Sparse Bayesian Learning for Spherical Deconvolution of Diffusion MRI Data

A dissertation presented
by
Jon Haitz Legarreta Gorroño
to
The Department of Physics

in partial fulfillment of the requirements
for the degree of
Master of Science
in the subject of
Medical Physics

Universidad Nacional de Educación a Distancia (UNED)
Madrid, Spain

September 2016

Advisors

Erick Jorge Canales Rodríguez, PhD
Fundació per a la Investigació i la Docència Maria Angustias Giménez (FIDMAG)
Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM)

Yasser Alemán Gómez, PhD
Laboratorio de Imagen Médica, Unidad de Cirugía y Medicina Experimental
Instituto de Investigación Sanitaria Gregorio Marañón (IISGM)
Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid
Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM)

Cristina María Santa Marta Pastrana, PhD
Departamento de Física Matemática y de Fluidos
Universidad Nacional de Educación a Distancia (UNED)

©2016 – Jon Haitz Legarreta Gorroño
all rights reserved.

Sparse Bayesian Learning for Spherical Deconvolution of Diffusion MRI Data

Abstract

Magnetic resonance diffusion imaging provides insights on the organizational microstructure of biological tissues. Fiber trajectories are determined by microscopic tissue heterogeneity. Regional differences in biological fiber orientation are revealed by water molecule diffusion displacement preference. This dependence of the diffusion coefficient on the direction in which it is measured is called anisotropy. Diffusion tensor imaging was the first technique proposed to model the underlying architecture and spatial orientation of the axonal fiber bundles in brain tissues. However, the assumptions of the tensor model on the homogeneous Gaussianity of the diffusion process in a voxel were found to fall short to model the complexity of fiber pathways in the human brain. Several alternatives have been proposed to overcome these limitations, yet the angular resolution of current state-of-the-art methods is insufficient for many applications.

In this thesis a new spherical deconvolution method to recover the fiber orientation distribution function from diffusion magnetic resonance imaging data is proposed. The new method does not require any assumptions on the underlying displacement probability density function and uses a Sparse Bayesian Learning approach to find the optimal solution. Results from synthetic data demonstrate that it is capable of resolving complex fiber crossings at lower angles than other state-of-the-art methods. The proposed method was also tested on real data. The superior ability to estimate the fiber bundle orientations may lead to improved results in whole brain fiber tractography and structural connectivity maps.

Keywords: diffusion magnetic resonance imaging (dMRI), spherical deconvolution (SD), sparse bayesian learning (SBL), compressed sensing (CS), intra-voxel orientational heterogeneity (IVOH), fiber orientation distribution function (fODF) reconstruction

Aprendizaje Bayesiano Disperso para la Deconvolución Esférica en Datos de IRM

Resumen

La imagen de difusión por resonancia magnética proporciona una representación de la disposición estructural de los tejidos. La heterogeneidad de los tejidos, aunque ésta sea apreciable únicamente a escala microscópica, determina la trayectoria que siguen las fibras. La preferencia en el desplazamiento por difusión de las moléculas de agua pone de manifiesto las diferencias regionales relativas a la orientación biológica de las fibras. Esta dependencia del coeficiente de difusión respecto a la dirección de medición se denomina anisotropía. La imagen mediante el tensor de difusión fue la primera técnica propuesta para modelar la arquitectura interna y la orientación espacial de las poblaciones de fibras axonales en el tejido cerebral. Sin embargo, las suposiciones del modelo tensorial sobre el carácter Gaussiano del proceso de difusión en cada vóxel presentan numerosas limitaciones a la hora de explicar la complejidad de las trayectorias de las fibras en el cerebro. Se han propuesto distintas alternativas para superar estas limitaciones, pero la resolución angular de los métodos existentes es aún insuficiente para numerosas aplicaciones.

Este trabajo propone un nuevo método para reconstruir la función de distribución de orientación de las fibras a partir de imágenes de difusión por resonancia magnética. Este nuevo método no realiza ningúna suposición acerca de la función de densidad de probabilidad de la orientación y emplea un algoritmo de aprendizaje Bayesiano disperso para obtener la solución óptima. Los resultados en datos sintéticos demuestran que el método es capaz de resolver configuraciones complejas de fibras con ángulos de cruce más pequeños que los que obtienen los métodos actuales. El método también ha sido aplicado a datos reales. La mejora en la capacidad de estimación de la orientación de las fibras podría conducir a resultados más precisos en tractografía de fibras y en mapas de conectividad estructural en el cerebro.

Palabras clave: imagen de difusión por resonancia magnética, deconvolución esférica, aprendizaje Bayesiano disperso, muestreado disperso, heterogeneidad orientacional intravóxel, reconstrucción de la función de distribución de orientación de las fibras

Contents

List of Figures

To my parents; aita, ama, zuei eskainita dago.

Acknowledgments

I wish to express my most sincere gratitude to many people who have in different ways contributed to the work presented in this Master's thesis.

First of all I would like to thank my advisors, Dr. Erick-Jorge Canales Rodríguez, Dr. Yasser Alemán Gómez, and Dr. Cristina María Santa Marta Pastrana. Thanks for your invaluable support, encouragement and guidance throught the process of conceiving, developing and writing this thesis dissertation. Thank you for introducing me to the exciting field of diffusion imaging. A special thank to Erick, with whom I have been working closely throughout this work, for your tremendous help. Thanks for always taking time for enlightening discussions, and for revealing the secrets of diffusion imaging. Your enthusiasm has been a source of inspiration. Thanks as well for detailed proofreading and writing advice.

Thanks to the colleagues of Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM) at Fundació per a la Investigació i la Docència Maria Angustias Giménez (FIDMAG) (Sant Boi de Llobregat, Barcelone) for your warm welcome during my stay in your lab. I hope we will meet again in the crossroads of fiber pathways!

Thanks to José Carlos Antoranz Callejo and the UNED University Medical Physics admissions committee for having given me the opportunity to pursue this Master's degree. Thanks for tons of advice over all these years. The path has proven to be thoroughly interesting, and the learning process enriching. It has been an amazing journey.

A special thanks to my parents. I am lucky to have you. Thank you for your understanding through the process. Understanding how research is like is not easy! This would not have been possible without your support.

Jon Haitz Legarreta Gorroño
September 2016

0

Introduction

0.1 Background

Magnetic Resonance Imaging (MRI) is a non-invasive and non-ionizing technique that provides excellent intensity contrast to study particular biological structures. Diffusion Magnetic Resonance Imaging (dMRI) is a technique where the diffusion of water molecules within tissues is measured through MRI excitation techniques. Diffusion phenomena refer to the Brownian motion described by particles within a medium. In the human body, diffusion is observed in different tissues, such as the brain, the muscles (e.g. the heart), and other tissues such as the kidneys, thanks to the transport of water molecules.

Diffusion, thus, can reveal the microstructure of a given tissue within the body at a macroscale. The diffusion trajectories are described by pathways, which provide a useful representation of how different parts of the tissue being imaged are connected. More specifically, these different parts may serve to distinct purposes, and hence, may constitute a functional region.

In the case of the brain, diffusion provides information about the directional structure of neural tracts found in the white matter and the cortex. Through the dMRI technique, a collection of T_2-weighted reference images, and a collection of diffusion-weighted images, acquired at different gradient directions, are acquired. The reconstruction of the diffusion information can be performed using a variety of methods that have been developed over the past three decades. Although the target function formulation can differ, the aim is to reconstruct the known fiber trajectories as faithfully as possible. The performance of the reconstruction techniques is affected by a number of parameters, such as the strength of the magnetic field, the number of gradients applied, or the achieved Signal-to-Noise Ratio.

At current resolution scales, multiple fibers are present in a voxel. In clinically available MRI scanners, a voxel may represent tissue volumes between $[1\times1\times1]$ and $[2\times2\times2]$ mm[39], while histology and microscopy studies estimate the diameters of fibers in the human brain being in the $[1-10]$ µm range[8]. Thus, intra-voxel fiber crossing, bending, kissing and branching configurations must be resolved in diffusion imaging.

0.2 Motivation

Diffusion magnetic resonance imaging provides a unique insight into the biological tissue microstructure. Anatomical structures where water molecule diffusion is preferentially

oriented in a given direction can be revealed thanks to this technique. Hence, the tissue properties can be inferred from the diffusion images, and be compared at different development stages or across different health states. White matter in the brain is one of the preferred target regions due to the restricted movement of water molecules along neurons' axonal pathways.

However, current diffusion orientation reconstruction methods account for a limited success in heterogeneous intra-voxel fiber configurations, namely fiber crossing, bending, twisting, fanning or kissing. Although numerous alternatives to the initial diffusion techniques, based on homogeneous Gaussian diffusion assumptions, have been proposed in literature, methods still find limitations in the lowest inter-fiber (crossing) angle they are capable of resolving (which is around 40°), and in regions with non-dominant fibers (i.e. voxels where a dominant fiber package's volume fraction is larger by a factor of two or greater than the non-dominant's).

This thesis is motivated by the analysis of the limitations of current state-of-the-art methods and the research of a method that could improve the estimation of the local fiber micro-geometry by resolving a range of intra-voxel configurations at significantly high angular resolutions. The technique proposed promotes a sparse solution to the fiber orientation distribution function within the spherical deconvolution framework. The method is found to be competitive under a variety of conditions, and is tested with both synthetic data, phantom and real datasets. The increased ability to disentangle complex white matter fiber pathways opens yet another possibility to improve the results of tractography algorithms to study the brain architecture.

0.3 Outline

This thesis is divided into three parts: in Part I (Chapter 1) the theoretical background on the imaging physics of Magnetic Resonance Imaging, and the diffusion phenomenon enabling diffusion Magnetic Resonance Imaging are revisited. A brief illustration of the fundamentals underpinning the diffusion signal acquisition, as well as a general perspective of the Diffusion Tensor Imaging (DTI) analysis framework, are also provided.

Part II (Chapters 2-3) builds upon the previous concepts to present a review of the different methods for the estimation of the fiber orientation distribution, as well as the associated mathematical tools (Chapter 2). This will prompt the development of the concepts, namely mathematical sparsity and Bayesian learning processes, involved by the proposed fiber orientation reconstruction method (Chapter 3). Chapter 4 includes the results of the method compared to other reconstruction algorithms, using synthetic datasets, and phantom and real whole-brain volumes. The discussion will follow the results in the same chapter.

Finally, Part III (Chapter 4) is dedicated to the conclusions drawn from the work presented, and the future research lines on the diffusion approach proposed in this thesis are presented.

0.4 Original Contributions

The work developed in this thesis contributes to the advance of the search of better methods for fiber orientation reconstruction in biological tissues.

The fundamental contributions to the state of the art in diffusion Magnetic Resonance Imaging can be summarized as follows:

- Improved angular sensitivity: the Spherical Deconvolution-Sparse Bayesian Learning method is capable of estimating complex intra-voxel configurations down to the range of 20° under some circumstances.
- Less demanding requirements for the imaging acquisition system: the proposed technique successfully resolves intra-voxel fiber heterogeneity at lower b-values than reviewed methods in literature. Using a lower b-value allows for faster acquisitions, and improved Signal-to-Noise Ratio figures.
-

If you want to understand function, study structure.

What Mad Pursuit: A Personal View of Scientific Discovery. 1988.
Francis Harry Compton Crick, British physicist, molecular biologist and neuroscientist (1916-2004). 1962 Nobel Prize in Physiology or Medicine

1

Principles of Diffusion MRI

The first part of this chapter consists in the formulation of the diffusion process for Brownian particles, in which the diffusion coefficient is related to the mean squared displacement of the particle. The excitation of nuclei in biological tissues through Magnetic Resonance techniques is later introduced as a necessary tool for the acquisition of diffusion images. Finally, the Diffusion Tensor Imaging technique is developed as the foundational basis for the description of the white matter micro-architecture in the brain.

1.1 Brownian Motion

Magnetic Resonance Imaging (MRI) is based on a microscopic phenomenon, called Brownian motion, by virtue of which molecules exhibit erratic trajectories in a medium. These kinematics are described by random thermal motion. Brownian motion is a process that occurs constantly in human tissues. It is named after the Scottish botanist Robert Brown, who in 1827 first described the movement of particles trapped inside pollen grains of the *Clarkia pulchella* flowering plant suspended in a water solution. It was not until 1905 when Albert Einstein published a paper[52] that the root causes of this behavior were reported, and it became clear that the pollen grains were being pulled by the random thermal agitation of the water molecules. Fig. 1.1 illustrates the Brownian motion phenomenon.

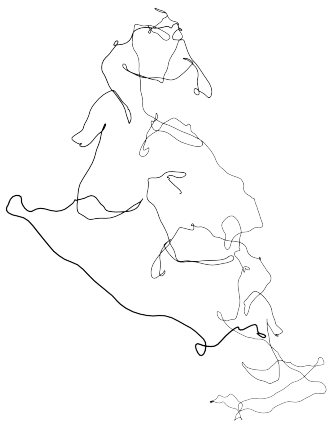


Figure 1.1 Brownian motion. The trajectory described by a particle in a free, unconstrained medium is a random walk caused by thermal effects. This phenomenon is called Brownian motion.

Although Brownian motion is mathematically expressed as a Wiener process, the physical process can be described by simple movement equations. At long timescales, the time evolution of the position of the Brownian particle can be mathematically described approximately by a Langevin equation, an equation which involves a random force field representing the effect of the thermal fluctuations of the fluid on the Brownian particle:

$$m\frac{d^2x}{dt^2} = -\gamma\frac{dx}{dt} + \xi(t) \tag{1.1}$$

where m denotes the particle mass; γ is the friction coefficient of the fluid; and $\xi(t)$ is a noise term or random force due to random density fluctuations in the fluid representing the effect of the collisions with the molecules of the fluid.

Thus, the expression in Eq. 1.1 has two terms: the first term represents the viscous force proportional to the particle's velocity, and the random force term, which has a Gaussian probability distribution.

The erratic motion of the Brownian motion is best described in statistical terms by the *displacement distribution*. The displacement distribution (statistically *displacement probability density function*) describes the proportion of molecules that undergo displacement in a specific direction and to a specific distance. The word 'displacement' is often replaced with the term *diffusion* within this context.

Einstein found[52] that the displacement probability density function p of Brownian particles diffusing in a medium could be expressed in the one-dimensional case as

$$p(x,t) = \frac{N}{\sqrt{4\pi Dt}}e^{-\frac{x^2}{4Dt}} \tag{1.2}$$

where x is the particle position; t is the time elapsed; N is the number of particles; and

D is the diffusivity of the medium.

The mobility of molecules can be characterized by a physical constant called the *diffusion coefficient*, D, which is related to the Root Mean Square (RMS) displacement of the molecules over a given time, t_{diff}, via the so-called Einstein equation

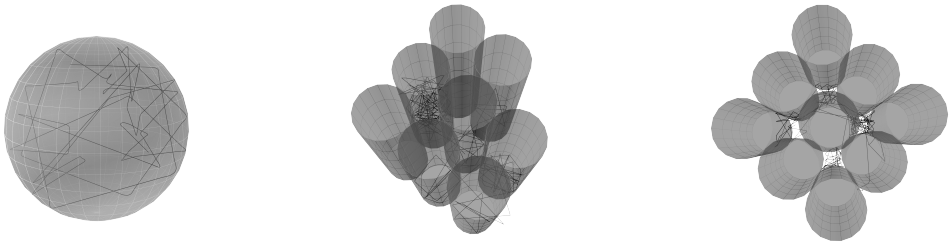
$$\hat{x}^2 = 2Dt_{diff} \tag{1.3}$$

In any given medium, the distance that a molecule diffuses in one direction in space may not be the same as in some other direction. Clearly, in a pure liquid, where there are no hindrances to diffusion or in a sample where the barriers are not coherently oriented, given a sufficient amount of time, diffusion is verified to be the same in all directions (i.e. there is no preferential direction). In this case, diffusion is said to be *isotropic*. However, if diffusion occurs preferably in a given direction (due to the presence of highly oriented barriers, compartments, etc.), it is termed *anisotropic* diffusion. In this case, molecules are said to undergo *restricted diffusion*. However, Beaulieau argued[13] that *hindered diffusion* may be a more appropriate term for biological systems, since membranes certainly have a finite permeability.

In biological samples, the underlying cellular microstructure of tissue influences the overall mobility of the diffusing molecules by providing numerous barriers and by creating various individual compartments. The degree of anisotropy being directly related to the geometry of these, different structural types (i.e. tissues) can be identified on the basis of their diffusion characteristics. The degree of hindrance to water diffusion will be determined by the size, shape, and composition of any physical constructions, as well as the spacing between them. Fig. 1.2 shows the trajectories of different particles subject to Brownian motion under different constraints that can be assimilated to those

imposed by biological structures.

Diffusion in a homogeneous medium is well described as having a Gaussian distribution. The spread of the Gaussian distribution (i.e. its variance) is determined by the type of molecule, the temperature of the medium, and the time allowed for diffusion, as described by Eq. 1.3.



(a) (b) (c)

Figure 1.2 Brownian motion in different restricted spaces. The diffusion phenomenon, described by Brownian motion or random walk, is influenced by the properties of the environment in which takes place. If the properties of the medium are equal along all directions in space, so will be the motion probability (i.e. there is no preferential direction); if the properties of the medium depend on the direction, the motion direction will not be equiprobable. In the first case, the medium is said to be "isotropic", whereas in the second case it is said to be "anisotropic". (a) Brownian motion in an isotropic medium (constrained to a spherical shape); (b) and (c) Brownian motion in anisotropic mediums: (b) constrained to a cylindrical shape; and (c) hindered by a cylindrical shape.

Biological tissues are highly heterogeneous media that consist of numerous, highly specific, structures, composed by permeable individual compartments and barriers that present different diffusivity characteristics.

A defining characteristic of neuronal tissue is its fibrillar structure. Neuronal structure consists of tightly packed and coherently aligned axons, surrounded by a complex extra-axonal environment containing astrocytes, glial cells, and ordered and randomly oriented extracellular matrix molecules. Axons are said to be organized in *fascicles* or

fiber bundles. As a result, the micrometric displacement of water molecules is hindered significantly perpendicular to the long axis of the fibers relative to the preferential parallel direction. Consequently, molecular displacement parallel to the fiber is typically greater than that perpendicular to it. When diffusive properties change with the direction of diffusion, the prevailing condition is anisotropy, and the associated displacement distribution is no longer isotropic or Gaussian, like that in unrestricted diffusion, but cigar-shaped. Restriction is thus at the core of axonal transport.

1.2 Nuclear Magnetic Resonance

Credit for the discovery of Nuclear Magnetic Resonance (NMR) goes to Isidor Isaac Rabi. Working at Columbia University in the 1930s, Rabi and his team were attempting to measure the magnetic properties of various nuclei including hydrogen, deuterium, and lithium. Rabi described how nuclei could be induced to flip their principal magnetic orientation by an oscillating magnetic field. Although the idea had been originally proposed by Dutch physicist Cornelius J. Gorter in 1936, Gorter was unable to validate this phenomenon due to limitations of his experimental setup.

The temporal behavior of the magnetization vector under the influence of an applied magnetic field can be described by the so-called phenomenological Bloch equations. The Bloch equations provide a valuable framework for describing many phenomena in NMR[20 19 102]. They constitute the mathematical model used for the description of most phenomena occurring in biological tissues during the acquisition of magnetic resonance images.

In the rotating frame, and in the absence of any Radio Frequency (RF) field, the

Bloch equations can be written as

$$\frac{\partial m}{\partial t} = -i\gamma \mathbf{r}(t)^T \mathbf{g}(t) - m/T_2 \tag{1.4}$$

where $m = m_x + im_y$ is the complex transverse magnetization vector; γ the gyromagnetic ratio (an intrinsic constant to a nucleus, $\mathrm{rad\,s^{-1}\,T^{-1}}$ in SI units); $\mathbf{r}(t)$ is the spin position as a function of time t, $\mathbf{g}(t) = |\nabla \mathbf{B}(t)|$ is the applied magnetic field gradient; $T_2 = 1/\rho$ is the spin-spin relaxation time (s); and ρ is the bulk relaxation rate ($\mathrm{s^{-1}}$).

The Bloch equations were modified [111] by Torrey to include the effects of molecular diffusion and flow to give the Bloch-Torrey equation

$$\frac{\partial m}{\partial t} = -i\gamma \mathbf{r}(t)^T \mathbf{g}(t) - m/T_2 + \nabla(\mathbf{D}(r)\nabla m) - \nabla \mathbf{v} m \tag{1.5}$$

where $\mathbf{D}$ is the diffusion tensor; and v is the flow vector.

The Bloch-Torrey equation's solution is expressed by (for a detailed derivation of the expression, see [111] and [107])

$$m = \exp\left[-i\gamma \int_0^t \mathbf{r}^T(t')\mathbf{g}(t')dt'\right] A(t) \exp\left(-t/T_2\right) \tag{1.6}$$

with

$$A(t) = \exp\left[-\int_0^t Tr\left(\mathbf{B}(t')\mathbf{D}\right) dt'\right] \exp\left[i\gamma \int_0^t \int_0^{t'} \mathbf{v}^T(t'')\mathbf{g}(t'')dt''dt'\right]$$

where $Tr\left(\cdot\right)$ is the matrix trace.

The above equation can be used to calculate the attenuation due to Gaussian diffusion and coherent flow for any effective gradient profile $\mathbf{g}^*(t)$.

1.3 Magnetic Resonance Imaging

While working on the relaxation times of different nuclei under the NMR, Armenian-American physicist Raymond Vahan (Johnson) Damadian, found that the hydrogen nuclei of water in cancerous and healthy tissues showed pronounced differences in relaxation times. Thus, on a paper published in 1971 he proposed[43] to distinguish cancerous tissue from healthy tissues on the basis of the cells' water structure by NMR.

The idea of superimposing small variations (called *gradients*) to the main magnetic field, and thus, change the field strength to affect the resonance frequency of nuclei to collect spatial information was suggested by three separate groups, including Paul Lauterbur[75], Richard R. Ernst[73] and Peter Mansfield[85]. The Magnetic Resonance (MR) image is the result of the application of gradients in different directions and with different intensities at specific moments of the acquisition sequence.

Paul Lauterbur is credited for demonstrating MRI[75] in 1973, extending upon the ideas behind NMR to create MR images using a back projection technique similar to that used in Computed Tomography (CT). Two years later, Richard R. Ernst proposed[73] collecting magnetic resonance images using phase and frequency encoding by means of the Fourier Transform (FT), which, together with gradient encoding, constitutes the basis of current MRI techniques. The values of the measured signal are organized in a coordinate system or formalism known as the **k**-space, where k refers to the sequence (image) index in Fourier space. Performing the acquisition enables the filling of **k**-space. To transform the raw MRI data from **k**-space into a position-encoded visual image, the FT operation is applied.

Although the first human MRI scans were carried out in 1976 by Mansfield's team, it was not until 1977 when the first MRI body exam was performed on a human subject[44].

In 1977, Peter Mansfield[84] developed the Echo Planar Imaging (EPI) technique, which enabled to collect MRI images many times faster than previously possible, and found applications in both Functional Magnetic Resonance Imaging (fMRI) and Diffusion Magnetic Resonance Imaging (dMRI). Fig. 1.3 displays a human brain MRI acquisition.

In Magnetic Resonance Imaging, an adequately applied magnetic field influences the phase of the spins, with the degree of influence depending on the strength of the field. MR data are collected by changing on and off an external magnetic field at a given intensity in a prescribed sequence, called *pulse sequence*. The pulse sequence determines the content, quality, contrast and resolution of the image.

MR signals are sensitive to the magnetic dipole orientation of the hydrogen nuclei in water and fat present in biological tissues. This magnetic dipole, known as *spin*, can align itself to an externally applied magnetic field. The MRI scanner maintains a constant, strong magnetic field B_0 (typically measured in Tesla (T)). A second magnetic field oscillating at RFs, known as the *RF pulse sequence*, or simply, *RF pulse*, is applied for a brief duration at regular intervals. In the presence of an external magnetic field, the net magnetization vector of spins $\mathbf{M}$ is parallel to the former. Spins will come back to this configuration some time after the application of the RF pulses. RF pulses are used primarily for excitation and refocusing. Following the application of a 90° ($\pi/2$) RF, spins will rotate away from their preferred orientation parallel to B_0 (usually depicted as parallel to the Z axis), and will precess in another orientation (typically the XY plane). Excited spins precess about B_0 at a frequency given by the Larmor equation $\omega = \gamma B$ (usually given in MHz). The net magnetization vector will be aligned parallel to the XY plane.

Due to the interaction between the spins (*spin-spin interaction*), each spin loses

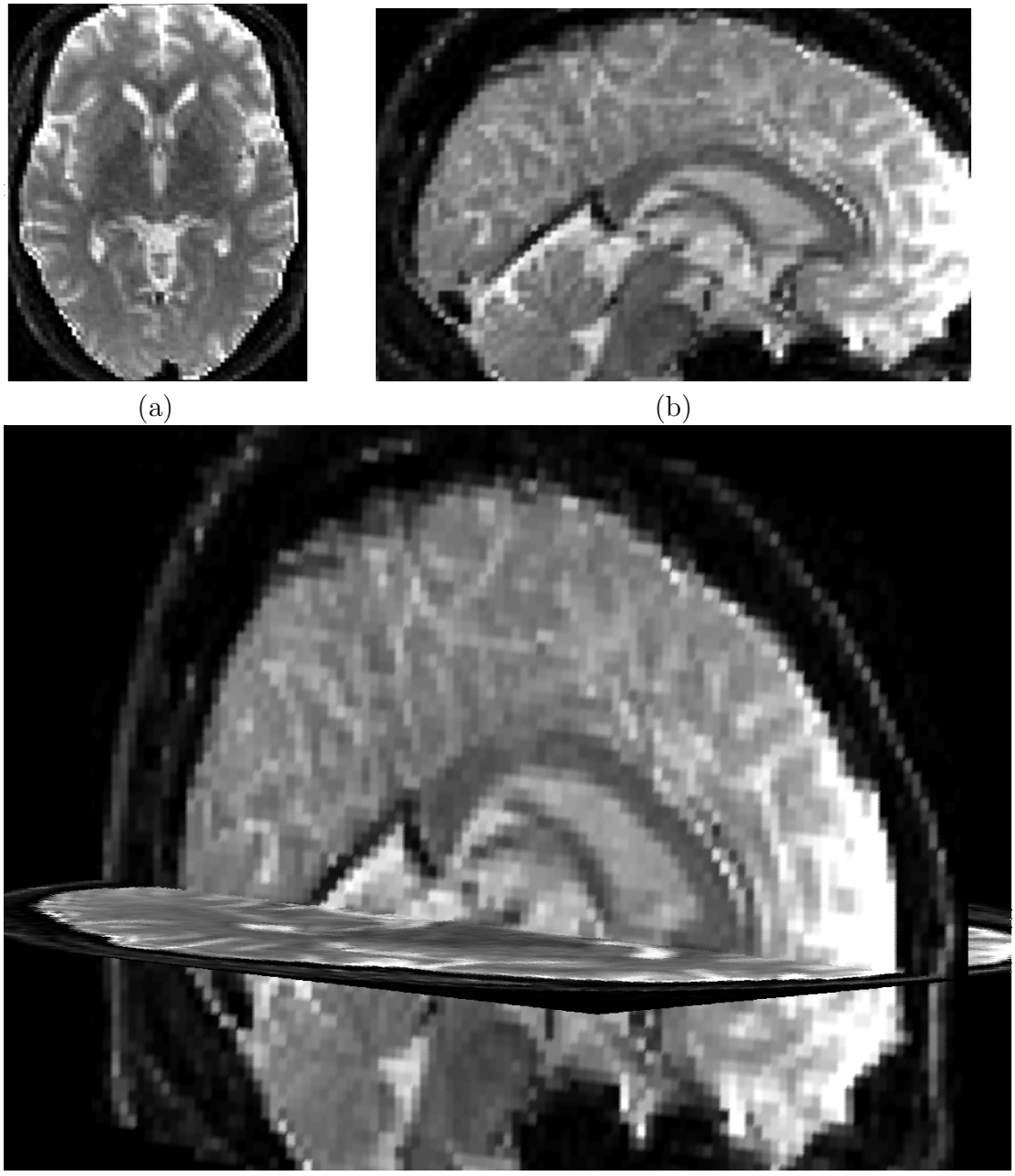


Figure 1.3 The brain under MRI. An MRI image of the human brain: (a) axial slice (inferior view); (b) sagittal slice (right lateral view); and (c) axial-sagittal (anterior-right) view.

coherence and will precess away from the Larmor frequency. The precessing part that is perpendicular to the direction of B_0 decays exponentially with a time constant T_2 (also known as *transverse relaxation*). At the same time, due to the so-called *spin-lattice interaction* spins realign themselves exponentially in the direction of B_0 with a time constant T_1 (commonly referred to as *longitudinal relaxation*). T_1 and T_2 are intrinsic properties of the tissues, satisfying $T_2 < T_1$. The generated magnetic field from the coherently precessing spins induces a current in the receiver coils; this current is used to generate the MR images, and corresponds to image brightness. The more coherent the phase of the precessing spins, the higher will be the brightness of the image picture elements (pixels). The observable NMR signal generated by non-equilibrium nuclear spin magnetization precessing about the magnetic field B_0 is called Free Induction Decay (FID).

The applied field is not perfectly homogeneous, though. This produces the spins to precess at different phases, which involves a loss of transverse magnetization, and thus a decreased signal intensity. The corresponding transverse relaxation constant is termed T_2^* ($T_2^* \ll T_2$).

Decoherence because of magnetic field inhomogeneity is not a true relaxation process; it is not random, but dependent on the location of the molecule in the magnet. For molecules that are not moving, the deviation from ideal relaxation is consistent over time, and the signal can be recovered by performing a Spin-Echo (SE) experiment. Echoes and the "spin echo effect" in nuclear magnetic resonance were first described[63] by Erwin Hahn in 1950. He discovered that the application of two successive 90° pulses separated by a short time period gave rise to the effect, but a signal, the "echo", was detected when no pulse was applied. The moment of spin refocus is called a *spin echo* (sometimes referred to as *Hahn echoes*), and creates the measured signal. Acquisition

sequences that employ a refocusing pulse are called *spin echo pulse sequences*, and *gradient echo sequences* otherwise.

The additional magnetic field generated by an NMR scanner are called magnetic field gradients or simply gradients (G). Gradients have a special role in diffusion imaging, as explained in Sec. 1.4.

Often, a second 180° (π) RF pulse is applied[28] at some time after excitation to recover the in-plane magnetization, since it inverts the phase of spins. This pulse is applied at half the time between the 90° RF excitation pulse and data sampling, which is termed Echo Time (TE). The time between two consecutive 90° excitation RF pulses is called Repetition Time (TR).

1.4 Diffusion Imaging

Diffusion phenomena present in biological tissues can reveal microstructural properties. dMRI or simply Diffusion-Weighted Imaging (DWI) is a technique that allows to capture such phenomena in various types of tissues *in-vivo* non-invasively. This idea was first suggested[124] by Wesbey *et al.* in the early 1990s. Potential clinical applications of water dMRI were suggested very early[76]. The first successful application of diffusion MRI to clinical routine was the detection of brain ischaemia[122] in the early 1990s, following the discovery[92] by Moseley *et al.* that water diffusion dropped at a very early stage of the ischaemic event in cat brains.

Anisotropic water diffusion is clearly not unique to neural fiber tracts, and it has been observed in other tissues such as kidneys[64], skeletal muscle[34], or the myocardium[59]. However, the degree of anisotropy tends to be much less than that found in neural fiber tracts.

Water diffusion is clearly restricted in neural fiber tracts, such as white matter (WM) and nerve. Fig. 1.4 shows the WM corpus callosum (cc) fiber bundles. Although in these last years there has been a rising interest in the scientific community (see, for example,[40]) to interpret and relate the dependency of the diffusion process (or any other metric computed from DWI) with the underlying tissue microstructure and properties, yet it remains an open issue.

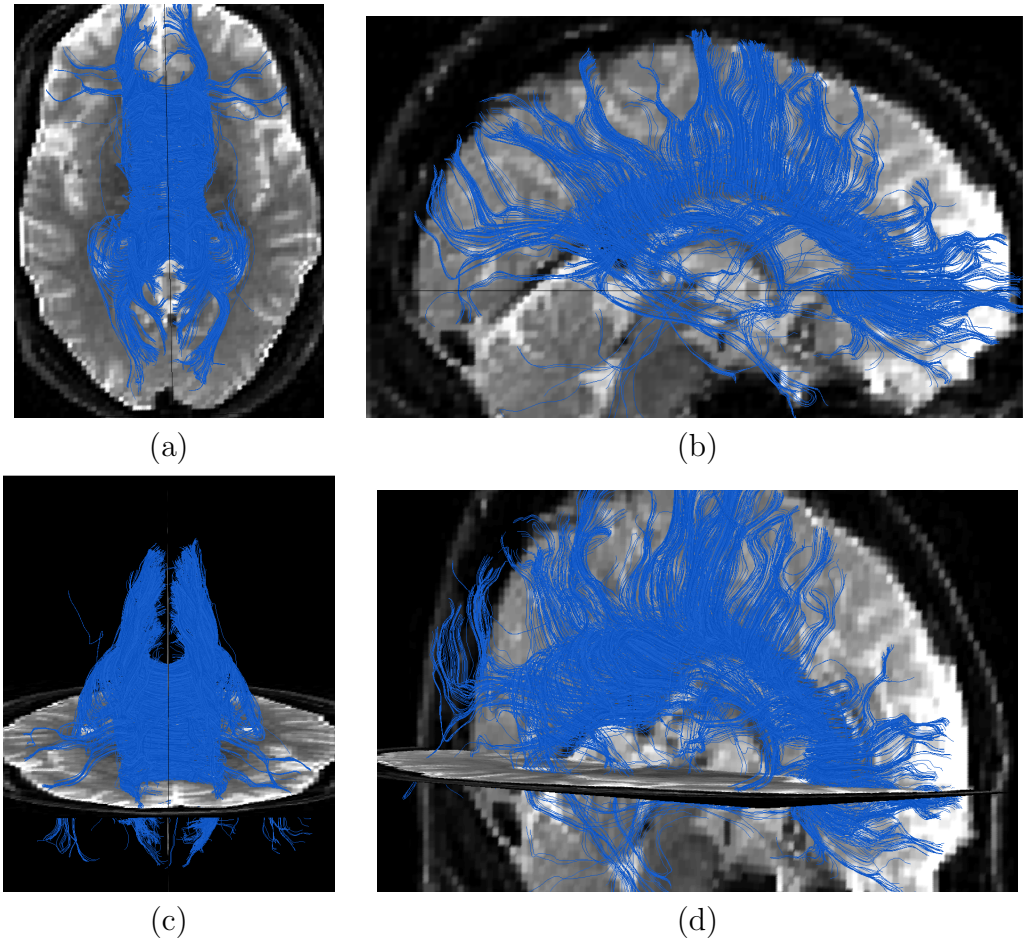


Figure 1.4 Tractography: the fiber bundles' pathways. Human brain white matter corpus callosum (cc) fiber bundle tracts or pathways: (a) axial slice (inferior view); (b) sagittal slice (right lateral view); (c) axial-coronal (anterior) view; and (d) axial-coronal-sagittal (anterior-right) view.

It is still not clear whether the observed diffusion anisotropy arises from the intra-axonal compartment (primarily restricted diffusion) or the extra-axonal compartment (primarily hindered diffusion) or some combination thereof. There is a body of experimental evidence that suggests[14] that myelin sheaths in neural axons are not necessary determinants for anisotropic water diffusion in neural fibers, structural features other than myelin being sufficient to give rise to anisotropy[13]. Evidence suggests that axonal membranes play a primary role in restricted water diffusion in white matter, while myelin modulates the degree of anisotropy. The complex and dense three-dimensional structure of the axons, composed of longitudinally oriented cylindrical neurofibrils, named microtubules and neurofilaments, which, at the same time, are inter-connected by microfilaments, presumably participate in the restriction of water diffusion along neural axons. The relative contributions of the various structural components of white matter to the anisotropy of the water diffusion in the brain are yet to be determined.

Gray matter (GM), on the other hand, does not have an oriented fiber structure, and thus, exhibits low diffusion anisotropy. At the white, gray matter interface some degree of structural transition is observed.

1.4.1 Pulsed-Gradient Spin Echo

Diffusion experiments are usually performed by spatially labeling spins at two different times during an MR experiment. These labeling periods are separated by a time interval (known also as the diffusion time, Δ) during which the spin's displacement is quantified. The measured signal decay will depend on the strength of the labeling (referred to as the q-value) and the diffusion time.

To describe diffusion, every voxel position is assigned a diffusion probability density function. The resulting image encodes the proportion of molecules in a given voxel

position that have been displaced a certain distance. To depict the displacement distribution, diffusion must be related to the signal intensity as measured by the magnetic resonance image. Magnetic resonance measurements of diffusion are sensitive to molecular displacements along the axis of the diffusion-sensitizing gradients applied in a standard Pulsed-Gradient Spin Echo (PGSE) sequence or Stejskal-Tanner method[107]. Spins moving in the magnetic field gradient direction are exposed to different magnetic field strengths depending on their position along the gradient axis. The diffusion of water along different directions in tissue can be readily assessed by varying the direction of the diffusion-sensitizing gradients.

In diffusion MR imaging, the application of a single PGSE sequence produces one diffusion-weighted image that corresponds to one position in **q**-space or, more precisely, that depicts the diffusion-wighted signal intensity in a specific position **q** for every brain position. Repeated applications of the sequence with gradients that vary in strength and direction allow data sampling throughout **q**-space.

In a PGSE sequence (Fig. 1.5), a 90°-180° spin echo pair of two additional diffusion-encoding gradient pulses G_{diff} are added to the standard SE sequence to introduce a phase shift proportional to the molecular displacement along the gradient direction. The key idea is to excite the spin system with a 90° pulse, encode the spin position with a time-constant magnetic field gradient of duration δ, invert the spin phase with a 180° pulse, apply a second magnetic field gradient with equal intensity and duration to the previous gradient at a time Δ after the first gradient pulse, and then acquire the image at time $t = \mathrm{TE}$.

The first gradient pulse introduces a phase shift that is dependent on the strength of the gradient at the position of the spin at $t = t_0 = 0$. Before the application of the second gradient pulse, which induces a phase shift depending on the spin position

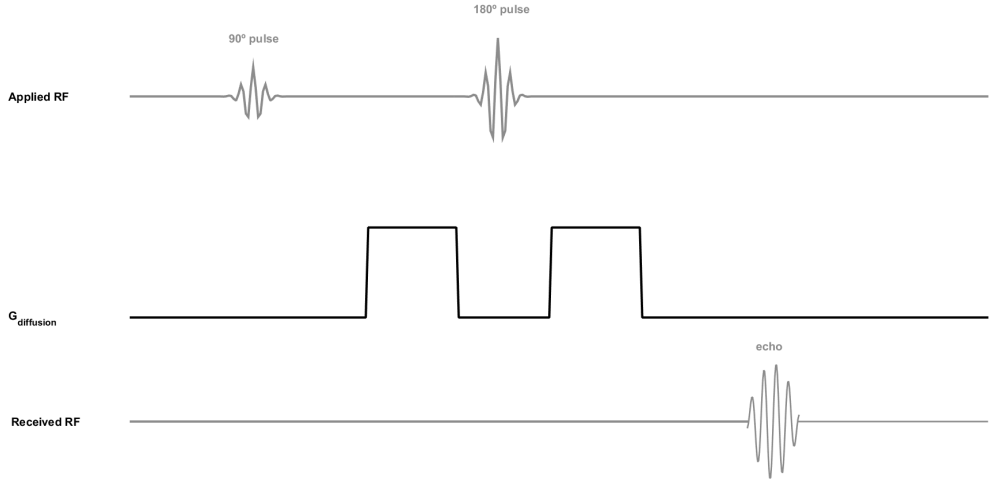


Figure 1.5 Pulsed-Gradient Spin Echo (PGSE) magentic resonance imaging acquistion sequence. A PGSE sequence introduces an additonal spin-echo pair of 90°-180° gradient pulses to the standard SE sequence to introduce a phase shift proportional to the molecular displacement along the gradient direction.

at time Δ, a 180° RF pulse is applied to reverse the phase shift introduced by the first gradient pulse. Since the diffusion-encoding gradient causes the field intensity (i.e. phase shift) to vary with position, all spins that remain at the same location along the gradient axis during the pulses will return to their initial state. However, if a given spin moves by diffusion between the time of the first and second gradients, then it will be subjected to a different field strength during the second pulse, and therefore, will not return to its original orientation. Due to the increased disorder of the spin phase distribution, the spin echo signal, which is proportional to the mean phase difference, will be smaller relative to the case in which there was no gradient. Hence, the decrease in the spin echo magnitude will reflect the amount of diffusion which occurred between the two diffusion gradient pulses. The longer the displacement distance, the higher the phase shift, and the greater the signal decrease. Hence, the resultant image shows low

signal intensity in regions where diffusion along the applied diffusion gradient is high.

A diffusion gradient can be represented as a 3D vector $\mathbf{q}$, called *spin displacement wavevector* (or simply *wavevector*) or *reciprocal vector*, whose orientation is in the direction of diffusion, and whose length, the *wavenumber*, is proportional to the gradient strength:

$$\mathbf{q} = \gamma \delta \mathbf{g} \tag{1.7}$$

with δ the diffusion-encoding gradient duration; and $\mathbf{g}$ the diffusion gradient wave vector.

The attempts to describe the DWI signal in terms of the q-value were originally developed by Callaghan[?] and Cory and Garroway[36]. In their original approach, no specific model of water diffusion was assumed. Instead, features of the measured displacement probability distribution were extracted by a Fourier transformation of the signal attenuation profile with respect to $\mathbf{q}$.

A single application of the pulsed-gradient SE sequence produces one brain image with a given diffusion weighting; multiple repetitions of this sequence, each with a different diffusion weighting (i.e. varying the gradient strength and direction), are necessary to sample the $\mathbf{q}$-space. The result is a collection of images, each of which reflects the particular diffusion weighting used, and corresponds to one point in $\mathbf{q}$-space. In every position, a Fourier transform relates the raw $\mathbf{q}$-space data to the diffusion probability density function.

Nevertheless, the gradient strength, or more often, the diffusion weighting, is more commonly expressed in terms of a scalar quantity named the *diffusion-sensitizing gradient factor* or b-value (b), which is proportional to the product of the square of the gradient strength $\mathbf{q}$ and the diffusion time interval Δ ($b \propto q^2\Delta$).

The b-value is derived from the Bloch-Torrey equations, and it is a function of the

strength, duration, temporal resolution and timing parameters of the specific acquisition scheme. Hence, it summarizes the amount of diffusion sensitizing gradient history.

In a PGSE experiment using rectangular diffusion-encoding gradient pulses of duration δ, and amplitude G, separated in time by Δ and applied, for example, along the X direction, the classical Stejskal-Tanner expression for the b-value in the X direction at the time of readout reads[107]:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right) \tag{1.8}$$

$$= q^2 \left(\Delta - \frac{\delta}{3} \right) \tag{1.9}$$

where G is the gradient strength; δ denotes the pulse width (timing of the diffusion-sensitizing gradients); and Δ is the diffusion time interval.

The Stejskal-Tanner equation

In his seminal paper on the NMR spin echo, Hahn[63] noted that molecular diffusion in the presence of a magnetic field gradient would reduce the magnitude of the observed echo signal. Quantitative measurement of a molecular diffusion coefficient was not made possible, however, until the introduction of the PGSE experiment by Stejskal and Tanner[107]. Stejskal and Tanner showed that the spin echo magnitude $\mathbf{S}(\mathbf{q}, \tau)$ from a PGSE experiment is directly related to the spin displacement Probability Density Function (PDF) $P(\mathbf{R}, \tau)$ by a Fourier relation

$$S(\mathbf{q}, \tau) = S0 \int P(\mathbf{R}, \tau) e^{iq^T \mathbf{R}} d\mathbf{R} \tag{1.10}$$

$$= S_0 \mathcal{F}[P(\Delta \mathbf{r})]$$

where $S_0 = S(0, \tau)$ is the signal in the absence of any applied diffusion gradient; $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ is the relative spin displacement from the initial position $\mathbf{r}'$ of the spin at the time of the first gradient to the position $\mathbf{r}$ of the spin at diffusion time τ; and $\mathbf{q}$ is the spin displacement wave vector introduced previously.

The PDF $P(\mathbf{R}, \tau)$, which is referred to as the Ensemble Average (Diffusion) Propagator (EAP), expresses the mean probability to observe a relative spin displacement $\mathbf{R}$ in the experimental diffusion time τ in a volume element (voxel).

The observed diffusion propagator represents a spatial sum over the microscopic environments existing in the voxel, where the resolution of the voxel is defined by the macroscopic spatial encoding as allowed by the available Signal-to-Noise Ratio (SNR). This averaged propagator is referred to as the EAP defined as the propagator averaged over all initial positions and expresses the average probability of a spin displacement $\mathbf{R}$

$$P(\mathbf{R}, \tau) = \int P(\mathbf{r}, \mathbf{r}', \tau) p(\mathbf{r}') dr' \tag{1.11}$$

where $p(\mathbf{r}')$ is the initial spin density.

The statistics of the spin displacement gives rise to the diffusion coefficient and, in the case of anisotropic materials, the $\mathbf{D}$. Consider an isotropic medium where the spin position at time $t = t_0 = 0$ is given by r', and the subsequent spin position at time τ is given by $\mathbf{r}$. The diffusion coefficient D is then given by Einstein's relation[52]

$$D = \frac{1}{6\tau} \langle \mathbf{R}^T \mathbf{R} \rangle \tag{1.12}$$

where $R = \mathbf{r} - \mathbf{r}'$ is the relative spin displacement as a function of time τ; and $\langle \cdot \rangle$ is the average over the spin ensemble.

D is often referred to as the *diffusivity value* or Apparent Diffusion Coefficient (ADC) (also denoted as D_{app} in diffusion literature). It is the diffusion coefficient measured by Nuclear Magnetic Resonance. The term *apparent* denotes that in a finite, non-homogeneous medium (such as human tissue) it represents a net effect, since in such environments the intrinsic diffusivity cannot be exactly defined or measured. Its value depends on the interactions of the diffusing molecule, in most cases water, with the cellular structures over a given diffusion time. The units of D are mm^2/s (for water at 37°C $D \approx 3 \times 10^{-3}$ mm^2/s), and b are s/mm^2, typically in the range from 0 to 5000 s/mm^2 for diffusion experiments, though some acquisition paradigms can call for values of b up to or beyond 10 000 s/mm^2 [123 23].

In cases where the medium is anisotropic, the Einstein relation generalizes to the diffusion tensor $\mathbf{D}$

$$\mathbf{D} = \frac{1}{6\tau}\langle \mathbf{R}\mathbf{R}^T \rangle \tag{1.13}$$

The diffusion term $\mathbf{D}$ in Eq. 1.13 is a three-dimensional rank-2 tensor. According to Onsager's hypothesis[95] the diffusion tensor is symmetric ($\mathbf{D} = \mathbf{D}^T$), and positive definite[118].

The spin echo attenuation for the PGSE experiment can be calculated using Eq. 1.6. The effective gradient for the PGSE sequence is:

$$\mathbf{g}^*(t) = -\theta(t) + \theta(t-\delta) + \theta(t-\Delta) - \theta(t-\Delta-\delta) \tag{1.14}$$

where $\theta(t) = \delta(t \geq 0)$ is the Heaviside step function. Substitution of this effective gradient into Eq. 1.5 gives

$$m = \exp\left[-\gamma^2\delta^2\mathbf{g}^T\mathbf{D}\mathbf{g}\left(\Delta - \frac{\delta}{3}\right)\right]\exp\left(-\gamma\delta\Delta\mathbf{g}^T\right)\exp\left(-\frac{TE}{T_2}\right) \tag{1.15}$$

which is the Stejskal-Tanner relation. The time constant $\tau = \Delta - \delta/3$ defines an effective diffusion time where the $\delta/3$ correction is due to the diffusion which occurs during the time in which the gradients are on.

The Stejskal-Tanner equation relates the observed diffusion signal to the underlying diffusion coefficient (or diffusion tensor) assuming that the diffusion is purely Gaussian. However, it is currently accepted[5] that the Gaussianity assumption is violated in biological tissue due to the presence of restrictions, such as permeability, or diffusion compartment heterogeneity.

The PGSE signal S is thus proportional to m, and can be understood as a superposition of the transverse spin magnetizations, each with a phase φ

$$S = S_0 \langle e^{i\varphi} \rangle \tag{1.16}$$

where $S_0 = S(0, \tau)$ is again the spin echo signal in the absence of any applied gradient, and corresponds to the term $\exp\left(-\frac{TE}{T_2}\right)$ in Eq. 1.15 (note that it does not depend on neither diffusion phenomena nor the b-value); and $\langle \cdot \rangle$ denotes the ensemble average. The ensemble average of any general function $f(\varphi)$ of the phase φ can be written explicitly in terms of the phase probability distribution

$$\langle f(\varphi) \rangle = \int f(\varphi) dP(\varphi) = \int f(\varphi) P(\varphi) d\varphi \tag{1.17}$$

where $P(\varphi)$ is the probability of a spin phase φ.

In the limit of infinitely narrow pulses ($\delta \ll \Delta$; see[72] and[128] for further details about the implications of the infinitely narrow pulse approximation) the effective gradient

Eq. 1.14 from the Stejskal-Tanner experiment is

$$\mathbf{g}^*(t) = \mathbf{g}(\delta(0) - \delta(\Delta)) \tag{1.18}$$

Note that $\mathbf{g}$ is independent of time. The spin phase φ is then

$$\begin{aligned}\varphi(t) &= \gamma \int_0^t \mathbf{r}(t')^T \mathbf{g}(t') dt' \qquad (1.19)\\ &= \gamma\delta\mathbf{g}^T(\mathbf{r}(0) - \mathbf{r}(\Delta))\\ &= \gamma\delta\mathbf{g}^T(\mathbf{r}' - \mathbf{r})\\ &= \gamma\delta\mathbf{g}^T\mathbf{R} \qquad (1.20)\end{aligned}$$

where $\mathbf{r}'$ is the spin position at the time of the first gradient; $\mathbf{r}$ is the spin position at the time of the second gradient, Δ is the time between the two gradients; and $\mathbf{R} = \mathbf{r}' - \mathbf{r}$ is the relative spin displacement.

The spin displacement wavevector can also be expressed in terms of the spin phase as $\varphi = \mathbf{q}^T\mathbf{R}$. The wavevector $\mathbf{q}$ plays the important role of the Fourier reciprocal vector to the relative spin displacement vector $\mathbf{R}$.

Owing to the linearity of the Larmor equation (see Sec. 1.3), the probability of a phase difference φ is proportional to the probability of a spin displacement $\mathbf{R}$. Hence,

$$P(\varphi, \tau) = P(\mathbf{R}, \tau) \tag{1.21}$$

Eq. 1.16 can then be written as

$$S(\mathbf{q},\tau) = S_0 \int P(\mathbf{R},\tau) e^{i\mathbf{q}^T\mathbf{R}} d\mathbf{R} \tag{1.22}$$

$$= S_0 \mathcal{F}[P(\mathbf{R},\tau)] \tag{1.23}$$

where $P(\mathbf{R},\tau)$ is the EAP, and $\mathcal{F}$ is the Fourier transform with respect to the relative spin displacement vector $\mathbf{R}$. By expanding $\mathbf{R} = \mathbf{r}' - \mathbf{r}$, the above equation can also be written in terms of the absolute spin positions as

$$S(\mathbf{q},\tau) = S_0 \int P(\mathbf{r}',\mathbf{r},\tau) p(\mathbf{r}) e^{i\mathbf{q}^T(\mathbf{r}-\mathbf{r}')} d\mathbf{r} d\mathbf{r}' \tag{1.24}$$

where $P(\mathbf{r}',\mathbf{r},\tau)$ is the diffusion propagator mentioned in Sec. 1.2; and $p(\mathbf{r})$ is the initial spin density.

Eq. 1.23 illustrates the Fourier reciprocal relationship between the PGSE signal and the EAP. This relationship is key and forms the basis of all diffusion imaging experiments. In particular, the Fourier relationship allows for direct reconstruction of the EAP by inverse Fourier transform of the diffusion signal with respect to the displacement wave vector

$$P(\mathbf{R},\tau) = S_0^{-1} \int S(\mathbf{q},\tau) e^{-i\mathbf{q}^T\mathbf{R}} d\mathbf{q}$$

$$= S_0^{-1} \mathcal{F}^{-1}[S(\mathbf{q},\tau)] \tag{1.25}$$

where the inverse Fourier transform $\mathcal{F}^{-1}$ is taken with respect to the reciprocal vector $\mathbf{q}$.

The EAP can be measured on the basis of PGSE experiments independent of any

assumptions on the form of the underlying diffusion process.

1.4.2 Diffusion Tensor Imaging

Diffusion Tensor Imaging (DTI) has been the most popular diffusion MRI reconstruction algorithm by far until recently. DTI was introduced[12] by Basser *et al.* in the mid 1990s. The key assumption in DTI reconstruction algorithms is that the direction of maximum diffusivity is an estimate of the major fiber orientation.

Figure 1.6 DTI neuron model. Brain white matter axon (or fiber) bundle model. The DTI model gives acceptable results when the fiber bundles are coherently aligned, although their diameter and spacing can be vary.

Fig. 1.7 illustrates the anisotropy concept germane to the diffusion tensor imaging theory.

Under the Brownian motion premise, the diffusion signal strength for isotropic diffusion is described by the model known as the Stejskal-Tanner[107] formula:

$$S(\mathbf{q}) = S_0 \exp(-bD) \tag{1.26}$$

The scalar Stejskal-Tanner relation in Eq. 1.26 contains two unknowns, namely, the unattenuated echo signal S_0 and the diffusion coefficient D. Hence, the diffusion coefficient can be quantitatively estimated from at least two measurements of the signal each with different b-value magnitudes. If we take the natural logarithm of the spin echo signal, then the diffusion coefficient can be obtained from a set of diffusion experiments by standard linear regression methods. Diffusion-weighted imaging cannot, however, describe the anisotropic Gaussian diffusion which is observed in fibrous biological tissues. In such cases, the DTI experiment is called for.

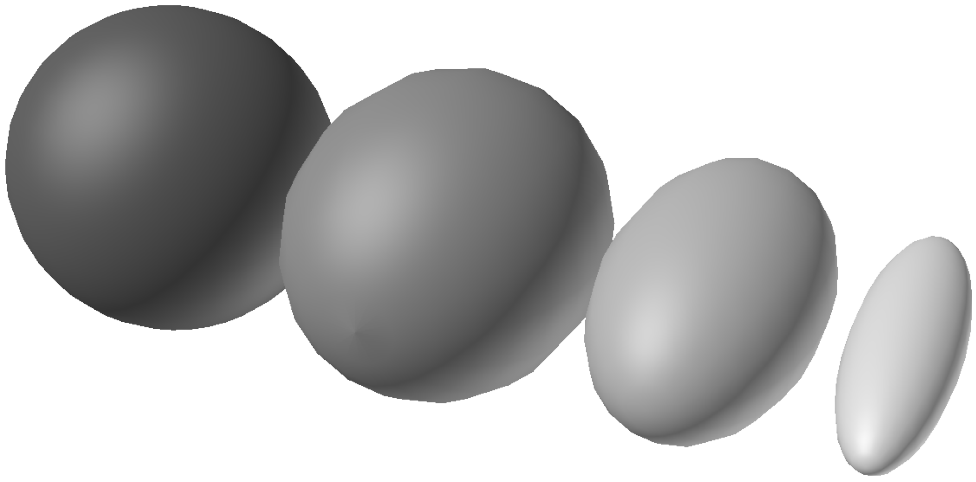


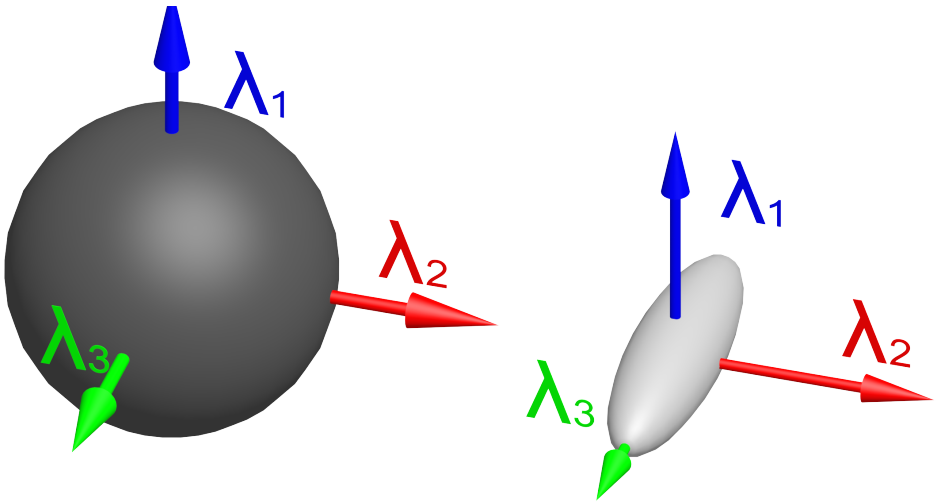
Figure 1.7 Anisotropy in diffusion tensor imaging. The displacement field of the diffusion process in an unrestricted, isotropic environment would describe a spherical boundary. Diffusion of water molecules inside axons is restricted by the geometry of the latter; since the preferential movement takes place along the long axis of the axons, the displacement field tends to be anisotropic, describing an ellipsoid. As the geometry gets more restrictive (i.e. the larger the ratio between the length and the radius in the transversal axis), the sharper will be the ellipsoid, and the higher the anisotropy.

It can be demonstrated that for free, anisotropic diffusion the spin echo signal is computed as

$$S(\mathbf{q}) = S_0 \exp(-b\mathbf{g}^T\mathbf{D}\mathbf{g}) \tag{1.27}$$

where $\mathbf{D}$ is the diffusion tensor we wish to measure; and $\mathbf{g}$ is a unit-vector pointing in the same direction as $\mathbf{q}$.

DTI provides two key insights into material microstructure that simple DWI does not[4]. First, it provides rotationally invariant statistics of the anisotropy of the probability density function of particle displacements over a fixed time, such as the Fractional Anisotropy (FA), which reflects the anisotropy of the microstructure. Second, the principal eigenvector of $\mathbf{D}$ can provide an estimate of the dominant orientation of



(a) (b)

Figure 1.8 The tensor model in diffusion tensor imaging. Under the DTI assumption, a fiber bundle is mathematically described as a second order tensor. Such a tensor is defined by its eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ and eigenvectors (spatial orientation of the eigenvalues). The shape of the ellipsoid is determined by the eigenvalues; the eigenvectors determine its orientation. DTI methods attempt to compute these in order to determine the fiber orientation. Under this model, the fiber orientation corresponds to the orientation of the dominant eigenvector. (a) shows the isotropic diffusion tensor model; and (b) shows anisotropic diffusion tensor model.

the microstructural fibers.

Diffusion tensor MRI measures the apparent water self-diffusion tensor under the premise of Gaussian diffusion (see Sec. 1.1). Based on the eigenstructure of the measured diffusion tensor it is possible to infer the orientation of the diffusion compartments within the voxel so that, for example, the major eigenvector of the diffusion tensor parallels the mean fiber orientation.

The diffusion tensor is a three-dimensional, symmetric, rank-2 tensor, and therefore has 6 unique coefficients: the 3 diagonal elements, and the 3 off-diagonal elements of

the tensor

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

where $D_{xy} = D_{yx}; D_{xz} = D_{zx}; D_{yz} = D_{zy}$.

The tensor reconstruction can be formulated as a linear inversion problem. The diffusion tensor can be decomposed into the eigensystem

$$\mathbf{D} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^T$$

where $R = (\mathbf{e_1}\mathbf{e_2}\mathbf{e_3})$ is a column matrix of the orthonormal diffusion tensor eigenvectors e_ν; and $\mathbf{\Lambda} = Diag\,(\lambda_1\lambda_2\lambda_3)$ is a diagonal matrix of the diffusion tensor eigenvalues λ_i. The eigenvectors $\mathbf{e_1}$, $\mathbf{e_2}$, $\mathbf{e_3}$ are often refereed to, respectively, as the major, medium, and minor eigenvectors.

The diffusion tensor eigensystem can be conceptualized in terms of the diffusion isoprobability surface which represents the surface on which a spin at the origin will diffuse to with equal probability. For a Gaussian diffusion process the isoprobability surface is a three-dimensional ellipsoid, as depicted in Fig. 1.7. The isoprobability ellipsoid represents the surface on which the Gaussian diffusion function has a constant value.

The axes of the isoprobability ellipsoid are oriented in the direction of the tensor eigenvectors and have lengths proportional to the diffusion distance along the corresponding eigenvectors. Since the diffusion distance is proportional to the square root of the diffusion eigenvalues, this is equivalent to scaling by the square root of the diffusion tensor eigenvalues.

All diffusion images must be compared to a reference image that is not diffusion weighted (i.e. a standard SE image): in other words, one for which the strength of the diffusion gradient is zero ($q = 0$ and $b = 0$). Consequently, at least 7 image acquisitions are required to reconstruct the diffusion tensor: 6 to obtain the 6 unique tensor elements, and 1 to estimate the unattenuated signal magnitude S_0.

Most DTI measurement schemes acquire more than the minimum seven measurements to reduce the effects of noise. The standard approach is to acquire M measurements with $q = 0$, and N measurements with non-zero wavenumbers $\mathbf{q_i}, i = 1, \ldots, N$. The $|\mathbf{q_i}|$ are all equal and the diffusion time t, and hence the diffusion weighting factor $b = t\left|\mathbf{q}^2\right|$ is fixed for all the $S(\mathbf{q_i})$. The gradient directions $\hat{\mathbf{q}}_\mathbf{i}$ are unique and distributed uniformly over the sphere. We refer to this kind of measurement scheme "spherical acquisition scheme", since the $\mathbf{q_i}$ all lie on a sphere in $\mathbf{q}$-space.

1.4.3 The Orientation Distribution Function

One of the main interests lies in knowing the way the measured diffusion signal depends on the underlying orientation of the fibers. One possible approach would be to replace the diffusion probability density function with an isosurface, a surface passing through all points of equal probability density value. For example, an isosurface of a 3D Gaussian distribution is an ellipsoid (see Fig. 1.7). A more commonly used technique that is less sensitive to noise involves the computation of the spin displacement Orientation Distribution Function (also referred to as Orientation Density Function) (ODF) (also referred to as orientation density function)[117] from the displacement distribution. An ODF may be considered a spherical polar plot whose radius in a given direction is proportional to the integral of the diffusion probability density function in that direction.

The ODF expresses the probability of a spin displacing into a differential solid angle

about a possible fiber direction $\hat{\mathbf{u}}$ at each individual MR image voxel: it is the projection of the diffusion function onto the sphere through the integration over the radial coordinate of the diffusion function. Thus, it sacrifices all the radial information but keeps the relevant directional information.

For a given propagator $P(r\hat{\mathbf{u}})$ the ODF (denoted ψ) is obtained via the radial projection transformation

$$\psi(\hat{\mathbf{u}}) = \int_0^\infty P(r\hat{\mathbf{u}})r^2 dr \tag{1.28}$$

where $\hat{\mathbf{u}}$ is a unit normal vector; and r is the radial coordinate in the diffusion space coordinate system.

The ODF is a function on the sphere. The sphere is usually sampled at evenly distributed points on a discrete spherical grid. The reconstruction task lies in correctly estimating the points (i.e. directions) where the value of the ODF is maximum to identify the direction of the underlying fibers.

Tuch[118] demonstrated the correspondence between the orientational maxima of the ODF and those of the underlying fiber orientation density at each voxel location.

The Fiber Orientation Distribution Function

Anderson suggested[6] that the diffusion ODF is only an approximation to the underlying fiber distribution, which is estimated more closely by the fiber Orientation Distribution Function (also referred to as fiber Orientation Density Function) (fODF). The fODF gives the fraction of fibers within the sample that are aligned along the direction (θ, ϕ), expressed in spherical coordinates, and therefore contains all the volume fraction information.

If we assume equal diffusion characteristics across all fiber populations in the brain,

the diffusion-weighted signal attenuation that would be measured from a single coherently oriented fiber population can be represented by an axially symmetric response function $R(\theta)$ (assuming a response function such that the fibers are aligned with the Z-axis). We can then express the observed diffusion signal $S(\theta, \phi)$ that would be measured from a sample containing several distinct fiber populations as the sum of the response functions of each population, weighted by their respective volume fractions, and rotated such that they are aligned along their respective orientations

$$S(\theta, \phi) = \sum_{i=1}^{N} f_i \hat{\mathbf{A}}_i R(\theta) \tag{1.29}$$

where f_i is the volume fraction for the i-th fiber population; $\hat{\mathbf{A}}_i$ is the operator representing a rotation onto the direction (θ_i, ϕ_i); and N is the total number of fiber populations.

This can be expressed as the convolution over the unit sphere of the response function $R(\theta)$ with a fiber orientation density function $F(\theta, \phi)$ (fiber ODF)

$$S(\theta, \phi) = F(\theta, \phi) \otimes R(\theta) \tag{1.30}$$

where $\otimes$ denotes the convolution operation.

The purpose of models is not to fit the data but to sharpen the questions.

The 11th R. A. Fisher Memorial Lecture, the Royal Society of London 20, April 1983.

Samuel Karlin, American mathematician (1923-2007)

2

Fiber Orientation Reconstruction Methods

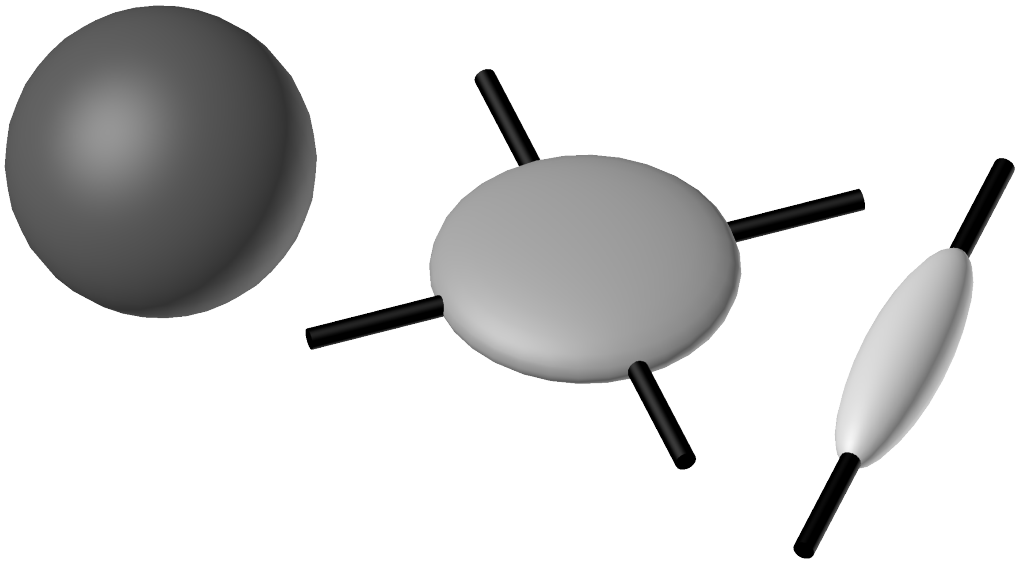
One of the main goals of current dMRI acquisitions is to determine the fiber orientations in each voxel from a set of measurements. The diffusion tensor framework can only describe diffusion phenomena occurring at a very limited class of biological conditions. Such diffusion excludes a vast range of biological environments which are commonly observed *in vivo* such as tissues with different diffusivity properties, or fiber populations with multiple preferential diffusion directions. Failure to resolve complex

tissue architecture with more than one significant fiber orientation within a voxel called into question the tensor model.

Figure 2.1 Intra-voxel orientational heterogeneity (IVOH). White-matter axonal fiber bundles can adopt a variety of configurations, such as crossing, bending, kissing or diverging. Fiber reconstruction algorithms are proposed to estimate as accurately as possible their orientations within a voxel.

As mentioned in , DTI was the first method to allow mapping a fiber orientation from a magnetic resonance image. Unfortunately, the diffusion tensor model has a number of important limitations that make it an inaccurate method to predict fiber orientations. The diffusion tensor can provide a fair approximation to a multivariate Gaussian diffusion process, but fails for heterogeneous *in vivo* fiber architectures. First, and most importantly, the tensor model assumes that there is a single eigensystem in each voxel (i.e. it only possesses a single maximum). According to[16], up to one third of white matter voxels contain more than one fiber bundle population. When a voxel contains contributions from different bundles with distinct orientations, the major eigenvector will, in general, no longer correspond to the orientation of any of the fiber tracts present[10], but to the mean orientation. Thus, the multi-modality of the observed diffusion signal precludes the standard tensor reconstruction from providing valuable orientational information, as depicted in Fig. 2.2. Second, it has been demonstrated[98] that the peaks of the diffusivity profile do not necessarily correspond to the orientations of the distinct fiber populations. Third, the mono-exponential signal decay model (see Eq. 1.26) has been shown to be invalid for high diffusion-weighting values *in vivo*[35]. And finally, it is observed that the diffusion of water molecules seen in the brain does not generally obey a Gaussian distribution law[71]. Thus, the Gaussianity assumption inherent to the diffusion tensor model may not be accurate enough. Hence, applications

relying on the diffusion tensor model will provide unreliable results for tractography applications.



(a) (b) (c)

Figure 2.2 The diffusion tensor imaging method pitfall. The DTI reconstruction method fails to determine accurately the orientation of fiber bundles when multiple fibers reside in the same voxel. Since the model assumes that there exists a single dominant fiber orientation at each voxel, when fibers' orientation is not coherent inside a voxel, DTI estimates an average orientation, and thus is unable to correctly resolve the true fiber orientations. (a) shows the isotropic diffusion model provided by the diffusion tensor; (b) shows fiber bundles (represented as black sticks) crossing at a given angle, and uninformative orientation information that the DTI model would provide in such a case (i.e. true fiber directions cannot be inferred from the ellipsoid maximum); and (c) shows a coherently oriented fiber bundle (black stick), whose orientation can be correctly estimated by the tensor model (ellipsoid).

Efforts to resolve this Intra-voxel Orientational Heterogeneity (IVOH) have been put in the last two decades. A good review can be found in[106]. Since the adoption of the High Angular Resolution Diffusion Imaging (HARDI) acquisition protocol, works attempting at solving IVOH have since focused on the development of new algorithms to

reconstruct the displacement distribution (or features thereof). The Diffusion Spectrum Imaging (DSI) (see) technique is probably the sole exception requiring a modified pulse sequence.

Although several criteria may be used to classify the fiber orientation reconstruction methods, the model criterion is used in this work. Algorithms are divided into two categories, namely, *model-based* reconstruction algorithms, and *model-free* reconstruction algorithms. As the names suggest, the former are based on the assumption of a model or distribution that is to be identified by the sampled data. The latter, on the other hand, make no assumption about the nature of the diffusion process.

This chapter presents and compares different model-based and model-free methods for the reconstruction of orientation distribution functions using sampled diffusion MRI data.

2.1 High Angular Resolution Diffusion Imaging

A popular diffusion data acquisition protocol, proposed by Tuch *et al.*[120,118] in the late 1990s/early 2000s, is High Angular Resolution Diffusion Imaging (HARDI), which involves acquiring diffusion information for a single b-value (single shell) in several gradient directions uniformly spread on a sphere[119].

The diffusion process may exhibit non-Gaussian behavior if the diffusion is restricted[36], or if there is slow exchange between partial volume elements exhibiting Gaussian diffusion. Traditional tensor imaging techniques cannot resolve such behavior, and higher angular resolution techniques are required. Tuch showed[120] evidence that human brain white matter regions containing heterogeneous fiber orientations show non-Gaussian diffusion when sampling the apparent diffusion coefficient at higher angular resolutions.

While this protocol allows for resolving the angular structure of the neural fibers, it does not provide information about the radial signal decay, which is sensitive to white matter anomalies[103].

2.2 Model-based Reconstruction Algorithms

2.2.1 Multi-Tensor Models

High b-value diffusion-weighted imaging revealed non-mono-exponential behavior of the signal attenuation[94 93 9] in white matter. This finding prompted some works in which the original DTI model was extended to contain two or more diffusion tensors[94 33]. In multi-tensor models (also referred to as Multiple-Diffusion Tensors model (MDT), or Gaussian Mixture Models (GMMs)), the diffusion signal is modeled as arising from a finite mixture of Gaussian diffusion processes in slow-exchange (i.e. the exchange between compartments should be extremely slow in relation to the experimental time scale).

If only one single fiber bundle is present in each voxel, the Gaussian model (Eq. 1.27) is accurate enough. For N fiber bundles without water exchange, the linearity of the Fourier transform (relating the diffusion propagator and the Stejskal-Tanner experiment, see) yields the multi-tensor model. In order to address the "partial volume effects" (i.e. contributions from multiple fibers in a single voxel), multi-tensor models[3] assume N diffusion compartments with little or no (material) exchange during measurement time. Each compartment is parametrized by a symmetric 3×3 diffusion tensor $\mathbf{D}_i$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$. The signal fractions $f_i \in [0, 1]$ sum to unity. The signal is

then $S(\mathbf{g})$ predicted as

$$S(\mathbf{g}) = S_0 \sum_{i=0}^{N} f_i e^{-b\mathbf{g}^T \mathbf{D}_i \mathbf{g}} \tag{2.1}$$

where S_0 is the non diffusion-weighted signal; N is the number of fibers; b is the diffusion weighting; $\mathbf{g}$ is the diffusion-sensitizing gradient; f_i is the partial volume fraction of the i-th fiber; and $\mathbf{D}_i$ is the corresponding diffusion tensor.

As it can be drawn from Eq. 2.1, the algorithm requires applying *a priori* knowledge about the underlying micro-architecture to select the appropriate number of fibers.

Multi-tensor models are not able to estimate simultaneously the diffusivity tensor and the relative volume fractions of fibers. Thus, in order to restrict the solution space, additional constraints are imposed. Many authors[118 6] assume equal eigenvalues on all diffusion tensors; others[82] prefer to fix the volume fraction instead. In either case, a very common assumption is axial symmetry ($\lambda_2 = \lambda_3$).

The "ball-and-stick" model[17] assumes that all diffusion tensors have equal principal eigenvalues (λ_1), a single "ball" (isotropic) compartment (i.e. $\lambda_1 = \lambda_2 = \lambda_3$), while the remaining "stick" compartments are infinitely anisotropic (i.e. perfectly linear) ($\lambda_2 = \lambda_3 = 0$).

The multi-tensor approach has a number of significant shortcomings, including the need for an iterative reconstruction scheme, nonlinear optimization schemes, which can be unreliable and inefficient, the possibility of model misspecification, the ability to describe only a discrete distribution of fiber orientations, and the inability to capture non-Gaussian behavior within each fiber compartment.

Extensions to this model have been proposed to allow the inference on multiple (i.e. more than two) fiber orientations, but they lead to complicated model selection problems, such as those adopted by Hosey *et al.* in[66] and Behrens *et al.* in[16]. Melie-García

et al. developed a Bayesian framework to solve the model selection problem, but the method was computationally expensive due to the use of Markov Chain Monte Carlo (MCMC) simulations.

Furthermore, the inverse problem of multiple-tensor fitting is ill-posed. Pasternak *et al.* suggested[101] to use regularization terms to stabilize the problem.

2.2.2 Multi-Compartment Models

Multi-compartment models consist in a variation of the multi-tensor models explained in . As opposed to the latter, where the multiplicity of the tensors models the intra-voxel fiber heterogeneity, in this case, the multiplicity in the tensors is given by a finite mixture model that describes an exchange model between the intracellular and extracellular compartments.

In these works, it is assumed that the "fast" (or rapidly decaying) diffusion component can be associated with the extracellular space, and the "slow" (or slowly decaying) diffusion component can be associated with the intracellular space. However, according to[10], if this slowly diffusing component arises from restricted diffusion (a non-Gaussian process), fitting the overall signal decay to a bi- or multi-exponential tensor model, which assumes Gaussian diffusion, is inappropriate.

The Composite Hindered and Restricted Model of Diffusion (CHARMED) proposed in[10] provides an implementation of a multi-compartment model. CHARMED expresses the signal decay observed in white matter in terms of Gaussian (hindered) and non-Gaussian (restricted) contributions. For noise-free environments, the diffusion model

can be written as

$$S(\mathbf{q},\Delta) = \sum_{i=1}^{M} f_i^h S_i^h(q,\Delta) + \sum_{j=1}^{N} f_j^r S_j^r(q,\Delta) \tag{2.2}$$

where

$$S^h(\mathbf{q},\Delta) = \exp\left[-4\pi^2\left(\Delta - \frac{\delta}{3}\right)\mathbf{q}^T\mathbf{D}\mathbf{q}\right], \tag{2.3}$$

$$\begin{aligned} S^r(\mathbf{q},\Delta) &= S_\perp(\mathbf{q}_\perp,\Delta)\cdot S_\parallel(\mathbf{q}_\parallel,\Delta) \qquad (2.4)\\ &= \exp\left[-4\pi^2|\mathbf{q}_\parallel|^2\left(\Delta - \frac{\delta}{3}\right)\mathbf{D}_\parallel\right] \\ &\quad\cdot \exp\left[\left(-4\pi^2 R^4|\mathbf{q}_\perp|^2\tau\frac{1}{\mathbf{D}_\perp}\right)\left(\frac{7}{96}\right)\left(2-\frac{99}{112}\right)\left(\tau\frac{R^2}{\mathbf{D}_\perp}\right)\right] \qquad (2.5)\end{aligned}$$

where h denotes the hindered compartment contribution, and r the restricted compartment contribution; $\perp$ denotes the contribution from the displacements perpendicular to the axis of the fiber, and $\parallel$ the contribution from the diffusion aligned with the fiber axis; M is the number of distinct hindered compartments, and N is the number of distinct restricted compartments (not necessarily equal to M); τ is half the echo time (TE); and R is the net displacement vector modulus.

The model assumes that one contribution to the net signal decay arises from hindered diffusion in the extra-axonal volume (including extra- and intracellular spaces), while another contribution to the net signal decay arises from restricted diffusion in the intra-axonal volume. The extra-axonal compartment's diffusion properties are characterized by an effective diffusion tensor, and an intra-axonal compartment, whose diffusion properties are characterized by a restricted model of diffusion within cylinders.

Gaussian signal attenuation is observed at low b-values; non-Gaussian attenuation arises from sampling at high b-values. To estimate various microstructural parameters

of the composite model, CHARMED thus requires sampling at both high and low b-values obtained along different directions.

Two important problems accompany the use of multi-compartment models. First, the number of such compartments has to be specified *a priori*, presenting a model-selection problem. Second, the parameter estimation requires a nonlinear fitting, which is sensitive to noise and to the number of measurements, and tends to perform poorly when more than two fiber orientations are present.

Although CHARMED seems to compare favorably to the standard DTI model or a dual tensor model for any given noise level (inter-fiber angles as small as 30° are reported[7]), this is at the cost of prohibitive scanning times (up to 18 h) due to the high b-values (as high as 44000 s/mm^2) required by the method.

2.3 Model-free Reconstruction Algorithms

2.3.1 Diffusion Spectrum Imaging

A mathematical alternative to the tensor model is the DSI method, which does not assume an analytical form of the diffusion process. In DSI, the ODF is reconstructed by sampling the diffusion signal on a Cartesian grid, Fourier transformation, and then radial projection.

The Fourier relationship between the spin echo magnitude and the PDF allows for direct reconstruction of the PDF by inverse Fourier transform of the diffusion signal with respect to the displacement wave vector (see Eq. 1.25 in Sec. 1.4.1). The DSI method exploits this relationship to reconstruct the PDF within each imaging voxel. The PDF is then radially projected onto the sphere (see Sec. 1.4.3), yielding the estimated ODF.

According to the results in[118], the Diffusion Spectrum Imaging method allows for the

resolution of complex fiber structures, such as intra-voxel fiber crossing and divergence. Nonetheless, results show multiple peaks in the detected ODF function in gray matter regions, and it is unclear whether these correspond to genuine anatomic structure. The main shortcoming of DSI is the long acquisition times required, since it requires an order of magnitude more measurements than DTI to get an equivalent fiber angular resolution power[4].

Furthermore, Canales-Rodríguez *et al.* showed[21] that the exact mathematical expression for the ODF was not correctly approximated by DSI.

2.3.2 Q-Ball Imaging

Instead of forcing the diffusion anisotropy data into a group of tensors, the HARDI signal can be reconstructed model-independently using a spherical inversion mathematical tool called the Funk-Radon transform, also known as the spherical Radon transform, resulting in a reconstruction method termed q-Ball Imaging (QBI). In q-ball imaging, the ODF reconstruction scheme is based on the finding that integration over a great circle in reciprocal **q**-space is equivalent to the Hankel transform of the planar projection of the diffusion function. QBI approximates the ODF by the Funk transform of the diffusion MRI signal at a fixed radius in **q**-space. The Funk transform, a mapping between functions of the sphere, is a generalization of the Radon transform to the sphere. For a given point on the sphere referred to as a *pole*, the value assigned by the Funk transform at the pole is the integral over the associated equator (i.e. the great circle):

$$\mathcal{G}(\mathbf{u})[f(\mathbf{u})] = \int_{S^2} f(\mathbf{u})\delta(\mathbf{u}^T\mathbf{w})d\mathbf{w} \tag{2.6}$$

where $\mathcal{G}$ denotes the Funk transform; $f(\mathbf{u})$ is a function on the sphere S^2; δ is the Dirac delta function; and $\mathbf{u}$ and $\mathbf{w}$ are unit vectors on the sphere.

The Funk transform of the diffusion signal $S(\mathbf{q})$, where $\mathbf{q}$ is sampled on a sphere, yields an ODF that approximates the ODF obtained by explicit radial projection:

$$\psi(\mathbf{r}) \approx \mathcal{G}_{q'}[S(\mathbf{u})] \tag{2.7}$$

where q' is the radius of the sampling shell.

One of the most salient contributions of QBI was spherical sampling. QBI requires sampling on one or more spherical grids. Tuch argued that the ability to sample directly on the sphere allows to spend the signal acquisition time more efficiently on *angular* resolution. As opposed to this, Cartesian Fourier *spatial* resolution contributes relatively inefficiently to the desired endgoal of angular resolution. In conventional Cartesian Fourier reconstruction, significant time and effort are invested in the acquisition near the origin and tails of the reciprocal space, where the angular contrast-to-noise ratio is relatively low. In the context of spherical sampling, the acquisition can be targeted to the spatial frequency band where the angular contrast-to-noise is allegedly greater [118].

The radial projection used to construct the ODF discards all of the radial information contained in the diffusion function. Hence, the ODF does not contain the radial information which was originally present in the Cartesian diffusion function, but preserves the salient angular contrast. Tuch states that spherical sampling provides a more natural framework for describing angular resolution, which can be defined as the upper-bound of the angular distances between the sampling points.

QBI can resolve multiple intra-voxel fiber orientations [118] [117], and does not require any assumptions on the diffusion process, such as Gaussianity or multi-Gaussianity.

Aganj *et al.* argued[1] that the definition of the ODF used in the original QBI, computed as a linear radial projection of the spin displacement PDF, does not take into account the quadratic growth of the volume element with respect to its distance from the origin. This inaccurate formulation reportedly distorts the ODF, and requires artificial post-processing (such as the introduction of normalization constants). Aganj *et al.* re-derived the ODF expression for QBI via Fourier analysis from the proper definition of the ODF in Constant Solid Angle (CSA), and used spherical harmonics (see) for the implementation. Although, results reported provide high angular contrast (close to 30°) and do outperform QBI, a high b-value was used ($b = 4800$ s/mm^2) (in the synthetic experiments), fiber dominance was not tested, and the method requires long scanning times.

An extension to the QBI algorithm, termed Exact q-Ball Imaging (EQBI), was provided[23] by Canales-Rodríguez *et al.*. EQBI provides a different method to calculate the ODF analytically using multiple spherical **q**-space shells.

As an alternative to the QBI method, Tristán-Vega *et al.*[116] introduced the Orientation Probability Density Function (OPDF) for fiber tracts in the brain white matter for HARDI data. In QBI, the ODF is estimated as the radial projection (integral) of the 3D PDF. Instead, Tristán-Vega *et al.* proposed to marginalize the radial part of the PDF to represent the orientation distribution, and named Orientation Probability Density Transform (OPDT) the estimator of such a function. In noise-free environments, the lower inter-fiber angle resolution power is close to 40° for $b = 3000$ s/mm^2, and is close to 50° for lower **b**-values ($b = 1200$ s/mm^2). Results on the effect of dissimilar relative fiber volume fractions were not reported by the authors. Furthermore, the OPDF estimator may yield negative values which have to be corrected.

2.3.3 Spherical Harmonics

Several studies proposed to represent multi-modal diffusivity (i.e. ADC) profiles using a Spherical Harmonics (SH) expansion[57 5 6 47]. For a formal definition of the SH basis functions, see Sec. A.1.

The diffusion measurements along different encoding directions can be expressed as rotations in three dimensions relative to the (unknown) principal axis system of the diffusion. In the HARDI technique, these measurements are along a set of directions covering the range of the spherical coordinates (θ, ϕ). The HARDI diffusion measurements have an inherent spherical symmetry because they are made by a series of 3D rotations. Systems with spherical symmetry are often more conveniently handled in the spherical basis $\mathbf{e} = r, \theta, \phi$. Rotation matrices transformed into the spherical basis are the *spherical harmonics*. Tensors transformed into this representation are called *spherical tensors*. Thus, the process of rotating a diffusion tensor can be reformulated by expressing the diffusion tensor in an irreducible form in which its individual components transform separately under rotations affected by spherical harmonic components.

This rationale can equally be applied to the general case of a HARDI measurement of a voxel of unknown fiber composition. In HARDI, the measured apparent diffusion coefficient $D(\Omega)$ is an arbitrary real function. Letting $Y_l^m(\theta, \phi)$ denote an l-order, m-degree SH basis function in the (θ, ϕ) direction, n_s be the number of discrete HARDI sampled data at a given (θ, ϕ) point on the sphere, N the number of SH used in the approximation of order $l (n_s \gg N)$, $\mathbf{X} = (x_1, \ldots, x_{n_s})^T$ the vector of measured ADCs

$D(\mathbf{g})$, $\mathbf{C} = (c_0^0, c_0^{-1}, c_1^0, \ldots, c_l^l)^T$ the vector of SH coefficients, and

$$\mathbf{B} = \begin{pmatrix} Y_0^0(\theta_1, \phi_1) & Y_1^{-1}(\theta_1, \phi_1) & \cdots & Y_l^l(\theta_1, \phi_1) \\ \vdots & \vdots & \ddots & \vdots \\ Y_0^0(\theta_{n_s}, \phi_{n_s}) & Y_1^{-1}(\theta_{n_s}, \phi_{n_s}) & \cdots & Y_l^l(\theta_1, \phi_{n_s}) \end{pmatrix} \tag{2.8}$$

the $n_s \times N$ matrix of discrete SH, the approach seeks for the spherical harmonic series that passes nearest to the discrete sampling points on the sphere. Hence, $\mathbf{X} = \mathbf{BC} + \mathbf{E}$ where the error vector $\mathbf{E}$ should be small. This system of overdetermined equations is solved with linear least square sums over the columns of $\mathbf{E}$ by minimizing $\|\mathbf{X} - \mathbf{BC}\|^2$, yielding

$$\mathbf{C} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{X} \tag{2.9}$$

The estimated ADC profile of order l is thus recovered by evaluating

$$D(\mathbf{g}(\theta, \phi)) = \sum_{k=0}^{l} \sum_{m=-k}^{k} c_k^m Y_k^m(\theta, \phi) \tag{2.10}$$

for any (θ, ϕ) outside the discrete measurements $\mathbf{X}$ or in the discrete linear case, by simple matrix multiplication,

$$\mathbf{X} = \mathbf{BC} \tag{2.11}$$

The above solution works so long as the noise level is small relative to the HARDI diffusion signal level. For cases where noise is larger, Descoteaux *et al.* proposed[47] a generalization of the standard least-squares evaluation method to include a regularization (smoothness maximization) criterion. Hence, they model the Apparent Diffusion Coefficient with Higher-Order Diffusion Tensor (HODT) described by Özarslan *et al.*[97]. That is, a linear transformation is found taking the coefficients of the spherical harmon-

ics series to the independent elements of the HODT.

This model has several advantages: no *a priori* knowledge about the diffusivity profile is assumed, and non-Gaussian diffusion can be potentially described.

Although results show[48] that it is possible to recover voxels showing isotropic, single fiber anisotropic, and multiple fiber anisotropic diffusion behavior, the recovered inter-fiber angle is never lower than 45° (for $b = 3000$ s/mm^2).

One of the drawbacks of the use of SH is the need of *a priori* assumptions (or statistical tests) to decide on the SH order (i.e. the complexity of the model). Characterization of sharper fODFs often requires contributions from higher order harmonics whose omission leads to ringing effects. Furthermore, higher order harmonic representations are more sensitive to noise since the magnitudes of individual SH coefficients are reduced[6].

The inability of the basis of SH to sparsely represent diffusion signals led to the proposal of Spherical Ridgelets (SR) in[90], where it was shown that the order of Spherical Ridgelets required to represent the HARDI signals with a precision exceeding the precision of their representation using SH is reduced by a factor of 5 on average. Spherical Ridgelets are derived from a set of wavelet functions through the application of the Funk-Radon transform.

2.3.4 Spherical Deconvolution

To avoid the problems of the multi-compartment models, Tournier *et al.* proposed[114] the Spherical Deconvolution (SD) framework to estimate the underlying arrangement and orientations of fiber populations. The key idea of the spherical deconvolution method is to assume that there is a distribution, rather than a discrete number, of fiber orientations in each voxel. Under this assumption, the sampled HARDI signal is the result of the (spherical) convolution of the underlying fODF, which is a real-valued function on the

unit sphere, with some kernel function representing the response derived from a single fiber bundle, which can be considered as a sort of system impulse response function.

Under this approach, the model of diffusion signal generation can thus be written as

$$S(\theta, \phi) = F(\theta, \phi) \otimes R(\theta) \tag{2.12}$$

$$= \int_0^{2\pi} \int_0^{\pi} F(\theta', \phi') R(\gamma') \sin \theta' d\theta' d\phi' \tag{2.13}$$

where $S(\theta, \phi)$ is the measured HARDI signal; $F(\theta, \phi)$ is the fODF sought along each direction (θ, ϕ); $R(\theta)$ is the fiber response function; γ' is the angle between directions given by θ, ϕ and θ', ϕ'.

SD approaches proceed by deconvolving this response from the observed diffusion signals to estimate the underlying orientation distribution function. As opposed to model-free techniques that estimate the diffusion ODF, the output from SD is directly the fiber ODF itself.

The kernel not only models the physical process, but also the degrading factors affecting the measured signal, such as the angular blurring introduced by HARDI acquisition[45]. Response functions may either be estimated on a voxel-by-voxel basis or assumed to be a constant across the image.

Spherical Harmonic Deconvolution (SHD)[114] and its constrained version (Constrained Spherical Harmonic Deconvolution (also referred to as Constrained Spherical Deconvolution) (CSHD))[112] are implementations of spherical deconvolution, proposed by Tournier *et al.*, in which it is assumed that both the observed diffusion signal and the single-fiber response functions may be adequately represented by a (truncated) linear combination of spherical harmonic basis functions. This reduces the number of parameters that need to be stored (reducing memory requirements) and, more importantly, reduces the de-

convolution process to a computationally trivial operation in SH space (analogous to convolution/deconvolution in Fourier space), reducing overall processing requirements.

Tournier *et al.* proposed[112] two alternatives to their initial SHD approach (also referred to as "filtered spherical deconvolution" –filtered SD– due to the use of a low-pass filtering on WM voxels of parallel fibers to minimize noise susceptibility effects, i.e. spurious negative lobes), named CSHD (or also Constrained Spherical Deconvolution (CSD)) and Super-resolved Constrained Spherical Deconvolution (super-CSD). Both introduce a regularization term using a modified version of the ℓ_2-norm (Tikhonov) penalization method.

The Tikhonov regularization involves the minimization of the weighted sum of two terms:

$$\|\mathbf{Af} - \mathbf{b}\|^2 + \lambda^2 \left\|\mathbf{L}(\mathbf{f} - \mathbf{f}')\right\|^2 \tag{2.14}$$

where $\mathbf{f}$ are the SH coefficients of the ODF; $\mathbf{b}$ are the measured signal intensities; $\mathbf{L}$ is a matrix that (indirectly) provides the amplitude of current estimate of the ODF for the set of directions; and $\mathbf{A}$ is the problem matrix $\mathbf{A} = \mathbf{QR}$, where $\mathbf{Q}$ maps the SH coefficients to the corresponding intensities along the HARDI diffusion signal, and $\mathbf{R}$ performs the spherical convolution with a response function (kernel).

The first part in Eq. 2.14 corresponds to the data-driven part, and is a simple linear least-square fit of the solution $\mathbf{f}$ (the parameters to be estimated) to the data $\mathbf{b}$, via the problem matrix $\mathbf{A}$. The second part is the regularization part, where the parameter λ controls the relative weighting between the two terms.

Assuming that $\mathbf{f}' = 0$, and discarding at each step fiber amplitudes below a given threshold, the problem reduces to iteratively finding an estimate of the set of fiber

orientations until convergence (i.e. no change in the matrix $\mathbf{L}$):

$$\mathbf{f}^{(i+1)} = \arg\min_{0} \left\{ \left\| \mathbf{A}\mathbf{f}^{(i)} - \mathbf{b} \right\|^2 + \lambda^2 \left\| \mathbf{L}\mathbf{f}^{(i)} \right\|^2 \right\} \tag{2.15}$$

The results provided by Tournier *et al.* under a number of conditions (different values of the weighting parameter, (truncation) order of the SH coefficients, SNR, etc.) show a lower bound for the inter-fiber angle at 40° (for a relatively standard b-value of $b = 3000$ s/mm^2).

Where the assumption that all single fiber populations within an image volume exhibit the same diffusion profile is valid, CSD can be shown to provide accurate estimates of the fiber orientations, as shown in[113]. In real datasets, however, there are many situations in which this fundamental assumption is not valid, reducing the method's performance and making it impractical.

A sensitivity analysis conducted by Parker *et al.*[100] in both single and crossing-fiber configurations showed that CSD produces spurious fODF peaks as the discrepancy between the estimated single-fiber response function and the target response increases.

Reconstruction of the fODF by spherical deconvolution is an ill-posed problem, and multiple solutions may exist, some of which may be physically implausible (such as negative peaks in the fODF). To address this, non-negativity constraints are often placed upon the deconvolution process. However, instabilities may persist in the deconvolution process.

One of the most notable contributions in overcoming these issues was proposed in[45]. In this work, Dell'Acqua *et al.* used a modified version of the Richardson-Lucy (RL) algorithm[104 79] for the deconvolution assuming a Gaussian model for the noise distribution. The RL algorithm follows a statistical Bayesian approach to the deconvolution

problem, and implements an iterative estimation scheme for approximating the solutions of a maximum-likelihood problem. The classical version of the RL algorithm was originally developed for astronomical image reconstruction and denoising in the presence of Poisson noise.

The RL algorithm for spherical deconvolution (SD-RL) can be compactly expressed in matrix-vector notation as[45]

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} \frac{\mathbf{H}^T \mathbf{S}}{\mathbf{H}^T \mathbf{H} \mathbf{f}^{(k)}} \tag{2.16}$$

where k refers to the k-th iteration step; $\mathbf{f}$ is the column-vector of length N which contains the values of fODF along N-directions uniformly distributed on the sphere; $\mathbf{S}$ is the column-vector of length M containing the values of the HARDI signal, acquired along M-directions uniformly distributed on the sphere; and $\mathbf{H}$ is an matrix which acts as a circulant matrix on a spherical surface; every column of length M contains the values of the fiber response profile oriented along one of the N-directions.

This framework intrinsically satisfies the non-negativity constraint, and controls the instabilities of the estimation process.

Due to the spherical symmetry of both the diffusion problem and the specific applied gradient direction set, the coupled symmetric sampled data can be averaged together in order to improve the diffusion-weighted SNR and reduce cross-term effects between imaging and diffusion gradients[45]. Hence, the size of the vectors and matrices in Eq. 2.16 can be halved to speed up the computation.

Convergence properties show[45] that this algorithm, in the case of Gaussian noise, converges to the Non-negative Least Squares (nnLS) solution.

The SD-RL approach fails in pure isotropic voxels, where only a number of randomly

oriented spurious spikes can be obtained. The same problem has also been reported for QBI and Persistent Angular Structure Magnetic Resonance Imaging (PASMRI)[45]. Dell'Acqua *et al.* proposed to workaround the problem by either discriminating isotropic voxels using their FA values[45] or using the algorithm's convergence properties[46].

The results provided by Dell'Acqua *et al.* showed that SD-RL outperforms other methods (such as QBI) in a number of situations (such as different diffusivity, fiber dominance, or SNR values), but the tested inter-fiber angles in the synthetic simulations was only 60°[45].

More recently, Dell'Acqua *et al.* proposed[46] a variation of the RL method to address the partial volume effects from isotropic tissue, such as gray matter or cerebrospinal fluid (CSF), and which may degrade the spherical deconvolution results, leading to the partial deactivation of the non-negativity constraint. The alternative approach, named Damped Richardson-Lucy (dRL) algorithm, is based on an adaptive regularization scheme.

The dRL variation can be written in matrix notation as

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} \left[\mathbf{1} + \mathbf{u}^{(k)} \left(\frac{\mathbf{H}^T\mathbf{S} - \mathbf{H}^T\mathbf{H}\mathbf{f}^{(k)}}{\mathbf{H}^T\mathbf{H}\mathbf{f}^{(k)}} \right) \right] \tag{2.17}$$

where $\mathbf{u}^{(k)}$ is a vector that performs the damping operation on each element of $\mathbf{f}$, and is defined as

$$\mathbf{u}^{(k)} = 1 - \mu\mathbf{r}^{(k)} \tag{2.18}$$

with

$$\mathbf{r}^{(k)} = 1 - \frac{\left(\mathbf{f}^{(k)}\right)^{\nu}}{\left(\mathbf{f}^{(k)}\right)^{\nu} + \eta^{\nu}} \tag{2.19}$$

where $\mathbf{r}$ is a vector controlling the damping for each fiber orientation function com-

ponent; ν is a geometrical parameter describing the damping curve and how fast the damping turns on and off; η acts as a threshold parameter controlling where the damping starts according to the fiber orientation function amplitude; and μ modulates the damping across voxels and is used to obtain a preliminary classification of brain tissue based on the voxel anisotropy ($\mu = \max\{0, 1 - 4\,\mathrm{std}(s)\}$).

Damping occurs when differences between data and recovered object are close to or lower than the noise level; for larger differences, the algorithm acts as the original RL form. Thus, spurious component amplification is prevented.

Fiber orientation recovery results with the damped version of the RL algorithm show a reduced false positive rate for a broad range of volume fraction values compared to the standard RL version. Furthermore, dRL exhibits a low overall target response mismatch sensitivity for sub-optimal choices of the single-fiber response function, with a calibration response function for a highly anisotropic fiber being optimal[100]. Nevertheless, dRL demonstrates a reduced ability to resolve low anisotropy crossing-fibers compared to CSD, and the lower bound of the inter-fiber angle resolved remains relatively high (50-60°, depending on the number of iterations allowed, and the allowed overestimation rate). Furthermore, the dRL algorithm exhibits tolerance to system impulse response imprecision.

The main issue for either the standard or the damped versions of the RL algorithm seems to be one of convergence. The algorithm exhibits the "semi-convergence" property; first the solution converges to the true value and then diverges as iterations proceed, leading to noise amplification and generation of artifacts such as spurious components. Thus, the algorithm returns the best results shortly before the fODF deconvolution fully converges, while actually reaching convergence can often result in over-fitting (i.e. an increased rate of false positives or the over-representation of the signal's noise compo-

nent)[46].

Compared to the multi-compartment models, the spherical deconvolution framework has two significant advantages. First, it does not require the specification of the number of underlying fibers before computing the fODF. Second, the spherical deconvolution methods often lead to a linear system which can be solved efficiently while the multi-compartment models usually involve a computationally expensive nonlinear fitting process.

Another advantageous property of the Spherical Deconvolution interpretive framework is that it works directly on the acquired data, therefore avoiding the use of other representation domains or basis, such as the SH. Furthermore, obtaining the fODF directly allows SD methods to resolve fiber crossings with small inter-fiber angles in datasets acquired within a clinically feasible scan time[53].

Contrary to SD methods using direct inversion to solve fiber crossing (see, for example,[114] and[6]) the iterative nature of the dRL method warrants that, thanks to semi-convergence, the solution found approximates well the true value, without unacceptable instabilities due to noise amplification.

Notwithstanding, the iterative nature of the algorithm makes it slower than direct inversion methods. Other computational issues that arise in the application of the deconvolution framework concern the stability of the deconvolution in the presence of noise, and the ability to accurately recover multiple fiber orientations from relatively low angular resolution imaging data[70].

The main assumption inherent in the Spherical Deconvolution technique is that the diffusion characteristics of fiber tracts can be assumed approximately constant, such that the response function measured for a typical coherently oriented fiber population is constant throughout the brain[114 6]. This premise may be broken in regions where the

white matter fibers have significantly different diffusion characteristics, due for example to (i) different myelination levels, (ii) axonal diameters, or (iii) axonal densities[13].

Using this simplifying hypothesis, the SD approach assume that the fODF within a voxel can be obtained by deconvolving a 'common' single-fibre response function from the observed set of diffusion signals. In practice, this common response function is not known *a priori*, and thus, an estimated fiber response must be used. Although SD methods produce inaccurate fODF estimates when the single-fiber response function does not estimate well the observed data, it only affects the estimated volume fractions of the various fiber populations, not their respective orientations[114 6]. The sensitivity to an incorrect single-fiber response function differs across SD methods[100].

(a) (b)

Figure 2.3 Spherical sampling. Spherical deconvolution methods use a sampling strategy that relies on probing the signal at evenly distributed points over a unit sphere: (a) sampled points; (b) faces defined by the edges linking the sampled points. There are many solutions to the issue of packing points on a sphere. In this case, the number of points (vertices) is $V = 492$ and the number of faces $F = 480$.

More recently, another SD method, named Robust and Unbiased Model-Based Spherical Deconvolution (RUMBA-SD), was introduced by Canales-Rodríguez *et al.*[53]. The method is intended to deal with realistic MRI noise, based on the dRL algorithm adapted

to the Rician likelihood model (see Sec. B.1). An iterative scheme based on solving

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} \frac{\mathbf{H}^T \left[\mathbf{S} \frac{I_n(\mathbf{SHf}^{(k)}/\sigma^2)}{I_{n-1}(\mathbf{SHf}^{(k)}/\sigma^2)} \right]}{\mathbf{H}^T \mathbf{Hf}^{(k)}} \tag{2.20}$$

is obtained. Here, I_n (I_{n-1}) is the modified Bessel function of the first kind of order n ($n-1$); and σ^2 is the noise variance.

It was demonstrated that, compared to the dRL method, RUMBA-SD offers increased ability to resolve fiber crossings at small inter-fiber angles and to better detect non-dominant fibers in both synthetic and human brain datasets.

2.4 Other methods

2.4.1 Diffusion Orientation Transform

One major difficulty with employing HARDI in studies involving orientation mapping has been that the peaks of the diffusivity profile do not necessarily yield the orientations of the distinct fiber populations[98]. Hence, Özarslan *et al.* propose[98] the Diffusion Orientation Transform (DOT), which transforms the diffusivity profiles into probability profiles whose peaks correspond to distinct fiber orientations.

The method expresses the average displacement probability (see 1.2) in spherical coordinates. Under the mono-exponential attenuation assumption, the radial part of the integral is evaluated analytically; the probability values on a fixed radius can be reconstructed either directly or parametrically in terms of a Laplace series / Least Squares (LS).

In either case, the average displacement probability $P(r)$ is re-written as the probability of finding the particle, initially at the origin, at the point $R_0 r$; hence, the interest

lies in computing the probability values on a sphere of radius R_0.

In the parametric reconstruction alternative, this probability can be re-written using a Laplace series expansion as

$$P(R_0 r) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} p_{lm} Y_l^m(\mathbf{r}) \tag{2.21}$$

with

$$p_{lm} = (-i)^l \alpha_{llm} = (-1)^{-1/2} \alpha_{llm} \tag{2.22}$$

where the α_{llm} coefficients are given by a spherical harmonic transform; and $Y_l^m(\theta, \phi)$ is the spherical harmonic function.

An alternative, non-parametric interpretation of the probability is given by

$$P(R_0 r) = \sum_{l=0}^{\infty} \int d\mathbf{u} (-1)^{-1/2} \frac{2l+1}{4\pi} P_l(\mathbf{u} \cdot r) I_l(\mathbf{u}) \tag{2.23}$$

where $\mathbf{u}$ is a unit vector specifying the direction of the diffusion sensitizing gradient; P_l is the Legendre polynomial of l-th order; and $I_l(\mathbf{u})$ is a function of orientation.

Thus, the parametric reconstruction allows expressing the probabilities in terms of a Laplace series, whereas the non-parametric reconstruction allows for direct probability computation. Although it is simpler to implement the latter scheme as no Spherical Harmonic Transform (SHT) is necessary, when the LS is truncated, the number of probability points to be evaluated is smaller than the directions along which the probabilities are estimated, which saves computation time.

DOT estimates a contour of the diffusion propagator estimated at a defined displacement length R_0. The ODF obtained through the DOT model (see[22] for mathematical details on the derivation) provides less angular contrast than the fODF, and provides a

nonnegligible rate of false positives.

2.4.2 Simple Harmonic Oscillator-based Reconstruction and Estimation

Özarslan *et al.* proposed[96] an alternative series representation for the **q**-space diffusion MR signals in terms of a complete set of orthogonal Hermite functions, named Simple Harmonic Oscillator based Reconstruction and Estimation (SHORE). The three-dimensional variant requires multiple q-shell acquisitions, which requires longer scanning times.

Results reported by the authors do not allow for an effective comparison of the algorithm's performance. Furthermore, results presented in[96] were performed *ex-vivo* with an ultra high-field scanner (14.1 T).

2.4.3 Higher-Order Diffusion Tensors

In the presence of multiple fibers, the diffusion tensor signal is oblate or planar and there is no unique principal direction. Additionally, the maxima of the Apparent Diffusion Coefficient profile do not correspond to the true fiber orientations. One natural generalization is to model the ADC with HODTs (also named Higher-Order Tensors (HOTs) or Generalized Diffusion Tensor (GDT) framework)[97 77]. This model does not assume any *a priori* knowledge about the diffusivity profile and has the potential to describe non-Gaussian diffusion. It constitutes an alternative basis of functions on the sphere, equivalent to Spherical Harmonics (see Sec. 2.3.3). However, the HODT makes it unnecessary to evaluate the Spherical Harmonic Transform from the diffusivity profiles, which is a computationally difficult task.

In the HODT statement of the reconstruction problem, the distribution of the diffusivities is assumed to be characterized by a tensor of arbitrary rank, and is re-written as

$$D(\mathbf{g}) = -\frac{1}{b} \ln \frac{S(\mathbf{g})}{S_0} \tag{2.24}$$

where $\mathbf{g}$ is the direction of the diffusion sensitizing gradient.

Özarslan *et al.* sought[97] to fit this expression to a rank-2 and higher order tensor model by using linear regression methods.

Results reported by the authors of the HODT methods are limited to either synthetic phantoms or qualitative comparisons with no real performance analysis.

2.4.4 Persistent Angular Structure MRI

This algorithm, introduced[67] by Jansons and Alexander, proposes to determine another feature, the (radially) Persistent Angular Structure (PAS), of the probability density function of particle displacements. The PAS is the function $\tilde{p}$ of the sphere whose Fourier transform, when embedded in the 3D space on a sphere of radius r (i.e. the $\mathbf{q}$-space), best explains the normalized sampled data.

The method is derived from the principle of maximum entropy to obtain an expression for the function that contains the least information (i.e. the one with the largest entropy) subject to the constraints from the data.

The information content of a probability density function p defined over a set Ω is given by

$$I[p] = \int_{\Omega} p(\mathbf{x}) \ln p(\mathbf{x}) d\mathbf{x} \tag{2.25}$$

The substitution

$$p(\mathbf{x}) = \tilde{p}(\hat{\mathbf{x}}) r^{-2} \delta(|\mathbf{x}| - r)$$

where δ is the standard one-dimensional δ distribution; r is a constant; $\hat{x}$ is a unit vector in the direction of $\mathbf{x}$; and $\tilde{p}$ is the (radially) Persistent Angular Structure, is equivalent to projecting the angular structure from all radii onto the sphere of radius r, and ignoring any information about the radial structure in the data.

The relative information of the probability density function p with respect to the probability density function p_0 is given by

$$I[p; p_0] = \int_\Omega p(\mathbf{x}) \ln \left(\frac{p(\mathbf{x})}{p_0(\mathbf{x})} \right) d\mathbf{x} \tag{2.26}$$

The constraints on p from the data can be incorporated into the expression above using the method of Lagrange multipliers to yield

$$I[\tilde{p}] = \int \left[\tilde{p}(\hat{\mathbf{x}}) \ln \tilde{p}(\hat{\mathbf{x}}) - \tilde{p}(\hat{\mathbf{x}}) \sum_{j=1}^{N} \left(\lambda_j \exp\left(i\mathbf{q}_j \cdot r\hat{\mathbf{x}} \right) \right) - \tilde{p}(\hat{\mathbf{x}})\mu \right] d\hat{\mathbf{x}} \tag{2.27}$$

where $\mathbf{q}_j, 1 \leq j \leq N$, are the non-zero wavenumbers for the MRI measurements; the λ_j are Lagrange multipliers for the constraints from the data; and the Lagrange multiplier μ controls the normalization of $\tilde{p}$. The information content, $I[\tilde{p}]$, has a unique minimum at

$$\tilde{p}(\hat{\mathbf{x}}) = \exp \left[\lambda_0 + \sum_{j=1}^{N} \lambda_j \exp\left(i\mathbf{q}_j \cdot r\hat{\mathbf{x}} \right) \right] \tag{2.28}$$

Taking advantage of the fact that the probability density function of particle displacements has antipodal symmetry,

$$\tilde{p}(\hat{\mathbf{x}}) = \exp \left[\lambda_0 + \sum_{j=1}^{N} \lambda_j \cos(\tilde{\mathbf{q}}_j \cdot r\hat{\mathbf{x}}) \right] \tag{2.29}$$

Finally, the method proposed by Jansons *et al.* seeks to minimize the error function

$$\sum_{j=1}^{N}\left[\frac{E(\mathbf{q}_i)}{E_0}-\int \tilde{p}(\hat{\mathbf{x}})\cos(\tilde{\mathbf{q}}_j\cdot r\hat{\mathbf{x}})\right] \tag{2.30}$$

with respect to the $N+1$ $\lambda_j, j=0,\ldots,N$ parameters of $\tilde{p}$.

The above nonlinear least squares problem is solved using a LevenbergMarquardt algorithm (also known as the damped least-squares) (LMA) (also known as the Damped Least-Squares (dLS) method).

The nonlinear optimization and numerical integration operations involved make the PASMRI method impractical due to the long computation times required. An alternative to ease this computational burden proposes to substitute the maximum-entropy parametrization of $\tilde{p}$ with a linear basis. This linearization makes the problem more tractable. Yet, authors do not provide quantitative results of the algorithm, or numerical comparisons to other methods in terms of its angular resolution power.

2.4.5 Hybrid Diffusion Imaging

The Hybrid Diffusion Imaging (HIDY) method[126] estimates DTI, DSI, and QBI diffusion properties in a single experiment using a concentric multi-shell diffusion weighting encoding scheme. Each shell has a constant diffusion weighting, but the number of encoding directions is increased with each encoding shell to increase the angular resolution. Thus, the most interior shells are suitable for a diffusion tensor analysis, whereas the outer shells allow for high angular resolution reconstruction frameworks such as QBI. The entire data set from all shells is combined for both fast and slow diffusivity estimation, using DSI and multi-exponential (nonlinear least squares bi-exponential) fit. The technique suffers from several shortcomings, such as long scan times (30 min) and

limited angular resolution (45°) despite the use of high b-values ($b > 6000$ s/mm^2) used in the outer shells.

2.4.6 Spherical Polar Fourier Transform

The Spherical Polar Fourier (SPF) transform was postulated by Assemlal *et al.*[11] and further investigated by Cheng *et al.*[32 31] and Caruyer *et al.*[30] as a model-free algorithm to recover the EAP and the ODF in **q**-space from the coefficients of the signal reconstructed in the SPF basis.

In the original approach by Assemlal *et al.*, the estimation is based on the approximation of the signal by a series expansion made of Gaussian-Laguerre and spherical harmonic functions followed by a projection on a finite dimensional space. Furthermore, the introduced method considers the Rician noise model of the MRI samples.

In the orthonormal basis SPF expansion, the MR signal attenuation **E** can be rewritten as the following series

$$E(\mathbf{q}) = \frac{S(\mathbf{q})}{S(0)} = \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{n,l,m} R_n(\|\mathbf{q}\|) y_l^m \left(\frac{\mathbf{q}}{\|\mathbf{q}\|} \right) \tag{2.31}$$

where $n \in \mathbb{N}$ is the radial index; $l \in \mathbb{N}, m \in \mathbb{Z}, -l \leq m \leq l$ are the angular indexes; $S(\mathbf{q})$ is the diffusion MR signal at a point $\mathbf{q}$ in the **q**-space; $a_{n,l,m}$ are the series coefficients; y_l^m are the real SH; and R_n is an orthonormal radial basis function made of Gaussian-Laguerre (GL) functions.

While the radial component of the MR signal is reconstructed a normalized basis of generalized GL polynomial functions, the angular component of the signal is reconstructed by elementary angular functions based on the complex SH basis (for further details, see[11]).

In practice, the expansions in Eq. 2.31 must be truncated to some finite $n_{max} = N$ and $l_{max} = L$ values. Furthermore, a variational framework is introduced to regularize the estimated diffusion functions during the estimation process.

Although authors argue that the SPF approach enables to accurately estimate various features of the tissue micro-architecture compared to other approaches (such as QBI), quantitative data concerning the resolvable inter-fiber angles is missing.

2.4.7 Microstructure Modeling Methods

Other approaches trying to provide a deeper insight into the white matter micro-architecture have also emerged. These models, namely the Diffusional Kurtosis Imaging (DKI) and the Neurite Orientation Dispersion and Density Imaging (NODDI), attempt to adopt a model that is closer to a real neuronal tissue model: the intra- and extra-axonal spaces are modeled separately, and new features are introduced to account for the uncovered dissimilarities between both compartments.

The diffusion kurtosis model is an expansion of the diffusion tensor model that, in addition to quantifying the diffusion tensor, provides an estimate of the degree to which water diffusion in biological tissues is non-Gaussian using a mathematical construct called the kurtosis tensor (KT)[69 78]. This diffusion kurtosis thus represents the extent to which the diffusion pattern of the water molecules deviates from a perfect Gaussian curve.

The diffusion kurtosis model expresses the diffusion-weighted signal as:

$$S(\mathbf{u}) = S_0 \exp\left[-\mathbf{b}\mathbf{D}(\mathbf{u}) + \frac{1}{6}\mathbf{b}^2\mathbf{D}^2(\mathbf{u})\mathbf{K}(\mathbf{u})\right] \tag{2.32}$$

where $S(\mathbf{u})$ is the signal intensity at the echo time; $\mathbf{b}$ is the applied diffusion weighting;

S_0 is the signal in the absence of diffusion gradient sensitization; $\mathbf{D}(\mathbf{u})$ is the value of diffusion (i.e. the apparent diffusion coefficient) along direction $\mathbf{u}$; and $\mathbf{K}(\mathbf{u})$ is the value of kurtosis along direction $\mathbf{u}$.

The directional diffusion $\mathbf{D}(\mathbf{u})$ and kurtosis $\mathbf{K}(\mathbf{u})$ can be related to the diffusion tensor and KT using the following equations:

$$D(\mathbf{u}) = \sum_{i=1}^{3}\sum_{j=1}^{3} u_i u_j D_{ij} \tag{2.33}$$

and

$$K(\mathbf{u}) = \frac{M_D^2}{D(\mathbf{u})^2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} u_i u_j u_k u_l W_{ijkl} \tag{2.34}$$

where D_{ij} are the elements of the second-order diffusion tensor (DT); W_{ijkl} are the elements of the fourth-order ($3 \times 3 \times 3 \times 3$), fully symmetric KT; M_D is the mean diffusivity and is independent of the direction ($M_D = \frac{1}{3}\sum_1^3 D_{ii}$); and i, j, k, and l are the indices of the directions in the physical space.

As the DT, KT has antipodal symmetry, and thus only 15 W_{ijkl} elements are independent and fully characterize the KT.

Hence, Eq. 2.34 is a linear equation system with 15 unknown parameters. Therefore, by applying the gradients in 15 or more non-collinear, non-coplanar directions, the solutions can be found and the complete diffusional kurtosis tensor can be estimated. Thus, this approach does not require the complete measurement of the displacement distribution function and, therefore, is more time efficient compared to the **q**-space imaging technique. Another advantage of the DKI model is the low b-value ($b \leq 2500$ s/mm^2) required for imaging.

Measurements of non-Gaussian diffusion from the diffusion kurtosis model are of

interest because they can be used to characterize tissue microstructural heterogeneity[68] and to derive concrete biophysical parameters, such as the density of axonal fibees and diffusion tortuosity[56]. In[56], Fieremans *et al.* further elaborated on the DKI model to separate, like the CHARMED model, the intra-axonal space and extra-axonal space, and to estimate the fraction of MRI visible water in the axons relative to the total visible water signal, i.e. the Axonal Water Fraction (AWF).

According to[65], DKI can be used to resolve crossing fibers in tractography, and to obtain invariant rotational measures not limited to well-aligned fiber populations. Authors report an inter-fiber angle resolution power of 35° for the ODF computed under the DKI framework. Real dataset tractography results[65 60] show that DKI compares favorably against DTI and DSI.

The NODDI method is an approach that models[130] dendrite and axon populations (collectively termed neurites) in the human brain WM. The method is not targeted at recovering the fiber orientation distribution in the WM; it rather seeks to provide more reliable descriptors of the water diffusion in the brain tissue. According to the authors, indices accounting for this neurite-structure relate more directly to and provide more specific markers of brain tissue microstructure than standard indices from DTI, such as Fractional Anisotropy.

The proposed technique enables such mapping by combining a three-compartment tissue model (intra- and extra-axonal spaces, and the CSF) with a two-shell HARDI protocol optimized for clinical feasibility.

Under this model, the normalized signal S reads

$$S = (1 - f_{iso})(f_{ic}S_{ic} + (1 - f_{ic})S_{ec}) + f_{iso}S_{iso} \tag{2.35}$$

where S_{ic} and f_{ic} are the normalized signal and volume fraction of the intracellular compartment; E_{ec} is the normalized signal of the extracellular compartment; and S_{iso} and f_{iso} are the normalized signal and volume fraction of the CSF compartment.

The intracellular compartment refers to the space bounded by the membrane of neurites. This space is modeled as a set of "sticks", i.e., cylinders of zero radius, to capture the highly restricted nature of diffusion perpendicular to neurites and unhindered diffusion along them. The normalized signal, S_{ic}, adopts a simplified version of orientation-dispersed cylinder model in [129], such that

$$S_{ic} = \int_{S^2} f(\mathbf{n}) e^{-bd_{\|}(\mathbf{q}\cdot\mathbf{n})^2} d\mathbf{n} \tag{2.36}$$

where $\mathbf{q}$ and b are the gradient direction and b-value of diffusion weighting, respectively; $f(\mathbf{n})d\mathbf{n}$ gives the probability of finding sticks along orientation $\mathbf{n}$; $e^{-bd_{\|}(\mathbf{q}\cdot\mathbf{n})^2}$ gives the signal attenuation due to unhindered diffusion along a stick with intrinsic diffusivity $d_{\|}$ and orientation $\mathbf{n}$.

The orientation distribution function $f : S^2 \to \mathbb{R}$ is modeled with a Watson distribution:

$$f(\mathbf{n}) = M\left(\frac{1}{2}, \frac{3}{2}, \kappa\right)^{-1} e^{\kappa(\mu\cdot\mathbf{n})^2} \tag{2.37}$$

where M is a confluent hypergeometric function; μ is the mean orientation; and κ is the concentration parameter that measures the extent of orientation dispersion about μ.

According to the authors, the Watson distribution is the simplest orientation distribution that can capture the dispersion in orientations.

The extracellular compartment refers to the space around the neurites, which is occupied by various types of glial cells and, additionally in gray matter, cell bodies (somas). In this space, the diffusion of water molecules is hindered (but not restricted)

by the presence of neurites; hence, it is modeled with simple (Gaussian) anisotropic diffusion.

The normalized signal, S_{ec}, again adopts the extracellular signal model of orientation-dispersed cylinders in [129], such that

$$\log S_{ec} = -b\mathbf{q}^T \left[\int_{S^2} f(\mathbf{n}) D(\mathbf{n}) d\mathbf{n} \right] \mathbf{q} \tag{2.38}$$

where $D(\mathbf{n})$ is a cylindrically symmetric tensor with the principal direction of diffusion n.

The CSF compartment models the space occupied by cerebrospinal fluid and is modeled as isotropic Gaussian diffusion with a given diffusivity.

Within this framework, a new fiber orientation index, named "orientation dispersion", is defined as

$$OD = \frac{2}{\pi} \arctan\left(\frac{1}{\kappa}\right) \tag{2.39}$$

which, according to the authors, provides a more complete physical description of the diffusion process in white matter.

Authors use a set of imaging protocols consisting of various sampled orientations and (relatively low) b-values ($b < 3000$ s/mm^2) to test their method. The results reported focus on the role of the orientation dispersion and the neurite density in the FA index. Although, according to the authors, the orientation dispersion may potentially quantify the bending and fanning of axons, and therefore, may classify voxels as containing crossing fiber populations, no evidence is provided.

Cubism had been an analysis of the object and an attempt to put it before us in its totality; both as analysis and as synthesis, it was a criticism of appearance. Surrealism transmuted the object, and suddenly a canvas became an apparition: a new figuration, a real transfiguration.

Corriente alterna. 1967.

Octavio Paz, Mexican poet and diplomat (1914-1998). 1990 Nobel Prize in Literature

3

Spherical Deconvolution-Sparse Bayesian Learning Diffusion Reconstruction

In 2 we described different methods to elucidate the orientation of white-matter fiber bundles in diffusion MRI images. The Spherical Deconvolution framework postulated by Tournier *et al.*[114] provided a robust foundation for the direct estimation of the fiber orientation distribution. Their approach relied on spatial regularization penalization. Successive works have incorporated different priors or penalization terms to the SD approach in search of the global minima of the cost function: *et al.* introduced the

Least Absolute Shrinkage and Selection Operator (LASSO) regularization, and Daducci *et al.* proposed the ℓ_0-norm.

Recently, the Sparse Bayesian Learning (SBL) framework was introduced by Zhang *et al.*[131], where minimization of penalties on the sparsity of the estimate was shown to provide notable results in signal recovery problems. The framework is able to handle underdetermined problems, and reduce the local minima of the cost function employed even in the presence of highly correlated samples. Solving these issues is essential for the fiber orientation reconstruction from diffusion MRI images employing SD methods.

3.1 Introduction

The diffusion signal $S(\mathbf{q})$ is a real-valued function, which determines the value of S at the location $\mathbf{q}$ in $\mathbf{q}$-space. In the HARDI acquisition protocol, the signal $S(\mathbf{q})$ is sampled at N discrete orientations on the sphere $\mathbf{u}_k{}_{k=1}^{N}$ for several different q-values. Thus, for each q-value, measurements are made along N directions uniformly distributed over a unit sphere.

The reconstruction schemes presented in attempt to estimate $S(\mathbf{q})$ (or its features thereof) through several strategies. Nonetheless, either model-based or model-free mathematical diffusion imaging methods postulated to date have achieved a relatively high lower bound in the Inter-fiber angle (IFA) their are capable to resolve. According to Behrens *et al.*[16], an estimated one third of white matter voxels in a diffusion acquisition contain more than one fiber population.

The ability of each method to resolve the micro-geometry of WM fibers critically depends on the data required. Many of the fiber orientation reconstruction algorithms in compute the diffusion ODF, which can be interpreted as a smoothed version of the

fODF[49]. Generally, a larger number of acquisitions directions, together with larger b-values, provide better angular resolution. However, that limits their practical application due to longer acquisition times. Thus, shorter acquisition schemes are privileged, which impacts the quality of the fiber orientation reconstruction.

Methods based on variations of the tensor model, such as DKI, DOT or Generalized Diffusion Tensor Imaging (GDTI), require measuring at different b-values. Thus, for a clinically feasible acquisition scheme, the number of gradient directions acquired at each b-value must be reduced. The resolution power of the methods may be influenced by this requirement. Furthermore, voxels containing complex geometries, such as fibers at low inter-fiber angles or a large number of fibers, may not fit in a single tensor order, or may required higher orders.

Spherical deconvolution methods tend to obtain stable reconstructions across wide fiber crossing angle ranges, and require less gradient acquisition directions and smaller b-values[39]. Furthermore, a higher angular contrast is ensured by the computation of the fODF compared to the diffusion ODF.

The SD methods that encourage sparsity are of particular interest. These methods tend to reconstruct the diffusion signal (or its features) while intrinsically promoting the solution that minimzes the number of estimated fibers (ideally being the number of ground-truth fibers) in the solution space. The underlying theory to sparse signal recovery is Compressed Sensing (CS).

3.2 Compressed Sensing

The theory of compressed sensing (also known as compressive sensing, compressive sampling, or sparse sampling) was introduced by Candès, Romberg, and Tao[25] and

Donoho[51] almost simultaneously in 2006. They provided evidence about the possibility of reconstructing a band-limited signal with fewer samples than the sampling theorem (also known as the Nyquist-Shannon sampling theorem) requires, given sufficient knowledge about the signal's *sparsity*. In numerical analysis, the term *sparsity* refers to the density quality or *compactness* of a construct or matrix (representing some sort of information source) in which most element (called *atom*) values are zero (or close to zero). Such a structure is then called to be *sparse*. A signal is called S-sparse if it has only S nonzero elements. However, natural signals are rather *compressible*. A signal is designated compressible if it has only a small proportion of large coefficients when the signal is transformed into a suitable domain. Mathematically, a signal is compressible if the decay rule of the coefficients in the transformed domain obey the power law[26].

Emphasis should be put on the fact that compressed sensing does not violate the sampling theorem, because the latter guarantees perfect reconstruction given sufficient, not necessary, conditions. Compressed sensing depends on the sparsity of the signal in question and not its highest frequency. Sparse signals with high frequency components can be highly under-sampled using compressed sensing compared to classical fixed-rate sampling[26 24].

Compressed sensing typically starts with taking a weighted linear combination of samples (also called "compressive measurements") in a basis different from the basis in which the signal is known to be sparse. Candès *et al.* demonstrated that the number of these compressive measurements can be small and still contain nearly all the useful information. Therefore, the task of converting the signal back into the intended domain involves solving an underdetermined matrix equation since the number of compressive measurements taken is smaller than the number of samples in the original domain signal. However, if there is a unique sparse solution to an underdetermined system, then the

compressed sensing framework allows the recovery of that solution. In order to choose a solution to such a system, extra constraints or conditions (such as smoothness) must be imposed as appropriate.

The least-squares solution to such problems is to minimize the ℓ_2-norm, that is, minimize the amount of energy in the system. Although mathematically simple, this leads to poor results for many practical applications, for which the unknown coefficients have nonzero energy. To enforce the sparsity constraint when solving for the underdetermined system of linear equations, one can minimize the number of nonzero components of the solution. The function counting the number of non-zero components of a vector is called the ℓ_0-norm. Candès *et al.*, proved that for many problems ℓ_1-norm is equivalent to the ℓ_0-norm. This equivalence result allows to solve the ℓ_1-norm problem, which is easier than the ℓ_0-norm problem.

Sparse signal representations from overcomplete (redundant) dictionaries have found increasing relevance in a large number of application domains[125]. Attaining such representations is central to solving regularized linear inverse problems that have far-reaching applications in signal processing, compression, and feature extraction.

Compressed Sensing in Diffusion Imaging

Compressed sensing has been used to shorten magnetic resonance imaging acquisition times measuring fewer Fourier coefficients[80,81]. Given that the intra-fiber angle resolution performance of many diffusion MR techniques comes at the expense of long acquisition times, the CS theory found an immediate application in to reconstruct the sparse diffusion directions in each voxel from a sub-Nyquist sampled data[91,87,88,89,115,103,127]. In the context of diffusion MRI, CS requires three components: an under-sampling strategy that can incoherently sample the q space, a sparse model that can represent the

multi-modal diffusion and a reconstruction scheme that can recover the diffusion peaks without aliasing artifacts.

Michailovich *et al.* proposed[91] to use CS to reduce the number of gradients required in the HARDI acquisition protocol without compromising the accuracy of the reconstruction. The use of spherical harmonics as a basis does not allow for a sparse representation of the HARDI signal, because the energies of SH are spread all over S^2, and as a result, a relatively large number of coefficients are needed to effectively encode the diffusion tensors. According to their rationale, using an appropriate basis would allow to concentrate the energies of the elementary signals alongside the great circles of S^2, which would reduce the number of coefficients required to represent the HARDI signal with a given precision comparable to SH.

Since the diffusion measurements are linear, the set of discrete samples S_k of the $S(\mathbf{u})$ diffusion signal can be expressed in the form of inner products $S_k = \langle S(\mathbf{u}), \varphi_{j_k}(\mathbf{u}) \rangle$, being $\{\varphi_{j_k}\}_{k=1}^{N}$ a subset of the so-called "Dirac sampling basis" $\{\varphi_i\}_{i \in \mathcal{I}}$.

Let $\{\psi_j\}_{j \in \mathcal{J}}$ be another basis in the signal space, used for the representation of $S(\mathbf{u})$[91]. Thus, S can be expressed according to:

$$S(\mathbf{u}) = \sum_{j \in \mathcal{J}} c_j \psi_j(\mathbf{u}), \quad \forall \mathbf{u} \in S^2 \tag{3.1}$$

The set $\{\psi_j\}_{j \in \mathcal{J}}$ is assumed to be finite, being the total number of elements M. Making the following assumptions

- S is assumed to be "sparsely representable" by $\{\psi_j\}_{j \in \mathcal{J}}$, which implies that the number K of non-zero coefficients c_j in Eq. 3.24 is significantly less than M.

- The bases $\{\varphi_i\}_{i \in \mathcal{I}}$ and $\{\psi_j\}_{j \in \mathcal{J}}$ are "incoherent", implying that the value of

$\mu = \sup_{i,j}|\langle\varphi_i(\mathbf{u}), \psi_j(\mathbf{u})\rangle|$ is relatively small.

Provided that such a basis $\psi_{j\,j} \in \mathcal{J}$ is available for which the above assumptions are valid, the theory of CS provides evidence that an accurate estimate of S is possible from $\mathcal{O}\big(\mu^2 \log(M)K\big)$, as opposed to $\mathcal{O}\big(M\big)$ in the case of the standard sampling theory.

The basis $\psi_{j\,j} \in \mathcal{J}$, called a basis of SR, is derived from a set of wavelet functions through the application of the Funk-Radon transform.

The same rationale was the foundational basis used by Rathi *et al.* to propose[103] the generalization of the use of (single-shell) spherical ridgelets to multi-shell acquisitions, thus introducing the Multi-Shell Imaging (MSI) acquisition protocol.

Tristán-Vega *et al.* postulated[115] the Spherical Wavelets (SW)[58] as a suitable basis to sparsely represent the fODF. The estimation problem becomes an inverse problem by relating the SW functions to their counterparts in **q**-space. To solve the optimization problem, $L1$ regularization was proposed as a first approach. The $L1$ regularization-based solution compared positively to the Spherical Ridgelet approach, providing a lower reconstruction error rate. However, the $L1$ reconstruction problem was proved to be intractable in reasonable time ranges. Thus, although sparsity is only enforced for $\ell_p \forall p \leq 1$, Tristán-Vega *et al.* also proposed reconstructing the ODF using an $L2$ regularization strategy. Reconstruction times were two orders of magnitude faster than with $L1$, providing comparable reconstruction perfomance. When considering the number of samples required for the reconstruction, Tristán-Vega *et al.* reported a $4:1$ reduction ratio compared to a HARDI acquistion scheme. However, synthetic experiments were only reported for a relatively high SNR (40 dB), and a lower inter-fiber angle bound of $45°$.

Landman *et al.* formulated[74] the compartment (fiber volume) fraction reconstruction problem as a LASSO problem. Since the computational complexity is a limitation of

compressed sensing techniques, Landman *et al.* proposed to use an acceleration technique based on a multi-resolution basis set. Furthermore, the non-negativity constraint on the relative volume fractions was intrinsically enforced by the optimization method used. The results reported show that fibers crossing at angles as low as 30° are reconstructed the with a remarkably low number of gradient acquisition directions (30) and low b-value ($b = 700$ s/mm^2).

Mani *et al.* postulated[83] that, for simultaneous improvement of resolution in spatial and angular domains in the CS sense, undersampling of the combined **k**-space-**q**-space was required. The diffusion signal in each voxel is modelled as a sparse linear combination of Gaussian basis functions, i.e. a GMM. The $L1$ penalty was used on the coefficients of the basis functions to enhance sparsity, while the total variation penalty was employed for the coefficients to exploit the spatial smoothness. The reconstruction problem was solved by an iterative re-weighted conjugate gradient (CG) algorithm. Reconstruction results are qualitatively comparable to the regular HARDI scheme, no fiber reconstruction performance data is provided though, and emphasis is put on the acceleration factor introduced by the proposed strategy.

Menzel[87] and Paquette[99] proposed different CS-based strategies to accelerate DSI sampling schemes and reconstruct the EAP and ODF features. Menzel *et al.* solve the CS-DSI problem by considering the gradient operation as sparse transform, i.e. using a total variation regularization. Paquette *et al.* represent the signal on the basis provided by the Discrete Wavelet Transform (DWT), and use the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) algorithm[15] to solve the optimization problem. Although experiments with real diffusion MRI data are also presented in[87], numerical results provided for synthetic dataset only refer to relatively high inter-fiber angle values (70°).

Another family of CS methods used in the context of diffusion signal feature reconstruction employ a "dictionary" or "kernel". A diffusion reconstruction dictionary consists of a set of atoms representing a set of (fiber) directions whose amplitude (representing the relative volume fraction) need to be estimated to identify the underlying fiber layout. Dictionary-based sparse-like algorithms are generally formulated as convex optimization problems and exploit some sorts of priors.

Merlet *et al.*[88,89] used a CS-based approach to reconstruct diffusion features from DTI signals. They employed a sparse dictionary that allowed them to estimate the EAP and ODF with fewer measurements than those required by a HARDI acquisition. In[88], the dictionary is learnt from a set of training diffusion data, and thus a pre-processing step prior to the estimation stage needs to be completed to build the dictionary. This increase in the computational burden prompted Merlet *et al.* to use[89] predefined sets of functions that form orthonormal bases commonly used in the dMRI field as the dictionary.

In either case, Merlet *et al.* opted to use the Levenberg-Marquardt algorithm (LMA) to solve the non-linear optimization problem posed by the dictionary update step. The need of non-linear optimization methods, which are slow, still impose a notable computational burden to the approach.

The results reported in[89] on the use of CS modelling the diffusion signal with a number of basis (such as SPF[11] or SHORE[96]) read that, in the most favourable experimental conditions, the maximum angular resolution solved with an acceptable error is around 30°.

Regularization

Suppose that for a known matrix $\mathbf{\Phi}$ and an observed vector of measurements $\mathbf{E}$, we wish to find a vector $\mathbf{c}$ such that

$$\mathbf{\Phi c} = \mathbf{E} \tag{3.2}$$

The standard approach is ordinary least squares linear regression. However, if no $\mathbf{E}$ satisfies the equation or more than one $\mathbf{E}$ does (i.e. the solution is not unique) the problem is said to be "ill-posed". In such cases, ordinary least squares estimation leads to an overdetermined (over-fitted), or more often an underdetermined (under-fitted) system of equations. Most real-world phenomena have the effect of low-pass filters in the forward direction. Therefore, in solving the inverse-problem, the inverse mapping operates as a high-pass filter that has the undesirable tendency of amplifying noise (eigenvalues / singular values are largest in the reverse mapping).

Ordinary least squares seeks to minimize the sum of squared residuals, which can be compactly written as

$$\|\mathbf{\Phi c} - \mathbf{E}\|^2 \tag{3.3}$$

where $\|\cdot\|$ is the Euclidean norm.

In order to give preference to a particular solution with desirable properties, a regularization term can be included in this minimization. Depending on the property of the observed data that is that is priviledged in the solution space, different regularization penalties exist. From a Bayesian point of view, many regularization techniques correspond to imposing certain prior distributions on model parameters.

For the purpose of signal and image reconstruction, ℓ_1-norm minimization models are used. In the CS reconstruction models using constrained ℓ_1-norm minimization, larger

coefficients are penalized heavily. A popular choice is to use an iterative algorithm for constructing the appropriate weights: each iteration requires solving one ℓ_1-norm minimization problem by finding the local minimum of a concave penalty function that more closely resembles the ℓ_0-norm.

In order to find $\mathbf{c}$, the convex optimization problem specified in Eq. 3.12 must be solved, i.e.

$$\min \mathbf{c} \in \mathcal{R}^{n_c} \|\mathbf{c}\| \quad \text{s.t.} \|\mathbf{E} - \mathbf{\Phi c}\| \leq \epsilon \tag{3.4}$$

where ϵ denotes the noise level.

L1 Regularization

The Least Absolute Shrinkage and Selection Operator (LASSO) or $L1$ regularization method was proposed by Tibshirani[108] and is a popular method for estimating generalized linear models. It is a "shrinkage estimator", i.e. generates coefficient estimates that are biased to be small, using to that end a penalty term which encourages the sum of the absolute values of the parameters to be small. The estimator can be stated as

$$\underset{\mathbf{c} \in \mathcal{R}^{n_c}}{\arg\min} \left\{ \|\mathbf{E} - \mathbf{\Phi c}\|^2 + \lambda \|\mathbf{c}_{\ell_1}\| \right\} \tag{3.5}$$

where λ is the regularization parameter; and the ℓ_1-norm is $\|\mathbf{c}_{\ell_1}\| = \sum_{i=1}^{N} |\mathbf{c_i}|$, being N the number of samples.

The regularization parameter λ replaces the noise level ϵ in Eq. 3.4, in governing the trade-off between the data consistency and its sparsity. As λ tends to zero, the estimate tends towards unregularized least-squares regression; as λ increases, the sparsity term dominates. Since the non-negative factor λ controls the trade-off between accuracy and

stability, it is important to choose λ in order to avoid the under-regularization or the over-regularization.

A number of algorithms exist to solve problems regularized with the LASSO method.

L2 Regularization

Least squares recovery with an $L2$ regularization is the preferred method used in order to solve systems of overdetermined equations. The $L2$ recovery problem can be stated as

$$\underset{\mathbf{c}\in\mathcal{R}^{n_c}}{\arg\min} \left\{ \|\mathbf{E} - \mathbf{\Phi c}\|^2 + \lambda \|\mathbf{c}_{\ell_2}\| \right\} \tag{3.6}$$

where λ is the regularization parameter; and the ℓ_1-norm is $\|\mathbf{c}_{\ell_2}\| = \sum_{i=1}^{N} \mathbf{c_i}^2$, being N the number of samples.

Tournier *et al.* used[112] the ℓ_2-norm regularization in their CSHD spherical deconvolution framework variant.

Tikhonov Regularization

The Tikhonov regularization scheme can can be formulated as finding the $\mathbf{c}$ that solves

$$\arg\min \left\{ \|\mathbf{\Phi c} - \mathbf{E}\|^2 + \lambda \|\mathbf{\Gamma c}\|^2 \right\} \tag{3.7}$$

where λ is a regularization parameter; and $\mathbf{\Gamma}$ is the regularization operator known as "Tikhonov regularization matrix".

In many cases, this matrix is chosen as a multiple of the identity matrix ($\mathbf{\Gamma} = \lambda \mathbf{I}$; $\|\mathbf{\Phi c} - \mathbf{E}\|^2 - \lambda \|\mathbf{c}\|^2$, giving preference to solutions with smaller norms; this is known as $L2$ regularization. In other cases, lowpass operators may be used to enforce smoothness.

Minimization of the objective function in Eq. 3.7 with $\mathbf{\Gamma} = \lambda\mathbf{I}$ yields the equation $\mathbf{\Phi}^T\mathbf{E} = (\mathbf{\Phi}^T\mathbf{\Phi} + \lambda\mathbf{I})\mathbf{c}$ whose solution is given by

$$\hat{\mathbf{c}} = (\mathbf{\Phi}^T\mathbf{\Phi} + \mathbf{\Gamma}^T\mathbf{\Gamma})^{-1}\mathbf{\Phi}^T\mathbf{E} \tag{3.8}$$

where $(\mathbf{\Phi}^T\mathbf{\Phi} + \lambda\mathbf{I})^{-1}\mathbf{\Phi}^T$ is called the damped least squares (DLS) inverse of $\mathbf{\Phi}$ and can be computed using the Singular Value Decomposition (SVD).

Total Variation Recovery

Total Variation (TV) can be seen as a non-negative real-valued functional defined on the space of real-valued functions (for the case of functions of one variable) or on the space of integrable functions (for the case of functions of several variables). In signal and image reconstruction, it is applied as "total variation regularization", where the underlying principle is that signals with excessive details have high total variation and that removing these details, while retaining important information such as edges, would reduce the total variation of the signal and make the signal subject closer to the original signal in the problem.

3.3 Sparse Bayesian Learning

SBL is a framework that was originally proposed[109] by Tipping to find robust solutions to sparse signal representation problems in the context of regression and classification. It was later applied by Wipf and Rao for sparse signal recovery[125]. A key feature of this development that is germane to the basis selection problem is the incorporation of a parameterized prior on the weights that encourages sparsity in representation, i.e., few nonzero weights. Furthermore, the global minimum solution is found to be the

sparsest one in SBL methods. Bayesian algorithms have received much attention since they generally achieve notable recovery performance [131].

The rationale for the SBL is to find the posterior probability $p(\mathbf{x}|\mathbf{y};\theta)$ via the Bayesian rule, where θ denotes the set of all the hyperparameters. Given the hyperparameters, the solution $\hat{\mathbf{x}}$ is given by the Maximum-A-Posteriori (MAP) estimate. The hyperparameters are estimated from the data by marginalizing over $\mathbf{x}$ and then performing evidence maximization or type-II Maximum Likelihood (ML)). The idea remains essentially unaltered for multiple observation vectors, but the ℓ_2 norm is applied on each source $\mathbf{X}_i$.

Following the notation in this work, SBL assumes a Gaussian likelihood model

$$p(\mathbf{E}|\mathbf{c};\sigma^2) = (2\phi\sigma^2)^{-N/2} \exp\left[-\frac{1}{2\sigma^2}\|\mathbf{E}-\mathbf{\Phi}\mathbf{c}\|^2\right] \tag{3.9}$$

where $\mathbf{\Phi}$ is the SBL dictionary matrix where $\text{rank}(\Phi) = N$; and σ^2 is the error variance.

The $\Phi \in \mathcal{R}^{N\times M}$ is a coherent matrix, one which has most of its mass in a relatively low number of elements, and whose columns represent a (possibly) overcomplete basis (i.e. $\text{rank}(\Phi) = N \wedge M > N$).

Obtaining the maximum likelihood estimates for c under these conditions is equivalent to finding the minimum ℓ_2-norm solution for Eq. 3.13. However, such solutions are known to produce non-sparse representations. Thus, some sort of weight prior needs to be incorporated to promote sparsity. Different priors can be invoked, each corresponding to a different hypothesis about the underlying truth. These hypothesis can be compared by evaluating the Bayesian evidence for the model prior.

SBL estimates a parametrized prior from the data as

$$p(\mathbf{c};\gamma) = \prod_{i=1}^{M} (2\pi\gamma_i)^{-\frac{1}{2}} \exp\left(-\frac{\omega_i^2}{2\gamma_i}\right) \tag{3.10}$$

where $\gamma = [\gamma_1, \ldots, \gamma_M]^T$ is a vector of M hyperparameters controlling the prior variance of each weight. These hyperparameters (along with the error variance σ^2) can be estimated from the data by marginalizing over the weights and performing a ML optimization, in a procedure referred to as evidence maximization or type-II maximum likelihood.

SBL is found to be globally convergent, i.e. each iteration os guaranteed to reduce the cost function until a fixed point is reached. The SBL cost function can potentially have multiple local minima, but all must be achieved at sparse solutions. When tasked with sparse linear inverse problems such as Eq. 3.13, cost functions whose minimization corresponds with maximally sparse solutions are priviledged. The SBL cost function is characterized by a global minimum that can produce the maximally sparse solution at the posterior mean[125].

SBL has been used for a variety applications in brain research, such as Electroencephalography (EEG) signal source localization[105], or the identification of the neuroanatomical basis of cognitive impairment in Alzheimer's Disease (AD)[121].

SBL methods have found a suitable application in the estimation of diffusion fiber features, such as the Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques (BEDPOSTX) implementation based on previous works by Behrens *et al.*[17][16], and involving sampling through Markov Chain Monte Carlo techniques to build up distributions on diffusion parameters at each voxel.

Bilgic *et al.* used[18] the k-means Singular Value Decomposition (k-SVD) algorithm[2]

coupled to the Focal Underdetermined System Solver (FOCUSS) algorithm[61] for the acceleration of DSI using CS and Bayesian learning techniques.

Although not enforcing signal sparsity, Melie-García *et al.* proposed[86] to use Bayesian techniques for model selection and parameter estimation, in combination with Reversible Jump Markov Chain Monte Carlo (RJMCMC) for probability and diffusion model parameter estimations, to reconstruct the fODF. Although authors grant the high computational burden of the method, its performance is notable in multi-fiber (containing as much as 5 fibers) synthetic datasets for a given band of SNR and b-value conditions. The authors used the same model used in the synthetic signal generation for the fiber reconstruction, which is an unlikely condition to be met in practice. Furthermore, the use of Monte Carlo simulations requiring a large number of iterations per voxel renders the method computationally inefficient.

3.4 Problem Statement

This work proposes to use of a SD-SBL algorithm to resolve the brain's fiber arrangement. In the context of the fiber orientation reconstruction, it is natural to impose some useful prior constraints based on the following considerations: the relative volume fractions of the fibers are expected to be nonnegative. Negative weights are not physically meaningful and should be penalized by adding a regularization term or excluded by imposing an explicit nonnegativity constraint. In addition, it is reasonable to assume that most white matter voxels only contain contributions from relatively few fiber bundles. Therefore, apart from a few significant peaks, we can assume that $\mathbf{c}$ has a sparse support, i.e. most of its entries are expected to be zero (or very small). The Sparse Bayesian Learning method implemented in the context of this work is based on

the version introduced by Zhang *et al.* in[131].

The signal recovery problem in the context of CS can be stated as follows. We want to recover a S-sparse signal $\mathbf{x} \in \mathcal{R}^n$ from an observation vector $\mathbf{y}_\Omega \in \mathcal{R}^m$ with $m < n$, such that

$$\mathbf{y}_\Omega = \mathbf{A}_\Omega \mathbf{x} \tag{3.11}$$

where $\mathbf{A}_\Omega^{m \times n}$ is called the CS matrix; and Ω defines a subset of indices corresponding to samples in which $\mathbf{A}_\Omega$ is evaluated and $\mathbf{y}_\Omega$ is acquired. This problem is ill-posed because we have fewer equations than unknowns ($m < n$). However, if x is *sufficiently* sparse, one can recover $\mathbf{x}$ in Eq. 3.11 by solving the following convex optimization problem[51 27]:

$$\arg\min_{\mathbf{x} \in \mathcal{R}^n} \{\|\mathbf{x}\|_\ell\} \quad \text{s.t. } \mathbf{y}_\Omega = \mathbf{A}_\Omega \mathbf{x} \tag{3.12}$$

where $\mathbf{y}_\Omega = \mathbf{A}_\Omega \mathbf{x}$ is the data consistency or the model-fitting constraint, $\|\mathbf{x}\|_\ell$ is the sparsity constraint. The data consistency constraint enables the solution to remain close to the raw data acquisition, whereas the minimization of the second term promotes sparsity.

When $\mathbf{x}$ is corrupted by noise $\mathbf{z} \in \mathcal{R}^m$, that is

$$\mathbf{y}_\Omega = \mathbf{A}_\Omega \mathbf{x} + \mathbf{z} \tag{3.13}$$

we relax the constraint in Eq. 3.12 and solve

$$\min \mathbf{x} \in \mathcal{R}^n \|\mathbf{x}\|_\ell \quad \text{s.t. } \|\mathbf{y} - \mathbf{A}_\Omega \mathbf{x}\|_\ell \leq \epsilon \tag{3.14}$$

where ϵ is the level of noise in the data.

In the single measurement vector (SMV) model, to ensure a unique global solution, the number of nonzero entries in the source vector has to be less than a threshold[61].

When a sequence of measurement vectors are available, a framework named the multiple measurement vector (MMV) model, the key assumption is that the support (i.e. indexes of nonzero entires) of every column of the source matrix is identical. This quality is referred to as "the common sparsity assumption" in the literature. In addition, similiar to the SMV model, the number of nonzero rows in the source matrix has to be below a threshold to ensure a unique global solution[38]. This leads to the sparsity constraint.

3.5 Numerical Experiment Design

Numerical experiments were performed to compare the performance of the proposed SD-SBL method under different experimental conditions against other methods. Synthetic datasets, a phantom model and a human dataset were used in the assessment. The minimization strategies compared to SBL were the nnLS, LASSO, and RUMBA-SD. According to Jian *et al.*[70], nnLS is arguably the technique of choice for the multi-fiber reconstruction problem in the presence of IVOH. Although nnLS is known to be less competitive than other methods, it was included for its simplicity. LASSO is a well-established method that provides notable angular contrast, although suffers from a number of drawbacks. Finally, RUMBA-SD is a spherical deconvolution approach known to outperform Richardson-Lucy-based methods that appeared recently[53]. It was elected on the basis of its ability to efficiently reconstruct the fODF for low inter-fiber angles. RUMBA-SD was used without spatial regularization in this work.

3.5.1 Synthetic Data Generation

In this work, we assume a generative multi-tensor signal model where the diffusion data in each voxel can be expressed as the sum of the signals from each intra-voxel compartment, as proposed in[118]. In this case, the term "compartment" is defined as a homogeneous region in which the diffusion process possesses identical properties in magnitude and orientation throughout, and which is different to the diffusion processes occurring in other compartments. From a modelling point of view, using a multi-compartment framework allows considering the partial volume effect in brain voxels with mixture of white matter, gray matter and cerebrospinal fluid. This strategy has been shown to be effective in reducing the occurrence of spurious fiber orientations[46].

Under this assumption, the canonical form of the diffusion signal in a voxel at the i-th gradient direction is expressed as

$$S_i = S_0 \left[\sum_{j=1}^{M} f_j e^{\left(-b_i \mathbf{v}_i^T \mathbf{D}_i \mathbf{v}_i\right)} + f_{GM} e^{(-b_i D_{GM})} + f_{CSF} e^{(-b_i D_{CSF})} \right] \tag{3.15}$$

where M is the total number of fiber bundles; f_j denotes the relative volume fraction of the j-th fiber compartment; f_{GM} and f_{CSF} are the relative volume fractions of the gray matter and cerebrospinal fluid compartments, respectively, so that $\sum_{j=1}^{M} f_j + f_{GM} + f_{CSF} = 1$; b_i is the diffusion-sensitizing gradient factor (i.e. b-value) used in the acquisition scheme to measure the diffusion signal S_i along the diffusion-sensitizing gradient unit vector $\mathbf{v}_i, i = 1, \ldots, N$; D_{GM} and D_{CSF} are respectively the mean diffusivity coefficients in GM and CSF; S_0 is the diffusion signal amplitude in the absence of diffusion sensitization gradients ($b_i = 0$); $\mathbf{D}_j = \mathbf{R}_j^T \mathbf{A} \mathbf{R}_j$ denotes the anisotropic diffusion tensor of the j-th fiber-bundle, where $\mathbf{R}_j$ is the rotation matrix that rotates a unit vector initially oriented along the X-axis towards the j-th fiber orientation (θ_j, ϕ_j), and $\mathbf{A}$ is

a diagonal matrix containing information about the magnitude and anisotropy of the diffusion process inside that compartment.

The observed signal at a given voxel will be the vector sum of the signals measured along different spatial gradient directions gradients:

$$S(\theta, \phi) = \sum_{i=1}^{N} S_i(\theta, \phi) \tag{3.16}$$

where (θ, ϕ) are, respectively, the polar and azimuth angles defining the applied gradient direction; and N is the total number of gradient directions.

The spherical deconvolution dictionary for SBL and RUMBA was generated using a diffusivity tensor with values $D = [1.7\ 0.4\ 0.4] \times 10^{-3}$ mm^2/s for white matter fiber bundles, and 3.0×10^{-3} and of 0.7×10^{-3} mm^2/s for the cerebrospinal fluid and gray matter isotropic compartments, respectively.

Without loss of generality, $S_0 = 1$ was assumed. A total number of 724 unique angular gradient directions uniformly distributed over the unit sphere were used for the simulations.

The extraction of the orientational peaks was based on the following procedure. Sparse methods directly return the fODF, being the candidates with a value greater than a given threshold the compartment estimates. Such threshold was set as a volume fraction ratio to the greatest component (i.e. $0.1 \cdot f_{max}$). This procedure allowed to discard spurious components. The orientations of the fibers and their relative volume fractions of continuous methods are computed from the maxima of the ODF functions. That is, the number of maxima, their magnitude and directions are computed, and the elements below the above threshold are similarly discarded.

Noise model

To be consistent with previous studies presenting new reconstruction methods, in this work the signal has been corrupted by Rician noise (see Sec. B.1) as follows:

$$S_{noisy} = \sqrt{(S + \eta_1)^2 + \eta_2^2} \tag{3.17}$$

where $\eta_1, \eta_2 \sim N(0, \sigma^2)$ and $\sigma^2 = S_0/SNR$ controls the level of the noise.

The noise level is thus determined by the on the image.

3.5.2 Fiber ODF model

Without loss of generality, under the spherical deconvolution framework, the measured diffusion signal S_i at a voxel i for N different sampling parameters (i.e., v_i and b_i, $i \in [1, \ldots, N]$) can be recast in matrix form as

$$\mathbf{S} = \mathbf{H}\mathbf{f} \tag{3.18}$$

where $\mathbf{S} = [S_1, \ldots, S_i, \ldots, S_N]^T$, and $\mathbf{H} = [\mathbf{H}^{WM}|\mathbf{H}^{iso}]$ comprises two sub-matrices. $\mathbf{H}^{WM}$ is an $N \times M$ matrix where every column of length N contains the values of the signal generated by the model given in Eq. 3.23 for a single fiber-bundle compartment oriented along one of the M directions, i.e., the (i, j)-th element of $\mathbf{H}^{WM}$ is equal to $\mathbf{H}^{WM}_{ij} = S_0 \exp\left(-b_i \mathbf{v}_i^T \mathbf{D}_i \mathbf{v}_i\right)$. Likewise, $\mathbf{H}^{iso}$ is an $N \times 2$ matrix where each of the two columns of length N contains the values of the signal for each isotropic compartment, i.e., $\mathbf{H}^{iso}_{i1} = S_0 \exp\left(-b_i D_{GM}\right)$ and $\mathbf{H}^{iso}_{i2} = S_0 \exp\left(-b_i D_{CSF}\right)$. Finally, the column-vector $\mathbf{f}$ of length $M + 2$ includes the volume fractions of each compartment within the voxel.

In the framework of model-based spherical deconvolution, $\mathbf{H}$ is created by specifying

the diffusivities, which are chosen according to prior information, and by providing a dense discrete set of equidistant M-orientations $\Omega = (\theta_j, \phi_j).i \in [1, \ldots, M]$ uniformly distributed on the unit sphere. Previous studies have used different sets and numbers of orientations . The goal is then to infer the volume fraction of all predefined oriented fibers, $\mathbf{f}$, from the vector of measurements $\mathbf{S}$ and the dictionary or kernel $\mathbf{H}$ of oriented basis signals. The underlying fiber bundle orientations will then correspond to the (θ_i, ϕ_i) orientation sets corresponding to the estimated relative volume fractions. Under this reconstruction model, $\mathbf{f}$ can be interpreted to as the fiber ODF evaluated on the set Ω. The matrix $\mathbf{H}$ is also known as the "diffusion basis functions" , or the "point spread function" that blurs the fiber ODF to produce the observed measurements.

The deconvolution problem given by Eq. 3.26 does not have a straightforward solution, since the resulting system of linear equations is ill-conditioned and ill-posed (i.e., there are more unknowns than measurements and some of the columns of $\mathbf{H}$ are highly correlated), which can lead to numerical instabilities. In order to solve such equation system, in this work we propose to use the SBL strategy.

This framework uses a sparse Bayesian learning algorithm to explore the solution space, and find the fiber configuration that minimizes the cost function.

3.5.3 Phantom Model

The data used for complex fiber configuration reconstruction performance assessment in this work corresponded to the data used at the 2nd HARDI Reconstruction Challenge[41], organized in the framework of the International Symposium on Biomedical Imaging (ISBI) 2013. The phantom used to generate these data simulates a 27 fiber-bundle complex architecture, with fibers aligned at distinct orientations, and comprising a wealth of configurations, including branching, crossing and kissing fibers. White-matter

fibers were simulated using a diffusion tensor model with cylindrical symmetry ($\lambda_1 = 1.7, \lambda_2 = \lambda_3 = 0.2 \times 10^{-3}$ mm^2/s). Isotropic compartments simulating the presence of and were also included in the data ($D_{GM} = 0.2; D_{CSF} = 1.7 \times 10^{-3}$ mm^2/s). The generated diffusion images had a total size of $50 \times 50 \times 50$ voxels.

Diffusion MRI images for methods using the HARDI protocol were generated using $N = 64$ sampling points on a sphere in **q**-space with constant $b = 3000$ s/mm^2, plus one additional image with $b = 0$. Images for methods using a DTI protocol were generated using $N = 32$ sampling points on a sphere in **q**-space at $b = 1200$ s/mm^2, plus one additional $b = 0$ image.

The spherical deconvolution dictionary for SBL and RUMBA was generated using a diffusivity tensor with values $D = [1.7\ 0.4\ 0.4] \times 10^{-3}$ mm^2/s for white matter fiber bundles, and 3.0×10^{-3} and of 0.7×10^{-3} mm^2/s for the cerebrospinal and gray matter isotropic compartments, respectively. Different values from ground truth values were used to avoid over-fitting.

The signal when no diffusion-sensitinzing gradient is present ($b = 0$) was assumed to be $S_0 = 1$.

3.5.4 Human Brain Dataset

The human brain diffusion MRI dataset studied corresponds to a whole-brain HARDI acquisition from a healthy subject. Data were acquired in a 3 T Philips Achieva scanner (Best, Netherlands) located at Hospital de la Santa Creu i Sant Pau (Barcelona, Spain) with an 8-channel head coil along 100 different gradient directions on the sphere in **q**-space with constant $b = 2000$ s/mm^2. Additionally, one unweighted volume ($b = 0$) was acquired with in-plane resolution of 2×2 mm and slice thickness of 2 mm. The acquisition was carried out without undersampling in the **k**-space (i.e., $R = 1$). The

final dimension of this dataset is $128 \times 128 \times 60 \times 101$ voxels.

The spherical deconvolution dictionary for SBL and RUMBA was generated using a diffusivity tensor with values $D = [1.4\ 0.4\ 0.4] \times 10^{-3}$ mm^2/s for white matter fiber bundles, and 3.0×10^{-3} and of 0.7×10^{-3} mm^2/s for the cerebrospinal fluid and gray matter isotropic compartments, respectively.

The signal when no diffusion-sensitizing gradient is present ($b = 0$) was assumed to be $S_0 = 1$.

3.5.5 Data Analysis

We have tested the angular contrast of the fiber orientation reconstruction methods method in a set of fiber-crossing configurations for the synthetic dataset. Two synthetic fibers were simulated at different compartment separation angles up to 90° in discrete steps of 5°. The ability to recover a single fiber (i.e. the 0° case) was also rated for the sake of completeness.

The sensitivity to miscalibration effects was assessed on the synthetic datasets, or according to available data on the phantom model and real datasets. The impact of different values of the diffusion-sensitizing gradient factor (i.e. b-value) was studied in the analytic diffusion signal model for the synthetic dataset: (i) $b = 3000$ s/mm^2 was used as a conservative value; (ii) $b = 1500$ s/mm^2 was elected to illustrate a more challeging value; and (iii) b was set to 800 s/mm^2 as the lower end case.

Two acquisition schemes were tested for the phantom models: (i) $b = 3000$ s/mm^2, measured along a total set of 64 uniformly distributed gradient directions, and (ii) $b = 1200$ s/mm^2, measured along a total set of 32 uniformly distributed gradient directions.

The response to differential fiber signal contribution (i.e. changing relative volume fractions) was tested on the synthetic dataset. The even relative volume case ($f_1 =$

$f_2 = 0.5$) was used as the baseline. The simulated fibers' relative volume fraction was then varied to test different contributions ($f_1 \neq f_2$) satisfying the $\sum_i f_i = 1$ condition.

As a general rule, the white matter fiber bundles were simulated with the same diffusivity properties corresponding to a tensor with values $D = D_1 = D_2 = [1.7\ 0.3\ 0.3] \times 10^{-3}$ mm^2/s. The influence of a larger mismatch with respect to the values used in the spherical deconvolution dictionaries was investigated varying the anisotropy of the signal using two additional sets of diffusivity tensors: $D = D_1 = D_2 = [1.7\ 0.2\ 0.2] \times 10^{-3}$, and $D = D_1 = D_2 = [1.5\ 0.3\ 0.3] \times 10^{-3}$ mm^2/s.

The performances of the reconstruction methods were also evaluated as a function of the noise magnitude for different levels, i.e., $SNR = 20, 30$ dB.

For each different configuration, a maximum of 2000 iterations were allowed for SD-SBL. For the rest of the methods tested, the number of iterations was decimated by a factor of 10, except for RUMBA, for which it was set to 600. For the phantom and human datasets, RUMBA was set to a maximum of 300 iterations, and a maximum value of 2000 was kept for SD-SBL.

In the case of the synthetic datasets, due to the random nature of the Rician likelihood noise model used, in order to quantify orientational uncertainty of the results, 100 estimates of the peak orientations were obtained, and their average and standard deviations were computed.

An angular cone of 12.5° was set to determine a fiber population orientation.

The performance of the algorithms on the synthetic datasets was assessed according to three criteria:

- **Detected fiber error**, that is, the number of peak orientations recovered by the methods.

- **Detected fiber penalty**, computed as the normalized penalty between the number of fibers estimated and the true fiber populations, thus providing an estimate of the failure rate.

- **Angular error**, computed as the angular deviation between the estimated and real fiber orientations.

- **Volume fraction error**, computed as the mean error in the estimated volume fractions with respect to the real volume fractions.

On the premise that the influence of all variables on the quality of fiber orientation reconstruction could be summarized by an absolute measure, a global penalty function was computed. This score was used to report the net algorithm performance, and included the following aspects:

- **Angular error**, computed as the mean angular deviation between the estimated and real fiber orientations.

- **Volume fraction error**, computed as the mean error in the estimated volume fractions with respect to the real volume fractions.

- **Success rate (SR)**, defined as the mean fraction of correctly estimated fibers (see[39] for a definition on multiple voxel datasets).

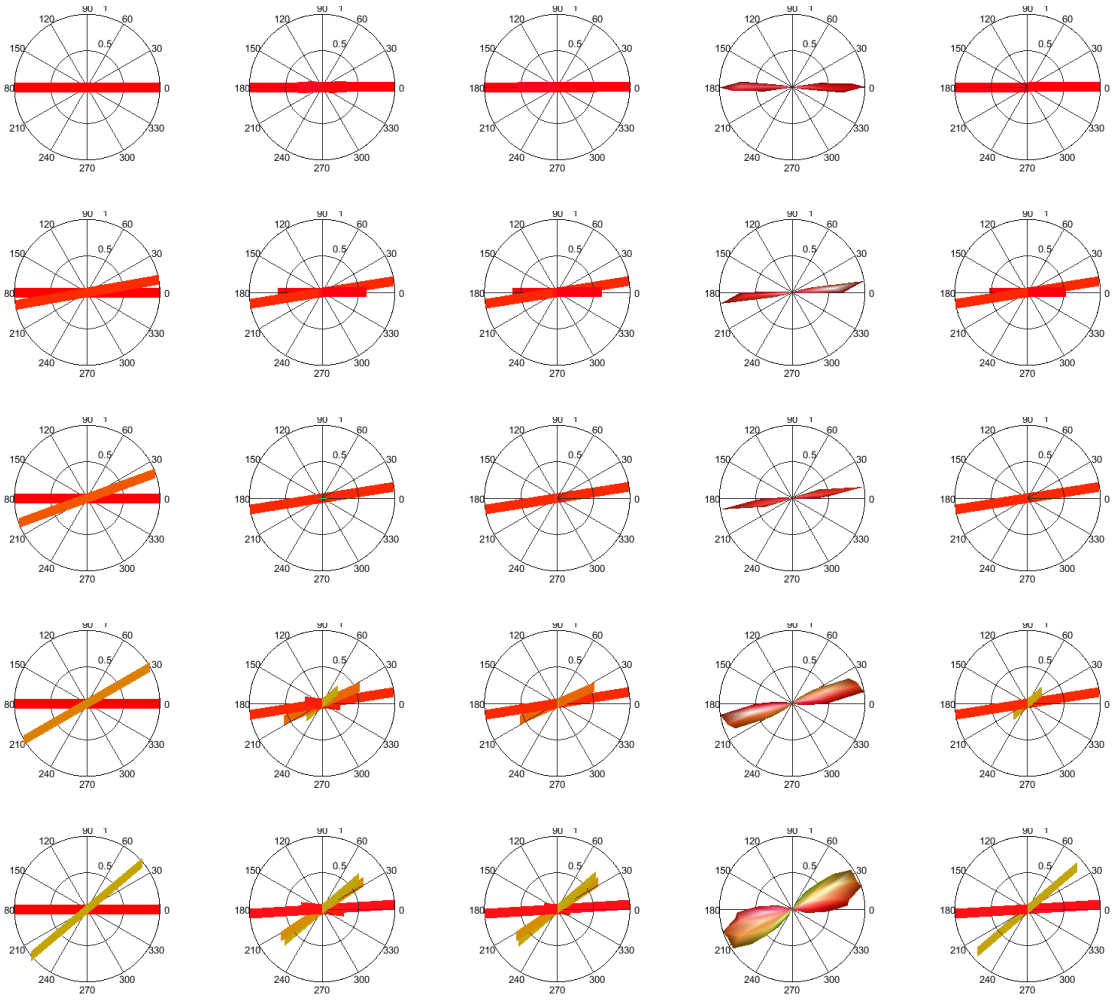
In order to provide a penalty sense to the function, the complementary to the SR (which can be termed as the "failure rate") was used.

3.6 Results

3.6.1 Synthetic Data

Fiber orientation reconstruction results for synthetic fibers are displayed in Fig. 3.1 and Fig. 3.2 for the diffusion value $b = 1500$ s/mm^2, for equal fiber volume fractions $f_1 = f_2 = 0.5$, at SNR 30 dB. Simulations show that methods are able to successfully solve fibre populations separated by inter-fiber angles as small as 40°. Some methods' performance decreases when the inter-fiber angle is below 40° (see, for example, RUMBA –column (d) in Fig. 3.1). Although other methods, such as nnLS, seem to succeed for angles as small as 30°, the number of overestimated fibers (i.e. false positives) is also increased. Fig. 3.1 shows that the angular resolution of SD-SBL is approximately 30°, but false positives are not present (compare nnLS and SD-SBL for 30° and 40° –columns (b) and (d) in Fig. 3.1). Furthermore, the fiber orientation diffusion function provides, approximately, the same signal for both fibers for angles as small as 40° in the SD-SBL case. For larger angles, nnLS and LASSO still add fibers that were not in the ground-truth dataset.

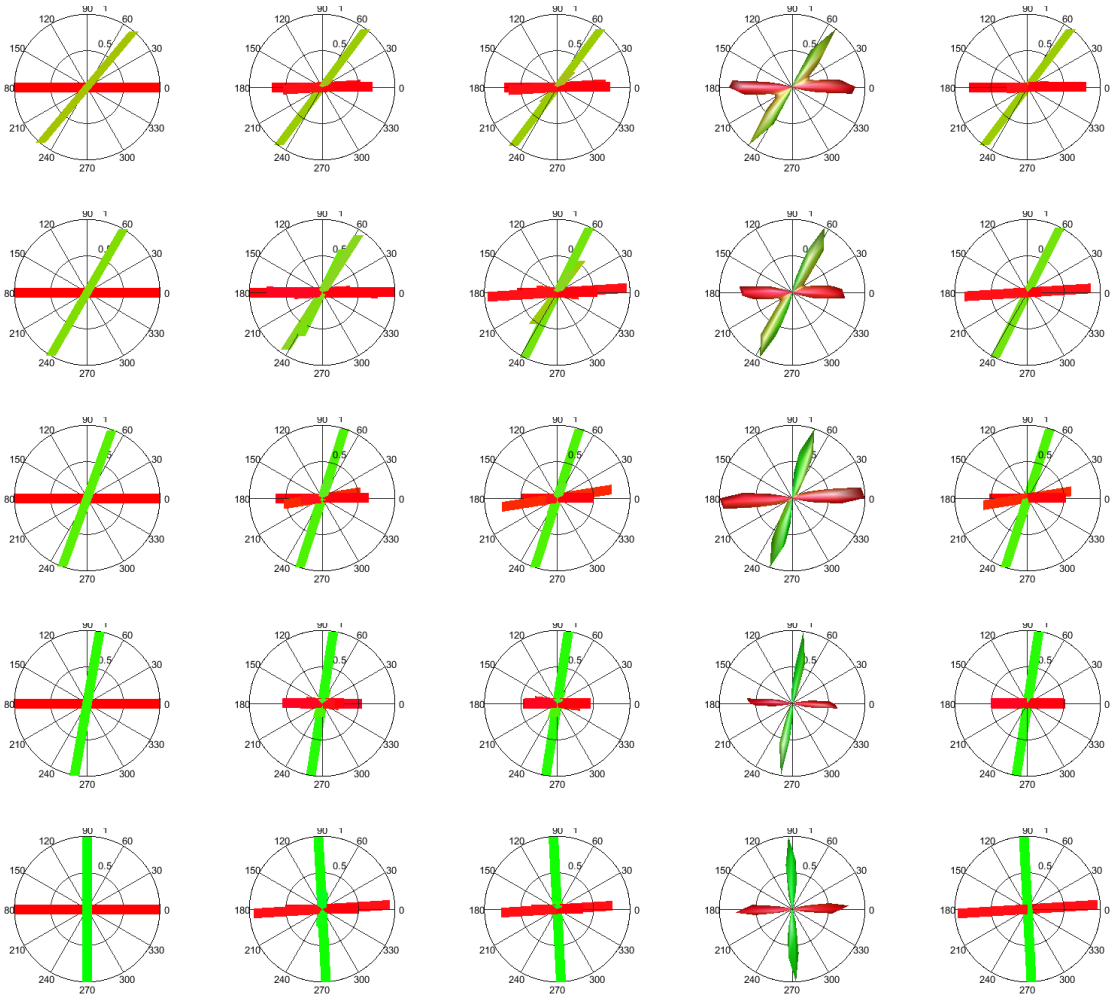
The effects of multiple realizations on the recovered fiber package orientations are depicted in Fig. 3.3, Fig. 3.4, Fig. 3.5, and Fig. 3.6. Thus, an intuitive insight into each method's compartment resolution power, success rate, and angular error is provided. Signals were generated for a diffusion value $b = 1500$ s/mm^2, for equal fiber volume fractions $f_1 = f_2 = 0.5$, at SNR 30 dB. In agreement with Fig. 3.1 and Fig. 3.2, nnLS' and LASSO's fiber population separation ability is acceptable for crossing angles above 40°. However, high overestimation rates are observed in the $[30 - 90°]$ inter-fiber angle range. For IFAs lower than 40°, nnLS and LASSO are not able to successfully separate the two fiber populations. RUMBA exhibits a higher underestimation rate



(a) (b) (c) (d) (e)

Figure 3.1 Fiber orientations for angles $[0-40°]$ recovered using different reconstruction methods. Fiber orientations reconstructed with different methods for inter-fiber crossing angles in the interval $[0-40°]$. Each row represents a different inter-fiber crossing angle, starting from 0° to 40°, in discrete steps of 10°. The ground truth is displayed in (a). Reconstructions methods are: (b) Non-negative Least-Squares (nnLS); (c) Least Absolute Shrinkage and Selection Operator (LASSO); (d) Robust and Unbiased Model-BAsed Spherical Deconvolution (RUMBA); and (e) Spherical-Deconvolution Sparse Bayesian Learning (SD-SBL) (continued).

than both nnLS and SBL, that is, for separation angles lower than 40°, RUMBA tends to detect a single fiber population, estimating the orientation to be the mean of the



(a) (b) (c) (d) (e)

Figure 3.2 Fiber orientations for angles $[50-90°]$ recovered using different reconstruction methods. Fiber orientations reconstructed with different methods for inter-fiber crossing angles in the interval $[50-90°]$. Each row represents a different inter-fiber crossing angle, starting from 0° to 40°, in discrete steps of 10°. As previously, (a) corresponds to the true fiber orientation; (b) nnLS; (c) LASSO; (d) RUMBA; and (e) SD-SBL. Synthetic signals were generated with $b = 1500$ s/mm^2, for equal fiber volume fractions $f_1 = f_2 = 0.5$, at SNR 30 dB, with fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s, and for a dictionary containing 724 directions. Differences in the depiction of RUMBA are due to the spatially continuous, smooth nature of the estimation algorithm.

two generating populations. SD-SBL exhibits a lower underestimation rate than the rest of the methods; while at 40° other methods may misestimate the number of fiber populations, SD-SBL still separates the simulated fiber populations. Furthermore, as the point clouds around the true fibers reveal, SD-SBL still separates the fibers when the crossing angle decreases to 30° for the majority of the cases. Although more dispersed across angles than for RUMBA, the number of overestimated fibers still remain low for SD-SBL.

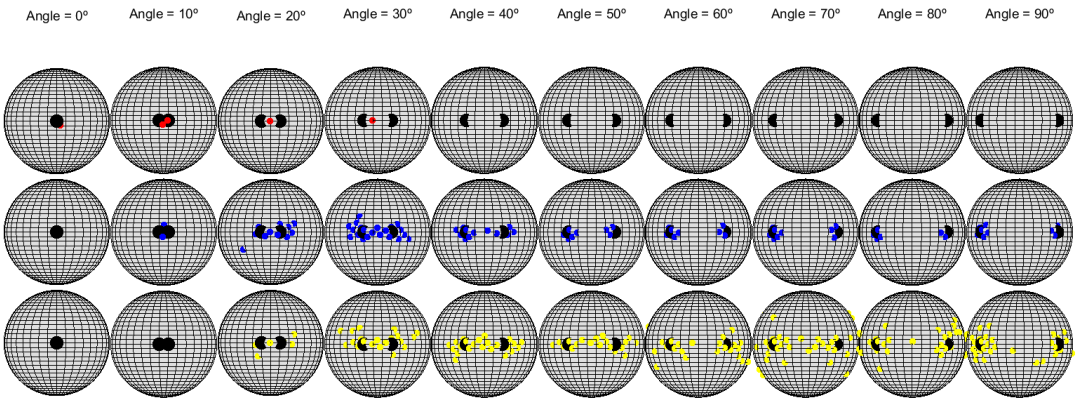


Figure 3.3 Angular distribution of estimated fibers recovered by nnLS. Angular distribution of estimated fibers for different inter-fiber angles as reconstructed by the nnLS method. Black dots represent the column-wise ground-truth fiber orientation. Each colored dot represents the orientation of an estimated fiber obtained in a single run. For a given inter-fiber angle, the population of estimated fibers is split into three groups: (i) the population of underestimated fibers (top row), which represent the orientations of the fibers (represented as red dots) for a realization that reconstructed fewer compartments compared to the ground-truth; (ii) the population of correctly estimated fibers (middle row), corresponding to the orientations of the fibers (blue dots) belonging to a run in which the right number of compartments was detected; and (iii) the population of over-estimated fibers (bottom row), which represent fiber orientations (yellow dots) corresponding to a realization in which the number of recovered fiber bundles exceeded the number of compartments in the model. The simulation was repeated 100 times with a b-value of 1500 s/mm^2, for equal fiber volume fractions $f_1 = f_2 = 0.5$, an SNR value of 30 dB, fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s, and for a dictionary containing 724 directions.

Reconstruction performance measurements for LASSO, RUMBA and SD-SBL are de-

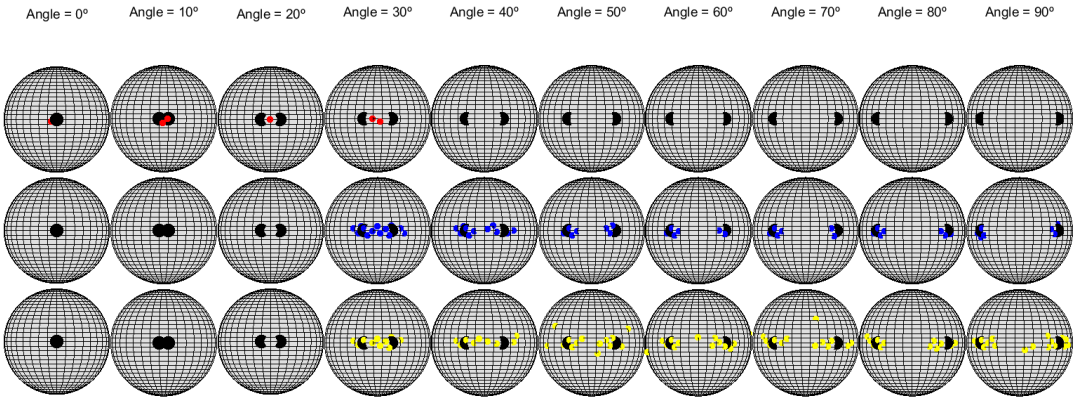


Figure 3.4 Angular distribution of estimated fibers recovered by LASSO. Angular distribution of estimated fibers for different inter-fiber angles as reconstructed by the LASSO method. Simulation (synthetic signal and dictionary) parameters were the same as those used for Fig. 3.3. Fiber orientation visual glyphs' meaning is the same as in Fig. 3.3.

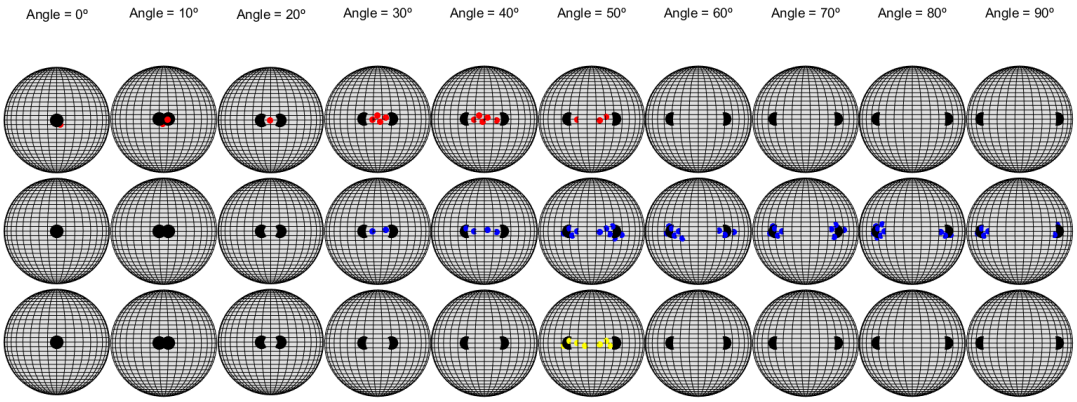


Figure 3.5 Angular distribution of estimated fibers recovered by RUMBA. Angular distribution of estimated fibers for different inter-fiber angles as reconstructed by the RUMBA method. Simulation parameters were the same as those used for Fig. 3.3. Fiber orientation visual glyphs' meaning is the same as in Fig. 3.3.

picted in Fig. 3.7 and Fig. 3.7. The mean and standard deviations were computed across the $[0 - 90°]$ IFA range in order to model uncertainty in two fibre-orientation estimates within a voxel. SD-SBL outperforms the LASSO method in the fiber orientation task

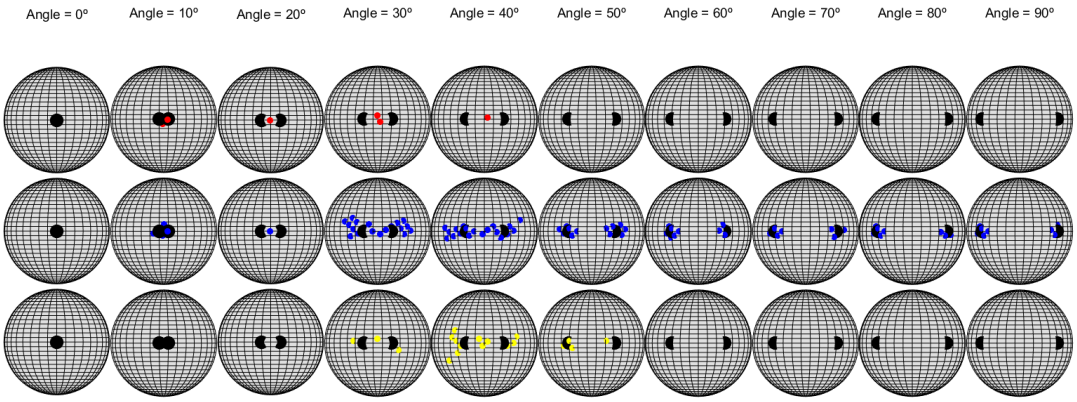


Figure 3.6 Angular distribution of estimated fibers recovered by SD-SBL. Angular distribution of estimated fibers for different inter-fiber angles as reconstructed by the SD-SBL method. Simulation parameters were the same as those used for Fig. 3.3. Fiber orientation visual glyphs' meaning is the same as in Fig. 3.3.

(Fig. 3.7). The LASSO algorithm overestimates the number of fiber compartments (top left), in agreemtn with Fig. 3.4, and shows a higher variance in the $[20-90°]$ range. LASSO detects the presence of two fiber populations around IFAs of 30°; SD-SBL is able to separate two fibers at angles as small as approximately 25°. The mean compartment misestimation penalty is consistently higher for LASSO; the standard deviation for SD-SBL is higher at low ($< 20°$) fiber separation angles. The overestimation rates of SD-SBL are limited, and higher in the $[20-40°]$ range, in agreement with Fig. 3.6. The angular error exhibited by both methods is similar across the whole range of simulated IFAs. The error's upper bound is 10°, with an approximately linear increase in the range of $[0-20°]$ for SD-SBL ($[0-25°]$ for LASSO), decreasing linearly to a value of 5° at 55°, and then exhibiting a low, nearly constant deviation from that mean value. The relative volume fraction error is 50% of the simulated fibers' volume fraction values for both methods until an IFA value of 20°. The error then decreases with a considerable gradient (from 0.5 to 0.2 in 10°), stabilizing then around a value of 0.1.

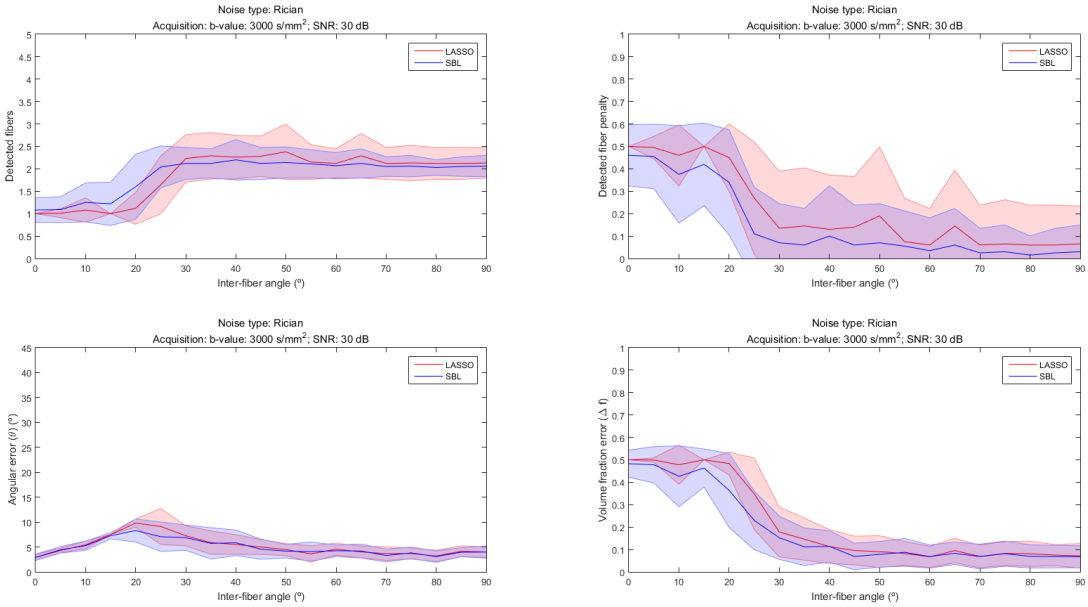


Figure 3.7 Inter-fiber angle resolution power performance for LASSO and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Quantitative measurements of the LASSO and the SD-SBL methods' fiber separation ability for an IFA range of $[0 - 90°]$. The top left plot shows the number of detected fibers; the top right plot shows the detected fiber penalty; the bottom left plot shows the angular error $\Delta\theta$ (in °); and the bottom right plot displays the volume fraction error Δf. Mean values are displayed as a dark, continuous line, and standard deviations are displayed as a shaded area around the mean. The simulation was repeated 100 times with a b-value of 3000 s/mm^2, for equal fiber volume fractions $f_1 = f_2 = 0.5$, an SNR value of 30 dB, fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s, and for a dictionary containing 724 directions. The angular step was 5°.

Although RUMBA achieves lower standard deviation figures than SD-SBL in the estimation of fiber populations, the mean angular resolution of SD-SBL is superior to RUMBA: the ability to separate the fiber compartments starts to decrease at an IFA value of 40° for RUMBA (Fig. 3.8). Except for a transition region between $20 - 40°$, where RUMBA features worse performances than SD-SBL, RUMBA exhibits a flat response (i.e. absence of any variation), for any of the parameters.

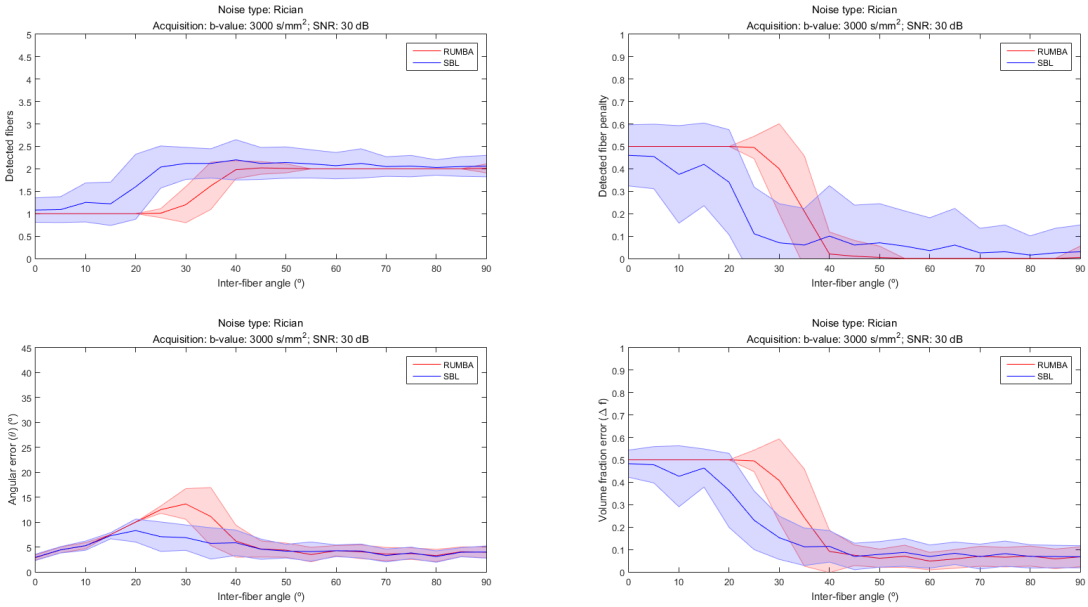


Figure 3.8 Inter-fiber angle resolution power performance for RUMBA and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Quantitative measurements of the RUMBA and the SD-SBL methods' fiber separation ability for an IFA range of $[0 - 90°]$. Simulation parameters were the same as those used for Fig. 3.7. Performance plots, mean and standard deviation visual representations are the same as in Fig. 3.7.

A summary of the individual results is displayed in Fig. 3.9, where the mean number of detected fibers, and the mean angular and volume fraction errors are displayed. The nnLS method is displayed for the sake of completeness. SD-SBL presents the desirable properties of: (i) high angular resolution (true fiber compartments are detected at lowest separation angles); (ii) low or virtually no overstimation (no false positives or non-existing fibers are estimated) across the whole IFA range; (iii) low, stable angular error after peaking at the resolution angle lower bound; and (iv) low relative volume fraction error.

The global performance figure achieved by each method at each simulated inter-fiber

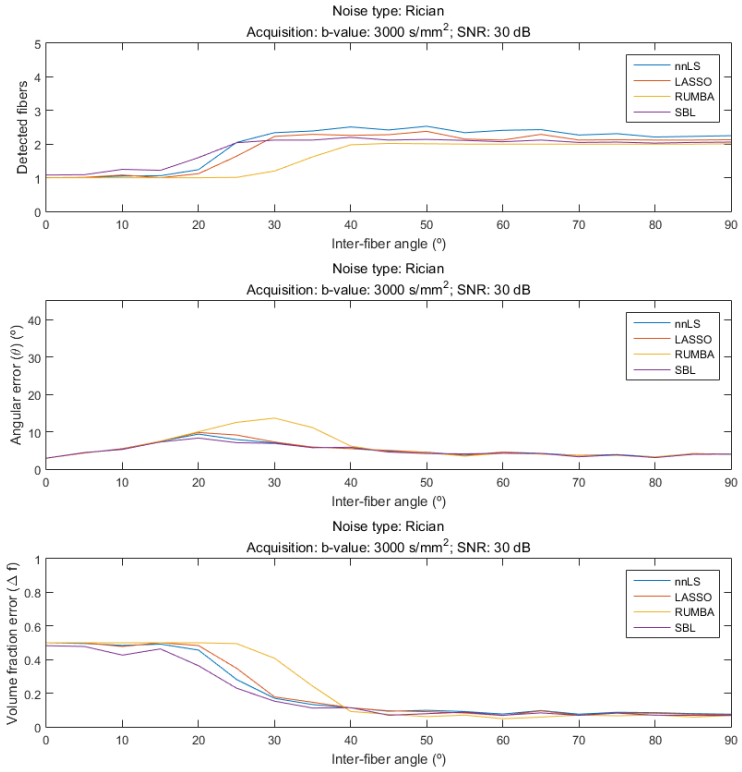


Figure 3.9 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0 - 90°]$. The top plot shows the number of detected fibers; the mid plot shows the angular error $\Delta\theta$ (in °); and the bottom plot displays the volume fraction error Δf. Statistics were computed for the recovered fibers in Fig. 3.7 and Fig. 3.8.

angle is displayed in Fig. 3.10. SD-SBL features the lowest global penalty for the $[0 - 40°]$ angle range. Interestingly enough, its performance is especially remarkable in the $[20 - 40°)$ range, where most diffusion orientation distribution function reconstruction have poorest performances. RUMBA demonstrates the best behavior in the $[40 - 90°]$ IFA subset.

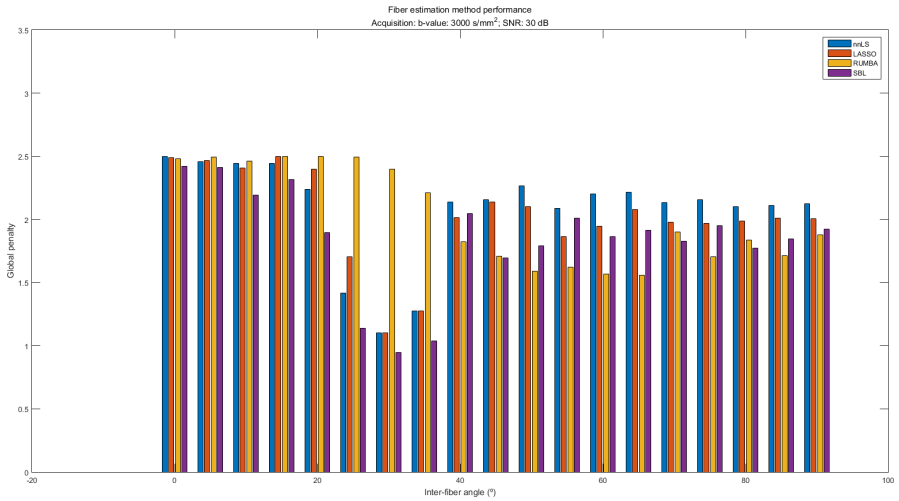


Figure 3.10 Global penalty for nnLS, LASSO, RUMBA, and SD-SBL computed on the reconstruction of a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** The global penalty is proposed to describe a method's overall performance taking into account: (i) the mean angular error; (ii) the mean volume fraction error; and (iii) the mean success rate. Each bar represents the penalty figure for the methods in the legend at each simulated inter-fiber angle. A method will have the best behaviour when its penalty figure is the lowest. The penalty figures were computed based on the statistics were obtained for a 100 cycle simulation, with a b-value of 3000 s/mm^2, equal fiber volume fractions ($f_1 = f_2 = 0.5$), an SNR value of 30 dB, fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s, and for a dictionary containing 724 directions. The angular step was 5°.

Evaluating the performance of the methods at low diffusion b-values is interesting due to the potential benefits this may have in faster acquisitions of the diffusion signal. A b-value of 1500 s/mm^2 can be considered sufficiently low compared to values reported by other works (see 2). The impact on the performance of the methods in this work of such a b-value are displayed in Fig. 3.11 and Fig. 3.12. The angular resolution power of any of the methods is affected by the variation; a decrease in anguar resolution of almost 20° is experienced by SD-SBL (10° for RUMBA) (see Fig. 3.8). However, both RUMBA and SD-SBL exhibit a flat response for a larger subset of inter-fiber angles

for the number of detected fibers, the detected fiber penalty, and the relative volume fraction error. The figures obtained by LASSO are also worse than when simulating for $b = 3000$ s/mm^2 (see Fig. 3.7).

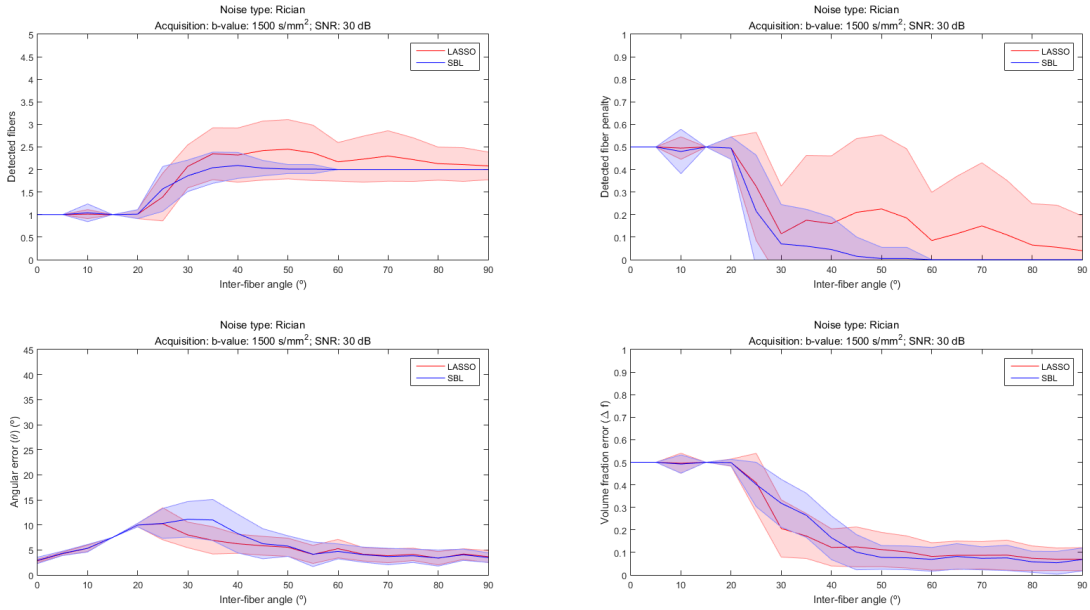


Figure 3.11 Inter-fiber angle resolution power performance for LASSO and SD-SBL on a synthetic dataset with $b = 1500$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Quantitative measurements of the LASSO and the SD-SBL methods' fiber separation ability for an IFA range of $[0 - 90°]$. The top left plot shows the number of detected fibers; the top right plot shows the detected fiber penalty, computed as the normalized penalty between the number of fibers estimated and the true fiber populations; the bottom left plot shows the angular error $\Delta\theta$ (in °); and the bottom right plot displays the volume fraction error Δf. Mean values are displayed as a dark, continuous line, and standard deviations are displayed as a shaded area around the mean. The simulation was repeated 100 times with a b-value of 1500 s/mm^2, equal fiber volume fractions ($f_1 = f_2 = 0.5$), an SNR value of 30 dB, fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s, and for a dictionary containing 724 directions. The angular step was 5°.

However, according to the mean value comparison (Fig. 3.13), SD-SBL's performance is still notable. Again, its mean overestimation rate is virtually zero. This statement os

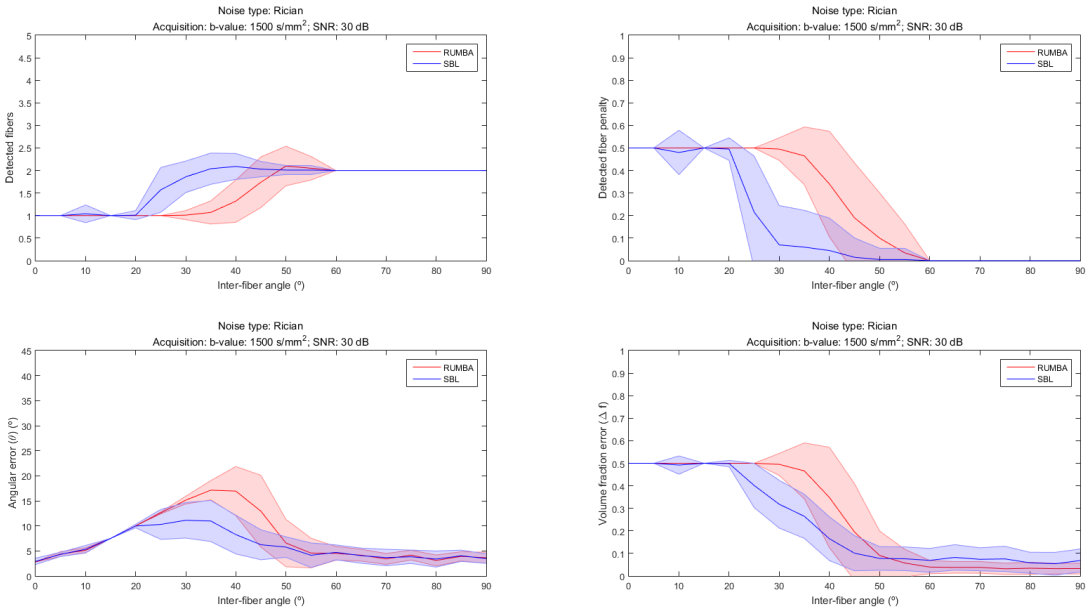


Figure 3.12 Inter-fiber angle resolution power performance for RUMBA and SD-SBL on a synthetic dataset with $b = 1500$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Quantitative measurements of the RUMBA and the SD-SBL methods' fiber separation ability for an IFA range of $[0 - 90°]$. Simulation parameters were the same as those used for Fig. 3.11. Performance plots, mean and standard deviation visual representations are the same as in Fig. 3.7.

confirmed by the global penalty figures (see Fig. 3.14): although an increase is registered in the global penalty of SD-SBL for the subset $0 - 30°$, it achieves lower global penalties with $b = 1500$ s/mm^2 above an IFA of 40°. And yet, its overall behavior is comparable to the $b = 3000$ s/mm^2 case.

At b-values lower than 1500 s/mm^2, the perfomance impact is larger. A $b = 800$ s/mm^2 can be considered as a limit case for current practice in diffusion imaging. At this rate, the angular resolution of SD-SBL is pulled back to 50°, and the angular and relative volume fraction errors are consequently higher (Fig. 3.15). However, SD-SBL

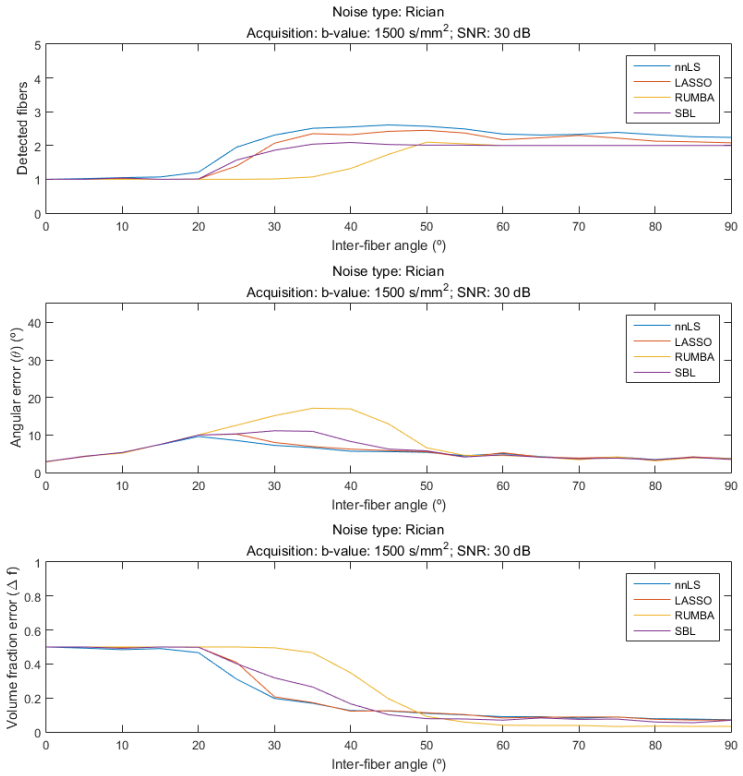


Figure 3.13 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 1500$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0 - 90°]$. Statistics were computed for a b-value of 1500 s/mm^2. The rest of simulation parameters were not varied with respect to Fig. 3.9. The plot layout reproduces the layout in Fig. 3.9.

still preserves close-to-zero overestimation rate.

Fiber bundle populations may not occupy the same relative volume fraction. Thus, studying the effects of uneven volume fractions is worthwhile. Fig. 3.16 shows the mean performance values of the fiber orientation reconstruction task across the $[0 - 90°]$ inter-fiber angle range for the $f_1 = 0.4; f_2 = 0.6$ relative volume fraction values. Compared

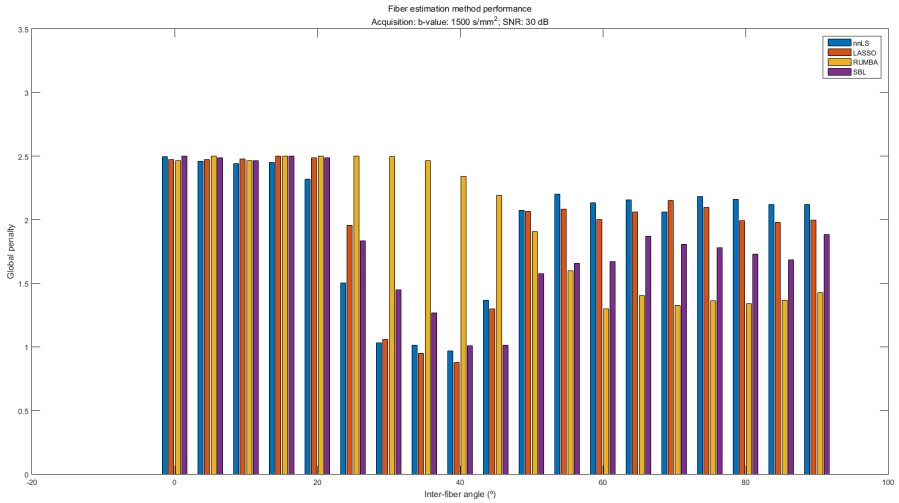


Figure 3.14 Global penalty for nnLS, LASSO, RUMBA, and SD-SBL computed on the reconstruction of a synthetic dataset with $b = 1500$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. The global penalty was computed as previously indicated for a b-value of 1500 s/mm^2. The rest of simulation parameters were not varied with respect to Fig. 3.10.

to the balanced ($f_1 = f_2 = 0.5$) case, the variations in the response are not significant for any of the methods studied.

A more extreme case for unbalanced relative volume fractions is represented in Fig. 3.18, where one of the populations was simulated with a relative volume fraction $f_1 = 1/3$, while $f_2 = 2/3$ was set for the second. Thus, one of the fiber bundles had twice the volume fraction of its counterpart. Although the response of the estimated fibers is less stable than for the ($f_1 = 0.4; f_2 = 0.6$) case, (see Fig. 3.16), and the angular error is worse, the net performance can be assimilated to the behavior in the ($f_1 = 0.4; f_2 = 0.6$) case.

The effect of differential white-matter fibre compartment diffusivities is depicted in Fig. 3.19 and Fig. 3.20. When the anisotropy is larger (Fig. 3.19), less uncertainty exists in the detection of the two fiber populations, and thus, as expected, less bias

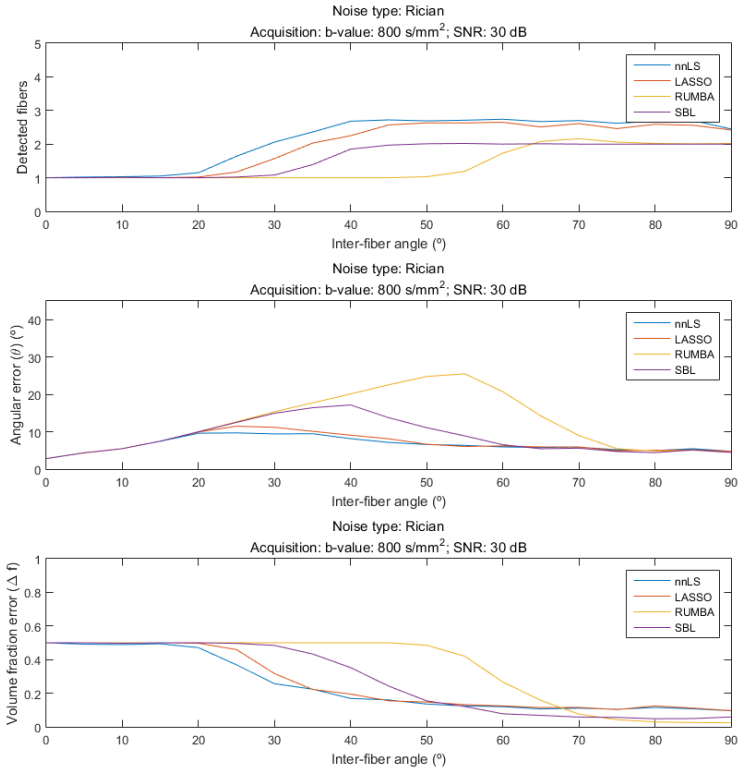


Figure 3.15 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 800$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0 - 90°]$ and a b-value of 800 s/mm^2. The rest of the simulation parameters remained unaltered with respect to the ones used for Fig. 3.9.

is observed. The net effect is a more coherent (i.e. flat) response across the entire angle range, with a sharper transition from the two-fiber population detection region to the singe-fiber detection region. On the contrary, when the mismatch between the dictionary and the diffusion signal is due to a reduced anisotropy (Fig. 3.20), a net degradation is observed in the performance, with increased overestimation rates and large variances for all methods except RUMBA. RUMBA exhibits a certain invariance

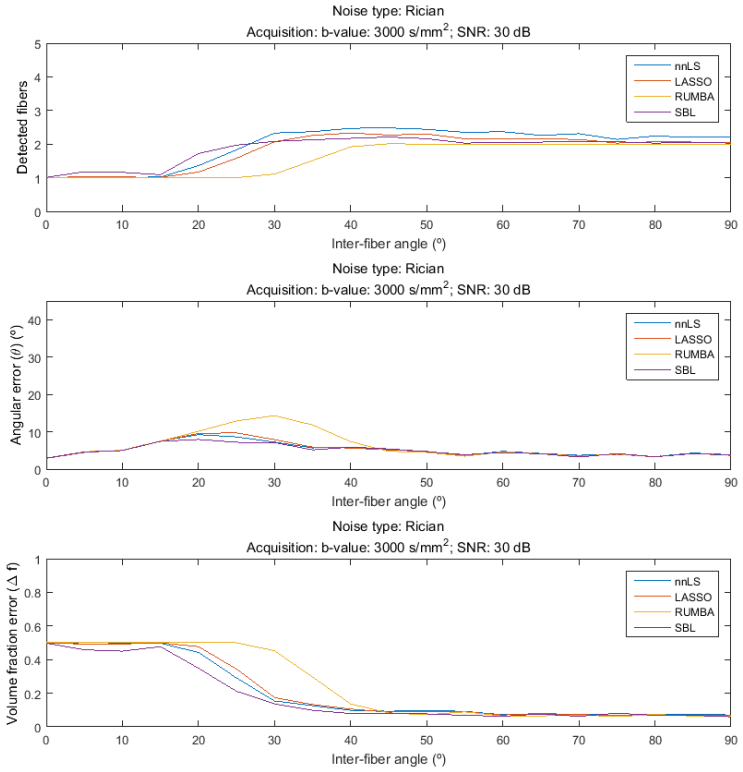


Figure 3.16 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2**,** $f_1 = 0.4; f_2 = 0.6$ **relative volume fractions, an SNR value of** 30 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0-90°]$. Statistics were computed for relative volume fractions $f_1 = 0.4; f_2 = 0.6$. The rest of simulation parameters were not varied with respect to Fig. 3.9. The plot layout reproduces the layout in Fig. 3.9.

to reduced anisotropy figures, and its response is still stable at the cost of a reduced angular contrast.

The precision of the estimated fiber orientations is also impacted by the quality of the acquistion, that is, by the Signal-to-Noise Ratio present in the samples. Previous simulations were performed at an SNR value of 30 dB, which may be considered as

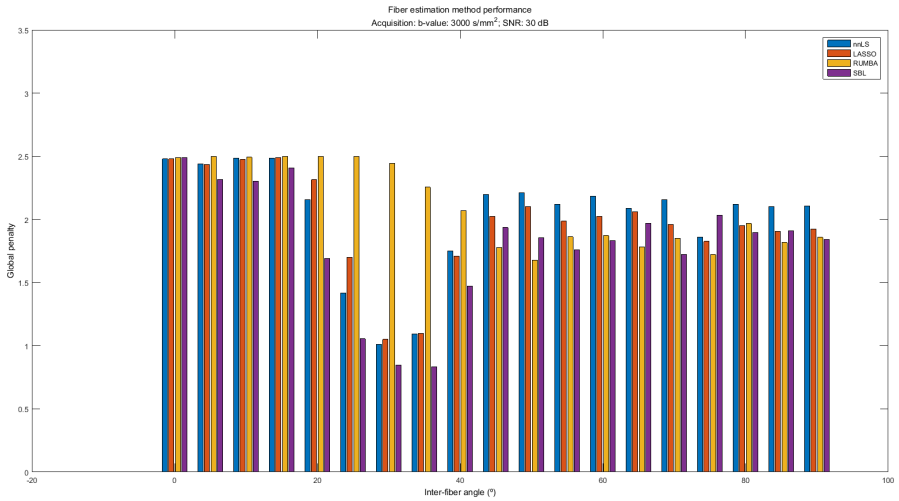


Figure 3.17 Global penalty for nnLS, LASSO, RUMBA, and SD-SBL computed on the reconstruction of a synthetic dataset with $b = 3000$ s/mm^2**,** $f_1 = 0.4; f_2 = 0.6$ **relative volume fractions, an SNR value of** 30 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** The global penalty was computed as previously indicated for relative volume fractions $f_1 = 0.4; f_2 = 0.6$. The rest of simulation parameters were not varied with respect to Fig. 3.10.

an optimum value in diffusion acquisitions. Higher noise levels negatively influence the ability to reconstruct the fiber population orientations. Fig. 3.21 and Fig. 3.22 show the performance comparisons between LASSO and SD-SBL and RUMBA and SD-SBL, respectively, in terms of the mean and standard deviation values across the $[0 - 90°]$ IFA range for an SNR value of 20 dB (a 33.3% decrease with respect to the previous value). The effect of a decline in the available SNR is a manifest increase of the standard deviations in the estimated fiber compartments, and therefore, its penalty. The effect is particularly noticeable for LASSO (Fig. 3.21), with a marked higher mean overstimation rate. The angular contrast of SD-SBL is somehow affected as well, but the two existing populations can be separated for angles as small as 35°. Even at this SNR value, SD-SBL outperforms the rest of the methods (Fig. 3.23), and maintaining a relatively low

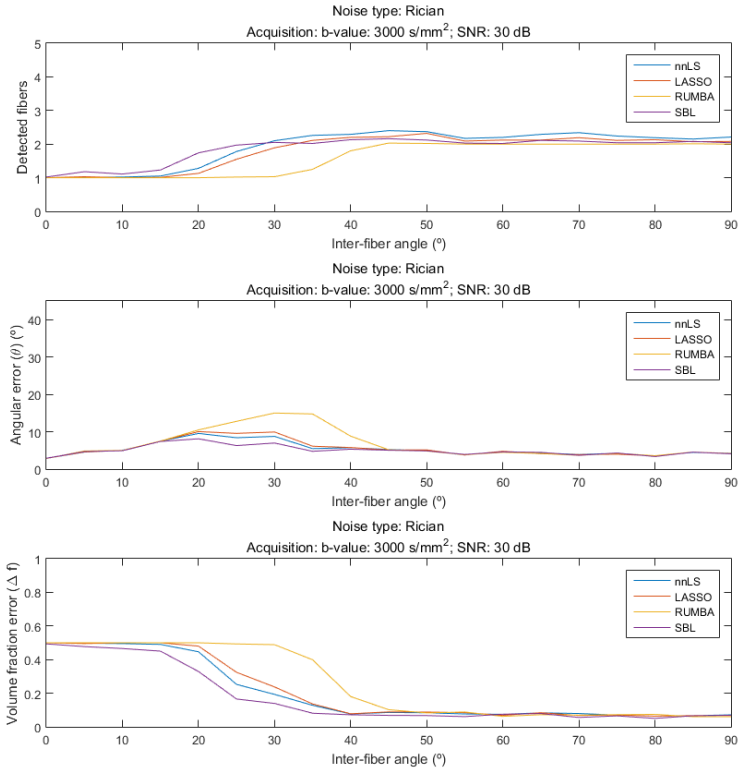


Figure 3.18 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, $f_1 = 1/3; f_2 = 2/3$ relative volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0 - 90°]$ and relative volume fractions $f_1 = 1/3; f_2 = 2/3$. The rest of the simulation parameters remained unaltered with respect to the ones used for Fig. 3.9.

global penalty (Fig. 3.24).

Finally, it is verified that SD-SBL does not introduce artefactual, physically meaningless negative relative volume fraction estimations.

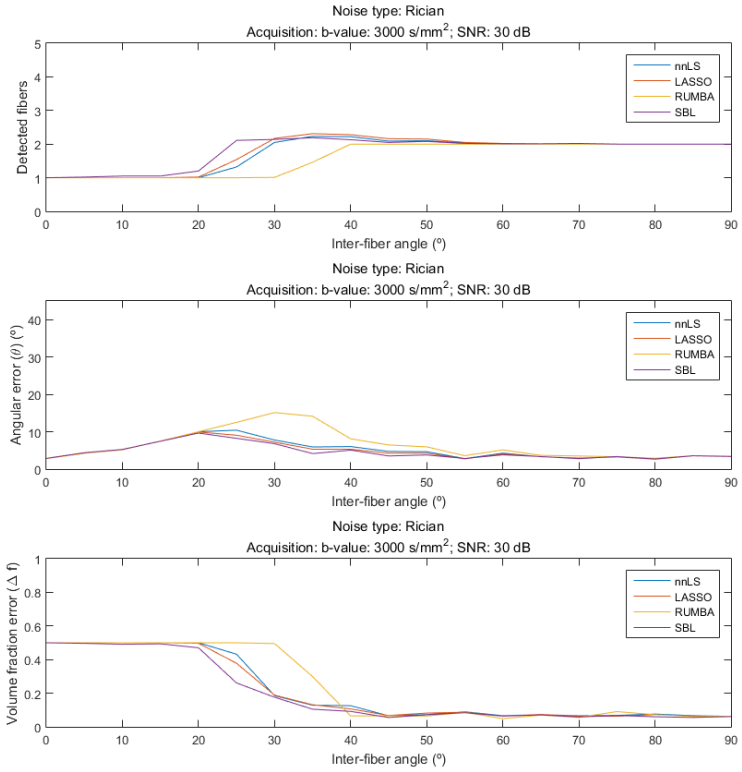


Figure 3.19 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.2×10^{-3} mm^2/s. Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0-90°]$. Statistics were computed for fiber diffusivity values of 1.7×10^{-3} and 0.2×10^{-3} mm^2/s. The rest of simulation parameters were not varied with respect to Fig. 3.9. The plot layout reproduces the layout in Fig. 3.9.

3.6.2 Phantom Model

The phantom model's fiber reconstruction was restricted to a volume of interest (VOI) of $50\times5\times50$ voxels centered around a mid-phantom coronal plane, where the compartment geometry showed a wealth of fiber configurations. For this purpose, a mask reducing the number of slices in the coronal direction was generated. Only the voxels from this subset

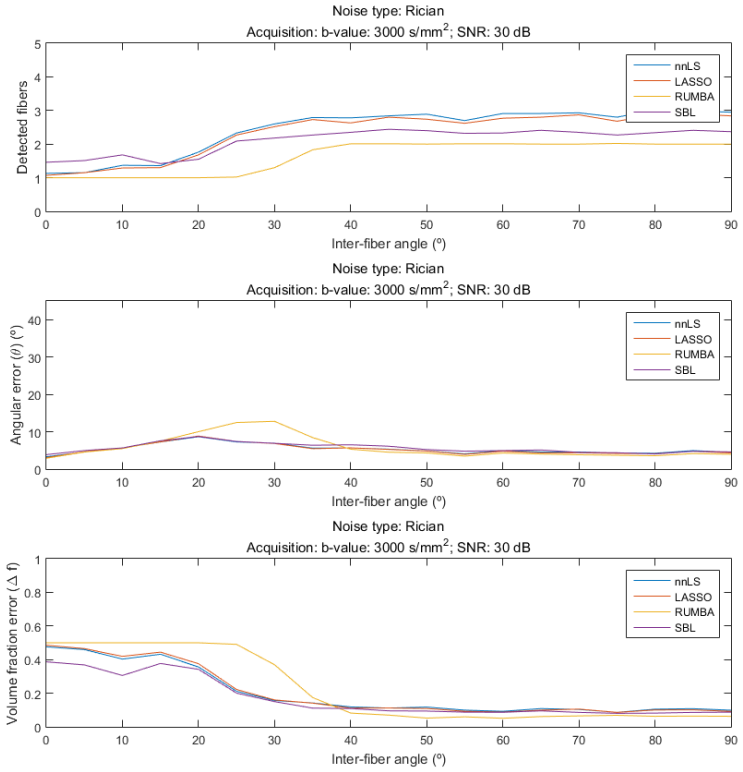


Figure 3.20 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 30 dB, and fiber diffusivity values of 1.5×10^{-3} and 0.3×10^{-3} mm^2/s. Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0-90°]$. Statistics were computed for fiber diffusivity values of 1.5×10^{-3} and 0.3×10^{-3} mm^2/s. The rest of simulation parameters were not varied with respect to Fig. 3.9. The plot layout reproduces the layout in Fig. 3.9.

were incorporated to the computation, excluding the voxels from other areas. Fig. 3.25 shows the ground-truth fiber geometry of the phantom in a coronal slice elected for fiber reconstruction analysis purposes.

Fig. 3.26 shows the fODF reconstruction for the $b = 3000$ s/mm^2 case with an applied noise level of 20 dB. Despite the heterogeneous number of fibers at each compartment,

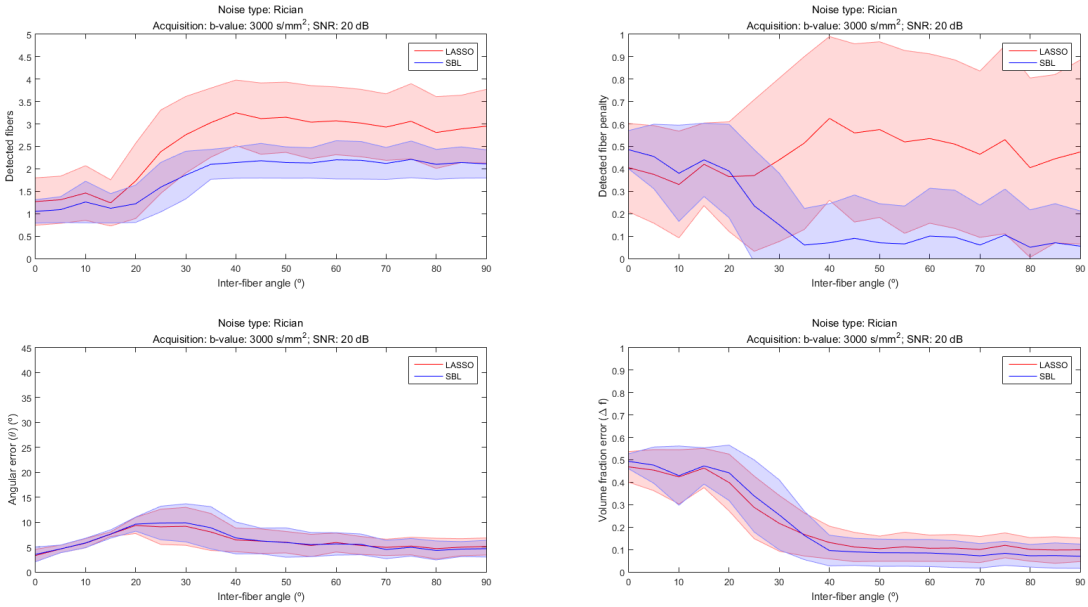


Figure 3.21 Inter-fiber angle resolution power performance for LASSO and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 20 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** Quantitative measurements of the LASSO and the SD-SBL methods' fiber separation ability for an IFA range of $[0 - 90°]$. Simulation parameters were the same as those used for Fig. 3.7 except for the SNR, fixed at 20 dB. Performance plots, mean and standard deviation visual representations are the same as in Fig. 3.7.

the figure highlights the ability of the method to resolve intra-voxel fiber orientations with a large inter-fiber angle variability. Although the method misses to reconstruct the orientation at a few single-fiber voxels with inapparent complexity (e.g. lower right region), the false positive rate is very low, i.e. no fiber bundles are detected in regions where the phantom does not contain them. A closer inspection of multi-fiber voxels indicates that the method scatteringly underestimates the number of fibers when more than two fibers conform the local micro-geometry. However, the orientations of the detected fibers is consistent with the ground truth. Thus, results are consistent with the synthetic data experiments, where the method systematically records virtually no

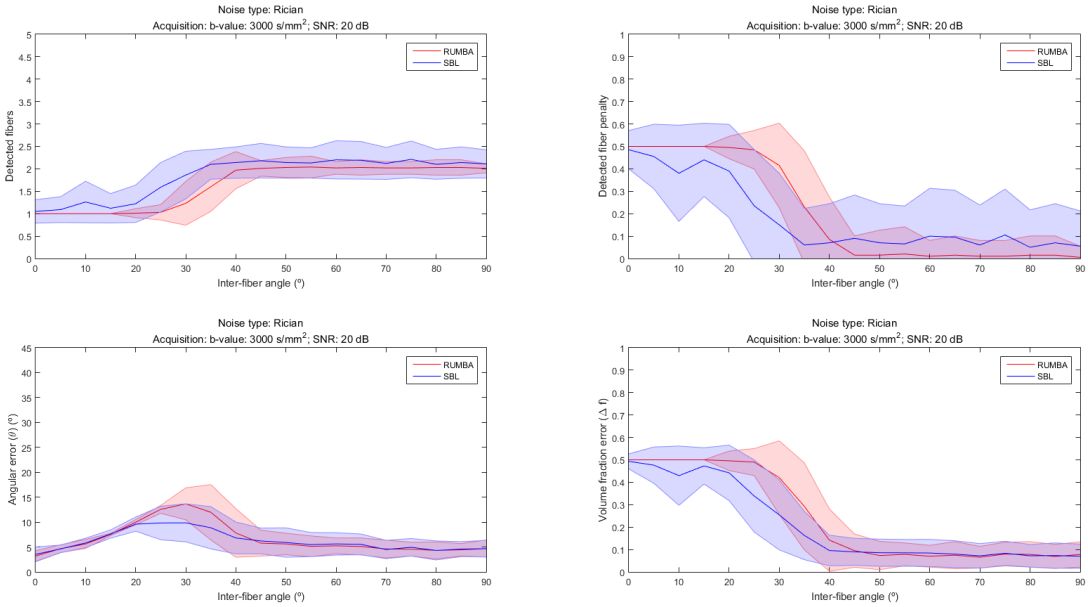


Figure 3.22 Inter-fiber angle resolution power performance for RUMBA and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 20 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** Quantitative measurements of the RUMBA and the SD-SBL methods' fiber separation ability for an IFA range of $[0 - 90°]$. Simulation parameters were the same as those used for Fig. 3.7 except for the SNR, fixed at 20 dB. Performance plots, mean and standard deviation visual representations are the same as in Fig. 3.7.

overestimation.

A 50% increase in the SNR increases the accuracy of the result, as verified in Fig. 3.27. The false positive rate (i.e. spurious fibers) is virtually zero at an SNR of 30 decibel. An increase in the accuracy of the detected fibers' orientation is also observed. Yet, the method still suffers from underestimation in regions where a three-fiber layout should be reported.

Fig. 3.28 demonstrates the performance of the RUMBA method for a SNR of 30 dB in the same region displayed in Fig. 3.27. From the image, it is apparent that RUMBA exhibits a number of limitations compared to SD-SBL. First, there is a clear penalization

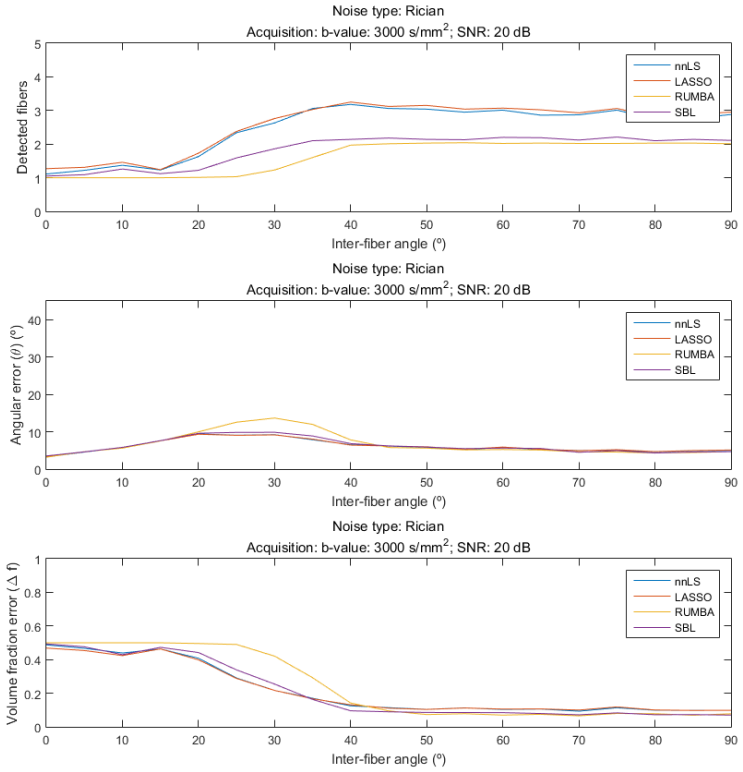


Figure 3.23 Group-wise performance comparison between nnLS, LASSO, RUMBA, and SD-SBL on a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 20 dB, and fiber diffusivity values of 1.7×10^{-3} and 0.3×10^{-3} mm^2/s. Mean fiber orientation recovery statistics for nnLS, LASSO, RUMBA and SD-SBL for an IFA range of $[0 - 90°]$ and an SNR value of 20 dB. The rest of the simulation parameters remained unaltered with respect to the ones used for Fig. 3.9.

concerning overestimation of fibers in regions where no compartment is present in the dataset (e.g. central superior part). Second, the orientation of detected fibers at low crossing angles (e.g. superior left compartments) seems to compare unfavourably with SD-SBL, including some understimation. Finally, RUMBA seems to also suffer from deficiencies in detection when at least three fiber compartments are present in a voxel.

Reconstruction of fiber orientation at a lower b-value (1200 s/mm^2), and halving

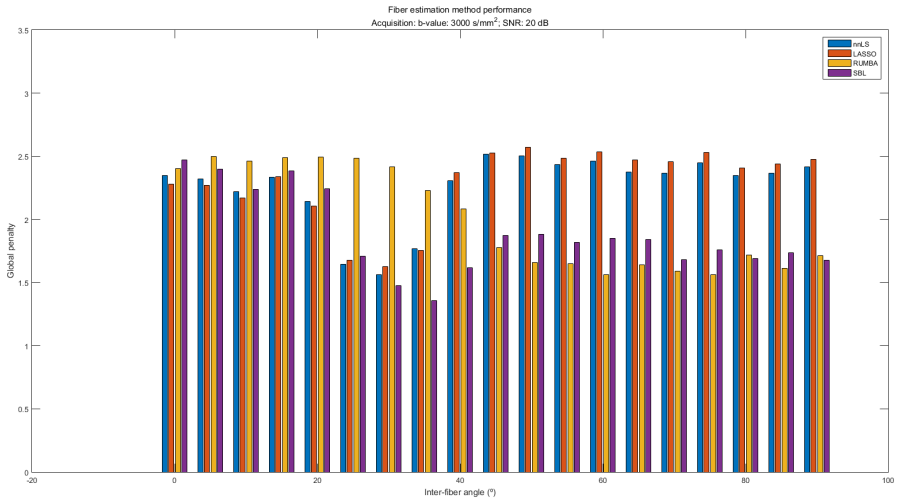


Figure 3.24 Global penalty for nnLS, LASSO, RUMBA, and SD-SBL computed on the reconstruction of a synthetic dataset with $b = 3000$ s/mm^2, equal fiber volume fractions, an SNR value of 20 dB**, and fiber diffusivity values of** 1.7×10^{-3} **and** 0.3×10^{-3} mm^2/s**.** The global penalty was computed as previously indicated for an SNR value of 20 dB. The rest of simulation parameters were not varied with respect to Fig. 3.10.

the number of gradient directions provides an expected decrease in the performance of SD-SBL, as shown by Fig. 3.29. Clearly, there is less evidence of the underlying ground truth fiber-compartment layout due to higher underestimation. Results are also confounded by the presence of spurious fibers. Effects on detected fiber orientation are also visible, with deviations from the true fiber orientation (e.g. upper left region).

The superior accuracy in fiber orientation detection and low overestimation rates comes at the cost of significantly longer computation times for SD-SBL. According to the simulations performed on the diffusion phantom, RUMBA uses 1 ms to compute the fODF at each voxel, while SD-SBL requires between 50 and 60 ms, depending on the SNR and gradient acquisition directions used.

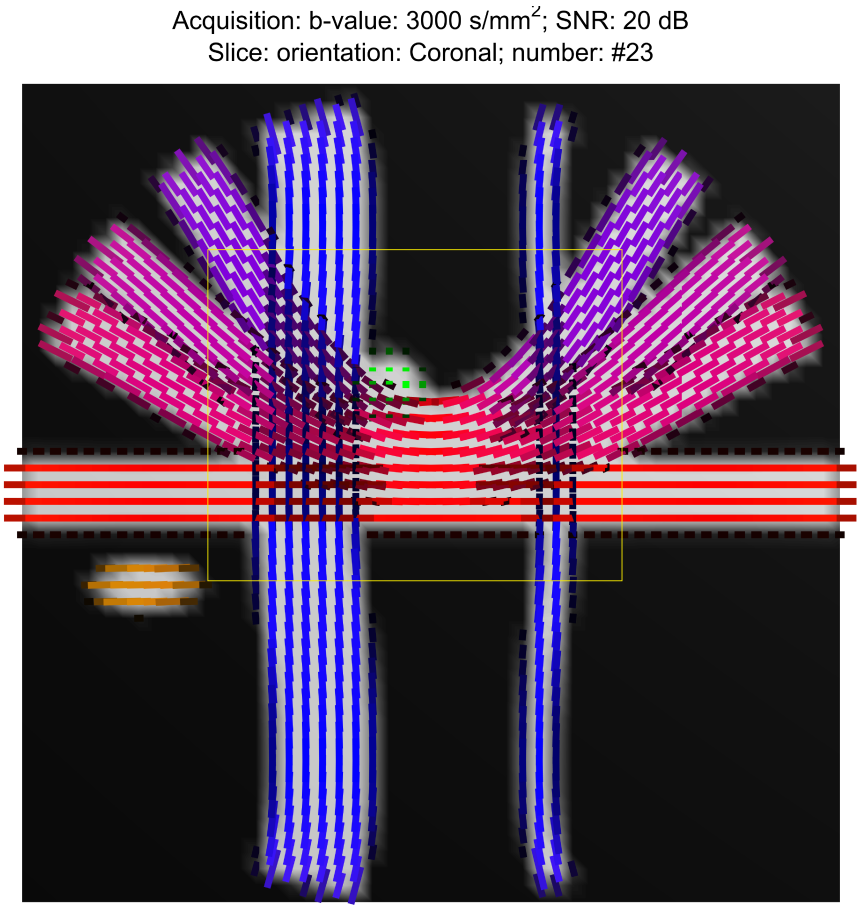


Figure 3.25 Ground truth fiber orientations for the 2nd HARDI Reconstruction Challenge phantom model. True fiber orientations for the 2nd HARDI Reconstruction Challenge[41] held at the ISBI 2013 conference. Coronal orientation central slice for the 3000 s/mm^2 b-value, 64-gradient direction, 20 dB SNR acquisition simulation. The yellow rectangle highlights a region of interest (ROI) area where the fiber orientation exhibits a significant number of fiber-crossings at different angles, together with coherently oriented fiber bundles. Fiber orientations are displayed by linear glyphs.

3.6.3 Human Brain Dataset

The results of the fiber reconstruction process on the human brain dataset are displayed in Fig. 3.30. The coronal view displays the fiber orientations according to SD-SBL. The

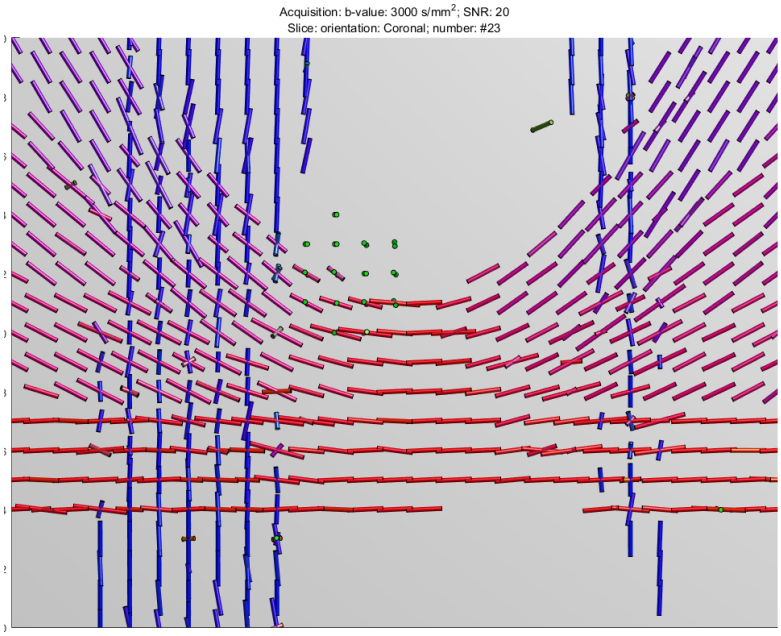


Figure 3.26 Phantom model fiber orientations recovered with SD-SBL at $b = 3000$ s/mm^2, SNR 20 dB. SD-SBL fiber orientation reconstruction result for the HARDI Challenge 2013 phantom for 64 acquisition directions, b-value 3000 s/mm^2, at SNR 20 dB. Detail from a coronal orientation central slice. Fiber orientations are represented by cylindrical glyphs.

slice comprises a variety of white-matter structures, such as the internal capsule (ic), the corpus callosum (cc) tracts, or the superior longitudinal fasciculus (slf). The first fiber package runs in an inferior-superior direction, and the cc tracts travel laterally at the medial longitudinal fissue level and continue superiorly; the slf travel mainly along the antero-posterior planes. Thus, the three fiber packages interleave their fibers in a wide region of the white matter.

Fig. 3.31 and Fig. 3.32 show fiber orientation reconstructions obtained by SD-SBL and RUMBA, respectively, in a close view of Fig. 3.30. Compared to RUMBA, SBL misses less slf fiber bundles. On the whole, both method estimates seem to be in good

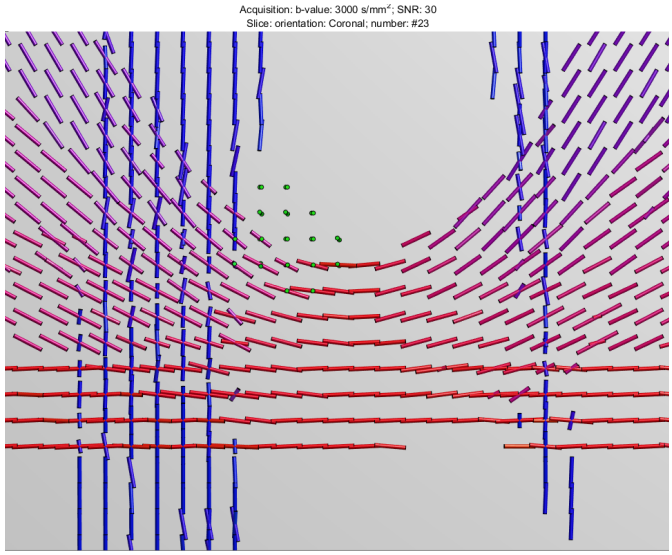


Figure 3.27 Phantom model orientations recovered with SD-SBL at reconstruction $b = 3000$ s/mm^2, SNR 30 dB. SD-SBL fiber orientation reconstruction result for the HARDI Challenge 2013 phantom for 64 acquisition directions, b-value 3000 s/mm^2, at SNR 30 dB. Detail from a coronal orientation central slice.

agreement with the known anatomy.

Again, as for the phantom, RUMBA requires one order of magnitude less computation time than SD-SBL.

3.7 Discussion

Compressed sensing has been previously used to accelerate diffusion MRI acquisitions and the recovery of diffusion features, such as the EAP[89]. The fiber orientation distribution function was introduced by Tournier *et al.*[114], along with the Spherical Deconvolution framework, as a reliable and accurate feature of the white matter axonal architecture. We have shown that the Spherical Deconvolution-Sparse Bayesian Learning fODF reconstruction technique, based on the CS theory, introduced in this work

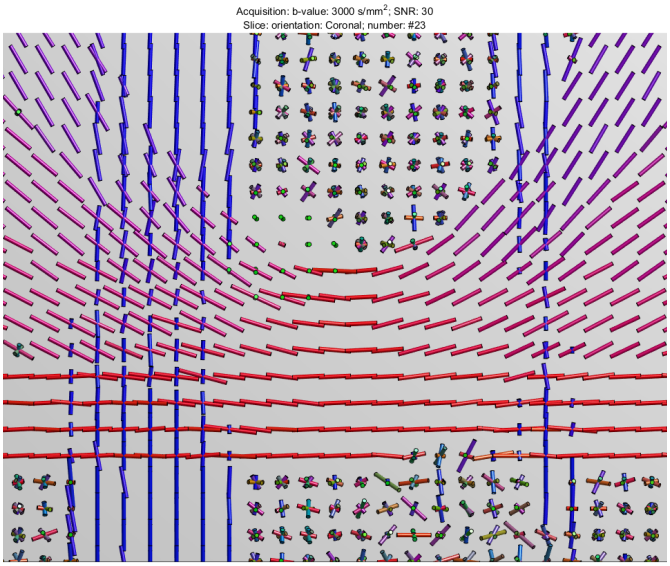


Figure 3.28 Phantom model fiber orientations recovered with RUMBA at $b = 3000$ s/mm^2, SNR 30 dB. RUMBA fiber orientation reconstruction result for the HARDI Challenge 2013 phantom for 64 acquisition directions, b-value 3000 s/mm^2, at SNR 30 dB. Detail from a coronal orientation central slice.

improves the angular resolution for the fiber population separation task. Furthermore, it provides a robust, stable performance under varying acquisition conditions in both synthetic datasets, phantom models, and human datasets.

Simulations show that nnLS, LASSO and RUMBA are able to successfully solve fibre populations separated by inter-fiber angles as small as 40°, which agrees previous results in literature[112,53,39]. Yet, SD-SBL scores a higher angular resolution (close to $[20-30°]$) under most experimental conditions.

For the validation of the method, it is essential to re-visit the often overlooked single-fiber case, as a significant proportion of imaged human white matter volumes consists of coherently oriented fiber bundles, as pointed by[100]. Although it may be argued that some of the methods separate the fiber populations at inter-fiber angles as small as 10°

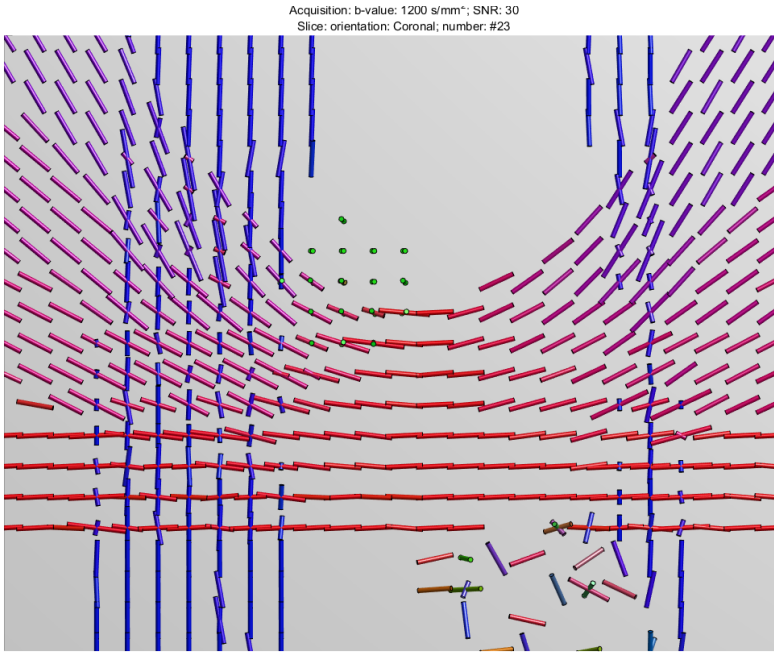


Figure 3.29 Phantom model fiber orientations recovered with SD-SBL at $b = 1200$ s/mm^2**, SNR 30** dB**.** SD-SBL fiber orientation reconstruction result for the HARDI Challenge 2013 phantom for 32 acquisition directions, b-value 1200 s/mm^2, at SNR 30 dB. Detail from a coronal orientation central slice.

(see nnLS, LASSO and SD-SBL in Fig. 3.1), they may not be statistically relevant, and it can be considered that for angles up to $[20 - 30°]$ most methods yield a single fiber. As shown by the results, when the inter-fiber angle is below $[20 - 30°]$ or when a single, coherent fiber exists in the synthetic dataset (i.e. fiber populations at $0°$) or the phantom dataset, a single fiber is detected (see, for example, Fig. 3.9). It is verified that SD-SBL's mean angular contrast is higher that its competitors', with a lower bound of close to $20°$. Although not investigated, other factors influencing the lowest achievable angular resolution may be related to the resolution of the spherical sampling grid used, or the angular cone used to estimate a fiber orientation. The angular error increases steadily

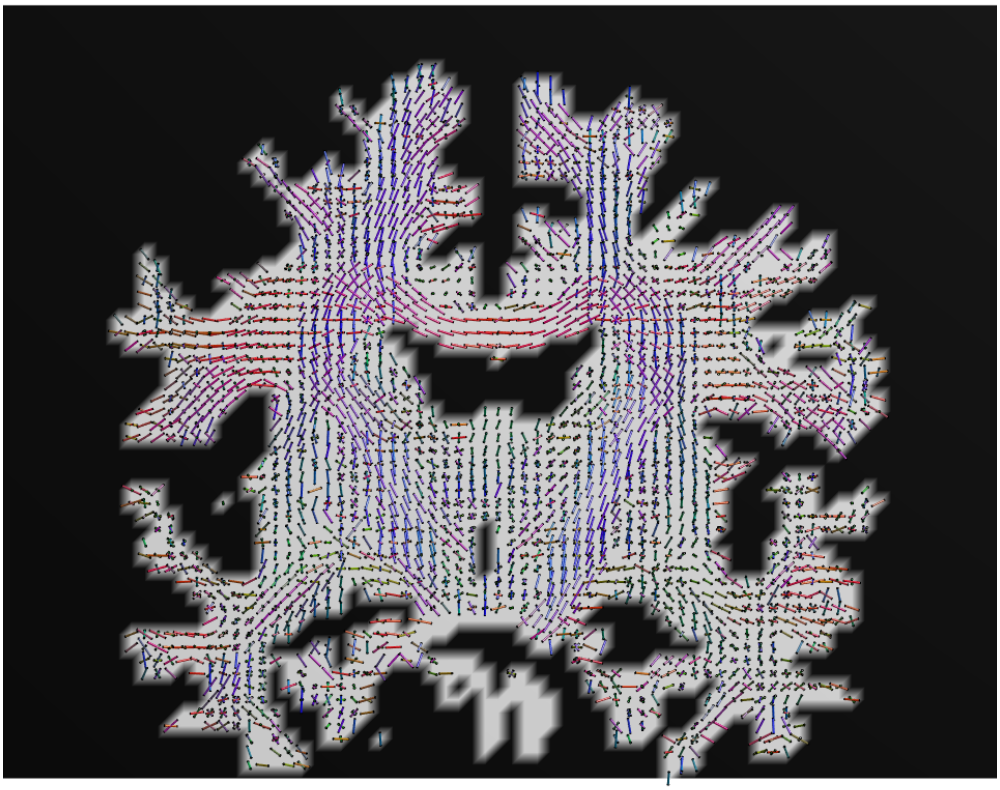


Figure 3.30 Coronal slice showing fiber orientations recovered by SD-SBL for a human brain dataset. Coronal view of the SD-SBL fiber orientation reconstruction result corresponding to a human brain dataset acquired at $b = 2000$ s/mm^2 and 100 gradient directions. The background image corresponds to the brain shape mask.

from a low value at 0° to its peak value at 20° (30° with the least favourable settings), i.e. the point where the two fiber populations are start to be successfully separated. This behavior is also mirrored in the fiber orientation distribution projections, where the detected single fiber points at half of the angle separating the two ground-truth directions (Fig. 3.6). When the simulations can be reduced to a single fiber population case, the volume fraction error is expected to be the mean value of the existing fiber packages. This is verified systematically for all simulations (see, for example, Fig. 3.18)).

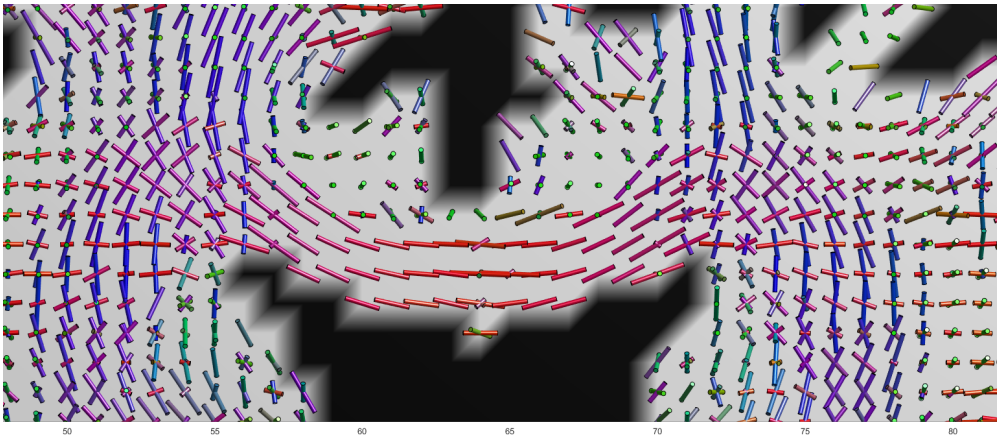


Figure 3.31 Detail of the fiber reconstruction in the corticospinal tract (cst) and corpus callosum (cc) tract regions corresponding to a human brain dataset using SD-SBL. Fiber population orientations recovered using the SD-SBL method a human brain dataset acquired at $b = 2000$ s/mm^2 and 100 gradient directions. Detailed view to a coronal slice where fibers corresponding to the corticospinal tract (cst) and the motor region of the corpus callosum (cc) tracts are visible.

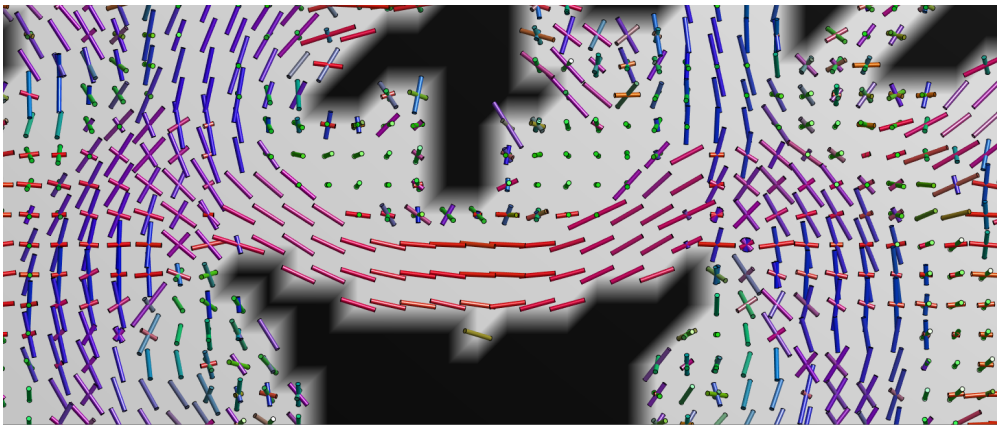


Figure 3.32 Detail of the fiber reconstruction in the cst and cc tract regions corresponding to a human brain dataset using RUMBA. Fiber population orientations recovered using the RUMBA method a human brain dataset acquired at $b = 2000$ s/mm^2 and 100 gradient directions. Detailed view to a coronal slice where fibers corresponding to the corticospinal tract (cst) and the motor region of the corpus callosum (cc) tracts are visible.

LASSO exhibits markedly low angular errors under all experimental conditions, but this is most probably driven by the excessively high overestimation rate it records, which leads, at the same time, to high detected fiber penalties. Such a low success rate explains LASSO's overall poor score across all inter-fiber angles (see Fig. 3.10). RUMBA records the lowest angular uncertainties, with flat responses at fiber crossing angles where the method consistently detects either a single fiber or the true fiber populations (see Fig. 3.8 and Fig. 3.12).

When faster acquisitions are required, lower diffusion-sensitizing gradient factors can be generally achieved. This may have a bearing on the fiber peak orientation estimation. As shown in Fig. 3.13, the angular contrast of any of the methods is affected by a lower b-value. Yet, SD-SBL's beats nnLS, LASSO and RUMBA in terms of resolution, success rate and volume fraction error. SD-SBL's resolution decreases approximately from 20° to 30°, with a smoother transition from the single-fiber to the two-fiber estimation case. However, it presents a remarkably low rate of noise amplification, comparable to RUMBA's, outside the transition zone. This allows SD-SBL to score a lower detected fiber penalty than in the $b = 3000$ s/mm^2 case. At a b-value as small as $b = 800$ s/mm^2, the deterioration in performances is yet more evident, with SD-SBL separating the true fiber populations at 40°, and consequently, recording higher angular errors at a broader range of IFAs. Again, the low angular errors scored by nnLS and LASSO are explained by the high number of extra compartments they add in the estimation.

Although the orientation contrast increases with the b-value, there is an associated decrease in the diffusion signal intensity. That is, the SNR goes down very quickly with increasing b-values. This may help understanding the substantial minimization of the standard deviation, along with a deterioration in the angular resolution, when using a lower b-value. Therefore, an optimal b-value should provide the best orientation

contrast to noise ratio.

Intra-voxel fiber heterogeneity may also account for fibers of varying relative volume fractions. The estimation of peak orientations of the bundles representing the smallest volume portion may be expected to be shaded by the largest fiber packages. When simulating uneven volume fractions ($f_1 = 0.4; f_2 = 0.6$, see Fig. 3.16), the methods' net performance remains largely invariable. When one of the population's volume fraction is 50% larger than its counterpart's, although flatness in the mean responses is affected, no relevant departures from the even volume fraction case are reported (see Fig. 3.18). Thus, it can be concluded that the methods are relatively insensitive to variations in the relative volume fraction, which agrees previous results reported in [114] and [6].

The effect of miscalibration on sensitivity to differential white-matter fiber compartment diffusivities largely depends on the sign of the signal diffusivity variation. Larger anisotropies provide a good angular contrast with reduced bias, which translates into a net improvement in performance for SD-SBL. An inverse impact is verified when the difference between the fiber diffusivities is reduced; the response is heavily penalized by the large overestimation rate.

Simulations with larger noise levels (i.e. lower SNR values) influence the results in more important ways. The impact on LASSO is easily recognizable by a larger average number of extra compartments (i.e. higher false positive rate), and the extreme penalty introduced by the large variations produced (see Fig. 3.21). As in previous results, its marked trend to estimate spurious compartments allows it to record very small angular errors. On the other hand, RUMBA's response is relatively robust to elevated noise levels (see Fig. 3.22). Although relevant, the degradation of the mean angular response of SD-SBL is not as extreme as for the LASSO, and is still close to the 30° boundary.

SD-SBL records very low overestimation rates across the whole inter-fiber simulation

range under most experimental conditions, as anticipated by results in Fig. 3.6. This is most probably explained by sparsity, which is intrinsic to the compressed sensing framework used by the SD-SBL method. Additionally, coherence shown by SD-SBL in orientations may also be an indicator of the absence of over-fitting, which is usually revealed as fibers in excess being detected at high cross angles.

Similar effects are verified for SD-SBL on the phantom dataset. Peak orientations recovered by SD-SBL are in agreement with the ground-truth at most voxels. As implied by the synthetic dataset simulations, the precision of the estimated peak orientations was found to be dependent on the acquistion b-value. When the diffusion-sensitizing gradient factor was reduced by a factor of 50%, the agreement between the ground-truth and estimated fODFs decreased. More importantly, the underestimation rate increased, with less IVOH resolved by SD-SBL. For the case of the phantom model, the decrease in the orientation accuracy is also influenced by the reduction of the available gradient directions.

The phantom model presented the additional challenge of pure isotropic voxels, where only a number of randomly oriented spurious peaks are obtained by some methods (see [45] for example). While RUMBA-SD seems to be prone to noise amplification at isotropic voxels, SD-SBL records a notable absence of fibers.

On the other hand, the reconstruction performed on the human brain dataset seems to correspond well to the knwon anatomy, with SD-SBL possibly misestimating less fiber bundles than RUMBA.

These results are in agreement with the recent evaluation carried out by Daducci *et al.*, where the sparse-like algorithms studied [39] exhibited notable performances. Under a set of experimental conditions, sparse-like methods showed the lowest number of underestimated compartments at low crossing angles.

The choice of different values for the fiber diffusivities in the dictionaries used for the phantom and the human brain dataset may be explained by the dependence of the kernel on the acquisition b-value. Although experiments show that RUMBA is robust with respect to the choice in the diffusivity values for the reconstruction dictionary, it is verified that SD-SBL's performance is influenced by the election. In particular, at a lower b-value, SD-SBL tends to provide better results (i.e. less uncertainty preserving high angular contrast at low inter-fiber angles) with higher anisotropy diffusivities, while it provides a potentially better result with lower anisotropy diffusivities at a higher b-value.

The Master, when he entered the grand temple, asked about everything. Someone said, "Who say that the son of the man of Tsâu knows the rules of propriety! He has entered the grand temple and asks about everything." The Master heard the remark, and said, "This is a rule of propriety."

Analects, III, ch. 15

Confucius, Chinese teacher, editor, politician, and philosopher (551-479 BC)

4

Conclusions and Future Work

We have presented a method to estimate the fiber orientation distribution function in diffusion Magnetic Resonance Imaging data employing a Sparse Bayesian Learning Spherical Deconvolution framework. We have demonstrated the method's ability to estimate the fiber bundles' orientations down to low crossing angles, both on synthetic datasets, phantom models and human datasets. We have compared the method to other state-of-the-art fiber orientation reconstruction methods, and have shown its improved recovery performance across a range of acquisition conditions.

4.1 Summary

SD-SBL reconstructs the fiber orientations with a high degree of accuracy, even at low diffusion-sensitizing gradient factor values. SD-SBL does not require solving any model selection problem, which may be an advantage when fiber orientations are to recovered in heterogeneous data (such as different acquisition parameters). SD-SBL's benefits draw on the correspondence of the global minimum solution estimated through sparsity with the target fiber population orientations.

The extensive simulations performed in this work leads us to highlight the following important advantages of SD-SBL:

- SD-SBL obtains the global minimum of the cost function with the minimum number of components that best describe the intra-voxel fiber population orientations. The sparsity on the solutions sought by SD-SBL allow the method to record better success rates (i.e. lower under/overestimation rates), and lower noise amplification at isotropic voxels, which is still an issue in fODF recovery.

- SD-SBL does not require nonlinear optimizations, and thus no instabilities arise from its numerical implementation.

- Non-negativity is intrinsic to the SD-SBL framework. Thus, no additional constraint needs to be imposed to avoid physically meaningless solutions (i.e. negative volume fractions).

Enhancements in the estimation of the local fiber micro-geometry at regions containing multiple fiber populations may improve the estimation of anatomical connections between distant voxels in the brain by means of fiber-tracking techniques. Such algorithms rely on accurate estimates of the white matter fiber population orientations. The

ability to produce sharper peaks at fiber orientations that are separated by smaller angles potentially improves the reliability of tracking results for tracts that pass through or close to other tracts that are similarly oriented. Furthermore, robustness on the estimated fiber orientations is critical to fiber pathway tracking.

In summary, the results obtained in this work point in the direction that the SD-SBL framework holds great promise for tractography applications.

One drawback of the SD-SBL algorithm is the computation time required, especially when being applied to large datasets or when considering clinical setting time-frames. The time elapsed to solve the fiber distribution is an order of magnitude greater than other comparable methods. However, SD-SBL's time ineffectiveness is not linked to the acquisition strategy or requirements, as opposed to other methods; it is hardware- and implementation-dependent. This work did not incorporate strategies to minimize the impact of large computation times. Thus, it may be expected that improvements in the implementation aspects and the hardware used might turn SD-SBL fit for on-line diffusion orientation reconstruction tasks.

4.2 Future work

Elucidating the brain's white matter micro-architecture through Magnetic Resonance Imaging will require new models capable of differentiating the individual compartments within the fiber pathways, as well as methods that obtain sharper fiber orientations. The method presented in this work offers improved angular sensitivity for fiber orientation reconstruction.

Some limitations and future extensions to the current work are related to the signal and noise generative models, as well as to the method's overall time performance.

First, several strategies exist for creating the signal dictionary, like using mixtures of intra-compartment models to capture different diffusivity profiles. Therefore, exploring the dictionary that provides the most effective signal representation deserves to be investigated. Second, we have not included additional strategies, such as spatial regularizations, to enhance the method's performance. Third, a broader validation of the method is required. A few efforts have emerged in the last years in order to provide a common testing ground (such as the "Phantomas" [29], an early version of which has been used as the phantom model in this work, or "Diffantom" [55] datasets), as well as universal performance reporting frameworks for tractography pipelines (such as "Tractometer" [37]) for diffusion reconstruction methods. Using such tools may help in obtaining standardized scores for method comparison purposes. Data provided by the two Human Connectome Project consortia [110] [54] may also be useful to test the method on real diffusion data acquired under different acquisition conditions. Furthermore, fiber-tracking methods applied to fODF reconstruction algorithms on both phantom and real, whole-brain datasets may help assessing a method's fitness for clinical practice.

And finally, issues linked to real data acquisition would require a closer study. On the one hand, although low diffusion-sensitizing gradients may provide sufficient angular contrast when using SD-SBL in synthetic datasets, issues stemming from fewer gradient direction acquisitions or different sampling strategies need further investigation. Certainly, an optimal value providing a sufficient SNR for a good angular resolution should exist for a given acquisition strategy. On the other hand, although we have used a Rician noise model, we have not investigated the factors affecting the noise. The observed noise on a Magnetic Resonance image depends on many factors, including the number of coils in the scanner and the multi-channel image combination method. Thus, further work is required to study the influence of the scanner settings into the proposed algo-

rithm. Second, although we have tested the method on synthetic datasets, phantom models and real datasets for different noise levels, we have not directly addressed the impact on the quality of reconstructions of other imaging artifacts such as eddy currents, B_0-field inhomogeneity and subject motion. In the same way, we have not studied the performance of the method when different **q**-space acquisition strategies (such as partial q-space acquisition) may be employed for accelerating the acquisition process, or when data from multiple shells is available. Yet it may be an interesting aspect to consider.

A

Appendix A

A.1 Spherical Harmonic Basis

The Spherical Harmonics (SH), normally indicated by $Y_l^m(\theta, \phi)$, where l denotes the degree, m the order or phase factor, and (θ, ϕ) the angular direction, are a basis for complex functions on the unit sphere satisfying the SH differential equation

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial F}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 F}{\partial\theta^2} + l(l+1)F = 0, \quad l \in \mathbb{Z}_+ \tag{A.1}$$

where θ and ϕ are the polar and azimuth rotation angles in spherical coordinates, respectively; and F is the function being analyzed.

The first two terms of this equation correspond to the Laplacian in spherical coordinates, also called the three dimensional Laplace-Beltrami operator Δ_b. It is a natural measure of smoothness for functions defined on the unit sphere, satisfying the relations $\Delta_b Y_l^m = -l(l+1)Y_l^m$.

For each nonnegative integer l there are exactly $2l+1$ spherical harmonics given by $Y_l^{-l}, \ldots, Y_l^0, \ldots, Y_l^l$, i.e., $m = -l, \ldots, 0, \ldots, l$. Explicitly, they are given as

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

where P_l^m is an associated Legendre polynomial of degree l and order m.

The set of spherical harmonics forms an orthonormal basis for all functions on the unit sphere. Since they form a basis, any spherical function $x : S^2 \to C$ can be written as $x(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_l^m Y_l^m$. This is called the Spherical Harmonic Transform (SHT). Due to orthonormality of the SH basis, the coefficients of the SH series c_m^l can be calculated by forming the inner product of x with the spherical harmonics, $c_l^m = \langle x(\theta, \phi) Y_l^m(\theta, \phi) \rangle = \int_0^{2\pi} \int_0^{\pi} x(\theta, \phi) Y_l^m(\theta, \phi) \sin(\theta) d\theta d\phi$. This is analogous to a Fourier decomposition of sinusoidal functions, but on the unit sphere.

Finally, even order spherical harmonics are antipodally symmetric, while the odd order spherical harmonics are antipodally antisymmetric.

B

Appendix B

B.1 Rician distribution

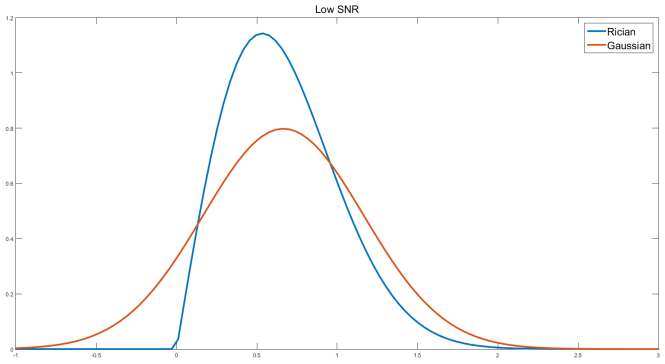
Magnetic Resonance Imaging magnitude data are known to be corrupted by a noise that follows a Rician (or Rice) distribution[62]. Rician noise produces, on average, a net increase in the signal which is different along every gradient direction. The probability density function of the Rician distribution follows the law in Eq. B.1

$$f(x|\nu,\sigma^2) = \frac{x}{\sqrt{\sigma^2}} e^{-\frac{-(x^2+\nu^2)}{2\sigma^2}} I_0\left(\frac{x\nu}{\sigma^2}\right) \tag{B.1}$$

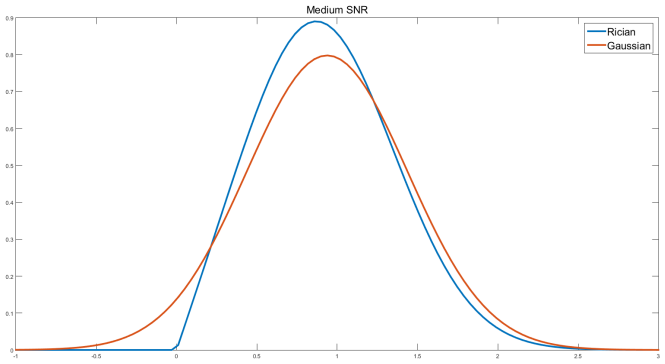
where $I_0(z)$ is the modified Bessel function of the first kind with order zero. Note that in this expression ν and σ are not the mean and standard deviation of the Rician distribution.

The Rician distribution is related to the Gaussian distribution: a distribution R has a Rice distribution ($R \sim \text{Rice}(\nu, \sigma)$) if $R = \sqrt{X^2 + Y^2}$, where $X \sim N\left(\nu \cos\theta, \sigma^2\right)$ and $Y \sim N\left(\nu \sin\theta, \sigma^2\right)$ are statistically independent normal random variables and θ is any real number.

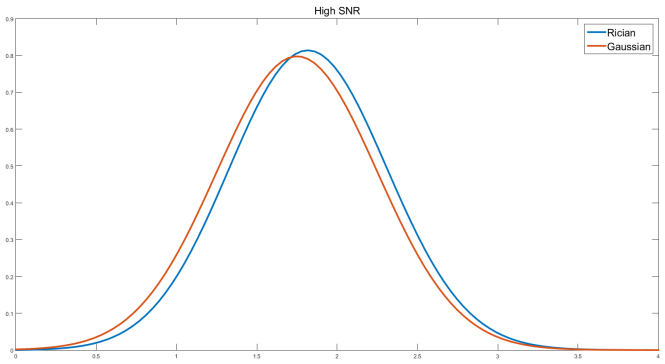
In many diffusion reconstruction works (see, for example,[46,114], and[117]), it is assumed that the noise obeys a Gaussian distribution law ($N(\mu, \sigma^2)$), which has the probability density $f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. However, the deviation of the Rician distribution from Gaussianity is particularly evident at low SNR values, as verified in Fig. B.1. The Gaussian distribution can be considered a good approximation of the Rician distribution only above an SNR of approximately 5:1[62]. According to Daducci *et al.*[39], reconstructions assuming zero-mean Gaussian noise produce small amplitude, spurious fibers to explain this discrepancy. Few works have studied (see, for example,[50,82] or[53]) the impact on fiber orientation reconstruction of Rician noise.



(a)



(b)



(c)

Figure B.1 Rician distribution at different SNR values. Comparison between the Rician distribution and the Gaussian distribution at different SNR values: (a) low SNR; (b) medium SNR; and (c) high SNR.

References

[1] Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., & Harel, N. (2010). Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2), 554–566.

[2] Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing*, 54(11), 4311–4322.

[3] Alexander, A. L., Hasan, K. M., Lazar, M., Tsuruda, J. S., & Parker, D. L. (2001). Analysis of partial volume effects in diffusion-tensor MRI. *Magnetic Resonance in Medicine*, 45(5), 770–780.

[4] Alexander, D. C. (2005). Multiple-Fiber Reconstruction Algorithms for Diffusion MRI. *Annals of the New York Academy of Sciences*, 1064(1), 113–133.

[5] Alexander, D. C., Barker, G. J., & Arridge, S. R. (2002). Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2), 331–340.

[6] Anderson, A. W. (2005). Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(5), 1194–1206.

[7] Assaf, Y. & Basser, P. J. (2005). Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27(1), 48 – 58.

[8] Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., & Basser, P. J. (2008). Axcaliber: A method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, 59(6), 1347–1354.

[9] Assaf, Y. & Cohen, Y. (1998). Non-Mono-Exponential Attenuation of Water andN-Acetyl Aspartate Signals Due to Diffusion in Brain Tissue. *Journal of Magnetic Resonance*, 131(1), 69 – 85.

[10] Assaf, Y., Freidlin, R. Z., Rohde, G. K., & Basser, P. J. (2004). New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52(5), 965–978.

[11] Assemlal, H.-E., Tschumperlé, D., & Brun, L. (2009). Efficient and robust computation of PDF features from diffusion MR signal. *Medical Image Analysis*, (pp. 1–16).

[12] Basser, P. J., Mattiello, J., & Bihan, D. L. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1), 259–267.

[13] Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in Biomedicine*, 15(7-8), 435–455.

[14] Beaulieu, C. & Allen, P. S. (1994). Determinants of anisotropic water diffusion in nerves. *Magnetic Resonance in Medicine*, 31(4), 394–400.

[15] Beck, A. & Teboulle, M. (2009). A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, 2(1), 183–202.

[16] Behrens, T. E. J., Johansen-Berg, H., Jbabdi, S., Rushworth, M. F. S., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1), 144–155.

[17] Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., Matthews, P. M., Brady, J. M., & Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5), 1077–1088.

[18] Bilgic, B., Setsompop, K., Cohen-Adad, J., Yendiki, A., Wald, L. L., & Adalsteinsson, E. (2012). Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries. *Magnetic Resonance in Medicine*, 68(6), 1747–1754.

[19] Bloch, F. (1946). Nuclear Induction. *Physical Review*, 70(7-8), 460–474.

[20] Bloch, F., Hansen, W. W., & Packard, M. (1946). The Nuclear Induction Experiment. *Physical Review*, 70(7-8), 474–485.

[21] Canales-Rodŕiguez, E. J., Iturria-Medina, Y., Alemán-Gómez, Y., & Melie-Garćia, L. (2010a). Deconvolution in diffusion spectrum imaging. *NeuroImage*, 50(1), 136–149.

[22] Canales-Rodŕiguez, E. J., Lin, C.-P., Iturria-Medina, Y., Yeh, C.-H., Cho, K.-H., & Melie-Garćia, L. (2010b). Diffusion orientation transform revisited. *NeuroImage*, 49(2), 1326–1339.

[23] Canales-Rodŕiguez, E. J., Melie-Garćia, L., & Iturria-Medina, Y. (2009). Mathematical description of q-space in spherical coordinates: Exact q-ball imaging. *Magnetic Resonance in Medicine*, 61(6), 1350–1367.

[24] Candès, E. J., Romberg, J., & Tao, T. (2006a). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2), 489–509.

[25] Candès, E. J., Romberg, J. K., & Tao, T. (2006b). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8), 1207–1223.

[26] Candes, E. J. & Tao, T. (2006). Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? *IEEE Transactions on Information Theory*, 52(12), 5406–5425.

[27] Candès, E. J. & Wakin, M. B. (2008). An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*, 25(2), 21–30.

[28] Carr, H. Y. & Purcell, E. M. (1954). Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical Review*, 94(3), 630–638.

[29] Caruyer, E., Daducci, A., Descoteaux, M., Houde, J.-C., Thiran, J.-P., & Verma, R. (2014). Phantomas: A flexible software library to simulate diffusion MR phantoms. In *Proceedings of the Joint Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM)-European Society for Magnetic Resonance in Medicine and Biology(ESMRMB)* (pp. 6407). Milan, Italy.

[30] Caruyer, E. & Deriche, R. (2012). Diffusion MRI signal reconstruction with continuity constraint and optimal regularization. *Medical Image Analysis*, 16(6), 1113–1120.

[31] Cheng, J., Ghosh, A., Deriche, R., & Jiang, T. (2010a). : (pp. 648–656). Berlin, Heidelberg: Springer Berlin Heidelberg.

[32] Cheng, J., Ghosh, A., Jiang, T., & Deriche, R. (2010b). : (pp. 590–597). Berlin, Heidelberg: Springer Berlin Heidelberg.

[33] Clark, C. A. & Le Bihan, D. (2000). Water diffusion compartmentation and anisotropy at high b values in the human brain. *Magnetic Resonance in Medicine*, 44(6), 852–859.

[34] Cleveland, G. G., Chang, D. C., Hazlewood, C. F., & Rorschach, H. E. (1976). Nuclear magnetic resonance measurement of skeletal muscle: Anisotrophy of the diffusion coefficient of the intracellular water. *Biophysical Journal*, 16(9), 1043–1053.

[35] Cohen, Y. & Assaf, Y. (2002). High b-value q-space analyzed diffusion-weighted MRS and MRI in neuronal tissues –a technical review. *NMR in Biomedicine*, 15(7-8), 516–542.

[36] Cory, D. G. & Garroway, A. N. (1990). Measurement of translational displacement probabilities by NMR: An indicator of compartmentation. *Magnetic Resonance in Medicine*, 14(3), 435–444.

[37] Côté, M.-A., Girard, G., Boré, A., Garyfallidis, E., Houde, J.-C., & Descoteaux, M. (2013). Tractometer: Towards validation of tractography pipelines. *Medical Image Analysis*, 17(7), 844–857. Special Issue on the 2012 Conference on Medical Image Computing and Computer Assisted Intervention.

[38] Cotter, S. F., Rao, B. D., Engan, K., Kreutz-delgado, K., & Member, S. (2005). Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing*, 53(7), 2477–2488.

[39] Daducci, A., Canales-Rodríguez, E. J., Descoteaux, M., Garyfallidis, E., Gur, Y., Lin, Y.-C., Mani, M., Merlet, S., Paquette, M., Rairez-Manzanares, A., Reisert,

M., Rodrigues, P. R., Sepehrband, F., Caruyer, E., Choupan, J., Deriche, R., Jacob, M., Menegaz, G., Přckovska, V., Rivera, M., Wiaux, Y., & Thiran, J.-P. (2014). Quantitative Comparison of Reconstruction Methods for Intra-Voxel Fiber Recovery From Diffusion MRI. *Medical Imaging, IEEE Transactions on*, 33(2), 384–399.

[40] Daducci, A., Canales-Rodŕiguez, E. J., Zhang, H., Dyrby, T. B., Alexander, D. C., & Thiran, J.-P. (2015). Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. *NeuroImage*, 105, 32–44.

[41] Daducci, A., Caruyer, E., Descoteaux, M., Houde, J.-C., & Thiran, J.-P. (2013a). HARDI Reconstruction Challenge 2013.

[42] Daducci, A., Palù, A. D., Lemkaddem, A., & Thiran, J.-P. (2013b). A convex optimization framework for global tractography. In *2013 IEEE 10th International Symposium on Biomedical Imaging* (pp. 524–527).

[43] Damadian, R. (1971). Tumor Detection by Nuclear Magnetic Resonance. *Science*, 171(3976), 1151–1153.

[44] Damadian, R., Goldsmith, M., & Minkoff, L. (1977). NMR in cancer: XVI. FONAR image of the live human body. *Physiological Chemistry and Physics*, 9(39176), 97–100.

[45] Dell'Acqua, F., Rizzo, G., Scifo, P., Clarke, R. A., Scotti, G., & Fazio, F. (2007). A Model-Based Deconvolution Approach to Solve Fiber Crossing in Diffusion-Weighted MR Imaging. *Biomedical Engineering, IEEE Transactions on*, 54(3), 462–472.

[46] Dell'Acqua, F., Scifo, P., Rizzo, G., Catani, M., Simmons, A., Scotti, G., & Fazio, F. (2010). A modified damped Richardson–Lucy algorithm to reduce isotropic background effects in spherical deconvolution. *NeuroImage*, 49(2), 1446–1458.

[47] Descoteaux, M., Angelino, E., Fitzgibbons, S., & Deriche, R. (2006). Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. *Magnetic Resonance in Medicine*, 56(2), 395–410.

[48] Descoteaux, M., Angelino, E., Fitzgibbons, S., & Deriche, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 58(3), 497–510.

[49] Descoteaux, M. & Deriche, R. (2015). From Local Q-Ball Estimation to Fibre Crossing Tractography. In N. Paragios, J. Duncan, & N. Ayache (Eds.), *Handbook of Biomedical Imaging: Methodologies and Clinical Research* (pp. 455–473). Boston, MA: Springer US.

[50] Descoteaux, M., Wiest-Daesslé, N., Prima, S., Barillot, C., & Deriche, R. (2008). Impact of Rician Adapted Non-Local Means Filtering on HARDI. In D. Metaxas, L. Axel, G. Fichtinger, & G. Székely (Eds.), *Medical Image Computing and Computer-Assisted Intervention –MICCAI 2008*, volume 5242 of *Lecture Notes in Computer Science* (pp. 122–130). Springer Berlin Heidelberg.

[51] Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4), 1289–1306.

[52] Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8), 549–560.

[53] Erick Jorge Canales-Rodŕiguez, Daducci, A., Sotiropoulos, S. N., Caruyer, E., Aja-Fernández, S., Radua, J., Mendizabal, J. M. Y., Iturria-Medina, Y., Melie-Garćia, L., Alemán-Gómez, Y., Thiran, J.-P., Sarró, S., Pomarol-Clotet, E., & Salvador, R. (2015). Spherical Deconvolution of Multichannel Diffusion MRI Data with Non-Gaussian Noise Models and Spatial Regularization. *PLoS ONE*, 10(10), e0138910.

[54] Essen, D. C. V., Ugurbil, K., Auerbach, E. J., Barch, D., Behrens, T. E. J., Bucholz, R. D., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W., Penna, S. D., Feinberg, D. A., Glasser, M. F., Harel, N., Heath, A. C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S. E., Prior, F., Schlaggar, B. L., Smith, S. M., Snyder, A. Z., Xu, J., & Yacoub, E. (2012). The Human Connectome Project: A data acquisition perspective. *NeuroImage*, 62(4), 2222–2231. Connectivity.

[55] Esteban, Ó., Caruyer, E., Daducci, A., Bach-Cuadra, M., Ledesma-Carbayo, M. J., & Santos, A. (2016). Diffantom: Whole-brain diffusion MRI phantoms derived from real datasets of the Human Connectome Project. *Frontiers in Neuroinformatics*, 10(4).

[56] Fieremans, E., Jensen, J. H., & Helpern, J. A. (2011). White matter characterization with diffusional kurtosis imaging. *NeuroImage*, 58(1), 177–188.

[57] Frank, L. R. (2002). Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47(6), 1083–1099.

[58] Freeden, W. & Schreiner, M. (1998). Orthogonal and Nonorthogonal Multiresolution Analysis, Scale Discrete and Exact Fully Discrete Wavelet Transform on the Sphere. *Constructive Approximation*, 14(4), 493–515.

[59] Garrido, L., Wedeen, V. J., Kwong, K. K., Spencer, U. M., & Kantor, H. L. (1994). Anisotropy of water diffusion in the myocardium of the rat. *Circulation Research*, 74(5), 789–93.

[60] Glenn, G. R., Kuo, L.-W., Chao, Y.-P., Lee, C.-Y., Helpern, J. A., & Jensen, J. H. (2016). Mapping the Orientation of White Matter Fiber Bundles: A Comparative Study of Diffusion Tensor Imaging, Diffusional Kurtosis Imaging, and Diffusion Spectrum Imaging. *American Journal of Neuroradiology*.

[61] Gorodnitsky, I. F. & Rao, B. D. (1997). Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Transactions on Signal Processing*, 45(3), 600–616.

[62] Gudbjartsson, H. & Patz, S. (1995). The Rician Distribution of Noisy MRI data. *Magnetic Resonance in Medicine*, 34(6), 910–914.

[63] Hahn, E. L. (1950). Spin Echoes. *Physical Review*, 80(4), 580–594.

[64] Henkelman, R. M., Stanisz, G. J., Kim, J. K., & Bronskill, M. J. (1994). Anisotropy of NMR properties of tissues. *Magnetic Resonance in Medicine*, 32(5), 592–601.

[65] Henriques, R. N., Correia, M. M., Nunes, R. G., & Ferreira, H. A. (2015). Exploring the 3D geometry of the diffusion kurtosis tensor—Impact on the development of robust tractography procedures and novel biomarkers. *NeuroImage*, 111, 85–99.

[66] Hosey, T., Williams, G., & Ansorge, R. (2005). Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6), 1480–1489.

[67] Jansons, K. M. & Alexander, D. C. (2003). : (pp. 672–683). Berlin, Heidelberg: Springer Berlin Heidelberg.

[68] Jensen, J. H. & Helpern, J. A. (2010). MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR in Biomedicine*, 23(7), 698–710.

[69] Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., & Kaczynski, K. (2005). Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53(6), 1432–1440.

[70] Jian, B. & Vemuri, B. C. (2007). A Unified Computational Framework for Deconvolution to Reconstruct Multiple Fibers From Diffusion Weighted MRI. *IEEE Transactions on Medical Imaging*, 26(11), 1464–1471.

[71] Jones, D. K. & Cercignani, M. (2010). Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR in Biomedicine*, 23(7), 803–820.

[72] Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, 73, 239–254.

[73] Kumar, A., Welti, D., & Ernst, R. R. (1975). NMR Fourier Zeugmatography. *Journal of Magnetic Resonance*, 18, 69–83.

[74] Landman, B. A., Bogovic, J. A., Wan, H., ElShahaby, F. E. Z., Bazin, P.-L., & Prince, J. L. (2012). Resolution of crossing fibers with constrained compressed sensing using diffusion tensor MRI. *NeuroImage*, 59(3), 2175–2186.

[75] Lauterbur, P. C. (1973). Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature*, 242, 190–191.

[76] Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., & Laval-Jeantet, M. (1986). MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2), 401–407.

[77] Liu, C., Bammer, R., Acar, B., & Moseley, M. E. (2004). Characterizing non-gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51(5), 924–937.

[78] Lu, H., Jensen, J. H., Ramani, A., & Helpern, J. A. (2006). Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, 19(2), 236–247.

[79] Lucy, L. B. (1974). An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79, 745.

[80] Lustig, M., Donoho, D. L., & Pauly, J. M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6), 1182–1195.

[81] Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed Sensing MRI. *IEEE Signal Processing Magazine*, 25(2), 72–82.

[82] Malcolm, J. G., Michailovich, O., Bouix, S., Westin, C.-F., Shenton, M. E., & Rathi, Y. (2010). A filtered approach to neural tractography using the Watson directional function. *Medical Image Analysis*, 14(1), 58–69.

[83] Mani, M., Jacob, M., Guidon, A., Magnotta, V., & Zhong, J. (2015). Acceleration of high angular and spatial resolution diffusion imaging using compressed sensing with multichannel spiral data. *Magnetic Resonance in Medicine*, 73(1), 126–138.

[84] Mansfield, P. (1977). Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics*, 10, L55–L58.

[85] Mansfield, P. & Grannell, P. K. (1973). NMR 'diffraction' in solids? *Journal of Physics C: Solid State Physics*, 6(22), L422.

[86] Melie-García, L., Canales-Rodríguez, E. J., Alemán-Gómez, Y., Lin, C.-P., Iturria-Medina, Y., & Valdés-Hernández, P. A. (2008). A Bayesian framework to identify principal intravoxel diffusion profiles based on diffusion-weighted MR imaging. *NeuroImage*, 42(2), 750–770.

[87] Menzel, M. I., Tan, E. T., Khare, K., Sperl, J. I., King, K. F., Tao, X., Hardy, C. J., & Marinelli, L. (2011). Accelerated diffusion spectrum imaging in the

human brain using compressed sensing. *Magnetic Resonance in Medicine*, 66(5), 1226–1233.

[88] Merlet, S., Caruyer, E., & Deriche, R. (2012). : (pp. 10–17). Berlin, Heidelberg: Springer Berlin Heidelberg.

[89] Merlet, S. L. & Deriche, R. (2013). Continuous diffusion signal, EAP and ODF estimation via Compressive Sensing in diffusion MRI. *Medical Image Analysis*, 17(5), 556 – 572.

[90] Michailovich, O. & Rathi, Y. (2010). On Approximation of Orientation Distributions by Means of Spherical Ridgelets. *IEEE Transactions on Image Processing*, 19(2), 461–477.

[91] Michailovich, O., Rathi, Y., & Dolui, S. (2011). Spatially Regularized Compressed Sensing for High Angular Resolution Diffusion Imaging. *IEEE Transactions on Medical Imaging*, 30(5), 1100–1115.

[92] Moseley, M. E., Cohen, Y., Mintorovitch, J., Chileuitt, L., Shimizu, H., Kucharczyk, J., Wendland, M. F., & Weinstein, P. R. (1990). Early detection of regional cerebral ischemia in cats: Comparison of diffusion- and T2-weighted MRI and spectroscopy. *Magnetic Resonance in Medicine*, 14(2), 330–346.

[93] Mulkern, R. V., Gudbjartsson, H., Westin, C.-F., Zengingonul, H. P., Gartner, W., Guttmann, C. R. G., Robertson, R. L., Kyriakos, W., Schwartz, R., Holtzman, D., Jolesz, F. A., & Maier, S. E. (1999). Multi-component apparent diffusion coefficients in human brain. *NMR in Biomedicine*, 12(1), 51–62.

[94] Niendorf, T., Dijkhuizen, R. M., Norris, D. G., Campagne, M. v. L., & Nicolay, K. (1996). Biexponential diffusion attenuation in various states of brain tissue: Implications for diffusion-weighted imaging. *Magnetic Resonance in Medicine*, 36(6), 847–857.

[95] Onsager, L. (1931). Reciprocal Relations in Irreversible Processes. I. *Physical Review*, 37(4), 405–426.

[96] Özarslan, E., Koay, C. G., Shepherd, T. M., Blackband, S. J., & Basser, P. J. (2009). Simple harmonic oscillator based reconstruction and estimation for three-

dimensional q-space MRI. In *Proceedings of the 17th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM)* (pp. 1396). Honolulu (HI, USA).

[97] Özarslan, E. & Mareci, T. H. (2003). Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50(5), 955–965.

[98] Özarslan, E., Shepherd, T. M., Vemuri, B. C., Blackband, S. J., & Mareci, T. H. (2006). Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, 31(3), 1086–1103.

[99] Paquette, M., Merlet, S., Gilbert, G., Deriche, R., & Descoteaux, M. (2015). Comparison of sampling strategies and sparsifying transforms to improve compressed sensing diffusion spectrum imaging. *Magnetic Resonance in Medicine*, 73(1), 401–416.

[100] Parker, G. D., Marshall, D. F., Rosin, P. L., Drage, N. A., Richmond, S., & Jones, D. K. (2013). A pitfall in the reconstruction of fibre ODFs using spherical deconvolution of diffusion MRI data. *NeuroImage*, 65, 433–448.

[101] Pasternak, O., Assaf, Y., Intrator, N., & Sochen, N. (2008). Variational multiple-tensor fitting of fiber-ambiguous diffusion-weighted magnetic resonance imaging voxels. *Magnetic Resonance Imaging*, 26(8), 1133–1144.

[102] Purcell, E. M., Torrey, H. C., & Pound, R. V. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review*, 69(1-2), 37–38.

[103] Rathi, Y., Michailovich, O. V., Setsompop, K., Bouix, S., Shenton, M. E., & Westin, C.-F. (2011). Sparse Multi-shell Diffusion Imaging. In *Proceedings of the 14th International Conference on Medical Image Computing and Computer-Assisted Intervention - Volume Part II*, MICCAI'11 (pp. 58–65). Toronto, Canada: Springer-Verlag.

[104] Richardson, W. H. (1972). Bayesian-Based Iterative Method of Image Restoration. *Journal of the Optical Society of America (OSA)*, 62(1), 55–59.

[105] Saha, S., de Hoog, F., Nesterets, Y. I., Rana, R., Tahtali, M., & Gureyev, T. E. (2015). Sparse Bayesian Learning for EEG Source Localization. *arXiv, Quantitative Biology, Quantitative Method*, (pp. 1–18).

[106] Sotiropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D. A., Yacoub, E., Lenglet, C., Essen, D. C. V., Ugurbil, K., & Behrens, T. E. J. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*, 80, 125–143. Mapping the Connectome.

[107] Stejskal, E. O. & Tanner, J. E. (1965). Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *The Journal of Chemical Physics*, 42(1), 288–292.

[108] Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society, Series B*, 58, 267–288.

[109] Tipping, M. E. (2001). Sparse Bayesian Learning and the Relevance Vector Machine. *The Journal of Machine Learning Research*, 1, 211–244.

[110] Toga, A. W., Clark, K. A., Thompson, P. M., Shattuck, D. W., & Horn, J. D. V. (2012). Mapping the Human Connectome. *Neurosurgery*, 71(1), 1–5.

[111] Torrey, H. C. (1956). Bloch Equations with Diffusion Terms. *Physical Review*, 104(3), 563–565.

[112] Tournier, J.-D., Calamante, F., & Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4), 1459–1472.

[113] Tournier, J.-D., Yeh, C.-H., Calamante, F., Cho, K.-H., Connelly, A., & Lin, C.-P. (2008). Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *NeuroImage*, 42(2), 617 – 625.

[114] Tournier, J.-D. D., Calamante, F., Gadian, D. G., & Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3), 1176–1185.

[115] Tristán-Vega, A. & Westin, C.-F. (2011). : (pp. 182–190). Berlin, Heidelberg: Springer Berlin Heidelberg.

[116] Tristán-Vega, A., Westin, C.-F., & Aja-Fernández, S. (2009). Estimation of fiber Orientation Probability Density Functions in High Angular Resolution Diffusion Imaging. *NeuroImage*, 47(2), 638–650.

[117] Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6), 1358–1372.

[118] Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., & Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4), 577–582.

[119] Tuch, D. S., Reese, T. G., Wiegell, M. R., & Wedeen, V. J. (2003). Diffusion MRI of Complex Neural Architecture. *Neuron*, 40(5), 885–895.

[120] Tuch, D. S., Weisskoff, R. M., Belliveau, J. W., & Wedeen, V. J. (1998). High Angular Resolution Diffusion Imaging of the Human Brain. In *Proceedings of the 7th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM)* Philadelphia (PA, USA).

[121] Wan, J., Zhang, Z., Rao, B. D., Fang, S., Yan, J., Saykin, A. J., & Shen, L. (2014). Identifying the Neuroanatomical Basis of Cognitive Impairment in Alzheimer's Disease by Correlation- and Nonlinearity-Aware Sparse Bayesian Learning. *IEEE Transactions on Medical Imaging*, 33(7), 1475–1487.

[122] Warach, S. J., Li, W., Ronthal, M., & Edelman, R. R. (1992). Acute cerebral ischemia: Evaluation with dynamic contrast-enhanced MR imaging and MR angiography. *Radiology*, 182(1), 41–47.

[123] Wedeen, V. J., Wang, R. P., Schmahmann, J. D., Benner, T., Tseng, W.-Y. I., Dai, G., Pandya, D. N., Hagmann, P., D'Arceuil, H. E., & de Crespigny, A. J. (2008). Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage*, 41(4), 1267–1277.

[124] Wesbey, G. E., Moseley, M. E., & Ehman, R. L. (November-December 1984). Translational Molecular Self-Diffusion in Magnetic Resonance Imaging: I. Effects on Observed Spin-Spin Relaxation. *Investigative Radiology*, 19(6), 484–490.

[125] Wipf, D. P. & Rao, B. D. (2004). Sparse Bayesian learning for basis selection. *Signal Processing, IEEE Transactions on*, 52(8), 2153–2164.

[126] Wu, Y.-C. & Alexander, A. L. (2007). Hybrid diffusion imaging. *NeuroImage*, 36(3), 617–629.

[127] Ye, W., Vemuri, B. C., & Entezari, A. (2012). An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 940–943).

[128] Yeh, C.-H., Tournier, J.-D., Cho, K.-H., Lin, C.-P., Calamante, F., & Connelly, A. (2010). The effect of finite diffusion gradient pulse duration on fibre orientation estimation in diffusion MRI. *NeuroImage*, 51(2), 743–751.

[129] Zhang, H., Hubbard, P. L., Parker, G. J. M., & Alexander, D. C. (2011). Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *NeuroImage*, 56(3), 1301–1315.

[130] Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4), 1000–1016.

[131] Zhang, Z. & Rao, B. D. (2011). Sparse Signal Recovery With Temporally Correlated Source Vectors Using Sparse Bayesian Learning. *IEEE Journal of Selected Topics in Signal Processing*, 5(5), 912–926.

Abbreviations

AD	Alzheimer's Disease
ADC	Apparent Diffusion Coefficient
cc	corpus callosum
CHARMED	Composite Hindered and Restricted Model of Diffusion
CS	Compressed Sensing
CSA	Constant Solid Angle
CSD	Constrained Spherical Deconvolution
CSF	cerebrospinal fluid
CSHD	Constrained Spherical Harmonic Deconvolution (also referred to as Constrained Spherical Deconvolution)
CT	Computed Tomography
DKI	Diffusional Kurtosis Imaging
dLS	Damped Least-Squares
dMRI	Diffusion Magnetic Resonance Imaging
DOT	Diffusion Orientation Transform
dRL	Damped Richardson-Lucy
DSI	Diffusion Spectrum Imaging
DT	diffusion tensor
DTI	Diffusion Tensor Imaging
DWI	Diffusion-Weighted Imaging
EAP	Ensemble Average (Diffusion) Propagator
EEG	Electroencephalography
EPI	Echo Planar Imaging
EQBI	Exact q-Ball Imaging
FA	Fractional Anisotropy
FID	Free Induction Decay

fMRI	Functional Magnetic Resonance Imaging
fODF	fiber Orientation Distribution Function (also referred to as fiber Orientation Density Function)
FT	Fourier Transform
GDT	Generalized Diffusion Tensor
GDTI	Generalized Diffusion Tensor Imaging
GL	Gaussian-Laguerre
GM	gray matter
GMM	Gaussian Mixture Model
HARDI	High Angular Resolution Diffusion Imaging
HIDY	Hybrid Diffusion Imaging
HODT	Higher-Order Diffusion Tensor
HOT	Higher-Order Tensor
ic	internal capsule
IFA	Inter-fiber angle
IVOH	Intra-voxel Orientational Heterogeneity
k-SVD	k-means Singular Value Decomposition
KT	kurtosis tensor
LASSO	Least Absolute Shrinkage and Selection Operator
LMA	LevenbergMarquardt algorithm (also known as the damped least-squares)
LS	Laplace series / Least Squares
MAP	Maximum-A-Posteriori
MCMC	Markov Chain Monte Carlo
MDT	Multiple-Diffusion Tensors model
ML	Maximum Likelihood
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSI	Multi-Shell Imaging
NMR	Nuclear Magnetic Resonance
nnLS	Non-negative Least Squares

NODDI	Neurite Orientation Dispersion and Density Imaging
ODF	Orientation Distribution Function (also referred to as Orientation Density Function)
OPDF	Orientation Probability Density Function
OPDT	Orientation Probability Density Transform
PAS	Persistent Angular Structure
PASMRI	Persistent Angular Structure Magnetic Resonance Imaging
PDF	Probability Density Function
PGSE	Pulsed-Gradient Spin Echo
QBI	q-Ball Imaging
RF	Radio Frequency
RJMCMC	Reversible Jump Markov Chain Monte Carlo
RL	Richardson-Lucy
RMS	Root Mean Square
RUMBA-SD	Robust and Unbiased Model-Based Spherical Deconvolution
SBL	Sparse Bayesian Learning
SD	Spherical Deconvolution
SE	Spin-Echo
SH	Spherical Harmonics
SHD	Spherical Harmonic Deconvolution
SHORE	Simple Harmonic Oscillator based Reconstruction and Estimation
SHT	Spherical Harmonic Transform
slf	superior longitudinal fasciculus
SNR	Signal-to-Noise Ratio
SPF	Spherical Polar Fourier
SR	Spherical Ridgelets
super-CSD	Super-resolved Constrained Spherical Deconvolution
SVD	Singular Value Decomposition
SW	Spherical Wavelets
TE	Echo Time
TR	Repetition Time

TV Total Variation
WM white matter

Notation

$\mathbf{A}$	problem matrix / compartment anisotropy characterization diagonal matrix
b	b-value
B_0	static magnetic field intensity
Δ	time between diffusion-encoding gradients (diffusion time)
δ	duration of diffusion-encoding gradient / Dirac delta function
$\mathbf{D}$	diffusion tensor
$\mathbf{E}$	spin echo signal attenuation
$\mathcal{F}$	Fourier transform
$\mathcal{F}^{-1}$	inverse Fourier transform
f_i	relative volume fraction of the i-th fiber population ($f_i \geq 0$)
$F(\theta, \phi)$	fiber orientation distribution function
$\mathbf{f}$	fiber orientation distribution function
G	magnetic field gradient intensity
$\mathbf{g}$	diffusion gradient wave vector
γ	gyromagnetic ratio
$\mathbf{H}$	spherical surface circulant matrix, kernel or dictionary
I_n	modified Bessel function of the first kind of order n
λ_i	diffusion tensor eigenvalues ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$)
$\mathbf{M}$	spin magnetization vector
m	particle mass / complex transverse magnetization vector
P_l^m	associated Legendre polynomial of degree l and order m
$P(r\hat{\mathbf{u}})$	diffusion propagator
$P(\mathbf{R}, \tau)$	ensemble-average propagator
ψ	orientation distribution function
$\mathbf{q}$	spin displacement wave vector

$\mathbf{R}$	relative spin displacement / diffusion tensor eigenvector matrix
$\mathbf{r}$	final spin position
$\mathbf{r}'$	initial spin position
$R(\theta)$	fiber response function
ρ	bulk relaxation rate
$\mathbf{R}_j$	rotation matrix rotatin a unit vector initially oriented along the X-axis towards the j-th fiber orientation (θ_j, ϕ_j)
$\mathbf{S}$	spin echo signal
S_0	spin echo signal in the absence of diffusion gradient
T_1	spin-lattice relaxation time (sometimes referred to as T1)
T_2	spin-spin relaxation time (sometimes referred to as T2)
τ	diffusion time
$\hat{\mathbf{u}}$	unit normal vector
φ	spin phase
$Y_l^m(\theta, \phi)$	spherical harmonic basis function of degree l and order m in (θ, ϕ) angular direction

Index

S

T

W