measure for non-local means filtering of MRI images. *Journal of Visual Communication and Image Representation*, 24(7):1040–1054, 2013.
- [38] S. Dolui, I.C. Salgado Patarroyo, O.V. Michailovich, and Y. Rathi. Reconstruction of HARDI using compressed sensing and its application to contrast HARDI. In *Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012 IEEE Workshop on*, pages 17–23, 2012.
- [39] D. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions in Information Theory*, 47(7), 2001.
- [40] P. Douek, R. Turner, J. Pekar, and N. Patronas D. Le Bihan. MR color mapping of myelin fiber orientation. *J. Comput. Assist. Tomogr.*, 15(6):923–9, 1991.
- [41] J. Eckstein. *Splitting Methods for Monotone Operators with applications to Parallel Optimization*. PhD thesis, Massachusetts Institute of Technology, 1989.
- [42] A. Fan, W. M. Wells, J.W. Fisher, M. Cetin, S. Haker, R. Mulkem, C. Tempany, and A. S. Willsky. A unified variational approach to denoising and bias correction in MR. *Inf. Process Med Imag*, 18:148–159, 2003.
- [43] M. Figueiredo, J. Bioucas-Dias, J. Oliveira, and R. Nowak. On total variation denoising: A new majorization-minimization algorithm and an experimental comparison with wavelet denoising. In *ICIP*, pages 2633–2636, 2006.

- [44] P. Fillard, M. Descoteaux, A. Goh, S. Gouttard, B. Jeurissen, J. Malcolm, A. Ramirez-Manzanares, M. Reisert, K. Sakaie, F. Tensaouti, T. Yo, J-F. Mangin, and C. Poupon. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. *NeuroImage*, 56(1):220–234, May 2011.
- [45] A. Foi. Noise estimation and removal in MR imaging: The variance-stabilization approach. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 1809–1814, 2011.
- [46] L. Frank. Anisotropy in high angular resolution diffusion-weighted MRI. *Magn. Reson. Med.*, 45(6):935–939, 2001.
- [47] J. Friedman, C. Tang, D. Carpenter, M. Buchsbaum, J. Schmeidler, L. Flanagan, S. Golembo, I. Kanellopoulou, J. Ng, P. Hof, P. Harvey, N. Tsopelas, D. Stewart, and K. Davis. Diffusion tensor imaging findings in first-episode and chronic schizophrenia patients. *Am J Psychiatry*, 165, 2008.
- [48] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11(2):221–232, June 1992.
- [49] A. Goh, C. Lenglet, Paul M. Thompson, and R. Vidal. Estimating Orientation Distribution Functions with Probability Density Constraints and Spatial Regularity. In *Medical Image Computing and Computer-Assisted Intervention MICCAI*, volume 5761, pages 877–885, 2009.
- [50] D. Goldfarb and S. Ma. Fast multiple-splitting algorithms for convex optimization. *SIAM Journal on Optimization*, 22(2):533–556, 2012.
- [51] B. Goldluecke, E. Strekalovskiy, and D. Cremers. The natural vectorial total variation which arises from geometric measure theory. 2012.
- [52] A. Guyton and J. Hall. *Textbook of medical physiology*. Elsevier Saunders, 11 edition, 2005.
- [53] E.M. Haacke. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley, 1999.
- [54] E. Hahn. Spin echoes. *Physical Review*, 80(4):580–594, 1950.
- [55] D. Healy and J. Weaver. Two applications of wavelet transforms in magnetic resonance imaging. *IEEE Transactions on Information Theory*, 38(2):840–860, March 1992.

- [56] B. Horwitz, K. Friston, and J. Taylor. Neural modeling and functional brain imaging: An overview. *Neural Networks*, 13:829–846, 2000.
- [57] J. Hsieh and SPIE (Society). *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE PM. SPIE, 2009.
- [58] P. Vandergheynst D. Van de Ville Y. Wiaux J. McEwen, G. Puy. Sparse image reconstruction on the sphere: implications of a new sampling theorem. *IEEE Transactions on image processing*, 22(6):2275–85, 2013.
- [59] K. Jansons and D. Alexander. Persistent angular structure: new insights from diffusion MRI data. *Inverse Prob*, 19:1031–46, 2003.
- [60] H. Johansen-Berg and T. Behrens. Just pretty pictures? what diffusion tractography can add in clinical neuroscience. *Curr. Opin. Neurol.*, 19(4):379–85, 2006.
- [61] H. Johansen-Berg and T. Behrens. *Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy*. Elsevier Science, 2009.
- [62] D. Jones and P. Basser. Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. *Magn. Reson. Med.*, 52(5):979–993, 2004.
- [63] DK Jones, A. Simmons, SC Williams, and MA Horsfield. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. *Magn Reson Med.*, 42(1):37–42, 1999.
- [64] Y. Kim, P. Thompson, A. Toga, L. Vese, and L. Zhan. HARDI denoising: Variational regularization of the spherical apparent diffusion coefficient sADC. In *Proceed. of IPMI*, pages 515–527, 2009.
- [65] K. Krissian and S. A. Fernández. Noise-driven anisotropic diffusion filtering of MRI. *IEEE Transactions on Image Processing*, 18(10):2265–2274, October 2009.
- [66] A. Kuurstra. *HARDI denoising using non-local means on the $\mathbb{R}^3 \times \mathbb{S}^2$ manifold*. MAsc. thesis, 2011.
- [67] Frank L. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magn. Reson. Med*, 47:1083–1099, 2002.
- [68] D Le Bihan, R Turner, C T Moonen, and J Pekar. Imaging of diffusion and micro-circulation with gradient sensitization: design, strategy, and significance. *J. Magn Reson Imaging*, 1(1):7–28, 1991.

- [69] M. Lysaker, A. Lundervold, and X. C. Tai. Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. *IEEE Transactions on Image Processing*, 12(12):1579–1590, December 2003.
- [70] J.J. Maller, R.H.S. Thomson, P.M. Lewis, S.E. Rose, K. Pannek, and P.B. Fitzgerald. Traumatic brain injury, major depression, and diffusion tensor imaging: making connections. *Brain research reviews*, 64(1):213–240, 2010.
- [71] J. V. Manjón, P. Coupé, A. Buades, D. L. Collins, and M. Robles. New methods for MRI denoising based on sparseness and self-similarity. *Medical Image Analysis*, 16:18–27, 2012.
- [72] T. McGraw, B. Vemuri, E. Ozarslan, Y. Chen, T. Mareci, and (Communicated by L. Vese). Variational denoising of diffusion weighted MRI. *Inverse Problems and Imaging*, 3(4):625648, 2009.
- [73] O. Michailovich and Y. Rathi. On approximation of orientation distributions by means of spherical ridgelets. In *ISBI*, pages 939–942, 2008.
- [74] O. Michailovich, Y. Rathi, and S. Dolui. Spatially regularized compressed sensing for high angular resolution diffusion imaging. *IEEE Trans. Med. Imaging*, 30(5):1100–1115, 2011.
- [75] D. Miglioretti, E. Johnson, A. Williams, and et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. *JAMA Pediatrics*, pages 1–8, 2013.
- [76] P. MomayyezSiahkal and K. Siddiqi. Rotation invariant completion fields for mapping diffusion MRI connectivity. In *Information Processing in Medical Imaging*, volume 22, pages 711–722, 2011.
- [77] S. Mori. *Introduction to Diffusion Tensor Imaging*. Elsevier Science, 2007.
- [78] C. Müller. *Spherical harmonics*. Lecture notes in Mathematics. Springer, 1966.
- [79] G. Houston J. Smith M. Smith M. James A. Parsons C. Huang L. Hall T. Carpenter N. Papadakis, D. Xing. A study of rotationally invariant and symmetric indices of diffusion anisotropy. *Magn. Reson. Imaging*, 17(6):881–92, 1999.
- [80] R.D. Nowak. Wavelet-based Rician noise removal for Magnetic Resonance Imaging. *IEEE Transactions on Image Processing*, 8(10):1408–1419, October 1999.

- [81] P.G.P. Nucifora, R. Verma, S. Lee, and E.R. Melhem. Diffusion-tensor MR imaging and tractography: exploring brain microstructure and connectivity. *Radiology*, 245(2):367–84, 2007.
- [82] E. Ozarslan, B. Vemuri, and T. Mareci. Generalized scalar measures for diffusion MRI using trace, variance and entropy. *Magn. Reson. Med.*, 53:866876, 2005.
- [83] RC McKinstry P. Mukherjee. Diffusion tensor imaging and tractography of human brain development. *Neuroimaging Clin N Am.*, 16(1):19–43, 2006.
- [84] S. Pajevic and C. Pierpaoli. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: Application to white matter fiber tract mapping in the human brain, 1999.
- [85] K. Pannek, J.L. Mathias, E.D. Bigler, G. Brown, J.D. Taylor, and S.E. Rose. An automated strategy for the delineation and parcellation of commissural pathways suitable for clinical populations utilising high angular resolution diffusion imaging tractography. *Neuroimage*, 50(3):1044–1053, 2010.
- [86] I. Patarroyo, S. Dolui, O. Michailovich, and E. Vrscaj. Reconstruction of HARDI data using a split Bregman optimization approach. In *Image Analysis and Recognition*, volume 7950, pages 589–596. 2013.
- [87] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
- [88] C. Pierpaoli, P. Jezzard, P. Basser, A. Barnett G., and Chiro. Diffusion tensor MR imaging of the human brain. *Radiology*, 201:637–648, 1996.
- [89] A. Pizurica, W. Philips, L. Lemahieu, and M. Acheroy. A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Transactions on Medical Imaging*, 22(3):323–331, March 2003.
- [90] A. Pizurica, A. M. Wink, E. Vansteenkiste, W. Philips, and J. B. T. M. Roerdink. A review of wavelet denoising in MRI and ultrasound brain imaging. *Current Medical Imaging Reviews*, 2(2):247–260, May 2006.
- [91] C. Poupon, B. Rieul, I. Kezel, M. Perrin, F. Poupon, and J. F. Mangin. New diffusion phantoms dedicated to the study and validation of high angular resolution diffusion imaging (HARDI) models. *Marg. Res. Med.*, 60(6):1276–1283, 2008.

- [92] J. Prince and J. Links. *Medical Imaging Signals And Systems*. Pearson Prentice Hall, 2006.
- [93] X. Qu, X. Cao, D. Guo, C. Hu, and Z. Chen. Compressed sensing MRI with combined sparsifying transforms and smoothed ℓ_0 norm minimization. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 626–629, 2010.
- [94] Y. Rathi, J. Malcolm, O. Michailovich, J. Goldstein, L. Seidman, R. McCarley, C. Westin, and M. Shenton. Biomarkers for identifying first-episode schizophrenia patients using diffusion weighted imaging. In *Proceedings of MICCAI'10*, pages 657–665, 2010.
- [95] M. Reisert and V. G. Kiselev. Fiber continuity: an anisotropic prior for ODF estimation. *IEEE Transactions on Medical Imaging*, 30(6):1274–1283, 2011.
- [96] P. Rousseeuw and A. Leroy. *Robust Regression and Outlier Detection*. Wiley, 2005.
- [97] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.
- [98] F. Rugg-Gunn, S. Eriksson, M. Symms, G. Barker, M. Thom, W. Harkness, and J. Duncan. Diffusion tensor imaging in refractory epilepsy. *Lancet*, 359(9319):1748–51, 2002.
- [99] G. Sapiro. Vector-valued active contours. In *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on*, pages 680–685, 1996.
- [100] P. Savadjiev, J. Campbell, G. Bruce Pike, and K. Siddiqi. 3D curve inference for diffusion MRI regularization and fibre tractography. *Medical Image Analysis*, 10(5):799–813, 2006.
- [101] P. Schaefer, P. Grant, and R. Gonzalez. Diffusion-weighted MR imaging of the brain. *Radiology*, 217(2):331–45, 2000.
- [102] D. Sinnaeve. The stejskal-tanner equation generalized for any gradient shape: an overview of most pulse sequences measuring free diffusion. *Concepts in Magnetic Resonance Part A*, 40A(2):39–65, 2012.
- [103] E. Stejskal and J. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. 1965.

- [104] E. O. Stejskal. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *The Journal of Chemical Physics*, 43(10):3597–3603, 1965.
- [105] P. Sundgren, Q. Dong, D. Gomez-Hassan, S. Mukherji, P. Maly, and R. Welsh. Diffusion tensor imaging of the brain: review of clinical applications. *Neuroradiology*, 46:339–350, 2004.
- [106] J. Tanner and E. Stejskal. Restricted Self-Diffusion of Protons in Colloidal Systems by the Pulsed-Gradient, Spin-Echo Method. *J. of Chemical Physics*, 49(4):1768–1777, 1968.
- [107] M. Termenon, Manuel Graña, A. Besga, J. Echeveste, and A. Gonzalez-Pinto. Lattice independent component analysis feature selection on diffusion weighted imaging for Alzheimer’s disease classification. *Neurocomputing*, 114:132–141, 2013.
- [108] T. Teuber and A. Lang. A new similarity measure for nonlocal filtering in the presence of multiplicative noise. *Computational Statistics and Data Analysis*, 56(12):3821–3842, 2012.
- [109] H. C. Torrey. Bloch equations with diffusion terms. *Phys. Rev.*, 104(3):563–565, 1956.
- [110] J. Tournier, F. Calamante, D. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23:1176–1185, 2004.
- [111] J-D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472, May 2007.
- [112] D. Tuch. *Diffusion MRI of complex tissue structure*. PhD thesis, 2002.
- [113] D. Tuch, R. Weisskoff, J. Belliveau, and V. Wedeen. High angular resolution diffusion imaging of the human brain. In *ISMRM*, 1999.
- [114] D. S. Tuch. Q-ball imaging. *Magn. Reson. Med.*, 52:1358–1372, 2004.
- [115] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med.*, 48:577–582, 2002.

- [116] A. Ulug and P. van Zijl. Orientation-independent diffusion imaging without tensor diagonalization: anisotropy definitions based on physical attributes of the diffusion ellipsoid. *Journal of Magnetic Resonance Imaging*, 9:804–813, 1999.
- [117] V. Varghees, M. Manikandan, and R. Gini. Adaptive MRI image denoising using total-variation and local noise estimation. In *Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on*, pages 506–511, 2012.
- [118] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4):600–612, April 2004.
- [119] J. B. Weaver, Y. Xu, D. M. Healy Jr., and L. D. Cromwell. Filtering noise from images with wavelet transforms. *Magnetic Resonance in Medicine*, 21(2):288–295, November 1991.
- [120] V. Wedeen, P. Hagmann, W. Tseng, T. Reese, and R. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn. Reson. Med.*, 54(6):1377–1386, 2005.
- [121] C. Westin, S. Peled, H. Gudbjartsson, R. Kikinis, and F. A. Jolesz. Geometrical diffusion measures for MRI from tensor basis analysis. In *ISMRM '97*, 1997.
- [122] N. Wiest-Daesslé, S. Sylvain, P. Coupé, S. Morrissey, and C. Barillot. Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI. In *Proceedings of MICCAI'07*, pages 344–351, 2007.
- [123] J. C. Wood and K. M. Johnson. Wavelet packet denoising of magnetic resonance images: Importance of Rician noise at low SNR. *Magnetic Resonance in Medicine*, 41(3):631–635, March 1999.
- [124] Y. Xu, J. B. Weaver, D. M. Healy, and J. Lu. Wavelet transform domain filters: A spatially selective noise filtration technique. *IEEE Transactions on Image Processing*, 3(6):747–758, November 1994.
- [125] Silvano Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics and Image Processing*, 33(1):116–125, January 1986.
- [126] M. Zhu, S. Wright, and T. Chan. Duality-based algorithms for total-variation-regularized image restoration. *Computational Optimization and Applications*, 2008.