Reconstruction of Complex White Matter Architecture from Diffusion MRI

Shahrum Nedjati-Gilani

A dissertation submitted in partial fulfillment of the requirements for the degree of Engineering Doctorate of University College London.

Department of Computer Science
University College London

2008
I, Shahrum Nedjati-Gilani, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

Diffusion MRI provides insight into the microstructural architecture of tissue by observing the restricted and hindered displacement of water molecules undergoing Brownian motion. Diffusion Tensor MRI (DT-MRI) is the most common diffusion MRI technique and is often used for reconstructing fibre population information. However, DT-MRI is only capable of recovering a single fibre orientation in each voxel. Other reconstruction algorithms can overcome this problem; however, these more complex algorithms have their own shortcomings. For example, they can not differentiate between fanning and bending structures, or correctly assign the spatial arrangement in voxels containing more than one fibre population.

The aim of my work is to find and evaluate methods capable of identifying and distinguishing complex fibrous microstructure accurately in each voxel of a 3D diffusion MRI acquisition.

This thesis proposes a method for generating a map of the number of fibre populations in each voxel of a 3D diffusion MRI acquisition and compares results from this new method with those obtained from an existing algorithm to assess the reliability of the generated fibre population map. A new regularized super-resolution method is presented for finding accurate fibre orientations and volume fractions of fibre populations on a sub-voxel scale. The method can be used to distinguish between various fibre configurations such as fanning and bending, ameliorate partial volume effects, The method is demonstrated on synthetic and human brain data. Finally, a new model is presented for generating fanning and bending structures on a sub-voxel scale, which provides measures for quantifying these structures. This model is used to create synthetic bending and fanning data, and explore the suitability of the model to distinguish between voxels with bending or fanning structures on human brain data.
Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Daniel Alexander. I cannot thank him enough for his advice, motivation, wisdom and support throughout the last few years.

I am grateful to the EPSRC and Philips Medical Systems for funding my studies. Thanks to Frank Hoogenraad and Ronald Holthuizen for their assistance during my time in Best. I am in particular grateful to Dr Elizabeth Moore for her help throughout for all Philips-related matters.

Many thanks to Geoff Parker, who has kindly helped provide the data used for the experiments in the thesis, and for all of his suggestions and guidance.

I would like to thank all of my friends at UCL who have made my being here so enjoyable. In particular, I would like to thank Yu Bai, Ai Wern Chung, Phil Cook, Jenny Edgar, Hubert Fonteijn, Matt Hall, Gemma Morgan, Laura Panagiotaki, Kiran Seunarine and Tony Shepherd. Whether going through drafts of papers or this thesis with me, keeping me company during the long lab hours, discussing the merits of pandas, squids and squirrels, or demonstrating their vocal abilities in karaoke, they have been instrumental in making my time at UCL constructive and a pleasure, and for that, I am incredibly grateful.

I would like to thank my examiners, Dr. Chris Clark and Prof. Daniel Rueckert for their time, and constructive feedback.

Special thanks go to my brother Shervin who has had to put up with me going on about all this ‘brain stuff’ for the last year, and to Gemma Morgan and Adrian Hadad, who ensured that I occasionally take time out and enjoy the non-academic world.

Finally, to my parents, Roya and Mehdi, who have guided me, encouraged me, loved me and inspired me. Thank you. I love you more than words could possibly express.
تقديم به پروین و رضا شهیدی
Contents

1 Introduction .......................................................... 11
   1.1 Problem Statement ............................................. 14
   1.2 Research Objective .......................................... 15
   1.3 Contributions ................................................. 15
   1.4 Report Structure ............................................. 16

2 An Introduction to MRI and Diffusion ......................... 17
   2.1 Magnetic Resonance: An Overview .......................... 17
       2.1.1 Properties of atomic nuclei and nuclear magnetism .... 17
       2.1.2 Precession, Resonance .................................. 18
       2.1.3 Pulse Sequences ......................................... 20
   2.2 Diffusion-Weighted MRI ...................................... 23

3 Fibre Population Reconstruction Techniques ................. 25
   3.1 Model-Based Methods ....................................... 25
       3.1.1 The Diffusion Tensor ................................... 25
       3.1.2 Multi-compartmental models ........................... 27
   3.2 Model-Independent Methods ................................ 30
       3.2.1 Diffusion Spectrum Imaging ............................ 30
       3.2.2 q-ball ..................................................... 31
       3.2.3 PAS-MRI .................................................. 32
       3.2.4 Spherical Deconvolution ................................. 33
   3.3 Voxel Classification Methods ............................... 34
       3.3.1 Spherical Harmonics .................................... 34
       3.3.2 K-means Clustering ..................................... 34
3.4 Conclusion ................................................................. 35

4 Mapping the number of fibre-orientations per voxel 37
4.1 Hierarchical Model Generation and Selection ................................. 37
  4.1.1 The Model Hierarchy ............................................. 37
  4.1.2 Model Selection .................................................. 39
4.2 Model Generation and Selection .............................................. 39
4.3 Conclusions and Discussion ............................................... 42

5 Regularized Super-Resolution for Diffusion MRI 46
5.1 Motivation ............................................................... 46
5.2 Super-Resolution Methods ............................................... 48
  5.2.1 Multiple-Image Super-Resolution .................................. 48
  5.2.2 Single-Image Super-Resolution .................................... 49
5.3 Regularization Methods .................................................. 50
  5.3.1 Conclusion ......................................................... 52
5.4 Problem Outline ....................................................... 52
  5.4.1 Model Fitting and Initialization .................................. 53
  5.4.2 Smoothing ......................................................... 53
  5.4.3 Optimization ..................................................... 55
5.5 Experiments and Results: Synthetic Data .................................. 57
  5.5.1 Experiment 1: Validation on synthetic data - fanning ............. 57
  5.5.2 Experiment 2: Validation on synthetic data - bending ............. 60
  5.5.3 Experiment 3: Validation on synthetic partial volume ............. 62
5.6 Experiments and Results: Brain Data .................................... 62
  5.6.1 Experiment 4: Reconstruction of the corpus callosum ............ 62
  5.6.2 Experiment 5: Reconstruction of the cortico-spinal tract .......... 64
  5.6.3 Experiment 6: Super-Resolution on the corpus callosum and cingulum ..................................................... 67
  5.6.4 Experiment 7: Super-Resolution on the pons ..................... 67
5.7 Conclusions and Discussion .............................................. 69
6 Fanning Sub-Voxel Structures in Diffusion MRI

6.1 Motivation .................................................. 72
6.2 Method .................................................. 73
  6.2.1 Minimization ......................................... 75
  6.2.2 Initialization of parameter values ................. 78
6.3 Experiments and Results ................................. 78
  6.3.1 Experiment 1: Synthetic Data ................. 78
  6.3.2 Experiment 2 ......................................... 79
6.4 Conclusions and Future Work ......................... 83
  6.4.1 Bending Structures and Classification of Voxels . 85

7 Industrial Relevance ........................................ 86

7.1 Introduction ................................................. 86
7.2 Mapping the number of fibre-orientations per voxel ......... 86
7.3 Regularized Super-Resolution ............................ 87
  7.3.1 Model Fitting and Initialization: Extending to two fibre-
       populations ........................................... 88
  7.3.2 Smoothing ............................................. 88
  7.3.3 Optimization ......................................... 90
  7.3.4 Experiments and Results ......................... 91
7.4 Conclusions and Discussion ............................. 92

8 Conclusions .................................................. 95

A Initialization of Fanning Model Parameters .................. 97

Bibliography ..................................................... 102
List of Figures

1.1 Confocal microscopy images of fibre microstructure in different types of brain tissue, and their corresponding probability density functions. . . 12

1.2 Examples of complex fibre microstructure . . . . . . . . . . . . . . . 15

2.1 Alignment of protons with and without the influence of an external magnetic field. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18

2.2 Precessing nuclei subjected to an external field, and the resulting net magnetization vector. . . . . . . . . . . . . . . . . . . . . . . . . . . 19

2.3 The effect of an RF pulse applied in the transverse plane . . . . . . 19

2.4 The Free Induction Decay. . . . . . . . . . . . . . . . . . . . . . . . 21

2.5 Spin-echo pulse sequence. . . . . . . . . . . . . . . . . . . . . . . . 22

2.6 The phasing and dephasing steps in the spin echo sequence. . . . . 22

2.7 Pulse-gradient spin-echo sequence. . . . . . . . . . . . . . . . . . . 23

3.1 Isotropic, prolate and oblate tensors . . . . . . . . . . . . . . . . . . 27

4.1 Histograms showing how often the global minimum was found out of 1000 attempts for the 1000 selected voxels for $N = 0, 1, 2, 3$. . . . . 40

4.2 Plots of the probability of finding the global minimum against the number of runs for the $N = 0, 1, 2$ and 3 models. . . . . . . . . . 41

4.3 Estimated number of fibre orientations per voxel. . . . . . . . . . . . 43

4.4 Estimated number of fibre orientations per voxel using spherical harmonics. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

5.1 Reconstructions of various complex fibre configurations. . . . . . . 47

5.2 Overview of the regularized super-resolution procedure. . . . . . . 54
5.3 Illustration of a sub-voxel with two voxels of its 6-neighbourhood, and associated fibre populations (one fibre-population version) ........................................... 56
5.4 Comparison of the performance of the smoothing functions on synthetic data with a fanning structure .......................................................... 58
5.5 Comparison of reconstructions from synthetic data with a fanning structure .......................................................... 59
5.6 Comparison of the performance of the smoothing functions on synthetic data with a bending structure .......................................................... 60
5.7 Comparison of reconstructions from synthetic data with a bending structure .......................................................... 61
5.8 Reconstruction of a synthetic dataset with partial volume effects .......................................................... 63
5.9 Reconstruction of the corpus callosum from a downsampled dataset using regularized super-resolution .................. 65
5.10 Reconstruction of the corticospinal tract from a downsampled dataset using regularized super-resolution .................. 66
5.11 Super-resolution on the corpus callosum and cingulum .......................................................... 68
5.12 Super-resolution on the pons .......................................................... 71
6.1 Overview of the model used to create fanning structures .......................................................... 74
6.2 Examples of sub-voxel grids and how the orientations are affected by the parameters used to describe the fanning model .......................................................... 76
6.3 Reconstruction of fanning structure in the corticospinal tract .......................................................... 84
7.1 Illustration of a sub-voxel with two voxels of its 6-neighbourhood, and associated fibre populations (two fibre-population version) .......................................................... 89
7.2 Comparison of reconstruction using regularized super-resolution from differently subsampled datasets .................. 93
A.1 Creating an estimate for $\psi$ and $u$ .................. 99
Chapter 1

Introduction

Diffusion Magnetic Resonance Imaging (MRI) provides an insight into the microstructural architecture of tissue by observing the restricted and hindered displacement of particles, usually water molecules, undergoing Brownian motion [13]. By looking at the probability density function \( p \) of displacements over a fixed period of time \( t \), inferences can be made about the tissue microstructure. In some regions of the brain, the barriers to diffusion have no preferred orientation, resulting in isotropic diffusion. An example of a region of the brain with isotropic diffusion is the ventricles, which contain cerebrospinal fluid (CSF). As there are few barriers to water-molecule mobility, water displacement is equally likely in all directions. Grey matter consists of dense tissue containing many barriers to water mobility, such as cell walls and membranes. However, the barriers in grey matter often have no preferred orientation and so hinder the water displacement equally in all directions, again resulting in isotropic diffusion.

In other regions of the brain, water molecules are constrained by organised fibrous structure, causing diffusion anisotropy.

White matter consists of myelinated axons connecting different areas of the brain. The axons are extensions of neurons, conducting electrical impulses from the neuron cell body, and are typically in the order of \( 10^{-6} \)m in diameter. Axons tend to form fibre bundles together. As a consequence of this organized structure, water molecules in white matter move on average further along fibres than across them, since mobility across the fibres will be hindered. This results in the anisotropic diffusion of water molecules, and \( p \) is thus elongated in the direction of the fibre. If we can determine the orientation of \( p \), we can infer the orientation of the fibres. Figure 1.1 shows different types of microstructure and their corresponding \( p \).
Figure 1.1: Confocal microscopy images of fibre microstructure (from [11, 12]) in different types of brain tissue (left column), and their corresponding probability density functions (right column). (a) Isotropic grey matter (putamen). (b) White matter with one dominant fibre-orientation, belonging to the anterior thalamic peduncle. (c) Two white matter fibre populations; the bundle of fibres in the middle represents fibres of the frontopontine tract. The horizontally cut fibres belong to the anterior thalamic peduncle.
Diffusion MRI measures the water diffusion over timescales in the order of 0.01s. The average displacement of water molecules over this time is in the order of a few micrometers (µm).

Applications of Diffusion MRI

One of the earliest and most successful uses of Diffusion MRI was in the detection of acute brain ischemia. Water diffusion drops at a very early stage of an ischemic event [73]. Diffusion-weighted images are very effective at detecting ischemic areas, and allow the potential for suitable treatment to be offered within the first few hours of the event, where brain tissue may be salvageable [73, 118, 103].

Diffusion MRI has also found many other uses, especially in the form of Diffusion Tensor MRI (DT-MRI) [13, 14]. DT-MRI offers quantitative information in evaluating the structural damage occurring in multiple sclerosis lesions and normal appearing white matter (NAWM). In [37], Dong et al show that lesions have higher mean diffusivity and lower anisotropy than NAWM, and higher mean diffusivity and lower anisotropy values in patient NAWM than white matter from normal control subject white matter. This suggests that MS lesions accompany a more generalised change in white matter microstructure [61, 119].

Generally, in conditions where the myelin or fibre structure is disrupted, the anisotropy of the structure is reduced. Conditions in which water content is altered are also detectable by the separation of mean diffusivity indices, such as the trace of the diffusion tensor. Diffusion MRI and DT-MRI have also been used to investigate other conditions such as schizophrenia [61, 18], Alzheimer’s disease [61, 48, 49], and epilepsy [37], and brain maturation in children, newborns and premature babies [61, 124, 53]. Although mainly used to examine white matter, diffusion MRI has been used to show that in patients with mild cognitive impairment, there is an increase in diffusivity in the temporal lobe grey matter [98]. Diffusion MRI has also been used to investigate areas of anatomy other than the brain; examples include the breast [65], uterine fibroids [55], and the kidneys and liver [75, 122].

An important application of Diffusion MRI is tractography. When the local fibre orientation of each voxel is found using the diffusion tensor (or any other means), one can examine the connectivity of the brain by tracing paths to and from areas that are
connected by white matter fibres. Many studies of tractography have been based on the use of the primary eigenvector of the diffusion tensor as an estimate of the fibre orientation, although other multi-fibre reconstruction methods can also be used (see Section 3.2). Tractography methods can be split into two broad classes, deterministic and probabilistic. Deterministic methods estimate a single maximum-likelihood fibre path by following fibre orientation estimates from voxel to voxel. Probabilistic tractography methods provide a quantitative measure of connectivity between two points in the brain. Applications of tractography include risk assessment of surgical procedures and monitoring development of the brain. Reviews of various tractography methods are provided in [72, 70].

DT-MRI is the most common diffusion MRI technique [14] and is often used for mapping fibre orientations for each voxel in a 3D diffusion MRI acquisition. However, DT-MRI assumes that $p$ follows a Gaussian distribution. This is not necessarily true; for example [42, 4] show that in regions known to contain crossing fibres, $p$ does not follow a Gaussian distribution. In such instances, the resulting fibre orientation estimate from DT-MRI will not be accurate. Other reconstruction algorithms exist that overcome this problem; however, these more complex algorithms are generally less successful in correctly identifying the orientation of single fibres and are more computationally demanding than DT-MRI. Also, existing reconstruction methods identify multiple fibre populations, but cannot correctly assign the spatial arrangement of fibre structure in such voxels, and thus have difficulties distinguishing various fibre configurations (examples of which are shown in Figure 1.2).

1.1 Problem Statement

The problem we address is to recover detailed microstructural fibre architecture in each voxel of diffusion MRI acquisitions. We can use this information to reconstruct white matter tracts more accurately, and reduce false positive connections. This allows for more accurate brain connectivity mapping. We could also use this information as a form of biomarker; if we can characterize white matter structure more accurately, we will be able to assess white matter integrity.
1.2 Research Objective

The objective of this project is to extract more accurate information about the spatial arrangement of fibre populations in each voxel of a 3D diffusion MRI acquisition.

Examples of complex microstructure (illustrated in Figure 1.2) of particular interest include:

- Crossing fibre populations
- Multiple fibre populations independently occupying segments of the same voxel (known as the partial volume effect)
- Fanning structures
- Bending structures

1.3 Contributions

The key contributions of this project are:

- a method that creates a map of the number of fibre populations in each voxel of a 3D diffusion MRI acquisition.
1.4. Report Structure

Chapter 2 contains background on MRI and diffusion MRI, and how we obtain the measurements we will use to extract fibre population information. Chapter 3 is a review of existing methods that find information about the fibre populations in a 3D diffusion MRI acquisition, such as the number of fibre populations, fibre orientations and their respective volume fractions. Chapter 4 describes a new method for mapping the number of fibre orientations per voxel, with experiments and results. In Chapter 5, we introduce a method for finding accurate fibre orientations and volume fractions of fibre populations on a sub-voxel scale using a new regularized super-resolution algorithm. Chapter 6 describes a model that describes fanning and bending structures on a sub-voxel scale, and shows preliminary results fitting the model to determine the degree of fanning in different structures. Chapter 7 provides an overview of the industrial relevance of this research. We discuss conclusions and future work in the final chapter.
Chapter 2

An Introduction to MRI and Diffusion

In this chapter, we discuss the principles of magnetic resonance, and provide background information on diffusion-weighted MRI, and how we obtain the measurements used to infer white matter microstructure.

2.1 Magnetic Resonance: An Overview

2.1.1 Properties of atomic nuclei and nuclear magnetism

The principles of magnetic resonance imaging depend on the spinning motion of specific nuclei present in biological tissues. Atomic nuclei consist of positively charged protons and chargeless neutrons. These particles have the intrinsic quantum mechanical property of spin. Certain nuclei have non-zero spin and acquire a magnetic moment $\mu$ characterizing the magnetic field surrounding the nucleus. For nuclei with spin $I$, we have

$$\mu = \gamma I,$$

where $\gamma$ is the gyromagnetic ratio. The value of $\gamma$ varies for different nuclei.

Typically, we consider hydrogen nuclei in clinical MRI for two reasons: hydrogen nuclei are abundant in the human body in the form of water and in other molecules such as fat. Also, the value of $\gamma$ for hydrogen is large compared to that of other nuclei, giving it a large magnetic moment [121]. Hydrogen nuclei consist of a single proton.

In the absence of an external magnetic field, the orientation of the magnetic moment is random. However, once in a magnetic field, the magnetic moments align themselves according to the magnetic field, shown in Figure 2.1. Some of the protons (or ‘spins’) will align themselves with the external field, and some will align themselves
2.1. Magnetic Resonance: An Overview

Figure 2.1: Left: Alignment of protons without the influence of an external magnetic field. Right: Alignment under the influence of external field $B_0$.

against the field. Alignment with the external field is a lower energy state than against. This means that although protons will continually move from one state to another, at any given point in time, there will be more protons aligned with the field than against. For a magnetic field $B_0$, the number of excess protons can be calculated by:

$$N_{excess} = \frac{N_{total} \gamma h B_0}{2 k_B T}$$

(2.2)

where $N_{total}$ is the total number of protons, $h$ is Planck’s constant, $k_B$ is the Boltzmann constant and $T$ is the temperature. From Equation 2.2, for every 2 million protons, there are 9 more protons aligned with the field than against at body temperature for field strength $B_0 = 1.5T$. The vector sum of the magnetic moments is the net magnetization vector $M_0$, and is aligned with $B_0$, shown in Figure 2.2. $M_0$ is a small quantity in the order of $10^{-6}$T, and is hence not directly measurable when aligned with $B_0$. Conventionally, we assume that $B_0$ is along the $z$-axis.

2.1.2 Precession, Resonance

Protons subjected to an external magnetic field $B_0$ will spin around the axis of $B_0$. This is referred to as precession, and is analogous to a spinning top wobbling about its axis. The frequency of precession $\omega_0$ is given by the Larmor equation:

$$\omega_0 = \gamma B_0.$$  

(2.3)

$\omega_0$ is known as the Larmor or resonance frequency. For values of $B_0$ typically used in MRI scanners (in the order of 1T) and for the value of $\gamma$ for hydrogen (42.57MHz/T), $\omega_0$ lies in the RF range of the electromagnetic spectrum.
2.1. Magnetic Resonance: An Overview

Figure 2.2: Precessing nuclei subjected to an external field $B_0$, and the resulting net magnetization vector, $M_0$. The circular paths show the path of precession.

Figure 2.3: When an RF pulse is applied in the transverse plane, perpendicular to $B_0$ (left), $M_0$ spirals from the $z$-axis towards the transverse plane (right).

To generate a measurable MR signal, we apply an electromagnetic pulse at the precessional frequency with a coil positioned such that the axis of the coil is perpendicular to $B_0$. From a quantum physics perspective, the protons absorb the energy from the pulse, and move from the lower energy state to the higher energy state. Also, the magnetic moments move in phase with one another in the precession circle. In the classical frame, $M_0$ spirals down from the $z$-axis to the transverse plane until the RF pulse is switched off. This is shown in Figure 2.3. The angle from which $M_0$ moves from the $z$-axis is known as the flip angle. We can calculate the flip angle $\alpha = \gamma B_1 t$, where $B_1$ is the strength of the RF magnetic field and $t$ is the pulse duration. Choosing values of $B_1$ and $t$ such that $\alpha = 90^\circ$ results in $M_0$ ending up in the transverse plane, and is
known as a 90° pulse. If $\alpha = 180°$, $M_0$ will be aligned with the $-z$-axis.

The precession of $M_0$ at the Larmor frequency in the transverse plane induces a current in a receiver coil, in accordance with Faraday's laws of induction. This creates a measurable signal that is proportional to the magnitude of $M_0$ [19].

Once the RF pulse is turned off:

- RF energy is retransmitted at the precession frequency, as some spins in the high energy state return to the lower energy state.

- $M_0$ gradually returns to the $z$-axis. This is known as $T_1$ relaxation, or longitudinal relaxation. The magnetization along the $z$-axis is $M_z = M_0(1 - e^{-t/T_1})$, where $t$ is the time after the RF pulse has been switched off. $T_1$ determines how quickly $M_0$ realigns with the $z$-axis, and is different for different tissues.

- The excited protons dephase. This is known as $T_2$ relaxation, or transverse relaxation. The dephasing results in the transverse contribution of $M_0$ to tend to zero. The magnetization in the transverse plane, $M_{xy}$, can be calculated by $M_{xy} = M_0e^{-t/T_2}$. $T_2$ determines how quickly $M_{xy}$ tends to zero, and depends on the tissue.

As the magnitude of $M_{xy}$ decreases, the current induced in the receiver coil also decreases. This signal called the free induction decay (FID), and is illustrated in Figure 2.4.

The voxel intensity of a given tissue type depends on the proton density of the tissue; the higher the proton density, the stronger the FID response signal. MR image contrast also depends on $T_1$ and $T_2$ values.

### 2.1.3 Pulse Sequences

A pulse sequence is a collection of defined RF and gradient pulses, signals and intervening periods of recovery. Two concepts used to define a pulse sequence are repetition time, $t_r$, and echo time, $t_e$. The repetition time is the period of the time from the application of one RF pulse to the application of the next RF pulse. $t_r$ determines the amount of $T_1$ relaxation that is allowed to occur between the end of one RF pulse and the next. The echo time is the time between which the RF pulse is applied and the
2.1. Magnetic Resonance: An Overview

Figure 2.4: The Free Induction Decay (FID). The signal envelope is equal to \( M_0 \sin \alpha e^{-t/T_2} \).

response signal is measured. \( t_e \) controls the amount of \( T_2 \) relaxation that is allowed to occur.

2.1.3.1 Spin Echo Sequence

In practice, the signal measured by the receiver coil decays faster than \( T_2 \) would predict. The combination of dephasing due to \( T_2 \) relaxation and inhomogeneities in \( B_0 \) is called \( T_2^{*} \) relaxation. The spin echo sequence, shown in Figure 2.5, is designed to eliminate \( T_2^{*} \) by cancelling dephasing due to field inhomogeneities, leaving just \( T_2 \) relaxation. The following events occur in the sequence, illustrated in Figure 2.6:

1. At \( t = 0 \), immediately after a \( 90^\circ \) RF pulse, \( M_0 \) is in the transverse plane.

2. For \( t = 0 \) to \( t = t_e/2 \), the spins dephase (\( T_2^{*} \) relaxation).

3. At \( t = t_e/2 \), a \( 180^\circ \) RF pulse is applied, which rotates the spins by \( 180^\circ \).

4. After a further \( t_e/2 \), at \( t = t_e \) the spins have rephased, and the signal is measured.

The spins are only rephased providing they are stationary throughout the sequence; if that were not true, the Larmor frequency for an individual spin would vary during
2.1. Magnetic Resonance: An Overview

Figure 2.5: Spin-echo pulse sequence.

Figure 2.6: The phasing and dephasing steps in the spin echo sequence (adapted from [69]).
the sequence. Therefore, the random motion of diffusing water molecules leads to the incomplete recovery of transverse magnetization at $t_e$. This is exploited to measure diffusion of water molecules in MRI.

### 2.2 Diffusion-Weighted MRI

The Diffusion-Weighted MRI pulse sequence is a modified version of the spin-echo sequence in Figure 2.5, known as the pulse-gradient spin-echo (PGSE) sequence [104], shown in Figure 2.7. The PGSE sequence adds two identical gradient pulses, $\Gamma_1$ and $\Gamma_2$, to the spin-echo sequence. The gradient pulses are each of duration $\delta$, and are separated by duration $\Delta$. The first gradient pulse causes the spins to dephase depending on their position, and the second gradient pulse rephases the spins. If there is no change in the position of the spin, $\Gamma_2$ has an equal and opposite effect to $\Gamma_1$. A change in position results in less rephasing of spin, and therefore a larger attenuation of signal.

In more detail, let us consider a spin at position $\mathbf{r}$. A pulse $\Gamma_1$ with constant value $\mathbf{g}$ for duration $\delta$ will offset the spin by $\mathbf{r} \cdot \mathbf{q}$, where $\mathbf{q} = \gamma \delta \mathbf{g}$ is known as the wavenumber. After the $180^\circ$ pulse, the offset will be $-\mathbf{r} \cdot \mathbf{q}$. If after $\Gamma_2$, the spin is at position $\mathbf{r}'$, the overall phase offset of the spin will be $\mathbf{q} \cdot (\mathbf{r}' - \mathbf{r}) = \mathbf{q} \cdot \mathbf{x}$, where $\mathbf{x}$ is the displacement of the spin.

The magnetization of the spin at the spin echo is

$$M_0 \exp (i \mathbf{q} \cdot \mathbf{x}).$$  \hspace{1cm} (2.4)
We sum over all displacements of \( n \) spins to find the signal:

\[
A^*(\mathbf{q}) = n M_0 \int p(\mathbf{x}) \exp(i \mathbf{q} \cdot \mathbf{x}) \, d\mathbf{x}
\]

\[
= A^*(0) \int p(\mathbf{x}) \exp(i \mathbf{q} \cdot \mathbf{x}) \, d\mathbf{x},
\]  \hspace{0.5cm} (2.5)

where \( A^*(0) \) is the signal with no diffusion-weighting gradients. Thus, normalized diffusion MRI measurements \( A(0) = A^*(\mathbf{q})/A^*(0) \) sample from the Fourier Transform of particle displacement density \( p \) at \( \mathbf{q} \).

The derivation of Equation 2.5 assumes that the spins have no displacement during the gradient pulses, which implies \( \delta \ll \Delta \), so diffusion time is approximately equal to \( \Delta \). However, \( \delta \) and \( \Delta \) are often of the same order, which complicates matters. However, if we assume that \( p \) is Gaussian and \( \Gamma_1 \) and \( \Gamma_2 \) are rectangular, we can assume the diffusion time is \( \Delta - \delta/3 \) [96].
Chapter 3

Fibre Population Reconstruction Techniques

In this chapter, we present and discuss various methods that can be used to reconstruct information about fibre populations in each voxel of a 3D diffusion MRI acquisition. The methods in this section are split into two main categories: those that use models for the underlying diffusion process, and those that do not. The model-based section covers Diffusion-Tensor MRI and its limitations, which motivates the development of the other models and algorithms.

3.1 Model-Based Methods

3.1.1 The Diffusion Tensor

Diffusion-Tensor MRI (DT-MRI) is the most common diffusion MRI technique \[14\] used for extracting fibre population information. DT-MRI computes the apparent diffusion tensor based on the assumption that \( p \) is a zero-mean tri-variate Gaussian distribution with covariance proportional to the diffusion tensor \( D \):

\[
p(x) = G(x; D, t) = \frac{1}{\sqrt{(4\pi t)^3 |D|}} \exp \left( -\frac{x^T D^{-1} x}{4t} \right).
\]  

The diffusion tensor, \( D \), is a positive-definite, symmetric and second-order tensor:

\[
D = \begin{pmatrix}
D_{xx} & D_{xy} & D_{xz} \\
D_{xy} & D_{yy} & D_{yz} \\
D_{xz} & D_{yz} & D_{zz}
\end{pmatrix},
\]  

where \( D_{xx}, D_{yy} \) and \( D_{zz} \) are diffusion coefficients along the \( x, y \) and \( z \) axes respectively in the coordinate frame of the magnetic gradients in the acquisition, and \( D_{xy}, D_{xz} \) and
3.1. Model-Based Methods

$D_{yz}$ are correlation coefficients between the axes.

With Gaussian $p$, we have

$$A^*(\mathbf{q}) = A^*(\mathbf{0}) \exp (-t \mathbf{q}^T \mathbf{D} \mathbf{q}), \quad (3.3)$$

where $A^*(\mathbf{q}) = A^*(\mathbf{0}) A(\mathbf{q})$ is the unnormalised measurement at $\mathbf{q}$ and $A^*(\mathbf{0})$ is the signal with no diffusion-weighting gradients. We require at least seven measurements to find the six free parameters of $\mathbf{D}$ along with the value of $A^*(\mathbf{0})$. $L$ measurements are made, comprising of $M$ measurements $A^*(\mathbf{q}_1) \ldots A^*(\mathbf{q}_M)$ with $\mathbf{q} = \mathbf{0}$ and $(L-M) > 6$ measurements $A^*(\mathbf{q}_{M+1}) \ldots A^*(\mathbf{q}_L)$ each with a unique $\mathbf{q}$. Normally, we fit $\mathbf{D}$ to the measurements in the least squares sense [14]. Typically, the value of $|\mathbf{q}_i|$ is constant for all $i > M$, and hence the diffusion-weighting factor $b = t|\mathbf{q}|^2$ is fixed. The gradient directions $\hat{\mathbf{q}}_i (i > M)$ are unique and distributed uniformly over the hemisphere [57]. We refer to this kind of measurement scheme as a spherical acquisition scheme, as all $\mathbf{q}_i$ lie on a sphere in $\mathbf{q}$-space.

One of the main benefits of DT-MRI is its ability to provide rotationally-invariant statistics of the anisotropy, which allows for quantitative comparison of different parts of the brain and of different subjects. The literature contains several such scalar indices that describe the size and shape of the DT. Two of the most common are $\text{Tr}(\mathbf{D})$, which is proportional to the mean squared displacement of particles, and the fractional anisotropy (FA) [13]:

$$FA = \left( \frac{3}{2} \sum_{i=1}^{3} \left( \frac{\lambda_i - \frac{1}{3} \text{Tr}(\mathbf{D})}{\sum_{i=1}^{3} \lambda_i^2} \right)^2 \right)^{1/2} \quad (3.4)$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of $\mathbf{D}$. The FA increases with directional dependence of particle displacements and is greatest in tissue with organized microstructure such as white-matter fibres. Eigen-decomposition of $\mathbf{D}$ also gives the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ corresponding to eigenvalues $\lambda_1, \lambda_2, \lambda_3$. In three dimensions, the Gaussian distribution has ellipsoidal contours, and the eigenvalues and eigenvectors determine the elliptical shape and orientation of the contours respectively. Figure 3.1 shows the various types of diffusion tensor:

- in isotropic tensors, $\lambda_1 \approx \lambda_2 \approx \lambda_3$; there is no preferred direction of diffusion.
- in prolate tensors, $\lambda_1 \geq \lambda_2 \approx \lambda_3$; diffusion is primarily in the direction of $\mathbf{e}_1$. 

3.1. Model-Based Methods

Figure 3.1: From left to right: isotropic, prolate and oblate tensors.

- in oblate tensors, $\lambda_1 \approx \lambda_2 \geq \lambda_3$; diffusion is primarily in the plane containing $e_1$ and $e_2$.

In the case of prolate diffusion tensors, $e_1$ provides an estimate of the fibre orientation. DT-MRI is popular as a reconstruction method; however DT-MRI will only provide meaningful results if $p$ can be modelled by a Gaussian distribution. The ellipsoidal contours of the a Gaussian function can only have a single peak; this limits the use of DT-MRI to voxels that contain a single dominant fibre orientation. For example, DT-MRI is not capable of resolving multiple fibre-orientations in a voxel; in such situations, the FA is underestimated, and the fibre orientation estimation is wrong. Therefore, DT-MRI is most suited for describing microstructure in voxels known to contain only one dominant fibre-orientation.

3.1.2 Multi-compartmental models

In order to handle multiple fibre populations in a voxel, various models have been proposed with multiple compartments for each voxel. Each compartment contributes a certain proportion to the measurements; if we assume $n$ compartments in a voxel, we have

$$p = \sum_{i=1}^{n} a_i p_i,$$

where $a_i$ represents the proportion of the signal contributed by the $i$th compartment, and $\sum_{i=1}^{n} a_i = 1$. These models allow for multiple fibre populations in a single voxel. If we assume each compartment can be modelled by a diffusion tensor

$$A(q) = \sum_{i=1}^{n} a_i \exp(-t q^T D_i q),$$

(3.6)
with the principal eigenvector for each tensor representing a fibre orientation. Because of practical considerations such as the number of measurements, the increase in the number of variables as \( n \) becomes larger and the measurement noise level, most methods consider a maximum of two fibre populations.

These methods have several limitations:

- Fitting such models to the voxel measurements requires non-linear optimization, which is unstable and dependent on finding good starting points to be successful.

- In cases where the voxel contains one fibre population, the multi-compartmental model can become unstable, and produce spurious results. To resolve this problem, we need to establish the number of fibre populations in each voxel, and then use the most appropriate model.

- Multi-compartmental models will not be able to accurately reconstruct fanning or bending fibre populations in a voxel. These models may even incorrectly fit two fibre populations for such configurations; one along the ‘mean’ fanning/bending orientation, and the other perpendicular to the first orientation.

- As there is no information about the spatial arrangement of the fibre populations within the voxel, we cannot distinguish between crossing and partial volumes, where two separate fibre populations occupy part of the same voxel.

Here we will discuss a few examples of multi-compartmental models.

### 3.1.2.1 Behrens’ Model

In [16], Behrens et al propose a simple multi-compartmental model, with two components: one modelling the isotropic diffusion of free water in the voxel, and the other modelling anisotropic Gaussian diffusion along the fibres:

\[
A^*(q_i) = A^*(0) \left( (1 - f) e^{-t|q_i|^2d} + f e^{-td(e \cdot q_i)^2} \right), \tag{3.7}
\]

where \( e \) is a unit vector in the direction of the fibre population, \( d \) is the diffusivity, and \( f \) is the volume fraction of the fibre. Behrens et al use MCMC to sample the posterior distribution on the single fibre orientation in Equation 3.7 and use that distribution as an estimate of the distribution of fibre orientations in a voxel, which they use in their probabilistic tractography algorithm. Behrens’ model extends to multiple fibre-orientation
cases by including \( n \) components modelling anisotropic diffusion along different fibre orientations:

\[
A^*(q_i) = A^*(0) \left( (1 - \sum_{n=1}^{N} f_n) e^{-t|q_i|^2 d} + \sum_{n=1}^{N} f_n e^{-t d (e_n \cdot q_i)^2} \right),
\]

(3.8)

where \( e_n \) is a unit vector in the direction of the \( n \)th fibre population, \( N \) is the number of fibre populations, and \( f_n \) is the volume fraction of the \( n \)th fibre.

In [52], Hosey et al classify voxels as containing either one or two fibre-orientations by fitting the parameters of Behrens’ model [16] for both \( N = 1 \) and \( N = 2 \) in each voxel. For each model, they find the probability distribution function (PDF) of the parameters given the measurements, assuming the model has been correctly chosen, using Monte Carlo Markov Chain. The more appropriate model is chosen by calculating a statistic \( \psi_{\text{model}} \), which is the average probability of the data given the model and its inferred parameters for each model, and finding the difference of the log values \( \Phi = \ln \psi_{2\text{-fibre}} - \ln \psi_{1\text{-fibre}} \). If \( \Phi \) is large, the 2-fibre model is chosen over the 1-fibre model. They also use Bayesian inference to incorporate neighbourhood information to regularize the distributions of the fibre orientations by making low deviations between the orientations of neighbouring voxels more likely; this was found to give tighter PDFs.

Although the method is capable of distinguishing between voxels with one and two fibre-orientations, it could be extended to cover zero and three (or greater) fibre-orientations, by fitting parameters to Behrens’ model for each case, and having a different \( \Phi \) for each comparison of zero vs. one, one vs two, and two vs three fibre-orientations.

3.1.2.2 CHARMED

In [8], Assaf et al describe CHARMED (Composite Hindered And Restricted Model of Diffusion). This model is similar to Behrens’ model in that both assume two types of diffusion in each voxel; in the case of CHARMED, the two types are hindered water in the extra-axonal space, and restricted water diffusion in the intra-axonal space. The net measured signal attenuation is equal to the weighted sum of the signals from hindered and restricted components:

\[
A^*(q_i) = A^*(0) \left( f_h A^*_{h}(q_i) + f_r A^*_{r}(q_i) \right) \quad \text{with} \quad f_h + f_r = 1,
\]

(3.9)
3.2. Model-Independent Methods

where \( f_h \) and \( f_r \) are the volume fractions of the hindered and restricted compartments respectively. Diffusion in the hindered compartment is assumed Gaussian and modelled by a diffusion tensor. In the restricted component, the diffusion is decomposed into diffusivities parallel and perpendicular to the fibre orientation:

\[
A^*_\parallel(q_i) = A^*_\perp(q_i)A^*_\parallel(q_i). \tag{3.10}
\]

The diffusion parallel to the fibre can be treated as one-dimensional free diffusion and the perpendicular component can be described as restricted diffusion within impermeable cylinders, for which analytical expression for \( p \) and \( A \) exist [83].

Results in [8, 9] show CHARMED finds the fibre orientation in cases with one fibre orientation, with a higher reliability than in DT-MRI. CHARMED is also capable of finding two fibre orientations by assuming two restricted components are present [9]. As with Behrens’ model, we could use the CHARMED model to form a hierarchy; by using a suitable model selection method, the most suitable model can be chosen, which would provide an estimate for the number of fibre-orientations for the voxel.

CHARMED has also been adapted in [10] to measure the axonal diameter probability density function of excised nerve fascicles. The model of restricted diffusion in water in the intra-axonal space is extended to include the contributions of the measured signal arising from axons of different inner diameter.

3.2 Model-Independent Methods

This section reviews methods that can be used to discriminate between different voxel microstructures without a model for the underlying diffusion process.

3.2.1 Diffusion Spectrum Imaging

The orientation distribution function (ODF) [114] is a function that captures the angular structure of \( p \). The ODF is

\[
\phi(\hat{x}) = \int_0^\infty p(\alpha\hat{x})d\alpha, \tag{3.11}
\]

where \( \hat{x} \) is a unit vector in the direction of \( x \). The ODF is the radial projection of \( p \) onto the unit sphere and has peaks in the directions in which \( p \) has most mass. It is assumed that the ODF contains information about the underlying microstructure in the diffusion environment. Most commonly, we assume that peaks in the ODF correspond to the orientations of the white matter fibre populations.
3.2. Model-Independent Methods

Diffusion Spectrum Imaging (DSI) [120] reconstructs a discrete representation of \( p \) directly from measurements on a regular grid of wavenumbers \( q \) via a fast Fourier Transform. The reconstruction gives values of \( p \) on a grid of displacements. In DSI, the ODF is computed numerically by interpolating the grid representation of \( p \). We assume that the peaks in \( p \) correspond to the orientations of the fibre populations. The ODF can have multiple pairs of equal and opposite peaks. Each pair provides a separate fibre-orientation estimate, which enables DSI to resolve the orientations of multiple fibre populations in a voxel. By counting the number of peaks, we can obtain an estimate for the number of fibre orientations in each voxel.

One major advantage of DSI over DT-MRI is its ability to measure the orientations of crossing fibres. Unlike the previously discussed models, DSI measures \( p \) without making assumptions on the form of the underlying diffusion function [112]. Also, DSI is capable of resolving multiple fibre orientations without the need for model selection. However, despite these advantages, DSI is not as commonly used as DT-MRI. The biggest disadvantage of DSI is that acquisition times are long since it requires an order of magnitude more measurements than DT-MRI to get sufficient detail in the reconstructed \( p \), and requires strong magnetic field gradients [112]. Also, much of the information in the measurements contributes only to the radial structure of \( p \), which means it is inefficient to use if the aim is simply to find the number and direction of fibre populations [2]. The Fourier Transform relationship between \( p \) and \( A(q) \) assumes an idealized pulse with infinitesimally short gradient pulses. In practice, this does not happen, which results in the blurring of \( p \) and \( \phi \), although strong peak directions are not affected much.

3.2.2 q-ball

Tuch’s q-ball algorithm [112, 115] finds an approximation to \( \phi(\hat{x}) \) by computing the Funk-Radon Transform [51] of measurements from a spherical acquisition scheme. The Funk-Radon Transform is a mapping between functions of the sphere. The value of the Funk-Radon Transform of a function \( f \) at a point \( \hat{x} \) is the integral of \( f \) over \( C(\hat{x}) \), the great circle perpendicular to \( \hat{x} \). We can write this as:

\[
\phi(\hat{x}) = \int_{C(\hat{x})} A(q) \, dq,
\]

(3.12)
where \( \hat{q} = q/|q| \). This is approximately equal to the summation of the signal around the circle. In the absence of noise, the approximation becomes closer as \(|q|\) increases, as does the ability to resolve distinct peaks in \( p \); however the SNR also decreases in real measurements. In [112], Tuch demonstrates how \( q \)-ball can handle more complex fibre structure in which DT-MRI fails. Qualitative results [114, 112] show good agreement between \( q \)-ball and DSI in a fibre-crossing region in the human brain. The acquisition requirements of \( q \)-ball are more manageable than those of DSI, although the approximation of \( \phi(\hat{x}) \) introduces some blurring, which may reduce angular resolution and precision of peak directions.

### 3.2.3 PAS-MRI

In [56], Jansons and Alexander define a feature of \( p \) called the Persistent Angular Structure (PAS). The PAS function \( \tilde{p} \) is a representation of the relative mobility of particles in each direction. More precisely, the PAS is the function \( \tilde{p} \) which, when embedded in three dimensional space on a sphere of radius \( r \), has Fourier Transform that best fits the measurements [3]. By substituting \( p(x) \) with \( \tilde{p}(\hat{x})r^{-2}\delta(|x| - r) \), the aim is to find \( \tilde{p} \) that minimises:

\[
\sum_{i=1}^{N} \left( A(q_i) - \int \tilde{p}(\hat{x}) \cos(rq_i \cdot \hat{x}) d\hat{x} \right)^2
\]

where \( \hat{x} \) is a unit vector in the direction of \( x \). Jansons and Alexander use a maximum entropy parametrization of \( \tilde{p} \). They fit the \( N + 1 \) parameters of \( \tilde{p} \) using a Levenberg-Marquardt algorithm and numerical approximations of the integrals in the above equation. The function \( \tilde{p} \) can have any number of pairs of equal and opposite peaks, and each pair provides a fibre-orientation estimate, so by counting the number of peaks, we can find the number of fibre-orientations. The variable \( r \) controls the smoothness of \( \tilde{p} \).

An advantage of PAS-MRI over DT-MRI is its capability to resolve the orientations of crossing fibres. However, it is also substantially slower to compute than other algorithms discussed in this section. By replacing the maximum-entropy parametrisation of \( \tilde{p} \) with a linear basis, \( \tilde{p} \) can be estimated via a linear transformation of the data [2]. PAS-MRI resolves orientations at fibre-crossings in some cases where \( q \)-ball fails and is more sensitive than \( q \)-ball to anisotropy in test functions and reconstructs directions more consistently [2]. The number of acquisitions required is similar to that...
of q-ball, making data faster and easier to acquire than DSI.

3.2.4 Spherical Deconvolution

In [106], Tournier et al suggest that the diffusion-weighted signal attenuation measured over the surface of a sphere can be expressed as the convolution over the sphere of the distribution function of fibre-orientations in a voxel, $f$, with the signal attenuation that would be measured from a typical single coherently oriented fibre population (known as the response function $R(q; \hat{x})$), i.e.

$$A(q) = \int f(\hat{x}) R(q; \hat{x}) d\hat{x}$$

(3.14)

If the response function $R$ is known a priori, $f$ can be calculated by performing the spherical deconvolution of $R$ from the measurements. From this, we can derive information such as the number of fibre populations by counting the number of peaks in the obtained distribution, and their associated orientations. Tournier [106] derives $R$ by taking the average signal from the most anisotropic voxels. In [6], Behrens’ ball and stick model [16] is used. The most significant underlying assumption in using this method is that the diffusion characteristics of all fibre populations found in the brain are identical. Making this assumption allows the response function measured for a typical coherently oriented fibre population to be constant throughout the brain. However this assumption may not hold in certain regions [105, 106, 15]. In cases where the assumption does not hold, the calculated volume fraction of the fibre populations will be inaccurate; it does not however alter the estimates of the fibre orientations, and hence the number of fibre orientations.

A major limitation of spherical deconvolution is its susceptibility to noise, resulting in spurious peaks in the recovered distribution $f$. Recent attempts to resolve this problem include using Tikhonov regularization [47], which removes negative lobes in $f$ while retaining high angular resolution [109]. In [5], Alexander uses a maximum entropy representation for $f$ that is naturally positive definite; it requires the same computational time as PAS-MRI, as it is based essentially on the same algorithm. In [36], Dell’Acqua et al use a multi-compartmental model and rewrite this in terms of a convolution process to provide a direct physical interpretation of the signal generation process.
3.3 Voxel Classification Methods

For voxels with a single fibre population, the diffusion tensor provides an appropriate model for mapping fibre population information, and it is unnecessary to use a more complex method. The following techniques identify voxels for which the DT is not suitable.

3.3.1 Spherical Harmonics

Spherical Harmonics are orthonormal basis functions on the unit sphere. Any function on a sphere $f(\theta, \phi)$ can be expressed using spherical harmonic basis functions

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{l,m} Y_{l,m}, \quad 0 \leq \theta < \pi, \quad 0 \leq \phi < 2\pi,$$

where $Y_{l,m} : S^2 \to \mathbb{C}$ is the spherical harmonic function with order $l$, index $m$, $S^2$ is the unit sphere in three dimensions, parameterised by the angles of colatitude and longitude, $\theta$ and $\phi$ respectively, $a_{l,m} \in \mathbb{C}$ are the complex coefficients.

A series containing orders up to $l = L$ is equivalent to an order $L$ polynomial on the sphere and vice versa. In [4], Alexander et al fit the spherical harmonics series to the log measurements at a fixed $|q|$. As the function $\log (A(q))$ is real-valued, $a_{l,m} = (-1)^m a_{l,-m}^*$. Also, as the function is symmetric, the spherical harmonic representation will only contain terms with even $l$. Once the spherical harmonic coefficients are found, a model selection algorithm based on the F-test for detection of variables [7] is used to decide whether a series of order $L + 2$ fits significantly better than a series with order $L$, and the series is truncated accordingly. This classifies voxels as those with isotropic, anisotropic Gaussian, and anisotropic non-Gaussian $p$, corresponding to orders 0, 2 and 4 respectively.

3.3.2 K-means Clustering

In [23], Carew et al discuss the use of the K-means clustering algorithm [50] to classify voxels by the number of fibre-orientations, up to a maximum of two per voxel. The K-means clustering algorithm [50] is as follows:

- Randomly assign each point in the data set to one of K clusters.
- For each data point:
  - Calculate the distance from the data point to each cluster.
3.4 Conclusion

- Assign the data point to the closest cluster.
- Repeat until no points in the data set move from one cluster to another.

The K-means clustering algorithm requires the specification of an appropriate distance measure between two data values that is sensitive to shape, and is rotationally and scale invariant. Carew et al use

\[ m(A(q)^i, A(q)^k) = \sqrt{\min_M \sum_{j=1}^{N} ||A(q_j)^i - MA(q_j)^k||^2}, \]

where \( M \) is a rotation matrix to allow rotational invariance, and normalise the measurements for scale invariance. In [23], four classifications were specified \( a \ priori \), which allowed isotropic, anisotropic Gaussian and two forms of 2-fibre crossing voxels to be classified.

3.4 Conclusion

In this chapter, we have presented and discussed various algorithms designed to extract information about the fibre populations in each voxel of a diffusion MRI acquisition, each with its advantages and disadvantages. Model-based methods explicitly assume certain features of the fibres, and if the model is a good approximation of the underlying structure, they can provide useful and accurate fibre population information.

Model-independent methods have the have the advantage of being able to reconstructing multiple fibre populations in each voxel without making prior assumptions about the number of fibre populations per voxel. They estimate features of the particle displacement density, \( p \), by spherical functions with peaks that provide fibre-orientation estimates. The relationship between \( p \) and such functions is complex and unclear. Also, these methods have been shown to produce spurious peaks in isotropic regions [2]. All of the methods discussed encounter difficulties trying to distinguish between certain fibre configurations, such as partial volumes and crossing fibres, or bending and fanning.

The shortcomings described motivate our work. Firstly, we present a method to create a map of the number of distinct fibre populations in each voxel. This will allow us to decide what reconstruction algorithm is most appropriate to use for the voxel in question. Next, we present a method for finding accurate fibre population information on a sub-voxel scale; this would allow us to distinguish between various fibre configurations such as bending, fanning and partial volumes. Finally, we discuss a model that
3.4. Conclusion

describes bending and fanning white matter structures on a sub-voxel scale.
Chapter 4

Mapping the number of fibre-orientations per voxel

In this chapter, we present a hierarchical model based on Behrens’ ball and stick model discussed in Section 3.1.2.1, and use the F-test to select the most appropriate model from the hierarchy, producing a map of the number of distinct fibre populations. Work from this chapter has been presented in [76, 77, 78].

4.1 Hierarchical Model Generation and Selection

4.1.1 The Model Hierarchy

We use the hierarchical model proposed by Behrens et al [16]. Each model consists of two components: one modelling the isotropic diffusion of free water in the voxel, and the other modelling the anisotropic diffusion along the fibres. Thus for voxel $l$ we have:

$$A^*(l, \mathbf{q}_i) = A^*(l, 0) \left( (1 - \sum_{n=1}^{N} f_n) e^{-t|\mathbf{q}_i|^2d} + \sum_{n=1}^{N} f_n e^{-td(e_n, \mathbf{q}_i)^2} \right),$$

where $e_n$ is a unit vector in the direction of the $n$th fibre population, $N$ is the number of fibre populations, $d$ is the apparent diffusion coefficient, and $f_n$ is the volume fraction of the $n$th fibre. The hierarchy contains four variations of the model, with $N = 0, 1, 2, 3$.

We use a Levenberg-Marquardt algorithm to fit Equation 4.1 to the data for $N = 0, 1, 2, 3$ by least-squares minimisation. The fitting process provides estimates of $d$, $A^*(0)$, $f_n$ and $e_n$ ($n = 1, 2, 3$). We choose starting values for $f$, $d$, $A^*(0)$ and $e_n$ from the fitted single diffusion tensor as follows:

- We initialise $A^*(0)$ in the model to the arithmetic mean of the $A^*(0)$ measure-
4.1. Hierarchical Model Generation and Selection

\begin{itemize}
  \item We initialise $d$ to $\frac{1}{3} \text{Tr}(D)$.
  \item We initialise the fibre directions $e_i$, $i = 1 \ldots N$, to the eigenvectors $v_i$, $i = 1, 2, 3$, of the DT, where $v_i$ has eigenvalue $\lambda_i$.
  \item The choice of starting values for the $f_n$ varies among the 1, 2 and 3-fibre cases. In each case, we initialise the volume fraction of the isotropic component to $1 - FA$. We then distribute the initial volume fractions of the fibre populations according to the distribution of eigenvalues of the DT:
    \begin{itemize}
      \item For $N = 1$, $f_1$ is set to the FA.
      \item For $N = 2$, we assume that the anisotropic diffusion is along $v_1$ and $v_2$, and diffusion along $v_3$ is purely from the isotropic component. This means that the contribution of the anisotropic diffusion along $v_1$ is $\lambda_1 - \lambda_3$. Similarly, the contribution of the anisotropic diffusion along $v_2$ is $\lambda_2 - \lambda_3$. The volume fraction of the anisotropic components are divided along these proportions, leading to the constraint $f_1/(\lambda_1 - \lambda_3) = f_2/(\lambda_2 - \lambda_3)$, which results in $f_1 = \frac{(\lambda_1 - \lambda_3)FA}{(\lambda_1 + \lambda_2 - 2\lambda_3)}$ and $f_2 = \frac{(\lambda_2 - \lambda_3)FA}{(\lambda_1 + \lambda_2 - 2\lambda_3)}$.
      \item For $N = 3$, we initialise the volume fractions as directly proportional to the $\lambda_i$, so that $f_n = \frac{\lambda_n FA}{\sum_{i=1}^{3} \lambda_i}$.
    \end{itemize}
\end{itemize}

The Levenberg-Marquardt algorithm does not guarantee to find the global minimum of the objective function. We run the fitting routine many times from slightly perturbed starting points to improve the chances of finding the global minimum. We perturb the starting points for $d$, $A^*(\emptyset)$ and $f_n$ by additive Gaussian noise with zero mean and standard deviation one tenth of the initial values from the procedure outlined above. Similarly, we perturb the starting points for $e_n$ by additive Gaussian noise on the angles of latitude and longitude with zero mean and standard deviation 0.1.

We use the repeated runs to determine the fraction of trials in which the fitting routine finds the global minimum or a value to within 0.1% of the global minimum, and thus determine the probability of finding the global minimum in a single run. We assume the run that produces the smallest fitting error finds the global minimum, although we cannot guarantee that this is the case in a finite number of runs. In practice,
for voxels where no other run produces a value within 0.1% of the smallest fitting error, we assume the minimum has not been found, in which case the assigned probability of finding the global minimum in a single run is zero. We use these probabilities to estimate the number of runs required to ensure that we find the global minimum to some degree of confidence for each voxel. We compute the number of runs required separately for each model in the hierarchy.

4.1.2 Model Selection

In order to decide the most suitable model for each voxel, the results obtained from each model must be compared with one another. We use an F-test [7] to select the best model for the data in each voxel. The F-test compares two nested models to decide whether the improvement of fit resulting from the more complicated model is worth the cost of including the additional variables. If \( M_a \) and \( M_b \), with respective degrees of freedom \( \nu_a \) and \( \nu_b \) (where \( \nu_a < \nu_b \)), are the two models to be compared, with the null hypothesis of \( M_a \) and \( M_b \) being equally suitable given a level of confidence \( c \), i.e. the simpler model \( M_a \) is sufficient, the F-test for this hypothesis is

\[
F(M_a, M_b) = \frac{\nu_a(Var(M_b) - Var(M_a))}{(\nu_b - \nu_a)E(M_b)},
\]

(4.2)

where \( Var(M) \) is the variance of model \( M \) about its mean value, and \( E(M) \) is the mean squared error between \( M \) and the sampled points. The value obtained from Equation 4.2 is compared to a threshold value \( T_c \) such that if \( F(M_a, M_b) > T_c \), the probability of the null hypothesis is less than \( c \). The algorithm used to select the best model using the F-test is as follows:

- Initialise \( a = 3 \)
- For \( b = 2, 1, 0 \)
  - Let null hypothesis \( H_0 \) be that \( M_a \) and \( M_b \) are equally suitable
  - If \( F(M_a, M_b) < T_c \), accept \( H_0 \) and let \( a = b \)
- Select \( M_a \).

4.2 Model Generation and Selection

**Hypothesis 1**: It is possible to find the values for the parameters in Equation 4.1, for each of \( N = 0, 1, 2, 3 \), that best fit the measurements in a finite number of runs of the model fitting algorithm with 95% certainty.
 Figure 4.1: Histograms showing how often the global minimum was found out of 1000 attempts for the 1000 selected voxels for $N = 0, 1, 2, 3$.

**Experiment:** For $N = 0, 1, 2, 3$, we run the Levenberg-Marquardt algorithm 1000 times per voxel on 1000 randomly selected voxels from the brain. From this, we can estimate the probability of finding the global minimum in each case. Figure 4.1 shows how often the smallest fitting error was found out of 1000 attempts for each of the 1000 voxels. If for voxel $l_i$ ($i = 1, \ldots, L$), the minimisation algorithm converges to the global minimum (or within 0.1% of the global minimum) for $r_i$ runs out of $R$ runs, the probability of the algorithm not finding the global minimum in one run is $1 - \frac{r_i}{R}$. Therefore the mean probability that the global minimum has not been found after $k$ runs of the algorithm is $\frac{1}{L} \sum_{i=1}^{L} \left( 1 - \frac{r_i}{R} \right)^k$. Therefore, the mean probability that the global minimum is found after $k$ runs is

$$1 - \frac{1}{L} \sum_{i=1}^{L} \left( 1 - \frac{r_i}{R} \right)^k . \quad (4.3)$$

**Results:** Figure 4.2 shows the probability of finding the global minimum ($y$-axis) after a certain number of runs ($x$-axis) of the minimisation algorithm on a voxel for each model in the hierarchy, when we consider $L = 1000$ randomly selected voxels for the
4.2. Model Generation and Selection

Experiment and run the minimization $R = 1000$ times on each voxel.

**Conclusion:** The results suggest for each voxel we can find the global minimum at least once with 95% certainty for $N = 0, 1, 2, 3$ fibre orientations using 4, 15 and 45 and 1626 runs of the Levenberg-Marquardt algorithm respectively. To reduce computation time for future experiments, we limit the $N = 3$ case to 162 runs, which provides a 90% certainty of finding the global minimum.

**Hypothesis 2:** For a given voxel, and a set of parameters for each case of $N = 0, 1, 2, 3$, the F-test, described in the previous chapter, classifies the voxels as containing 0, 1, 2 or 3 fibre-orientations.

**Experiment:** We use the F-test as described in the previous section on a $128 \times 128 \times 10$ image with 61 diffusion weighted images with a $b$-value of 1270 s mm$^{-2}$ and eight measurements at $q = 0$, producing the results shown in Figure 4.3. The values of the thresholds were chosen manually. We compare the results qualitatively and quantitatively to those inferred from the model selection of the more established spherical harmonics-based voxel classification [4], discussed in Chapter 3, on the same slices (shown in Figure 4.4) by assuming voxels classified as order 0 to have zero fibre crossings, voxels classified as order 2 to contain one fibre orientation, and voxels of order 4 to have two.

**Results:** The results are similar in both cases in terms of differentiating areas of isotropic and anisotropic diffusion, although some 2-fibre clusters appear in different locations. The spherical harmonic method looks for non-Gaussian behaviour rather
4.3 Conclusions and Discussion

I = Number of isolated voxels

<table>
<thead>
<tr>
<th>No of fibre-orientations</th>
<th>New method</th>
<th>Spherical Harmonics</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>89</td>
<td>96</td>
</tr>
<tr>
<td>1</td>
<td>218</td>
<td>208</td>
</tr>
<tr>
<td>2+</td>
<td>394</td>
<td>553</td>
</tr>
<tr>
<td>Total number of voxels in class</td>
<td>13067</td>
<td>12778</td>
</tr>
<tr>
<td>I/T</td>
<td>0.68%</td>
<td>0.75%</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2058</td>
<td>7775</td>
</tr>
<tr>
<td>2+</td>
<td>7491</td>
<td>2063</td>
</tr>
<tr>
<td>I/T</td>
<td>2.91%</td>
<td>2.68%</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>13067</td>
<td>7775</td>
</tr>
<tr>
<td>2+</td>
<td>7491</td>
<td>2063</td>
</tr>
<tr>
<td>I/T</td>
<td>19.15%</td>
<td>28.61%</td>
</tr>
</tbody>
</table>

Table 4.1: Comparison of the new method with spherical harmonics.

than distinct fibres and so may pick out regions of fibre divergence that the new method
does not. The results of the new technique appear to show a less dispersed distribution
of 2-fibre voxels, which may indicate that a higher fraction of these voxels are classified
as such due to noise.

A more quantitative method of assessing the classification is to count the number
of isolated voxels. Table 4.1 provides a comparison of the two methods based on the
number of isolated voxels using 6-connectivity to decide adjacency of voxels.

The results support qualitative conclusions that the new method provides denser clus-
ters of fibre-crossing voxels. On the other hand, spherical harmonics provides a more
dispersed distribution of voxels with 2 fibre-orientations, which may indicate that a
higher fraction of voxels in this case are classified as such due to noise.

Conclusion: We can use the F-test to classify the results obtained from Behrens’ model
generation method as containing 0, 1, 2 or 3 fibre orientations. The combination of
model generation and model selection described does have some limitations, as it does
not allow for oblate fibre orientation distributions, which could be expected in various
brain regions in the presence of fanning (e.g. corona radiata) and of high curvature
(e.g. optic radiation). The results generated from this method are similar to those ob-
tained with spherical harmonics used by Alexander et al [4] as can be seen by compar-
ing the images in Figure 2 with their counterparts in Figure 3.

4.3 Conclusions and Discussion

We have demonstrated a method for mapping the number of fibre orientations in each
voxel of a 3D diffusion MRI acquisition. The results generated from this method share
similarities to those obtained using spherical harmonics [4], as can be seen by com-
paring the images in Figure 4.3 and 4.4. The new technique provides a less dispersed
distribution of crossing-fibre voxels compared to spherical harmonics, which suggests
that the new method is less sensitive to noise; it also has the benefit of distinguishing
Figure 4.3: Estimated number of fibre orientations per voxel.
Figure 4.4: Estimated number of fibre orientations per voxel using spherical harmonics.
between 2-fibre and 3-fibre voxels.

The results from this method can be used for selecting the most appropriate algorithm for finding the orientations of fibres in each voxel. This has industrial application, and is discussed further in Chapter 7.

The approach described does have some limitations, as it does not readily take into account configurations such as fanning and bending. We might improve the method by using global optimization techniques such as simulated annealing for fitting.
Chapter 5

Regularized Super-Resolution for Diffusion MRI

In this chapter, we present a new method for finding accurate fibre orientations and volume fractions of fibre populations on a sub-voxel scale from a 3D diffusion MRI acquisition in order to distinguish between various fibre configurations such as bending, fanning and partial volumes. Work from this chapter has been presented in [79, 80, 81, 82].

5.1 Motivation

In Chapter 3, we have discussed various methods that can be used to find fibre population information such as orientations of fibre populations in each voxel. If we use DT-MRI to estimate fibre population information in a voxel, we encounter problems in more complex intra-voxel fibre configurations. We can see examples of such configurations in Figure 5.1. In fanning and bending structures, the principal orientation of the DT tends to correspond to the ‘mean’ orientation of the configurations, but we lose information about more detailed structure. Furthermore, the two are indistinguishable from just the DT. The diffusion tensor is also not appropriate for partial volumes and crossing fibres; the best fit to either configuration will result in an oblate tensor, which does not contain useful information about the fibre populations. We can use alternative fibre reconstruction methods, discussed in Chapter 3, that would be capable of finding multiple fibre populations in a single voxel, but these have their own shortcomings, shown in Figure 5.1. In the case of fanning and bending configurations, we would like the output distribution of our reconstruction to have a broad peak. This may not nec-
Figure 5.1: Top row: from we have from left to right, fanning, bending, partial volume and crossing fibre configurations. Second row: Reconstruction using DT-MRI. Third row: Reconstruction using the ball and stick model, with two anisotropic components. Bottom row: Distribution of fibre orientations. Images courtesy of Kiran Seunarine.
essarily happen, and the result could imply two distinct fibre populations, rather than fanning or bending structure. Even if we were to get a broad peaked distribution, we would not be able to differentiate between bending and fanning. With partial volumes and crossing fibres, we can obtain accurate results for the orientations of the fibre populations in the voxel; however we have no information about the spatial arrangement of the populations, and therefore can not differentiate between the two configurations. The aim of this work is to able to accurately reconstruct fibre populations on a sub-voxel scale in order to be able to distinguish such configurations.

5.2 Super-Resolution Methods

The precise definition as to what constitutes a super-resolution (SR) algorithm varies from source to source, but for the purpose of this review, we will define them simply as algorithms that produce as output an image with a higher resolution than the image (or images) provided as input.

We can increase the resolution of an image by one of the following methods [45]:

- Interpolation.
- Aggregation of multiple low-resolution images to extract a super-resolution resolution image.
- Single image super-resolution.

Interpolation is the simplest of the suggested methods, and are also the least effective. Interpolation (which can be defined as fitting the original data to a continuous function and resampling at a finer resolution [26]) typically results in a blurring of edges and image details, and whilst image sharpening can be useful in some scenarios by amplifying existing high-frequency image details, it will not work when there is too much noise [45]. We shall now look in slightly more depth at the other two SR methods.

5.2.1 Multiple-Image Super-Resolution

If we have multiple low-resolution images captured from the same scene, we can reconstruct an SR image, assuming that the low-resolution (LR) images are naturally
shifted with sub-voxel precision. There are numerous variants of multiple-image super-resolution reconstruction methods; non-uniform interpolation [87] is a multi-image generalization of simple interpolation that divides the task of reconstructing an SR image into three successive stages (registration, non-uniform interpolation onto a high-resolution grid and deblurring). Image regularization has also been used to reconstruct SR images [40, 39, 22]. Peled and Yeshurun [91, 92] adapt Irani and Peleg’s iterative back-projection method [54], a form of multiple-image SR reconstruction, to create SR diffusion images. In this approach, the SR image is estimated by back-projecting the error between simulated LR images created by downsampling the SR image, and the observed LR images; the error is the difference of the two sets of images. This is repeated iteratively to minimize the error. Other methods include adaptive filtering and Projection onto Convex Sets (POCS) [87].

5.2.2 Single-Image Super-Resolution

With this method, we estimate missing high-resolution detail not present in the original image. The most prominent example of single image techniques is example-based SR [45, 43, 44]. The algorithm uses a training set to learn the fine details of an image at low resolution, and uses those learned relationships to predict fine details in other images. The training set is generated from a collection of high resolution images; each image is blurred and subsampled to create an LR image. An analytical interpolation (e.g. cubic spline) is applied to the LR image to produce an image with the same number of pixels as the high-resolution image, but without the high-resolution detail. The lowest frequency components of the LR image are filtered out, as it is the higher-frequency components that are more important to predict the extra details in the high-resolution image. The filtered LR image and high-resolution image are now contrast normalized, and the resulting images are broken into patches; each training pair contains a low-resolution patch (typically $7 \times 7$), and a corresponding high-resolution patch (typically $5 \times 5$). The relationship between the low- and high-resolution patches and between neighbouring high-resolution patches can be modelled using a Markov network. Alternatively, Freeman et al [45, 44] also present a one-pass algorithm that uses the same local relationship as the Markov network, and provides a fast approximate solution to the Markov network.
5.3 Regularization Methods

Finding fibre populations and their associated estimated orientations using the methods mentioned in Chapter 3 can be difficult due to noisy measurements. Also, partial volume effects can cause problems if we do not use a reconstruction method that is capable of resolving multiple fibre orientations in a voxel (e.g. DT-MRI). To overcome these problems, a variety of denoising, smoothing and regularization methods have been proposed. In general, regularization methods are used to obtain solutions to ill-posed inverse problems. Regularization restricts the solution obtained such that it displays both fidelity to the data, and fidelity to an a priori solution model. With diffusion MRI data, regularization uses information available from neighbouring voxels and assumptions of smoothness to restore noisy and discontinuous data. Regularization is important for tractography; noisy data can cause errors in fibre tract orientation and false positive results. Improving the orientations will allow for better results.

Some nonlinear smoothing techniques use variational principles. With variational approaches, we define a cost function $F(I)$ for an initial image $I_0$ in the image domain $\Omega$, and solve the PDE resulting from minimizing the function. In general, the cost function is of the form:

$$F(I) = \int_{\Omega} \alpha R(\nabla I) + C(I, I_0)d\Omega$$

(5.1)

where $R$ is the regularization term, $C$ constrains denoised image $I$ to $I_0$, and $\alpha$ is a weighting coefficient determining the relative contribution of each term. Different variational approaches can be categorized based on the data used in the regularization process. Parker [89] and Vemuri [116] smooth the individual diffusion-weighted images prior to the computation of anything else. In these cases, the regularization occurs at the scalar level. This type of variational approach can be expressed as:

$$F(I) = \int_{\Omega} \alpha (I - I_0)^2 + \Phi(|\nabla I|)d\Omega$$

(5.2)

where $\Phi$ is a function that smooths insignificant edges and preserves significant edges. Parker’s method uses the non-linear smoothing technique of Perona and Malik [93] on the images ($\Phi(|\nabla I|) = e^{-(|\nabla I|/k)^2}$), and Vemuri use the total-variation (TV) normalization framework [25] ($\Phi(|\nabla I|) = |\nabla I|$). The benefit of this approach is that we have
noise reduction in the raw data. However, the diffusion weighted images are smoothed independently of one another and the constraints imposed by tensors; the structural information provided by the tensor is also ignored.

Alternatively, the diffusion tensor can be regularized. This can be expressed as:

\[
F(D) = \int _\Omega \alpha ||D - D_0||^2 + \Phi(|\nabla D|) d\Omega. \tag{5.3}
\]

Examples of such methods include those by McGraw [68], and Chefd’hotel [27]. Some methods perform the regularization on the primary eigenvector [111]. Several methods such as Wang [117] and Tschumperlé [110] have simultaneous fitting and regularization:

\[
F(D) = \sum _\Omega \alpha (M(D) - I_0)^2 + \Phi(|\nabla D|) d\Omega \tag{5.4}
\]

In [34], Coulon et al extend the total variation framework introduced by Chan and Shen [25] to regularize the principal diffusion direction, and use an anisotropic diffusion process to regularize the three eigenvalues. In [35], they regularize the primary diffusion direction and reconstruct the DT. They then use this information as a prior to regularize image intensity.

A notable example of a non-variational technique is the method proposed by Poupon et al [95]. Rather than using the principal eigenvectors directly as fibre orientations, this method computes a regularized vector field that allows a trade-off between diffusion tensor data and the \textit{a priori} knowledge of assuming low curvature of most fibres. However, this model does not use all of the structural information provided by the diffusion tensor, and like many other models, is limited to a single fibre orientation per voxel. Other examples include the work of Martín-Fernández et al [66], which also incorporates a Bayesian framework, in order to regularize the diffusion tensor. Pajevic et al [86] compute a continuous tensor field by repeatedly performing B-spline transforms on the diffusion tensor data, hence smoothing the full tensor field.

All of the methods mentioned so far assume that we only have one fibre population per voxel. Some regularization methods take multiple fibres into account; Cointepas [29] use a spin-glass framework to minimize the global energy of the fibre map, given the diffusion data and infer multiple fibre directions in cases of partial volume averaging. Campbell [21] and Savadjiev [99] model fibre tracts locally as segments of 3D helix curves, which have constant curvature and torsion. A tract is a concatenation of
helix segments, and curvature and torsion can vary along the tract. The algorithm uses relaxation labeling to maximize average local support for ODF directions. Ramirez-Manzanares et al [97] express the observed tensors as linear combinations of a set of highly anisotropic basis tensors and minimize a cost function to recover the coefficients of the combination. Pasternak et al [90] describe a variational method for multiple tensor fields, which they show to be a general framework for DTI fitting and regularization.

5.3.1 Conclusion

We have discussed a variety of super-resolution methods, whose aim is to achieve high-resolution enlargements of pixel-based images. We have also discussed regularization methods used in diffusion MRI, whose aim is to maintain some form of coherence between the information in neighbouring voxels. Our goal is to use the concepts introduced in both fields to accurately reconstruct fibre population information on a sub-voxel level.

5.4 Problem Outline

We treat the task of finding accurate fibre orientations and volume fractions on a sub-voxel scale as a general inverse problem, which we solve by regularization and optimization. An overview of the method is shown in Figure 5.2.

For a set of image voxels \(l_i, i = 1, \ldots, L\), and wavenumbers \(q_k, k = 1, \ldots, M\), we have measurements \(A(l_i, q_k)\). From these measurements, we want to find \(p(s_h)\), a set of model parameters in each of a set of high-resolution voxels \(s_h\), where \(h = 1, \ldots, H\). The set of parameters depend on the model we choose to fit to the measurements. The forward problem is to estimate the measurements \(A(l_i, q_k)\) from \(p(s_h)\). Measurement \(A(s_h, q_k)\) estimates on the high-resolution grid come directly from the model parameters \(p(s_h)\), and we can estimate the measurements at \(l_i\) by

\[
\hat{A}(l_i, q_k) = \sum_{h=1}^{H} \mu_{hi} A(s_h, q_k)
\]  

(5.5)

where \(\mu_{hi}\) is a weighting coefficient that accounts for partial overlap between \(s_h\) and \(l_i\) and could also account for factors such as the point-spread function (the kernel with which an underlying image is convolved to generate an observed image) and the slice profile (the spatial distribution of sensitivity of the imaging process in the direction perpendicular to the plane of the slice).
The inverse problem finds the model parameters on the high resolution grid from $A(l_i, q_k)$. We solve the inverse problem with an optimization procedure to minimize an error metric between the observed and estimated measurements, subject to a spatial coherence constraint. For image $I$, we minimize the objective function:

$$J(I) = \alpha T(I) + E(I)$$

(5.6)

where $T(I)$ is a smoothing term ensuring that the transitions of model parameters controlling fibre populations in neighbouring sub-voxels are smooth, $\alpha$ is a weighting coefficient, and $E(I)$ is the error component defined as

$$E(I) = \sum_{i=1}^{L} \sum_{k=1}^{M} \left( A(l_i, q_k) - \tilde{A}(l_i, q_k) \right)^2.$$  

(5.7)

### 5.4.1 Model Fitting and Initialization

We use Behrens’ model [16], and assume one anisotropic fibre population in each voxel, so that

$$A(s, q_k) = (1 - f)e^{-td|q_k|^2} + fe^{-td(e \cdot q_k)^2},$$

(5.8)

where $f$ is the volume fraction of the anisotropic fibre population with orientation $e$, $d$ is the diffusivity, and $t$ is the diffusion time. The model parameter set for $s_h$ is $p(s_h) = \{d, f, e\}$. We fit the model to the data by minimizing Equation 5.7 directly with a Levenberg-Marquardt algorithm to find initial values $p(l)$ for each large voxel $l$, and use nearest neighbour interpolation to find a starting point for the optimization of $p(s_h)$.

### 5.4.2 Smoothing

For all sub-voxels $s$, we define the smoothing term of the objective function $T(s)$. The objective function rewards similarities between fibre populations in $s$ and its neighbours, and is minimized when $p(n) = p(s)$, $\forall n \in N(s)$, where $N(s)$ is the 6-neighbourhood of $s$. 

Figure 5.2: Overview of the procedure. We start with a set of measurements for each voxel, and our aim is to find fibre population information on the sub-voxel level. Here, we are effectively doubling the resolution.
5.4. Problem Outline

We propose the following candidate functions for \( T(s) \):

\[
T_1(s) = \sum_{n \in N(s)} (f_n - f_s)^2
\]
\[ (5.9) \]

\[
T_2(s) = \sum_{n \in N(s)} (1 - |e_n \cdot e_s|)^2
\]
\[ (5.10) \]

\[
T_3(s) = \sum_{n \in N(s)} |m_{s,n} \cdot e_s|^2 (1 - |e_n \cdot e_s|)^2
\]
\[ (5.11) \]

\[
T_4(s) = \sum_{n \in N(s)} (f_n - f_s)^2 + (1 - |e_n \cdot e_s|)^2
\]
\[ (5.12) \]

\[
T_5(s) = \sum_{n \in N(s)} (f_n - f_s)^2 + |m_{s,n} \cdot e_s|^2 (1 - |e_n \cdot e_s|)^2
\]
\[ (5.13) \]

Each of the objective functions take some combination of the following factors into consideration:

- The term \((f_n - f_s)^2\) captures similarity of volume fraction of \( s \) with those of its neighbours. This is used in \( T_1, T_4 \) and \( T_5 \).

- The term \(|e_n \cdot e_s|^2\) captures similarity of the fibre orientation of \( s \) with its neighbours. This is used in \( T_2, T_3, T_4 \) and \( T_5 \).

- The alignment of the voxels relative to the fibre orientation, quantified by \(|e_s \cdot m_{s,n}|^2\). If a neighbouring voxel is more closely aligned to the fibre orientation, we penalize differences in the fibre orientation more. This reflects the idea that it is more important for voxels following the path of the fibre orientation to have a similar fibre orientation. This is used in \( T_3 \) and \( T_5 \).

5.4.3 Optimization

For computational tractability, we break down the objective function into \( L \) patches, corresponding to voxels \( l_i \), which we minimize one by one. To solve the inverse problem, we minimize the objective function

\[
J(l_i) = \alpha T_i(l_i) + E(l_i) \quad \text{for } i = 1, \ldots, L
\]
\[ (5.14) \]

separately in each voxel, with the smoothing function defined as

\[
T_i(l_i) = \sum_{h=1}^H \mu_h T(s_h).
\]
\[ (5.15) \]
Using $T(s_h)$ allows each sub-optimization to still take information from neighbouring voxels into account. The error function is defined as

$$E(l_i) = \sum_{k=1}^{M} \left( \hat{A}(l_i, q_k) - \tilde{A}(l_i, q_k) \right)^2.$$  \hspace{1cm} (5.16)

We minimize the objective function with respect to $p(s_h)$ for each voxel $l_i$ by least-squares minimization using a Levenberg-Marquardt algorithm. Once we have performed the optimization for all $l_i$, we repeat this process to refine the parameters further until convergence, which is usually within a few iterations.
5.5 Experiments and Results: Synthetic Data

5.5.1 Experiment 1: Validation on synthetic data - fanning

Hypothesis: We can reconstruct a fanning fibre structure on a sub-voxel level in a synthetic dataset.

Experiment: We create a synthetic high spatial resolution dataset of $8 \times 6 \times 6$ voxels. We assume a ball and stick model, and for all voxels, we use the parameters $d = 1.5 \times 10^{-9} \text{m}^2\text{s}^{-1}$ and $f = 0.75$. We choose these values for $f$ and $d$ as they are typical values for voxels containing white matter with one anisotropic component. To create the fanning structure, we define the orientation of each anisotropic component to be along the radial line through the centre of the voxel from a point at distance of five voxels from the centre of the voxel grid\(^1\). We use the set of 61 gradient directions provided with Camino [30] computed by electrostatic energy minimization [33], and use one $q = 0$ measurement. We use Equation 5.8 to create measurements $A(s_h, q_e)$. We add Rician noise with zero mean and standard deviation of the inverse of the SNR, using SNR values of 8, 12, 16, 20, 50 and 100. We downsample the measurements by doubling the slice thickness to create a $4 \times 6 \times 6$ voxel grid and run the algorithm on the downsampled dataset. We repeat the experiment for each of the smoothing functions $T_1, T_2, T_3, T_4$ and $T_5$. We set the weighting coefficient in Equation 5.14 as $\alpha = 1$.

Results: To evaluate how the method has performed in each scenario, we use two metrics:

- Comparing the parameters found for each voxel $\tilde{p}(s_h)$ with the original parameters used to generate the voxels $p(s_h)$:

$$
C_P = \sum_{h=1}^{H} (f_{s_h} - \tilde{f}_{s_h})^2 + f_{s_h} (1 - |e_{s_h} \cdot \tilde{e}_{s_h}|^2). \quad (5.17)
$$

The first term captures similarity of $\tilde{f}_{s_h}$, the value found for the volume fraction of each sub-voxel, with $f_{s_h}$, the original value. The second term captures similarity of $\tilde{e}_{s_h}$, the value found for the orientation of the anisotropic component in each sub-voxel, with the original value $e_{s_h}$; this is weighted by the volume fraction; capturing the similarity of the orientations is more important if the associated volume fraction is large.

---

\(^1\)This model will be extended and discussed in more detail in Chapter 6.
5.5. Experiments and Results: Synthetic Data

- Comparing $\tilde{A}(s_h, q_k)$, the measurements generated from the parameters found for each voxel and $A(s_h, q_k)$, the measurements from the original parameters used to generate the voxels.

\[ C_m = \sum_{h=1}^{H} \sum_{k=1}^{M} (A(s_h, q_k) - \tilde{A}(s_h, q_k))^2 \]  \hspace{2cm} (5.18)

For both $C_p$ and $C_m$, a lower score is better. Figure 5.4 shows the performance of the optimization with each objective function for a selection of SNR values after three iterations to allow the parameters to converge. The results show the smoothing functions $T_1$, $T_4$ and $T_5$ outperform $T_2$ and $T_3$ in both $C_m$ and $C_p$ metrics. This indicates that it is important for the smoothing function to take into account the similarity of volume fractions between voxels. Figure 5.5 compares the results of the regularized super-resolution with the ground truth, initial estimates from fitting the ball and stick model to the downsampled measurements, and interpolation of the downsampled measurements. We can see that with the initial estimates, taken from fitting the model to the downsampled measurements, the values of $f$ found are slightly lower than that in the ground truth. We also see an averaging effect with the orientations when compared to the orientations of the fibre populations in the ground truth. Linear interpolation also slightly underestimates $f$, although the orientations are closer to those in the ground truth. The super-resolution however, produces an accurate reconstruction both for values of $f$ and the orientations.

**Conclusion:** We can reconstruct a fanning fibre structure on a sub-voxel level in a
Figure 5.5: (a) Ground truth of the third slice of the $8 \times 6 \times 6$ high resolution synthetic dataset with fanning structure. (b) Initial fibre population estimates, from reconstruction on the downsampled dataset. (c) Sub-voxel reconstruction with linear interpolation. (d) Reconstruction with regularized super-resolution with smoothing function $T_4$. The orientations are colour coded with red indicating left-right, green posterior-anterior and blue in-out, and the length of the orientations are in proportion to the corresponding volume fraction.
5.5. Experiments and Results: Synthetic Data

Figure 5.6: Performance of the five objective functions on the synthetic bending dataset, measured by the $C_p$ and $C_m$ metrics for a selection of SNR values.

synthetic dataset. We also see that some candidate smoothing functions ($T_1, T_4, T_5$) perform better than others ($T_2, T_3$).

5.5.2 Experiment 2: Validation on synthetic data - bending

Hypothesis: We can reconstruct a bending fibre structure on a sub-voxel level in a synthetic dataset.

Experiment: We construct synthetic a high resolution dataset of $8 \times 6 \times 6$ voxels and add noise in the same manner as described in the previous experiment. To create a bending structure, we define the anisotropic component to be perpendicular to the radial line described in the previous experiment.

Results: Figure 5.6 shows the performance of the optimization with each objective function for the datasets with various SNR values after three iterations. As with the previous experiment, we find that $T_1, T_4$ and $T_5$, outperforms $T_2$ and $T_3$ in both metrics. Figure 5.7 compares the results of the regularized super-resolution with the ground truth, initial estimates from fitting the ball and stick model to the downsampled measurements, and interpolation of the downsampled measurements. The results here echo the findings of the previous experiment; from the same downsampled dataset, the super-resolution method is more successful at accurately reconstructing fibre population information than linear interpolation.

Conclusion: We can reconstruct a bending fibre structure on a sub-voxel level in a synthetic dataset. For future experiments, we shall only consider $T_4(s) = \sum_{n \in N(s)} (f_n - f_s)^2 + (1 - |e_n \cdot e_s|)^2$ as the smoothing function, as $T_4$ has provided the best scores in
Figure 5.7: (a) Ground truth of the third slice of the $8 \times 6 \times 6$ high resolution synthetic dataset with bending structure. (b) Initial fibre population estimates, from reconstruction on the downsampled dataset. (c) Sub-voxel reconstruction with linear interpolation. (d) Reconstruction with regularized super-resolution with smoothing function $T_4$. 
5.6 Experiments and Results: Brain Data

5.5.3 Experiment 3: Validation on synthetic partial volume

Hypothesis: We can use our method to resolve fibre populations in synthetic voxels with partial volume effects.

Experiment: For this experiment, we construct a high resolution dataset of $8 \times 8 \times 10$ voxels using the parameters $\{d, f, e\} = \{1.5 \times 10^{-9} \text{m}^2 \text{s}^{-1}, 0.8, [0.9239, 0.3827, 0]\}$ for the first five slices, and $\{d, f, e\} = \{1.5 \times 10^{-9} \text{m}^2 \text{s}^{-1}, 0.8, [0.3827, 0.9239, 0]\}$ for the other five slices. The angle between the orientations of the anisotropic components of the fibre populations is $45^\circ$. We downsample this dataset in the $z$-direction, so that the middle slice now contains two fibre populations per voxel. Our initial fibre population estimates for voxels in the sub-sampled middle slices find one fibre population in each voxel with orientation of approximately $[0.7071, 0.7071, 0]$, and reduced anisotropy values, shown in Figure 5.8.

Results: We run our optimization algorithm to reconstruct the fibre populations at the original resolution. We find that the optimization is capable of correctly separating the fibre populations to each sub-voxel in the slice, thus demonstrating that we can use this method to resolve partial volume effects. We also compare the results of the regularized super-resolution to those obtained by fitting Behrens’ model to linearly interpolated measurement values, and find that the orientations do not accurately represent the actual fibre orientations.

Conclusion: We can use regularized super-resolution to accurately reconstruct fibre population information in synthetic data to resolve partial volume effects. The results obtained with this method are better than those obtained by linear interpolation.

5.6 Experiments and Results: Brain Data

5.6.1 Experiment 4: Reconstruction of the corpus callosum

Hypothesis: We can use our method to accurately reconstruct bending fibre populations in the corpus callosum.

Experiment: In this experiment, we use our regularized super-resolution method to reconstruct fibre orientations and volume fractions from a low resolution dataset and compare our results to those obtained from directly fitting the model to the measure-
Figure 5.8: (a) Ground truth of the middle four slices of the synthetic dataset. (b) Subsampled middle slices of the dataset with initial fibre population estimates. Because the middle slices contain two fibre populations, the original reconstruction is unable to resolve them correctly. (c) Fibre population estimates from linear interpolation. (d) Fibre population estimates after regularized super-resolution. The two fibre populations have been correctly separated in the middle slices.
ments of a high resolution dataset. We use diffusion-weighted human brain data from a 128 $\times$ 128 $\times$ 32 image with 61 diffusion-weighted images with a b-value of 1200 s mm$^{-2}$ and one measurement at q=0, with eight repeats of each measurement, acquired in a Philips 3T Achieva scanner. The region of interest we consider is 8 $\times$ 14 $\times$ 6 voxels in size and includes part of the corpus callosum. This ROI is our high resolution dataset. We sub-sample the region of interest by doubling the thickness of the axial slices, and use our method to reconstruct the fibre populations at the original resolution. We also fit Behrens’ model to the measurements on the original voxel space for comparison purposes.

**Result:** Figure 5.9 shows the results from using our super-resolution method on the low resolution data, and from fitting Behrens’ model to the measurements of the original data. A visual comparison of the results shows that the reconstruction accurately reproduces the fibre orientations and volume fractions in the majority of voxels.

**Conclusion:** We have demonstrated that our method can accurately reconstruct fibre population information in the corpus callosum from a sub-sampled dataset.

### 5.6.2 Experiment 5: Reconstruction of the cortico-spinal tract

**Hypothesis:** We can use our method to accurately reconstruct fanning fibre populations in the cortico-spinal tract.

**Experiment:** This experiment uses the same brain data as described in the previous experiment. The region of interest we now consider is 14 $\times$ 14 $\times$ 4 voxels in size and includes part of the cortico-spinal tract. This ROI is our high resolution dataset. We sub-sample the region of interest by doubling the thickness of the sagittal slices, and use our method to reconstruct the fibre populations at the original resolution. We also fit Behrens’ model to the measurements on the original voxel space for comparison purposes.

**Result:** Figure 5.10 shows the results from using our super-resolution method on the low resolution data, and from fitting Behrens’ model to the measurements of the original data. A visual comparison of the results shows that the reconstruction accurately reproduces the fibre orientations and volume fractions in the majority of voxels.

**Conclusion:** We have demonstrated that our method can accurately reconstruct fibre population information in the cingulum from a sub-sampled dataset.
Figure 5.9: (a) FA map with the corpus callosum ROI highlighted and the resulting fibre population estimates in (b) the ground truth (c) the reconstructed dataset. The orientations are colour coded with red indicating left-right, green anterior-posterior and blue inferior-superior, and the length of the orientations are in proportion to the corresponding volume fraction.
Figure 5.10: (a) FA map with the corticospinal tract ROI highlighted and the resulting fibre population estimates in (b) the ground truth (c) the reconstructed dataset.
5.6.3 **Experiment 6: Super-Resolution on the corpus callosum and cingulum**

**Hypothesis:** We can use our method to accurately reconstruct fibre population information in the corpus callosum, the cingulum and the voxels containing fibre populations from both structures, where we have a partial volume effect.

**Experiment:** We use the method on the same human brain dataset used in the previous two experiments. For our experiment, we consider a region of interest of $12 \times 18 \times 14$ voxels in size, which includes part of the corpus callosum and the cingulum bundle, illustrated in Figure 5.11. We use the super-resolution method to quarter the axial slice thickness, thereby quadrupling the spatial resolution. We reconstruct the fibre populations at this higher resolution. We set $\mu_{hi} = 1$ when sub-voxel $s_h$ in contained in voxel $l_i$, and zero otherwise.

**Result:** Figure 5.11 compares the results of the super-resolution method to those obtained by fitting Behrens’ model to linearly interpolated measurement values. The region highlighted in the yellow box contains two distinct fibre populations: the corpus callosum (coloured red, indicating in-out of the figure and left-right in the brain) and in the cingulum (green, indicating left-right in the figure and posterior-anterior in the brain. At the original resolution (b), partial volume effects artificially reduce the strength of the anisotropic component. Linear interpolation (c) does not help in the partial volume region, and simply interpolates the low $f$. However, the super-resolution (d) correctly retains strong orientations by separating the two directional components and identifies the boundary between the two structures with sub-voxel accuracy.

**Conclusion:** Our method is capable of accurately reconstructing fibre population information and resolving partial volume effects. We have also shown that our method succeeds where linear interpolation fails.

5.6.4 **Experiment 7: Super-Resolution on the pons**

**Hypothesis:** We can use our method to accurately reconstruct fibre population information in the pons, where we have a partial volume effect between the cortiospinal tract (inferior-superior) and transpontine fibres (left-right).

**Experiment:** We use the method on the same brain data used in previous experiments in this chapter. For this experiment, we consider a region of interest of $14 \times 14 \times 8$
Figure 5.11: (a) FA map with the ROI highlighted. Because of the partial volume effect, the region between the cingulum and corpus callosum has a low FA, and does not accurately represent the true anisotropy of the fibre populations. (b) Initial fibre population estimates at the original spatial resolution. (c) Sub-voxel fibre populations reconstructed with linear interpolation. (d) Reconstruction with regularized super-resolution.

In (b), (c) and (d), the coloured lines show the volume fraction and orientation of the anisotropic component; the lengths of the lines in 3D are proportional to $f$. The background intensities are the same in each figure and are the FA of the large voxels.
voxels in size, which includes the pons, illustrated in Figure 5.12. We use the super-resolution method to quarter the coronal slice thickness, thereby quadrupling the spatial resolution. We reconstruct the fibre populations at this higher resolution.

**Result:** Figure 5.12 compares the results of the super-resolution method with those obtained using linear interpolation and using PAS-MRI [56], which reconstructs in each voxel a continuous function on the sphere that reflects the distribution of the fibre orientations.

In the region highlighted in yellow, two distinct fibre populations are present; one coloured red in Figure 5.12 (left-right in the figure and the brain), and the other coloured blue (in-out in figure, and superior-inferior in the brain) because of partial volumes of the corticospinal tract (blue) and trans-pontine fibres (red). At the original resolution (b), we find again that partial volume effects artificially reduce the strength of the anisotropic component. Linear interpolation (c) does not help in the partial volume region, and simply interpolates the low $f$. Although the PAS-MRI reconstruction (d) is capable of reconstructing the two fibre populations in the voxels, it cannot assign spatial arrangement of the fibre populations within each voxel, much like related methods such as q-ball [115] and spherical deconvolution [105]. However, the super-resolution (e) correctly recovers stronger orientations by separating the two directional components and identifies the boundary between the two structures with sub-voxel accuracy.

**Conclusion:** Our method is capable of accurately reconstructing fibre population information and resolving partial volume effects. We have also shown that our method provides more accurate results than linear interpolation, and more information than PAS-MRI.

### 5.7 Conclusions and Discussion

We have presented a method capable of reconstructing fibre population information on a sub-voxel scale. We can use this method to reconstruct fibre configurations such as bending, fanning, and partial volume effects. In particular, we have shown that this method can resolve partial volume effects on a sub-voxel scale. The method will allow improved visualization of fine fibre configurations and improve the accuracy and precision of tractography and connectivity mapping techniques, where different fibre populations are in close proximity of one another. Note that the method is similar
to fitting Behrens’ model with multiple fibre orientations as in [17], but allows additional spatial separation of distinct directions. Future work will consider alternative smoothing functions and use multiple fibre models for each sub-voxel to distinguish more complex fibre configurations such as genuine crossings from partial volume effects. We discuss multiple fibre populations per sub-voxel in the Industrial Relevance chapter. The method extends naturally to increase resolution in multiple directions simultaneously, although the experiments here only consider one dimension for clarity. Also worth considering are alternative methods of smoothing and regularization such as [90, 100, 68, 34] that have been developed, which we may draw on for future inspiration.
Figure 5.12: (a) FA map with the ROI highlighted, (b) initial fibre population estimates at the original spatial resolution, (c) fibre population reconstruction with linear interpolation, (d) PAS-MRI reconstruction and (e) reconstruction with regularized super-resolution (we recommend viewing the electronic document at increased magnification for clearer visualization).
Chapter 6

Fanning Sub-Voxel Structures in Diffusion MRI

In this chapter, we present a new model for fanning white matter structures on a sub-voxel scale. We use the model to create synthetic diffusion MRI data from fanning structures, and explore the suitability of the model for estimating the degree of fanning in each voxel of a human brain diffusion MRI acquisition.

6.1 Motivation

In Chapter 5, we presented a method that can find complex fibre configurations on a sub-voxel scale. We now direct our interest more specifically towards sub-voxel fanning configurations. As discussed in previous chapters, reconstruction algorithms have difficulty in identifying and reconstructing fanning structures in a voxel. In particular, fanning structures and bending structures are difficult to differentiate from one another, as the fibre ODFs resulting from these configurations are the same, as shown in Figure 6.1.

Recently, some work has been done on modelling and reconstructing fanning fibre structures. In [101], Savadjiev et al use a 3D curve inference algorithm to infer differential geometric information over neighbouring voxels, and use this information to differentiate between fanning and bending configurations in voxels. In [58], Kaden et al model fibre populations in each voxel as Bingham distributions of orientations, which allows fanning and bending structures to be captured better, although it does not distinguish between the two.

Here, we present a simple model for fanning white matter structures on a sub-
voxel scale. This method allows us to directly quantify the degree of fanning. Our eventual aim is to use the model presented in this chapter (and a closely related model for bending structures) to classify voxels containing bending or fanning structures.

6.2 Method

We devise a parametric model of how the fibre orientation varies spatially over the voxel in fanning structures. Figure 6.1 illustrates the concepts used for the method. The fanning structure of the sub-voxels is determined by the position of the voxel grid relative to the centre of a series of concentric circles. The orientation at each sub-voxel is determined by the radial line passing through the centre of the sub-voxel. The closer the grid is to the centre, the greater the degree of fanning will be. Changing the position of the grid relative to the centre affects the orientation of the fanning structure. We shall now define the parameters used to create the fanning structures.

In this model, we divide each voxel into a grid of sub-voxels, and assume Behrens’ ball and stick model [16] in each sub-voxel, giving the parameter set \( p_s = \{e_s, f_s, d_s\} \) for each sub-voxel \( s_h \), where \( e_s \) is the fibre orientation, \( f_s \) is the volume fraction and \( d_s \) is the diffusivity. The orientation of the fibre bundle in each sub-voxel is along the radial line through the centre of the sub-voxel from a point \( P \). The position of \( P \) relative to the sub-voxel grid determines the fibre orientations of the sub-voxel fanning grid. Figure 6.2 illustrates the parameters used to describe the model.

We define the orientation in each sub-voxel \( s_h \) as \( e_s = \frac{\mathbf{s} - \mathbf{p}}{||\mathbf{s} - \mathbf{p}||} \), where \( \mathbf{s} \) is the vector from the centre of the high-resolution voxel grid to the centre of \( s_h \), and \( \mathbf{p} \) is the vector from the centre of the high-resolution voxel grid to \( P \), a fixed point at distance \( r = ||\mathbf{p}|| \) from the centre of the grid. Figure 6.2 (a) and (b) shows how changing \( r \) affects the orientations of the sub-voxel grid. We control the global orientation of the configuration by rotating the whole system through angle \( \psi \) about the unit axis \( \mathbf{u} \), with the centre of the voxel grid as the origin:

\[
\mathbf{p} = R(\mathbf{u}, \psi) \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}.
\] (6.1)
Figure 6.1: Overview of the model used to create fanning structures. The radial line passing through the centre of each sub-voxel determines the orientation of the associated fibre population. By comparing sub-voxel grids (a) and (b), we can see that the latter fans to a greater degree. Grid (c) shows how the orientations change when the position of the grid is altered.
The rotation matrix \( \mathbf{R} \) is

\[
\mathbf{R}(\mathbf{u}, \psi) = \begin{pmatrix}
  u_x^2(1 - c_\psi) + c_\psi & u_x u_y (1 - c_\psi) - u_z s_\psi & u_x u_z (1 - c_\psi) + u_y s_\psi \\
  u_x u_y (1 - c_\psi) + u_z s_\psi & u_y^2 (1 - c_\psi) + c_\psi & u_y u_z (1 - c_\psi) - u_x s_\psi \\
  u_x u_z (1 - c_\psi) - u_y s_\psi & u_y u_z (1 - c_\psi) + u_x s_\psi & u_y^2 (1 - c_\psi) + c_\psi
\end{pmatrix},
\]

(6.2)

where \( c_\psi = \cos \psi \) and \( s_\psi = \sin \psi \) [74]. Figure 6.2 (c) shows how \( \psi \) and \( \mathbf{u} \) affects the orientations.

To obtain configurations that fan out more in one direction than the other we reduce the size of the grid in one dimension so only the central slices remain. We call the ratio of the grid size in two dimensions the grid anisotropy, \( a \). Figure 6.2 (d) shows the effect of \( a \) on the fanning structure.

We can now describe \( \mathbf{e}_s \) in terms of the parameters \( \{a, r, \mathbf{u}, \psi\} \) and the position of the sub-voxel centre relative to the centre of the sub-voxel grid. The values of these parameters are common for all sub-voxels \( s_h \) in voxel \( l \). If we assume \( f_s \) and \( d_s \) are constant for all \( s_h \) in \( l \), we can define the parameters for all sub-voxels in the voxel grid with one set of parameters \( \mathbf{p}(l) = \{f, d, a, r, \mathbf{u}, \psi\} \), where:

- \( f \) controls volume fraction for all sub-voxels in the voxel grid
- \( d \) controls diffusivity for all sub-voxels in the voxel grid
- \( a \) controls the voxel grid anisotropy
- \( r \) controls the spatial variation of the fibre orientations.
- \( \mathbf{u} \) and \( \psi \) control the orientation of the fanning structure.

In Table 6.2, we show a selection of voxel grids and their associated parameter combinations.

### 6.2.1 Minimization

We fit the model to the data by minimizing

\[
E(l) = \sum_{k=1}^{M} \left( A(l, \mathbf{q}_k) - \hat{A}(l, \mathbf{q}_k) \right)^2,
\]

(6.3)
directly with a Levenberg-Marquardt algorithm to find \( \mathbf{p}(l) \), where \( A(l, \mathbf{q}_k), k = 1 \cdots M \) are the voxel measurements, and \( \hat{A}(l, \mathbf{q}_k) \) are the measurements reconstructed.
Figure 6.2: Examples of sub-voxel grids and how the orientations are affected by the parameters used to describe the fanning model. The left column shows the effect on the orientation in a single sub-voxel; the right column shows the orientations on a slice of the sub-voxel grid. Comparing (a) and (b) shows how varying $r$ changes the sub-voxel grid. In (c), $P$ is rotated by $\psi$ about $u$, perpendicular to the plane. In (d), we increase the value of $a$, the grid anisotropy.
Table 6.1: Examples of synthetic fanning structures. Note that the parameters $f$ and $d$ are not needed to create the sub-voxel orientations, and are hence not included in the table.
by summing sub-voxel measurements over the sub-voxel grid. The sub-voxel measurements are reconstructed from $p(l)$ using the model. To improve the fit of the model to the measurements, we perform the minimization in several stages, minimizing $E(l)$ with respect to some of the parameters of $p(l)$ whilst keeping other parameters constant at each stage of the minimization.

### 6.2.2 Initialization of parameter values

The model requires careful initialization for the fitting procedure above to converge on sensible parameter estimates. We estimate starting positions from the fitted DT, the shape of which is sensitive to the degree and anisotropy of the fanning structure. The orientation of the DT provides a direct estimate of the orientation parameters of the model, $u$ and $\psi$. The relationship between the DT and other parameters of the model, however, is less clear. We fix $f$ as 0.75; we choose this value as the model is only applicable in white matter regions and the intra-cellular volume fraction in dense white matter is around 75% [1]. For the shape parameters, $d$, $a$ and $r$, we learn a crude linear mapping from the DT eigenvalues by using a synthetic training set derived from our model. We discuss the initialization of $d$, $a$, $r$, $\psi$ and $u$ in more detail in Appendix A.

### 6.3 Experiments and Results

#### 6.3.1 Experiment 1: Synthetic Data

**Hypothesis:** We can use our method to accurately reconstruct fanning structures from synthetic data.

**Experiment:** The purpose of this experiment is to ensure that the method can recover known parameter settings in the same model. We use various configurations of $p(l)$ to create a fanning fibre structure on a $5 \times 5 \times 5$ high resolution voxel grid, and create voxel measurements from these. We use the set of 61 gradient directions as used in experiments in Chapter 5 and use one $q = 0$ measurement. We add Rician noise with zero mean and standard deviation of the inverse of the SNR, using SNR values of 12 and 20. From the measurements, we reconstruct the fanning model parameter set $\tilde{p}(l)$ by fitting the model using the procedure outlined above.

**Results:** Table 6.3.1 shows the results obtained from 12 configurations of $p(l)$. The table shows the parameter sets $p(l)$ used to create the synthetic data, the parameter set
values recovered from the data with SNR values 12 and 20, the distribution of the fibre populations, and the error metric $E(l)$ described in Equation 6.3 in each case. From the results, we can see that the reconstructed parameter values are in general agreement with the initial parameter values. As we would expect, the reconstructions from the higher SNR datasets provide more accurate results and a smaller error. There are a few cases where the optimization does not converge to the global minimum and hence does not find the correct result; an example of such an occurrence is in Experiment #7.

**Conclusion:** We have shown that our method is capable of reconstructing synthetic fanning structures. We can avoid converging to local minima either by perturbing the starting point for our optimization, or by altering the stages of the minimization process.

### 6.3.2 Experiment 2

**Hypothesis:** We can use the fanning model to reconstruct the fanning structure in each voxel of a region of interest containing the fanning structure in the corticospinal tract.

**Experiment:** We use diffusion-weighted human brain data as described in the experiments in Chapter 5. The region of interest we consider is $14 \times 14 \times 4$ voxels in size and includes part of the corticospinal tract (CST), shown in Figure 6.3. The ROI contains three distinct sections. The first section, highlighted in yellow in 6.3, contains a crossing of the CST with the superior longitudinal fasciculus (SLF). The section highlighted in red consists mainly of the fanning corona radiata, but may contain some crossings with fibres parallel to the SLF. In the third region, highlighted in blue, the CST intersects the corpus callosum. We use our method to find fanning parameters for each voxel in the ROI, and assume each voxel consists of a $5 \times 5 \times 5$ sub-voxel grid.

**Results:** Figure 6.3 shows the resulting fibre population distributions in a sagittal slice of the ROI, and compares the results of the new reconstruction method (Figure 6.3f) to those obtained from fitting the diffusion tensor (6.3b) and using constrained Spherical Deconvolution (cSD) [108] (Figure 6.3c), PAS-MRI [56] (Figure 6.3d) and q-ball (Figure 6.3e). In the case of the diffusion tensor reconstruction, the results show the ‘mean’ orientation of the fanning structure in each voxel, but the DT simply assumes a single anisotropic fibre population for the voxel, which does not reflect the full complexity of the structures.

The results from cSD reconstruction are sometimes misleading as they suggest
Table 6.2: Comparison of synthetic fanning structures for various configurations of $p(l)$ and reconstructions after adding noise. The sub-voxel reconstructions show the spatial arrangement of the middle slice; the distributions of orientations shows the orientations for all sub-voxels in the grid.
### 6.3. Experiments and Results

#### Original SNR = 12

<table>
<thead>
<tr>
<th>Original</th>
<th>SNR = 12</th>
<th>SNR = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td>Distribution</td>
<td>Distribution</td>
</tr>
</tbody>
</table>

#### Original SNR = 20

<table>
<thead>
<tr>
<th>Original</th>
<th>SNR = 12</th>
<th>SNR = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td>Distribution</td>
<td>Distribution</td>
</tr>
</tbody>
</table>

#### Parameters

- \( f \) (x) \times 10^{-9}
- \( a \)
- \( \psi \)
- \( r \)
- \( u_x \)
- \( u_y \)
- \( E \)

#### Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Original</th>
<th>SNR = 12</th>
<th>SNR = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>( a )</td>
<td>1.5</td>
<td>2.0115</td>
<td>1.6109</td>
</tr>
<tr>
<td>( \psi )</td>
<td>1.3598</td>
<td>1.1809</td>
<td>1.1177</td>
</tr>
<tr>
<td>( r )</td>
<td>8</td>
<td>3.5290</td>
<td>6.9382</td>
</tr>
<tr>
<td>( u_x )</td>
<td>0.75</td>
<td>0.8187</td>
<td>0.9148</td>
</tr>
<tr>
<td>( u_y )</td>
<td>-0.125</td>
<td>0.0140</td>
<td>0.0574</td>
</tr>
<tr>
<td>( E )</td>
<td>0</td>
<td>0.3403</td>
<td>0.119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Original</th>
<th>SNR = 12</th>
<th>SNR = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>( a )</td>
<td>4</td>
<td>0.2903</td>
<td>6.4945</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.3142</td>
<td>0.2793</td>
<td>0.4348</td>
</tr>
<tr>
<td>( r )</td>
<td>15</td>
<td>107.5965</td>
<td>27.2997</td>
</tr>
<tr>
<td>( u_x )</td>
<td>0.1429</td>
<td>0.2223</td>
<td>0.2192</td>
</tr>
<tr>
<td>( u_y )</td>
<td>0.8</td>
<td>0.9739</td>
<td>0.6459</td>
</tr>
<tr>
<td>( E )</td>
<td>0</td>
<td>0.5517</td>
<td>0.108</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Original</th>
<th>SNR = 12</th>
<th>SNR = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>( a )</td>
<td>14</td>
<td>38.1849</td>
<td>45.6666</td>
</tr>
<tr>
<td>( \psi )</td>
<td>2.7925</td>
<td>1.6233</td>
<td>-2.2311</td>
</tr>
<tr>
<td>( r )</td>
<td>15</td>
<td>70.7171</td>
<td>51.9204</td>
</tr>
<tr>
<td>( u_x )</td>
<td>0.1430</td>
<td>0.7295</td>
<td>0.8689</td>
</tr>
<tr>
<td>( u_y )</td>
<td>0.8</td>
<td>-0.5342</td>
<td>0.2051</td>
</tr>
<tr>
<td>( E )</td>
<td>0</td>
<td>0.5991</td>
<td>0.146</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Original</th>
<th>SNR = 12</th>
<th>SNR = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>( a )</td>
<td>0.6</td>
<td>0.0000</td>
<td>-0.0598</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.3</td>
<td>1.4541</td>
<td>0.4696</td>
</tr>
<tr>
<td>( r )</td>
<td>7</td>
<td>5.6823</td>
<td>6.2210</td>
</tr>
<tr>
<td>( u_x )</td>
<td>-0.5</td>
<td>-0.2005</td>
<td>-0.2766</td>
</tr>
<tr>
<td>( u_y )</td>
<td>-0.8</td>
<td>-0.0865</td>
<td>-0.2281</td>
</tr>
<tr>
<td>( E )</td>
<td>0</td>
<td>0.4257</td>
<td>0.1385</td>
</tr>
</tbody>
</table>
### 6.3. Experiments and Results

<table>
<thead>
<tr>
<th>#</th>
<th>Original</th>
<th>$\text{SNR} = 12$</th>
<th>$\text{SNR} = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>$\begin{array}{c} f \ d \times 10^{-9} \ a \ \nu \ r \ u_x \ u_y \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.5 \ 1.5 \ 10 \ 1.3 \ 0.3183 \ 0.4286 \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.4306 \ 0.8306 \ 0.9519 \ 6.2178 \ 0.3147 \ 0.5811 \end{array}$</td>
</tr>
<tr>
<td>10</td>
<td>$\begin{array}{c} f \ d \times 10^{-9} \ a \ \nu \ r \ u_x \ u_y \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.5 \ 2 \ 10 \ 1.3 \ 0.3183 \ 0.4286 \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.4026 \ 1.4688 \ 6.8915 \ 2.7258 \ 0.4476 \ 0.1390 \end{array}$</td>
</tr>
<tr>
<td>11</td>
<td>$\begin{array}{c} f \ d \times 10^{-9} \ a \ \nu \ r \ u_x \ u_y \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.5 \ 3 \ 7.5 \ 1.3 \ 0.5776 \ 0.5776 \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.6372 \ 3.4947 \ 5.8062 \ 3.1663 \ 0.0435 \ -0.0833 \end{array}$</td>
</tr>
<tr>
<td>12</td>
<td>$\begin{array}{c} f \ d \times 10^{-9} \ a \ \nu \ r \ u_x \ u_y \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 1.5 \ 8 \ 0.4 \ 25 \ 0.5776 \ 0.5776 \end{array}$</td>
<td>$\begin{array}{c} 0.75 \ 2.1459 \ 1.8756 \ 4.6181 \ 0.3409 \ -0.6531 \ -0.7432 \end{array}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E(\parallel)$</th>
<th>Original</th>
<th>$\text{SNR} = 12$</th>
<th>$\text{SNR} = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0</td>
<td>0.4162</td>
<td>0.1491</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0.3412</td>
<td>0.2094</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0.3412</td>
<td>0.2094</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0.3094</td>
<td>0.1496</td>
</tr>
</tbody>
</table>
fibre crossings in voxels we expect to contain fanning. Results from PAS-MRI suggest fanning structures in these voxels in a qualitative manner, but interpreting the results to obtain quantitative information about the degree of fanning is a non-trivial task. The peaks of the PAS reconstructions are anisotropic in this region [102], which suggests fanning qualitatively, but it is difficult to determine the precise degree of fanning from the peak shape [102].

PAS and cSD both suggest crossings in some fanning areas, but cSD often suggests 3-way crossings. These are highly unlikely in this region. The two-fibre crossing configurations that PAS often shows are possible; however it is likely to be spurious in lower regions of the ROI, and more likely to be correct higher up where the SLF intersects in the red ROI. \( q \)-ball does show single peaks in the fanning region, but captures crossing less well and does not have peak anisotropy.

The new method suggests plausible fanning structures for voxels where we expect fanning, although we can see that in a minority of voxels, the optimization process visibly fails to find the global minimum, resulting in unlikely fanning distributions given the distributions of neighbouring voxels.

**Conclusion:** The fanning model recovers plausible fanning structures on sub-voxel scale in known regions of fanning in human brain data. Since the model is designed to model fanning structures, it is not surprising to see the model behave unusually when trying to deal with other configurations.

### 6.4 Conclusions and Future Work

We have presented a model for generating fanning structures on a sub-voxel scale. Preliminary work suggests that we can use the model to reconstruct fanning structures in real brain data and provide quantitative information about the fanning structure. Defining the structure more accurately allows more appropriate action to be taken by tractography algorithms, resulting in fewer false positive and false negative tracts. Further work on fitting the model to the measurements would involve

- improving the procedure used to find the starting point and the minimization process.
- investigating the effects of the sub-voxel grid size, as some voxels in our results
Figure 6.3: (a) FA map with the corticospinal tract ROI highlighted and reconstructions from (b) DT, (c) cSD, (d) PAS-MRI, (e) q-ball and (f) sub-voxel fanning model, showing the fanning distribution in each voxel.
show the effects of finite sampling.

### 6.4.1 Bending Structures and Classification of Voxels

Referring back to our conceptual diagram of 6.1, by using tangential, rather than radial directions to define the sub-voxel fibre orientations, we can extend our model to generate and reconstruct sub-voxel structures with bending. We could use this to create and reconstruct mixtures of bending and fanning structures in voxels. From our model, we envisage an algorithm that uses fanning and bending models to the voxel measurements to classify voxels as either containing a fanning or bending structure. Such an algorithm will require knowledge from neighbouring voxels; if the set of sub-voxel orientations produced by fanning and bending models are the same, the spatial arrangement of the orientations does not affect the sum in Equation 6.3.
Chapter 7

Industrial Relevance

7.1 Introduction

Much of the work presented in this thesis has been motivated by relevance to industry. The Engineering Doctorate (EngD) is designed in response to the needs of industry and the demand for industrial qualifications coming from students, during which time the student (officially known as a ‘Research Engineer’) undertakes a project that presents research challenges whilst solving problems relevant to the sponsoring company. The sponsor of this project, Philips Medical Systems, created the scanner from which the first clinical MR images were acquired. Their current MR scanners include software for use with diffusion-weighted data. One of the main selling-points of the software is the use of tractography toolkits, such as FiberTrak [41]; the benefits highlighted for users include safe surgery under tractography guidance and using tractography results for successful resectioning. Being able to accurately reconstruct complex white matter structures is therefore of genuine interest to the company.

In this chapter, we shall discuss which parts of work previously discussed have been motivated by the industrial sponsor, and how the work has been brought to use for the sponsor.

7.2 Mapping the number of fibre-orientations per voxel

The work discussed in Chapter 4 was strongly motivated by industrial interest. In clinical environments, the diffusion tensor is by some margin the most popular method used to reconstruct fibre population information and for tractography. More recently, software has been written for the scanners that can fit multi-tensor models to the mea-
surements such as [123], but as discussed in the literature review, such methods will encounter several problems. In particular, for cases where the voxel contains one fibre population, the multi-tensor model can become unstable, and produce spurious results. Therefore, if we can correctly assign the number of fibre-orientations in a voxel, we would be able to use the diffusion tensor model for voxels with one anisotropic fibre population, and more complex models in other voxels.

Work was done at the industrial partner’s site to implement the work described in Chapter 4 for their scanners. One of the main considerations that needed to be taken into account is speed; fast processing is paramount for clinical environments, and is one of the reasons the diffusion tensor is so commonly used, in spite of the alternative reconstruction methods available. The code we had started to write is being optimized from the original incarnation, written in Matlab, to run as quickly as possible by re-writing the code in C++ as part of Philips’ diffusion imaging software package.

7.3 Regularized Super-Resolution

The work discussed in Chapter 5 is also of industrial interest. Two main points of interest were brought up from this work:

- Can we extend the model and smoothing function to incorporate two anisotropic fibre populations? This would allow us to more accurately reconstruct crossings on the sub-voxel level, and also distinguish crossings from partial volumes.

- Can we use this method to with differently sub-sampled datasets to find the best way to reconstruct accurate fibre population information? If we can use a sub-sampled dataset to accurately reconstruct fibre population information at a higher resolution, we can spend less time scanning, which can be very useful. This is a significant application of the regularized super-resolution method, as it offers the potential to get high-resolution images with less acquisition.

To answer the questions, we modified the method discussed in Chapter 5 to be able to take two fibre-populations per sub-voxel into account.
7.3. Regularized Super-Resolution

7.3.1 Model Fitting and Initialization: Extending to two fibre-populations

The model fitting and initialization process is similar to that discussed in Chapter 5. As before, we use Behrens’ model [16]; we now assume a maximum of two anisotropic fibre populations in each voxel, so that

\[ A(s_h, q_k) = f_0 e^{-t|q_k|^2d} + f_1 e^{-td(e_1 \cdot q_k)^2} + f_2 e^{-td(e_2 \cdot q_k)^2} \]  

(7.1)

where \((f_1, e_1)\) and \((f_2, e_2)\) are the respective volume fractions and orientations of the fibre populations in the voxel, \(d\) is the diffusivity, \(t\) is the diffusion time, and \(f_0 = 1 - (f_1 + f_2)\). The model parameter set for \(s_h\) is now \(p(s_h) = \{d, f_1, f_2, e_1, e_2\}\). We fit the model to the data to find initial values for \(p(l_i)\), and interpolate linearly for \(p(s_h)\) as in Chapter 5:

\[ p(s_h) = \sum_{i=1}^{L} \rho_{hi} p(l_i) \]  

(7.2)

where \(\rho_{hi}\) is the partial overlap fraction of sub-voxel \(s_h\) in voxel \(l_i\).

7.3.2 Smoothing

We need to re-define the smoothing function discussed in Chapter 5 to take two fibre populations into account. For each sub-voxel \(s_h\), we define the smoothing function \(T_1\) as:

\[ T_1(s_h) = \sum_{\forall n \in N(s)} \left[ \sqrt{f_{s_h1}f_{n1}} |e_{s_h1} \cdot m_{s_h,n}|^2 |e_{n1}|^2 \\
+ \sqrt{f_{s_h2}f_{n2}} |e_{s_h2} \cdot m_{s_h,n}|^2 |e_{n2}|^2 \\
+ \frac{1}{3} \sqrt{f_{s_h0}f_{n0}} (1 - |f_{n0} - f_{s_h0}|) \right] \]  

(7.3)

where \(N(s_h)\) is the 6-neighbourhood of \(s_h\), \(m_{s_h,n}\) is the vector distance of the centres of voxels \(s_h\) and \(n\), and \(f_{s_h0}\) and \(f_{n0}\) are the isotropic volume fractions of \(s_h\) and \(n\) respectively (shown in Figure 7.1).

The value of \(T_1(s_h)\) is invariant to the model parameters, and is constant when the fibre populations and orientations of all neighbouring voxels of \(s_h\) are identical to that of \(s_h\), i.e. \(p(n) = p(s_h)\) for \(\forall n \in N(s_h)\). The factors taken into consideration for \(T_1(s)\) are similar to those contributing to the smoothing function with one fibre population per sub-voxel in Chapter 5, but now take two fibre populations into account:
7.3. Regularized Super-Resolution

Figure 7.1: Illustration of a sub-voxel $s$ with two voxels of its 6-neighbourhood, and associated fibre populations.

- Similarity of fibre orientations of the fibre populations of $s$ with those of its neighbours, measured by $|e_{s_1} \cdot e_{n_1}|^2$ and $|e_{s_2} \cdot e_{n_2}|^2$.

- Similarity of volume fractions in $s$ with those of its neighbours, measured by $1 - |f_{n_0} - f_{s_{00}}|$.

- The alignment of the voxels relative to the fibre orientation, measured by $|e_{s_{n_1}} \cdot m_{s_{n_1}}|^2$ and $|e_{s_{n_2}} \cdot m_{s_{n_1}}|^2$. If a neighbouring voxel is more closely aligned to the fibre orientation, we penalize differences in the fibre orientation more. This reflects the idea that it is more important for voxels following the path of the fibre-orientation to have a similar fibre-orientation.

We also define a second smoothing function based on the normalized difference of $d$ compared to neighbouring values:

$$T_2(s_h) = \frac{|d_{s_h} - d_n|}{\sqrt{d_{s_h}d_n}} \quad (7.4)$$

We can now define the smoothing term:

$$T(l) = \sum_{h=1}^{H} (T_1(s_h))^{-2} + \beta (T_2(s_h))^2 \quad (7.5)$$

where $\beta$ is a weighting coefficient. $\beta$ allows us to control the relative contribution of $T_2(s_h)$ to the smoothing term.
7.3. Regularized Super-Resolution

7.3.2.1 Fibre Matching

To compute $T_1(s_h)$, we need to associate the fibre populations of $s_h$ with those in the neighbouring voxels to decide which pairs of fibres match. If $\{(f_{x_1}, e_{x_1}), (f_{x_2}, e_{x_2})\}$ are the fibres in voxel 1 and $\{(f_{y_1}, e_{y_1}), (f_{y_2}, e_{y_2})\}$ are the fibres in voxel 2, we consider the two available possibilities:

- **Possibility 1**: $(f_{x_1}, e_{x_1})$ corresponds to $(f_{y_1}, e_{y_1})$ and $(f_{x_2}, e_{x_2})$ corresponds to $(f_{y_2}, e_{y_2})$.

- **Possibility 2**: $(f_{x_1}, e_{x_1})$ corresponds to $(f_{y_2}, e_{y_2})$ and $(f_{x_2}, e_{x_2})$ corresponds to $(f_{y_1}, e_{y_1})$.

We evaluate both possibilities with:

$$H_1 = (1 - |f_{x_1} - f_{y_1}|)(1 - |f_{x_2} - f_{y_2}|)(f_{x_1} f_{y_1} |e_{x_1} \cdot e_{y_1}| + f_{x_2} f_{y_2} |e_{x_2} \cdot e_{y_2}|)$$

$$H_2 = (1 - |f_{x_2} - f_{y_1}|)(1 - |f_{x_1} - f_{y_2}|)(f_{x_2} f_{y_1} |e_{x_2} \cdot e_{y_1}| + f_{x_1} f_{y_2} |e_{x_1} \cdot e_{y_2}|).$$

In Equation 7.6, the first and second terms in parentheses increase the value of the result if the associated voxel fractions are close together. The third term will have a high value if the corresponding fibre orientations from the pair of voxels are similar; the similarity is evaluated by finding the vector product of the orientations. As the importance of the similarity of fibre orientations depends on the contributions of the associated volume fractions, we also multiply the vector product by the volume fractions. If $H_1 > H_2$, we assume that possibility 1 is true, otherwise we assume possibility 2 to be true.

7.3.3 Optimization

For computational tractability, we break down the objective function into $L$ patches corresponding to voxels $l_i$, as was the case in Chapter 5. We solve the inverse problem by minimizing the objective function

$$J(l_i) = \alpha T(l_i) + E(l_i) \quad \text{for } i = 1 \cdots L$$

separately in each voxel, with the smoothing function now defined as

$$T(l_i) = \sum_{h=1}^{H} \mu_{hi} \left[ (T_1(s_h))^{-2} + \beta (T_2(s_h))^2 \right]$$
and the error function defined as

\[ E(l_i) = \sum_{k=1}^{M} \left( \frac{A(l_i, q_k) - \tilde{A}(l_i, q_k)}{A(l_i, q_k)} \right)^2 \]  

(7.9)

We minimize the objective function for each voxel \( l_i \) by least-squares minimization using a Levenberg-Marquardt algorithm with respect to \( p(s_h) \). Once we have performed the optimization for all \( l_i \), we repeat this process to refine the parameters further until convergence, which is usually within a few iterations.

### 7.3.4 Experiments and Results

**Hypothesis:** A ‘jittered’ acquisition (described in more detail below) of real brain data, provides better fibre population reconstruction compared to a low-spatial high-angular resolution acquisition or a high-spatial low-angular resolution acquisition.

**Experiment:** We test our method on three datasets that are sub-sampled in different ways from the image data used in Chapter 5 (128 × 128 × 32 image, 61 diffusion weighted images, a \( b = 1200 \) s mm\(^{-2}\) and one \( q = 0 \) measurement, with eight repeats of each measurement). Each of the datasets contain the same number of measurements, and simulate acquisitions that we expect to require approximately the same time. We run our optimization method on each dataset to determine which set provides us with the most accurate information about fibre populations in voxels. From one repeat, we select a 18 × 18 × 8 ROI (shown in Figure 7.2) and sub-sample to create three datasets:

- For dataset 1, we double the thickness of each slice to create a 18 × 18 × 4 voxel array with 61 diffusion weighted images and one measurement at \( q = 0 \) to create a dataset with low-spatial high-angular resolution.

- For dataset 2, we only use 30 diffusion weighted images and one measurement at \( q = 0 \) while retaining the spatial resolution; we choose the subset of 30 directions from the original set of 61 that are optimally spread over the unit sphere, using the method proposed by Cook et al [33], which is implemented in the Camino package [31]. This is done by modelling each direction as the axis of a pair of identically charged particles on the sphere, and searching for a configuration that minimizes the electrostatic energy of pairs in the same subset using simulated-annealing optimization. This provides us with a high-spatial low-angular resolution dataset.


- Dataset 3 consists of two subsets; both have the slice thickness doubled to create a $18 \times 18 \times 4$ voxel array, and consist of 30 diffusion weighted images and one measurement at $q = 0$. The two subsets are offset by half a slice width in the $z$-direction, and each of the ‘jittered’ subsets uses 30 of the 61 original directions with no overlap.

**Results:** We run our optimization algorithm on all three datasets to reconstruct fibre population information over the $18 \times 18 \times 8$ voxel grid. We evaluate the reconstruction from each dataset using the metric

$$C = \sum_{h=1}^{H} \sum_{k=1}^{M} \left( A(s_h, q_k) - \tilde{A}(s_h, q_k) \right)^2$$  \hspace{1cm} (7.10)

where $A(s_h, q_k)$ is now the average measurement of all 8 repeats; lower $C$ is better. We show the reconstructed fibre populations and orientations on a slice from each dataset in Figure 7.2. After five iterations, the scores for each reconstruction are 46.17, 36.78 and 57.56 respectively, indicating that dataset 2 provides more accurate information.

**Conclusion:** We find that a high-spatial low-angular resolution acquisition provides better reconstruction compared to a low-spatial high-angular acquisition. The jittered acquisition appears to offer no advantage; this is most likely to be because it is harder to find a good starting point. Further work is required to test the dependence on starting point and algorithm parameters. In order to perform this experiment, we had actually acquired a ‘jittered’ dataset with Philips; however motion problems rendered it unusable. We would like to continue our efforts to acquire a workable dataset.

### 7.4 Conclusions and Discussion

We have discussed how the research has been motivated by industrial interest, and the knowledge transfer involved. More specifically, we have discussed the relevance of the fibre-population counting method for using an appropriate fibre reconstruction method in clinical environments. We have also discussed the industrial application of our regularized super-resolution algorithm, discussed first in Chapter 5. We have extended the method to take two fibre populations into account in each sub-voxel, and we have used the method to compare fibre reconstructions from differently sub-sampled acquisition datasets. From our comparison, we found that a high-spatial low-angular resolution
Figure 7.2: (a) FA map with the ROI highlighted, and the resulting fibre populations estimates in (b) the ground truth, (c) low-spatial high-angular (LSHA), (d) high-spatial low-angular (HSLA), and (d) ‘jittered’ (bottom-right) datasets.
acquisition provides better reconstruction compared to a low-spatial high-angular acquisition. The jittered acquisition appears to offer no advantage, although further work is required to test the dependence on starting point and algorithm parameters.
Chapter 8

Conclusions

We have presented a method for mapping the number of fibre orientations in each voxel of a 3D diffusion MRI acquisition. The results generated from this method share similarities to those obtained using spherical harmonics as described by Alexander et al in [4]. However, the new technique provides a less dispersed distribution of crossing-fibre voxels compared to spherical harmonics, and also has the benefit of distinguishing between 2-fibre and 3-fibre voxels. The resulting information can be used to select the most appropriate algorithm for finding the orientations of fibres in each voxel. This work was motivated by industrial interest, as discussed in Chapter 7. We might improve the method by using global optimization techniques such as simulated annealing for fitting. However, if this method is to be used in a clinical environment, the algorithm needs to run as close to real-time as possible, and using simulated annealing is a step in the opposite direction in terms of speed. Finding an appropriate balance between speed and accuracy will be an interesting task.

We have presented a method that can find accurate fibre population information on a sub-voxel scale. This allows us to help distinguish between various fibre configurations such as bending, fanning and partial volumes. Results from our experiments demonstrated that we can use this method to resolve partial volume effects. The method will allow improved visualization of fine fibre configurations and improve the accuracy and precision of tractography and connectivity mapping techniques. The method extends naturally to increase resolution in multiple directions simultaneously. Future work would also include considering more alternative smoothing functions. In Chapter 7, we presented work done on extending the model to two fibre populations per sub-voxel, and how the method can be used to reconstruct fibre population information from
sub-sampled datasets, although further work is required to extract as much information as possible from the sub-sampled datasets.

Finally, we have presented a model for generating fanning structures on a sub-voxel scale. This model extends naturally to modelling bending structures, and we intend to use these models to be able to reconstruct combinations of fanning and bending in each voxel. The models could also be used, in conjunction with neighbourhood information, to classify voxels as containing fanning or bending white matter architecture. Further work on fitting the model to the measurements would involve improving the procedure used to find the starting point and the minimization process.

All of the work completed makes progress towards addressing the research objective of extracting more accurate information about the spatial arrangement of fibre populations; the regularized super-resolution method in particular allows us to reconstruct various sub-voxel structures and configurations, and the other contributions of this thesis target more specific problems that allow us to obtain more information about fibre populations in each voxel.

We believe that the approach considered here shows promise for further application in the field of diffusion imaging and can make further interesting contribution to the study of tissue microstructure using diffusion MRI.
Appendix A

Initialization of Fanning Model Parameters

Initialization of \(d\), \(a\) and \(r\)

To initialize \(d\), \(a\) and \(r\), we assume that:

- we can approximate \(d\) from a linear combination of the eigenvalues of the diffusion tensor \(\lambda_1, \lambda_2, \lambda_3\), and the ratios of these values.

\[
d = \alpha_0 + \alpha_1 \lambda_1 + \alpha_2 \lambda_2 + \alpha_3 \lambda_3 + \alpha_4 \frac{\lambda_1}{\lambda_2} + \alpha_5 \frac{\lambda_1}{\lambda_3} + \alpha_6 \frac{\lambda_2}{\lambda_3} \quad (A.1)
\]

- we can approximate \(a\) from a linear combination of the log values of \(\lambda_1, \lambda_2, \lambda_3, d\) and \(r\).

\[
\log a = \beta_0 + \beta_1 \log \lambda_1 + \beta_2 \log \lambda_2 + \beta_3 \log \lambda_3 + \beta_4 \log d + \beta_5 \log r \quad (A.2)
\]

- we can approximate \(r\) from a linear combination of the log values of \(\lambda_1, \lambda_2, \lambda_3, d\) and \(a\).

\[
\log r = \gamma_0 + \gamma_1 \log \lambda_1 + \gamma_2 \log \lambda_2 + \gamma_3 \log \lambda_3 + \gamma_4 \log d + \gamma_5 \log a \quad (A.3)
\]

If we can find values for the \(\alpha\), \(\beta\), and \(\gamma\), we can calculate \(d\) using Equation A.1. We can then find \(a\) and \(r\) by:

\[
a = \exp\left(\frac{c_1 + \beta_5 c_2}{1 - \beta_5 \gamma_5}\right) \quad \text{and} \quad r = \exp\left(\frac{c_2 + \beta_5 c_1}{1 - \beta_5 \gamma_5}\right) \quad (A.4)
\]
where

\[ c_1 = \beta_0 + \beta_1 \log \lambda_1 + \beta_2 \log \lambda_2 + \beta_3 \log \lambda_3 + \beta_4 \log d \]
\[ c_2 = \gamma_0 + \gamma_1 \log \lambda_1 + \gamma_2 \log \lambda_2 + \gamma_3 \log \lambda_3 + \gamma_4 \log d \]

To find suitable values of \( \alpha, \beta \) and \( \gamma \), we create synthetic measurements from \( L \) sets of parameters \( p(l_i) = \{ f_i, d_i, a_i, r_i, u_i, \psi_i \}, i = 1 \ldots M \). For each \( p(l_i) \), we fit the diffusion tensor to the measurements to obtain the eigenvalues \( \lambda_{i1}, \lambda_{i2} \) and \( \lambda_{i3} \). We can rewrite Equations A.1, A.2 and A.3 to include all the sets of parameters:

\[
\begin{pmatrix}
  d_1 \\
  d_2 \\
  \vdots \\
  d_M
\end{pmatrix} =
\begin{pmatrix}
  1 & \lambda_{11} & \lambda_{12} & \lambda_{13} & \frac{\lambda_{11}}{\lambda_{12}} & \frac{\lambda_{11}}{\lambda_{13}} & \frac{\lambda_{12}}{\lambda_{13}} \\
  1 & \lambda_{21} & \lambda_{22} & \lambda_{23} & \frac{\lambda_{21}}{\lambda_{22}} & \frac{\lambda_{21}}{\lambda_{23}} & \frac{\lambda_{22}}{\lambda_{23}} \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  1 & \lambda_{M1} & \lambda_{M2} & \lambda_{M3} & \frac{\lambda_{M1}}{\lambda_{M2}} & \frac{\lambda_{M1}}{\lambda_{M3}} & \frac{\lambda_{M2}}{\lambda_{M3}}
\end{pmatrix}
\begin{pmatrix}
  \alpha_0 \\
  \alpha_1 \\
  \alpha_2 \\
  \alpha_3 \\
  \alpha_4 \\
  \alpha_5 \\
  \alpha_6
\end{pmatrix}
\]

\( S_d = LA \) \hspace{1cm} (A.5)

\[
\begin{pmatrix}
  \log a_1 \\
  \log a_2 \\
  \vdots \\
  \log a_M
\end{pmatrix} =
\begin{pmatrix}
  1 & \log \lambda_{11} & \log \lambda_{12} & \log \lambda_{13} & \log d_1 & \log r_1 \\
  1 & \log \lambda_{21} & \log \lambda_{22} & \log \lambda_{23} & \log d_2 & \log r_2 \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  1 & \log \lambda_{M1} & \log \lambda_{M2} & \log \lambda_{M3} & \log d_M & \log r_M
\end{pmatrix}
\begin{pmatrix}
  \beta_0 \\
  \beta_1 \\
  \beta_2 \\
  \beta_3 \\
  \beta_4 \\
  \beta_5
\end{pmatrix}
\]

\( S_a = L'B \) \hspace{1cm} (A.6)
Figure A.1: Creating an estimate for $\psi$ and $u$.

$$
\begin{bmatrix}
\log r_1 \\
\log r_2 \\
\vdots \\
\log r_M
\end{bmatrix} =
\begin{bmatrix}
1 & \log \lambda_{11} & \log \lambda_{12} & \log \lambda_{13} & \log d_1 & \log a_1 \\
1 & \log \lambda_{21} & \log \lambda_{22} & \log \lambda_{23} & \log d_2 & \log a_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & \log \lambda_{M1} & \log \lambda_{M2} & \log \lambda_{M3} & \log d_M & \log a_M
\end{bmatrix}
\begin{bmatrix}
\gamma_0 \\
\gamma_1 \\
\gamma_2 \\
\gamma_3 \\
\gamma_4 \\
\gamma_5
\end{bmatrix}
$$

(A.7)

From this, we calculate least square estimates for $A$, $B$, and $\Gamma$:

$$
A = L^{-1}S_d \quad B = L'^{-1}S_a \quad \Gamma = L''^{-1}S_r
$$

(A.8)
Initialization of $\psi$ and $u$

To find initial estimates for parameters $u$ and $\psi$ for our fanning model for voxel $l$, we use Equation 6.1:

$$
p = R(u, \psi) \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}.
$$

We assume that $\hat{p} = p/\|p\| = p/r$, which corresponds to the fibre orientation of the central voxel in the voxel grid, can be approximated by $e_1$, the primary eigenvalue of the diffusion tensor that fits the measurements of voxel $l$. By dividing both sides of Equation 6.1 by $r$, we have

$$
e_1 = R(u, \psi) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{(A.9)}
$$

Similarly, we assume

$$
e_2 = R(u, \psi) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{(A.10)}
$$

for $e_2$, the second eigenvector of the tensor. To find the rotation $R(u, \psi)$ that satisfies Equations A.9 and A.10, we perform the inverse rotation of $R(u, \psi)$ in two stages, as shown in Figure A.1. $R_1(u_1, \psi_1)$ rotates $e_1$ to $[1, 0, 0]^T$; the result of this rotation is subsequently rotated by $R_2(u_2, \psi_2)$. For rotation $R_1$, the vector that $e_1$ is rotated about is

$$
u_1 = e_1 \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{(A.11)}
$$

and the angle of rotation is

$$
\psi_1 = \arccos \left( e_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right). \quad \text{(A.12)}
$$
Similarly, for $R_2$, we will have

\[
\mathbf{u}_1 = \mathbf{e}'_2 \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \psi_2 = \arccos \left( \mathbf{e}'_2 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right),
\]

where $\mathbf{e}'_2 = R_1 \mathbf{e}_2$. From this, we can now find $R = (R_2 R_1)^{-1}$. The real eigenvector of $R$ gives $\mathbf{u}$, and $\psi$ is the inverse cosine of the corresponding eigenvalue.
Bibliography


[78] S. Nedjati-Gilani, G. J. M. Parker, D. C. Alexander. “Mapping the Number of Fibre Orientations per Voxel in Diffusion MRI.” In: Proceedings of the 14th Annual Meeting of the International Society for Magnetic Resonance in Medicine, 2006;3169


