# SIGNAL PROCESSING IN DIFFUSION MRI: HIGH QUALITY SIGNAL RECONSTRUCTION

BARTOSZ P. NEUMAN, MSc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

DECEMBER 2014

To my wife, parents, and family.

# Abstract

Magnetic Resonance Imaging (MRI) is a medical imaging technique which is especially sensitive to different soft tissues, producing a good contrast between them. It allows for *in vivo* visualisation of internal structures in detail and became an indispensable tool in diagnosing and monitoring the brain related diseases and pathologies. Amongst others, MRI can be used to measure random incoherent motion of water molecules, which in turn allows to infer structural information.

One of the main challenges in processing and analysing four dimensional diffusion MRI images is low signal quality. To improve the signal quality, either denoising algorithm or angular and spatial regularisations are utilised. Regularisation method based on Laplace–Beltrami smoothing operator was successfully applied to diffusion signal. In this thesis, a new regularisation strength selection scheme for diffusion signal regularisation is introduced. A mathematical model of diffusion signal is used in Monte–Carlo simulations, and a regularisation strength that optimally reconstructs the diffusion signal is sought. The regularisation values found in this research show a different trend than the currently used L-curve analysis, and further improve reconstruction accuracy.

Additionally, as an alternative to regularisation methods a backward elimination regression for spherical harmonics is proposed. Instead of using the regularisation term as a low-pass filter, the statistical $t$-test is classifying regression terms into reliable and corrupted. Four algorithms that use this information are further introduced. As the result, a selective filtering is constructed that retains the angular sharpness of the signal, while at the same time reducing corruptive effect of measurement noise.

Finally, a statistical approach for estimating diffusion signal is investigated. Based on the physical properties of water diffusion a prior knowledge for the diffusion signal is constructed. The spherical harmonic transform is then formulated as a Bayesian regression problem. Diffusion signal reconstructed with the addition of such prior knowledge is accurate, noise resilient, and of high quality.

# Publications

## Conference Proceedings

- Neuman, B.P. and Tench, C. and Bai, L. Regularisation vs elimination: How to accurately reconstruct fibre orientation distribution function, *International Conference on Medical Imaging Using Bio-Inspired and Soft Computing*, 2013.
- Neuman, B.P. and Tench, C. and Bai, L. On computation of diffusion and fibre orientation distribution functions in high angular resolution diffusion imaging, *Medical Image Understanding and Analysis*, 2012.
- Neuman, B.P. and Tench, C. and Bai, L. Laplace–Beltrami regularisation for diffusion weighted imaging, *Medical Image Understanding and Analysis*, 2012.
- Neuman, B.P. and Tench, C. and Bai, L. Direct reconstruction of fibre orientation using discrete ground truth interpolation, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2012.
- Neuman, B.P. and Bai, L. and Tench, C. Reliably estimating the diffusion orientation distribution function from high angular resolution diffusion imaging data, *Medical Image Understanding and Analysis*, 2011.

## Journals

- Neuman, B.P. and Tench, C. and Bai, L. Tikhonov regularisation in diffusion signal estimation, *Annals of the BMVA*, 2013.

# Acknowledgements

I would like to thank my PhD supervisors, Dr Li Bai and Dr Christopher Tench, for supporting me during these past three years.

This research has been supported by the European Commission under the Medical Imaging using Bio-inspired and Soft Computing (MIBISOC) Marie Curie Initial Training Network (contract No. 238819).

Finally, I would like to thank my colleagues, fellow residents of B39 office. You were always (except when I was alone in the office) there to answer my questions, however trivial they had been. Without you and this constant exchange of ideas (as well as problems) this thesis would have taken considerably more time.

# Contents

# List of Figures

# 1

# Introduction

Magnetic Resonance Imaging (MRI) is a medical imaging technique which allows for *in vivo* visualisation of internal structures in detail. It is especially sensitive to different soft tissues, producing a good contrast between them. Since 1973 MRI has become indispensable in diagnosing and monitoring diseases and pathologies. Over the course of time various MRI applications have been introduced, some of which are linked to the diffusion of water in the brain. Diffusion imaging was first introduced for measuring the random incoherent motion of water molecules, and quickly became a reliable tool for estimating local fibre orientations and white matter tracts (tractography).

Unlike a typical MRI scan, a diffusion weighted scan produces few dozen different slices of the same brain section. Each slice is obtained using a different diffusion encoding gradient, which allows to measure diffusion in the direction it represents. The final image contain hundreds of thousands of voxels, each containing a sparse set of signal samples distributed in a three dimensional (3D) space.

To make use of this data, sparse and noisy signal within each voxel has to be analysed using either some parametric model (*e.g.* ball and stick, tensor model) or further transformed into diffusion orientation distribution function (ODF). It is then possible to derive clinically usable information (like fractional anisotropy or mean diffusivity) or to infer fibre orientations and volume fractions within voxels. In the latter case the extracted fibre structures can be used to perform tractography resulting in a full connectivity map of the brain.

The estimation of the diffusion signal, either through a parametric or non-parametric model, is the initial step of diffusion signal analysis (eddy current correction, and brain extraction are considered as preprocessing). As the measured signal is of poor quality (low signal to noise ratio), those estimates might be inaccurate. More importantly though, any errors introduced in the estimation step will be propagated further and hamper other processing and analysis. It is therefore imperative that estimated diffusion profiles are of highest accuracy. In this thesis, novel diffusion signal processing and reconstruction methods will be presented.

## 1.1 Research Aim

The focus of this research is on non-parametric methods, as they do not rely on a model selection. Additionally, with a sufficient number of samples, non-parametric methods can resolve any potential fibre configuration within a voxel. Unfortunately, at the same time they are susceptible to noise, and with low SNR can become inaccurate. The aim of this research therefore, is to develop a non-parametric and noise resistant algorithm for diffusion signal processing and reconstruction.

## 1.2 Software

The algorithms presented in this thesis were implemented using Matlab 2013a (`http://www.mathworks.co.uk/`). For visualisation of results Matlab 2013a, DSI Studio v.1 (`http://dsi-studio.labsolver.org/`), and TrackVis v0.6 (`http://www.trackvis.org/`) were used. Human

brain images were pre-processed using FMRIB Software Library v5.0 (`http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/`).

## 1.3 Thesis Outline

In Chapter 2, magnetic resonance imaging, water diffusion, and methods for measuring and processing of diffusion signal are introduced.

Chapter 3 describes a new algorithm for selecting regularisation strength in diffusion signal reconstruction. The improved regularisation parameter is based on Monte–Carlo simulations, correctly reacting to the change in diffusion weighting applied and signal-to-noise ratio.

Chapter 4 investigates the statistical significance of spherical harmonic basis functions used to reconstruct the measured signal. A new method that uses backward elimination and statistical testing is presented, improving the reconstruction accuracy. Furthermore, four methods to combine the regularisation scheme introduced in Chapter 3 with statistical inference are proposed and analysed.

Chapter 5 introduces a novel reconstruction method in which prior knowledge build from a Monte–Carlo simulations is utilised. The linear regression used in Chapters 3 and 4 is substituted with a Bayesian regression.

Chapter 6 summarises methods and results described in Chapters 3, 4, and 5 and outlines future work.

Appendix A contains the Matlab code used to generate diffusion signal for any diffusion weighting gradients, any fibre configuration, any signal-to-noise ratio, and at any diffusion weighting factor. This code is used in all Monte–Carlo simulations thorough this thesis.

Matlab implementations of contributed algorithms are presented in Appendices B-D. Code snippets important for optimising Tikhonov regularisation are given in Appendix B. Appendix C shows the code for backward elimination method described in Chapter 4. Code relevant to Bayesian regression is given in Appendix D.

# 2

# Background

## 2.1 Magnetic Resonance

Magnetic Resonance Imaging (MRI) is an *in-vivo* imaging technique that can be used to visualise internal structures of the human body. It is based on a nuclear magnetic resonance (NMR) of molecules, a phenomenon described by Bloch [11] and Purcell [67].

When exposed to a strong, constant magnetic field, the average magnetic momentum (spin) of protons align with the direction of that field. This results in a net magnetisation for nuclei that have an unpaired protons or neutrons. By applying electromagnetic field of a certain frequency it is possible to "flip" spins. After the electromagnetic field is turned off, spins rotate with the same frequency that was previously applied, while returning to the previous direction enforced by the constant magnetic field. The frequency of both electromagnetic field and spin rotation is called the resonating frequency, which depends on the number of protons in the atom, and the strength of the external magnetic

field. For hydrogen located in a field strength of 1.5T, this resonating frequency is in the radio frequency (RF) bandwidth.

To use the NMR phenomenon for imaging purposes it is necessary to position the source of the electromagnetic energy. Additional magnetic fields are introduced, and a full 3D spatial encoding is achieved by varying the magnetic frequency, phase, and field strength. The signal is mathematically reconstructed, and finally represented as a 3D greyscale image, the contrast of which depends on the strength of the measured signal.

## 2.2 Brownian Motion

Brownian motion (called molecular diffusion, or just diffusion) is a random thermal motion of all molecules. The molecules are propelled by the thermal energy alone, owning the random movement to the frequent collisions with other molecules.

Molecular diffusion can be characterised by a physical constant $D$, called a diffusion coefficient. The diffusion coefficient defines the average displacement of molecules over a unit time. The rate of diffusion is described as [21]:

$$< r^2 >= 6Dt, \tag{2.1}$$

where $< r^2 >$ is a mean squared displacement of molecules during the time $t$.

In a barrier-free environment an unrestricted, random movement of molecules can occur, and the mean squared displacement is equal in all directions. Whenever a barrier is introduced though, molecule movement perpendicular to the barrier becomes restricted, and the mean squared displacement is dependent on the direction. Figure 2.1 visualises the molecules distribution in an unrestricted (isotropic) and restricted (anisotropic) diffusion.

## 2.3 Diffusion Imaging

Diffusion MRI (dMRI) is a non-invasive imaging technique that allows to measure the movement of water molecules. As human body consists

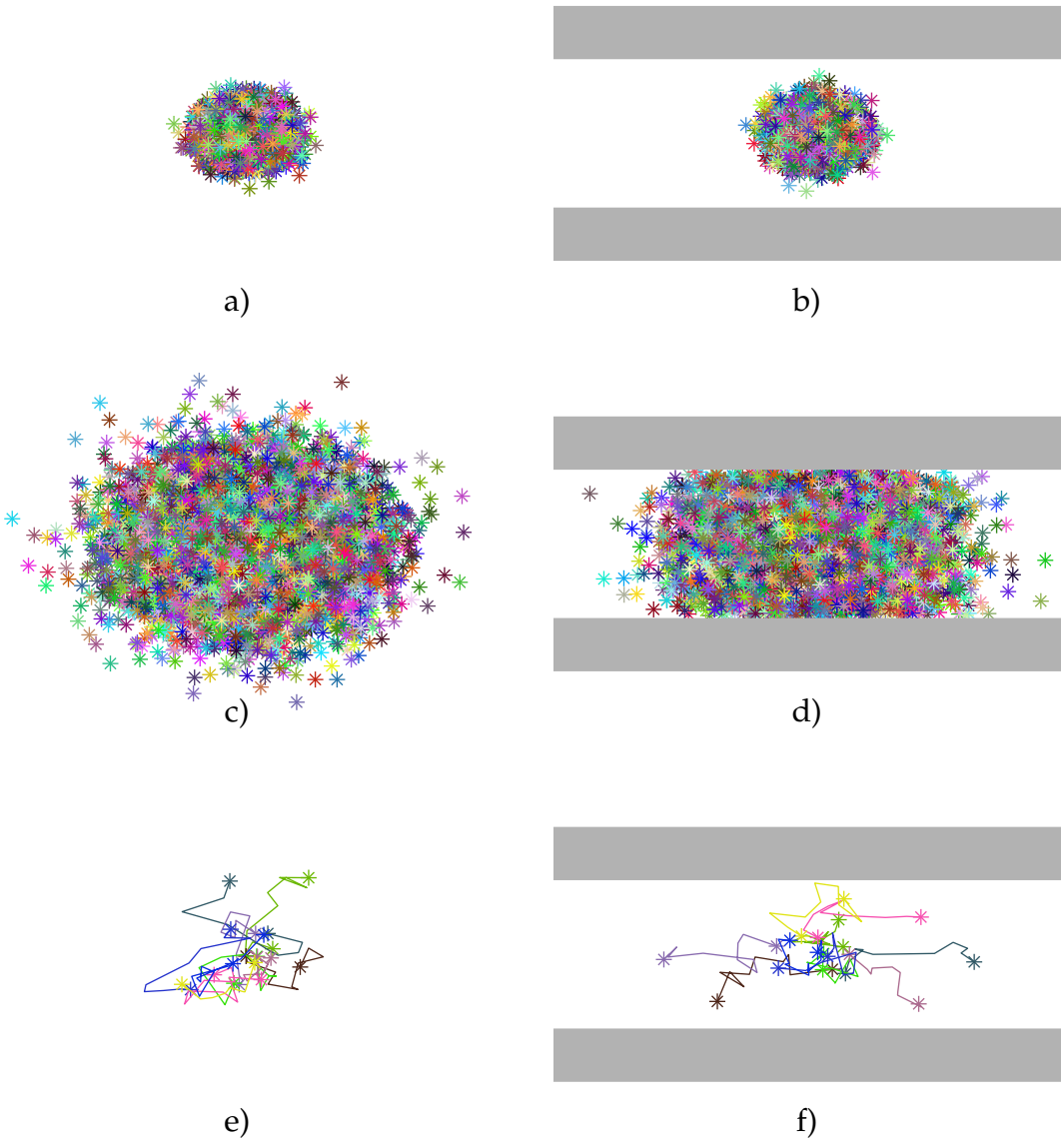



**Figure 2.1:** Simulation of random movement of water molecules. Start position (a,b), end position (c,d), and trace of movement for ten selected molecules (e,f) in isotropic (a,c,e) and anisotropic (b,d,f) media. Gray horizontal line indicates the barrier.

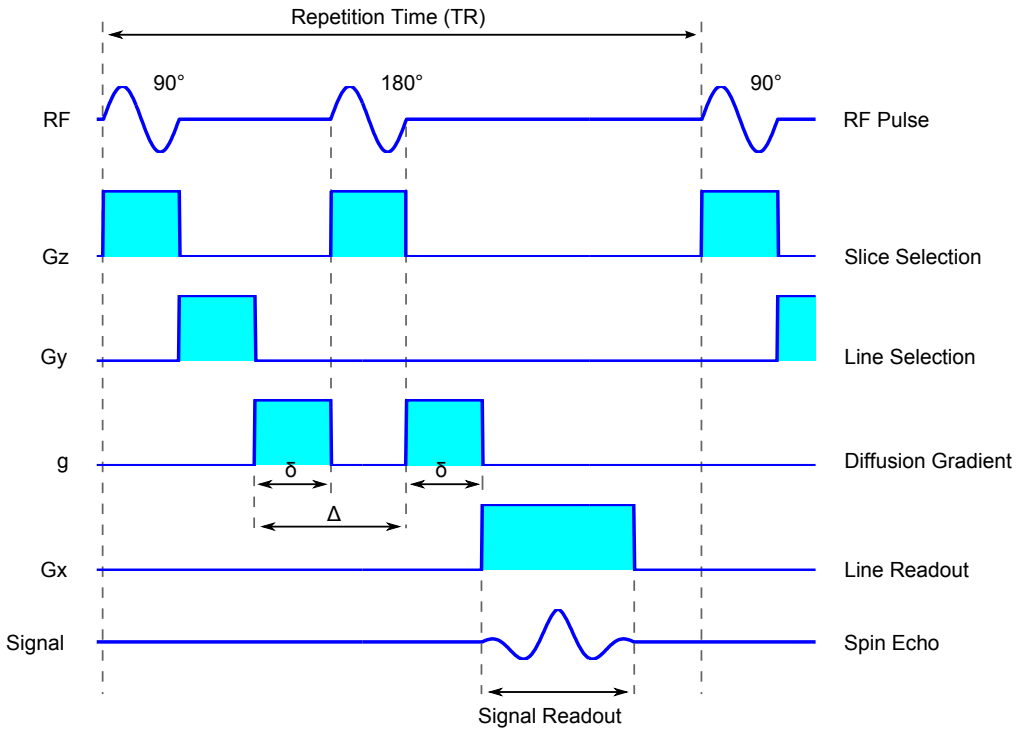



**Figure 2.2:** Schematic representation of Stejskal-Tanner pulse gradient spin echo imaging sequence. Two gradient pulses $g$ are placed on either side of the refocussing pulse 180°. First 90° pulse flips spins in transverse plane and first gradient pulse $g$ shifts spins phase. Second pulse gradient, after the 180° refocussing pulse, shifts spins phase in the opposite direction. If there is no change in spins location, the first phase shift is cancelled by the second. Any change in spin location during $\Delta$ will result in a T2 signal attenuation. Figure recreated based on [46].

mostly of water, the measured signal comes almost exclusively from hydrogen diffusion. Applied to the brain, dMRI can be used to reconstruct the white matter tracts [5], study brain connectivity [72], and detect early changes in the cerebral tissue [44].

In dMRI, in addition to the homogeneous magnetic field, two additional pulsed field gradients are introduced (Figure 2.2, line $g$). They have the same magnitude but are in opposite directions. The first gradient linearly varies the homogeneous field, so that the protons rotate at different frequencies at different locations. The second gradient, refocuses the spins. If there was any movement during the time between those two pulses ($\Delta$), the refocussing will not be ideal, resulting in a reduction of the measured signal.

Figure 2.2 shows the diffusion weighting sequence used by Stejskal and Tan-

ner [73]. With the use of rectangular pulsed gradients, the reduction of the spin echo signal can be expressed as:

$$\frac{E}{E_0} = \exp\left[-\gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right) D\right] . \tag{2.2}$$

In this equation, $E_0$ is the signal intensity without the diffusion weighting, $E$ is the signal with the applied gradient, $\gamma$ is the gyromagnetic ratio, $G$ is the strength of the gradient pulse, $\delta$ is the duration of the pulse, $\Delta$ is the time between the two pulses, and $D$ is the diffusion coefficient. For convenience, scanner-related parameters are combined into a single scalar $b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right)$ called diffusion weighting coefficient. The simplified Stejskal–Tanner equation becomes thus:

$$E = \frac{\exp\left(-bD\right)}{E_0} . \tag{2.3}$$

Perhaps the most important breakthrough in diffusion MRI occurred in 1990, when the diffusion weighted signal in white matter was found to be relative to the orientation of the diffusion weighting gradient [51]. The diffusion signal could no longer be explained with Equation 2.3 alone and a tensor based mathematical model was established. The discovery also opened a path for new applications, like anisotropy measures, estimating orientation of white matter tracts, and inferring the fibre tracts.

## Visualisation of diffusion signal

There are two practical ways of visualising spherical 3D data such as diffusion signal – either through a greyscale mapped sphere (Figure 2.3a), or through a sphere deformation (Figure 2.3b). Early works used the former, in which sections of sphere were shaded based on the signal magnitude in that direction. With the advent of tensor imaging, the signal (tensor) can be visualised using an ellipsoid, and multiple tensor models and higher-order tensor models can be visualised using an antipodally symmetric spherical function. Tensor imaging introduced also a direction-specific colouring, where the principal direction of the tensor (X, Y, Z) was transformed into RGB space.

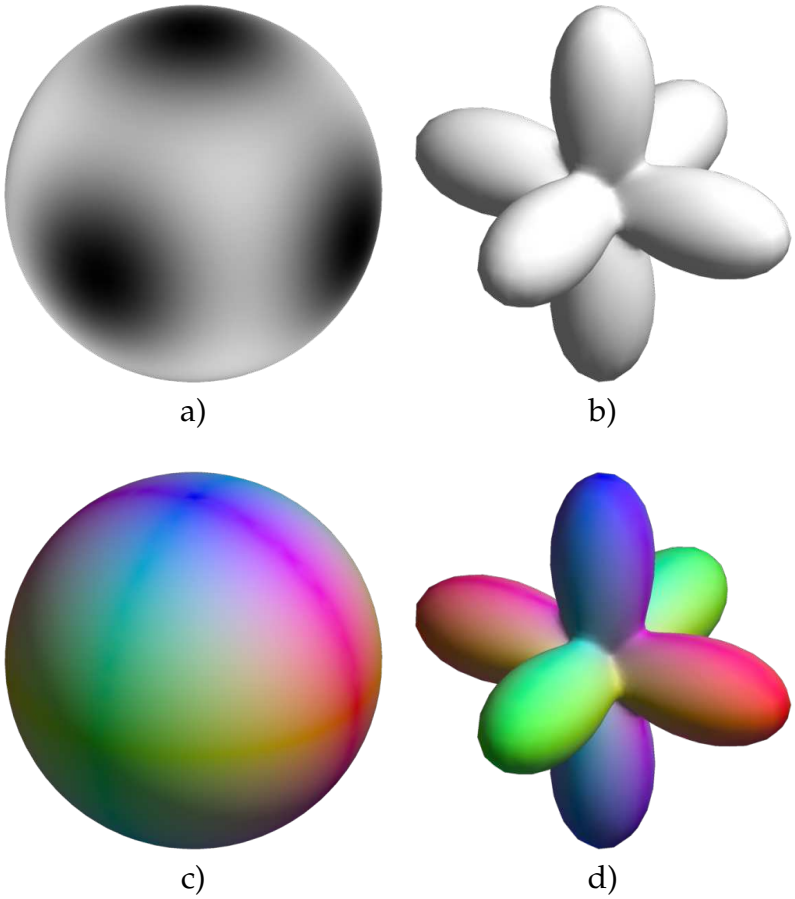



**Figure 2.3:** Visualisation schemes for spherical signal: sphere shading (a) and deformation (b) visualisation schemes. By combining sphere deformation with direction-dependent colour coding (c), the shapes become more intuitive to interpret (d).

A spherical signal with high magnitude in X, Y, and Z directions is visualised in Figures 2.3a, 2.3b, and 2.3d. The sphere shading method is shown in Figure 2.3a, with the greyscale intensity linearly dependent on the signal magnitude. Similarly, the shape deformation used in Figure 2.3b is linearly scaled by signal magnitude. Finally, Figure 2.3d represents a modification of sphere deformation scheme, where each sample is colour coded based on its location ($XYZ \Rightarrow RBG$ colour mapping shown in Figure 2.3d).

In this thesis, the modified visualisation method is used (Figure 2.3d): 3D shapes are colour coded, and the sphere is deformed. Colour coding is based on the direction of the sample, and the direct mapping of XYZ coordinates onto RGB colour space is used. In addition, the sphere is deformed accordingly to the signal magnitude in that given direction. All figures depicting 3D shapes will follow this format.

### 2.3.1 Diffusion Tensor Imaging

As explained in Section 2.2, the apparent diffusion coefficient, and consequently the diffusion weighed signal depends on the direction of a measurement. In an isotropic environment the signal will be equal in all directions, while in an anisotropic environment it will depend on the physical barriers present.

By measuring a diffusion weighted signal in different directions, the 1D data can be extended to 3D. Diffusion imaging (DTI) reconstructs the 3D diffusion profile using a 2nd order tensor [6, 7]. The measured diffusion signal $E_i$ is expressed as:

$$E_i = E_0 \exp(-b\vec{g}_i^{\mathrm{T}}\mathbf{D}\vec{g}_i), \tag{2.4}$$

with $\vec{g}_i$ being the direction of the $i$th diffusion sensitising gradient, and $\mathbf{D}$ a diffusion tensor.

Since the measured signal cannot be negative, $\mathbf{D}$ is a positive definite matrix. Additionally, as water molecules are uncharged, $\mathbf{D}$ is symmetric, meaning it has 6 unique elements. Therefore, in order to determine tensor parameters 6 non-collinear DW images are required. To reduce the effect of low signal-to-noise ratio (SNR) of a single image [36] often more than 6 images are acquired. Acquired images should be evenly distributed in space in order to make the statistical properties of the computed parameters rotationally invariant, and to reduce noise bias [29, 36, 37].

Of particular interest in diffusion MRI is the 3D distribution of water diffusion, which is necessary to infer information about barriers (or lack of such) within the medium. This distribution is characterised by orientation distribution function (ODF). Formally, the diffusion ODF $\Psi$ is defined as the Funk–Radon transform of the diffusion signal $E$:

$$\Psi(\mathbf{u}) = \int_{\mathbf{q}\perp\mathbf{u}} E(\mathbf{q})dq, \tag{2.5}$$

where both $\mathbf{u}$ and $\mathbf{q}$ are unit directions. Figure 2.4 shows relationship between $E$ and $\Psi$.

Typically, the applied diffusion weighting varies between 1000 s/mm$^2$ and 5000 s/mm$^2$. The higher $b$ is, the more sensitive to molecular motion the image

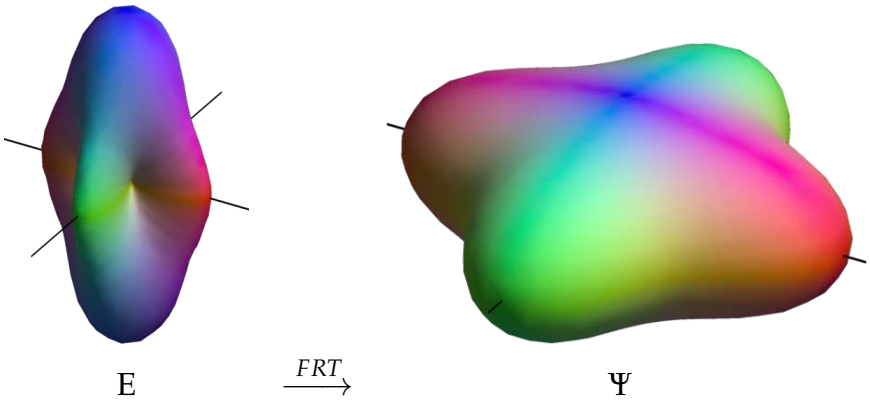



**Figure 2.4:** Relationship between the diffusion signal (left) and ODF (right). Simulation of a 90° fibre crossing using $b = 3000$ s/mm$^2$ and 8th order SH series.

becomes. At the same time, with higher $b$ values the signal amplitude becomes lower and the image is more affected by noise. Figure 2.5 shows two 3D profiles of diffusion signal, simulated with $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, as well as their corresponding ODFs. Maximal simulated diffusion signal magnitude was 0.6065 for low, and 0.1350 for high $b$-value.

For the single fibre configuration, DTI can be used to accurately model the ODF. The single diffusion tensor model is inaccurate for more complex fibre configurations, such as multiple fibre crossing, kissing, and bending [3, 4, 25]. The diffusion signal within voxels containing multiple fibres can be better explained as a sum of single tensors responses [2, 25]:

$$E_i = E_0 \sum_{k=1}^{n} v_k \exp(-b\vec{g}_i^{\mathrm{T}} \mathbf{D}_k \vec{g}_i). \tag{2.6}$$

$\mathbf{D}_k$ is a diffusion tensor of the $k$th fibre and $v_k$ a volume fraction ($\sum_{k=1}^{n} v_k = 1$), which tells how strongly the fibre affects the total diffusion within the voxel. An $n$-fibre crossing will have $6n$ unique parameters, and to be correctly resolved requires at least that many different diffusion weighted images.

Because of implications of Equation 2.6, it is not possible to model a multi fibre response with a DTI. The model is not flexible enough to cope with the higher degree of freedom brought by complex fibre configurations. For example a DTI fitted to a 90$^o$ fibre crossing will have infinite number of peaks (Figure 2.6a), while a 60$^o$ fibre crossing will have one peak instead of two (Figure 2.6c).

Unfortunately, without additional constraints the multi tensor model (MTM) is

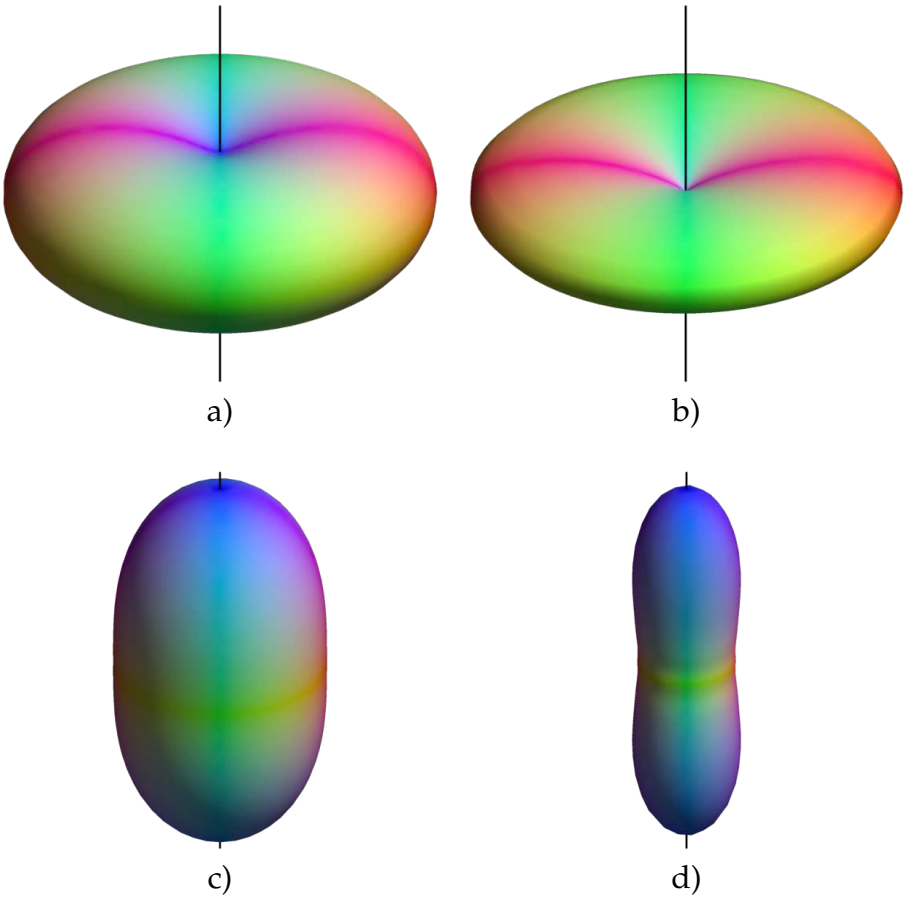



**Figure 2.5:** Visualisation of diffusion signal (top) and ODF (bottom) for low (left, $b = 1000\ \mathrm{s/mm^2}$) and high (right, $b = 3000\ \mathrm{s/mm^2}$) diffusion weighting applied. Higher $b$-values are more sensitive to molecular displacement, resulting in a sharper ODF.

numerically unstable even for $N = 2$ as multiple $\mathbf{D}_1$ and $\mathbf{D}_2$ can satisfy Equation 2.6. As such, only two-tensor models with positive semi-definite and symmetrical constraints are of possible use. With $N > 2$ the situation becomes much worse, and as of now a 3 tensor model has not yet been used.

Alternatively to expanding $N$, Özarslan proposed to increase $\mathbf{D}$ in size [60]. The method, called higher order tensor (HOT), is stable regardless of the size of $\mathbf{D}$. Like MTM though, it has to ensure that $\mathbf{D}$ is symmetric and positive semi-definite.

### Data Visualisation and Assessment

One of the early problems with using DTI was how to visualise and assess the tensor data. With diffusion MRI, the measured signal $E$ is dependent on the

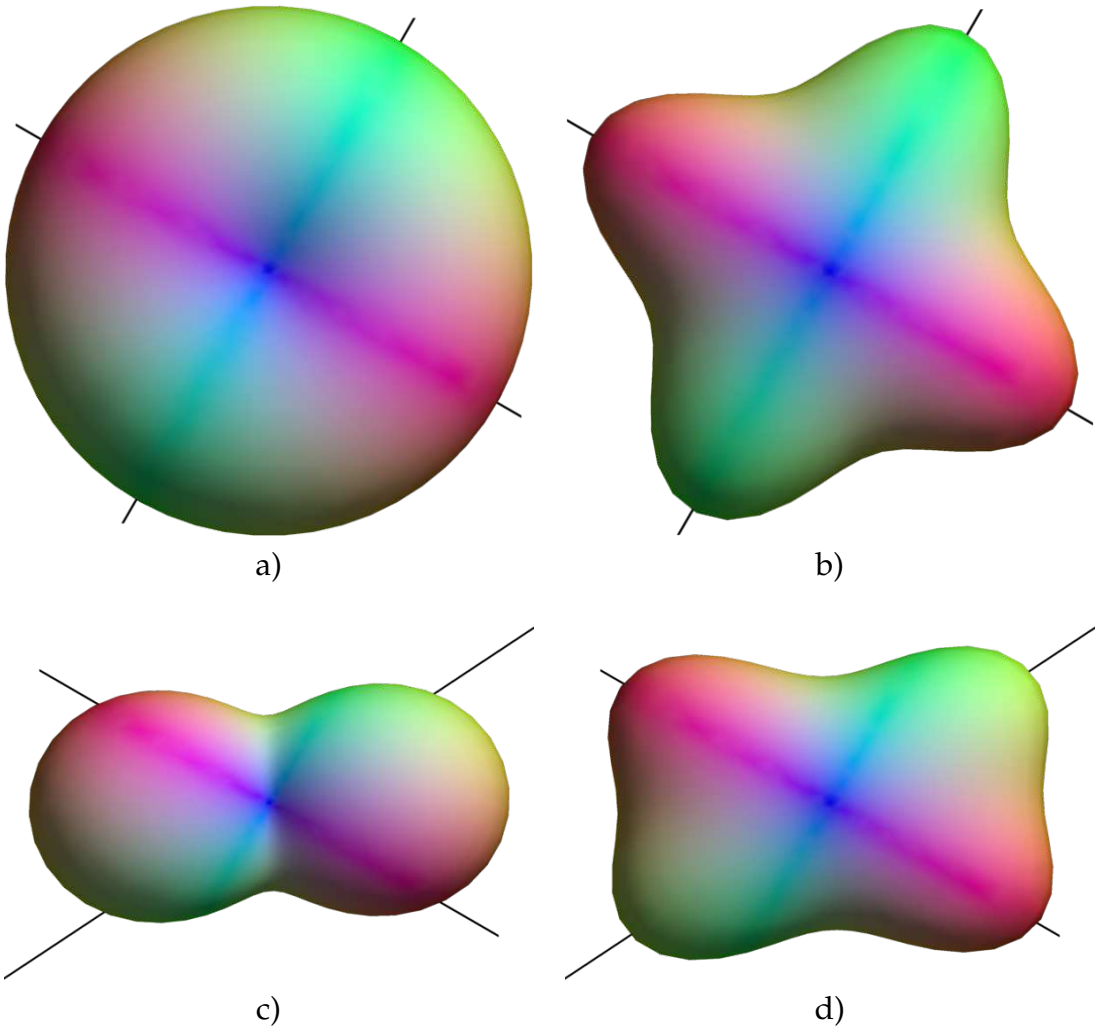



**Figure 2.6:** Visualisation of a 90° (a) and 45° (c) fibre crossings modelled using a DTI, and their true ODFs (b and d); $b = 3000$ s/mm$^2$, view from top.

measuring direction which makes it rotationally dependent. Additionally, the diffusion tensor **D** is of high dimensionality, which makes it difficult to directly use for the whole image. As such neither could be directly used to visualise or assess the data, and the need for simplified tensor analysis emerged.

The 3x3 matrix **D** itself, can be decomposed into 3 eigenvalues ($\lambda_1$, $\lambda_2$ and $\lambda_3$) and eigenvectors ($\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$) that show the strength and the direction of molecular diffusion. The eigenvalues can be used to create 3D greyscale maps (Figure 2.7a-c). By transforming $x$, $y$ and $z$ coordinates to red, green and blue colour, the primary eigenvector ($\vec{v}_i$ corresponding to the highest eigenvalue $\lambda_i$) can be used to create an 3D, RGB image (Figure 2.7d).

The eigenvalues $\lambda_1$, $\lambda_2$ and $\lambda_3$ can be used to compute rotational invariant

diffusion anisotropy measures, like mean diffusivity (Equation 2.7), fractional (Equation 2.8) and relative (Equation 2.9) anisotropy, and volume ratio (Equation 2.10). Each of those measures range from 0 to 1, and can be used to create a 3D grey scale map (Figure 2.7e-h).

Mean diffusivity (MD) summarises the average diffusion within the voxel, and is computed as the average of all three eigenvalues of **D**:

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} . \tag{2.7}$$

Larger MD values are found in cerebrospinal fluid (CSF), while smaller in more organised structures of the brain. As a consequence, MD maps can be used to remove the CSF voxels from the image. MD maps can also be used to study CSF-related diseases [52].

Fractional anisotropy (FA), perhaps due to its intuitive interpretation, is the most widely used diffusion anisotropy measure in dMRI research. The magnitude of FA is dependent on the anisotropy of **D** [40], and is given by:

$$FA = \frac{\sqrt{3}\sqrt{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} . \tag{2.8}$$

FA varies from 0 for isotropic ($\lambda_1 = \lambda_2 = \lambda_3$) to 1 for anisotropic ($\lambda_1 \geq \lambda_2 = \lambda_3$) diffusion, and is sensitive to low anisotropy [74].

Relative anisotropy (RA) is a normalised version of FA [45]. It represents the ratio of the anisotropic part of **D** to its isotropic part:

$$RA = \frac{\sqrt{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{3}MD} . \tag{2.9}$$

RA ranges from 0 for isotropic to $\sqrt{2}$ for anisotropic diffusion, and scales linearly with anisotropy [74].

Volume ratio is defined as a ratio of the ellipsoid volume and sphere of MD-diameter [83]:

$$VR = \frac{\lambda_1\lambda_2\lambda_3}{MD^3} \tag{2.10}$$

VR changes from 1 for isotropic to 0 for extremely anisotropic diffusion [74].

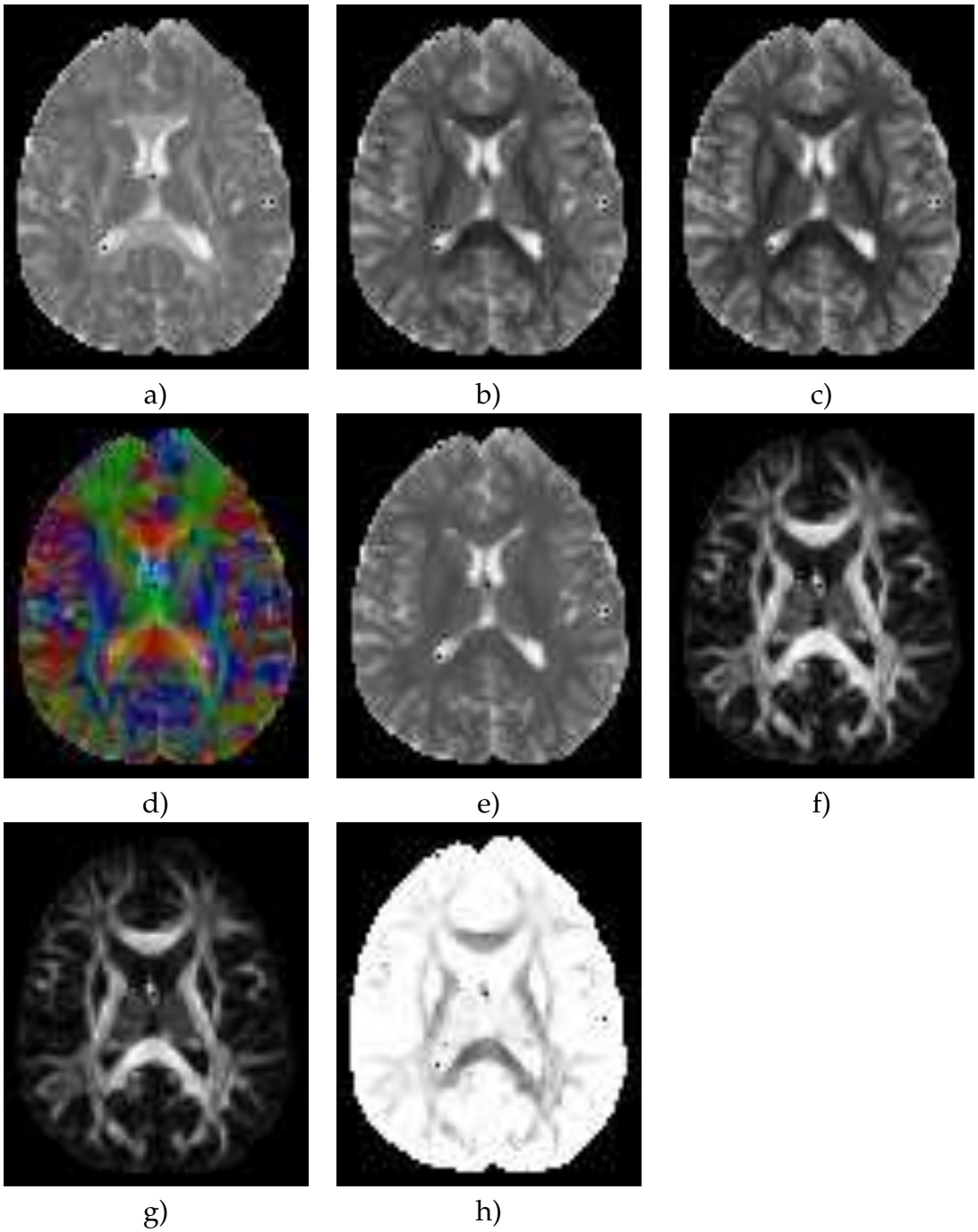



**Figure 2.7:** Visualisation of information derived from eigenvalues and primary eigenvector: primary eigenvalue (a), secondary eigenvalue (b), tertiary eigenvalue (c), RGB visualization of primary eigenvector (d), mean diffusivity (e), fractional anisotropy (f), relative anisotropy (g), and volume ratio (h).

### 2.3.2 High Angular Resolution Diffusion Imaging

Both MTM and HOT models have more parameters to estimate than classical DTI, and as such require more data to work with. The number of acquisitions is increased, and the diffusion is measured in many non-collinear directions. The resulting high angular resolution diffusion image [80] (HARDI) can be used to not only solve MTM and HOT models, but also to perform DTI analysis.

In HARDI analysis, both number and direction of spherical sampling gradients have to be selected. The more of the directions that are sampled the better the reconstruction accuracy becomes. The upper limit is decided by the acquisition time, as every volume requires extra time to scan. Finally, the angular resolution of HARDI is balanced with other conditions, like desired spatial resolution, diffusion weighting applied, and single volume SNR. In a clinical environment, depending on application, this translates to 30 to 60 diffusion gradients being used, with a scanning time up to 15 minutes.

Like in DTI, the gradient encoding scheme should be chosen in a way that reduces the noise bias, and keeps the statistical properties of the computed parameters rotationally invariant.

### 2.3.3 Q-Ball Imaging

As an alternative to tensor analysis (DTI, HOT, MTM), Tuch introduced a model-independent HARDI reconstruction scheme [79, 81] called Q-ball imaging (QBI). The method was based on a numerical approach to calculate the equator integral from Equation 2.5. Since the equator points do not always coincide with the diffusion sampling points, the HARDI signal had to be first interpolated using the spherical radial basis function [22]. Anderson [4], Hess [31, 32], and Descoteaux [18] have independently and in parallel developed analytical solutions for the ODF reconstruction in QBI using spherical harmonic series [59] and Funk–Hecke theorem.

With a model-free representation of ODF it is now possible to calculate more accurate fractional anisotropy. The generalised fractional anisotropy (GFA) is

an extension to FA, and is computed as a standard deviation divided by the root mean square of the ODF [79]:

$$GFA = \sqrt{\frac{n\sum_{i=1}^{n}(\mathbf{\Psi}(u_i) - \bar{\mathbf{\Psi}})^2}{(n-1)\sum_{i=1}^{n}\mathbf{\Psi}(u_i)^2}}\ , \tag{2.11}$$

where $\mathbf{\Psi}(u_i)$ is the ODF sampled in $u_i$ direction, $\bar{\mathbf{\Psi}}$ is the mean of the ODF, and $n$ is the number of ODF samples.

## 2.4 HARDI Processing

### 2.4.1 Spherical Harmonics Analysis

Spherical harmonics (SHs) are globally supported, complete, orthonormal spherical basis function. They allow for a non-parametric analysis of the diffusion signal, and can be seen as the extension of Fourier basis functions to the sphere. The basis functions, defined as a solution to Laplace equation, are given by:

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)\exp(im\phi)\ , \tag{2.12}$$

with $P_l^m$ being the associated Legendre function of order ($l \geq 0$) $l$ and phase $m$ ($-l \geq m \geq l$)[1], and $\theta$ and $\phi$ a colatitude and longitude.

The coefficients of the series are found through the SH transform:

$$c_{lm} = \int_0^{2\pi}\int_0^{\pi} E(\theta,\phi)Y_l^m(\theta,\phi)\sin\theta d\theta d\phi\ . \tag{2.13}$$

In practice, the number of signal samples is limited, and rarely equal to the number of basis functions. As such, instead of interpolating the measured signal, the signal is approximated with a truncated $n$th order SH series [3]:

$$\hat{E}(\theta,\phi) = \sum_{l=0}^{n}\sum_{m=-l}^{l} c_{lm}Y_l^m(\theta,\phi)\ . \tag{2.14}$$

[1] In Mathematics called band and degree.

Here $\hat{E}(\theta, \phi)$ is the approximation of signal $E$ measured in direction $\theta, \phi$. The signal is approximated with an $n$th order SH series, $Y_l^m$ is a SH function defined in Equation 2.12, and $c_l^m$ is a SH coefficient of order $l$ and band $m$.

Classical SH transform is no longer used to find SH coefficients and instead, an efficient method based on matrix operations is preferred. Having infinite number of samples, and infinite number of SH basis functions the diffusion signal can be expressed as:

$$\mathbf{E} = \mathbf{Yc} \, . \tag{2.15}$$

With a sufficient SNR to assume a normal distribution of noise[2], the SH representation of the signal can be found by minimising the least squares difference:

$$\arg\min_{\mathbf{c}} \|\mathbf{Yc} - \mathbf{E}\|^2 \, , \tag{2.16}$$

which can be solved using the ordinary least squares (OLS) method:

$$\mathbf{c} = (\mathbf{Y^T Y})^{-1} \mathbf{Y^T E} \, . \tag{2.17}$$

Having the SH coefficients, it is now possible to recreate the missing sections of the spherical signal. The inverse transform, which simply becomes another matrix multiplication, is used to reconstruct the signal:

$$\hat{\mathbf{E}} = \mathbf{Y}_{large} \mathbf{c} \, , \tag{2.18}$$

where $\mathbf{Y}_{large}$ is a SH design matrix generated using a dense spherical sampling.

The diffusion weighting gradients are constant across the whole image, and therefore SH design matrix $\mathbf{Y}$ does not change. This allows to compute the term $(\mathbf{Y^T Y})^{-1} \mathbf{Y^T}$ only once, and use it over the whole image significantly speeding up the processing time.

Finally, the diffusion signal that is analysed is antipodally symmetric, which allows to remove antipodally asymmetric odd-order basis functions from Equations 2.14-2.17. The total number of basis functions used to describe a signal is thus $(n+1)(n+2)/2$ and should be no greater than the number of acquired samples (number of diffusion weighting gradients used).

[2]Otherwise a maximum likelihood method should be used.

Having a SH representation of the signal, the diffusion ODF can be directly estimated [4, 18, 31, 32] by a linear transformation of coefficients $c$:

$$\Psi(\theta,\phi) \approx \sum_{l=0}^{n} \sum_{m=-l}^{l} 2\pi c_{lm} P_l(0) Y_l^m(\theta,\phi) \; . \tag{2.19}$$

Here $P_l(0)$ being the associated Legendre function of order $l$ evaluated at 0. Combining Equations 2.17 and 2.5 yields:

$$\mathbf{\Psi} = 2\pi \mathbf{P}(\mathbf{Y^T Y})^{-1} \mathbf{Y^T Y E} \; , \tag{2.20}$$

where $\mathbf{P}$ is a diagonal matrix with elements $P_l(0)$. Again, the term $\mathbf{P}(\mathbf{Y^T Y})^{-1}\mathbf{Y^T Y}$ needs to be computed only once.

The ODF should be relatively smooth with a few maxima oriented along the direction of underlying fibres. Unfortunately, due to noise the ODF has a lot of sharp spikes and needs to be smoothed. Noise related peaks can be reduced by filtering SH coefficients [76], including a regularization scheme in the signal approximation [19, 32], or by selectively removing the noise-infested basis functions [53]. In all cases smoothing of ODF function is produced at the cost of a lower angular resolution.

Finally, the GFA can be directly calculated in the SH domain [13]:

$$GFA = \sqrt{1 - \frac{(c_0^0)^2}{\sum_l \sum_{m=-l}^{l} (c_l^m)^2}} \; . \tag{2.21}$$

### 2.4.2 Fibre Orientation Distribution Function

Fibre orientation density (FOD) and fibre orientation distribution function (FODF) are sharper versions of ODF. The spherical deconvolution introduced by Tournier [76] allows to compute FOD directly from HARDI data. The measured signal $E$ is a convolution of unknown FOD $F$ with the signal $R_E$ coming from a single fibre population (Figure 2.8, top):

$$E = F \otimes R_E. \tag{2.22}$$

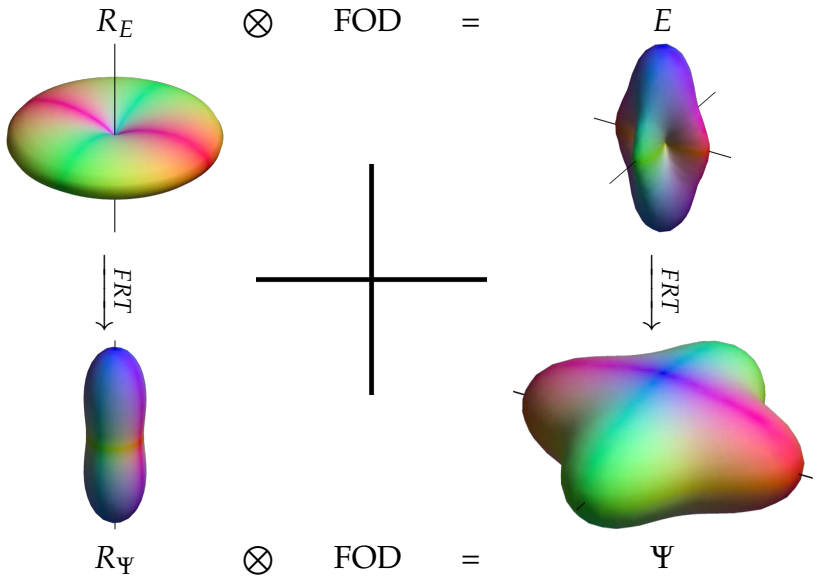



**Figure 2.8:** Relationship between diffusion signal, ODF, and FOD. Diffusion signal for a single fibre (top left) can be converted to ODF (bottom left) using a Funk-Radon Transform. Diffusion signal for a 90° fibre crossing (right top) is a convolution of diffusion signal for a single fibre and the FOD (fibre structure, two delta functions crossing at 90°; middle). The ODF for a 90° fibre crossing (bottom right) is both a convolution of the ODF for a single fibre crossing and the FOD, as well as a Funk-Radon Transform of diffusion signal for a 90° fibre crossing. HARDI and ODF profiles are reconstructed using 8th order SH series; diffusion signal was simulated for $b = 3000\ \mathrm{s/mm^2}$.

The single fibre response $R_E$ is approximated from the most anisotropic voxels [76], or on a voxel-by-voxel basis [4], and represented with rotational harmonics [30]. Since the SH transform is a Fourier transform on the sphere the convolution can be efficiently represented as a matrix multiplication, or linear transformation of SH coefficients:

$$f_l^m = c_l^m / r_l \ . \tag{2.23}$$

Here $f_l^m$ is a FOD spherical harmonic coefficient of $l$ order and $m$ band, and $r_l$ a rotational harmonic coefficient of a single fibre response $R_E$. The rotational harmonic coefficients of the deconvolution kernel are calculated from $m = 0$ coefficients of spherical harmonic representation of delta function $d$ and single fibre response $s$:

$$r_l = s_l^{m=0} / d_l^{m=0} \ , \tag{2.24}$$

With an infinite SH series the signal will deconvolve to a sum of delta functions oriented along the underlying fibre tracts (Figure 2.8, FOD). Unfortunately, the number of samples is limited, and the series expansion of the signal has to be truncated. Additionally, SHs are globally supported complete functions and when truncated are unsuitable to represent sharp signals. As such, using truncated SH series will result in an oblate FOD, as well as introduce unwanted "ringing" effect near the centre of the lobe (Figure 2.9a). Due to the noisy signal reconstruction the deconvolution will pollute the FOD with false peaks (Figure 2.9b).

To remove the ringing, and partially reduce the false FOD peaks that are caused by noise, Schultz proposed deconvolution using a non-ringing cosine power lobe [69]. The deconvolution kernel is built in the same way as in Tournier's spherical deconvolution, but using the cosine power lobe ($\cos^h$):

$$r_l = s_l^0 / p_l^0 \,, \tag{2.25}$$

where $p$ are coefficients of cosine power lobe SH transform, and $h$ is used to select the sharpness of the resulting FOD. Depending on the $h$, signal will be deconvolved to a more oblate (low $h$) or prolate (high $h$) FOD. With higher $h$ though, cosine power lobe kernel starts to behave like delta function kernel, introducing ringing and susceptibility to noise. In practice, a good sharpness-to-artefacts ratio can be achieved with $h$ being two to three times larger than the SH series order used [56]. Figure 2.9c and Figure 2.9d show the 8th order SH series reconstruction of the signal deconvolution to a cosine power lobe with $h$ set to 16.

Another way of acquiring FOD is based on sharpening the diffusion ODF. Using the same spherical deconvolution method, it is possible to deconvolve diffusion ODF to FOD using a single fibre response ODF (Figure 2.8, bottom). For this, Descoteaux provided a formal relationship between ODF and FOD (called fibre ODF, or FODF, as it was derived directly from ODF) [20]. The method, called spherical deconvolution transform (SDT), like cosine power lobe, produces more noise resilient FOD (Figure 2.9e and Figure 2.9f).

Like in an SD, the deconvolution kernel is based on the assumed single fibre response. In SDT though, the deconvolution kernel is built from an ODF profile of a single fibre response. The kernel is defined using $\lambda_1$ and $\lambda_2$, a major and

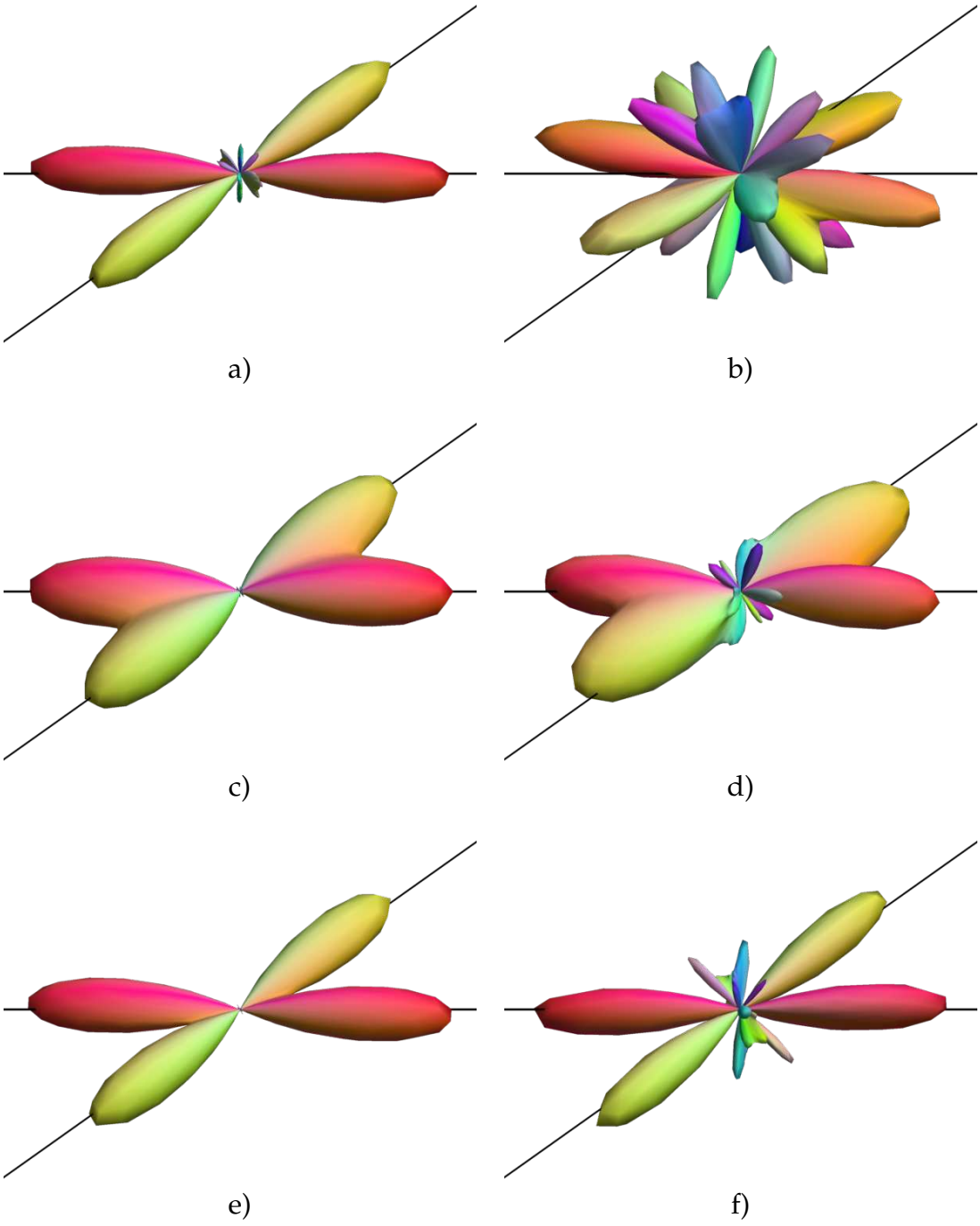



**Figure 2.9:** Spherical deconvolution to a delta function (top row), cosine power lobe (middle row), and spherical deconvolution transform (bottom row), of a simulated 45° fibre crossing; 8th order SH series, $b = 3000$ s/mm$^2$ without (left column) and with (right column) added Rician noise (SNR $= 35$). The spherical deconvolution to a delta function produces sharp FOD profile with ringing (a). In a noisy signal this method introduces numerous false peaks (b). FOD produced with a cosine power lobe kernel and spherical deconvolution transform do not introduce ringing (c and e), and suppress false peaks at the cost of angular resolution (d and f).

minor eigenvalue of a tensor representing assumed single fibre response:

$$r_l = \frac{2\pi}{\sqrt{\sum_i^N \left((\lambda_2/\lambda_1 - 1)z_i^2 + 1\right)}} \sum_i^N \left( P_l(z_i) \left( (\lambda_2/\lambda_1 - 1)z_i^2 + 1 \right) \right). \quad (2.26)$$

With $z_i = -1 + 2i/N$, $N$ should be chosen large enough to densely sample the $< -1, 1 >$ range.

A different approach was sought by Kezele [39], who reconstructed the sharp diffusion ODF by incorporating a spherical wavelet transform into the Funk–Radon transform. Also Tristan-Vega [77, 78] modified the Funk–Radon approximation to the radial integral. By including the Jacobian of the spherical coordinates in FRT he computed a true orientation probability density function. Similar approach, but with a different orientation function (both with and without SH transform) was proposed by Özarslan [61] in a diffusion orientation transform.

It is important to note that the spherical harmonic basis functions are globally supported and thus are not well suited to describe sharp FODs. A recent study by Michailovich, in which the HARDI signal (and subsequently ODF) is modelled using multiresolution bases of spherical ridgelets [48] can match the SH-based QBI accuracy-wise with just a few basis functions (45 basis functions vs 4 to 8). The advantages of SH series are in robustness to noise and ability to use linear least squares methods for fast computationally inexpensive implementation.

### 2.4.3 Regularisation of Spherical Harmonic Series

In order to obtain a sharp FOD it is necessary to use large number of SH basis function. With the increasing order of SH series though, the SH transform becomes more affected by noise. This can lead to increase in ringing artefacts as well as introduce numerous weak false peaks.

One of the early works that addressed noise susceptibility of longer SH series was done by Tournier [76]. A straightforward coefficient attenuation of SH series was performed after initial SH transform of measured signal. Each coefficient of $n$th order was weighted (attenuated) prior to spherical deconvolution.

As the weights ($w = [1, 1, 1, 0.8, 0.1, 0.02, 0.002]$ for $n = [0, 2, 4, 6, 8, 10, 12]$) were chosen empirically, this method became known as ad-hoc low-pass filtering.

More formal solution to noise suppression was sought by Hess [32] and Descoteaux [19] through the use of Tikhonov regularisation. In a linear regression, the Equation 2.16 is modified by adding a regularisation term:

$$\arg\min_{\mathbf{c}_\lambda} \|\mathbf{Y}\mathbf{c}_\lambda - \mathbf{S}\|^2 + \lambda^2 \|\mathbf{\Gamma}\mathbf{c}_\lambda\|^2 \,, \tag{2.27}$$

with $\mathbf{\Gamma}$ a chosen Tikhonov matrix and $\lambda^2$ a parameter controlling the strength of regularization; for $\lambda^2 = 0$ it minimises the least squares norm. The regularised solution is therefore defined as:

$$\mathbf{c}_\lambda = (\mathbf{Y}^T\mathbf{Y} + \lambda^2\mathbf{\Gamma}^T\mathbf{\Gamma})^{-1}\mathbf{Y}^T\mathbf{S} \,. \tag{2.28}$$

The difference in Hess and Descoteaux regularisation lies in the Tikhonov matrix used. Hess used a diagonal matrix which only improved the matrix condition. Descoteaux on the other hand, built the regularisation matrix on a Laplace–Beltrami operator, which is a natural measure of smoothness for functions defined on a sphere. In SH representation, for this operator $\mathbf{\Gamma}^T\mathbf{\Gamma}$ simply becomes a diagonal matrix with $l_j^2(l_j + 1)^2$ along the diagonal ($l_j$ being the order associated with the $j$th coefficient).

An important issue in using Tikhonov regularisation is the selection of the regularisation strength $\lambda^2$. Stronger regularisation, while removing more noise will also reduce angular resolution of the FOD. Analogically, with lower parameter the FOD is sharper but at the same time high frequencies, which are polluted by noise, are being preserved. Figure 2.10 shows the effect of increasing regularisation strength on noise free and noisy FOD reconstruction.

The study by Descoteaux [18] focuses on an L-curve analysis and its use for an automatic regularisation parameter selection. Later on, Sakaie [68] work based on a general cross validation (GCV) produce similar results – the values found through GCV mostly coincides with L-Curve method. Figure 2.11 shows the optimal regularisation parameter found using the L-curve analysis. As expected, the increase in SNR means that less regularisation is needed, due to the measured signal becoming more similar to the real. The optimal parameter

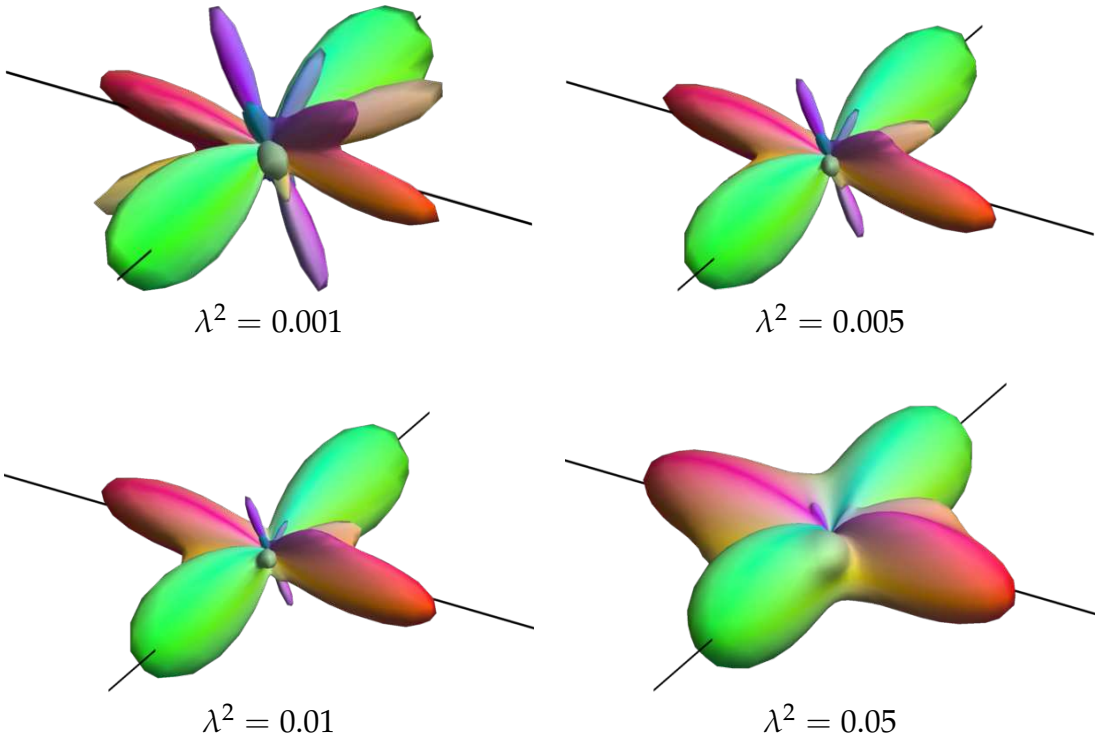
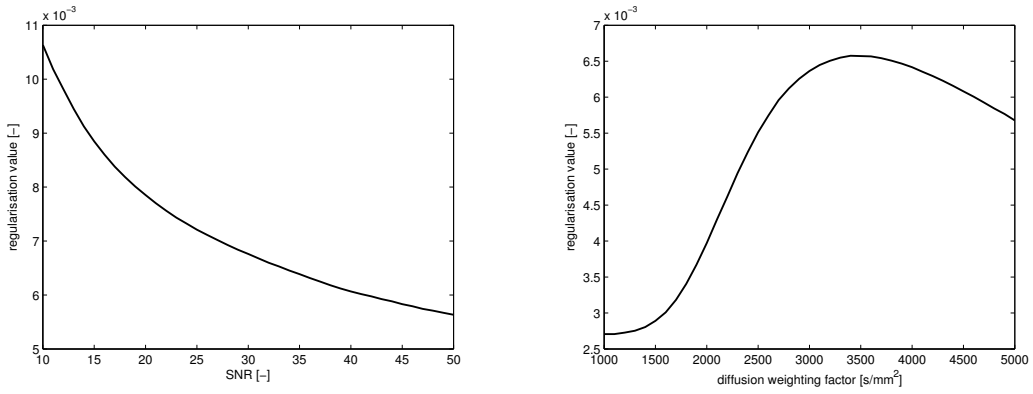



**Figure 2.10:** The effect of increasing a regularisation strength in SH transform. With the stronger regularisation applied, false peaks are suppressed at the cost of a lowered angular resolution.

value as a function of diffusion weighting though is more complex tough. Initially, the parameter value increases along with the diffusion weighting applied. With higher $b$, absolute signal values $E_i$ become smaller (see Equation 2.3), and more regularisation is needed to deal with noise. At roughly $b = 3000\ s/mm^2$, for a given SNR (SNR $= 35$) this trend reverses, most likely due signal becoming sharp enough to be mistaken with noise.

As both $L$-curve and GCV methods work by finding the optimal compromise between the data preservation and the improvement of the design matrix condition they might not be optimal when the signal is corrupted. As such, the data preservation should not be explicitly sought. In Chapter 3, Section 3.2 an optimisation scheme will be introduced, which focuses on selecting a regularisation strength that assures the highest quality of reconstructed profile.

Unfortunately, due to the ringing introduced in the deconvolution transform ALF and Tikhonov regularisation methods cannot guarantee the non-negativity of the FOD function. The improved deconvolution algorithms proposed by Dell'Acqua [17], Sakaie [68], and Tournier [75] based




a) b)

**Figure 2.11:** L-curve based optimal regularisation parameter as a function of: a) SNR, and b) diffusion weighting factor for a two fibre crossing.

on iterative approach (with Dell'Acqua using a modified Richardson-Lucy deconvolution [15], and Sakaie and Tournier a spherical deconvolution) address this at the additional computational cost (*e.g.* extra iterations, no possibility to use precomputed terms).

## 2.5 Human Brain and Tractography

The brain is the most complex organ in human. It is divided into five regions: telencephalon (cerebral hemispheres), diencephalon (thalamus, hypothalamus, pituitary gland, and pineal body), mesencephalon (midbrain), metacephalon (the pons and cerebrellum), and myelencephalon (medulla oblongata). The cerebrum, which is made of four lobes (frontal, temporal, occipital, and parietal), is the largest structure in the human brain. It consists of white and grey matter. Neuron cell bodies, dendrites and synapses form grey matter, while white matter is made of myelinated nerve fibers. Grey matter is found both on the outside (cerebral cortex), and inside (basal ganglia, thalamus, and hypothalamus) of the brain.

The fibers in white matter are arranged in tracts, connecting together different sections of brain. Depending on the location of regions that those tracts connect, they can be classified as commissural, projection, or association tracts. Commissural tracts connect left and right hemispheres; two largest commissures are the corpus callosum, and anterior commissure. Association fibres

are local, short connections located within one hemisphere. Projection fibres are ascending and descending fibres, that connect cerebrum with other parts of the brain and spinal cord (and vice versa). In these organised sections of the brain, the diffusion is constrained by myelin sheaths and axonal membranes [9] of fibre bundles. Therefore, by tracking the path of the least resistance it is possible to infer the orientation and configuration of those bundles, and thus reconstruct the white matter.

The whole family of algorithms used to resolve fibre bundles, is called tractography, and can be divided into three groups: local methods, global methods, and methods based on diffusion simulations. Local methods that track the curve sequentially through neighbouring voxels come in two variants: deterministic [5, 14, 49] and probabilistic [10, 62] streamline tractography. Global approaches on the other hand, find the best path between two points based on the optimisation criteria. They utilise front evolution techniques [63], weighted graphs [34, 47, 86], or minimise a global energy function [24, 42]. Finally, diffusion simulation methods [27, 41] either directly solve the diffusion equation, or simulate the fluid flow. While being able to resolve fibre crossing [26] and branching [8, 38] using DTI data alone, they are considerably slower and require many parameters and thresholds to tune.

# 3

# Robust Tikhonov Regularisation

Both parametric (*e.g.* DTI) and non-parametric (*e.g.* SH analysis) methods can benefit from using Tikhonov regularisation. In both cases, to automatically select the amount of regularisation applied either an L-curve or GCV analysis has to be performed.

In this chapter, a new method for finding optimal regularisation value is proposed [55, 58]. It can be formulated as an optimisation problem, and solved through minimising (or maximising) a fitness function. As this function is selected based on the similarity of real and recreated signal, the optimal parameter is thus found through ground truth verification (GTV). It is expected that the recommended regularisation values found with the GTV method will be different from parameters found with *L*-curve analysis and GCV. Furthermore, using the GTV optimised values should result in a more accurate diffusion signal reconstruction. With the GTV found values the regularisation should be able to more accurately reduce the noise corruption and not over or under regularise the signal.

Since the SH transform can be implemented efficiently, even exhaustive algorithms, which tests all possible solutions are viable. Enough iterations and wide variety of fibre configurations should be simulated, so that the averaged regularisation parameter stays unbiased. The parameters found, and accuracy of signal reconstruction using those parameter will be compared with the current state of the art parameter selection methods. Fortunately, as both L-curve and GCV analyses result in the same "optimal" parameter values [16], only one will have to be compared. Due to the use of Monte-Carlo simulation in the validation step, the L-curve will be compared as it is computationally cheaper than GCV.

This chapter is organised as follows. In Section 3.1 different fitness functions which can be used to compare the reconstruction of diffusion signal accuracy are discussed. Section 3.2 focuses on the optimal procedure for regularisation strength selection. Results are then validated using numerical simulations in Section 3.3. Tractography of a Fibre Cup phantom is performed and the results are discussed in Section 3.4. In Section 3.5 real brain image was analysed using optimally selected regularisation parameters. Discussion and further improvements are provided in Section 3.6.

## 3.1 Fitness Function

In order to optimise a reconstruction algorithm, a method of comparing two shapes should first be described. This function, called fitness function, has to be able to tell how similar two reconstructed shapes are. Perhaps the most intuitive way of comparing two spherical shapes is by computing the volumetric difference between the two shapes. This difference is 0 for two identical shapes, and becomes larger as more dissimilar they become. The fitness function to minimise the absolute difference between the real and reconstructed profiles is thus:

$$\arg\min_{\lambda} \|\mathbf{Y}_{large}\hat{\mathbf{c}}_{\lambda} - \mathbf{Y}_{large}\mathbf{c}\|^2 . \tag{3.1}$$

Two SH series are computed: of real, noise free signal $\mathbf{c}$ and of regularised, noisy signal $\hat{\mathbf{c}}_{\lambda}$. Both signals are upsampled using a dense spherical sampling $\mathbf{Y}_{large}$. Number of those points has to be large enough to accurately approximate the volumetric difference between those two profiles. By utilising the

Fourier transform addition theorem, same error measure can be obtained in a frequency domain. This allows for efficient implementation of fitness function, which is necessary for large Monte-Carlo simulations. By removing $\mathbf{Y}_{large}$ from Equation 3.1 the inverse transform is thus avoided:

$$\arg\min_{\lambda} \|\hat{\mathbf{c}}_{\lambda} - \mathbf{c}\|^2 . \tag{3.2}$$

Revising Equation 2.19 is is now possible implement a fitness function based on an ODF:

$$\arg\min_{\lambda} \|\mathbf{P}\hat{\mathbf{c}}_{\lambda} - \mathbf{P}\mathbf{c}\|^2 , \tag{3.3}$$

where the constant matrix $\mathbf{P}$ implements a Funk-Radon transform. Similarly, using Equation 2.26 an FODF based fitness function can be built:

$$\arg\min_{\lambda} \|\mathbf{R}\hat{\mathbf{c}}_{\lambda} - \mathbf{R}\mathbf{c}\|^2 . \tag{3.4}$$

The constant matrix $\mathbf{R}$ is used to perform a spherical deconvolution transform.

Finally, two other measures can be used to directly differentiate between two SH series – correlation coefficient [4] and symmetric Kullback–Leibler (sKL) divergence [12]. The correlation coefficient is given by:

$$\mathrm{r}_{\lambda} = \frac{\sum_l \sum_m c_{lm}\hat{c}_{lm}}{\sqrt{\sum_l \sum_m c_{lm}c_{lm}}\sqrt{\sum_l \sum_m \hat{c}_{lm}\hat{c}_{lm}}} . \tag{3.5}$$

The sKL divergence measure is based on the DTI analysis, which in SH representation relates to using the coefficients of the ODF:

$$\mathrm{sKL}_{\lambda} = \frac{1}{2}\{\frac{1}{c_{00}}\sum_l \sum_{m=-l}^{l} [c_{lm}d_{lm} - c_{lm}\hat{d}_{lm}] + \frac{1}{\hat{c}_{00}}\sum_l \sum_{m=-l}^{l} [\hat{c}_{lm}\hat{d}_{lm} - \hat{c}_{lm}d_{lm}]\} , \tag{3.6}$$

with $\mathbf{c}$ and $\mathbf{d}$ being SH coefficients of $\mathbf{\Psi}$ and $\log(\mathbf{\Psi})$, and $\hat{\mathbf{c}}$ and $\hat{\mathbf{d}}$ of noise corrupted $\hat{\mathbf{\Psi}}$ and $\log(\hat{\mathbf{\Psi}})$. In Equations 3.5 and 3.6, SH coefficients $\hat{c}_{lm}$ and $\hat{d}_{lm}$ are regularised with strength $\lambda$ (for clarity the $\lambda$ index is dropped in the scalar notation).

Figure 3.1 shows the aptitude of different error measures as a function of cross-

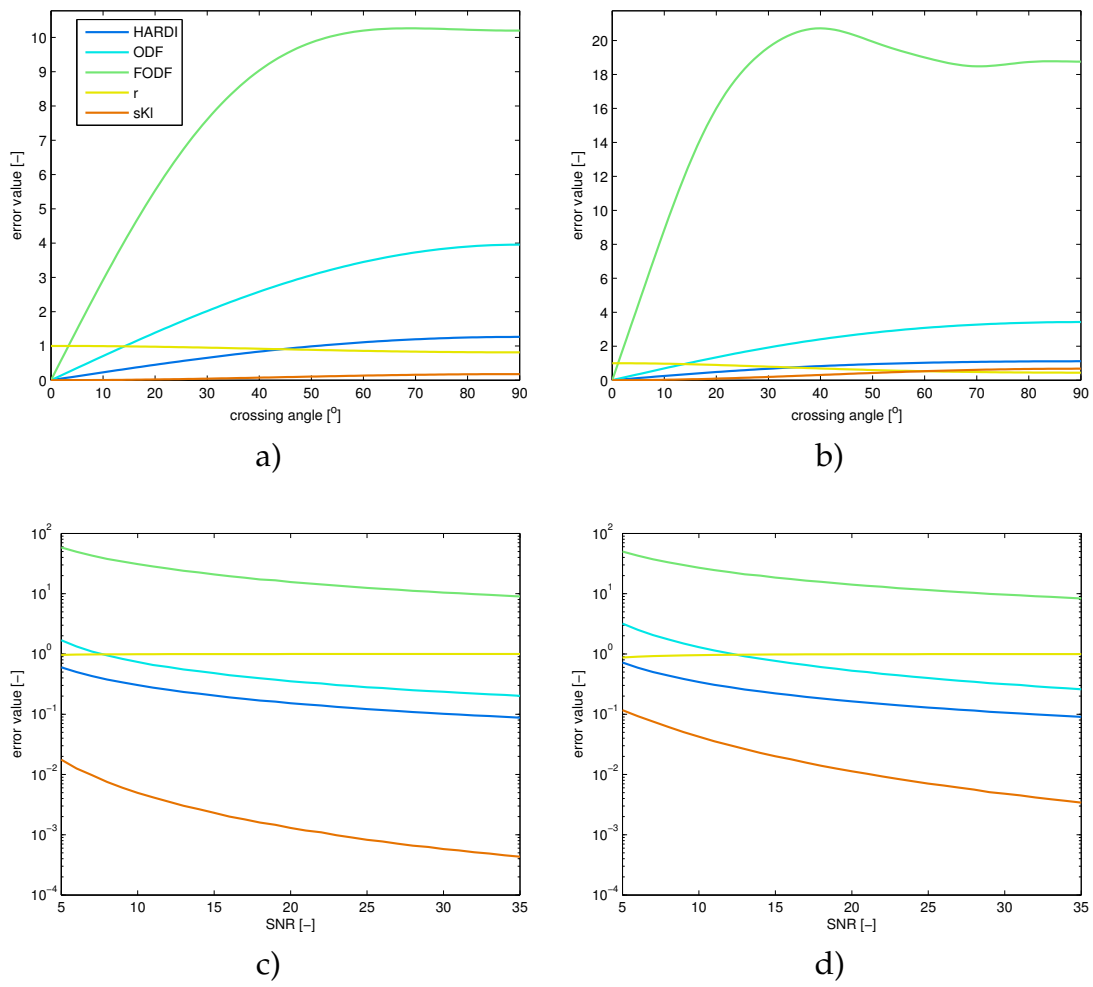



**Figure 3.1:** Aptitude of different error measures as a function of crossing angle (top) and SNR (bottom). Signal was simulated using low ($b = 1000\ s/mm^2$, left) and high ($b = 3000\ s/mm^2$, right) diffusion weighting factor.

ing angle (top) and SNR (bottom). Two diffusion weighting factors are visualised – low ($b = 1000\ s/mm^2$, left) and high ($b = 3000\ s/mm^2$, right).

The L2 norm of FODF differences, does not scale linearly with regards to the crossing angle. For example, for $b = 3000\ s/mm^2$ it peaks at around 45°, falls down till around 80°, and then remains constant. As such, the FODF cannot be used to reliably asses similarity of two reconstructions.

The ideal error measure should be a linear function of both crossing angle and SNR, and be sensitive to changes. As none of the remaining error measures exhibit such behaviour, all four will be used for optimisation, and numerical validation.

## 3.2 Optimal Parameter Selection

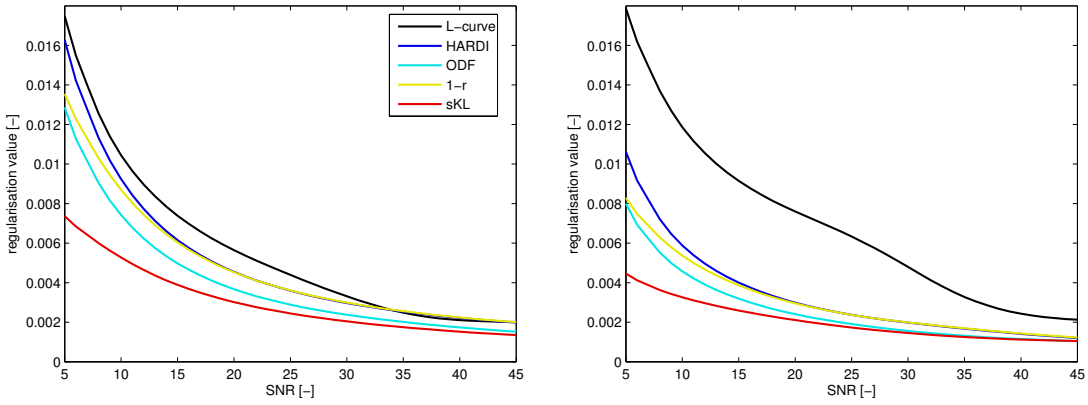
The algorithm for creating the phantom signal and analysing the regularisation fitness can be summarised as follows:

1. For a given diffusion weighting factor $b$ and signal to noise ratio $1/\sigma$ do:
2. Select a number of $n$ fibre crossings with random volume fractions.
3. Generate a synthetic signal $E$ using a multi-tensor model.
4. Generate a noise corrupted signal $\hat{E} = \sqrt{(E + N(0, \sigma^2))^2 + N(0, \sigma^2)^2}$.
5. Investigate the fitness function for various regularisation parameters $\lambda^2$.
6. Repeat from 2 for $k$ times.

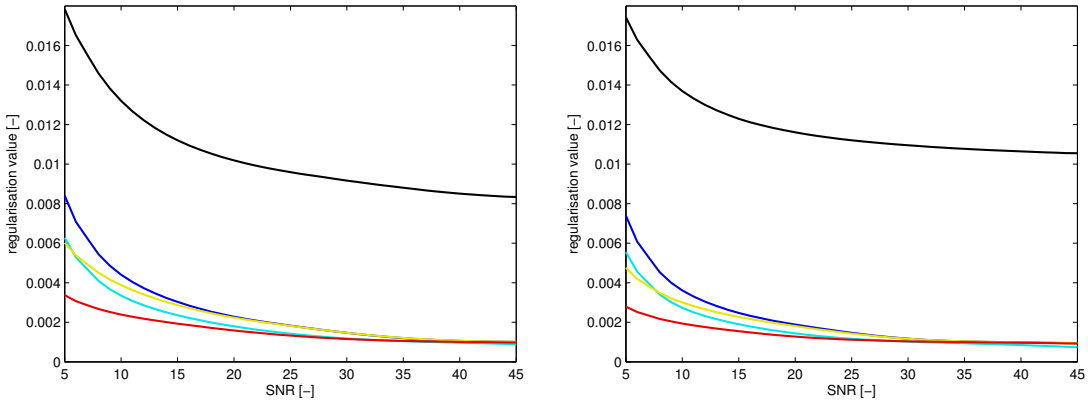
For a set of diffusion weighting factors and SNRs (both feasible and non-feasible combinations), a range of $\lambda^2 \in [0, 0.5]$ is tested. No limits on minimal crossing angle or volume fractions are placed. For each iteration the best parameter is selected based on the current fitness function. Four fitness functions from Section 3.1 are tested: $HARDI$, $ODF$, $1 - r$ and $sKL$. Each simulation is tested $k = 10,000$ times, and the final parameter is averaged.

The optimal parameter as a function of SNR and diffusion weighting, found using an exhaustive search (GTV method) and L-curve analysis, is presented in Figures 3.2-3.9. Fibre configurations up to three fibre crossings, as well as an unknown (random) fibre configurations were tested, and the best parameter for a single fibre (Figures 3.2 and 3.3), two fibre crossing (Figures 3.4 and 3.5), and three fibre crossing (Figures 3.6 and 3.7) reconstructions is provided. Additionally, the optimal parameter for an unknown (random) fibre configuration is given by Figures 3.8 and 3.9.

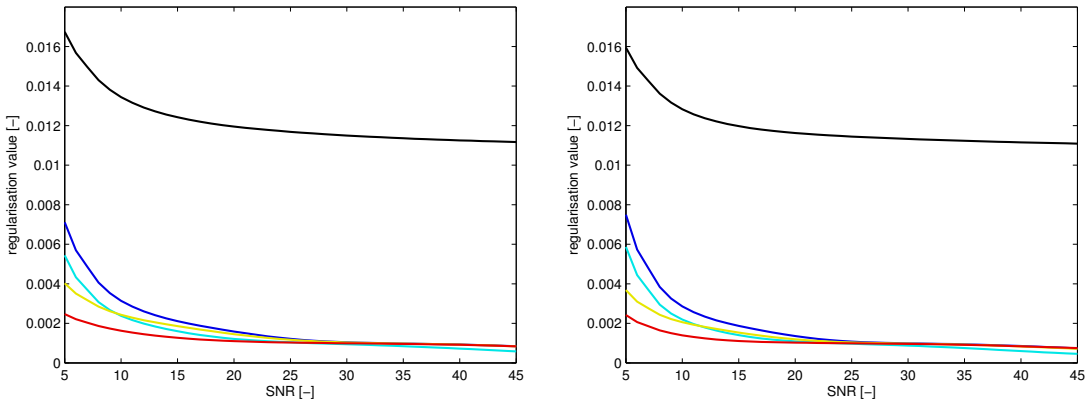
**Optimal parameter as a function of SNR** All tested fitness functions, as well as an L-curve analysis react correctly to the change in SNR (Figures 3.2, 3.4, 3.6, and 3.8). As the signal-to-noise increase, less smoothing is required. The L-curve based parameter though, especially for a single fibre reconstruction with a high diffusion weighting, does not decrease to 0.




a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

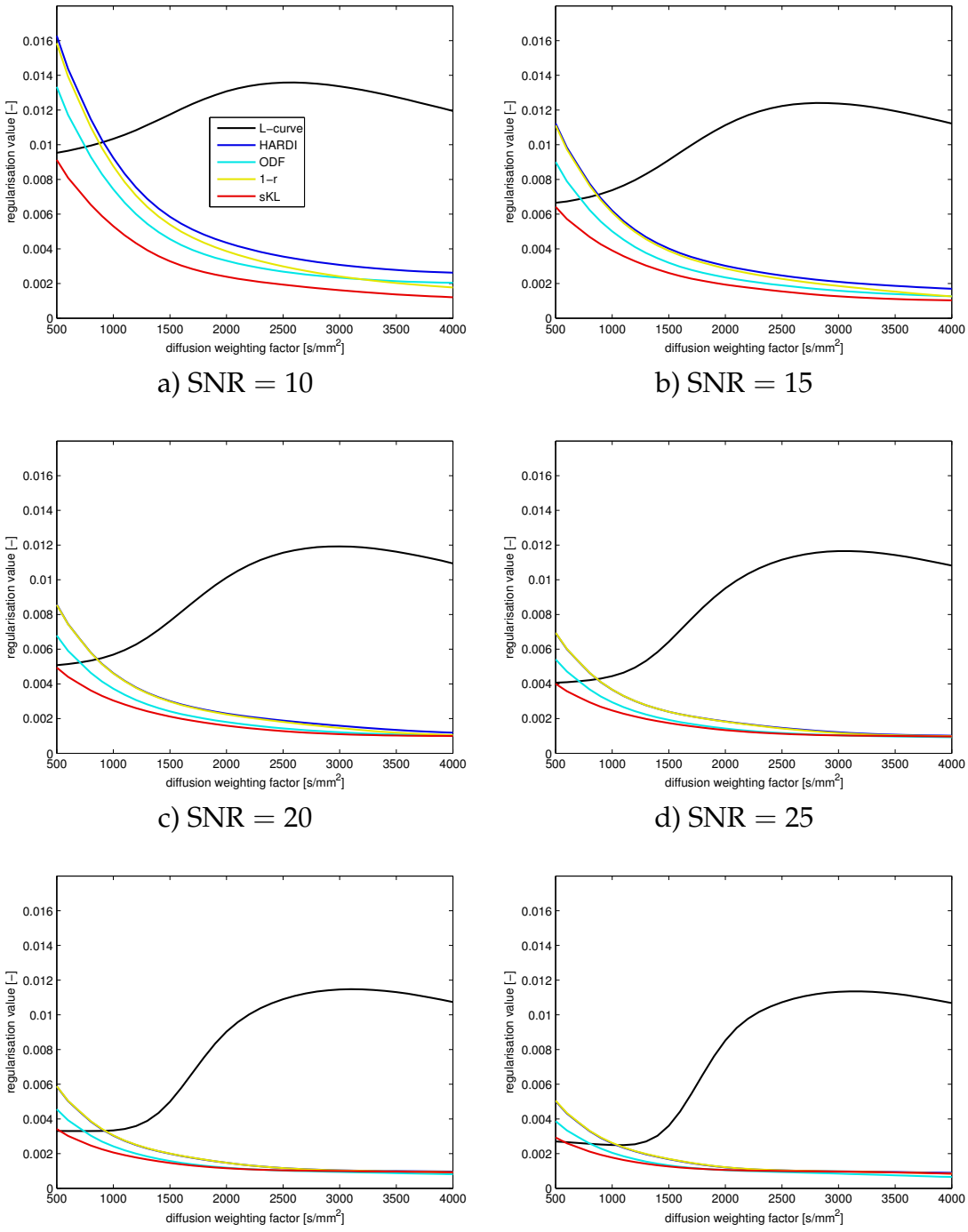



c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$




e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.2:** Optimal parameter values for a single fibre reconstruction as a function of SNR for six diffusion weighting factors.

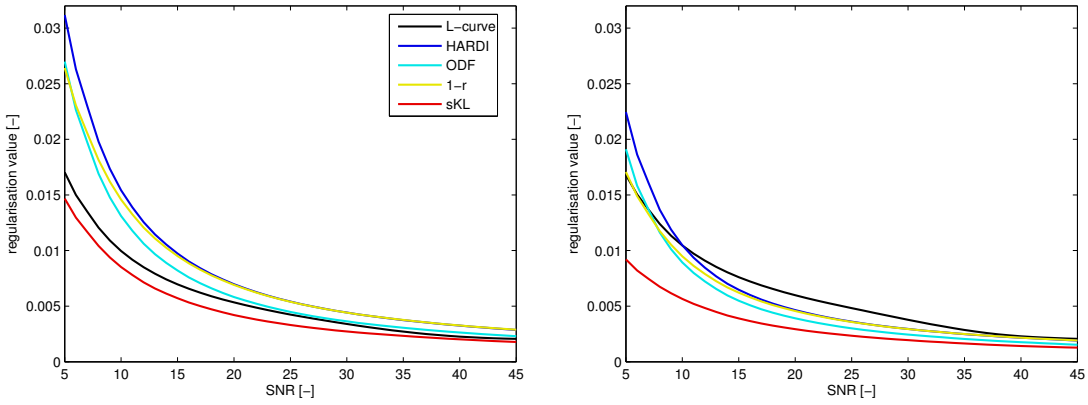



a) SNR = 10 b) SNR = 15

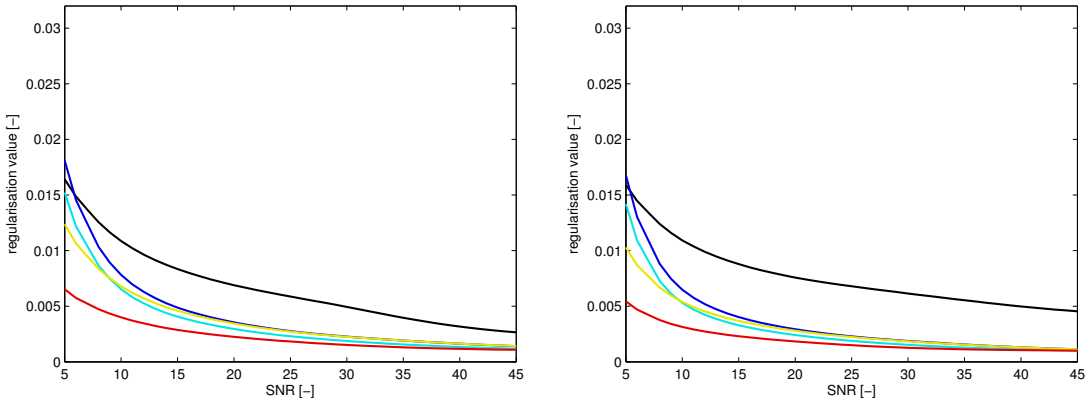
c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

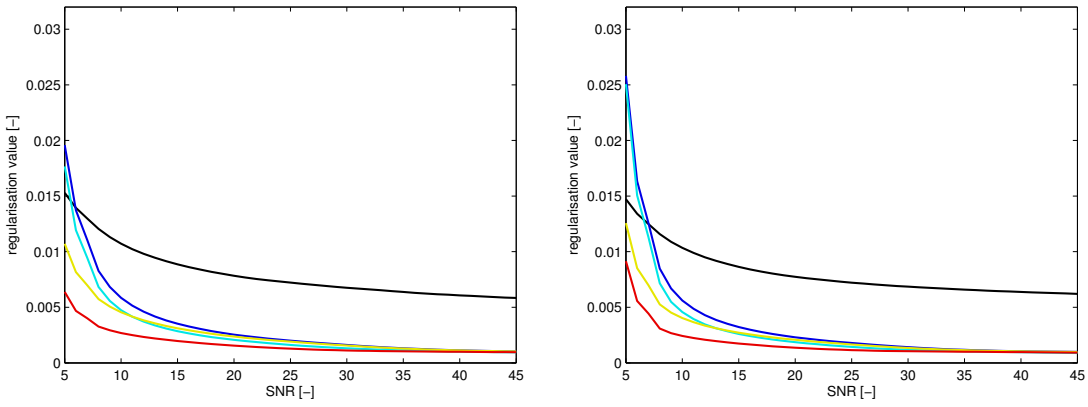
**Figure 3.3:** Optimal parameter values for a single fibre reconstruction as a function of diffusion weighting factor for six SNRs.




a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

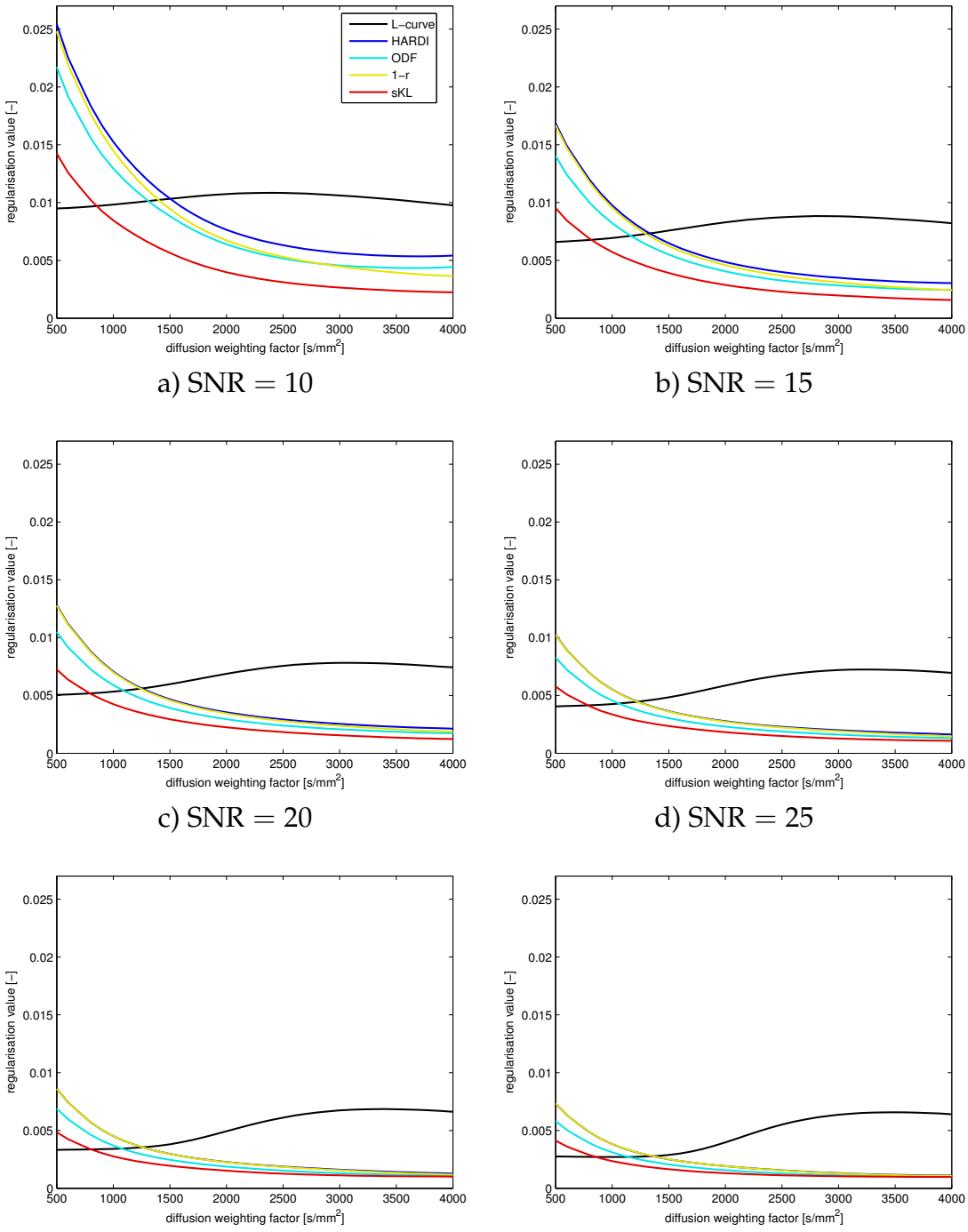



c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$




e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.4:** Optimal parameter values for a two fibre reconstruction as a function of SNR for six diffusion weighting factors.

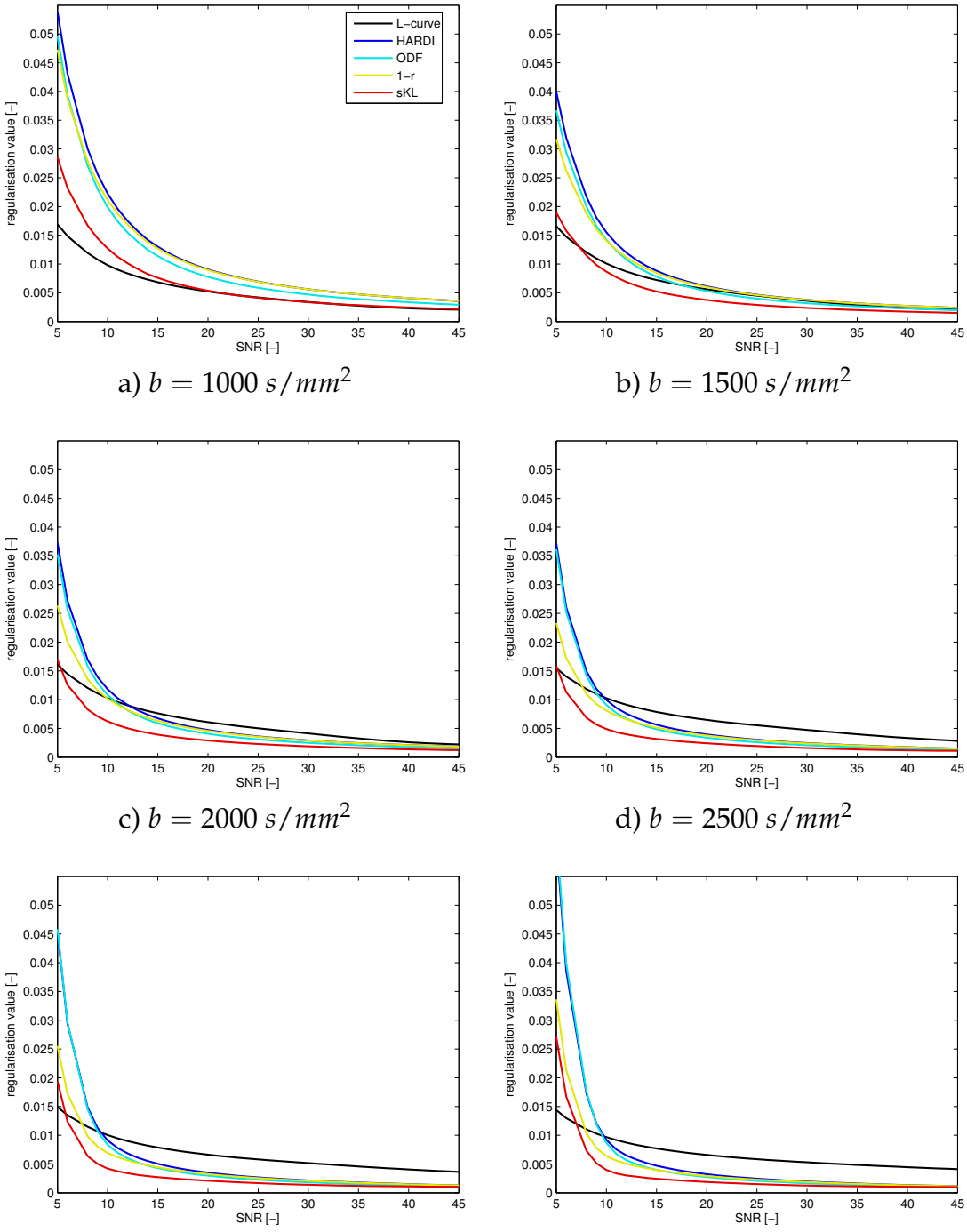



a) SNR = 10 b) SNR = 15

c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

**Figure 3.5:** Optimal parameter values for a two fibre reconstruction as a function of diffusion weighting factor for six SNRs.

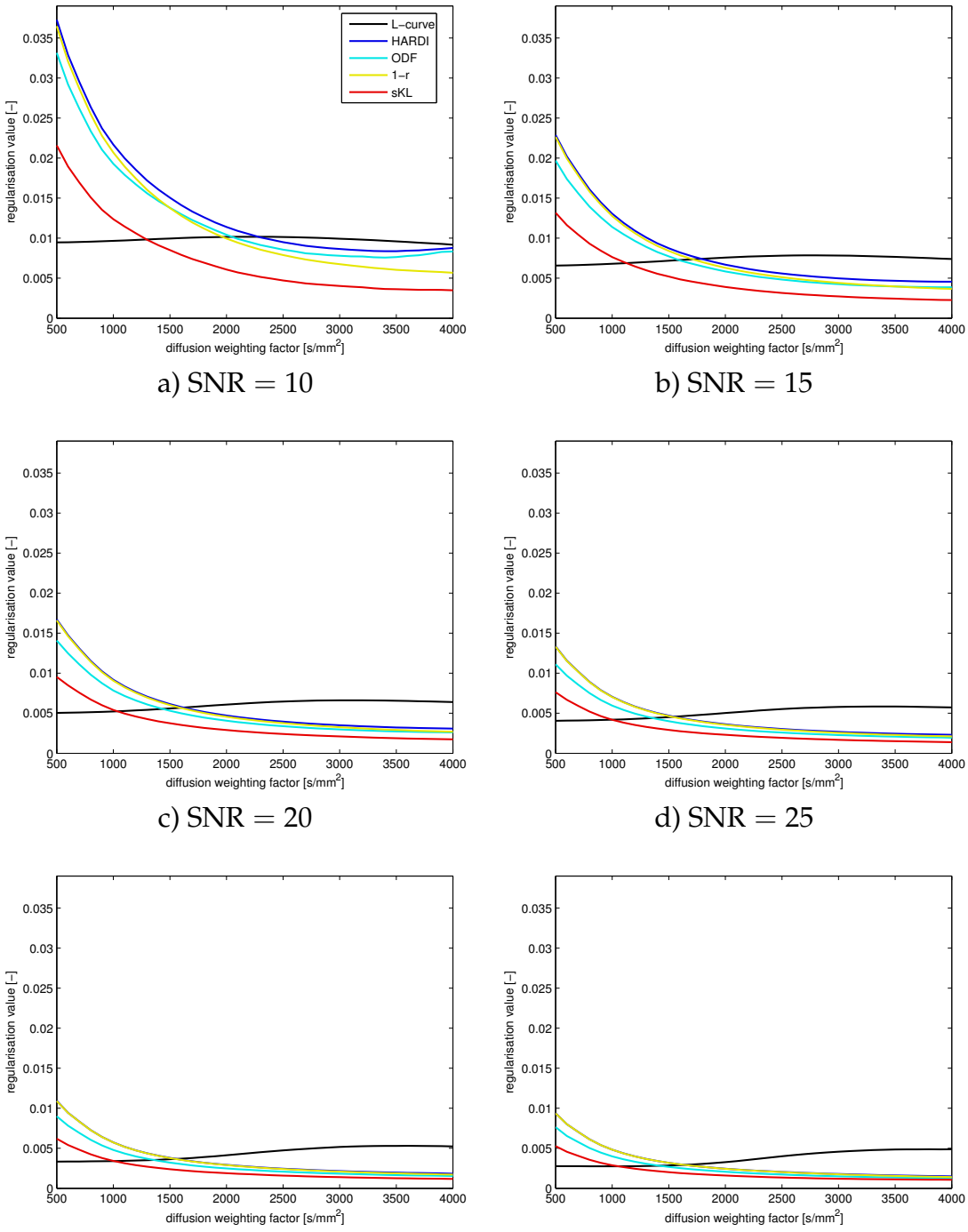



a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$

e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.6:** Optimal parameter values for a three fibre reconstruction as a function of SNR for six diffusion weighting factors.

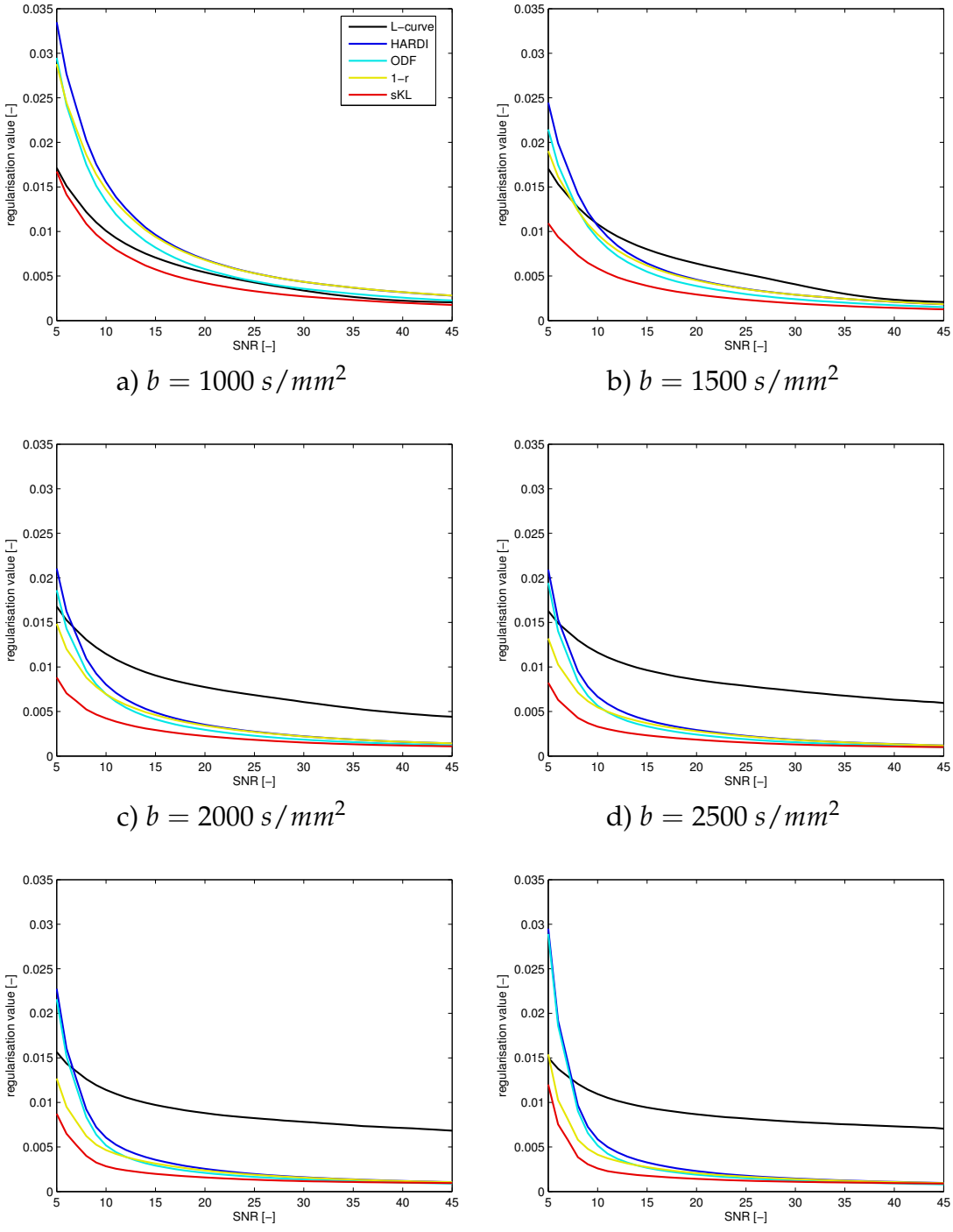



a) SNR = 10 b) SNR = 15

c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

**Figure 3.7:** Optimal parameter values for a three fibre reconstruction as a function of diffusion weighting factor for six SNRs.

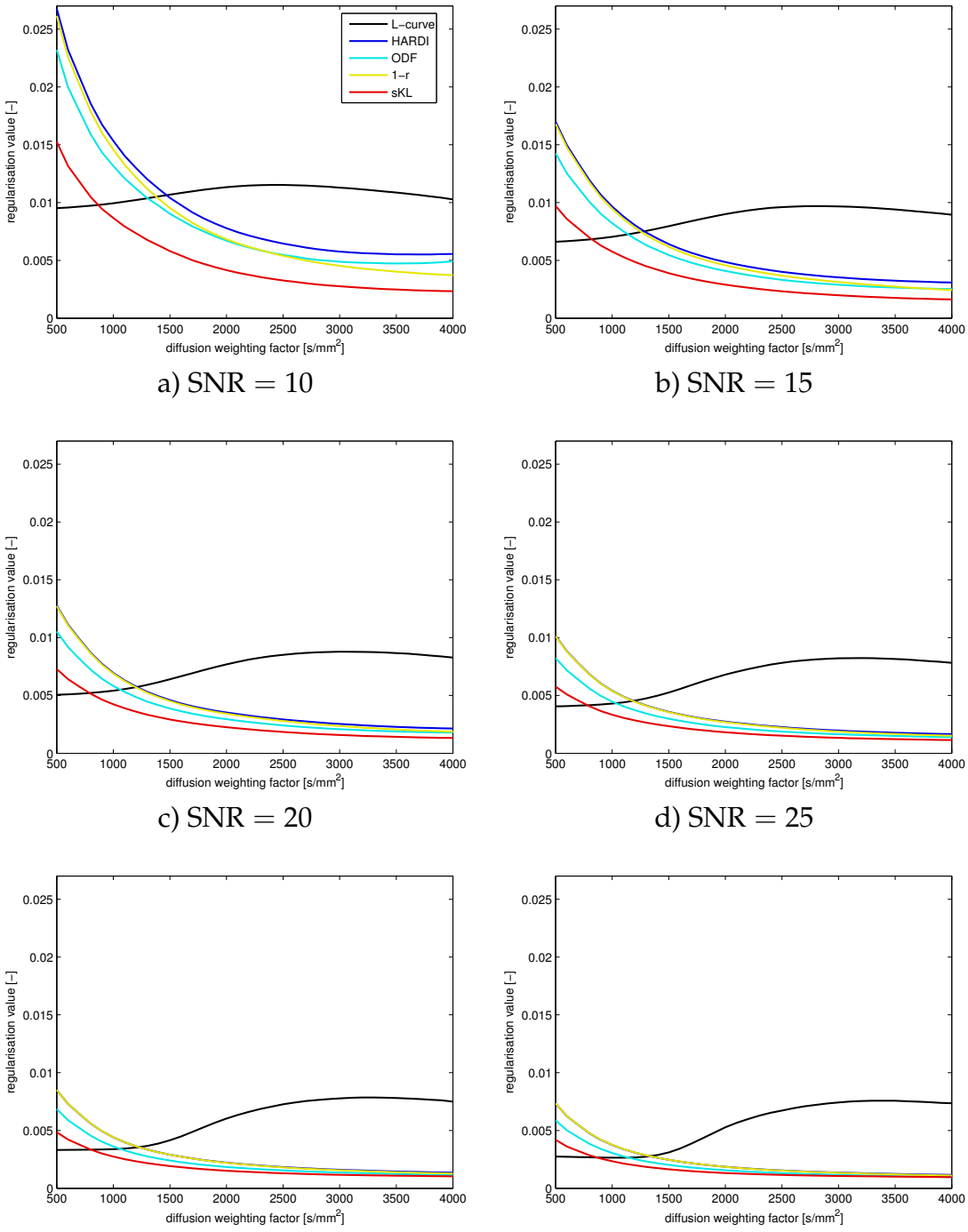



a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$

e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.8:** Optimal parameter values for a random fibre reconstruction as a function of SNR for six diffusion weighting factors.




a) SNR = 10 b) SNR = 15

c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

**Figure 3.9:** Optimal parameter values for a random fibre reconstruction as a function of diffusion weighting factor for six SNRs.

This means that even a noise-free signal will be smoothed out, decreasing the angular resolution. Other fitness functions – HARDI, ODF, coefficient correlation, and symmetric Kullback-Leibler divergence – produce parameter that decreases towards 0 as SNR reaches infinity.

**Optimal parameter as a function of b-value** The GTV method also reacts properly to the change in diffusion weighting factor (Figures 3.3, 3.5, 3.7, and 3.9). As the diffusion signal becomes sharper (higher b-values), it becomes impossible to apply strong regularisation without affecting the underlying signal. This trend is maintained across all fitness functions tested.

As mentioned, at high diffusion weighting applied, the measured signal becomes very sharp (fast changing). As such, with low SNR it becomes impossible to distinguish signal from noise, as both are in a high frequency spectrum. The stronger regularisation that is used in this case (Figures 3.5a, 3.7a, and 3.9a at $b = 4000\ \mathrm{s/mm^2}$), is smoothing both signal and noise, and degrades the reconstructed signal towards an oblate shape (or even sphere, if the SNR is too low).

The L-curve based parameters exhibit more complex trend. For all tested SNRs, the parameter value found using this method increases along with the diffusion weighting applied. Depending on a diffusion weighting factors, it peaks at around 3000 $\mathrm{s/mm^2}$ (earlier for lower and later for higher b-values), and then starts to decay. Both the decay rate and peak magnitude is unaffected by the SNR.

With the exception of single fibre simulations, the L-curve analysis is less sensitive to the change in diffusion weighting factor than the GTV method. The difference between the recommended parameter value for low and high diffusion weighting applied is close to 0, often with the high diffusion applied parameter being larger then low.

**Optimal parameter and the number of fibre crossing** Both the regularisation parameter found with an L-curve and GTV methods follow their respective trends regardless of the number of simulated fibre crossing. Depending on the underlying fibre configuration, the values found to be optimal are different though.

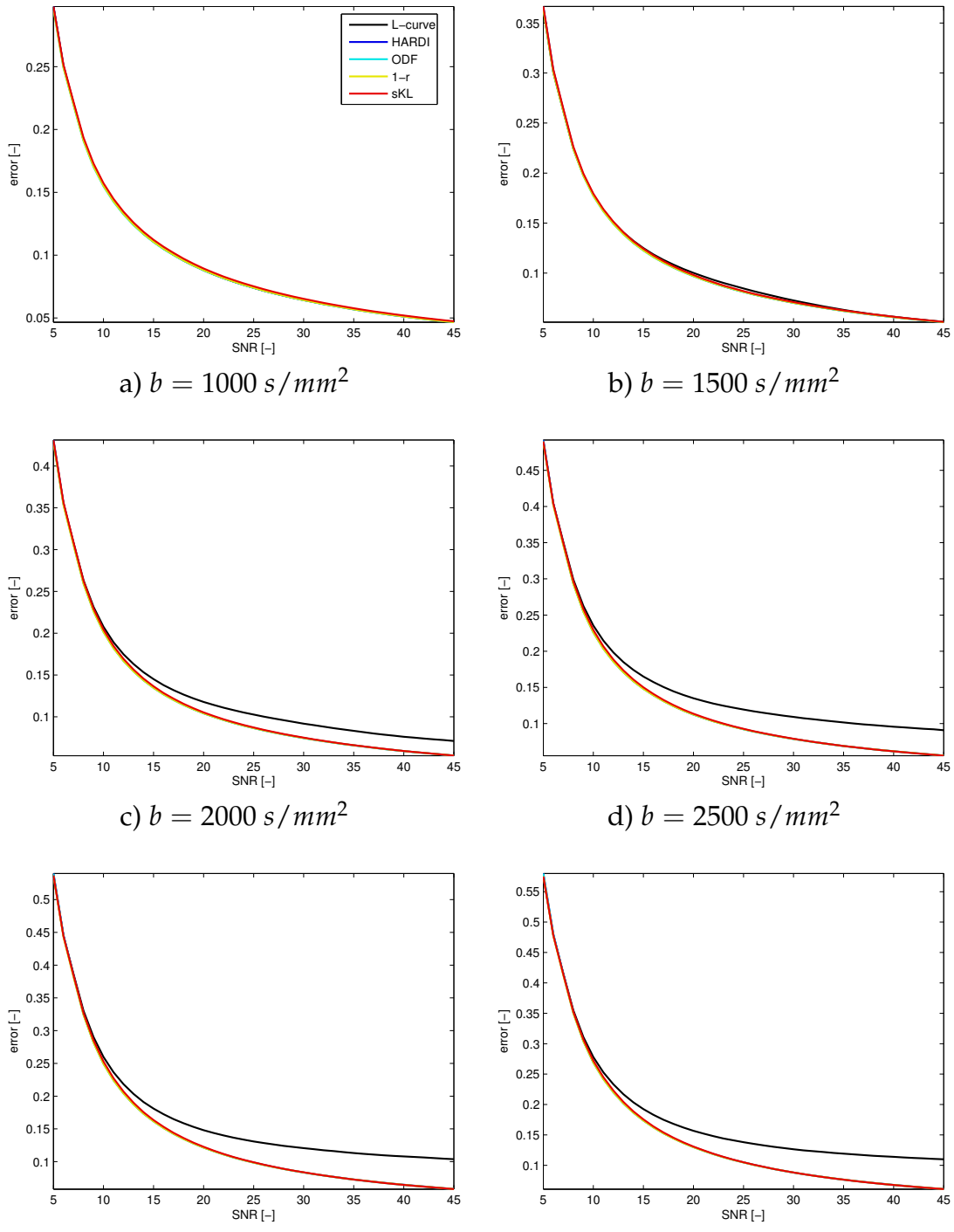
For all SNRs and diffusion weighting factors, optimal parameter value was the largest for a three fibre crossing simulation. The single fibre simulation need the smallest regularisation to be applied, and the optimal regularisation for a two fibre crossing is in-between. The simulations involving random fibre configurations provide regularisation values similar to the simulation of a two fibre crossing. This is concise with Descoteaux's research [18]. When model selection on a per-voxel basis is not viable (*e.g.* most clinical cases), regularisation values provided found through a random fibre configuration should be used.

## 3.3 Numerical Validation of Optimised Reconstruction

To assess the effect of new regularisation parameter on regularisation accuracy, another set of Monte-Carlo experiments is designed. For all tested diffusion weighting factor and SNR pairs in Section 3.2, diffusion signal of random fibre configuration is reconstructed with regularisation values found by L-curve and GTV methods. The reconstruction error is measured using the four fitness function selected in 3.1, and for each fitness function, is averaged over all iterations. Similar to Section 3.2, number of iterations for each test was set to $10,000$.

Figures 3.10-3.17 show the signal reconstruction error of an L-curve and GTV based regularised SH transform. The regularisation value was selected based on the results shown in the Figures 3.2-3.9. In many figures, only two functions are visible - black, representing the error of L-curve based regularised reconstruction, and red, of GTV optimisation using symmetric Kullback–Leibler divergence fitness function. In those cases the missing functions are fully, or partially, obstructed by GTV-sKL.

The error of regularised reconstruction of a random fibre configuration is measured using four fitness functions: HARDI (Figures 3.10 and 3.11), ODF (Figures 3.12 and 3.13), correlation coefficient (Figures 3.14 and 3.15), and symmetric Kullback–Leibler diverge (Figures 3.16 and 3.17). Again, similar to Section 3.2, all errors are plotted as a function of SNR and diffusion weighting factor. The larger the SNR or diffusion weighting factor become (or both), the bigger the difference in regularisation accuracy using the parameters recommended by L-curve and GTV.

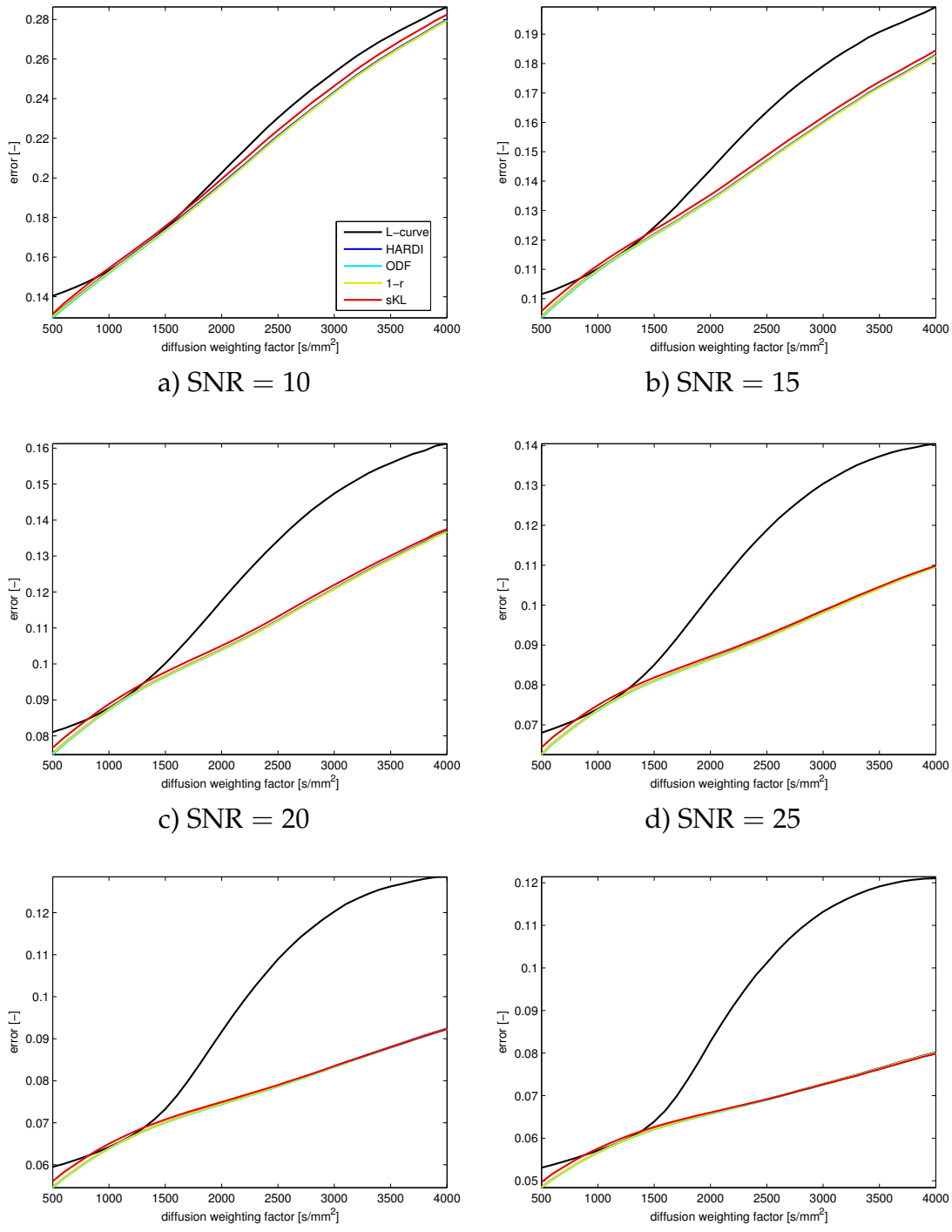



a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$

e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.10:** Reconstruction error represented using a HARDI error measure as a function of SNR for six diffusion weighting factors.

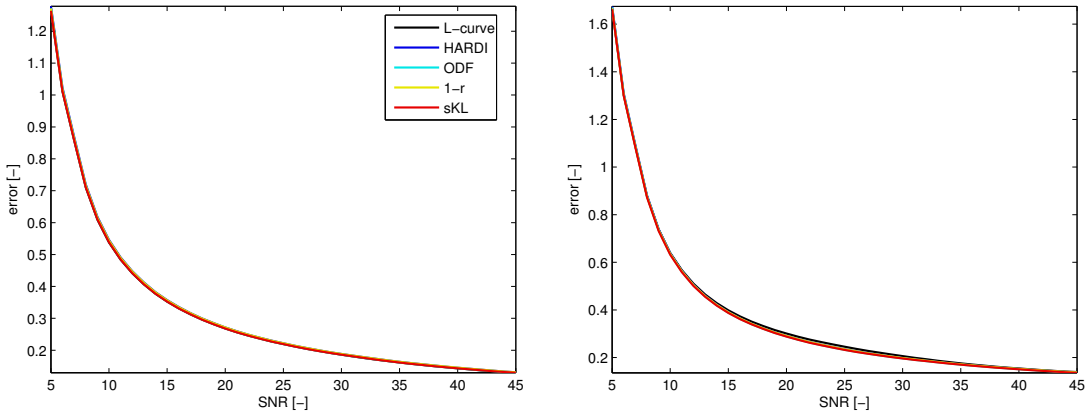



a) SNR = 10 b) SNR = 15

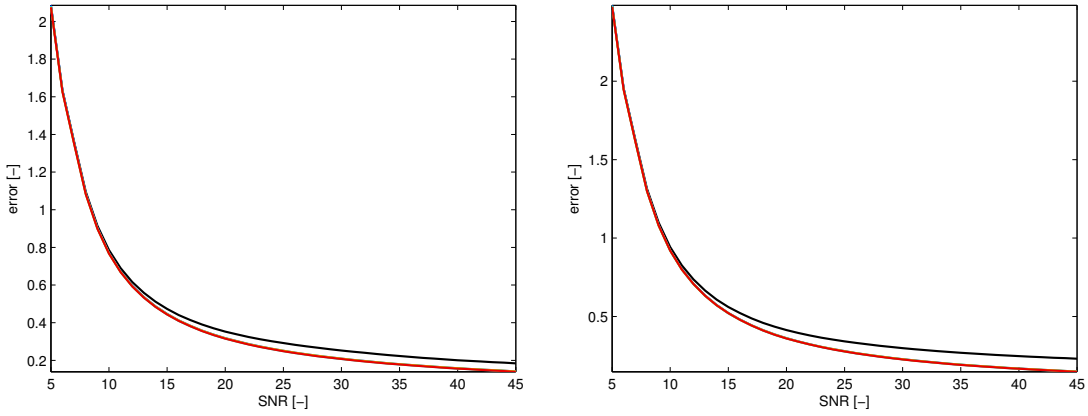
c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

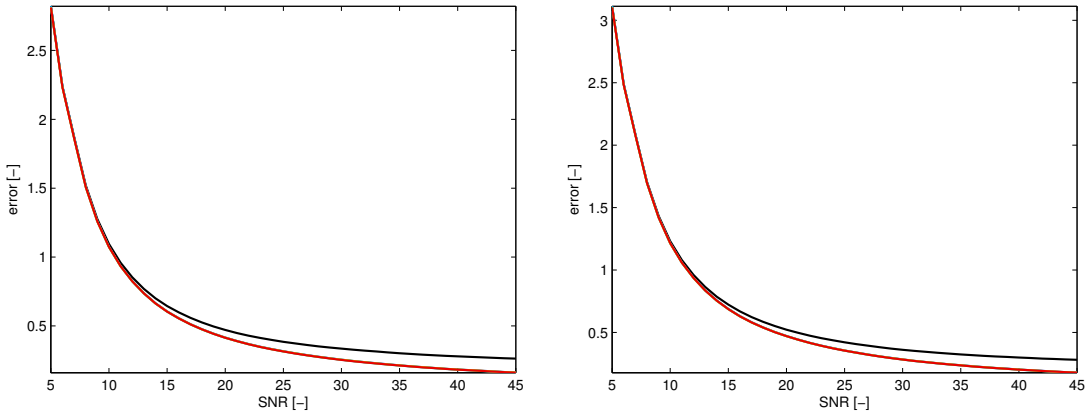
**Figure 3.11:** Reconstruction error represented using a HARDI error measure as a function of diffusion weighting factor for six SNRs.




a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

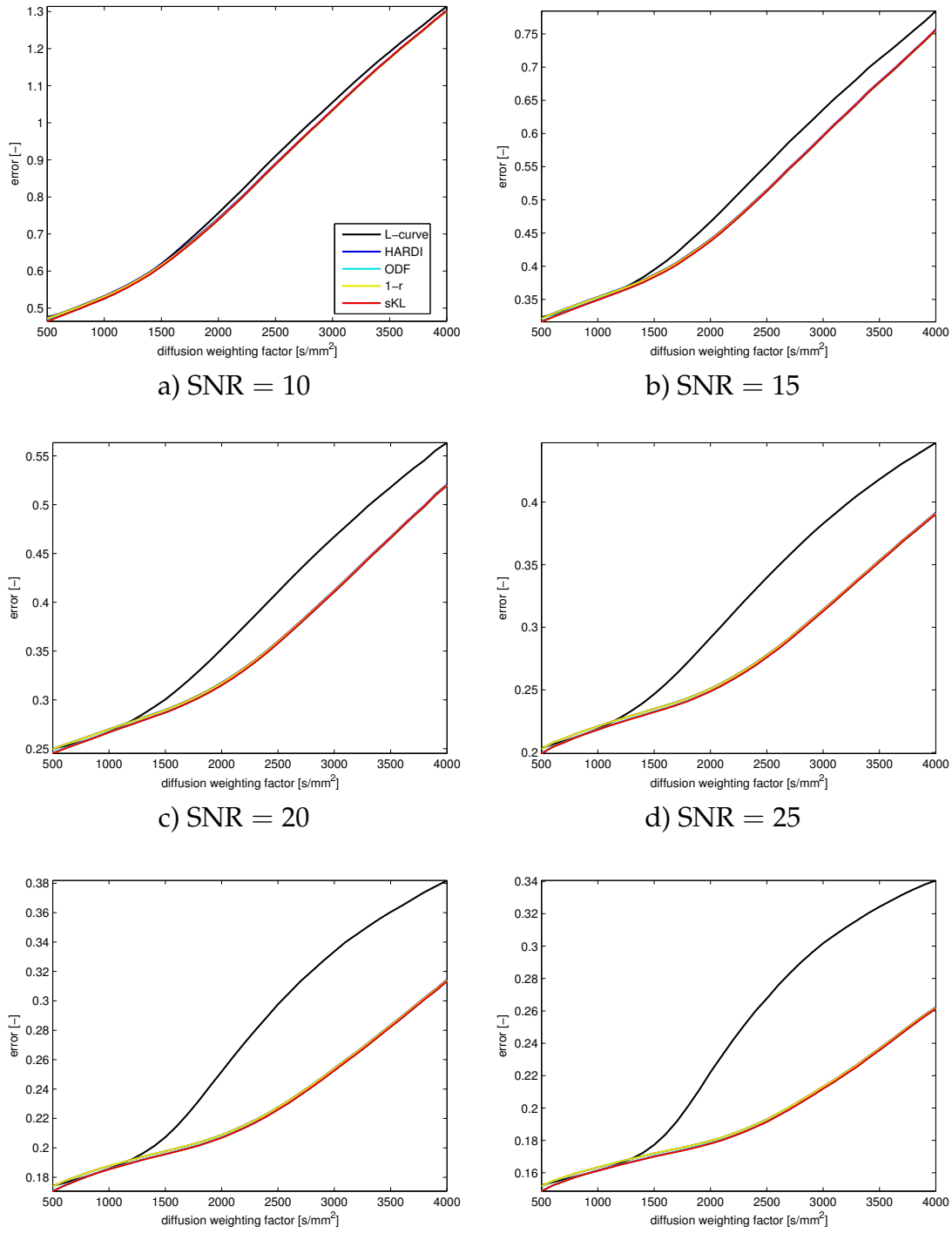



c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$




e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.12:** Reconstruction error represented using an ODF error measure as a function of SNR for six diffusion weighting factors.

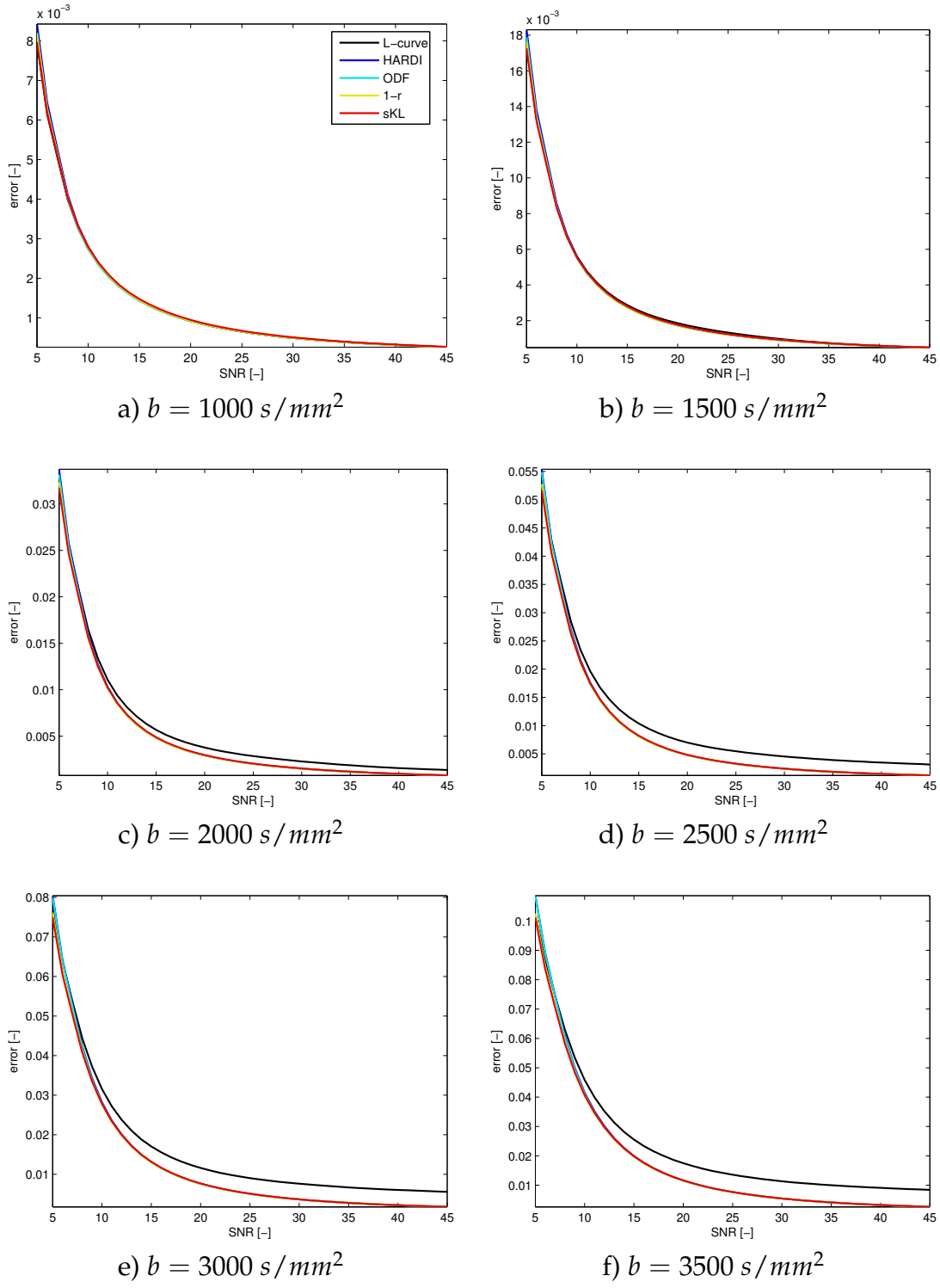



a) SNR = 10 b) SNR = 15

c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

**Figure 3.13:** Reconstruction error represented using a ODF error measure as a function of diffusion weighting factor for six SNRs.

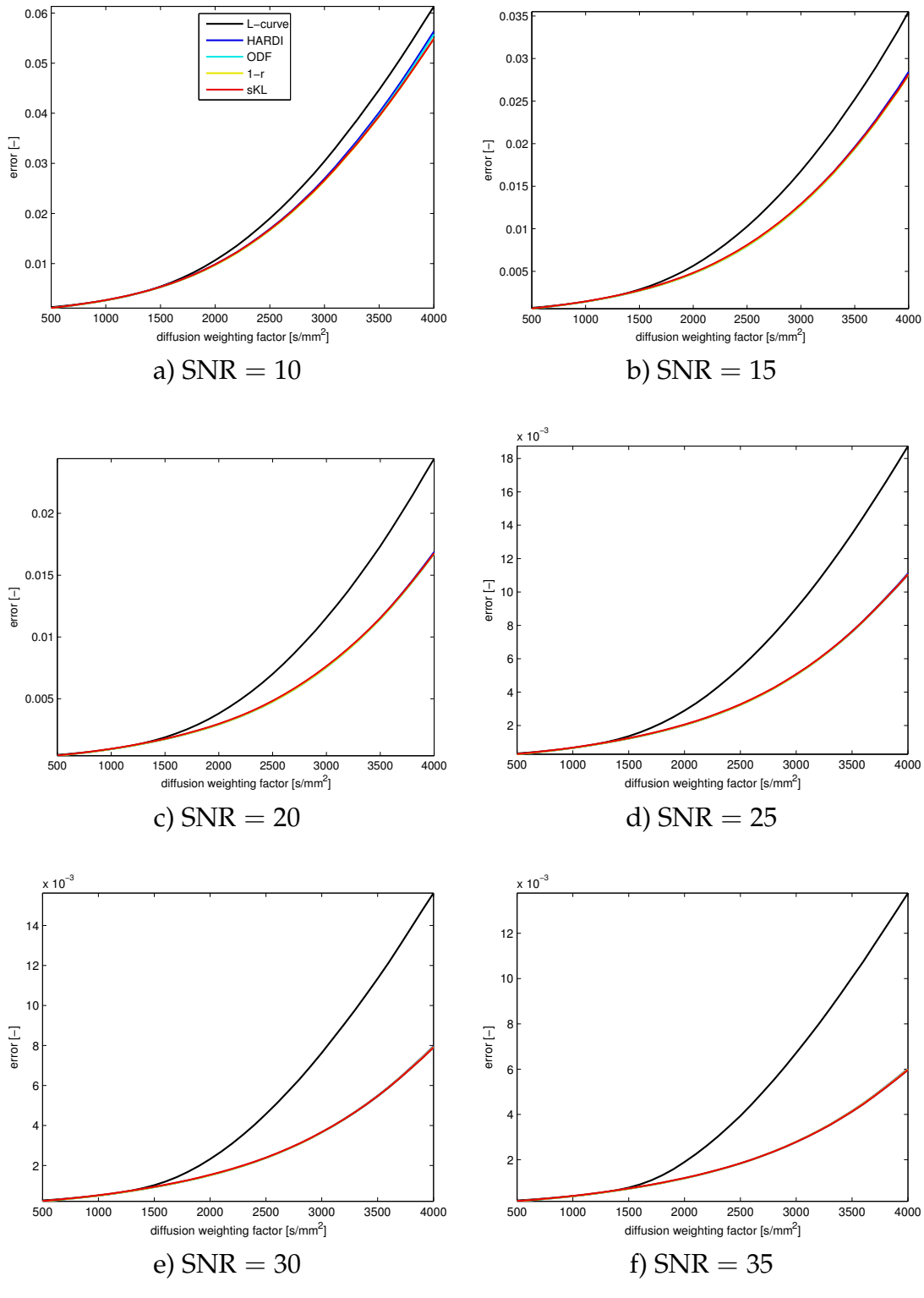



a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$

e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

**Figure 3.14:** Reconstruction error measured using a correlation coefficient as a function of SNR for six diffusion weighting factors.

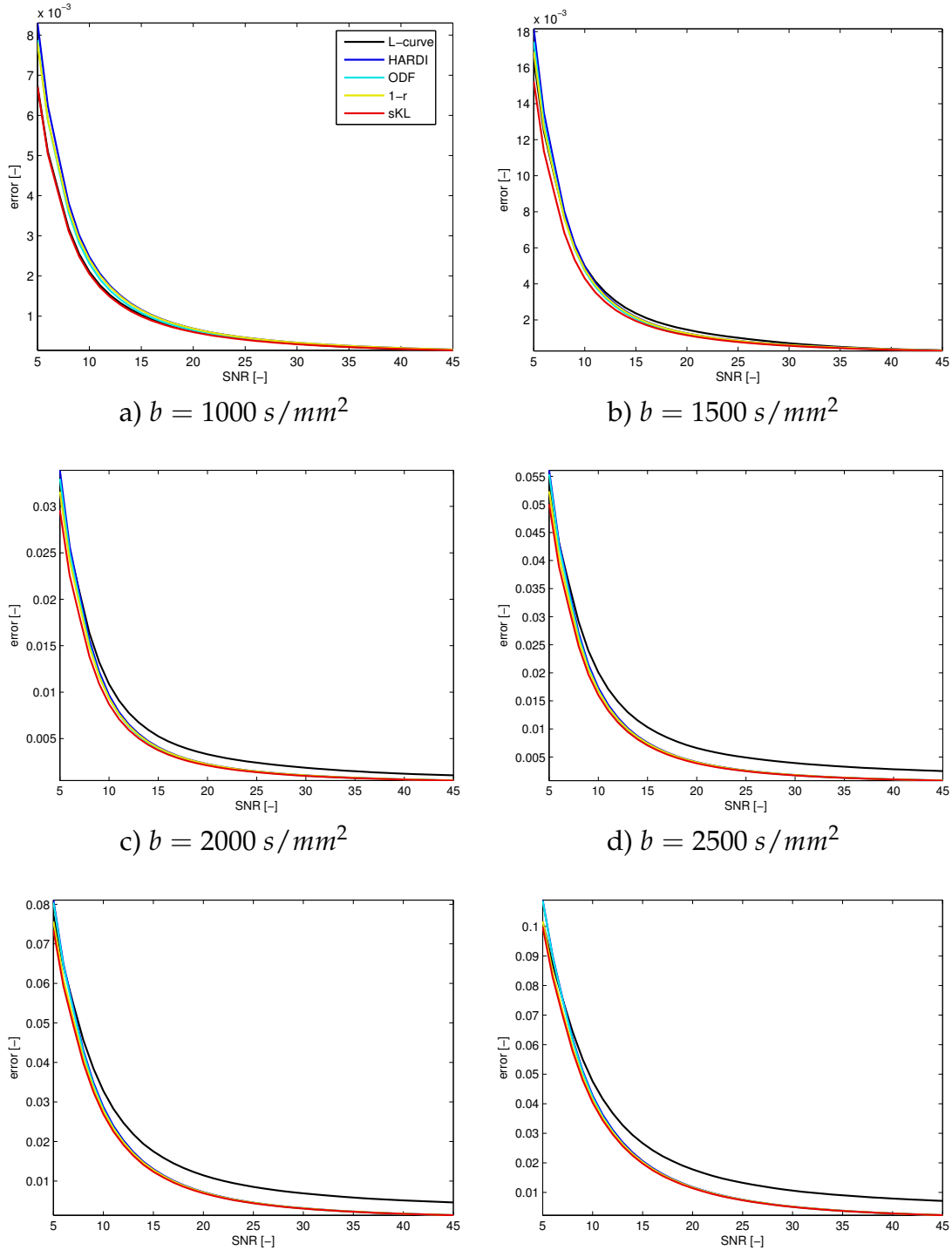



a) SNR = 10 b) SNR = 15

c) SNR = 20 d) SNR = 25

e) SNR = 30 f) SNR = 35

**Figure 3.15:** Reconstruction error measured using a correlation coefficient as a function of diffusion weighting factor for six SNRs.

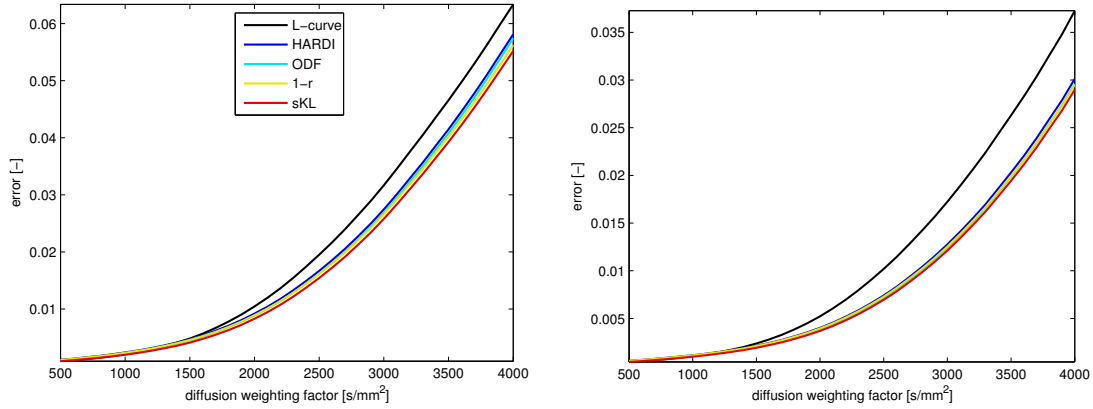



a) $b = 1000\ s/mm^2$ b) $b = 1500\ s/mm^2$

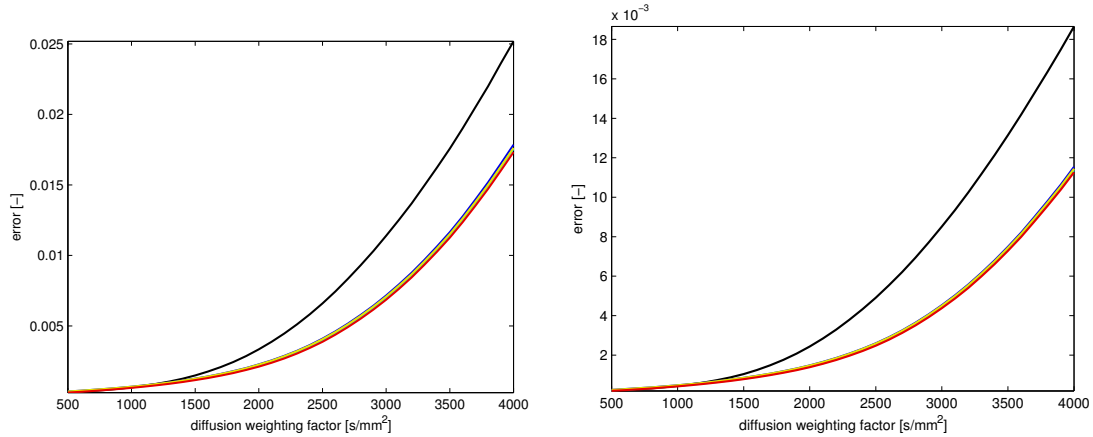
c) $b = 2000\ s/mm^2$ d) $b = 2500\ s/mm^2$

e) $b = 3000\ s/mm^2$ f) $b = 3500\ s/mm^2$

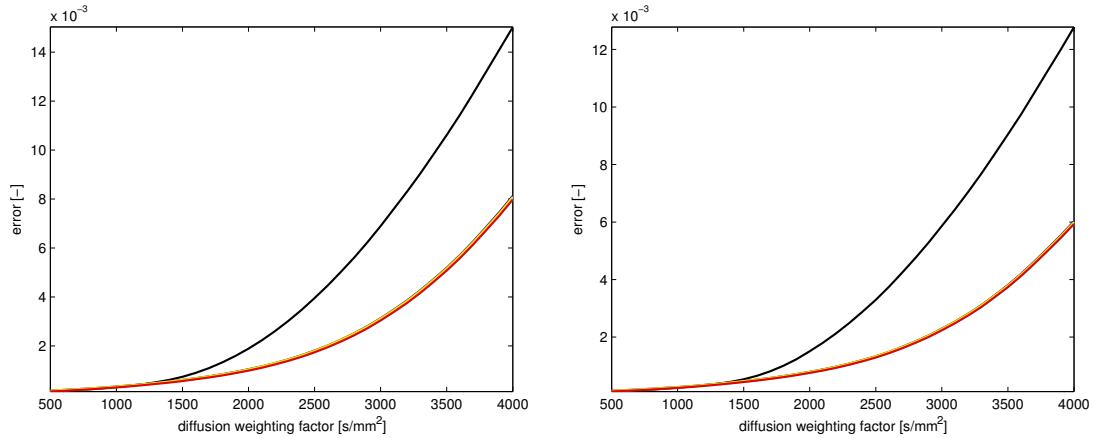
**Figure 3.16:** Reconstruction error measured using a symmetric Kullback–Leibler divergence as a function of SNR for six diffusion weighting factors.




a) SNR = 10 b) SNR = 15




c) SNR = 20 d) SNR = 25




e) SNR = 30 f) SNR = 35

**Figure 3.17:** Reconstruction error measured using a symmetric Kullback–Leibler divergence as a function of diffusion weighting factor for six SNRs.

In all tests, the reconstruction with GTV based regularisation value performs no worse than L-curve based value. As expected, the GTV based regularised reconstruction becomes more accurate as the SNR increases. This is due to GTV method correctly assessing the signal to be of good quality, and gradually decreasing the applied regularisation towards zero. The L-curve based parameter on the other hand, regardless of the SNR or diffusion weighting applied, always is greater than zero.

## 3.4 Tractography using Fibre Cup Phantom

To test if the improved regularisation produces globally viable results, a streamline tractography on a Fibre Cup phantom was performed. The Fibre Cup, a tractography contest first held at MICCAI 2009 in London, is a competition in which mechanical phantom [23, 65] is used to evaluate tractography algorithms performance. The phantom consists of many possible fibre structures, like fibre crossing, kissing, splitting, or bending, and is therefore a good visualisation tool for qualitative tractography results. Same seeds as in Fibre Cup competition were used, and reconstructed fibres were compared with Fibre Cup ground truth.

The diffusion images of the Fibre Cup phantom come in two resolutions – 3 mm and 6 mm isotropic. Each voxel size was scanned using three different diffusion weighting factors – low ($b = 650$ s/mm$^2$), medium ($b = 1500$ s/mm$^2$), and high ($b = 2000$ s/mm$^2$ or $b = 2650$ s/mm$^2$ for $3mm$ and $6mm$ isotropic). All images were acquired on the 3T Tim Trio MRI systems. A single-shot diffusion-weighted twice refocused spin echo echoplanar pulse sequence was used. The Eddy current was corrected to the first order. A single T2 image was acquired, as well as a 64 diffusion sensitised images. The gradients were uniformly distributed over the sphere. There were two repetitions for 3 mm isotropic images, and one of 6 mm isotropic.

In order to qualitatively assess the viability of the optimised reconstruction, a streamline tractography [5, 82] was performed on the $b = 1500$ s/mm$^2$, 3 mm isotropic image. The image SNR was estimated to be 26, and as a consequence, accordingly with the Figure 3.9d, the regularisation parameter $\lambda^2$ was set to 0.002. For tractography algorithm, the minimal fibre length was constrained to

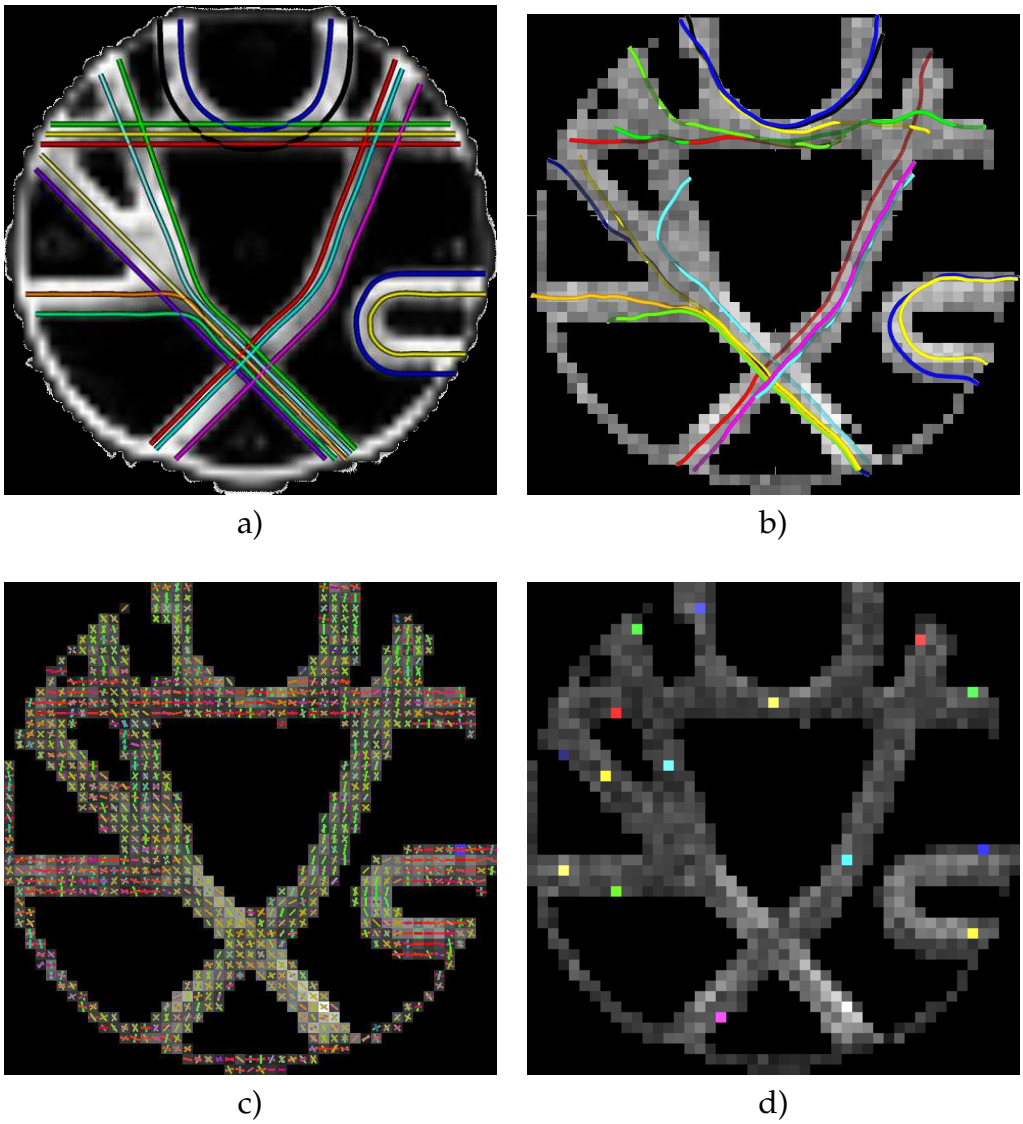



**Figure 3.18:** The Fibre Cup phantom ground truth (a) and tractography results (b). Reconstructed FODFs using optimised regularisation scheme (c) and seeds used to perform the tractography (d).

75 mm, and maximal to 500 mm (never reached due to phantom size). Other parameters include: maximal angle set to 45$^o$, subvoxel seeding, trilinear interpolation, and smoothing (momentum) set to 0.5.

Figure 3.18 shows the ground truth (a), tractography results (b), reconstructed FODFs (c), and seeds used to initiate the tractography (d). There is one fibre per seed, and each fibre has the same colour as its seed. Fibres are coloured based on the original Fibre Cup ground truth image (Figure 3.18a). Fibres representing ground truth (Figure 3.18a) are overlayed on top of GFA, while the reconstructed images (Figure 3.18b-d) are plotted on top of the quantitative anisotropy [85].

The phantom consist of a bending fibre (right, blue and yellow), a bending fibre that touches (kisses) another fibre (top, blue and black), a top horizontal fibre (green, yellow, and red), and four diagonal fibres. Three top-left to bottom-right diagonal fibres cross at one point, and then cross with the remaining top-right to bottom-left diagonal fibre.

As seen in Figure 3.18b, majority of seeds were resolved correctly. There is no problem with a right bending fibre (blue and yellow ) and both seeds produce expected fibres. Also both top bending and kissing fibres (blue and black) are correct. Furthermore, both top-left to bottom-right diagonal fibres (purple and yellow), and both middle-left to bottom-right fibre (orange and green) are correctly located, but end prematurely. One out of three top-right to bottom-left diagonal fibres (red) is resolved. The two remaining fibres, while correctly located, are truncated. The upper top-left to bottom-right fibre (cyan), although crudely shaped and slightly truncated, can be considered as a proper fibre. Also two horizontal fibres (green and red) are correctly, but not fully found. The upper-top to bottom-right diagonal fibre (green), and horizontal fibre (yellow) incorrectly leaks into wrong directions.

## 3.5 Diffusion Signal Reconstruction from Human Brain Image

In-vivo images of a 20 year old male patient were acquired using Philips MR Achieva R3.2; local research ethics committee approval and informed consent

were obtained. Image acquisition matrix was set to $224 \times 224$ with in-plane resolution of $2 \times 2$ mm$^2$. 48 slices were acquired with thickness of 2 mm. 32 diffusion weighted images with $b = 1000$ s/mm$^2$ ($TE = 54$ ms, $TR = 7803.283$ ms), and a single non-DW image were taken.

Images were corrected for the eddy current distortion using FSL [35, 71, 84], and the brain was extracted using the Brain Extraction Tool [70] (BET). Diffusion signal was reconstructed using a regularised 6th order SH series[1] and transformed into FODF (Figures 3.19-3.21). FODF profiles were colour-coded as explained in Section 2.3. Each anatomical direction is associated with one principle colour: blue red with medial-lateral direction, green with anterior-posterior direction, and blue with inferior-superior direction.

Three crossings in the brain have been highlighted: between the superior corpus callosum and the cingulum (Figure 3.19), between the corona radiata and superior longitudinal fasciculus at the centrum semiovale (Figure 3.20), and between the corona radiata and the corpus callosum at the centrum semiovale (Figure 3.21). Finally, a streamline tractography was performed to verify a global correctness of the reconstructed signal (Figure 3.22).

Both the superior corpus callosum and the cingulum can be seen in the Figure 3.19. The green FODFs that run posterior-anteriorly represent the cingulum (middle section of bottom Figure 3.19), while the red-blue FODF (left and right edges of bottom Figure 3.19) represent the corpus callosum and run medial-laterally and inferior-superiorly. Some voxels seem to contain both cingulum and corpus callosum – FODF is shaped as a crossing in superior-inferior and posterior-anterior directions. This is an apparent crossing caused by a limited MRI resolution.

The true crossing at the centrum semiovale is shown in the Figure 3.20b. The red-blue FODFs that mainly run in the superior-inferior direction represent the corona radiata (right section of bottom Figure 3.20). The superior longitudinal fasciculus, represented by green FODF profiles, is seen on the left and runs mostly in anterior-posterior direction.

The FODFs of apparent fibre crossing between the corona radiata and the corpus callosum at the centrum semiovale are shown in Figure 3.21. The corpus

[1] Longest possible SHS – there were 32 non-collinear diffusion encoding gradients and to use 8th order SHS at least 45 are necessary.

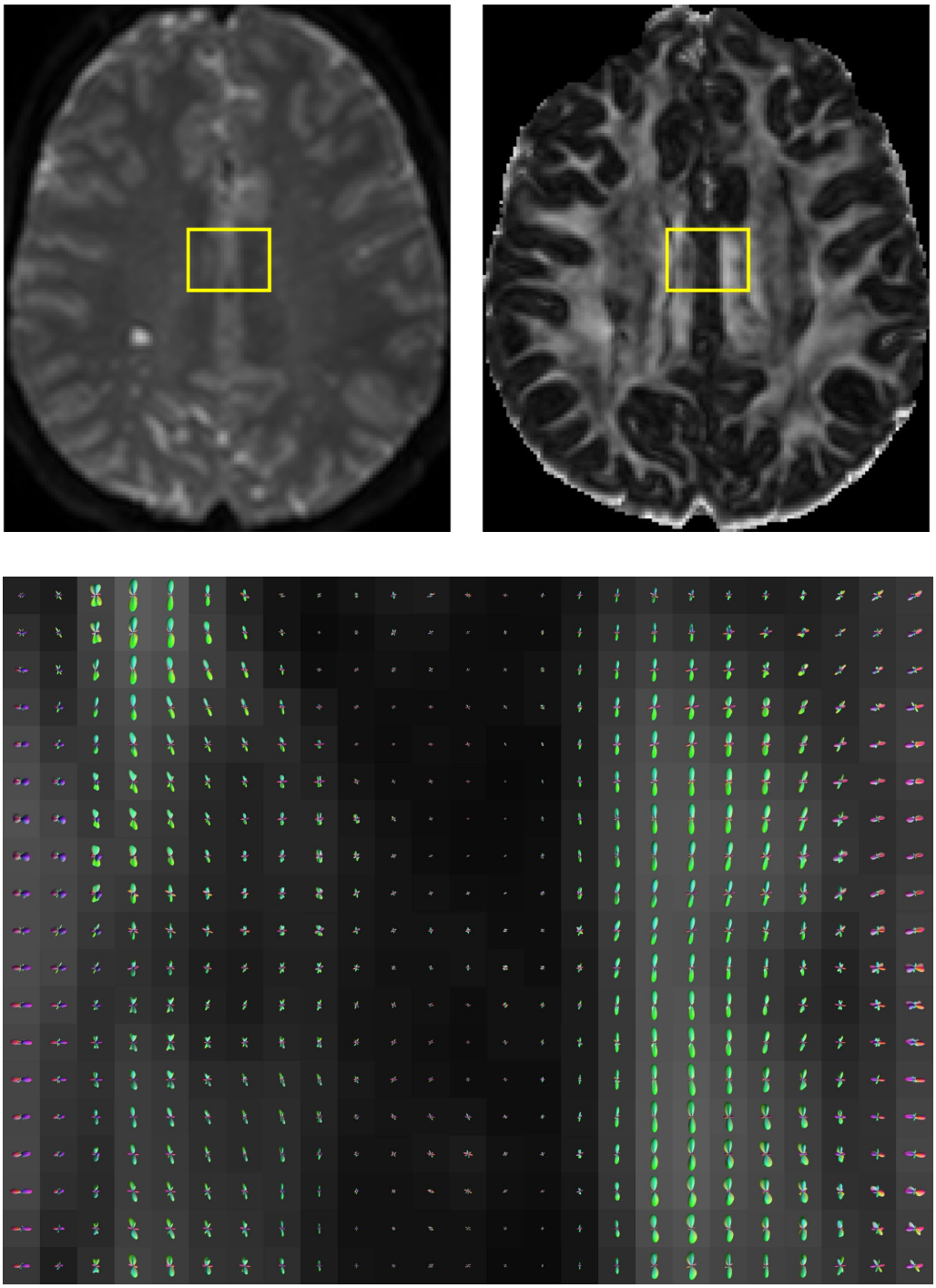


**Figure 3.19:** Fibre crossing between the superior corpus callosum and the cingulum: T2 (top-left) and GFA (top-right) images with a highlighted region of interest, and reconstructed FODFs (bottom).

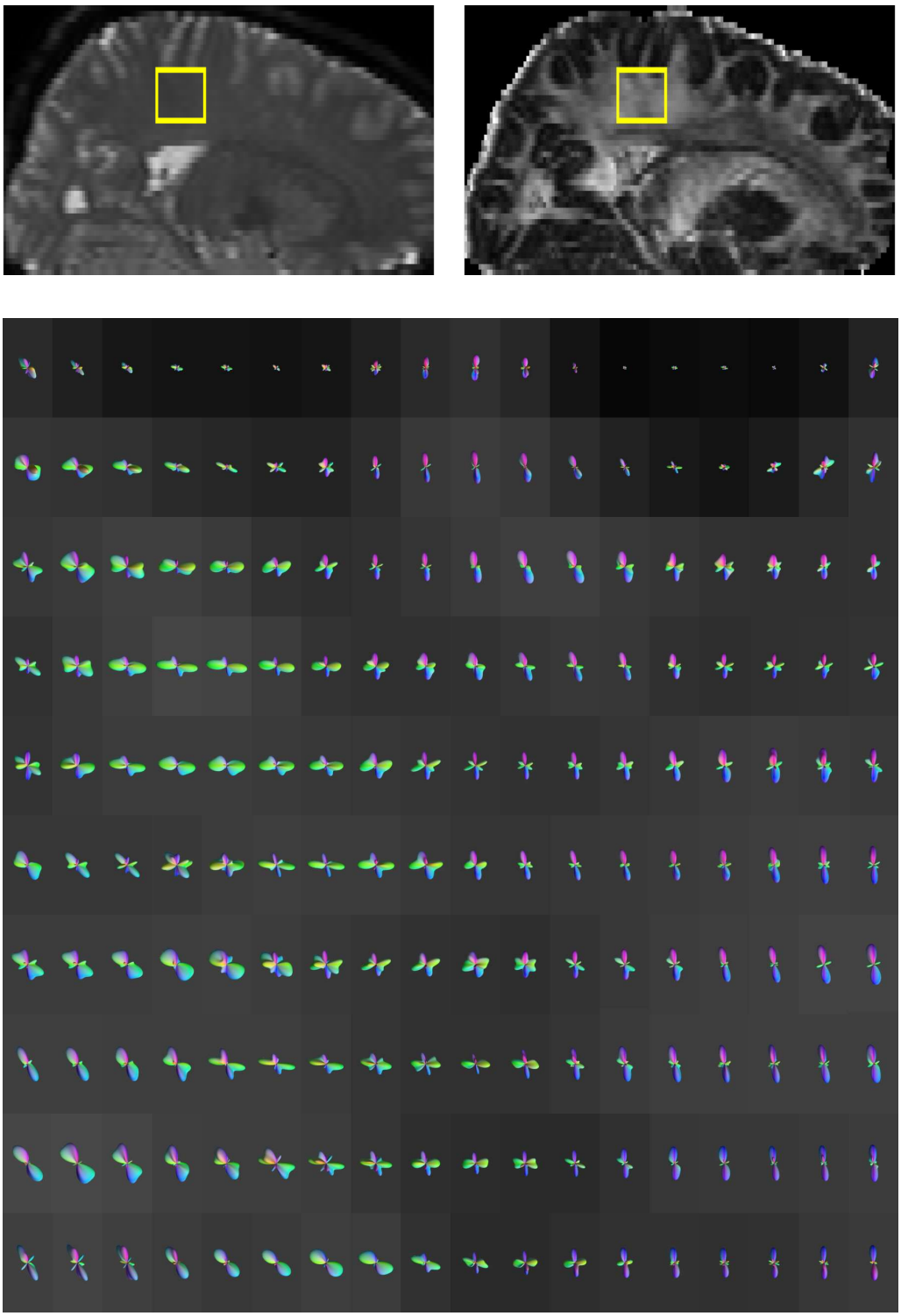


**Figure 3.20:** Fibre crossing between the corona radiata and longitudinal fasciculi at the centrum semiovale: T2 (top-left) and GFA (top-right) images with a highlighted region of interest, and reconstructed FODFs (bottom).

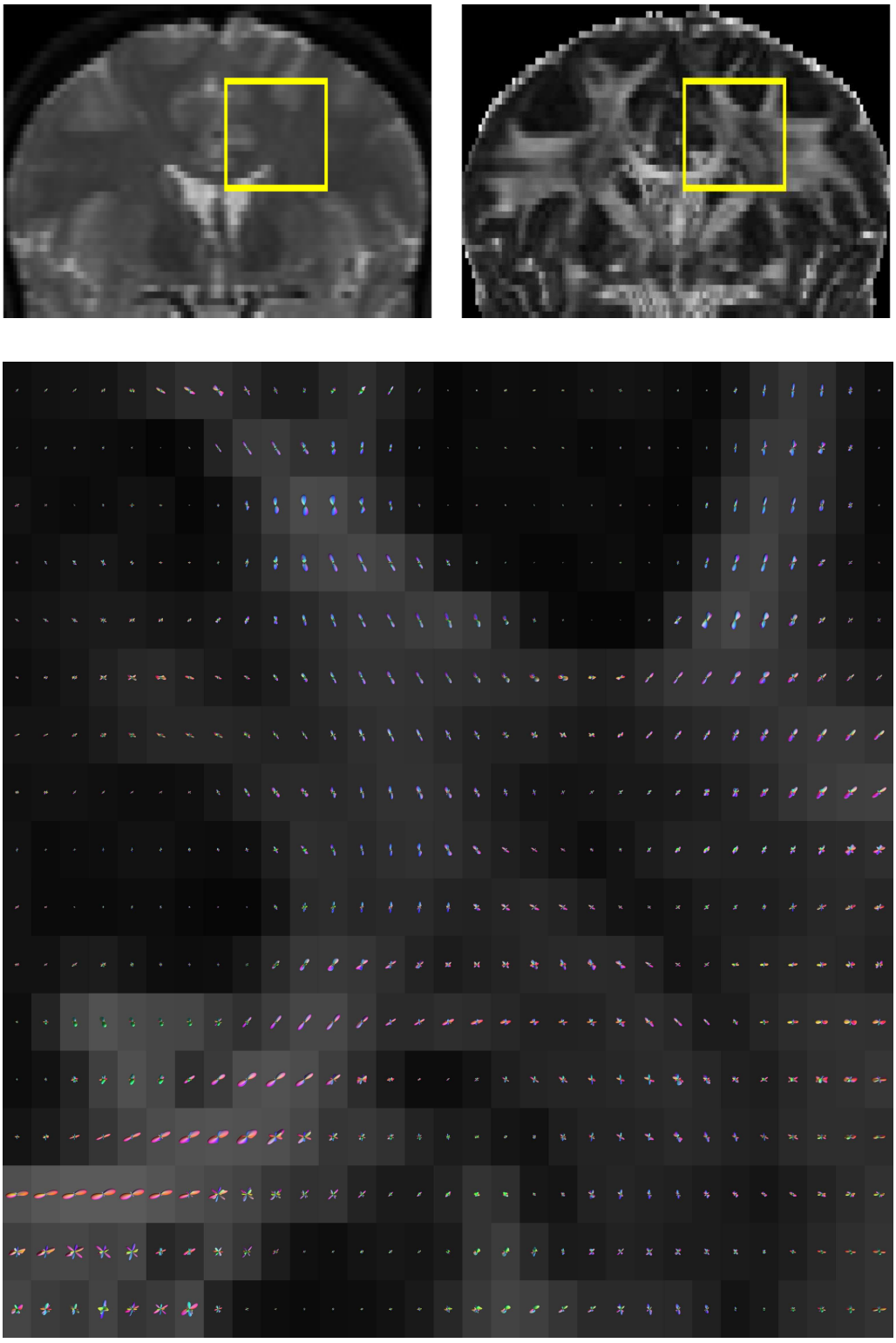


**Figure 3.21:** Apparent fibre crossing between the corona radiata and the corpus callosum at the centrum semiovale: T2 (top-left) and GFA (top-right) images with a highlighted region of interest, and reconstructed FODFs (bottom).

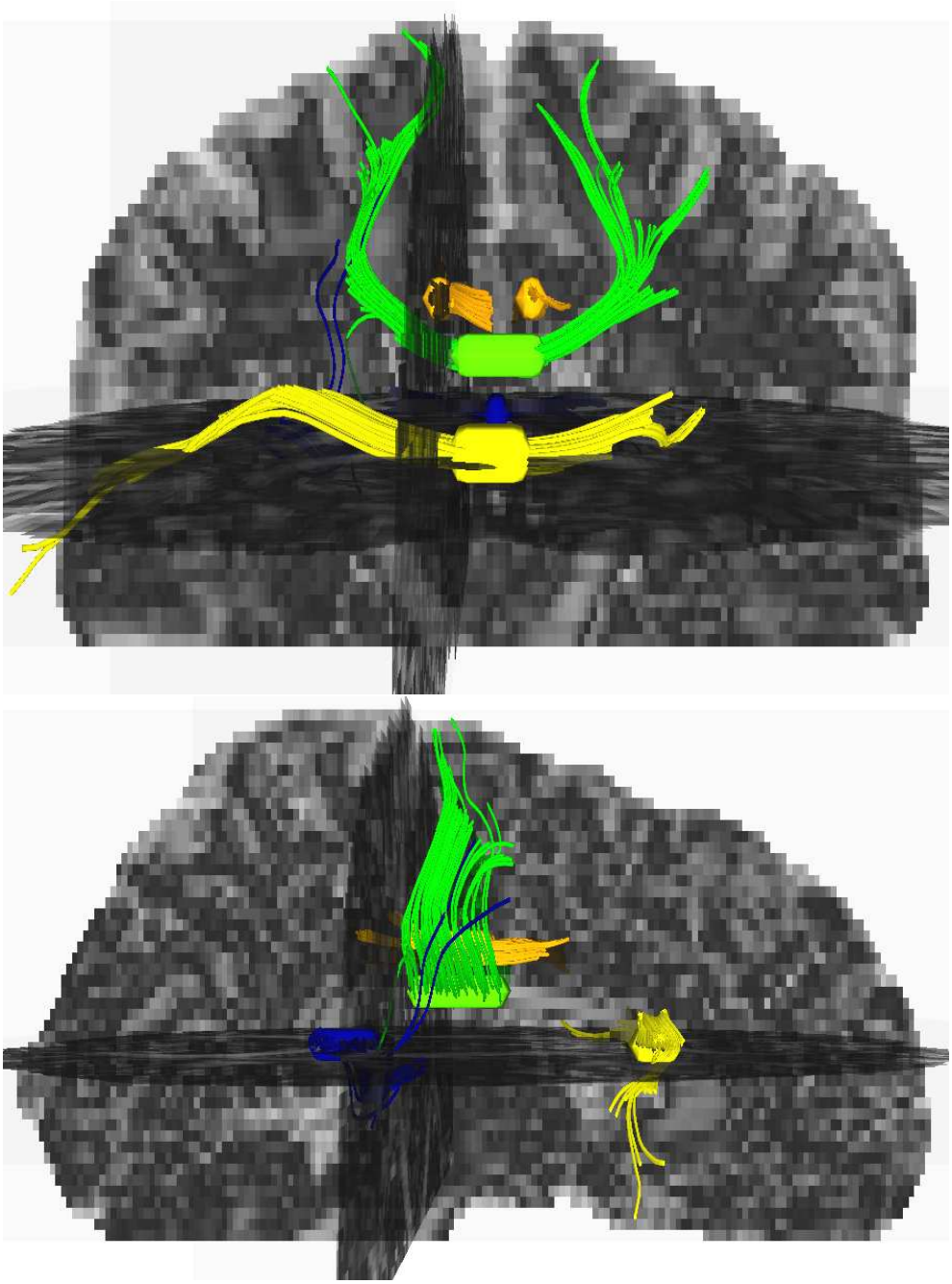


**Figure 3.22:** Fibre tracts reconstructed at cingulum (orange), and corpus callosum (green, blue, and yellow).

callosum is seen through red FODFs running in a medial-lateral direction (bottom left section of Figure 3.21). Corona radiata on the other hand, radiates in inferior-superior direction and medial-lateral directions. It can be seen as fanning, bending fibres represented as blue, red, and mixed (purple) FODFs.

In all three cases, the FODFs are oriented in anatomically correct directions. The direction transition across voxels is smooth, and continuous.

Finally, Figure 3.22 shows reconstructed fibre tracts of two brain structures: cingulum (orange), and corpus callosum (green, blue, and yellow). Orange and blue seeds are located within the region shown in the Figure 3.19, while yellow and blue are located in the genu and splenium sections of corpus callosum. For each seed 100 tracts were reconstructed using a streamline tractography. The maximal angle was set to $45^{o}$, minimal tract length to 25 mm, and forward momentum to 0.5 with trilinear interpolation, and subvoxel seeding.

## 3.6 Discussion

In this Chapter, the optimal method for selecting regularisation parameters was presented. The method fine-tunes the regularisation parameter for a given diffusion weighting factor and SNR. It performs a large number of Monte Carlo simulations in order to find the numerically optimal reconstruction. The regularisation values found through multiple exhaustive searches were different to those currently utilised ($L$-curve [18] or GCV [68]). Diffusion signal reconstruction using SH transform with regularisation controlled by the proposed method was on average more accurate (in worst case it was equally accurate) than with regularisation set up by other methods (namely $L$-curve).

The optimisation described in this chapter is performed independently of the image analysis. It can be therefore applied to any application where diffusion signal is already reconstructed using a regularised SHT. The only change for these applications would be in adjusting the regularisation parameter, by either using Figures 3.2-3.9 or by performing own optimisations similar to ones explained in Section 3.2.

The Tikhonov regularisation with the Laplace–Beltrami operator is an elegant way of regularising the measured signal within each voxel. In the Fourier do-

main, the method is computationally efficient and therefore viable for large data sets (up to few million voxels, with several dozen samples in each voxel). The GTV optimisation method offers a cost-free accuracy boost to any processing pipeline that involves the before mentioned regularisation.

The framework for selecting the regularisation strength presented in this chapter can be extended in many directions. The most obvious direction would be to explore different optimisation functions. For example, if one is interested in more accurate tractography, a complex ground truth that models intra-voxel behaviour should be used. On the other hand, if one has the access to a hardware phantom [66], it is possible to use the GTV method to "train" the optimisation for a given MRI scanner. The ground truth would be based on the hardware phantom, and the regularised reconstruction of that phantom scan should be optimised. This "profiling" of the scanner should theoretically result in a higher accuracy than the simulated GTV described in this chapter.

The angular regularisation is one of several techniques that can be used to improve the reconstruction accuracy. Spatial regularisation that forces to maintain signal continuity across voxels can be used in conjunction with angular regularisation. Apart from intra and inter voxel regularisation, modifications to SD are also possible. The major issue with both SD and SDT is the lack of non-negativity constraint. The reconstructed FODFs may, due to noise and ringing, exhibit negative regions. Another improvement in the accuracy of the reconstructed signal is thus possible by constraining the negativity of FODF, and propagating it back to HARDI signal [75]. This can be implemented on top of the angular and spatial regularisation, but requires iterative re-evaluation of each voxel and slows down the reconstruction step linearly.

# 4

# Statistical Inference for Reconstructing Diffusion Signal

The diffusion weighted signal acquired using clinically viable parameters suffers from low SNR, and poor spatial and angular resolution. Improving any of these properties increases the scanning time which, while acceptable in research environment, becomes very expensive in clinics. Recent research in a non-parametric diffusion signal processing was focused on a intra and inter voxel reconstruction of the signal, and involved either spatial or angular regularisation, and sometimes both.

In this chapter, an alternative to angular smoothing is presented. The proposed method is based on a stepwise linear discriminant analysis [1], which allows to select the most important predictors while retaining the highest discrimination power. Instead of modifying SH coefficients (*e.g.* regularisation schemes described in Section 2.4.3), a statistical approach to assess them in terms of significance is utilised [64]. Even without the ground truth data, it is possible to estimate how necessary each basis function is to represent the signal, and how

much it is affected by noise. As such, depending on the selected threshold, the frequency can be deemed significant to signal representation and retained, or insignificant and removed from the signal. The resulting SH series will therefore contain only most useful basis functions and be able to filter noise from the signal. Furthermore, the remaining SH basis function can still be regularised, further improving signal reconstruction accuracy.

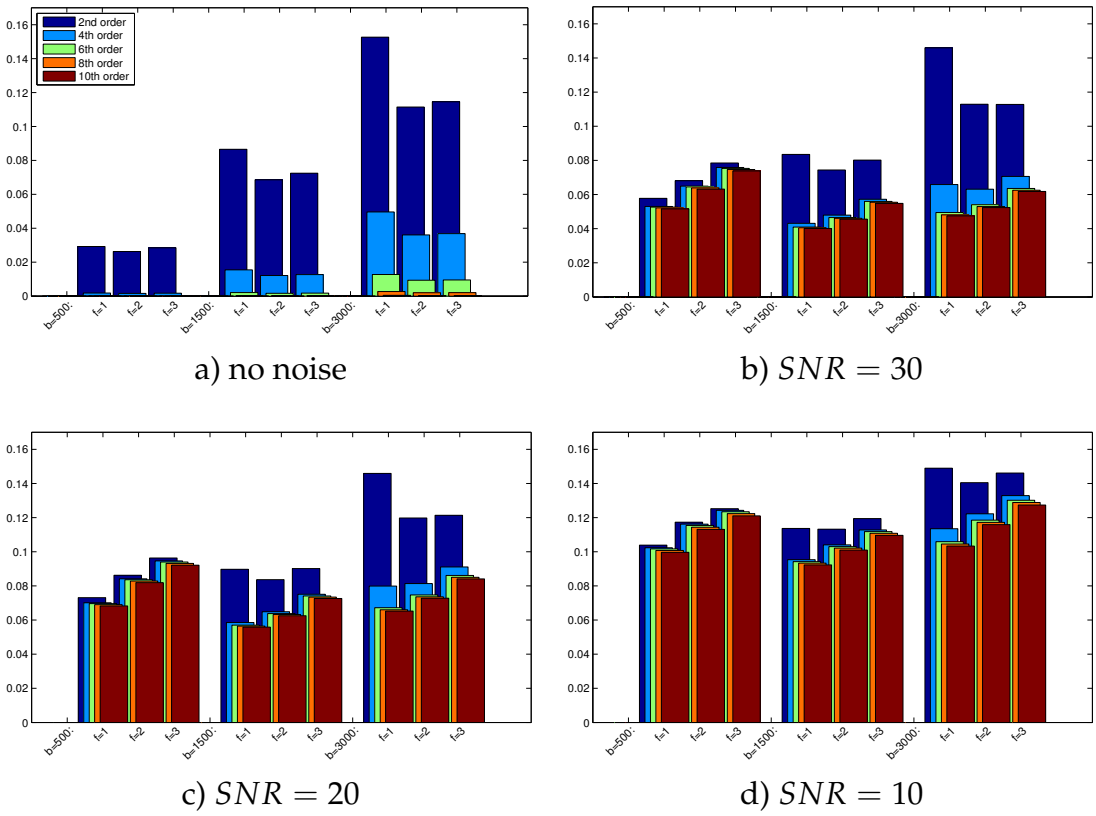
This chapter is organised as follows. Section 4.1 highlights the issue of using long SH series to represents diffusion signal. Monte Carlo simulations are used to test the average accuracy of the signal representation as a function of the number of SH basis functions. Additionally, the average magnitudes of SH coefficients, and a noise influence on SH transform is explored. In Section 4.2, a statistical method for elimination of the least significant SH basis function using a Student's *t*-test is introduced. Section 4.3 introduces stepwise regression, as well as three methods for combining the statistical inference with angular regularisation is described in Chapter 3. Additionally, the accuracy of backward elimination regression in diffusion signal reconstruction is tested. GFA maps and full brain tractography of a human brain are presented in Section 4.4, and show the viability of backward elimination regression in medical imaging. Finally, conclusions and discussion are presented in Section 4.5.

## 4.1 Spherical Harmonic Analysis of Noisy Signal

Spherical harmonics, as globally supported spherical functions, can be used to approximate any spherical signal. The signal is represented as a superposition of basic waves, more often called basis functions. By having enough basis functions $\mathbf{Y}$, it is therefore possible to represent a diffusion signal $\mathbf{E}$ with a certain error $\boldsymbol{\epsilon}$:

$$\mathbf{E} = \mathbf{Y}\mathbf{c} + \boldsymbol{\epsilon} \, . \tag{4.1}$$

At the same time though, the SH analysis is a generalisation of a Fourier analysis, and the SH basis functions represent a spherical frequencies of the signal. As in the Fourier transform, the higher the order of the basis function, the higher the frequency of the signal this basis function represents. Therefore, any truncated SH series can be considered as an accurate representation




a) no noise b) $SNR = 30$

c) $SNR = 20$ d) $SNR = 10$

**Figure 4.1:** Normalised root-mean-square deviation for SHT of diffusion signal with low (left triplets), medium (middle triplets), and high (right triplets) diffusion weighting applied. Each triplet shows NRMSD for a 2nd to 10th order SHT of diffusion signal. Diffusion signal was generated with no noise (a), and with increasing corruption (b-d).

of the signal, up to the truncated frequency. The truncation of the series is enforced by the limited number of samples, from which the transform is computed. Due to noise, in order to have a numerically stable linear regression the number of parameters (SH coefficients) must be no greater than the number of linearly independent samples.

Figure 4.1 shows the mean of normalised root-mean-square deviation (NRMSD) of approximated signal as a function of SH order. The diffusion signal was generated for three diffusion weighting factors $b = 500$ s/mm$^2$, $b = 1500$ s/mm$^2$, and $b = 3000$ s/mm$^2$, each 10000 times for 1, 2 and 3 fibre crossings independently. The discrete signal was simulated at 1000 uniformly distributed points on a sphere. Four cases are presented: one with no noise (Figure 4.1a), and three with increasing noise (Figure 4.1b-d, SNR of 10, 20, and 30).

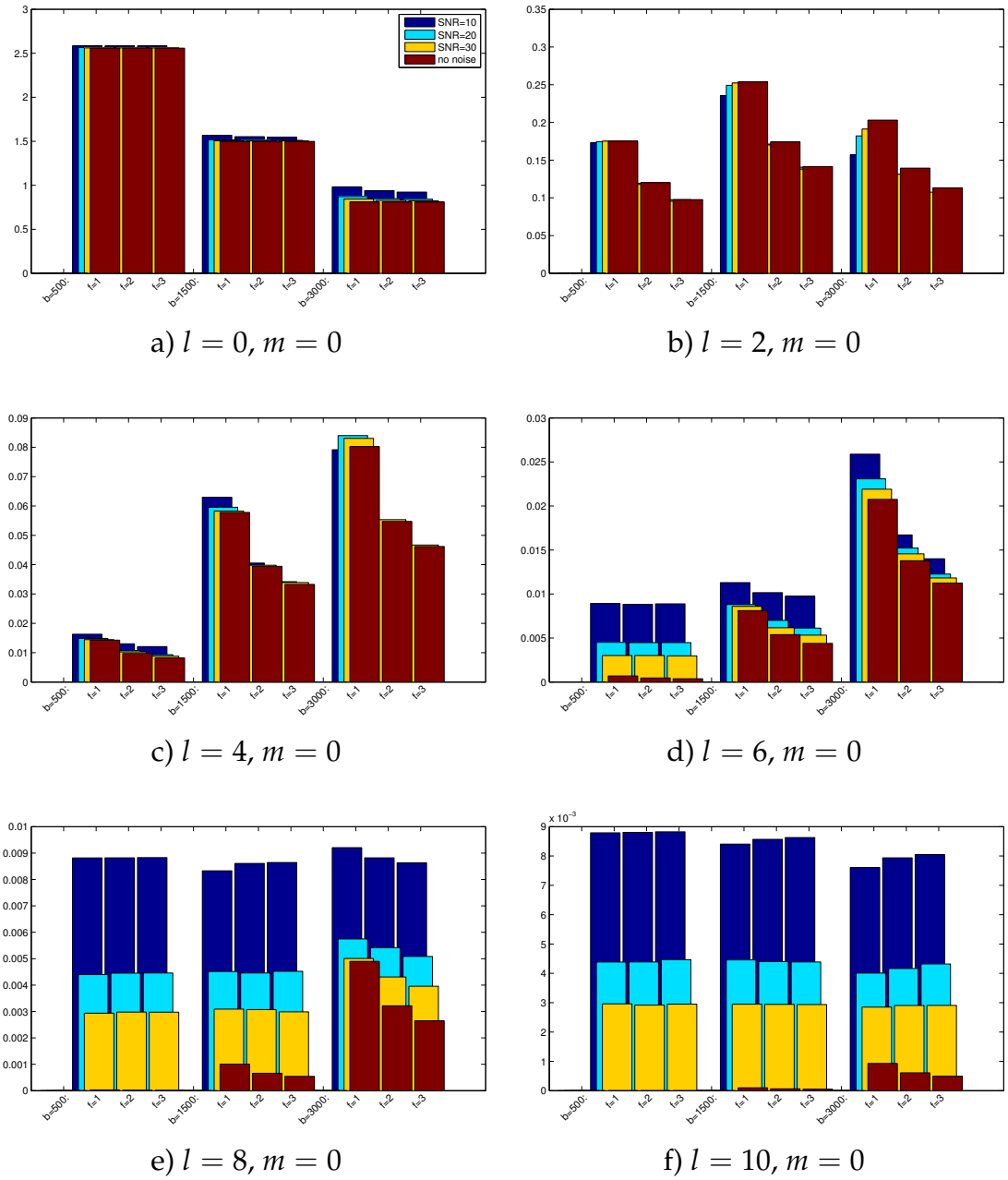
Even with a low diffusion weighting factor, a 2nd order SH series does not

allow to accurately approximate the diffusion signal. Across all 9 tests (low, medium, and high diffusion weighting factor and one, two, and three fibre crossings) the single fibre configuration proved to be the most difficult to accurately approximate (highest error). As the diffusion weighting factor increases, the signal becomes sharper, meaning more basis functions are needed to accurately approximate the signal. With a medium diffusion weighting factor, 4th order SH series, and with a high, the 6th order series become inaccurate. The difference in NRMSD across the number of fibres crossing is marginal. In general, 6th and 8th order SH series are sufficient to accurately approximate the diffusion signal simulated using a middle and high diffusion weighting factor.

With the addition of noise, the error drastically increases. Even with a high SNR, any higher order SH series ($l > 2$) becomes, at best, only twice more accurate than the 2nd order SH series. As the noise is always random (following Rician distribution), higher order basis functions become more affected. For low diffusion weighting factor, regardless of the fibre configuration simulated, the higher order basis functions no longer help in signal representation (same error as for 2nd order SH series). With medium and high diffusion weighting applied, the 4th order basis functions are still useful (reduce the error). Even with reliable measurements ($SNR = 30$), any SH series above 6th order will not improve the approximation accuracy.

Depending on SNR, the 6th or 8th order SH series can be considered as optimal. However, even when truncating the approximation to just several basis functions the approximation error is still several times higher than ground truth (approximation of noiseless signal). To some degree, medium SH basis functions react to noise present in signal introducing false regularities in approximation. Figure 4.2 presents the average magnitude of the main SH coefficient $m = 0$ for a set of orders $l \in \{2, 4, 6, 8, 10\}$. Both noise free and corrupted signals for low, medium, and high diffusion weighting factors are investigated, as well as 1, 2 and 3 fibre crossing configurations. To produce statistically significant results, each magnitude is averaged from SH transform of 10000 independent voxels.

The initial SH basis functions of up to 4th order can be considered stable. Regardless of the presence and strength of noise, the average magnitude of those frequencies remains constant. With higher frequencies though ($l \geq 6$), noise influence starts to be seen in the SH coefficients. It can be seen that at low diffusion weighting factor, even with a high SNR, the average magnitude of the




a) $l = 0, m = 0$ b) $l = 2, m = 0$

c) $l = 4, m = 0$ d) $l = 6, m = 0$

e) $l = 8, m = 0$ f) $l = 10, m = 0$

**Figure 4.2:** Average magnitudes of main SH basis functions, at increasing frequencies. Noiseless (red), and corrupted signal analysis (blue, cyan, and yellow).

6th order coefficients is several times larger than those of noise free equivalent. This noise susceptibility escalates at higher frequencies, especially when lower diffusion weighting is applied.

With low diffusion weighting factors, the average magnitude of 10th order basis functions for SH transform of noisy signal is over 1000 times larger than it should be. This means that those basis functions are mostly representing noise, not the underlying diffusion signal. As a result, the higher order terms need to be saturated (see Chapter 3) or not used at all.

## 4.2 Statistical Inference and Linear Regression

### 4.2.1 Rationale

Stepwise regression [33] is a set of regression models that allow for automatic selection of predictive variables. The algorithms can be divided into three approaches: forward selection, backward elimination, and bi-directional elimination. Forward models usually start with no predictors other than a constant value, and over iterations add variables that improve the model the most. The process stops when there are no more useful variables. Backward elimination methods work exactly the opposite, they start with a full model, and over iterations remove those variables that are the least useful. Bi-directional algorithms combine both approaches, either adding the most useful variable, or removing the least significant one in each iteration. In this Section, an alternative to series truncation, based on backward elimination method is proposed.

Instead of removing all corrupted $l$th order basis function, a selective procedure through the use of $t$-test can be utilised [53]. The initial regression model is found using a classical, unregularised least squares method. Next, for each predictive variable $c_i$, $t$-scores are computed. The decision whether the variable is insignificant or not is made based on the lowest $t$-score. If the variable is deemed insignificant, it is removed from the regression model, and $t$-scores for remaining variables are recomputed. The procedure is repeated until only significant variables are left.

As mentioned in Section 2.4.1, the SH transform can be implemented and solved as a linear regression (Equation 2.16). It is therefore possible to use one of the linear regression $t$-tests – test for regression slope – to tell whether there is no correlation between $E$ and $c$ [1]. The null hypothesis states that the slope (SH coefficients $c$) is equal to 0 (is insignificant), and alternative that it is not equal to 0 (is not insignificant):

$$\begin{aligned} H_0 &: c_i = 0 \\ H_1 &: c_i \neq 0 \,. \end{aligned} \tag{4.2}$$

The null hypothesis has to be tested for each coefficient $c_i$.

Before $t$-values for each coefficient $c_i$ can be found, it is necessary to obtain the standard error for each of the least-squares estimators:

$$SE(c_i) = \sqrt{\frac{MSE}{\sum_{k=1}^{N}(X_{ik} - \bar{X})^2}} \,, \tag{4.3}$$

where $MSE$ is the mean squared error, which in regression analysis is an unbiased estimate of noise variance:

$$MSE = \frac{\sum_{k=1}^{N}(E_k - \hat{E}_k)^2}{N - L} \,. \tag{4.4}$$

In the formula above, the nominator is simply a sum of squared residuals, while the denominator is the sample size $N$ reduced by the number of parameters $L$ used in the regression. Having the standard error, it is now possible to compute $t$-scores for each coefficient $c_i$:

$$t_i = \frac{c_i}{SE(c_i)} \,. \tag{4.5}$$

The last step of the $t$-test is to use the $t$-scores to obtain the $p$-values. In statistics, the $p$-value is a probability of obtaining a result equal to, or more extreme, than what was observed. If the $p$-value for $i$th coefficient is lower then the predetermined significance value $\alpha$ (usually 0.05), the null hypothesis is rejected. As a result, the basis function corresponding to the coefficient $c_i$ is considered as significant in signal representation, and should be kept.

It is important to note that the $t$-test is a test for marginal significance of the

[1] Alternatively, the partial $F$-test can be used for the same results.

predictor $c_i$. Since it takes into account all other $c_j$ ($i \neq j$) used to represent the variable $E$, any change in $c_j$ will change the significance of $c_i$. However, the SH basis functions are orthonormal and the potential removal of $c_i$ does not change the actual values of $c_j$ [2]. As a consequence, the least significant basis function will stay least significant, even if other basis functions are removed (the $p$-values will change). This allows forming the iterative backward elimination method, where in each iteration the least significant basis functions is removed from the signal representation.

### 4.2.2 Regularised Backward Elimination Regressions

There are multiple ways of combining the backward elimination regression with regularisation. The most straightforward approach is to use regularised coefficients in Equation 4.5. For further references, the SH series obtained using classical backward elimination will be referred to as backward elimination (BE), and the regularised backward elimination will be called regularised backward elimination (RBE).

Alternatively, backward elimination can be used to classify basis functions as either significant or insignificant. Based on this classification, two possible algorithms emerge. First, only significant basis functions are used, but the coefficients itself are regularised [57]. Another way to combine regularisation with elimination is to use full SH series, but regularise only the insignificant coefficients. For further reference, former method will be called significant-regularised backward elimination (SBE), while latter insignificant-regularised backward elimination (IBE). Their respective formulations are given by:

$$\mathbf{c}_{SBE} = \mathbf{c}_{TR} \circ \text{sgn}|\mathbf{c}_{BE}| \,, \tag{4.6}$$

$$\mathbf{c}_{IBE} = \mathbf{c}_{TR} \circ (1 - \text{sgn}|\mathbf{c}_{BE}|) + \mathbf{c}_{BE} \,, \tag{4.7}$$

where $\mathbf{c}_{TR}$ are regularised SH coefficients, and $\circ$ denotes a Hadamard product.

[2] The numerical orthonormality of the discrete SHT is enforced through the sampling scheme and, depending on the sampling scheme used, may not be maintained.

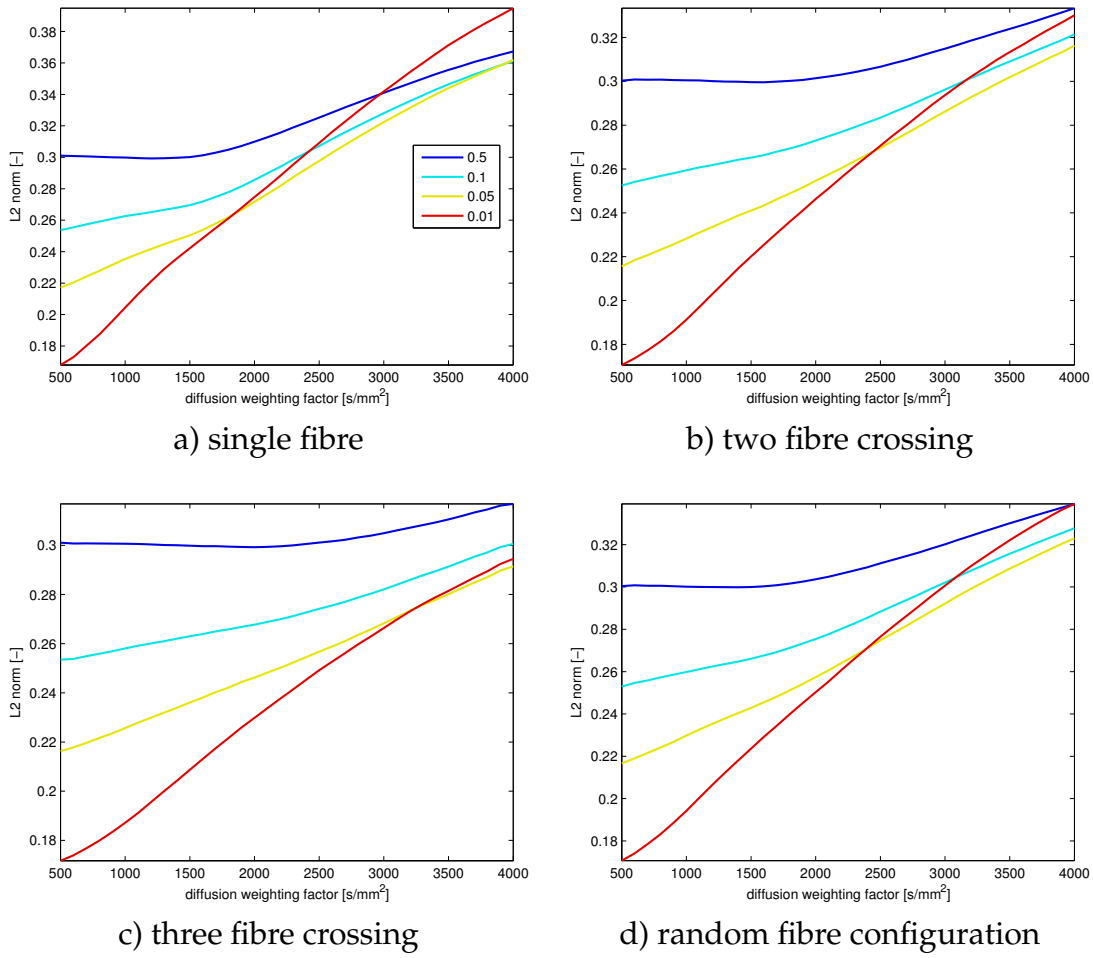



**Figure 4.3:** Reconstruction error as a function of diffusion weighting factor for low SNR at different significance levels.

## 4.3 Quantitative Evaluation

To test the influence of backward elimination on accuracy of diffusion signal representation, exhaustive numerical simulations were performed. Diffusion signal was generated for multiple fibre configurations (one, two, three, and random fibre crossing), different SNRs, and diffusion weighting factors. Simulated, noisy signal was analysed with a SHT, and unreliable basis functions were rejected using backward elimination procedure. Reconstruction error of reduced SH series as a function of diffusion weighting factor is presented in Figures 4.3 and 4.4. Four significance levels were tested, $\alpha \in \{0.5, 0.1, 0.05, 0.01\}$ for both low (10) and high SNR (30). In each case, error was averaged over 10000 independent simulations.

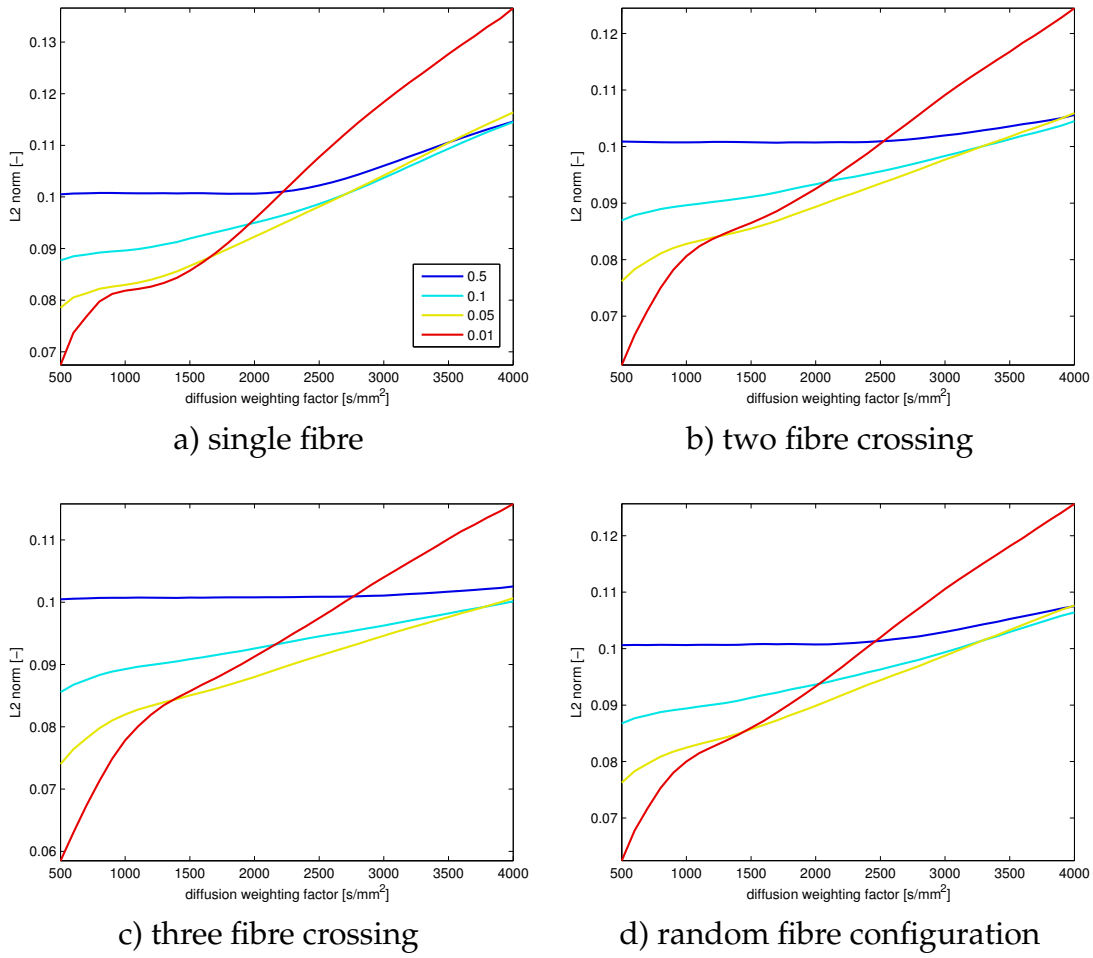



**Figure 4.4:** Reconstruction error as a function of diffusion weighting factor for high SNR at different significance levels.

When diffusion signal is simulated with low diffusion weighting applied it is more beneficial to use low significance level. The lower the probability of incorrectly rejecting the null hypothesis is, the more basis functions are removed from signal representation. In a low changing signal this improves the accuracy of signal reconstruction. Keeping a high significance level with a high diffusion weighting though, increases the probability of Type II error. When the null hypothesis is incorrectly accepted, the accuracy of signal representation suffers, and a higher significance level should be used. For further reference, a point at which two lines representing the errors of two significance levels cross is called "critical" point. To the left of critical point a lower significance level should be used, to the right – higher.

The location of critical point depends on the SNR use to generate diffusion sig-

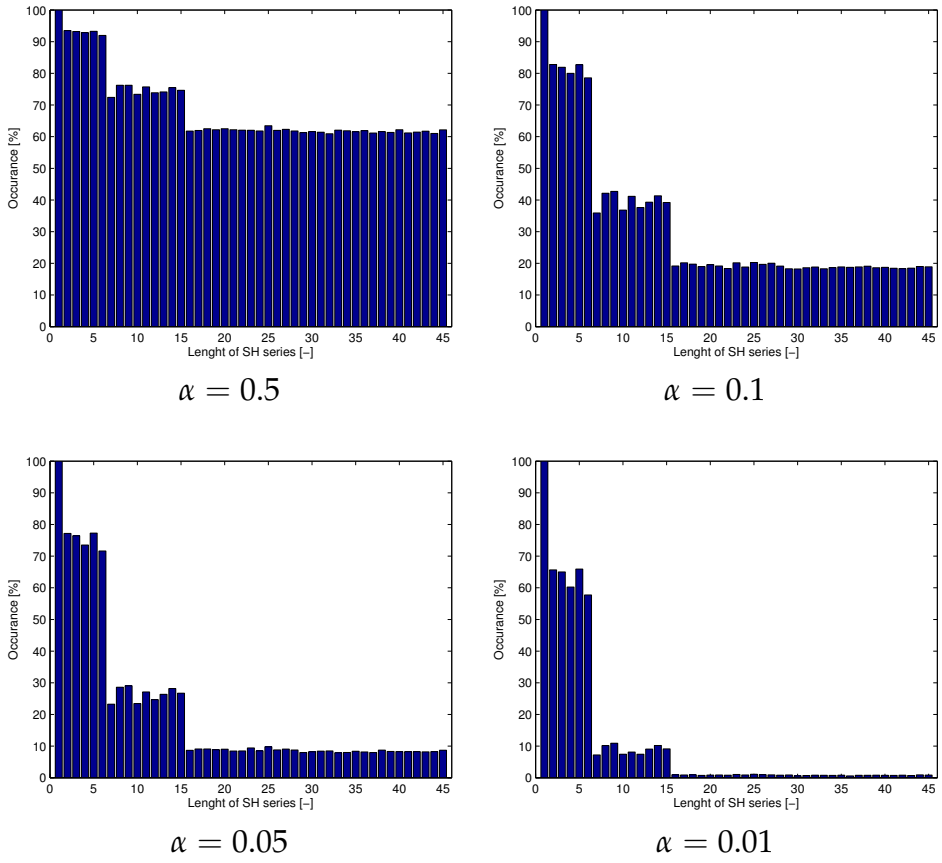



**Figure 4.5:** Basis function relevance for four significance levels, low SNR.

nal, and underlying fibre configuration. At a given diffusion weighting factor, the higher the SNR becomes the less basis functions can be safely removed from Equation 2.18, meaning a higher significance level has to be selected. For example, it becomes more beneficiary to use $\alpha = 0.05$ over $\alpha = 0.01$ starting from $b = 2300$ s/mm$^2$ when reconstructing a low quality ($SNR = 10$) diffusion signal coming from two fibre crossing. With a high quality signal though ($SNR = 30$), it is better to use higher significance level as soon as $b = 1200$ s/mm$^2$. Additionally, depending on the fibre configuration, the critical point is located at different location, *e.g.* as the number of fibre crossings increase, the point moves toward right.

Percentage of how often each SH basis functions is used after backward elimination process is applied (basis function relevance) is shown in Figures 4.5 and 4.7, while the distribution of a number of basis functions used in a signal representation (polynomial degree) in Figures 4.6 and 4.8. Backward elimination regression with significance value $\alpha \in \{0.5, 0.1, 0.05, 0.01\}$ was performed for low and high SNRs on a 10000 diffusion signal reconstruction of a ran-

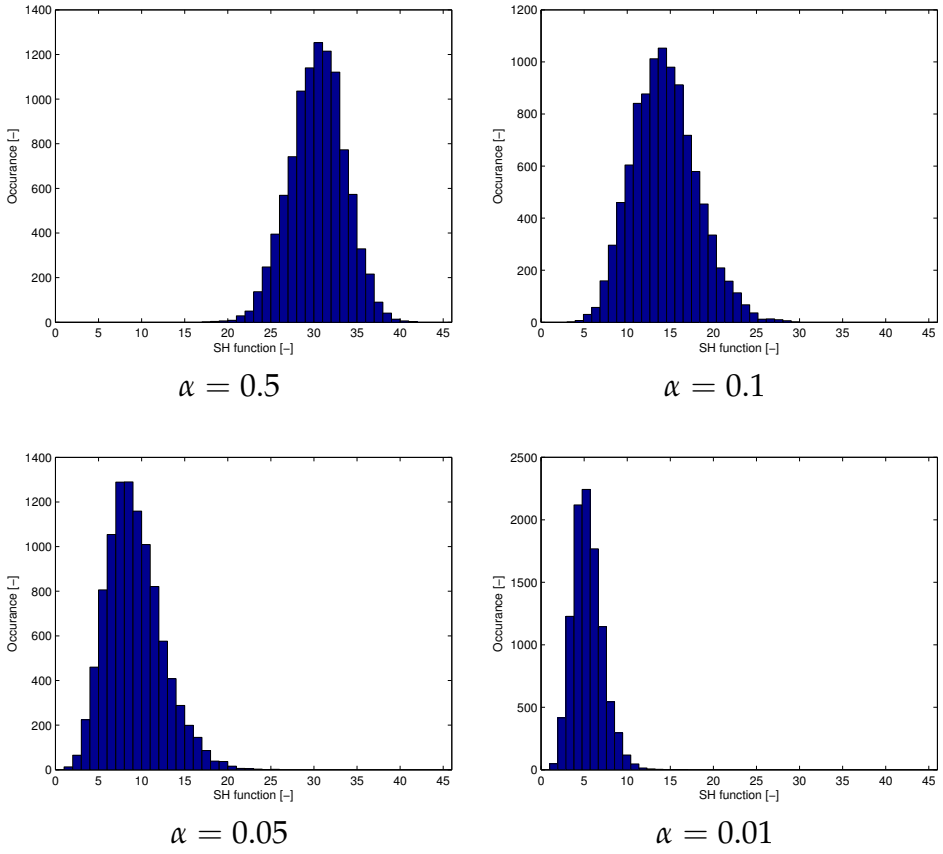



**Figure 4.6:** Distribution of the polynomial degree for four significance levels, low SNR.

dom fibre configuration. Diffusion signal simulated using a medium diffusion weighting factor $b = 1500$ s/mm$^2$.

The relevance of a basis function depends on its order, significance level used, as well as SNR. The 2nd order basis functions (2–6) are robust to noise, and regardless of SNR or significance level, are used in 90% of cases. Higher order SH basis though, are more sensitive to noise, and thus less relevant. The relevance of 4th order basis functions (7–15) is, depending on the significance level, 10% to 70% for low, and 40% to 70% for high SNR simulations. The higher order basis functions become less relevant.

When SH functions are used to represent the diffusion signal, only antipodal symmetric bases (bases of even order) are needed. As such, any full $n$th order SHS will have $(n+1)(n+2)/2$ terms. This means there are 45 different basis functions used when the signal is approximated using 8th SH series. Furthermore, due to backward elimination, SH series can be reduced to just few basis functions (Figure 4.6 and Figure 4.8, polynomial degree). The lower the signif-

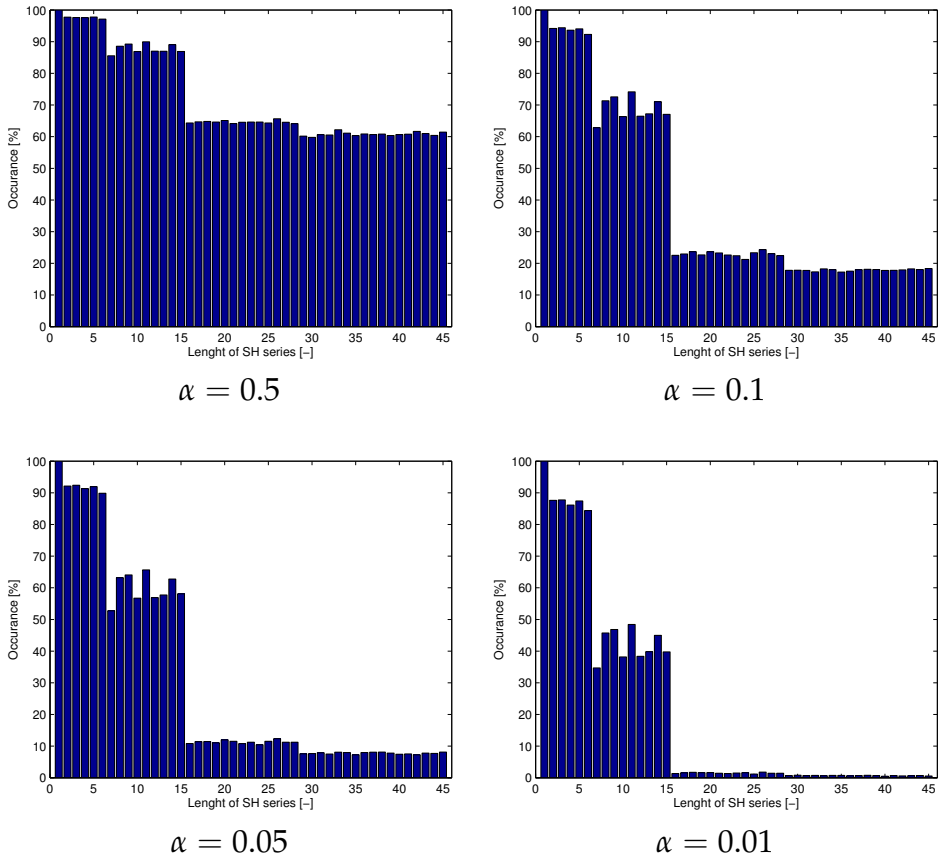



**Figure 4.7:** Basis function relevance for four significance levels, high SNR.

icance value is, the less basis functions will be used. Additionally, the number of statistically significant bases is tied to SNR; and with low SNR, more basis functions are rejected from the signal representation.

Although the backward elimination regression improves the accuracy of signal reconstruction, in terms of absolute error it is not as good as Tikhonov regularisation. The optimised Tikhonov regularisation described in Section 3.2 can be up to two times more accurate than backward elimination (Figure 3.11a vs Figure 4.3d and Figure 3.11e vs Figure 4.4d). Since these two methods work in different domains (backward elimination removes frequencies, while Tikhonov regularisation saturates them) they can be combined into regularised backward elimination regression.

Finally, Figures 4.9 and 4.10 show the reconstruction error for all four methods: BE (solid), RBE (dash and dot), SBE (dash), and IBE (dot). Significance levels are coded with different colours, red is used for 0.01, yellow for 0.05, cyan for 0.1, and blue for 0.5. Black solid line is used to represent the error

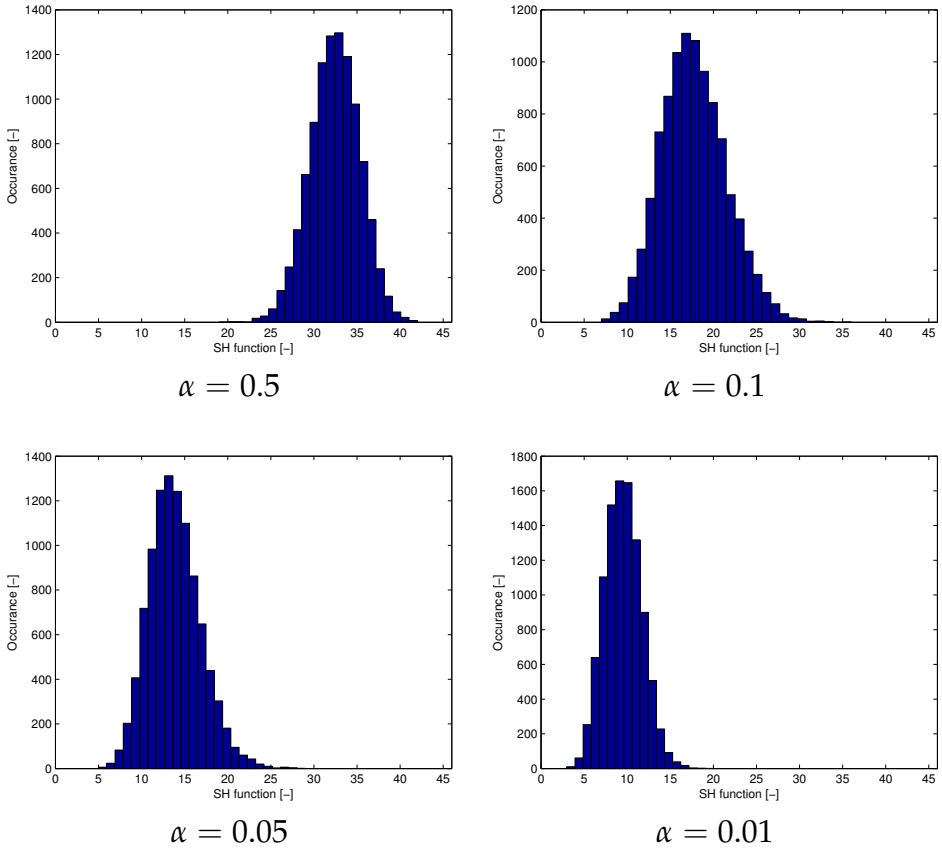
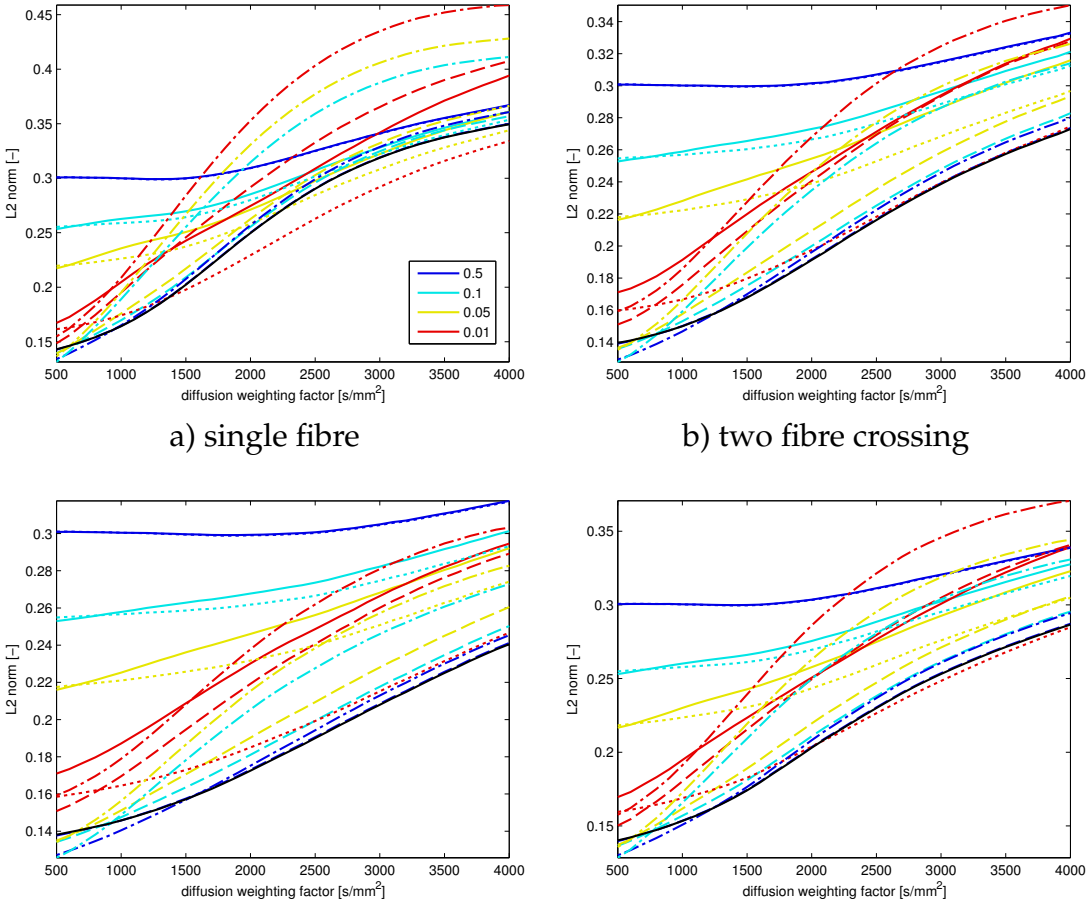



**Figure 4.8:** Distribution of the polynomial degree for four significance levels, high SNR.

of a full regularised signal reconstruction.

In general, the Tikhonov regularisation by itself produces good reconstruction accuracy for either low SNR or low diffusion weighting factors. As soon as SNR increases and a medium diffusion weighting is applied the SBE method becomes the most accurate (dot lines). Lowest significance values for this method result in the most accurate reconstruction.

Out of all methods, direct application of Tikhonov regularisation to backward elimination regression almost always gives the worst results. This can be seen especially with a high SNR (Figure 4.10), when RBE performs even worse than BE. The only application for this method may be found with high significance values (blue dash and dot line) at low diffusion weighting factors ($b < 1700\,\mathrm{s/mm^2}$), as in these cases it can accurately represent the diffusion signal.



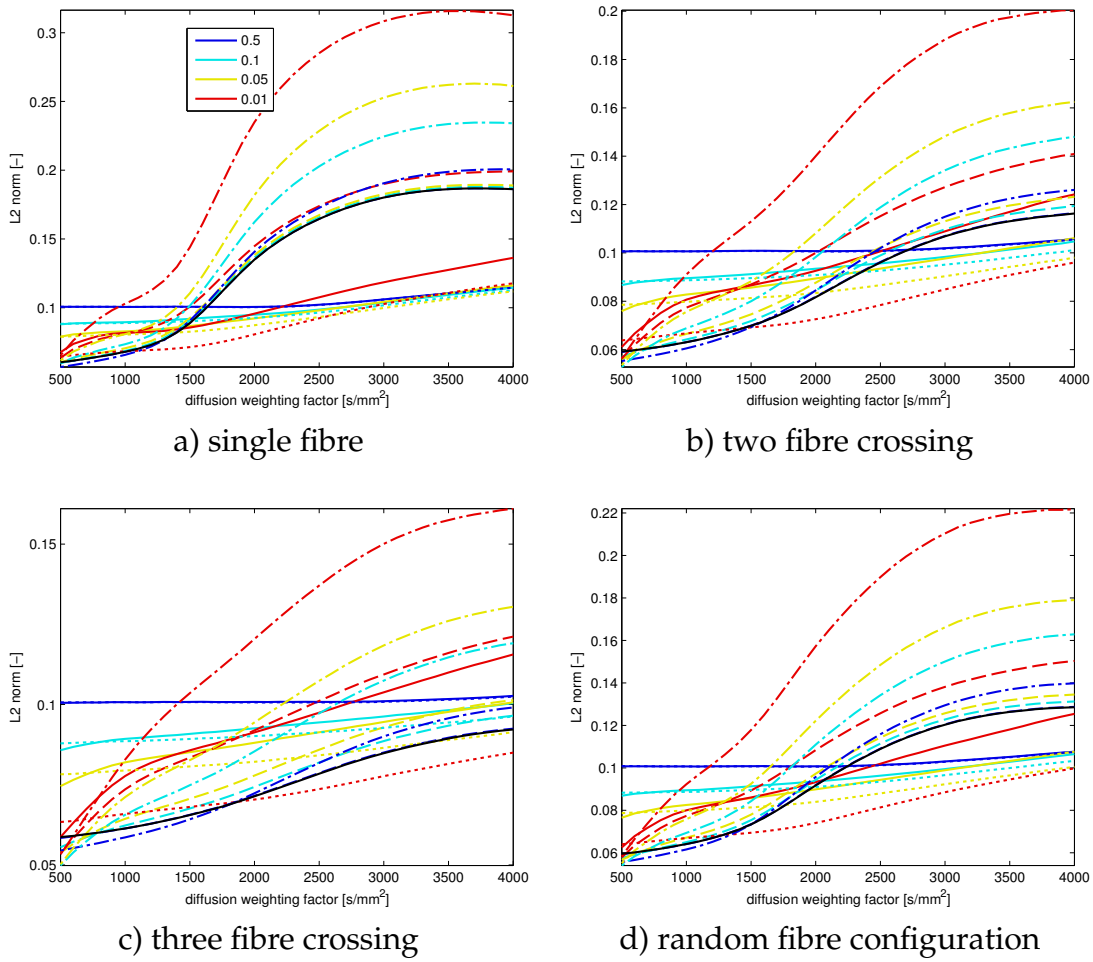

a) single fibre b) two fibre crossing

c) three fibre crossing d) random fibre configuration

**Figure 4.9:** Reconstruction error for backward elimination methods (various coloured) and Tikhonov regularisation (solid black), low SNR. Four backward elimination methods are compared: BE (solid coloured), RBE (dash and dot coloured), SBE (dash coloured), and IBE (dot coloured).

## 4.4 Backward Elimination Regression in Brain Imaging

In-vivo images of a healthy human brain were acquired using Philips MR Achieva R3.2; local research ethics committee approval and informed consent were obtained. Image acquisition matrix was set to $224 \times 224$ (in-plane resolution of $2 \times 2$ mm$^2$, $TE = 72$ ms, $TR = 15,292$ ms), and 52 slices were acquired with thickness of 2 mm. 61 diffusion weighted images with $b = 3000$ s/mm$^2$, and 6 non-DW images were taken. Eddy current correction and brain extraction was done using FSL and BET software. Diffusion signal was reconstructed using an 8th order SH series. SNR was estimated to be between 18 and 21,




a) single fibre

b) two fibre crossing

c) three fibre crossing

d) random fibre configuration

**Figure 4.10:** Reconstruction error for backward elimination methods (various coloured) and Tikhonov regularisation (solid black), high SNR. Four backward elimination methods are compared: BE (solid coloured), RBE (dash and dot coloured), SBE (dash coloured), and IBE (dot coloured).

and as such significance level $\alpha$ for backward elimination was set to 0.01, and regularisation value $\lambda^2$ for Tikhonov regularisation was set to 0.002.

Figure 4.11 shows GFA maps computed from two SH series – full regularised and classical BE regressions, as well as the number of non-zero SH basis functions after performing backward elimination (white representing high count). BE removes corrupted frequencies from signal representation, which means that GFA maps generated from reduced SH series have better contrast. The number of non-zero basis functions used to reconstruct the signal for voxels containing isotropic signal is very low, and often SH series is reduced to a single basis function (sphere). The highest number of basis functions is seen in the voxels with medium to high anisotropy. In these voxels SNR is high and more

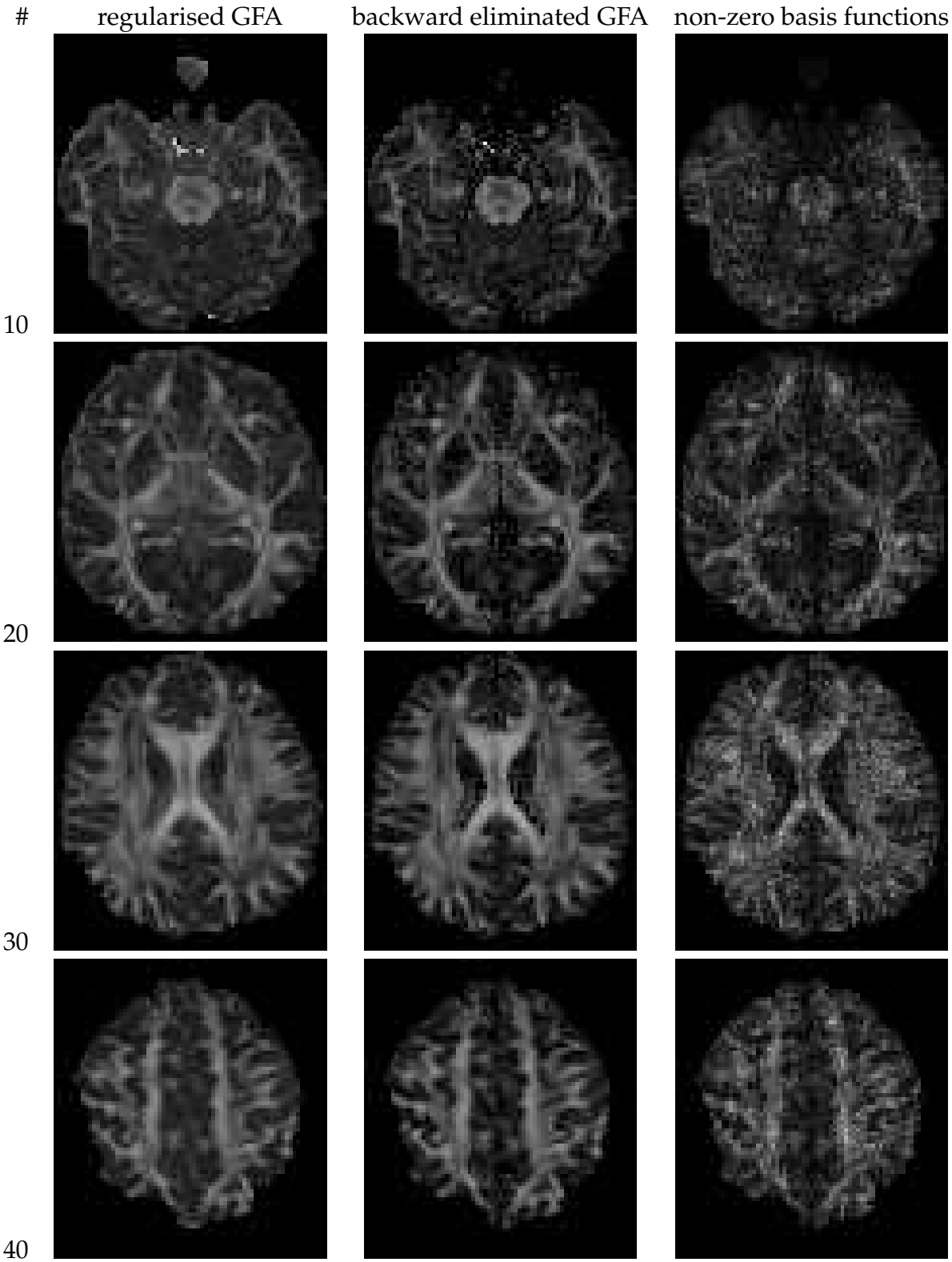



**Figure 4.11:** GFA maps computed from regularised and backward elimination regressions, and the non-zero basis functions map for backward elimination. Inferior-superior view, 10th, 20th, 30th and 40th slice.

basis functions are deemed significant. As a result, non-zero basis functions map closely correlates with the GFA map, and thus brain structure.

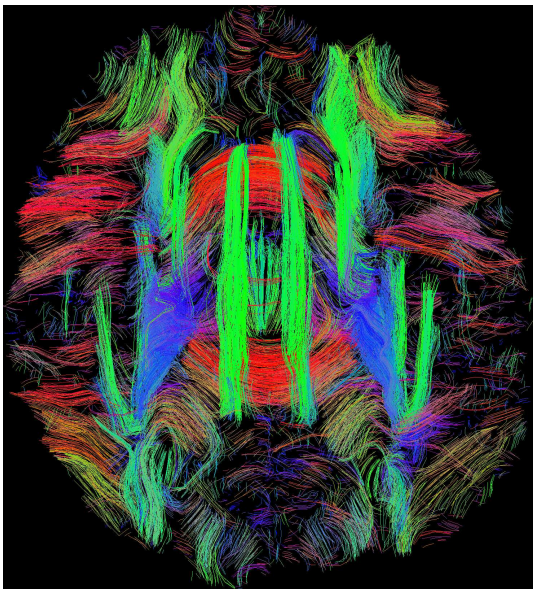
Full brain tractography is presented in Figure 4.12. Four methods were used to reconstruct the signal: full regularised regression and three backward elimination algorithms (BE, SBE, and IBE). All tractography result can be considered correct and look similar to each other, although there is a difference in quantities of reconstructed tracts for regularised and IBE regressions, and BE and SBE.

The first two methods are less restrictive, and produce longer tracts. Mean tract length is 11.4688 mm and 11.4578 mm for regularised and IBE methods, while mean tracts for BE and SBE methods are 8.78166 mm and 8.75327 mm long (maximal tract are 105.71 mm, 115.916 mm, 107.549 mm, and 107.577 mm for regularised, IBE, BE, and SBE methods respectfully). The longer tracts produced by regularised and IBE regressions are often resolved as multiple shorter tracts when using BE and SBE regressions, most likely due to their restrictive signal representation (use of reduced or reduced and regularised SH series instead of full regularised or full and partially regularised SH series). As a result, the number of tracts resolved for regularised and IBE regressions are 156282 and 159482, while for BE and SBE is almost twice as much – 247611 and 244768.

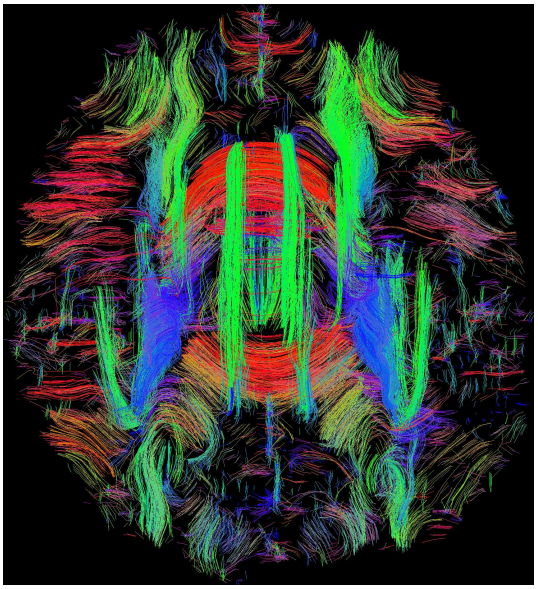
## 4.5 Discussion

In this chapter, an alternative to the widely used angular regularisation schemes based on backward elimination regression was presented. The method, which is part of the stepwise regression family, uses Student's $t$-test to automatically select the most reliable basis functions. In an iterative process, the least significant basis function is removed from the signal representation one at a time, or more precisely, the coefficient associated with it is set to 0. The algorithm terminates once only significant basis functions are left.

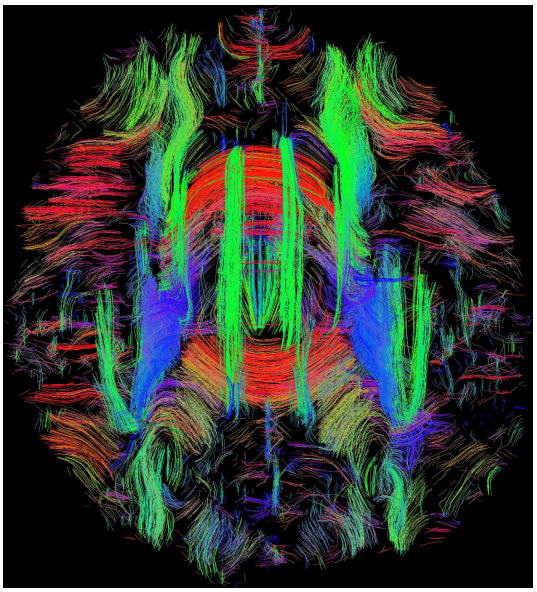
In Section 4.3 numerical evaluation of this method was presented. The method, by removing most noise affected basis functions, consistently improves the signal representation accuracy. However, its performance is not as good as popularly used regularisation schemes based on Tikhonov regularisation [19,



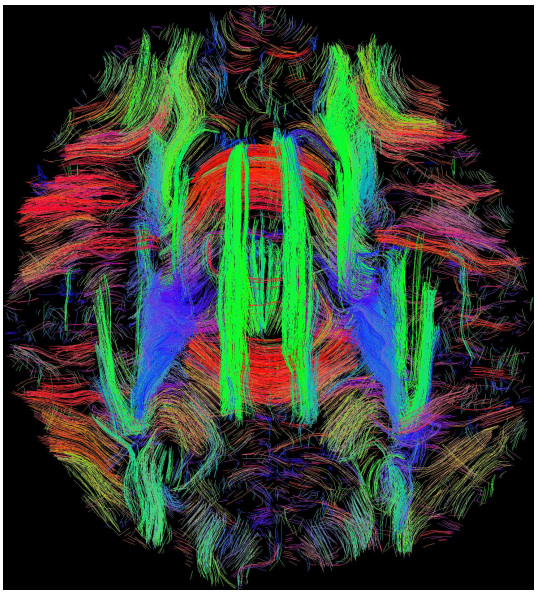
a) full regularised regression



b) backward elimination regression – BE



c) backward elimination regression – SBE



d) backward elimination regression – IBE

**Figure 4.12:** Full brain tractography, inferior-superior view of 27th slice.

32]. Therefore, three additional methods that incorporate angular regularisation into backward elimination regression had been developed. Depending on the quality and parameters of the scan, either first (RBE), or last (IBE) algorithm should be used for best results.

Additionally, a new measure based on the number of non-zeroed coefficients (number of basis functions used) was introduced. Maps generated using this measure coincidence with GFA maps, and in general with white matter organisation in the brain (Figure 4.11). They could potentially be used in a further processing. One of the possible application could be in developing a reliability-aware algorithms. For example, a tractography algorithm can use this information to decide how strongly the signal in each voxels should affect tract reconstruction. Another possible use for this information could be a selective denoising algorithm, where the denoising strength is correlated with the reliability of signal in each voxel.

Finally, in Section 4.4 the viability of the backward elimination regression and the developed methods in real image processing is addressed. Diffusion signal of a full brain scan (SNR between 18 and 22, and diffusion weighting factor $b = 3000$ s/mm$^2$) was represented using a backward elimination reduced 8th order SH series. GFA maps computed from backward eliminated SHS, thanks to the partially removed noise, have a better contrast than those obtained from just regularised SH transform. Three out of four introduced methods, due to SNR and diffusion factor requirements, were deemed viable: BE, SBE, and IBE. Full brain tractography that was performed on the signal represented using those methods proved to be consistent with the tractography performed on the signal represented using classically regularised SH series.

It is important to note that backward elimination methods require significant processing power, as each voxel has to be analysed multiple times. Due to the high computational requirements and only marginal increase of reconstructed signal accuracy this method may not be viable for clinical applications. The method can also be used for the new measure it introduces – number of non-zero coefficients. However, as there are no clinical experiments which use this measure its usefulness is not known.

# 5

# Bayesian Regression and Spherical Harmonic Analysis

In Chapter 3 the reconstruction of diffusion signal was improved by regularising higher frequencies with a Laplace–Beltrami smoothing operator. The signal reconstruction was then further improved by introducing statistical inference and significance testing in Chapter 4. In both cases though, the coefficients of SH transform were found using ordinary least squares (OLS) method (Equation 2.17). In this Chapter, an alternative method using Bayesian inference is used.

As stated in previous chapters, signal reconstruction (approximating signal values in-between measuring points) can be regarded as a linear regression problem. The dependent variable (regressand) is expressed by explanatory variable (regressors). In diffusion signal reconstruction the measured signal is a regressand, and SH basis functions are regressors. The relationship between regressand and regressors is defined by a parameter vector (regression coefficients), often found using OLS. Bayesian regression extends the linear regression by introducing a additional knowledge about the relationship between dependent

and explanatory variables. Regression coefficients will be therefore affected by the dependent variable and a prior knowledge, which should result in a superior accuracy of the diffusion signal reconstruction.

This chapter is organised as follows. In Section 5.1 Bayesian regression and the construction of prior knowledge are introduced. Numerical evaluation and phantom simulations are presented in Section 5.2. Section 5.3 shows the application of Bayesian regression and prior knowledge to diffusion MRI image of human brain. Finally, Section 5.4 concludes this chapter discussing advantages and shortcomings of described method as well as further work.

## 5.1 Bayesian Regression

By using Bayesian regression it is possible to infer about unknown parameters using some known (prior) information. The information can be based either on data (*e.g.* past observations) or beliefs (*e.g.* predicted behaviour). For a linear regression model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \,, \tag{5.1}$$

with unknown regressors $\boldsymbol{\beta}$, and $\epsilon \sim N(0, \sigma^2)$ with unknown $\sigma^2$. The conjugate prior for $\boldsymbol{\beta}$ and $\sigma^2$ is the multivariate normal inverse gamma distribution, $N - \Gamma^{-1}(\mathbf{m}, \mathbf{V}, a, b)$:

$$f(\boldsymbol{\beta}, \sigma^2 | \mathbf{m}, \mathbf{V}, a, b) = |\mathbf{V}|^{-1/2} (2\pi)^{-p/2} \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2}\right)^{p/2+a+1} \exp\left(-\frac{2b + (\boldsymbol{\beta}-\mathbf{m})^{\mathrm{T}} \mathbf{V}^{-1} (\boldsymbol{\beta}-\mathbf{m})}{2\sigma^2}\right) \,, \tag{5.2}$$

with mean $\mathbf{m}$, covariance $\sigma^2\mathbf{V}$, shape $a$, scale $b$, and $p$ being the number of regressors.

The posterior distribution is simply another multivariate normal inverse gamma distribution $N - \Gamma^{-1}(\mathbf{m}^*, \mathbf{V}^*, a^*, b^*)$, with parameters:

$$a^* = a + n/2 \,, \tag{5.3}$$

$$b^* = b + \{\mathbf{m}^T\mathbf{V}^{-1}\mathbf{m} + \mathbf{y}^T\mathbf{y} - (\mathbf{m}^*)^T(\mathbf{V}^*)^{-1}\mathbf{m}^*\}/2\,, \tag{5.4}$$

$$\mathbf{m}^* = \mathbf{V}^*(\mathbf{V}^{-1}\mathbf{m} + \mathbf{X}^T\mathbf{y})\,, \tag{5.5}$$

$$\mathbf{V}^* = (\mathbf{V}^{-1} + \mathbf{X}^T\mathbf{X})^{-1}\,, \tag{5.6}$$

with $n$ being the number of observations.

Alternatively, the conjugate prior can be expressed as a multiplication of two distributions: normal defined by mean $\mathbf{m}$ and covariance $\sigma^2\mathbf{V}$, and inverse gamma defined by shape $a$ and scale $b$. Similarly, the posterior distribution can be decomposed into two distributions: $N(\mathbf{m}^*, \mathbf{V}^*)$ and $\Gamma^{-1}(a^*, b^*)$.

When a prior knowledge is weak, the diagonal elements of matrix $\mathbf{V}$ tend to go to infinity, rendering $\mathbf{V}^{-1}$ close to zero. The posterior then becomes:

$$\mathbf{m}^* = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}\,, \tag{5.7}$$

$$\mathbf{V}^* = (\mathbf{X}^T\mathbf{X})^{-1} \tag{5.8}$$

which is equal to ordinary least squares solution. Similar behaviour should be established, when the data is completely reliable ($\sigma^2 = 0$). It can be achieved by using the unbiased estimate of noise $\hat{\sigma}^2$ to scaling the $\mathbf{V}^{-1}$:

$$\mathbf{m}^* = \mathbf{V}^*(\hat{\sigma}^2\mathbf{V}^{-1}\mathbf{m} + \mathbf{X}^T\mathbf{y})\,, \tag{5.9}$$

$$\mathbf{V}^* = (\hat{\sigma}^2\mathbf{V}^{-1} + \mathbf{X}^T\mathbf{X})^{-1}\,. \tag{5.10}$$

The posterior distribution is therefore affected by both data and prior knowledge, with data reliability controlling the ratio. It represents how likely regressors $\boldsymbol{\beta}_x$ are, given the observed data. The highest probability is obtained for $\boldsymbol{\beta}_x = \mathbf{m}^*$, and therefore posterior mean is a natural candidate for unknown regressors. Combining Equations 5.9 and 5.10 it can thus be written:

$$\boldsymbol{\beta} = (\hat{\sigma}^2\mathbf{V}^{-1} + \mathbf{X}^T\mathbf{X})^{-1}(\hat{\sigma}^2\mathbf{V}^{-1}\mathbf{m} + \mathbf{X}^T\mathbf{y})\,. \tag{5.11}$$

### 5.1.1 Prior Knowledge for Diffusion Signal Reconstruction

In order to apply Bayesian regression to diffusion signal reconstruction prior information has to be identified. Since the diffusion signal can be accurately

simulated using mathematical models, a numerical simulation is used to generate a set of noise free SH transforms. For a given diffusion weighting factor and a set of diffusion coding vectors, a random fibre configuration is selected. Diffusion signal is generated using a multiple tensor model (Equation 2.6) and the SH transform of the signal is computed. SH coefficients $c$ are stored, and the process is repeated $N = 1,000,000$ times. The prior knowledge is expressed in the form of prior distribution defined by its mean $\mu$ and variance $V$ for each coefficient $k$:

$$\mu_k = \frac{1}{N}\sum_{i=1}^{N} c_{ki} \,, \tag{5.12}$$

$$V_k = \frac{1}{N}\sum_{i=1}^{N} (c_{ki} - m_k)^2 \,. \tag{5.13}$$

Figure 5.1 shows prior distribution, computed for low ($b = 1000\ s/mm^2$) and high ($b = 3000\ s/mm^2$) diffusion weighting factors. The first basis function ($l = 0$, $m = 0$) is omitted for visualisation purposes ($\mu = 1.81$ for low, and $\mu = 0.69$ for high diffusion weighting factors). Both mean and variance exhibit the same trend regardless of the diffusion weighting applied. The variance of coefficients decreases as the order of basis functions increases. Additionally, most coefficients are equally likely to be positive or negative ($\mu = 0$). Some coefficients however, are always either positive or negative ($\mu \neq 0$ and $\mu \gg V$). For example, 9th and 11th coefficients stay negative regardless of the fibre configuration and diffusion weighting factor that was used to generate diffusion signal.

The prior distribution defined by $\mu$ and $V$ gives general information on the behaviour of SH transform of the diffusion signal. Therefore, it can be used to constrain noise influence on diffusion signal reconstruction (Equation 5.11).

### 5.1.2 Localised Prior Knowledge

The prior distribution generated in Section 5.1.1 describes a global behaviour of SH basis functions. It was created from a set of all possible fibre configurations and therefore does not capture any local properties that exist. However, by restricting the fibre orientation for which diffusion signal is created, it is possible to check for potential local properties of SH analysis.

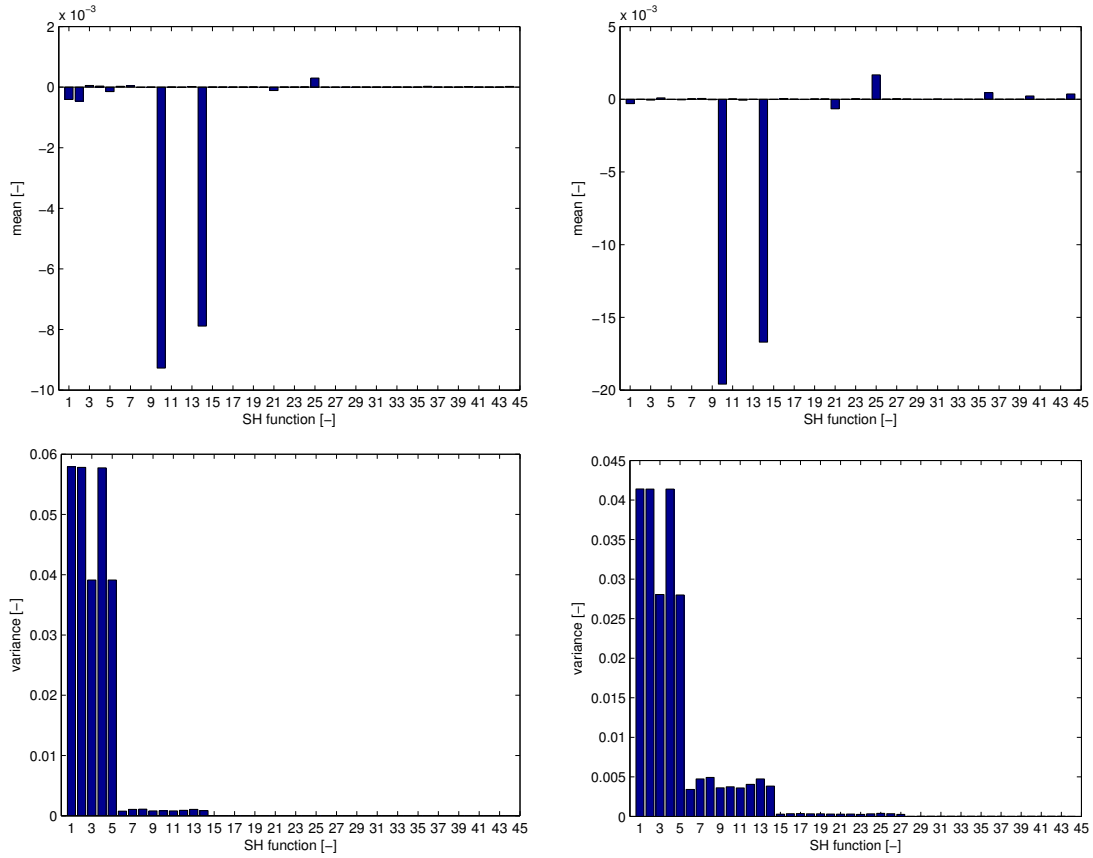



**Figure 5.1:** Mean (top) and variance (bottom) of prior distribution for low (left) and high (right) diffusion weighting factors.

An experiment can be constructed, which explores some properties of SH analysis. For a set of single fibre voxels, where fibre directions are all on the same plane, it can then be observed that the coefficients of SH analysis of the diffusion signal for this set of voxels will change in a predictable manner. For example, consider five fibres in the same plane, each rotated by 5° from the previous one. For each fibre, diffusion signal using MTM is generated, and SH analysis is performed. The middle fibre is selected as a reference, and every SH coefficient is normalised using the largest coefficient around the reference:

$$\hat{c}_k = \frac{c_k - c_{10}}{\max(|c_0|, |c_5|, |c_{10}|, |c_{15}|, |c_{20}|)} \,. \tag{5.14}$$

In the Equation 5.14, $c_k$ is the coefficient of SH transform of diffusion signal generated for $k^o$ rotated fibre, and $\hat{c}_k$ is the normalised and reference shifted version. With such notation, all reference coefficients ($k = 10$) become 0, while other coefficients vary from $-1$ (lower than reference)

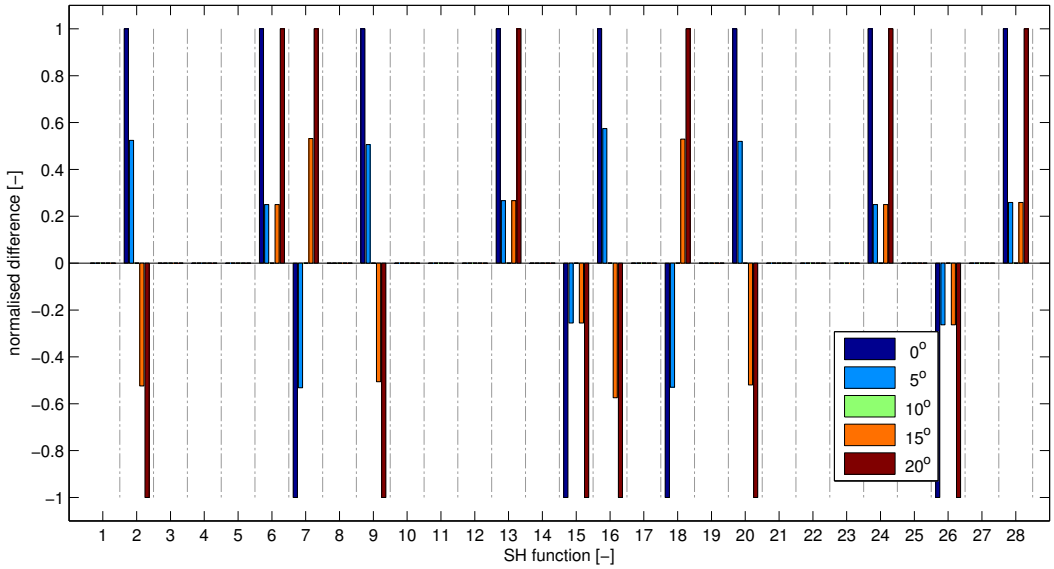



**Figure 5.2:** Normalised difference in SH coefficients of diffusion signal generated for five single fibres belonging to the same plane.

to 1 (bigger than reference).

The predictability of the coefficient change is shown in the Figure 5.2. The normalised difference in SH coefficients magnitude between the reference (green) and rotated fibres changes in a monotonous way (or remain unchanged). For example, 2nd, 9th, 16th, and 20th coefficients decrease, while 7th and 18th increase in magnitude along with the increasing rotation.

It is observed that SH coefficients of signals from fibres within a predefined cone shape are similar. Averaging the SH coefficients will result in a non-zero mean and low variance. The narrower the cone shape becomes the lower the variance. A prior built from these simulations will therefore capture local properties of a diffusion signal coming from a single fibre in a direction defined by the cone. Again, this can be extended to more complex fibre configurations.

Finally, an algorithm to build a localised prior knowledge can be formed. The SH analyses of one million random fibre configurations (Section 5.1.1) can be reused. However, to compute mean and variance for each coefficient only a selection of configurations should be used, *e.g.* since the SH analyses of diffusion signal generated for a single fibre $\vec{r}_x$, $\vec{r}_1$, $\vec{r}_2$, and $\vec{r}_3$ are similar, when analysing an unknown fibre $\vec{r}_x$ the prior for $\mathbf{c}_x$ needs to be strongly influenced by known coefficients $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$. As a result, from the precomputed one million simulations only a subset of most similar SH coefficients are taken, and

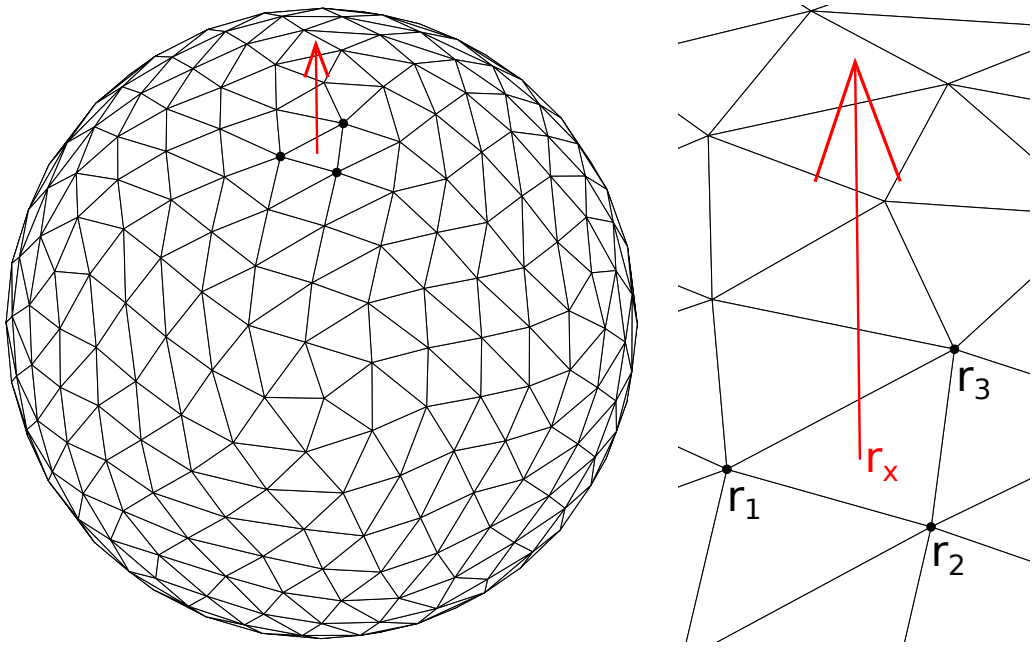



**Figure 5.3:** Spherical sampling scheme. Vertices represent fibre directions for which the SH transform of diffusion signal is known ($\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$ for directions $\vec{r_1}$, $\vec{r_2}$, and $\vec{r_3}$). Red vector $\vec{r_x}$ represents unknown direction for which SH transform of diffusion signal is known as well ($\mathbf{c}_x$). Vector $\vec{r_x}$ is a linear combination of $\vec{r_1}$, $\vec{r_2}$, and $\vec{r_3}$, while SH coefficients $\mathbf{c}_x$ are linear combination of coefficients $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$.

used to compute new mean and variance. The new, localised prior is thus used in the same way as outlined in Section 5.1.

## 5.2 Numerical Evaluation

To numerically test the performance eight fibre configurations were selected: single fibre without (1) and with isotropic compartment (2), two orthogonal fibre crossings without (3) and with isotropic compartment (4), two 45° fibre crossings without (5) and with isotropic compartment (6), and three 90° fibre crossings without (7) and with isotropic compartment (8). Figures 5.4 and 5.5 show the error of HARDI and FODF reconstructions defined as normalised root mean squared deviation (lower means better). Both low diffusion weighting with high SNR (Figure 5.4) and high diffusion weighting with low SNR (Figure 5.5) are presented. Number of iterations was set to 10000.

For all multiple fibre configurations Bayesian regression using both global

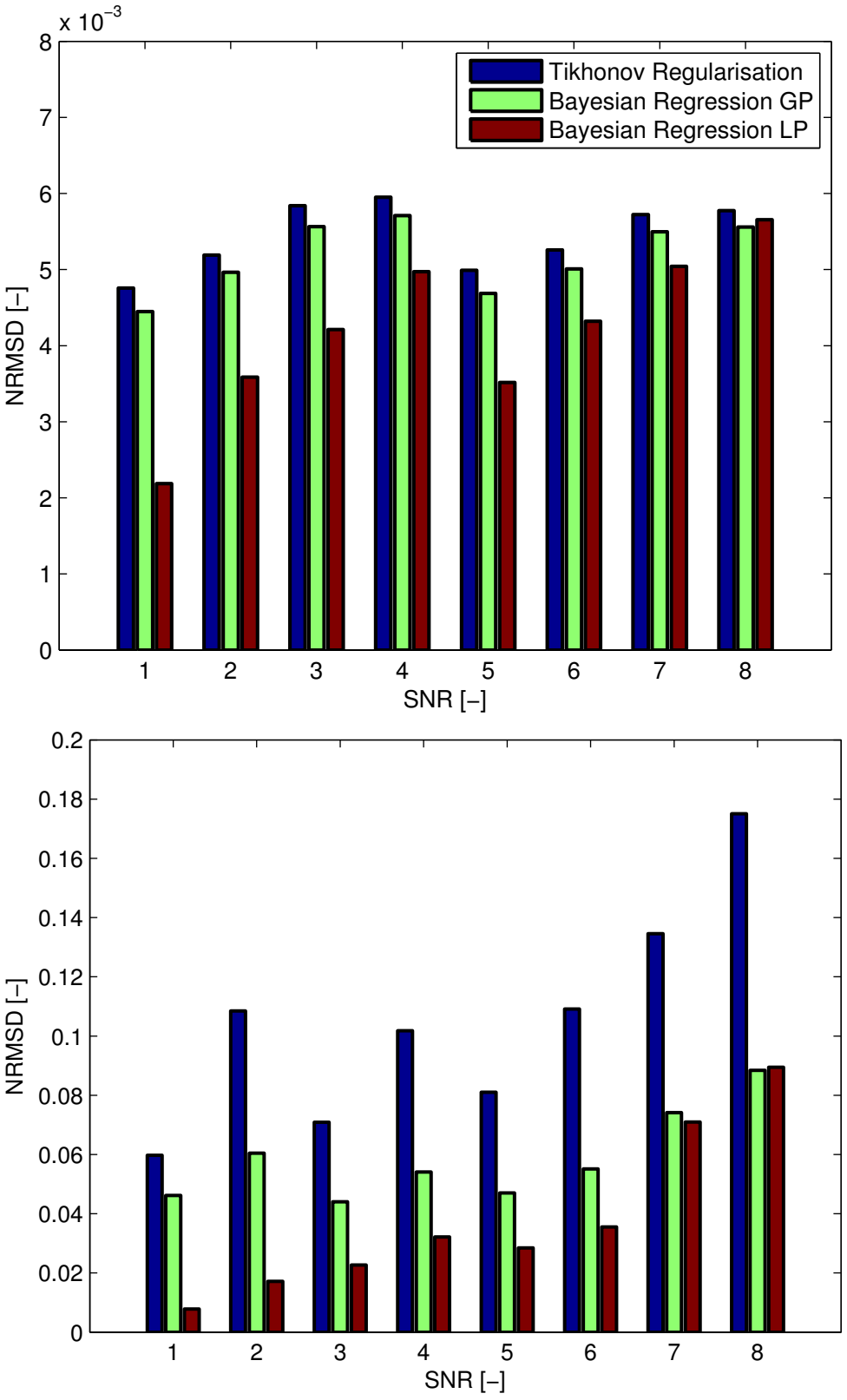



**Figure 5.4:** Reconstruction accuracy of HARDI signal (top) and FODF (bottom) for low diffusion weighting; $b = 1000$ s/mm$^2$, $SNR = 35$.

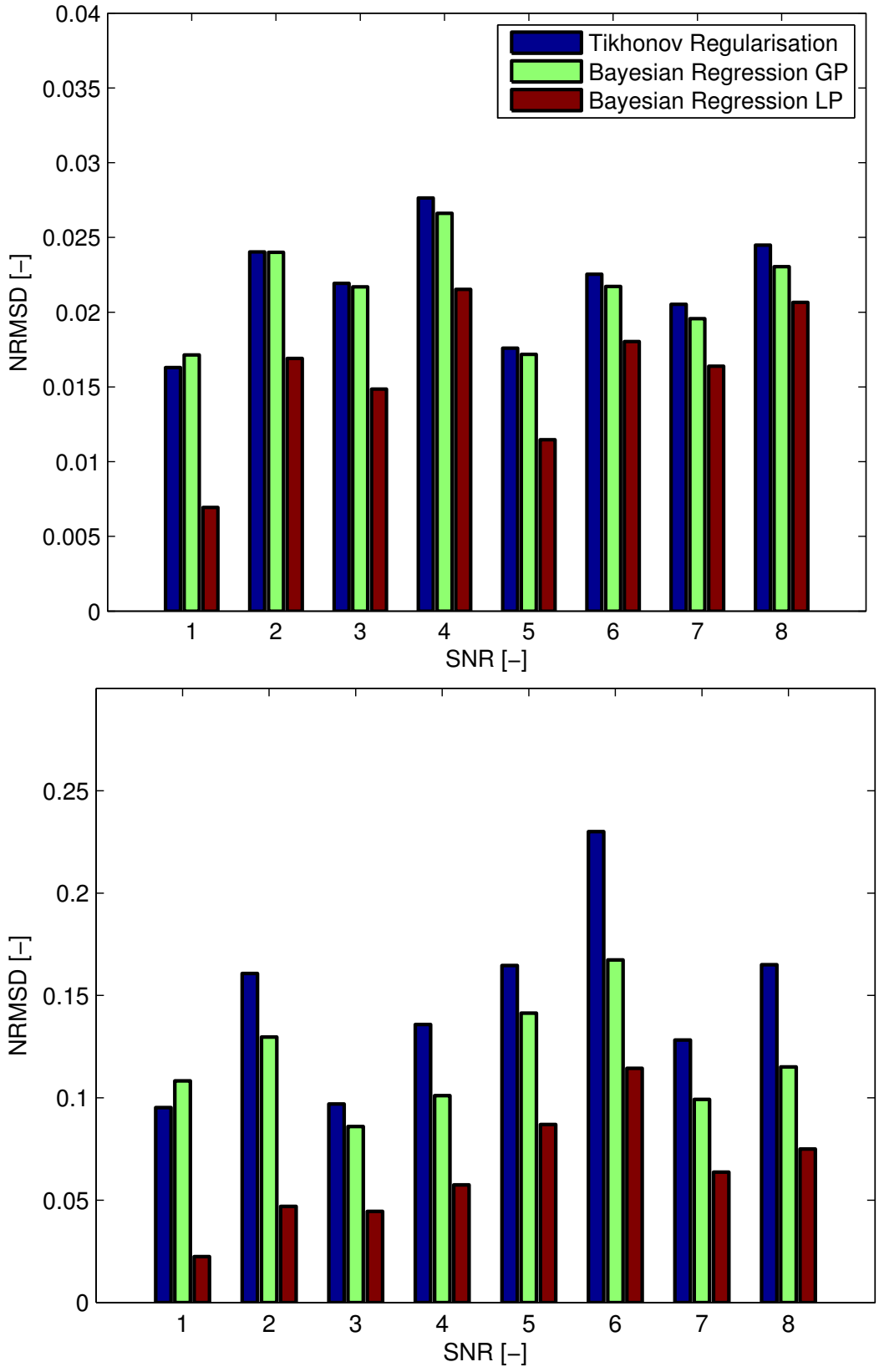



**Figure 5.5:** Reconstruction accuracy of HARDI signal (top) and FODF (bottom) for high diffusion weighting; $b = 3000$ s/mm$^2$, $SNR = 20$.

and localised prior performed better than optimised Tikhonov regularisation (Chapter 3). The improved reconstruction accuracy can be observed for low and high diffusion weighted applied. By utilising prior knowledge, the magnitude of SH coefficients is restrained to physically possible. Refinement of prior knowledge (Section 5.1.2) further increases the SH transform accuracy and provides strong noise invariance. FODF which is more sensitive to noise benefits more from introducing prior knowledge to the SH transform.

With the increased diffusion weighting and the decreased SNR it becomes more preferable to use Tikhonov regularisation to reconstruct the diffusion signal of a single fibre. The diffusion signal can change quickly, and the global prior knowledge, which now includes more high frequencies, can no longer suppress as much noise as regularisation. Ultimately, this is the only case in which regularisation methods can outperform Bayesian as the SNR becomes too low, and prior knowledge too general. In case of more complex structures like two and three fibre crossing, the prior knowledge is flexible enough to aid signal reconstruction, while selectively removing noise. The regularisation on the other hand, indiscriminately saturates higher frequencies removing both signal and noise.

Finally, refining the prior knowledge can improve the accuracy of signal reconstruction. For simple fibre configurations localising prior knowledge results in the most accurate, almost noise free reconstruction. Consequently, FODFs consist of only true peaks, significantly reducing the error (configurations 1 and 2). As the configuration becomes more complex though, the benefits of prior knowledge localisation diminish. At low diffusion weighting factors and three fibre crossings it is no longer desired to refine the prior knowledge.

Figures 5.6 and 5.7 visualise reconstructed HARDI signal (top sub-figure) and FODF (bottom sub-figure) from simulations with low (Figure 5.6) and high (Figure 5.7) diffusion weighting factor. In all sub-figures, five reconstructions are visualised (from bottom): noise free, least squares, optimised Tikhonov regularisation, global Bayesian regression, and localised Bayesian regression.

For both low and high diffusion weighting factors, Bayesian regression methods reconstruct more regular and noise-free signal than optimised Tikhonov regularisation. The near-perfect reconstruction can be seen in case of FODF, especially when using a localised prior. With localised prior

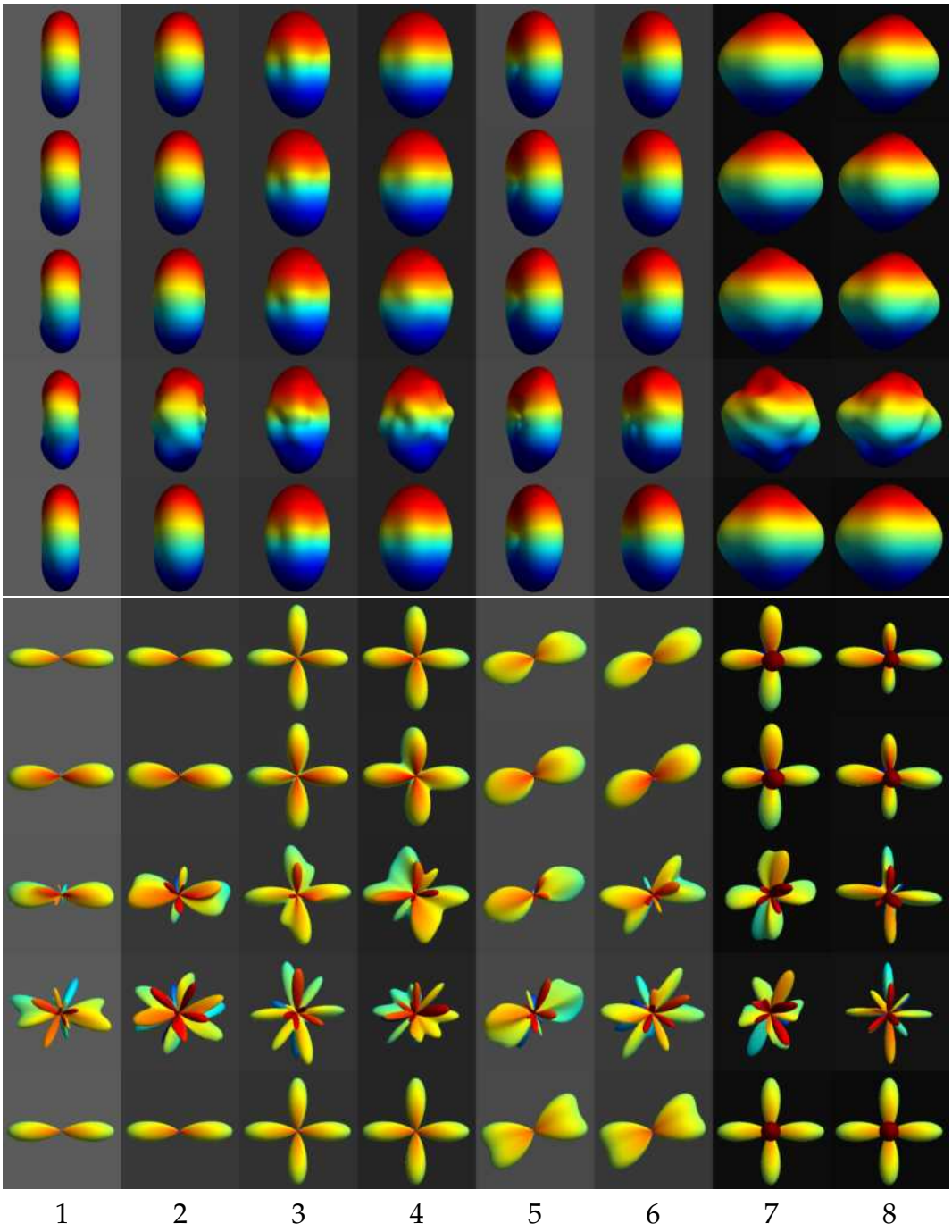



**Figure 5.6:** Visualisation of HARDI signal (top image) and FODF (bottom image) for low diffusion weighting. From bottom to top row: real signal, least squares reconstruction, regularised reconstruction, global and local Bayesian reconstruction; $b = 1000\ \mathrm{s/mm^2}$, $SNR = 35$.

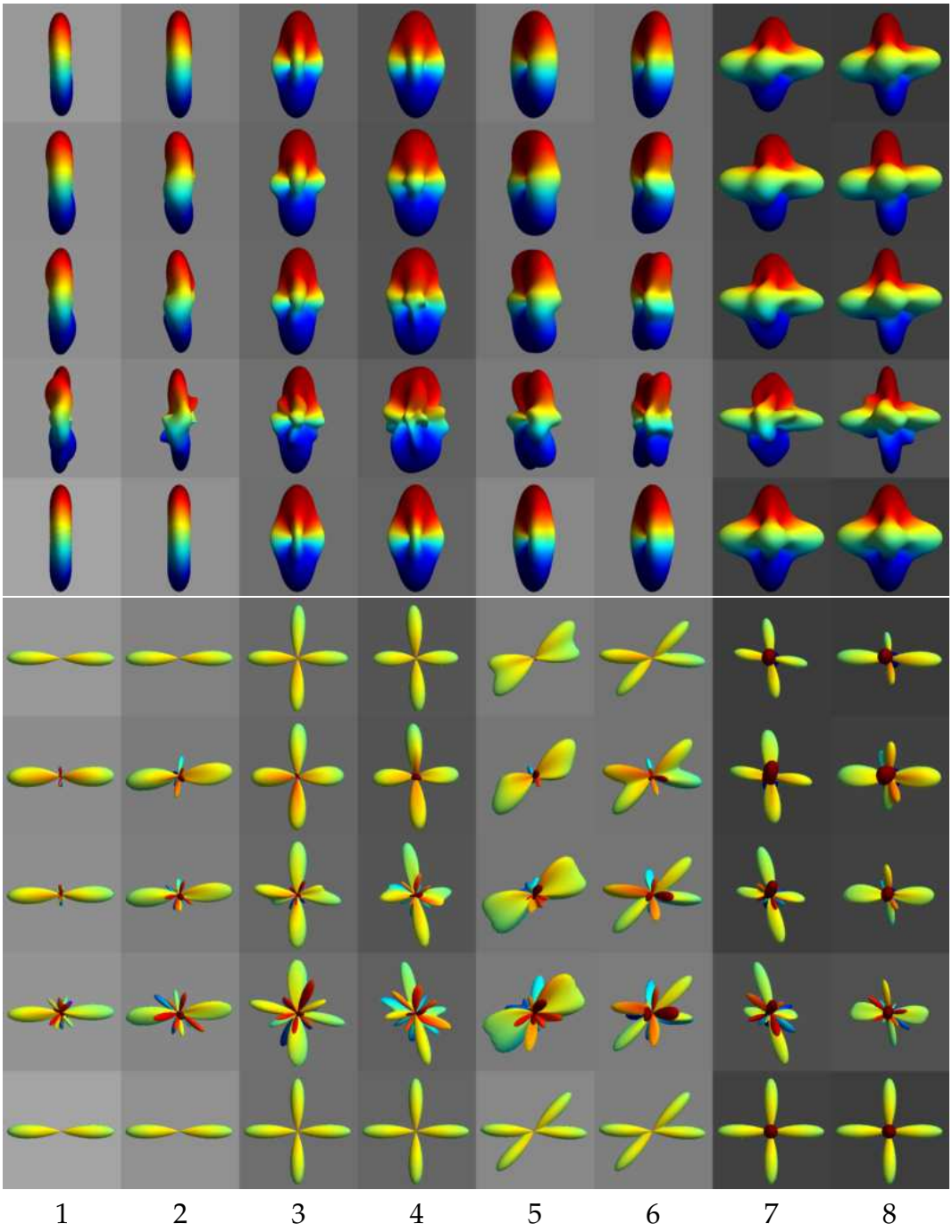



**Figure 5.7:** Visualisation of HARDI signal (top image) and FODF (bottom image) for high diffusion weighting. From bottom to top row: real signal, least squares reconstruction, regularised reconstruction, global and local Bayesian reconstruction; $b = 3000\ \mathrm{s/mm^2}$, $SNR = 20$.

there are no false, noise imbued peaks in the reconstructed signal. This high noise robustness is obtained at the cost of a small orientation tilt (Figure 5.7, top rows of both sub-figures).

## 5.3 Brain Image Application

The same brain scan from Section 4.4 was used to test the application of Bayesian regression model in diffusion signal reconstruction. In accordance with Chapter 4, signal reconstruction was done using 8th order SH series as well. This will allow for a direct comparison with the results presented in Section 4.4.

Figure 5.8 shows GFA maps computed from three SH series – regularised series, as well as Bayesian series using two priors (global and localised). Bayesian regression using either prior, similarly to backward elimination, provides a way of removing high frequency noise without lowering angular resolution of diffusion signal. GFAs produced from Bayesian obtained SH series contain higher contrast and as a result look "sharper" than regularised GFA. The GFA computed from a Bayesian regression using a localised prior has higher contrast than the GFA computed using a global. Additionally, the GFA computed using a localised prior is sharp enough to observe the separation between two hemispheres.

Similar to Chapter 4.4, TrackVis was used to perform a full brain tractography. All tracts passing through 27th slice are shown in Figures 5.9 and 5.10. Bayesian regression with global prior reconstruct accurate diffusion signal, partially invariant to noise. Both fibre lengths and number of fibres resolved are similar to Tikhonov regularisation (average fibre length of 11.3454 mm and 11.4688 mm, and 157609 and 156282 fibres resolved). Fibre tracks resolved from the signal reconstructed using a localised prior though, are not always correct. Despite superior signal reconstruction introduced by the use of localised prior, fibre tracks resolved (Figure 5.10) were not on par with fibres resolved when global prior (Figure 5.9) was used. Among correctly resolved fibre tracks there is a large number of artificial tracts (grey tracts), both inside and on the surface of brain.

The localised prior always provides the best candidate for a prior, even for a voxel that does not contain any potential fibres. In the current implementation

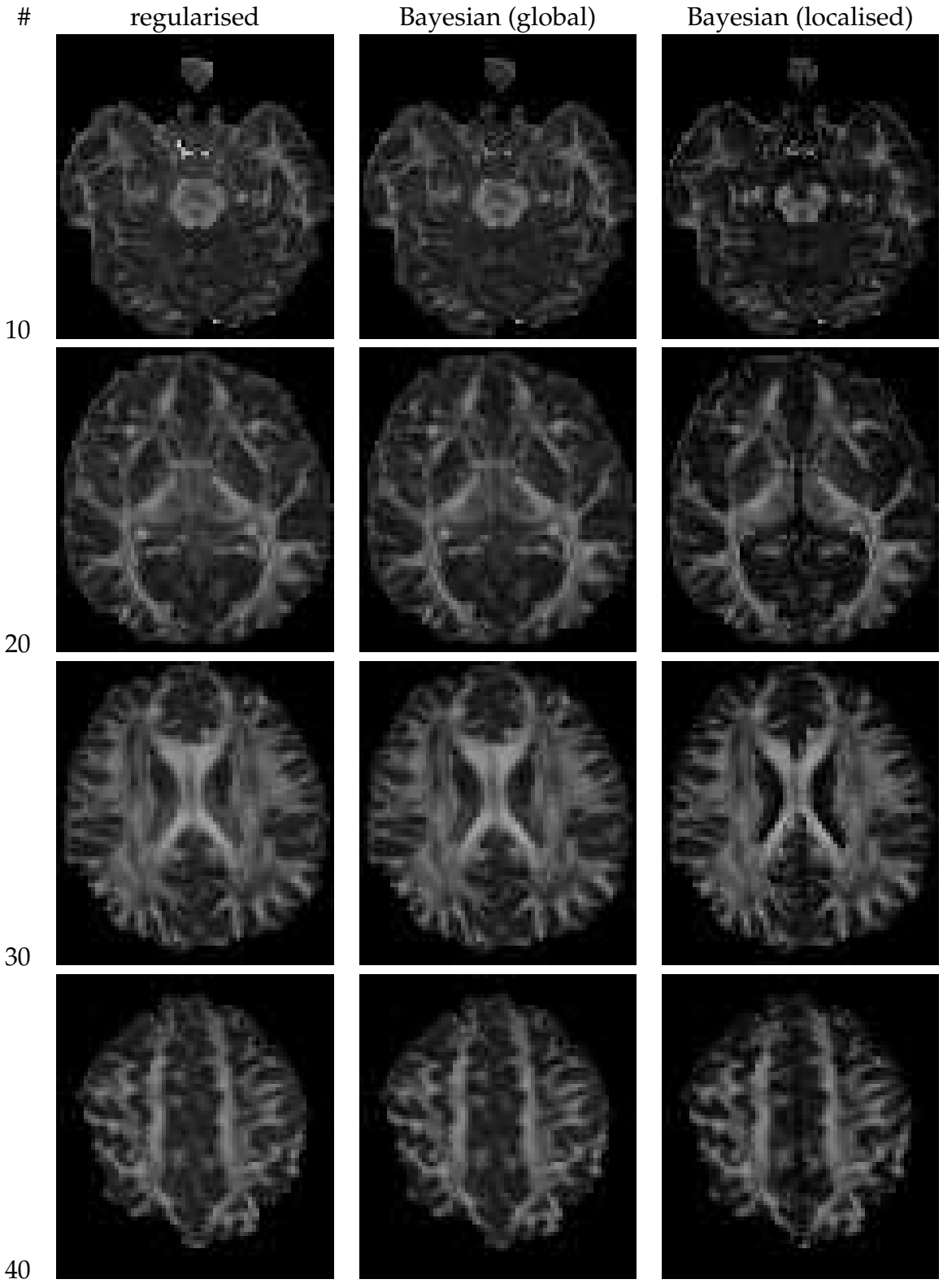



**Figure 5.8:** GFA maps computed from regularised (left) and Bayesian regressions (middle and right). Bayesian regression using a global (middle) and localised (right) prior. Inferior-superior view, 10th, 20th, 30th and 40th slice.

the list of candidates is fixed (10000 most similar SH series). In case of isotropic data, the list will contain SH series of isotropic data, as well as most similar anisotropic SH series. Due to low magnitudes of isotropic based candidates, such localised prior will have a tendency to "fiberise" the isotropic voxel. As the method is stable, the localised prior for an isotropic voxel is always computed in the same way. This results in an isotropic voxel being "fiberised" in a systematic way. The tracking algorithm, fibre assignment by continuous tracking [43, 50], finds a group of same voxels and assigns fibres through them ($90^o$ fibre crossing, seen as grey fibres in Figure 5.10). In a non-localised reconstruction (*e.g.* Tikhonov regularisation or when using a global prior) these voxels are still reconstructed as either isotropic, or randomly oriented anisotropic. As a result, tractography will discard them as possible fibre holders due to their non-uniformity. This unexpected consequence of prior localisation could be corrected either in the reconstruction step, or during the tractography itself.

## 5.4 Discussion

In this Chapter, Bayesian regression model and its application to diffusion signal reconstruction was introduced. This regression model supplements the measured data with some beliefs about it. The stronger and more accurate the beliefs are, the better theoretical results it can deliver. Those beliefs come in a form of a prior knowledge, and the main challenge in creating a Bayesian regression model is in providing a proper prior.

In Section 5.1.1, a Monte–Carlo method for creating a physically realistic prior knowledge was proposed. Diffusion signal for a set of random fibre configurations was created and represented using SH series. The mean and variance of the SH coefficients are calculated as a 'global' prior, as it was computed from SH coefficients of diffusion signal from all possible fibre configurations. Additionally, only a subset of fibre configurations was used, and mean and variance were computed from the respective SH coefficients. The subset was selected in a way that minimised the difference between the Monte–Carlo obtained SH coefficients and the least squares obtained SH coefficients of a unknown voxel. The prior obtained in this way was called "localised" as it adapted to the observed data.

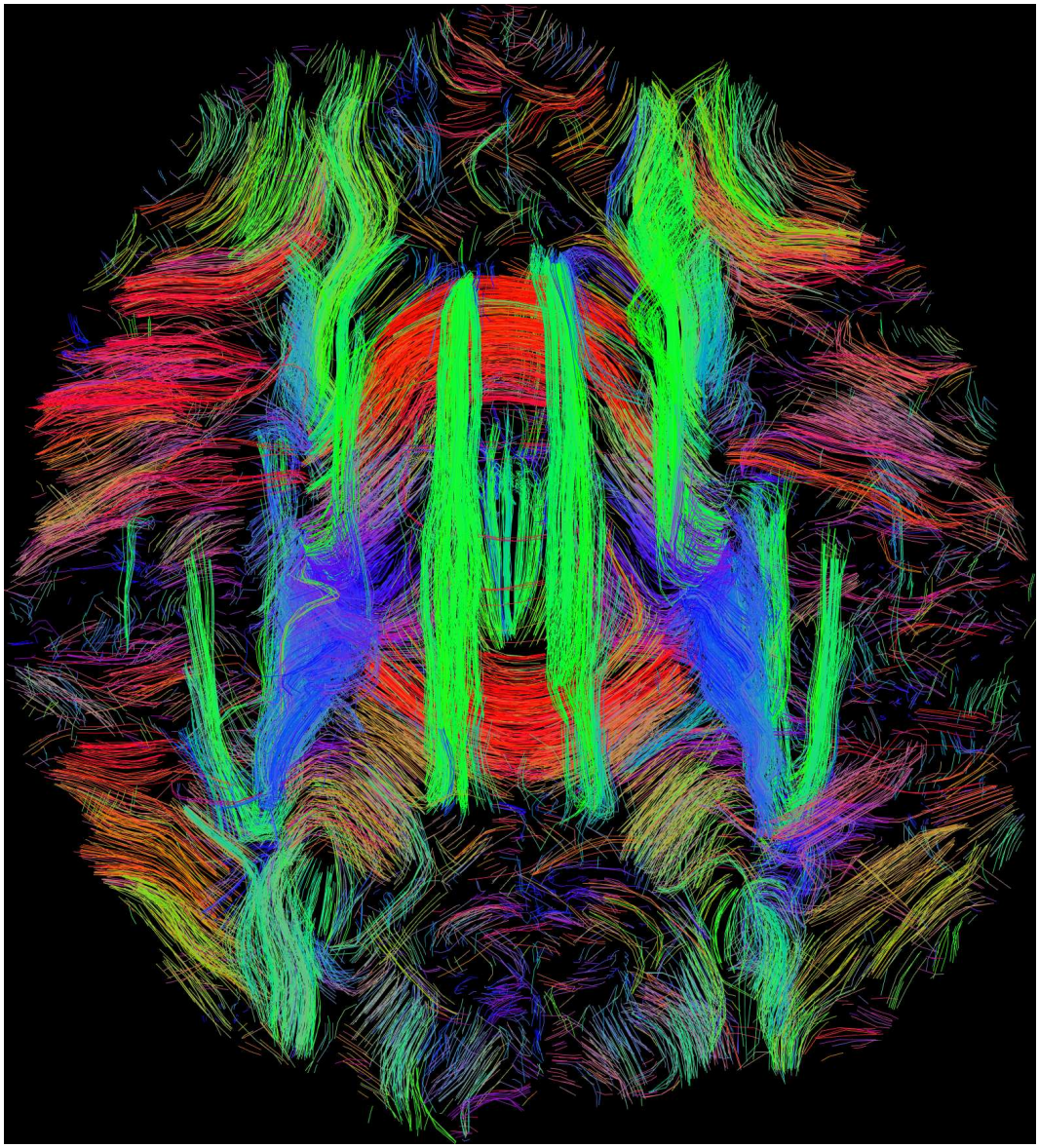


**Figure 5.9:** Full brain tractography, inferior-superior view of 27th slice, Bayesian reconstruction of diffusion signal using a global prior.

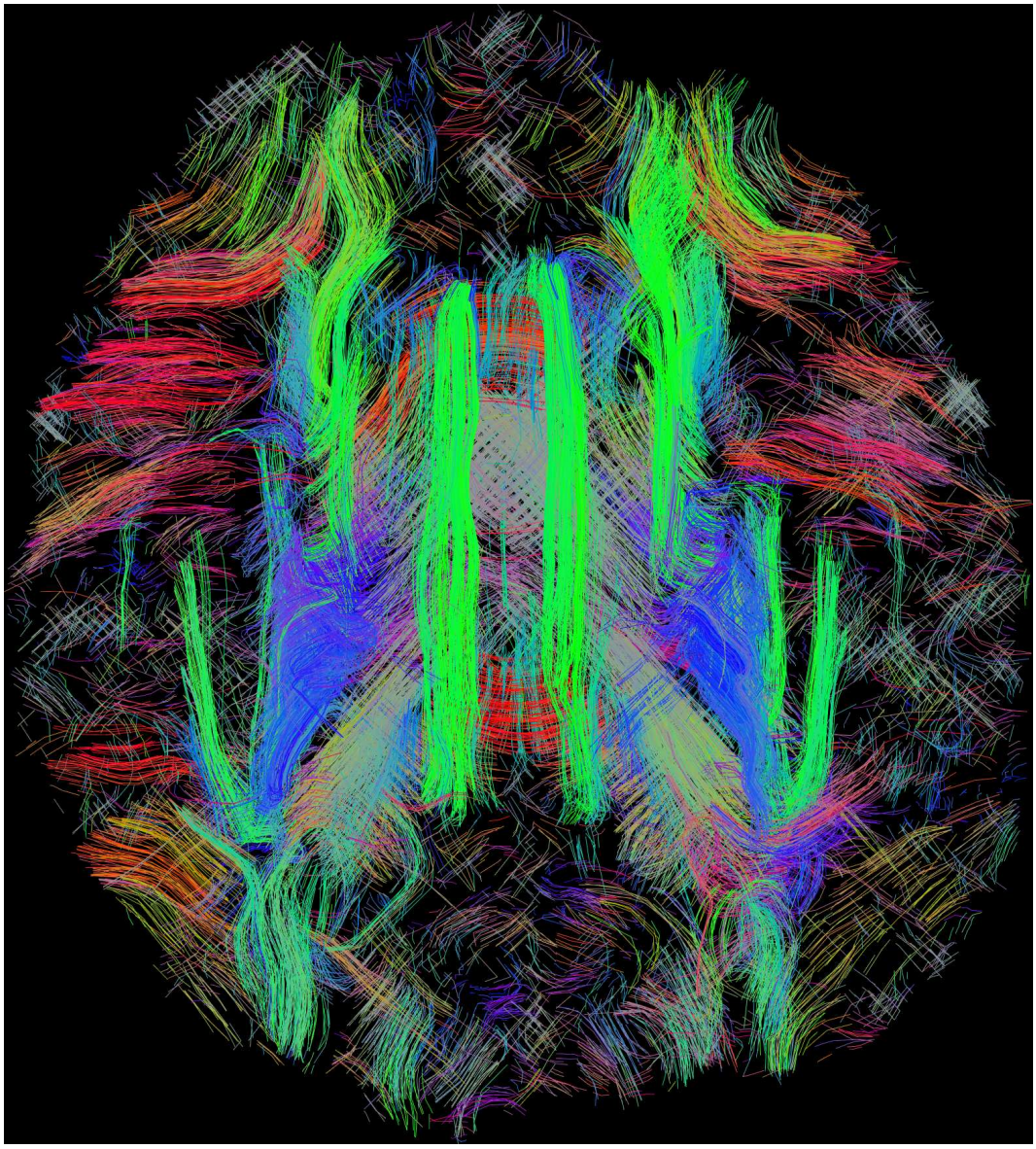


**Figure 5.10:** Full brain tractography, inferior-superior view of 27th slice, Bayesian reconstruction of diffusion signal using a localised prior.

Numerical superiority of the Bayesian regression was shown in Section 5.2. For both low and high diffusion weightings applied, the application of Bayesian regression improved signal reconstruction accuracy. The only case in which Bayesian regression performed worse than the current state-of-the-art (regularisation schemes) was with high diffusion weighting factor and simple fibre configuration (single fibre) when using a global prior. Global prior for high diffusion weighting applied has a high variance, as it needs to accommodate for high changes in the magnitude of higher order SH coefficients. As a result, it could not successfully remove as much noise as a classical regularisation scheme. This was corrected by the localisation of the prior – variance was lowered, and non-zeroed mean was introduced.

In Section 5.3 both priors were tested on an *in vivo* human brain image. Tractography using the signal reconstructed through Bayesian regression with both priors was performed and visually inspected. The signal reconstructed using a global prior can be used as is by the off-the-shelf software (*e.g.* TrackVis). The fibre tracts resolved are visually similar to tracts resolved from the signal reconstructed using current state-of-the-art. However, the localised prior had the tendency to overrule the observed data, especially at isotropic voxels, resulting in anisotropy being given to isotropic voxels. This confuses the tractography algorithm to construct fibres passing through those voxels.

This can be traced back to the way the localised prior is computed. A fixed number of similar SH coefficients is used to compute the new mean and variance. The subset of SH coefficients contains not only coefficients for isotropic but also anisotropic data. This skews the prior towards anisotropic prior, and ultimately introduces anisotropic bias in the reconstruction of isotropic signal. The most straightforward solution would be to enforce a fixed number of purely isotropic voxels in prior generation. However, this does not guarantee that any isotropic signal will contribute to the prior, *e.g.* that anisotropic SH coefficients will not be involved in computing the mean and variance. Alternatively, as the anisotropy in those voxels is relatively low, any fibres passing through them could thus be rejected.

The main novelty of this research is the application of Bayesian regression model to diffusion signal reconstruction and introduction of methods to compute the prior knowledge. There are many possibilities to generate and use prior knowledge in diffusion signal reconstruction which were not

explored in this research. For example, different methods for building a localised prior should be investigated. While the "most similar" approach works in simulated environment, the tractography algorithms and real brain reconstruction proved to be challenging. Alternative methods could include, among others, prior weighting or ensuring a correct distribution in the global prior (*e.g.* enough voxels with isotropic signal).

Finally, the Bayesian regression opens a path to many new applications. For example, the inclusion of prior knowledge in regression model allows to work with insufficient data. The higher order SH series can be used, with more regressors than regressands, possibly to obtain sharper and more accurate FODFs.

# 6

# Conclusions and Future Directions

## 6.1 Summary and Conclusions

This thesis is concerned with improving the quality of the diffusion weighted signal obtained from an MRI scanner. The random motion of water molecules is used to introduce contrast to DW images. In organised structures, the motion becomes restricted in a systematic way. As a consequence, by using the measured diffusion signal it is possible to infer information about the underlying structure.

In medicine, diffusion MRI is most commonly used to produce images of human brain. The most organised structure within the brain is the white matter. Bundles of axons that run in the same direction form fibre tracts, and restrict water motion in perpendicular direction. The measured signal can be used to establish directions in which the signal least hindered, and fibre tracts, which usually span across many voxels, can be reconstructed using information derived from the measured signal.

The resolution and quality of acquired signal is restricted by the scanning time. Having an image with high spatial and angular resolution, as well as high SNR, imposes clinically implausible scanning times. As a result, the spatial and angular resolution of the image is limited: voxels are usually 1 or 2 mm wide, and up to 60 sampling directions represented by diffusion weighting gradients are used. The parameters, along with the diffusion weighting factor, are selected in such a way that an image of acceptable quality (SNR wise), in an reasonable time frame (usually up to 15 minutes) can be obtained.

There are few consequences stemming from limited spatial and angular resolutions. First, restricting the spatial resolution to 1 mm isotropic voxels means that thousands of axons, which are 10 $\mu$m thick, are contributing to that voxels signal. Not all of those axons will pass in the same directions, *e.g.* some will cross, fan, or touch (kiss). This means that the obtained diffusion profile can potentially have a non-trivial shape. Moreover, same fibre bundles may span across few neighbouring voxels. As a consequence of this non-trivial shape and limited angular resolution, a lot of significant data can be missed. Without enough sampling directions, a voxel in which three fibre bundles cross can be mistaken for a simpler configuration (*e.g.* two fibre bundles crosses, or even a single fibre).

Considering all these challenges, a successful signal analysis or fibre reconstruction algorithm needs to be able to retain from a noisy scan as much original signal as possible. As a result, signal denoising algorithms, regularisation schemes, or constrained models were proposed in the past. This research, investigates the applicability of spherical harmonic basis functions to represent as well as denoise diffusion signal.

Application wise, the focus of this thesis is on improving the quality of *in-vivo* images of the human brain. The algorithms presented here though can not only be applied to images of different regions of the body, but also extended to non-medical applications.

## 6.2 Notable Achievements

A systematic approach to optimally select a regularisation parameter was presented in Chapter 3. Using Monte–Carlo simulations, diffusion signal of differ-

ent parameters is generated, corrupted, and represented with SH series. The regularisation strength that on average provides the most accurate signal representation is then selected. The recommended regularisation parameters were found to significantly differ from currently utilised *L*-curve analysis [18], as well as to improve the accuracy of SH signal representation.

Additionally, a novel algorithm to fit SH basis function to HARDI signal was described in Chapter 4. The method, based on backward elimination regression, uses statistical testing to discriminate between accurate and corrupted SH basis functions. Combined with Tikhonov regularisation [19], the method further improves diffusion signal reconstruction accuracy.

Finally, an application of Bayesian method for solving linear regression problems was introduced in Chapter 5. The linear regression problem is usually solved using ordinary least squares method. However, instead of adding the regularisation term prior information is used. The prior is constructed through Monte–Carlo simulations. A mathematical model was used to simulate a diffusion signal, which was then analysed using SH transform. Two methods of using a prior knowledge were tested, and compared using numerical phantoms. Finally, the method was used to process the diffusion image of a human brain.

## 6.3 Known Limitations

As was mentioned in Section 2.4.1, with a sufficient SNR the Rician and Gaussian distributions start to match, which allows to solve linear regression problems with ordinary least squares method. In practice though, the measured signal is not always of high enough quality to use the OLS solution. When SNR is too low, the correct solution should be found using maximum likelihood method [28]. As a consequence, the SH transform used in Chapters 3, 4, and 5 should be reimplemented using maximum likelihood method.

As a part of research described in Chapter 3, 4, and 5 a limited clinical study was performed. In each chapter, one diffusion MRI image was analysed – either partially or in full. These experiments served as a proof of concept and demonstrated that the proposed methods can potentially become clinically feasible. However, due to the limited and incomplete scope of this proof further extensive clinical study is needed.

## 6.4 Future Directions

While the SH allow for a convenient computation of ODF and FODF, unlike parametric models they cannot distinguish the signal from noise. With a higher diffusion weighting applied more basis functions are needed to represent the fast changing signal. The higher order basis functions represent higher frequencies, and are naturally more susceptible to corruption. Adding mathematical (*e.g.* non-negativity) and physical (*e.g.* Gaussian-like behaviour) constraints should further reduce the noise influence, and significantly improve the accuracy of SH transform.

With a higher diffusion weighting applied, the multi tensors model becomes inaccurate. Recent research, especially involving high field MRI scanners, has shown two components of diffusion signal: intra and extra cellular. Water diffusion becomes restricted inside axons, while outside it is merely hindered. This observation led to development of more advanced diffusion models, like diffusion Kurtois, or biexponential. The foundation of this research is based on SH basis functions, which are globally supported, and therefore can be used to represent any spherical signal. As a consequence, any change in the mathematical model used to generate the diffusion signal should have a marginal effect. Nevertheless, further investigation at high diffusion weighting factors should be made, in order to validate the findings described in this thesis.

When using the multi tensor model, all scanner parameters are conveniently combined into a diffusion weighting coefficient $b$. With more complex models though, it is no longer possible to rely just on $b$, as the same value can be obtained from different scanning parameters. Most often, the complex model will separate extra and intra cellurar diffusion, and model each one as a function of a different, scanning related parameter (*e.g.* echo time $\Delta$). As a result, all functions of $b$ (like optimal regularisation value) should be recalculated as a function of those additional parameters.

Finally, super resolution technique could be added to presented research. By simply performing sub-voxel interpolation of the measured signal the accuracy of tractography algorithms has been improved. Adding this technique to Tikhonov regularisation or backward elimination regression should improve the signal reconstruction accuracy, and as a result, the

accuracy of ODF and FODF as well. Since the research presented here is focused on the SH transform, it is desired that the sub-voxel interpolating algorithm is implemented in SH domain.

# Appendices

# A

# Diffusion Signal Simulation

The Monte–Carlo simulations described in this thesis, unless otherwise specified, were all based on a phantom signal created using Equation 2.6. The MTM implementation in Matlab language is shown in Listing A.1. Function *create_signal* returns a vector with diffusion signal for a single voxel, and takes 5 mixed arguments. The 5 arguments are as follows: diffusion weighting factor $b$ (scalar), T2 value at voxel $E0$ (scalar), diffusion weighting gradients $g$ ($n$ by 3 matrix), $m$ diffusion coefficients (3 by 3 by $m$ matrix), and volume ratios $vol$ (vector).

```
1  function E = create_signal(b, E0, g, D, vol)
2     n = size(g, 1);
3     m = size(D, 3);
4     E = zeros(n, 1);
5     % generate signal, assuming size(vol, 2) == m
6     for in = 1:n
7        gi = g(in,:)';
8        for im = 1:m
9           Di = D(:,:,im);
```

```
10          E(in) = E(in) + vol(im) * ...
11              exp(-b * (gi' * Di * gi));
12        end
13        E = E0 * E;
14    end
15    % normalise, as it is possible for sum(vol) != 1
16    E = E / sum(vol);
17 end
```

**Listing A.1:** Diffusion signal simulation at a voxel level.

The generated signal $E$ can further be corrupted, by adding a Rician distributed noise with standard deviation $st$:

```
1 function Es = corrupt_signal(E, st)
2    Es = sqrt((E + st*randn(n, 1)).^2 + (st*randn(n, 1)).^2);
3 end
```

**Listing A.2:** Signal corruption using Rician distributed noise.

# B

# Ground Truth Verification

The GTV method can be explained using two functions: main function *optimal_parameters* (Listing B.1) in which the optimal parameter is selected based on a pre-defined criteria, and sub function *investigate_lambda* (Listing B.2) where regularisation parameters are evaluated. Both functions *optimal_parameters* and *investigate_lambda* expect the first argument to be a structure (*simParams*). In addition, function *investigate_lambda* requires a diffusion weighting factor (*bval*), and standard deviation of noise (*st*).

The structure *simParams*, represents simulation parameters and is constant. The structure defines the range of SNRs (*simParams.snrs*), diffusion weighting factors (*simParams.bvals*), and regularisation parameters (*simParams.lambdas*) to be tested, number of iteration (*simParams.iterations*), SH order (*simParams.order*)) and design matrix (*simParams.aMatrix*) to be used, diffusion sensitising gradients (*simParams.gradients*) and the fibre configuration (*simParams.dNum* and *simParams.dIso*) for which the diffusion signal is simulated.

For each SNR and diffusion weighting factor, *simParams.iterations* voxels are generated. For each of those voxels, the optimal regularisation parameter using both L-Curve (function *lcurve_parameters*) and GTV method (function *investigate_lambda*) is found. For GTV method, in this implementation the regularisation parameter that minimises the minimal error is selected (Listing B.1, lines 18-23).

```
1  function opt = optimal_parameters(simParams)
2     a = size(simParams.snrs, 2);
3     b = size(simParams.bvals, 2);
4     opt = zeros(a, b, 5);
5     i = 1;
6     for snr = simParams.snrs
7        j = 1;
8        for bval = simParams.bvals
9           % compute l-curve parameters
10          r0 = lcurve_parameters(simParams, bval, 1/snr);
11          % investigate lambdas using GTV
12          [r1, r2, r3, r4] = investigate_lambda(simParams,
               bval, 1/snr);
13          % minimise minimal error
14          tmp1 = zeros(simParams.iterations, 1);
15          tmp2 = zeros(simParams.iterations, 1);
16          tmp3 = zeros(simParams.iterations, 1);
17          tmp4 = zeros(simParams.iterations, 1);
18          for k = 1:simParams.iterations
19             tmp1(k) = simParams.lambdas(r1(k,:) ==
                  min(r1(k,:)));
20             tmp2(k) = simParams.lambdas(r2(k,:) ==
                  min(r2(k,:)));
21             tmp3(k) = simParams.lambdas(r3(k,:) ==
                  min(r3(k,:)));
22             tmp4(k) = simParams.lambdas(r4(k,:) ==
                  min(r4(k,:)));
23          end
24          % store data
25          opt(i,j,1) = mean(r0);
26          opt(i,j,2) = mean(tmp1);
27          opt(i,j,3) = mean(tmp2);
28          opt(i,j,4) = mean(tmp3);
29          opt(i,j,5) = mean(tmp4);
30          j=j+1;
31       end
32       i=i+1;
33    end
```

```
34 end
```

**Listing B.1:** Optimal parameter selection.

The second function returns 4 error measures for each of the regularisation parameter tested. Functions *r_correlation* and *skl_divergence* are used to calculate correlation coefficient and symmetric Leibler–Kullback divergence, and implemented according to their respective publications. Functions for generating and corrupting diffusion signal are described in Appendix A.

```
1  function [res1, res2, res3, res4] =
       investigate_lambda(simParams, bval, st)
2      X = simParams.aMatrix;
3      L2 = l_tikhonov(simParams.order, 2);
4      n = size(simParams.gVectors, 1);
5      ol = odf_legendre(simParams.order);
6      lsize = size(simParams.lambdas, 2);
7      % preallocate output
8      res1 = zeros(simParams.iterations, lsize);
9      res2 = zeros(simParams.iterations, lsize);
10     res3 = zeros(simParams.iterations, lsize);
11     res4 = zeros(simParams.iterations, lsize);
12     % OPTIMISATION: precompute SHT
13     inversePrecomputed0 = (X' * X) \ X';
14     inversePrecomputed = cell(lsize, 1);
15     for j=1:lsize
16        lambda = simParams.lambdas(j);
17        inversePrecomputed{j} = (X' * X + lambda * L2) \ X';
18     end
19     % do the simulation
20     for i=1:simParams.iterations
21        D = diffusion_random(simParams.dNum, simParams.dIso);
22        yp = create_signal(bval, 1, simParams.gVectors, D);
23        y = corrupt_signal(yp, st);
24        % perfect reconstruction
25        cp = inversePrecomputed0 * yp;
26        ocp = ol .* cp;
27        % test all lambda
28        for j=1:lsize
29           c_lb = inversePrecomputed{j} * y;
30           oc_lb = ol .* c_lb;
31           % and save
32           res1(i,j) = norm(cp - c_lb);
33           res2(i,j) = norm(ocp - oc_lb);
34           res3(i,j) = 1-r_correlation(cp, c_lb);
```

```
35          res4(i,j) = skl_divergence(inversePrecomputed0,
               inversePrecomputed{j}, ocp, oc_lb, X*ocp, X*oc_lb);
36        end
37     end
38  end
```

**Listing B.2:** Calculation of fitness functions.

# C

# Backward Elimination Method

The actual implementation of the $t$-test introduced in Chapter 4, Section 4.2 does not use $p$-values. Instead, the inverse cumulative distribution functions of the $t$-distribution is used to find $t$-values at the $1 - \alpha/2$ [1] percentile for a range of degrees of freedom (from $N - L$ to $N$). The minimal $t$-score is compared with the threshold, and if lower, the null hypothesis is accepted, and the basis function represented by the coefficient $c_i$ is considered insignificant and removed from the signal representation. These threshold values ($icdf\_arr$) need to be computed only once for the whole image, and allows for a noticeable reduction in the computation time.

The function to perform backward elimination (Listing C.1) expects 3 arguments to be passed, and returns 2 parameters. The arguments are the diffusion signal $y$ (vector), SH design matrix $X$ (matrix), and inverse cumulative distribution function for $t$-distribution $icdf\_arr$ (vector) evaluated at $1 - \alpha/2$ percentile. The returned parameters are new SH coefficients

[1] One-sided test due to taking the absolute value of the coefficient $c_i$.

where eliminated frequencies are zeroed *newcoeffs* (vector), and a binary list of basis functions used *cc* (vector).

```
function [newcoeffs, cc] = backward_elimination(y, Xorg,
    icdf_arr)
    N = size(Xorg, 1);
    P = size(Xorg, 2);
    Morg = inv(Xorg' * Xorg);
    fullcoeffs = Xorg \ y;
    coeffs = fullcoeffs;
    E_eval = Xorg * coeffs;
    nc = 1:P;
    M = Morg;
    X = Xorg;
    % start the main loop
    while true
        L = size(coeffs, 1);
        res = sum( (y - E_eval).^2 ) / (N-L);
        signif = abs(coeffs ./ sqrt(res * M(logical(eye(L)))));
        % OPTIMISATION: use precomputed inverse cdf 'icdf_arr'
        msignif = min(signif);
        if msignif > icdf_arr(N-L) || L == 1 || isnan(msignif)
            == 1
            % stopping criteria reached, or just one SH basis
                function left
            break;
        else
            m = find(signif == msignif);
            if size(m, 1) == 0
                error('Something went terribly wrong!');
            end
            % remove the least significant and start again
            nc(m) = [];
            X(:,m) = [];
            M(m,:) = [];
            M(:,m) = [];
            coeffs(m) = [];
            E_eval = X * coeffs;
        end
    end
    % convert 'nc' (vector with index of basis functions used)
        to 'cc' (binary mask of basis functions used)
    cc = zeros(1, P);
    for c = nc
        cc(c) = 1;
    end
```

```
40      % create reduced SHT, the difference between 'coeffs' and
           'newcoeffs' is that the latter contains '0'
41      newcoeffs = fullcoeffs .* cc';
42  end
```

**Listing C.1:** Backward Elimination Method.

# D

# Bayesian Regression

Computation of posterior distribution using a measured data and prior knowledge is shown in Listing D.1. Function *bayesian_regression* expects five arguments: diffusion signal $E$ (vector), SH design matrix $X$ (matrix), estimated noise variance $sig2$ (scalar), and prior means $mpr$ (vector) and variances $Vpr$ (diagonal matrix). Function returns posterior means $mpo$ (vector) and variances $Vpo$ (diagonal matrix). When using a global prior, this step can be optimised by pre-computing matrices $Vpo$, as well as term $sig2 * inv(Vpr) * mpr$. However, when using a localised prior this optimisation is no longer possible.

```
1 function [mpo, Vpo] = bayesian_regression(E, X, sig2, mpr, Vpr)
2    Vpo = inv(sig2 * inv(Vpr) + X' * X);
3    mpo = Vpo * (sig2 * inv(Vpr) * mpr + X' * E);
4 end
```

**Listing D.1:** Computation of posterior distribution.

# Bibliography


[1] A. A. Afifi and V. Clark. *Computer-Aided Multivariate Analysis*. Chapman & Hall, Ltd., 1999.

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

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

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

[5] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44:625–632, 2000.

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

[7] P. J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.

[8] P. G. Batchelor, D. L. Hill, D. Atkinson, and F. Calamante. Study of connectivity in the brain using the full diffusion tensor from MRI. In *Information Processing in Medical Imaging*, pages 121–133. Springer, 2001.

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

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

[11] F. Bloch, W. W. Hansen, and M. Packard. Nuclear induction. *Physical review*, 70:460–474, 1946.

[12] M. C. Chiang, A. Klunder, K. McMahon, G. de Zubicaray, M. Wright, A. Toga, and P. Thompson. Information-theoretic analysis of brain white matter fiber orientation distribution functions. In *Information Processing in Medical Imaging*, 2007.

[13] J. Cohen-Adad, M. Descoteaux, S. Rossignol, R. Hoge, R. Deriche, and H. Benali. Detection of multiple pathways in the spinal cord using q-ball imaging. *NeuroImage*, 42:739–749, 2008.

[14] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences*, 96(18):10422–10427, 1999.

[15] M. E. Daube-Witherspoon and G. Muehllehner. An iterative image space reconstruction algorthm suitable for volume ECT. *IEEE Transactions on Medical Imaging*, 5:61–66, 1986.

[16] I. J. Day. On the inversion of diffusion NMR data: Tikhonov regularization and optimal choice of the regularization parameter. *Journal of Magnetic Resonance*, 211:178 – 185, 2011.

[17] F. Dell'Acqua, G. Rizzo, P. Scifo, R. Clarke, G. Scotti, and F. Fazio. A model-based deconvolution approach to solve fiber crossing in diffusion-weighted MR imaging. *IEEE Transactions on Biomedical Engineering*, 54:462–472, 2007.

[18] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. A fast and robust ODF estimation algorithm in q-ball imaging. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2006.

[19] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58:497–510, 2007.

[20] M. Descoteaux, R. Deriche, T. Knosche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions on Medical Imaging*, 28:269–286, 2009.

[21] A. Einstein. *Investigations on the theory of the Brownian movement*. Dover Pubns, 1956.

[22] G. E. Fasshauer and L. L. Schumaker. Scattered data fitting on the sphere. In *International Conference on Mathematical Methods for Furves and Surfaces II*, 1998.

[23] P. Fillard, M. Descoteaux, A. Goh, S. Gouttard, B. Jeurissen, J. Malcolm, A. Ramirez-Manzanares, M. Reisert, K. Sakaie, F. Tensaouti, et al. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. *NeuroImage*, 56:220–234, 2011.

[24] P. Fillard, C. Poupon, and J.-F. Mangin. A novel global tractography algorithm based on an adaptive spin glass model. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 927–934. Springer, 2009.

[25] L. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47:1083–1099, 2002.

[26] N. S. Hageman, A. W. Toga, K. L. Narr, and D. W. Shattuck. A diffusion tensor imaging tractography algorithm based on navier–stokes fluid mechanics. *IEEE Transactions on Medical Imaging*, 28(3):348–360, 2009.

[27] P. Hagmann, J.-P. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder, and R. Meuli. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *Neuroimage*, 19(3):545–554, 2003.

[28] J. W. Harris and H. Stocker. *The Handbook of Mathematics and Computational Science*. Springer-Verlag New York, Inc., 1997.

[29] K. Hasan, D. Parker, and A. Alexander. Comparison of gradient encoding schemes for diffusion-tensor MRI. *Journal of Magnetic Resonance Imaging*, 13:769–780, 2001.

[30] D. Healy, H. Hendriks, and P. Kim. Spherical deconvolution. *Journal of Multivariate Analysis*, 67:1–22, 1998.

[31] C. Hess, P. Mukherjee, E. Han, D. Xu, and D. Vigneron. A spherical harmonic approach to q-ball imaging. In *International Society for Magnetic Resonance in Medicine*, 2005.

[32] C. Hess, P. Mukherjee, E. Han, D. Xu, and D. Vigneron. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine*, 56:104–117, 2006.

[33] R. R. Hocking. The analysis and selection of variables in linear regression. *Biometrics*, 32:1–49, 1976.

[34] Y. Iturria-Medina, E. Canales-Rodriguez, L. Melie-Garcia, P. Valdes-Hernandez, E. Martinez-Montes, Y. Aleman-Gomez, and J. Sánchez-Bornot. Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. *NeuroImage*, 36(3):645–660, 2007.

[35] M. Jenkinson, C. Beckmann, T. Behrens, M. Woolrich, and S. Smith. FSL. *NeuroImage*, 62:782–790, 2012.

[36] D. Jones. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. *Magnetic Resonance in Medicine*, 51:807–815, 2004.

[37] D. Jones, M. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42:515–525, 1999.

[38] N. Kang, J. Zhang, E. S. Carlson, and D. Gembris. White matter fiber tractography via anisotropic diffusion simulation in the human brain. *IEEE Transactions on Medical Imaging*, 24(9):1127–1137, 2005.

[39] I. Kezele, M. Descoteaux, C. Poupon, F. Poupon, and J. Mangin. Spherical wavelet transform for ODF sharpening. *Medical Image Analysis*, 14:332–342, 2010.

[40] P. Kingsley. Introduction to diffusion tensor imaging mathematics: part II. anisotropy, diffusion-weighting factors, and gradient encoding schemes. *Concepts in Magnetic Resonance*, 28(2):123–154, 2006.

[41] M. A. Koch, D. G. Norris, and M. Hund-Georgiadis. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage*, 16(1):241–250, 2002.

[42] B. Kreher, I. Mader, and V. Kiselev. Gibbs tracking: a novel approach for the reconstruction of neuronal pathways. *Magnetic Resonance in Medicine*, 60(4):953–963, 2008.

[43] B. Kreher, J. Schneider, I. Mader, E. Martin, J. Hennig, and K. Il'yasov. Multitensor approach for analysis and tracking of complex fiber configurations. *Magnetic Resonance in Medicine*, 54:1216–1225, 2005.

[44] D. LeBihan. Molecular diffusion, tissue microdynamics and microstructure. *NMR in Biomedicine*, 8:375–386, 1995.

[45] D. LeBihan, J. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.

[46] C. Lenglet. *Geometric and variational methods for diffusion tensor mri processing*. PhD thesis, Harvard University, USA, 2010.

[47] S. Lifshits, A. Tamir, and Y. Assaf. Combinatorial fiber-tracking of the human brain. *neuroimage*, 48(3):532–540, 2009.

[48] O. Michailovich and Y. Rathi. On approximation of orientation distributions by means of spherical ridgelets. *IEEE Transactions on Image Processing*, 19:461–477, 2010.

[49] S. Mori, B. J. Crain, V. Chacko, and P. Van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of neurology*, 45(2):265–269, 1999.

[50] S. Mori and P. van Zijl. Fiber tracking: principles and strategies–a technical review. *NMR in Biomedicine*, 15(7-8):468–480, 2002.

[51] M. Moseley, Y. Cohen, J. Kucharczyk, J. Mintorovitch, H. Asgari, M. Wendland, J. Tsuruda, and D. Norman. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology*, 176(2):439–445, 1990.

[52] K. Narr, N. Hageman, R. Woods, L. Hamilton, K. Clark, O. Phillips, D. Shattuck, R. Asarnow, A. Toga, and K. Nuechterlein. Mean diffusivity: a biomarker for CSF-related disease and genetic liability effects in schizophrenia. *Psychiatry Research: Neuroimaging*, 171:20–32, 2009.

[53] B. Neuman, L. Bai, and C. Tench. Reliably estimating the diffusion orientation distribution function from high angular resolution diffusion imaging data. In *Medical Image Understanding and Analysis*, 2011.

[54] B. Neuman, C. Tench, and L. Bai. Direct reconstruction of fibre orientation using discrete ground truth interpolation. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2012.

[55] B. Neuman, C. Tench, and L. Bai. Laplace–Beltrami regularization for diffusion weighted imaging. In *Medical Image Understanding and Analysis*, 2012.

[56] B. Neuman, C. Tench, and L. Bai. On computation of diffusion and fibre orientation distribution functions in high angular resolution diffusion imaging. In *Medical Image Understanding and Analysis*, 2012.

[57] B. Neuman, C. Tench, and L. Bai. Regularisation vs elimination: How to accurately reconstruct fibre orientation distribution function. In *International Conference on Medical Imaging Using Bio-Inspired and Soft Computing*, 2013.

[58] B. Neuman, C. Tench, and L. Bai. Tikhonov regularisation in diffusion signal estimation. *Annals of the BMVA*, 8:1–14, 2013.

[59] A. Nikiforov, V. Uvarov, and R. Boas. *Special functions of mathematical physics*. Birkhäuser BaselBoston, 1988.

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

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

[62] G. J. Parker, H. A. Haroon, and C. A. Wheeler-Kingshott. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. *Journal of Magnetic Resonance Imaging*, 18(2):242–254, 2003.

[63] G. J. Parker, C. A. Wheeler-Kingshott, and G. J. Barker. Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *IEEE Transactions on Medical Imaging*, 21(5):505–512, 2002.

[64] P. T. Pope and J. T. Webster. The use of an F-statistic in stepwise regression procedures. *Technometrics*, 14:327–340, 1972.

[65] C. Poupon, B. Rieul, I. Kezele, M. Perrin, F. Poupon, and J. Mangin. New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. *Magnetic Resonance in Medicine*, 60:1276–1283, 2008.

[66] P. Pullens, A. Roebroeck, and R. Goebel. Ground truth hardware phantoms for validation of diffusion-weighted MRI applications. *Journal of Magnetic Resonance Imaging*, 32:482–488, 2010.

[67] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 69:37–38, 1946.

[68] K. Sakaie and M. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted MRI. *NeuroImage*, 34:169–176, 2007.

[69] T. Schultz and H. Seidel. Estimating crossing fibers: A tensor decomposition approach. *IEEE Transactions on Visualization and Computer Graphics*, 14:1635–1642, 2008.

[70] S. Smith. Fast robust automated brain extraction. *Human brain mapping*, 17:143–155, 2002.

[71] S. Smith, M. Jenkinson, M. Woolrich, C. Beckmann, T. Behrens, H. Johansen-Berg, P. Bannister, M. De Luca, I. Drobnjak, D. Flitney, et al. Advances in functional and structural mr image analysis and implementation as FSL. *Neuroimage*, 23:S208–S219, 2004.

[72] S. Sotiropoulos, L. Bai, and C. Tench. Fuzzy anatomical connectedness of the brain using single and multiple fibre orientations estimated from diffusion MRI. *Computerized Medical Imaging and Graphics*, 34:504–513, 2010.

[73] E. Stejskal and J. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42:288, 1965.

[74] P. C. Sundgren, Q. Dong, D. Gomez-Hassan, S. K. Mukherji, P. Maly, and R. Welsh. Diffusion tensor imaging of the brain: review of clinical applications. *Neuroradiology*, 46:339–350, 2004.

[75] J. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35:1459–1472, 2007.

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

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

[78] A. Tristán-Vega, C. Westin, et al. A new methodology for the estimation of fiber populations in the white matter of the brain with the Funk-Radon transform. *NeuroImage*, 49:1301–1315, 2010.

[79] D. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52:1358–1372, 2004.

[80] D. Tuch, T. Reese, M. Wiegell, N. Makris, J. Belliveau, and V. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48:577–582, 2002.

[81] D. Tuch, T. Reese, M. Wiegell, and V. Wedeen. Diffusion MRI of complex neural architecture. *Neuron*, 40:885–895, 2003.

[82] V. J. Wedeen, R. Wang, J. D. Schmahmann, T. Benner, W. Tseng, G. Dai, D. Pandya, P. Hagmann, H. D'Arceuil, and A. J. de Crespigny. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage*, 41(4):1267–1277, 2008.

[83] C. A. M. Wheeler-Kingshott, G. J. Barker, S. C. A. Steens, and M. A. van Buchem. *Quantitative MRI of the brain: measuring changes caused by disease*, chapter 7, pages 203–257. Willey, 2005.

[84] M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. Smith. Bayesian analysis of neuroimaging data in fsl. *Neuroimage*, 45:S173–S186, 2009.

[85] F.-C. Yeh, V. J. Wedeen, W. Tseng, et al. Generalized q-sampling imaging. *IEEE transactions on medical imaging*, 29(9):1626, 2010.

[86] A. Zalesky. DT-MRI fiber tracking: a shortest paths approach. *IEEE Transactions on Medical Imaging*, 27(10):1458–1471, 2008.