COMPARTMENT MODELS AND MODEL SELECTION FOR *IN VIVO* DIFFUSION MRI OF HUMAN BRAIN WHITE MATTER

URAN FERIZI

A dissertation submitted in partial fulfilment
of the requirements for the degree of

Doctor of Philosophy
of the
University College London

November 2014

DECLARATION

I, Uran Ferizi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

November 2014
London, England

Uran Ferizi

Primary Supervisor

DANIEL C. ALEXANDER

(Professor of Imaging Science, Department of Computer Science)

Secondary Supervisor

CLAUDIA A.M. WHEELER-KINGSHOTT

(Reader in Magnetic Resonance Physics, Institute of Neurology, Queen Square, London)

Tertiary Supervisor

OLGA CICCARELLI

(Professor of Neurology, Institute of Neurology, Queen Square, London)

TO MY MOTHER AND BEST FRIEND, *RESMIE LENCKA*.

ACKNOWLEDGEMENTS

I want to thank my supervisor, Prof. Danny Alexander, for supervising and supporting me during my work. Despite his superb research credentials (which I've come to believe are a risky liability in this age of PhD work) and many other engagements, his exemplary academic dedication and open-door attitude towards me, just as to all his students, was crucial to and made my work more productive.

I will also miss Dr. Claudia Wheeler-Kingshott who, beside spending considerable time with my data acquisition and MRI issues on a weekly basis, gave me a very warm and much needed moral support throughout my study.

Among many fantastic friends, I like to mention two, Mr Artan Tafilaj and Sir Noel Malcolm, both of whose virtues are too many to list, and without whom this PhD would have been dull.

Behind me stands an excellent family: my mother, who has been my unwavering friend since childhood, my two sisters Ardiana Ferizi Olloni and Labeate Ferizi Zeneli, brothers, nieces and nephews.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

Ferizi U, Schneider T, Panagiotaki E, Nedjati-Gilani G, Zhang H, Wheeler-Kingshott CAM, Alexander DC: *A ranking of diffusion MRI compartment models with in vivo human brain data.* Magn Reson Med, 2013

Ferizi U, Schneider T, Tariq M, Wheeler-Kingshott CAM, Zhang H, Alexander DC: The Importance of Being Dispersed: *A Ranking of Diffusion MRI Models for Fibre Dispersion Using In Vivo Human Brain Data.* Lecture Notes in Computer Science, Springer Berlin Heidelberg, Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2013, vol. 8149, pp.74–81

Ferizi U, Schneider T, Panagiotaki E, Nedjati-Gilani G, Zhang H, Wheeler-Kingshott CAM, Alexander DC: *Using in-vivo human brain data to select diffusion MRI compartment models.* International Society for Magnetic Resonance in Medicine (ISMRM), 2013

Ferizi U, Schneider T, Panagiotaki E, Nedjati-Gilani G, Zhang H, Wheeler-Kingshott CAM, Alexander DC: *Ranking Diffusion-MRI Models with In-Vivo Human Brain Data.* Proceedings of the International Society for Biomedical Imaging (ISBI), 2013

Ferizi U, Panagiotaki E, Schneider T, Wheeler-Kingshott CAM, Alexander DC: *White Matter Models of In Vivo Diffusion MRI Human Brain Data: A Statistical Ranking.* Proceeding of the 16th Conference on Medical Image Understanding and Analysis (MIUA), 2012

ABSTRACT

Diffusion MRI microstructure imaging provides a unique noninvasive probe into tissue microstructure. The technique relies on mathematical models, relating microscopic tissue features to the MR signal. The assumption of Gaussian diffusion oversimplifies the behaviour of water in complex media. Multi-compartment models fit the signal better and enable the estimation of more specific indices, such as axon diameter and density. A previous model comparison framework used data from fixed rat brains to show that three compartment models, designed for intra/extra-axonal diffusion, best explain multi-b-value datasets.

The purpose of this PhD work is to translate this analysis to *in vivo* human brain white matter. It updates the framework methodology by enriching the acquisition protocol, extending the model base and improving the model fitting.

In the first part of this thesis, the original fixed rat study is taken *in vivo* by using a live human subject on a clinical scanner. A preliminary analysis cannot differentiate the models well. The acquisition protocol is then extended to include a richer angular resolution of diffusion-sampling gradient directions. Compared with *ex vivo* data, simpler three-compartment models emerge. Changes in diffusion behaviour and acquisition protocol are likely to have influenced the results.

The second part considers models that explicitly seek to explain fibre dispersion, another potentially specific biomarker of neurological diseases. This study finds that models that capture fibre dispersion are preferred, showing the importance of modelling dispersion even in apparently coherent fibres.

In the third part, we improve the methodology. First, during the data pre-processing we narrow the region of interest. Second, the model fitting takes into account the varying echo time and compartmental tissue relaxation; we also test the benefit to model performance of different compartmental diffusivities. Next, we evaluate the inter- and intra-subject reproducibility of ranking.

In the fourth part, high-gradient *Connectom-Skyra* data are used to assess the generalisability of earlier results derived from a standard *Achieva* scanner. Results showed a reproducibility of major trends in the model ranking. In particular, dispersion models explain low gradient strength data best, but cannot capture *Connectom* signal that remains at very high b-values.

The fifth part uses *cross-validation* and *bootstrapping* as complementary means to model ranking. Both methods support the previous ranking; however, the leave-one-shell-out *cross-validation* supports less difference between the models than bootstrapping.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AIC Akaike Information Criterion . . . 25

BIC Bayesian Information Criterion . . . 13

CHARMED Composite Hindered and Restricted Model of Diffusion . . . 12

CSF Cerebrospinal Fluid . . . 14

CC Corpus Callosum . . . 12

DT Diffusion Tensor . . . 13

DTI Diffusion Tensor Imaging . . . 56

DW Diffusion-Weighted . . . 23

DWI Diffusion-Weighted Imaging . . . 28

EPI Echo-Planar Imaging . . . 31

FA Fractional Anisotropy . . . 51

FID Free Induction Decay . . . 29

FOV Field-of-View . . . 59

GPD Gaussian Phase Distribution . . . 35

Part I

GENERAL INTRODUCTION

1

MOTIVATION

In medicine, treating a disease requires good knowledge of the changes in the physiology and pathology of the organs involved. Because of its complexity and importance, the brain (arguably) poses the biggest challenge of all organs: to understand its intricate structure on both scales, at the cell-level and as an entire network.

The brain white matter provides connexion between the more peripheral and task-processing cortical centres of grey matter and other parts of the cortex or body muscles. In many brain pathologies, changes in the microstructure integrity or packing of these fibres, known as *axons*, occur as the disease progresses. For example, in Multiple Sclerosis, the degeneration of white matter is one observed finding; other neuro-inflammatory diseases, such as encephalopathy or brain ischaemia, are associated with oedema, which indirectly affects the white matter structure. Identifying these changes with medical imaging, however, is not straightforward considering the very small scale involved: a typical imaging area unit contains around half a million axons.

Diffusion Magnetic Resonance Imaging (MRI) measures the water dispersion in biological tissue and can therefore be used to probe the microstructure. Though useful in other tissue types, this technique is most often applied in the brain, especially where parallel fibres restrict water mobility; it thus provides putative measures of white matter integrity and connectivity. The earliest technique of diffusion MRI, the Diffusion-Weighted (DW) MRI, measures the displacement of water molecules at a scale of a micrometer. From this we learn the restriction that would have provided the pattern of particle dispersion observed. More complicated models can provide more information. The simplest of these is the Diffusion Tensor and, at present, the most widely used. Its assumption of Gaussian diffusion oversimplifies the diffusive behaviour of water in complex media and is known to break down for relatively large diffusion weights (b-values) which can provide higher tissue contrast. This has been addressed by constructing more complex models which provide more specific biomarkers, such as axon diameter, packing density, or dispersion. From these we can learn about the state of tissue integrity.

Modelling approximates the medium which produces the signal and, because of inherent noise, it is guaranteed to be imperfect. A good model tuning should nevertheless balance the danger of over-fitting with the ability to capture most of the signal features. Many models relating microscopic tissue features to the Magnetic Resonance (MR) signal were recently collected and compared with fixed rat brain data [Panagiotaki et al., 2012]. This work compared the models using the BIC, ranking them in order of how well they explain data acquired from the fixed White Matter (WM) of rat CC. The study concluded that three compartment models with

non-zero axon diameter, an anisotropic extracellular compartment, and an isotropic restriction model perform best. However, the results do not directly inform in vivo human imaging experiments because: a) the tissue sample is from a small animal; b) the tissue is fixed, which affects water diffusion significantly [Shepherd et al., 2009] and, therefore, different models may perform better; and c) the experiment used an animal scanner that can achieve higher gradient strengths than human imaging systems.

1.1 PROBLEM STATEMENT

Given so many candidate models for diffusion MRI *brain signal, which describes best the signal from the white matter of the* in-vivo *human brain?*

We are interested not simply in the quality of how well the model fits the data, but also in how robust the model performance and intrinsic parameters are to variations in the data; i.e. how stable the estimated model parameters are to variations in the data noise, how stable the model comparisons are across intra-subject scans, and inter-subject scan/rescan sessions. Considering the amount of experimentation done on fixed and animal tissue, as a proxy for live human tissue, it is also informative to know how different the results are from previous work.

1.2 PROJECT AIMS

- To apply the original model comparison framework [Panagiotaki et al., 2012] to live human data, and evaluate the effects of transition from fixed rat brain;
- to explore other parametric/geometric models which can potentially describe better the signal from and structure of the live human tissue;
- to improve the methodology by adjusting and optimising the data acquisition, improving the model fitting, and expanding and comparing different model selection techniques.

1.3 CONTRIBUTIONS MADE

In Experiment 1, we apply the original model comparison framework of [Panagiotaki et al., 2012] to *in vivo* human data to avoid, and evaluate, the effects of the fixation process. We collect data acquired *in vivo* on the human brain corpus callosum, where fibres bundles are very homogenous and coherently oriented. The multi-compartment geometric models are ranked using standard model selection criteria. Results show that the data acquired cannot differentiate the models well.

In Experiment 2 we explore enriching the acquisition protocol with more diffusion-sensitising gradient directions. This is motivated by the fact that the previous experiment sampled signal from the tissue in only three directions. Additionally, the gradient strengths it could sample were much lower than in the original protocol and, because the scanning is performed on a live human, the scanning time is relatively restricted. This experiment indicated that enhancing gradient angular resolution does indeed help in differentiating between the models.

In Experiment 3, we collect data over two non-stop 4h sessions, using the protocol with the higher angular resolution. Specifically, we use a rich, multi-shell HARDI protocol, to probe a wide range of gradient orientations, diffusion times, gradient pulse times, and gradient magnitudes. As with fixed rat tissue, three compartment models explain the data best. However, a clearer hierarchical structure and simpler models emerge.

We enhance the taxonomy with more models in Experiment 4. These models explicitly seek to explain fibre dispersion in the brain. We drew some models from earlier work in the field; others were adapted from combinations of existing model compartments. This class of models ranked higher than the previous models, and can potentially provide more specific biomarkers of disease. The results demonstrated the importance of modelling dispersion, even in apparently coherent fibres. (With this dataset we organised a challenge, as part of the MICCAI'15 CDMRI workshop, which sought other potentially better candidate models. The other contestants trained their models to three-quarters of our dataset and the model performance was evaluated on the missing quarter.)

We introduce a few methodological enhancements in Experiment 5. First, we fit the models voxel-wise over more homogeneous regions of the genu, midbody and splenium, rather than data averaged across the whole of the corpus callosum, to reduce artificially inflated dispersion. Second, we account explicitly for variable TE among measurements. Third, we study models with compartmentally different T_2 and diffusivity.

In Experiment 6 we test the inter- and intra-subject model ranking reproducibility. We acquire data with the long 8h protocol across the same subject (the "inter-subject" reproducibility), but we reduce the scanning for use across four other healthy subjects. Broadly, the ranking between model groups remains the same, while there are variations within each group.

In Experiment 7 we collect data from the Connectom scanner and repeat the model comparison exercise. Here we explore the generalisability of earlier model comparison results, which use standard human scanners with 60mT/m gradients, to the wider measurement space of human data accessible with the Connectom scanner, affording up to 300mT/m gradients.

Experiment 8 compares the techniques for model selection, in this case for diffusion MRI data. While the method of choice has mainly been the BIC and Akaike Information Criterion (AIC) criteria, we include *bootstrapping* and *cross-validation* as alternative methods.

Part II

BACKGROUND

2

THE BRAIN; THE PHENOMENA OF DIFFUSION

2.1 DISEASES AFFECTING THE BRAIN

In many diseases, a change in the integrity or packing of the cells is a good indicator of disease progression. For example, in cancerous cells the nucleus grows in size very fast, and cytoplasm reduces, affecting the density and hence pressure on fluids to permeate across membranes. In the brain, some neuro-inflammatory diseases, such as encephalopathy or brain ischaemia, are associated with brain oedema (swelling): accumulating fluid which affects the function of other cells. In others, such as Multiple Sclerosis, the degeneration of information-transmitting brain fibres, the neurons, is one observed phenomenon.

2.2 ANATOMY OF NEURONAL FIBRES

In a living organism, the neurons process and transmit information; they make about 10% of the whole central nervous system. As shown in fig.2.1, most neurons have three main parts: the axon, the cell body (soma), and the dendrites. The long tail of the neuron, the axon, is wrapped in lipid-rich myelin, giving it and the whole brain the white colour; functionally the myelin magnifies the conduction of the axonal electrical signal by a factor of 10 to 100 [Trapp and Kidd, 2004].

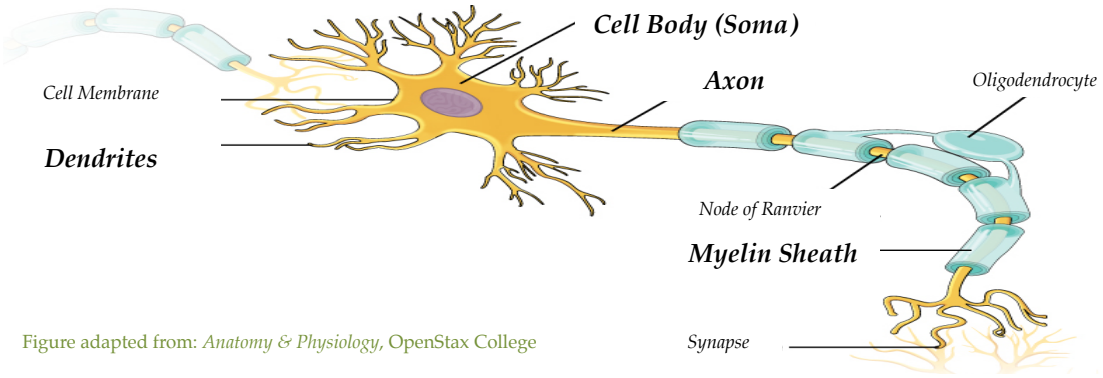


Figure 2.1: A sketch of the neuron, showing its three main parts: the soma, the axon, and the dendrites.

Loosely speaking, the axon lies in the *white matter* part of the brain, the cell body in the *grey*. The *myelin* sheath is rich in fat (about 42%), in addition to protein (18%) and water. The axon thickness in the brain varies from 0.3 to 10 µm. This is roughly proportional to the amount of myelin it is wrapped in, so as to optimise the conduction of the signals passing through. However, the real mechanism as to how the axon regulates the level of myelin is not yet known Nave and Salzer [2006].

3

MAGNETIC RESONANCE IMAGING

3.1 THE PHYSICS OF THE HYDROGEN ATOM

Diffusion-Weighted Imaging (DWI) tracks water hydrogen protons, which are charged particles. In general, as a charged body moves, it changes the magnetic field surrounding it; vice-versa, a changing magnetic field will induce a charge movement that can involve rotational or translational motion. (The exact relationship in space and time of these two fields is described by Maxwell's Equations.) The proton, while moving about chaotically, also spins on its own axis, much like the earth spins on its axis while going around the Sun. Under the influence of a strong external magnetic field, the proton aligns itself with the field.[1]

A Quantum Mechanics treatment of the proton considers the spin to be in one of two energy states, "up" or "down", and accepting energy in photons (packets/quanta) with energy $\hbar\omega_0$ to transition from a lower energy state into a higher one, where Planck's constant $\hbar = 6.626\text{x}10^{-34}$ J s and ω_0 is the photon's frequency. A simpler Classical Physics interpretation can be used because of the number of protons present in the voxel [Liang and Lauterbur, 2000]. As an example to illustrate this, the smallest unit of image, the voxel, contains[2] about $24\text{x}10^{23}$ hydrogen nuclei and, though only about 1 in 200 million hydrogen nuclei/protons would not have a homologous spin with the opposite orientation, there would still be [3]well in excess of 10^{15} excess spins in one orientation than the opposite, thus resulting in a net magnetisation.

3.2 NUCLEAR MAGNETIC RESONANCE

The convention is that the direction of the main magnetic field, used to initially align the protons, is that of z-axis (the vertical axis in fig.3.1). Usually, any quantity in this direction is given the adjective longitudinal; perpendicular to that is called transverse. When the magneti-

1 This behaviour is similar to that of a spinning top/gyroscope. If, and importantly while the top is spinning perfectly straight (say, with the vertical) it is tipped, such that the gravity and the normal reaction force of the surface it stands on are no longer on the same line, a torque (a moment of the forces) will act on the top, and it will then start to go round, or precess, about the vertical axis.

2 *The voxel* is of size $(2\text{mm})^3$, or $8\text{x}10^{-3}$ml. This volume would be filled by about 8g of water, and with water's Molar Mass of 18gr/mol, this is about 2.25 mol of water/hydrogen molecules, or 2.25mol x 2 H-atoms x Avogadro's constant = $24.3\ \text{x}10^{23}$ hydrogen atoms.

3 *Boltzman's Equation:* $N^{spin-up}/N^{spin-down} = e^{\triangle E/KT}$, where Boltzmann's Constant $K = 1.4 \times 10^{-23} JK^{-1}$, and (Einstein+Larmour's) $E = \hbar\omega_0 = \hbar\gamma B_0$, where Planck's constant $\hbar = 1 \times 10^{-34} Js$, the gyroscopic ratio $\gamma = 267.5 \times 10^6 rad \cdot s^{-1} \cdot T^{-1}$. Thus, for $B_0 = 3T$, $N^{spin-down}/N^{spin-up} \sim 1 + 10^{-8}$, meaning that only 1 in every 200 million protons does not have a homologue with an opposite spin. Further, using the result given in the footnote above, a voxel would contain about 10^{15} excess spins aligned with the external field.

sation vector is fully aligned with the z-direction, we say that the system is fully magnetised. Conversely, when this vector is completely tipped to the transverse plane the system is said to have reached saturation.

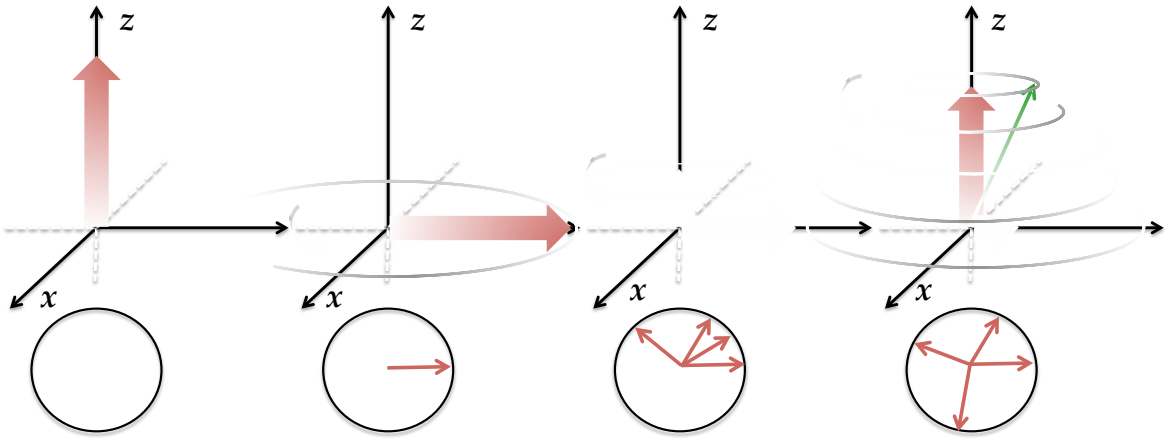


Figure 3.1: Spin magnetisation before/during/after an RF pulse. The first plot from the left shows the total magnetisation of the spins aligned in the B_0 field; the circle below shows the vector on the transverse plane which we measure. The second plot shows the spins tipped by the RF pulse to align with the transverse/horizontal axis; the spins are in-phase. The third and fourth plots illustrate the start and continuation of the de-phasing of the spins: transverse magnetisation decreases, longitudinal magnetisation increases.

The tissue-specific response of the individual nuclei/protons' magnetisation is termed *relaxation*. There are two main parameters that quantify this: T_1 and T_2 relaxation times.

T_1 (longitudinal or spin-lattice relaxation) time measures how fast the original longitudinal component is recovered: after energy is transmitted at Larmour RF, the spin vector turns by 90°; later, as energy diffuses through the structure/lattice, the vector returns to its original position, and its magnitude is given by:

$$M_z(t) = M_0 \cdot \left[1 - \exp\left(\frac{-t}{T_1}\right)\right] \tag{3.1}$$

T_2 (transverse or spin-spin relaxation) time concerns de-phasing of the different spins. While on the transverse plane, the net vector will also shrink in size. This comes from the loss of the synchronisation of the spins which were initially in-step. In T_2, B_0 is assumed to be homogeneous, and the interaction of the particles is assumed to be of a purely random nature, and its magnitude dependence on time is:

$$M_{xy}(t) = M_0 \cdot \left[1 - \exp\left(\frac{-t}{T_2}\right)\right] \tag{3.2}$$

The reality is that the (B_0) field will not be homogeneous, so another T_2 time, the T_2^* is introduced, which also accounts for these inhomogeneities. T_2^* is faster than T_2 decay. The signal affected by T_2^* decay is called the Free Induction Decay (FID) signal.

3.3 SPIN-ECHO SEQUENCE

In the work we will be using the basic Nuclear Magnetic Resonance (NMR) sequence of SE. Figure 3.2 illustrates the SE experiment. The first half is similar to the process in fig.3.1: the 90° pulse flips the spins onto the transverse plane, which then start to precess at slightly different phases, thus also losing their spatial coherence and total magnetisation. The second RF pulse is what reverses this process, and thus recovers some of the signal back, the *spin-echo*, at time TE. Though in some experiments this 180° pulse can be repeated to achieve multiple spin echoes, usually, the SE sequence is repeated again after time Repetition Time (TR) (TR being the time between two 90° pulses).

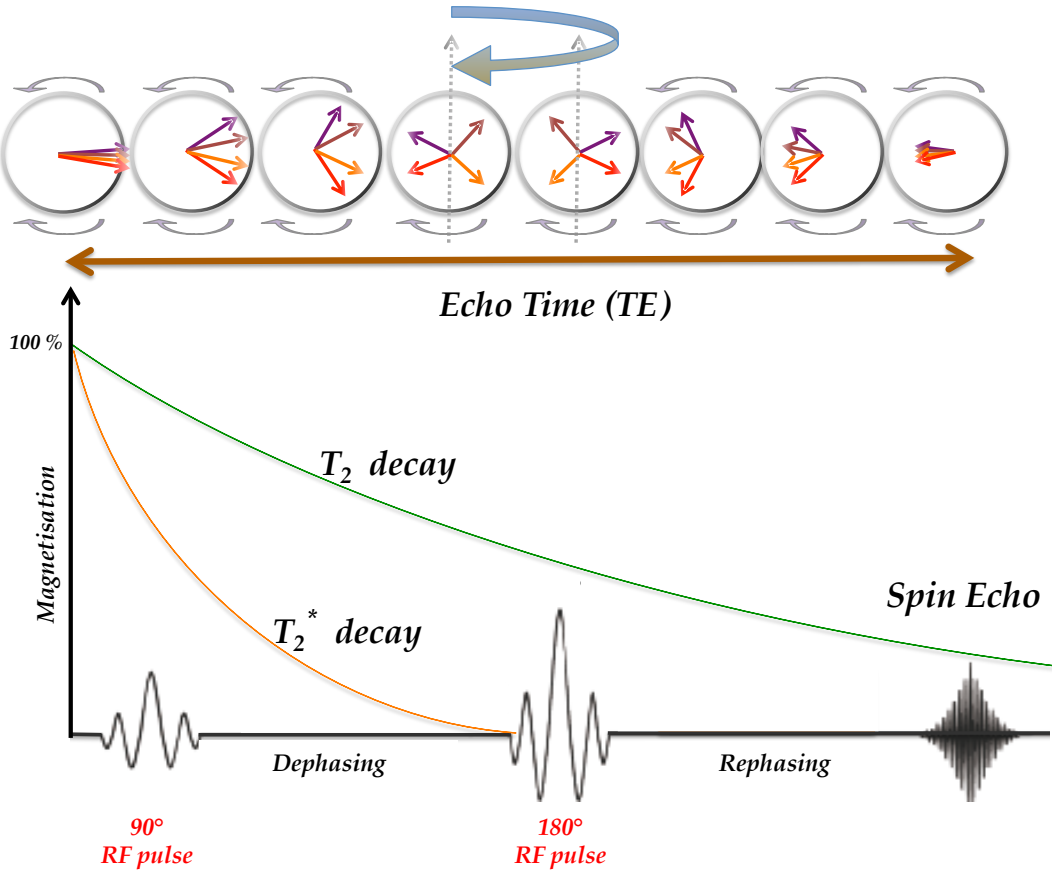


Figure 3.2: The SE sequence. The first 90° excitation pulse flips the spins on the plane perpendicular to the B_0 field. With time, phase differences reduce in the resultant magnetisation, as shown by the circles on the stop panel; the T_2 also causes a reduction in the magnetisation. At time TE/2, the second refocusing 180° RF pulse flips the spins, leaving the faster (precessing) spins behind the slow ones. At time TE the spins are in phase with each-other; this is *spin-echo*.

Different kinds of tissue have different characteristic T_2, e.g. white matter in the brain has much shorter T_2 than the CSF. This provides a natural means of contrasting between different tissue types. In this type of imaging, a balance needs to be struck between low TE, which produces good signal, and high TE which increases the contrast.

3.4 IMAGE FORMATION

The hydrogen proton NMR signal is used in MRI to construct body images. Hydrogen is mostly present in body water or fat[4]. There are three steps to localise the signal from which part of the body it came from. This is done by applying RF pulses in turn, to gradients in three directions, to select the slice and encode frequency and phase [Smith, 1985]; the diffusion-weighting will be discussed in the next chapter.

SLICE SELECTION To encode a slice, a gradient is applied in the z-direction at the same time as the sinc-shaped RF pulse, giving spins of different z-coordinate a different frequency. [5]

IMAGE ENCODING Applying a uniform magnetic field B_0 will give all particles the same Larmour frequency:

$$\omega_0 = \gamma B_0$$

where $\gamma = 267.5 \times 10^6 \ \frac{rad}{sT}$. This is the basic law in MRI, and a more detailed derivation is given in sections A.1 and A.2 of the Appendix.

Adding a gradient $\mathbf{G}$, in the B_0-direction, will set all the spins at their respective Larmour frequency

$$\omega(x) = \omega_0 + x\mathbf{G}(x)$$

The resonance frequency is proportional to the position of the spin and, further, the plane perpendicular to the gradient will have a signal proportional to its number of spins.

For phase encoding, the aim is to give the transverse magnetisation vector of different spins a different phase angle. The same is applied as for frequency, up to the point where along x direction spins have different Larmour frequencies. But when this phase-gradient is stopped, the spins would be left with only the B_0 magnetic field, hence an ω_0 frequency but, crucially, also with a characteristic phase.

K-SPACE If the image to be constructed has, for example, 64 points in the phase encoding direction, its gradient will require 64 different magnitudes (which will produce 64 FID curves). The same applies to the frequency encoding. The raw coil signals $F(k_x, k_y)$ are Fourier transforms of the x-y location image function $f(x, y)$, frequency encoding in the (say) x-direction and phase encoding in the y-direction.[6] Inverse Fourier transformation then produces a MR image of spin location.

4 Some background information on this chapter and the next are taken from *http://www.cis.rit.edu/htbooks/mri/*

5 Here, Fourier Transform techniques are used to translate the information from the dense time domain to the frequency one. The importance of the sinc-shape, as opposed to a simple sinusoidal, is to make the slice selection possible; sinc in the time domain translates via Fourier Transform to a hat function in the frequency domain.

6 Single-shot Echo-Planar Imaging (EPI) is used to acquire the full k-matrix in one 'shot' Mansfield and Pykett [1978].

4

DIFFUSION MRI

4.1 DIFFUSION PHYSICS: THE FIRST STUDIES

In 1822 Fourier [1878] gave a mathematical framework for the laws of heat dissipation, observing that between two neighbouring solid particles, "the most heated molecule communicates to the less heated a quantity of heat expressed by the product of the instant duration, of the small difference of the temperatures, and of a certain function of the distance of the molecules" Philibert [2005].

Coming from a more practical angle, in the autumn of 1826, Robert Brown [1828] set about investigating the "mode of action of the pollen in the process of impregnation". Through his modest microscope, he reported the behaviour of pollen grains suspended in water as giving a pattern of motion which "arose neither from currents in the fluid, nor from its gradual evaporation, but belonged to the particle itself".

To reinforce Brown's point that the particles had an innate ability to diffuse out and into other media, in 1845, the Glaswegian chemist Thomas Graham, famous[1] for his work on the diffusion of gases, described the gaseous particles, "when brought into contact, do not arrange themselves according to their density, the heaviest undermost, and the lighter uppermost, but they spontaneously diffuse, mutually and equally, through each other, and so remain in the intimate state of mixture for any length of time".

In 1855, the Zurich-based physiologist Adolf Fick connected the conduction of heat in solids with the diffusion of particles, thus adopting Fourier's mathematical formulation of heat dissipation as the standard model for general diffusion. He expressed his phenomenological relation between change of concentration ϕ in time t, diffusivity D, and spatial variation x via:

$$\frac{\partial \phi}{\partial t} = -D \frac{\partial^2 \phi}{\partial x^2}$$

He verified these results on the diffusion of salt in water, and reiterated another point made by Graham, which was that the diffusivity increased as the temperature increased.

Einstein later derived the relationship for n-dimensional space between diffusivity D and the mean squared displacement, $< R(t)^2 >$:

$$D = \frac{< R(t)^2 >}{2n\,t}$$

This also provided a relationship between the microscopic scale, through the 'mean squared displacement', and the macroscopic scale, through Fick's diffusivity.

1 Incidentally, Graham is the discoverer of dialysis, the method of separating particles of various dimensions, hence different rates of diffusion, through a semi-permeable membrane.

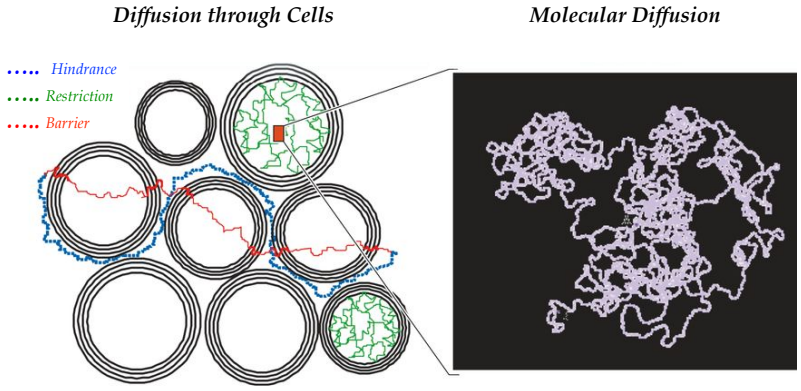


Figure adapted from: **Denis Le Bihan**, *'Diffusion MRI: what water tells us about the brain'*, EMBO Molecular Medicine (2014) -- with publisher's permission

Figure 4.1: The types of diffusion in brain tissue. While water molecules individually perform a random walk, as shown on the right, the cells' boundaries can alter the shape of this diffusion, as shown on the left.

4.2 DIFFUSION IN THE BRAIN

The phenomenon of diffusion has emerged as a powerful tool for probing the microstructure of the brain. Moseley et al. [1990] showed the potential of this emerging DWI technique in clinical practice by detecting ischaemic stroke earlier than with other techniques (T_1 and T_2 weighted imaging). Later, Le Bihan et al. [1986] showed that MRI of water diffusion can be used to image brain tumours. This is made possible by the scale of restriction that the tissue barrier imposes on the water diffusion.

Water diffusion can provide information about the underlying structure. At least within the DWI community, the different types or 'scales' of diffusion, also shown in fig.4.2, are termed as *free, hindered, and restricted*. As the names suggest, within a fibre, a water particle can be thought of as being free to diffuse along the fibre, but restricted to move across it; between the fibres the movement is rather tortuous, so this is termed 'hindered' diffusion.

4.3 THE START OF DIFFUSION IMAGING

It was Hahn [1950] who recognised that (water/hydrogen proton) spin echoes were affected by water diffusion and, later, Carr and Purcell [1954] proposed a direct way to measure this diffusion, by adding diffusion-sensitising gradients to Hahn's SE sequence. However, it is the next adaptation, the PGSE sequence used by Stejskal and Tanner [1965], which is now most often used in diffusion imaging. (It was also in this period Stejskal [1965] that we saw the first use of a *tensor* to map the dispersion pattern of particles.)

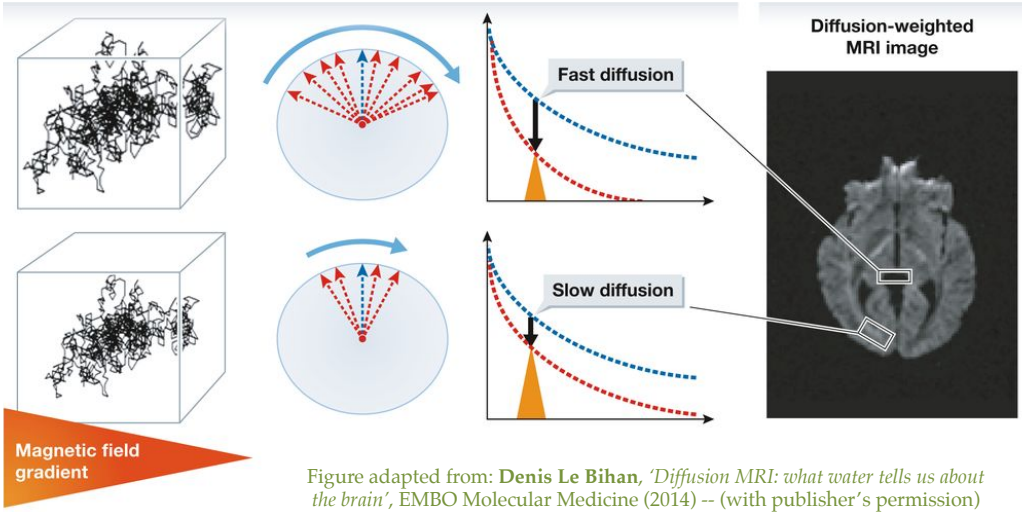


Figure adapted from: **Denis Le Bihan**, *'Diffusion MRI: what water tells us about the brain'*, EMBO Molecular Medicine (2014) -- (with publisher's permission)

Figure 4.2: The restricted and hindered diffusion can give different apparent patterns: the restriction slows mobility, but diffusion free of boundaries is comparably faster.

4.4 DIFFUSION SENSITISATION

Compared with the SE sequence of Carr and Purcell [1954], the PGSE sequence does not apply the gradients throughout the diffusion encoding, but in pulses, so effectively distinguishing between diffusion and its encoding. The gradients can be applied in any of three directions, x, y or z, so as to obtain images in those directions. The gradients, placed either side of the 180° pulse (see fig.4.3), are applied in the same direction, and are of equal magnitude $|\mathbf{G}|$ and duration δ (desired to be as short as possible).

While the two gradients (before and after the 180° pulse) produce no net phase offset on stationary spins, for those spins which have diffused (i.e. changed location) during time Δ there will be a net phase remaining. So, if a spin is initially at position $\mathbf{r}$ and after some time has diffused to $\tilde{\mathbf{r}}$, then the phase change associated to this spin has changed [2] by $\mathbf{q}\cdot\mathbf{x}$, where *wavevector* $\mathbf{q} = \gamma\delta\mathbf{g}$ and displacement $\mathbf{x} = \mathbf{r} - \tilde{\mathbf{r}}$. Therefore, the signal at the spin-echo would be:

$$\mathbf{S}(q) = \mathbf{S}(0)\int p(\mathbf{x})\exp(-i\mathbf{q}\cdot\mathbf{x})d\mathbf{x} \tag{4.1}$$

where $\mathbf{S}(0)$ is the signal obtained in the absence of diffusion-sensitising gradients , $p(\mathbf{x}; D, t)$ is the conditional probability that a spin arriving at $\tilde{\mathbf{r}}$ originated from $\mathbf{r}$ during this diffusion time; it is this p which we will try to capture via various models [Alexander, 2006]. For example, one simple model for probability $p(\mathbf{x}; D, t)$ can follow an isotropic zero-mean Gaussian (Normal) Distribution

$$G(\mathbf{x}; \mathbf{D}, t) = \frac{1}{\sqrt{(4|\mathbf{D}|\pi t)}}\exp\left(-\frac{\mathbf{x}^t\mathbf{D}^{-1}\mathbf{x}}{4t}\right) \tag{4.2}$$

2 This comes from wave mechanics jargon where a wave's position function-vector is $\Psi(\mathbf{x}, t) = \Psi_0\exp[i(\mathbf{k}\cdot\mathbf{x} + \omega t)]$, where ω is the temporal frequency (so ωt is the temporal phase) and $\mathbf{k}$ is termed *wavevector* and its magnitude k *wavenumber* (a sort of "spatial frequency" to give a spatial phase $\mathbf{k}\cdot\mathbf{x}$).

where **D** is either the scalar fluid diffusivity D or the 3x3 matrix DT. We will be discussing other models in the next chapter.

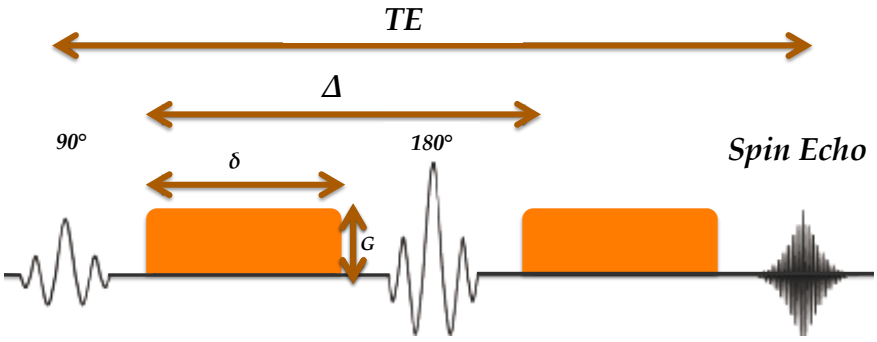


Figure 4.3: The PGSE sequence [Stejskal and Tanner, 1965]. The first RF pulse tips all the spins by 90°, and the second 180° RF 'refocusing' pulse reverses the phases of those spins. Differently to Hahn [1950], the gradients are not applied throughout TE, but are split into two, so as to distinguish between diffusion Δ and its encoding δ (which is desired as short as possible). At time TE the spin echo forms.

4.5 ONE APPROXIMATION; AND ONE ASSUMPTION

The derivation of equation 4.1 assumes that the pulse duration δ is very short compared with diffusion time Δ, so that the spins do not diffuse during the application of the pulse. This is known as the Short Gradient Pulse (SGP) approximation. However, this is very hard to achieve in practice, especially with standard clinical scanners; the assumption is often violated in order to achieve b-values high enough to give adequate diffusion contrast. This inability to keep $\delta \ll \Delta$ implies the need for careful modelling interpretation of the process, as the true particle displacement may in fact be underestimated [Mitra, 1995].

The Gaussian Phase Distribution (GPD) assumption [Murday and Cotts, 1968] models analytically the effects of finite δ. When p in equation 4.1 is Gaussian and the pulses are rectangular, then the phases of the spins due to the magnetic field gradients are Gaussian-distributed.

Relevant to this section, we will introduce one variable which will be abundant for the rest of this thesis. Generally in PGSE diffusion imaging, the experimental tuneables, gradient strength $|\mathbf{G}|$, diffusion time Δ and gradient pulse δ, are often combined and reported as a "diffusion weighting" factor:

$$b = \gamma^2 |\mathbf{G}|^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right) \tag{4.3}$$

where $\gamma \sim 42\text{MHz/T}$ is the proton's gyromagnetic ratio. As explained in Stejskal and Tanner [1965], for the PGSE experiment, unhindered Gaussian diffusion would attenuate the spin-echo signal by a factor of $\exp(-bD)$, where D is the diffusivity.

5

MODELLING DIFFUSION-MRI SIGNAL

5.1 THE DIFFUSION TENSOR

Magnetic resonance (MR) microstructure imaging uses mathematical models to relate MR signals in each image voxel to microscopic tissue features, and thus estimate and map histological features. Diffusion MRI measures water diffusion in biological tissue, which can be used to probe the microstructure.

In brain imaging, the standard model for water dispersion in tissue is the DT [Basser et al., 1994], which assumes a trivariate Gaussian dispersion pattern (there is more on *tensors* in the Appendix A.3). This assumption of Gaussian diffusion oversimplifies the diffusive behaviour of water in complex media, and is known experimentally to break down for relatively large b-values. DT derived indices, such as *mean diffusivity* or *fractional anisotropy*, can correlate with major tissue damage, e.g. in ischaemic brain injury [Sotak, 2002] or Multiple Sclerosis [Castriota-Scanderbeg et al., 2002], but lack sensitivity and specificity to subtle pathological changes. Indices of such changes may include axon radius, density, orientation, dispersion and permeability, which potentially give much greater insight into tissue architecture and pathology. In diffusion MRI, the standard DT model has two key limitations: first it is too simple to explain the data over a wide range of b-values and orientations; second, it lacks specificity to particular tissue features.

5.2 MULTICOMPARTMENT MODELS

To address the limitations of the DT, a variety of alternative biophysical diffusion *MRI* models have emerged over the last decade to address these limitations; we give their mathematical formulation in Appendix A.4. These models underpin the emerging generation of microstructure imaging techniques that are now starting to replace DT-imaging in a range of biological and clinical studies into tissue microstructure variation.

5.2.1 *Straight Fibres*

Stanisz et al. [1997] pioneered the multi-compartment representation of separate diffusive processes in nervous tissue. As shown in fig.5.1, the model had three geometric compartments: ellipsoids for restricted intra-axonal water, anisotropically hindered extracellular water (with diffusivity and relaxation constants different to the intracellular space) and isotropically restricted glial cell water. Permeability was included via exchange terms from each of the three

compartments (effectively, as an extra diffusion coefficient). They justified the need for more than one compartment by conducting an experiment that demonstrated the breakdown of mono-exponential signal decay (of the DW model) as diffusion time Δ was varied. They then used an analytical model, instead of Monte-Carlo simulations, to generate data that calibrates the model for each medium separately. Next, they fitted the model to signal from ex-vivo bovine optic nerve fibre to deduce volume size and diffusivity, and concluded that three was the minimum number of compartments to adequately model the tissue.

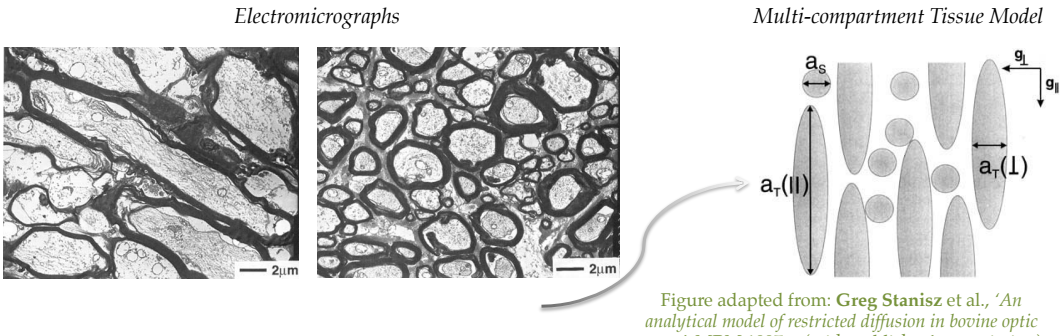


Figure adapted from: **Greg Stanisz** et al., *'An analytical model of restricted diffusion in bovine optic nerve'*, MRM 1997 -- (with publisher's permission)

Figure 5.1: The Stanisz et al. [1997] compartments in the model for bovine optic nerve are shown on the right. The motivation for such choice of geometry is shown on the left: electro-micrograph shots of tissue in the parallel (left) and perpendicular (middle) to the axis defined by the orbit and optic chiasm.

The Ball-Stick, a 'simple partial volume model', [Behrens et al., 2003] is the simplest two-compartment model designed to capture the fibre structure. To describe voxel signal there are two compartments. The first, the Stick compartment, models restricted diffusion inside and around the axons; in theory, it can accommodate any distribution of fibre orientations, but, in this study, the fibre bundle is simplified to be coherent and straight. The second compartment, the Ball, captures any free-water in the voxel; this includes diffusion across the fibres, which is radially symmetric. The model assumes no inter-compartmental diffusion exchange.

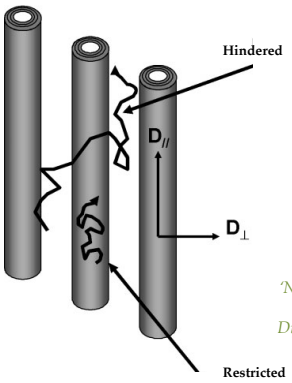


Figure adapted from: **Yaniv Assaf** et al., *'New Modeling and Experimental Framework to Characterize Hindered and Restricted Water Diffusion in Brain White Matter'*, MRM 2004 -- (with publisher's permission)

Figure 5.2: The CHARMED model for brain white matter. The two types of diffusion are 'hindered' outside the cylinders, represented via a diffusion tensor, and 'restricted' inside the cylinders, which is decomposed into parallel and perpendicular diffusivities.

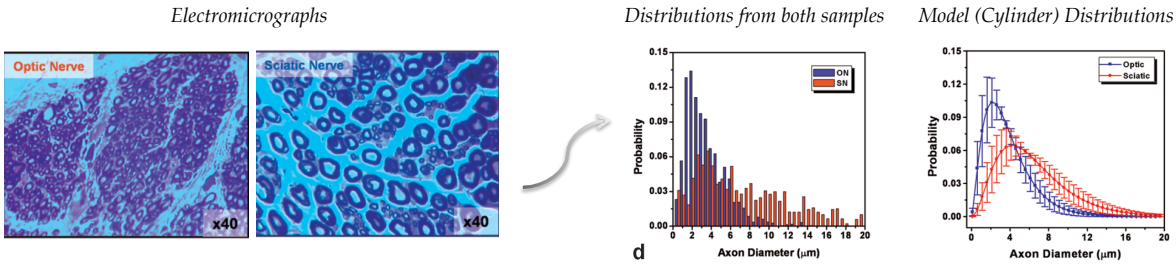


Figure adapted from: **Yaniv Assaf** et al., *'AxCaliber: A method for measuring axon diameter distributions from diffusion MRI'*, MRM 2008 -- (with publisher's permission)

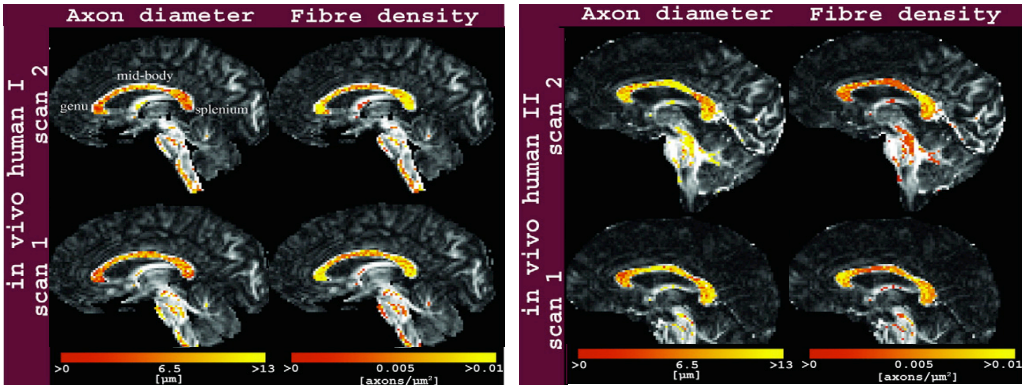
Figure 5.3: The AxCaliber model aims to capture the distribution of fibre thickness. In this application, the model-derived axon-thickness distributions (shown on the right) show a good reproducibility of the measured distribution of axon diameters (in the middle) derived from both the optic nerve (top-left) and the sciatic nerve (bottom-left).

The Composite Hindered and Restricted Model of Diffusion (CHARMED) by Assaf et al. [2004] is designed to capture diffusion explicitly inside and outside the axons. It replaces Behrens' intracellular stick with cylindrical impermeable fibres, and replaces the extracellular Ball with a full DT. As shown in fig.5.2, initially the model used fixed diameter distribution to estimate fibre orientation and volume fraction. Later work on AxCaliber [Assaf et al., 2008] additionally estimates the parameters of the gamma distribution (see fig.5.3). The first study [Assaf et al., 2004] used simulations and excised spinal cord data to validate the technique but, subsequently, Assaf and Basser [2005] applied the technique on healthy human subjects, using a single TE of 133 ms, a diffusion-sensitising gradient of 40mT/m, and a multi-shell gradient sampling protocol which sought to address the decreasing Signal-to-Noise Ratio (SNR) as b-values increased up to 10,000 s/mm^2.

ActiveAx [Alexander et al., 2010; Dyrby et al., 2013] combines elements of Stanisz's model and AxCaliber to obtain the simplest model called the Minimal Model of White Matter Diffusion (MMWMD) that adequately fits data while providing estimates of axon density and diameter. The four compartment model includes a single axon diameter and cylindrically-symmetric extracellular diffusion, which makes it parsimonious enough to obtain orientationally-invariant parameter estimates. The initial ActiveAx application by [Alexander et al., 2010] used a previously optimised protocol [Alexander, 2008] to collect a four-HARDI-shell dataset from *in vivo* healthy subjects to validate the technique. Figure 5.4 shows that it can recover known distributions of white matter indices obtained from histology. E.g. axon diameters are estimated as lower in the genu and splenium than in the midbody; the opposite applies to the axon density.

5.2.2 *Dispersed Fibres*

The simple models above do not account for fibre direction inhomogeneity which is abundant in the brain even at a sub-voxel level [Jeurissen et al., 2010; Zhang et al., 2011, 2012]. A wide



Figures adapted from: **Daniel C. Alexander** et al., *'Orientationally invariant indices of axon diameter and density from diffusion MRI'*, NeuroImage 2010 --- with publisher's permission.

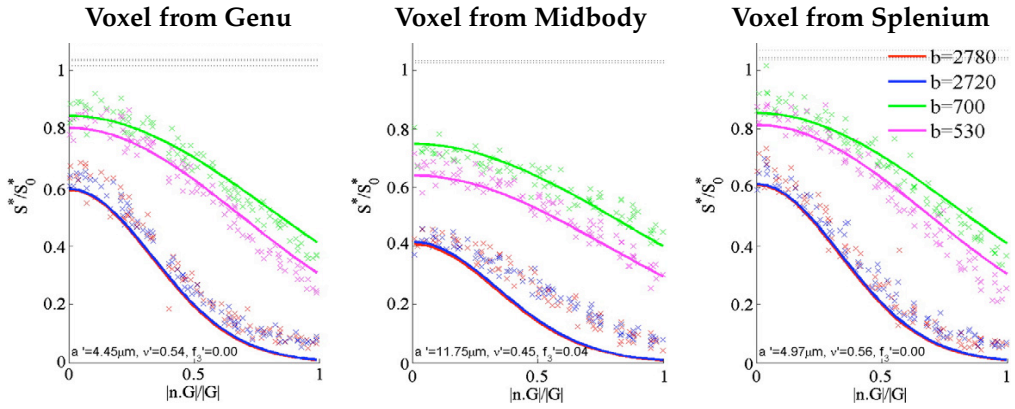
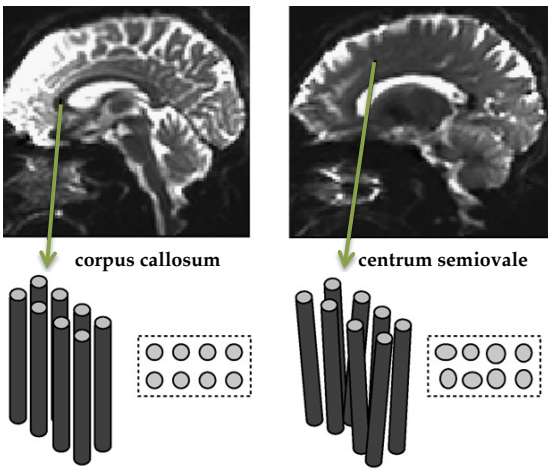


Figure 5.4: The top row shows voxel-wise indices from the *ActiveAx* model: axon thickness and density. The second row shows how well the model captures the underlying signal in three CC regions.

family of multiple fibre reconstruction algorithms [Seunarine and Alexander, 2009; Tournier et al., 2011] aim to recover multiple fibre orientations or the fibre orientation distribution, but these are not directly relevant here as they do not separate different tissue compartments. However, various compartment models incorporate the idea. Hosey et al. [2005] and Behrens et al. [2007] extend the Ball-Stick model to multiple Sticks, capturing multiple fibre populations with distinct orientation. Similarly, Assaf and Basser [2005] extend Assaf et al. [2004] to multiple intracellular compartments. The DIAMOND model [Scherrer et al., 2013] also employs discrete fibre populations and the number of fibres within each voxel is determined via a model selection framework based on the generalization error. Jeurissen et al. [2010] combine the constrained spherical deconvolution model of the fibre orientation distribution from Tournier et al. [2007] with additional grey matter and CSF compartments.

Zhang et al. [2011] extend the MMWMD to relax the assumption of straight parallel fibres to a Watson distribution of orientations. This is particularly important in regions where the fibres as not as coherently aligned with each other as, say, in the corpus callosum, and which can lead to inaccurate axon diameter estimation; see also fig.5.5 for an illustration of these differences. This point was also made in the study by Nilsson et al. [2012]. Through Monte Carlo simulations of water diffusion it showed that axonal undulations, abundant in structures such



		Maximum angle between $\vec{\mu}$ and $\vec{n}$						
		5°	10°	15°	30°	45°	60°	75°
κ	4	2.5%	9.5%	20%	56%	80%	91%	97%
	8	5.5%	20%	39%	84%	97%	99%	100%
	16	11%	37%	65%	98%	100%	100%	100%
	32	21%	61%	88%	100%	100%	100%	100%
	64	38%	85%	99%	100%	100%	100%	100%
	128	62%	98%	100%	100%	100%	100%	100%

Figure and table adapted from: **Hui Zhang** et al., *'Axon diameter mapping in the presence of orientation despersion with diffusion MRI'*, NeuroImage 2011 -- (with publisher's permission)

Figure 5.5: Modelling dispersed cylinders. On the right, the figure illustrates the potential for overestimation of axon indices, such as diameter, if the fibres are not coherently oriented. On the left, the table gives an intuitive indication of the (Watson) dispersion parameter κ: each percentage gives the probability that an orientation $\vec{n}$ sampled from a κ-specific distribution and the mean orientation $\vec{\mu}$ are within the given angle.

as the spinal cord and optic nerve, can distort the diffusion MRI signal and model estimates, as shown in fig.5.6.

The dispersed cylinders of Zhang et al. [2011] led to the simpler Neurite Orientation Dispersion and Density Imaging (NODDI) model [Zhang et al., 2012], which further assumes zero axon diameter (Watson distributed Sticks rather than Cylinders). To make the technique simple and practical for clinical practice, the optimisation produced a two-HARDI-shell scanning based on the original four-shell protocol (and showed that reducing the protocol further, to one shell, sacrificed the specificity of the neurite density). The method was evaluated on *in vivo* human brain, and the results produced sensible estimates for axon dispersion and density, potentially disentangling the two contributing factors to the DT *fractional anisotropy* index. Later work [Tariq et al., 2014] incorporates dispersion anisotropy by replacing the Watson distribution with a Bingham distribution.

Concurrently with NODDI, Sotiropoulos et al. [2012] extended the Ball-Stick model to Ball-Rackets also using a Bingham distribution of Sticks, as in fact proposed earlier in Kaden et al. [2007]. The Ball-Rackets study uses simulations to evaluate the model accuracy (see fig.5.7) and fixed macaque brain data; the scanning acquisition uses a single 120-direction HARDI shell of b-value 8,000s/mm^2 on which to validate the model, and then compare the results with five other non-parametric techniques.

5.3 HIGHER GRADIENTS, HIGHER DIFFUSION CONTRAST

Higher gradients make possible faster diffusion encoding, giving shorter diffusion times Δ, hence shorter TE, and higher SNR. They also provide a higher-contrast pattern of the water

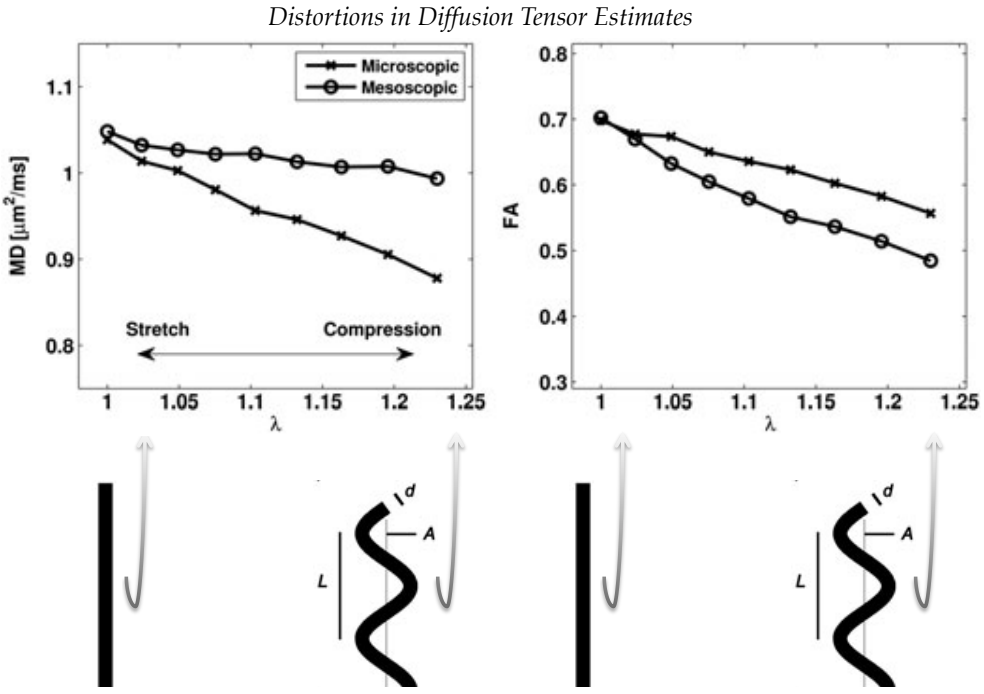


Figure adapted from: **Markus Nilsson** et al., *'The importance of axonal undulation in diffusion MR measurements: a Monte Carlo simulation study.'*, NMR Biomed. (2011) -- with publisher's permission

Figure 5.6: The Monte Carlo simulation of water diffusion showed that both the signal and, hence, the model parameters fitted to the data are distorted by the undulation.

diffusion; in a PGSE experiment, the b-weighting is proportional to the square of the gradient strength. (Increasing the field strength, on the other hand, affects the measured T_2 [Cox and Gowland, 2010] and SNR, but it does not affect the diffusion contrast.) While there are limitations on human imaging systems, which ordinarily reach up to 40 or 60mT/m gradients, animal systems can ramp up to 1,000mT/m.

By applying gradients of up to 282 mT/m only across the corpus callosum fibres, Barazany et al. [2009] uses AxCaliber to recover in the *in vivo* rat corpus callosum a gamma distribution of axon diameters that correlates well with findings from histology. Recovering these tissue anatomical trends supports the technique, but a secondary issue is that the mean AxCaliber diameters are slightly higher than histological estimates and the Cylinders/axonal volume fraction is only estimated at one-fifth of the total volume. The authors attribute the mismatch between axon diameters to the fixing of the tissue during histology, and different compartmental T_1 and T_2 weighting for the low intra-axonal volume fraction. The study by Panagiotaki et al. [2012] also collects data using an animal imaging system. The experiment applies to fixed rat brains a wide range of b-values, using up to 400mT/m gradient strength. This data is then used for comparing models similar to, among others, CHARMED and MMWMD; it additionally finds that models with a single axon diameter distribution, rather than a two-parameter gamma distribution, provide a more stable fitting.

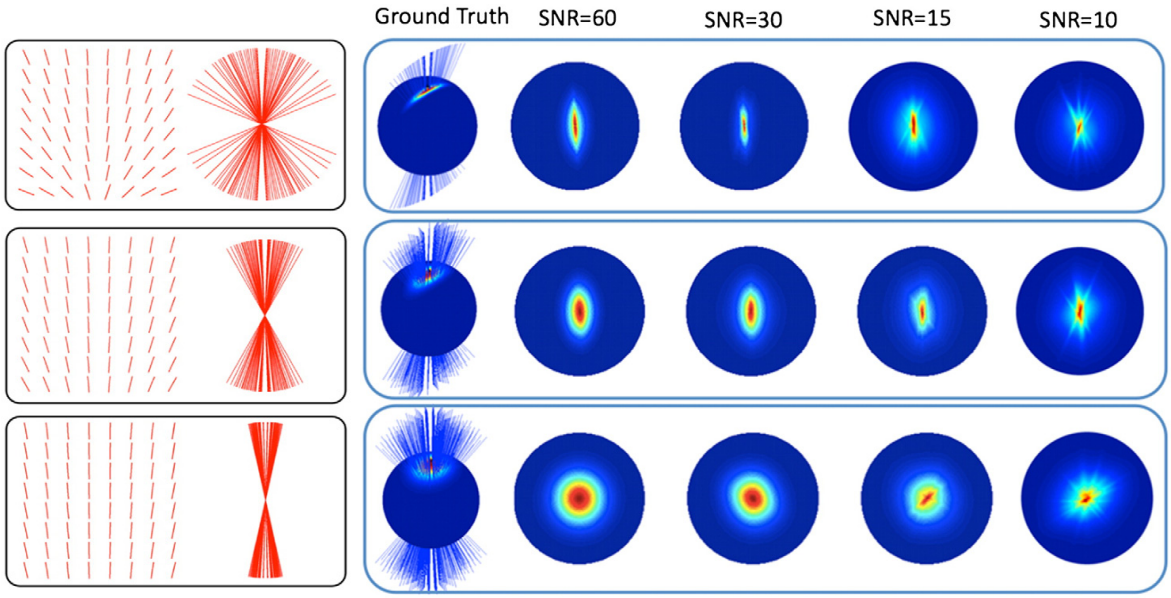


Figure adapted from: **Stamatios N. Sotiropoulos** et al., *'Ball and Rackets: Inferring fiber fanning from diffusion-weighted MRI'*, NeuroImage (2012) -- with publisher's permission

Figure 5.7: Dispersed sticks via a Bingham distribution; the top row corresponds to the most anisotropic distribution (maximum fanning); the bottom row corresponds to an isotropic distribution (equivalent to the Watson distribution). On the left, in red, are the sub-voxel grids used to simulate the fanning; on the right are the simulated distributions at various noise levels.

Using *ex-vivo* monkey brain data, Dyrby et al. [2013] illustrate the benefits of stronger gradients, which provide higher water diffusion sensitivity. In particular, the study showed that higher gradients provide higher sensitivity of axon diameter distributions in white matter, especially at the lower end (less than ~ 2.5µm), using fixed post-mortem tissue and a small-bore animal imaging system. The study increases the gradients applied from 60, which are approximately the maximum achieved in standard clinical scanners, to 300 mT/m.

The recent development of human MR systems with 300mT/m gradients, in particular the Connectom scanner [Setsompop et al., 2013], is a major step towards the long-term translation of microstructure imaging techniques to widespread clinical practice. The first experiments verifying those findings on live human subjects are now beginning to emerge [McNab et al., 2013; Duval et al., 2014; Huang et al., 2014]. McNab et al. [2013] provides three initial applications, two in diffusion tractography and another on axon diameter estimation in the *in-vivo* healthy human brain. The latter application, which is most relevant to this work, revealed the feasibility and challenges of extracting microstructural information from the living tissue.

5.4 COMPARING DIFFUSION-MRI MODELS

The wide-ranging set of available models relating the diffusion signal to microstructural tissue features raises the question about which are most appropriate in different situations. A

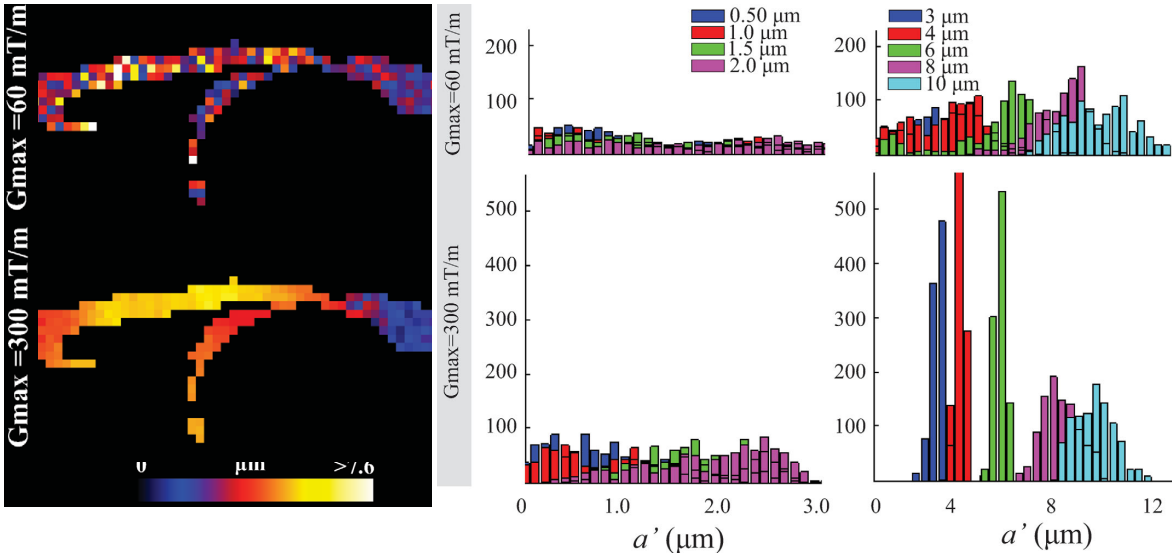


Figure adapted from: **Tim B. Dyrby** et al., *'Contrast and Stability of the Axon Diameter Index from Microstructure Imaging with Diffusion MRI.'*, MRM (2012) – with publisher's unrestricted permission

Figure 5.8: Shown on the left is the distribution of estimated axon diameters in fixed monkey brain, illustrating the benefit of higher gradients: 60 vs. 300 mT/m (top row and second row, respectively). While both gradients recover the CC trend of axonal thinkness, the higher gradients give a better contrast. By using simulations, the histograms on the right investigate the potential of estimating axon diameters, identifying 2.5-10 µm as a feasible range.

series of studies have aimed to identify the combination of compartments that best explain the diffusion MRI signal from WM over the accessible range of the measurement space. When aiming to describe diffusion MRI signal, many earlier studies have considered alternative hypotheses/models. For example, the study by Stanisz et al. [1997], mentioned in sub-section 5.2.1, compares two- and three-compartment models, favouring the latter, because the two-compartment model failed to capture adequately the non-restricted signal. Alexander et al. [2010] also compares two competing models for the fixed monkey data it used, explaining the preference of the fitting for a fourth Dot compartment in terms of the type of tissue used (i.e. the 'trapped' water). Ball-Sticks [Behrens et al., 2007], CHARMED [Assaf and Basser, 2005], and DIAMOND [Scherrer et al., 2013] have also had to make choices about the number of fibres supported by the data.

The first and most comprehensive model comparison framework was given by Panagiotaki et al. [2012]. The study provided a taxonomy of simple multi-compartment models for non-dispersive fibres, by collecting and enriching earlier models. It then compared the parsimonious performance of the models on data from fixed rats through a standard model comparison criterion, the BIC (see the next chapter for more), which balances the ability to explain data with model complexity. Richardson et al. [2013] repeated the study using non-fixed tissue in 'viable' *in vivo* state showing consistency of model ranking but inconsistency of parameter estimates with fixed tissue. Each study concludes that, to capture the broad trends in the signal, three compartments are required.

6

MODEL SELECTION

Model selection criteria aim to balance the model complexity with goodness-of-fit, to identify the simplest model that explains some data. A model that is too simple will under-fit the data and produce high bias (low accuracy); a model too complex will over-fit the data and produce high variance (low precision). All model selection techniques we introduce below aim at producing estimators which have low bias[1] and low variance. In practice, this is very hard to achieve, and such methods more commonly 'balance' bias with variance.

We start with methods which are based on the Kullback-Leibler (K-L) information, and then proceed to cross-validation and bootstrapping, and discuss briefly Model Averaging (which we will not be using in this work, but it will be later referred to).

6.1 BAYESIAN APPROACH TO MODEL COMPARISON

For a given model, the *Bayes'* theorem expresses its posterior odds (given the dataset) as proportional to the prior belief about the model (which is assumed to have generated the given dataset) and the likelihood (the probability of obtaining the data, given the model). This theorem forms the basis for comparing two competing models; specifically, the *Bayes factor* [Jeffreys, 1961] measures the ratio between two models' likelihoods; roughly, a factor above 3 is good evidence in favour of one model.

Computing the Bayes factor involves cumbersome integrals, but its approximation can be achieved through *Laplace* approximations, *Markov chain Monte Carlo* methods, or the BIC [Jeffreys, 1961; Wasserman, 2000; Burnham and Anderson, 2002].

Schwarz et al. [1978] defined the BIC as:

$$\mathrm{BIC} = -2\log(L(\hat{\theta}|data)) + K\log(N) \tag{6.1}$$

for a model of K parameters $\hat{\theta}$ which maximise the likelihood of obtaining the given N *independent and identically distributed* measurements. The lower the score, the more predictive the model is.

An advantage of the BIC/*Bayes'* factors (over traditional model comparison methods, such as likelihood ratio tests) is that they can be applied to non-nested models [Burnham and Anderson, 2002] (p.88).

1 For an estimator $\hat{\theta}$, the bias measures the deviation of its expectation from its true value θ, i.e. $E[\hat{\theta}] - \theta$, where E is the expectation; the variance measures the estimator's variability, i.e. $E[(E[\hat{\theta}] - \hat{\theta})^2]$.

6.2 INFORMATION-THEORETIC APPROACHES

FROM K-L INFORMATION TO AIC Kullback and Leibler [1951] introduced the notion of information lost[2] every time we approximate the reality through a model. If information, denoted by I, measures how far the θ-parameterised model g is from reality f, then

$$I(f, g) = \int f(x) \ln\left(\frac{f(x)}{g(x|\theta)}\right) dx \tag{6.2}$$

The aim then is to minimise this difference; in the model comparison language, the model with lower I is the better model.

Using the K-L information, Akaike [1974] produced a new index to minimise I in equation 6.2; this simple formulation combined parameter estimation with model comparison, and is:

$$AIC = -2\ln(L) + 2K \tag{6.3}$$

AIC assumes that the models are nested, and that the selected "best" model is the ideal/true model [Burnham and Anderson, 2002] (p.293). However, these assumption can be relaxed [Ripley, 2004]. One other assumption is that the fitting is by maximum likelihood methods (ibid).

Unlike AIC, BIC does take into account the size N of the dataset; on the other hand, BIC assumes that this "true" model is fixed as the dataset is increased, unlike in the case for AIC [Burnham and Anderson, 2002](p.301). In general, the AIC is less conservative in penalising complexity. This can be seen from the coefficient of the parameter variable K, a.k.a. the complexity penalising term. In eq. 6.1, the coefficient increases logarithmically in the number of measurements, but is fixed at 2 in the case of eq.6.3; more on this in section 6.4.

LOG-LIKELIHOOD WHEN THE ERRORS ARE NORMALLY DISTRIBUTED: When fitting models through a *least-squares* procedure, assuming that the errors ϵ_i across measurements i are normally distributed and independent, the error of prediction for measurement i is:

$$g(\epsilon_i|\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{1}{2}\left(\frac{\epsilon_i}{\sigma}\right)^2}$$

which means that, over all n measurements, the likelihood (of the parameters of the model, given the data) will be:

$$L(\theta|x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\epsilon_i}{\sigma}\right)^2}$$

Maximising the likelihood is equivalent to maximising the logarithm of the likelihood; this is equivalent to minimising the negative of the function:

$$\ln L(\theta|x) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\epsilon_i}{\sigma}\right)^2 \tag{6.4}$$

2 This measure is a generalisation of Shannon [2001] *entropy*, and similar to Boltzmann [1877] *entropy* H, related to probability $P = e^H$.

which, when we know σ, the first two terms become constants and, the log-likelihood then is:

$$\ln L(\theta|x) \sim -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\epsilon_i}{\sigma}\right)^2 \tag{6.5}$$

but when we do not know σ, and use its Maximum Likelihood Estimation (MLE),

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}{\epsilon_i}^2$$

then the non-constant term left is:

$$\ln L(\theta|x) \sim -\frac{n}{2}\ln(\hat{\sigma}^2) = -\frac{n}{2}\ln\left(\sum_{i=1}^{n}\frac{(\epsilon_i)^2}{n}\right) \tag{6.6}$$

6.3 OTHER TECHNIQUES

NON-PARAMETRIC BOOTSTRAPPING This technique for finding the variance of model parameters was introduced by Efron [1979]. For want of more distinct subjects' datasets, bootstrapping [Efron, 1979] sub-samples the original dataset repeatedly. Bootstrapping is said to provide 'optimistic' estimators, meaning the technique does not adequately punish over-fitting, because each dataset is used to both train and test the model.

CROSS-VALIDATION This method [Stone, 1974] is less prone to over-fitting than AIC/BIC: the model is trained on a part of the dataset, and the prediction error is evaluated on the missing data. When there is over-fitting, the model will fit the training data well, but then fail on the testing set.

However, selecting the size of the training/testing datasets can be tricky; the method is not completely immune to over- and under-fitting: decreasing the test set means higher variance. For example, to favour (i.e. increase) the test set, the original dataset can be split in half, i.e. use half of the dataset for training and the other half for testing. Though this may provide low variance on the estimators, it is rather wasteful of the data, and results in large bias. At the other end is the computationally intensive Leave-One-Out Cross-Validation (LOO-CV), which produces n datasets (of n-1 elements) on which to train the data, where n is the number of elements in the original dataset; this reduces the bias but increases the variance [Efron, 1983].

6.4 COMPARING THE METHODS

When the number of measurements is large, LOO-CV is not that different from, and is said to be asymptotic to, AIC and bootstrapping [Stone, 1977]. In common with AIC and bootstrapping, it is not efficient in penalising redundant model complexity. Shao [1993] shows that this is indeed the case for linear models, so increasing the data infinitely would not give certainty to the best predictive model (this can be seen in the AIC formulation vs. BIC, in eqns.6.3 and 6.1).

The recommendation is that coarser dataset splitting can be more efficient for model selection, that is leaving more than one element out for fitting. But, even then, to exhaustively try all combinations would mean fitting to $\binom{n}{k}$ datasets, where k is the number of elements left out of the original dataset. Often it is recommended [Kohavi, 1995; Diamantidis et al., 2000] to split the dataset randomly into 10 or 20 `folds`.

Both bootstrapping and cross-validation are used widely in error prediction and model selection, with some authors preferring one above the other. E.g. Kohavi [1995] justifies through examples why k-fold cross-validation is a better method for accuracy prediction and model selection than LOO-CV and bootstrapping. On the other hand, through a simple experiment with a small (14-element) dataset, Efron and Gong [1983] show that 10-fold cross-validation is also a low-bias high variance method, and that bootstrapping with preferential sampling of the elements provides better balance between bias and variance. Further, Efron and Tibshirani [1997] generally regard cross-validation as a low-bias high-variance method, and improves to .632+ the previous .632 method[3] to address this[4].

6.5 MODEL AVERAGING

We saw above that every technique for model comparison above has its benefits and drawbacks - largely dictated by how they balance bias with variance.

Model Selection can seem to be conservative in often assuming that one model from the set of candidate models is the true model, and then do parameter estimation. Other methods have sought to address this by combining homologous estimators in each model. In this linear combination, each element is weighted by the confidence in each model.

In *bootstrap smoothing*, or *bagging* [Efron and Tibshirani, 1996; Breiman, 1996], the standard errors are averaged across all the bootstrap datasets so as to ensure low variance. In AIC model averaging, the parameters are weighted by the normalised AIC score [Burnham and Anderson, 2002].

3 In the .632 method, 0.632*n is the expected number of distinct original dataset elements appearing on the unseen/testing set. When sampling with replacement, the probability of not selecting any data element after n times is $(1\text{-}1/n)^n$, with a limit $1/e \approx 0.368$ as n becomes large. Therefore, 0.632 is the accuracy weight for the models trained on the seen data but tested on the unseen data.

4 The .632+ essentially assigns weights to the error prediction on each, seen and unseen, observation. This weight, the *relative over-fitting rate* factor R, ranges from no over-fitting factor 0 to no-information factor 1.

Part III

MODEL SELECTION FRAMEWORK

7

INTRODUCTION

The following three experiments repeat the work of Panagiotaki et al. [2012].This study provides the taxonomy of models, which are then compared with fixed rat brain, using model selection criterion BIC. The difference here comes from the type of data used: we scan *in vivo* the human brain. This also means changes in the acquisition protocol.

In Experiment 1, similar to Panagiotaki et al. [2012], we acquire rich signal by probing the corpus callosum tissue fibres in three directions, once along and twice across the fibre, with many combinations of diffusion and gradient times and gradient magnitudes so as to fit the models to as wide a fraction of the measurement space as possible. The models are compared using BIC, and confirmed with the AIC (for their definition, see eqns. 6.3 and 6.1).

In Experiment 2, we examine the effect of increased angular sampling. We see that a multi-shell HARDI protocol is necessary to capture the complexity of the models.

In Experiment 3, we repeat Experiment 1, but this time using a richer, 45-direction multi-shell HARDI protocol. We add to the model comparison framework two new methods. The first is bootstrapping, which constructs datasets from the original one by sampling with replacement. The testing is then done on the training dataset. The second method is via four-fold cross-validation. This involves splitting the original datasets into four quarters, and keeping each quarter in turn for testing while training the models on the rest of the data. They both confirmed the ranking obtained by AIC/BIC, which was consistent throughout the datasets. The main finding from all three experiments was that, compared with the fixed tissue study [Panagiotaki et al., 2012], simpler three-compartment models emerge.

The work in Experiment 1 has previously been published as:

Ferizi U, Panagiotaki E, Schneider T, Wheeler-Kingshott CAM, Alexander DC: *White Matter Models of In Vivo Diffusion MRI Human Brain Data: A Statistical Ranking*. Proceeding of the 16th Conference on Medical Image Understanding and Analysis (MIUA), 2012

The work in Experiment 3 has previously been published as:

Ferizi U, Schneider T, Panagiotaki E, Nedjati-Gilani G, Zhang H, Wheeler-Kingshott CAM, Alexander DC: *A ranking of diffusion MRI compartment models with in vivo human brain data.* Magn Reson Med, 2013

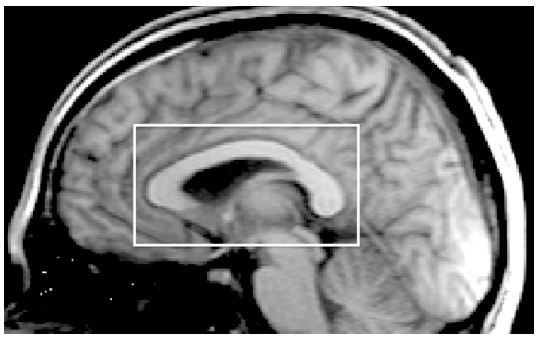
8

EXPERIMENT 1: FROM *EX VIVO* RATS TO *IN VIVO* HUMAN BRAIN

In this experiment we follow as closely as possible the experiment design used in Panagiotaki et al. [2012]. Any model extracted from this kind of work will ultimately be used for human brain diseases, so we will use *in vivo* data. This means having to adapt the scanning protocol to the lower gradient strengths available on the clinical MR scanners. Below is a description of this protocol and an outline of the preprocessing done to obtain a set of measurements for model fitting. We also make use of averaging across the most representative CC white matter voxels so as to enhance the signal SNR. Then follow details of the fitting procedure and the techniques used for model selection and ranking.

8.1 METHODS

DATA ACQUISITION: Using a 3T Philips scanner we scan a 30-year old healthy man, using single-shot EPI with cardiac gating. Three gradient-encoding directions are used: one along the corpus callosum main fibre direction and two in its perpendicular plane. The images consist of eight 4mm-thick sagittal slices, an image size of 64x64 and in-plane resolution of 2mm x 2mm.



fibre

Figure 8.1: This picture shows the scanned volume (boxed)

A PGSE sequence is used to probe:

- gradient strengths $|\mathbf{G}|$ = 30, 40, 50, 60 mT/m;
- pulse widths δ = {5, 15, 25} ms;
- 9 diffusion times Δ = {20, 30, 40, ..., 100} ms.

This produced a total of 63 diffusion weightings, excluding any combinations where $\delta > \Delta$, with maximum weight b=8,300s/mm^2. Because of the varying echo time TE, in addition to

every diffusion-weighted acquisition, a corresponding non-diffusion-weighted (b=0) image was obtained. Also, a separate HARDI acquisition with the same image resolution was performed, having 32-gradient encoding directions and b=711s/mm^2. The total acquisition time was 2.5hrs.

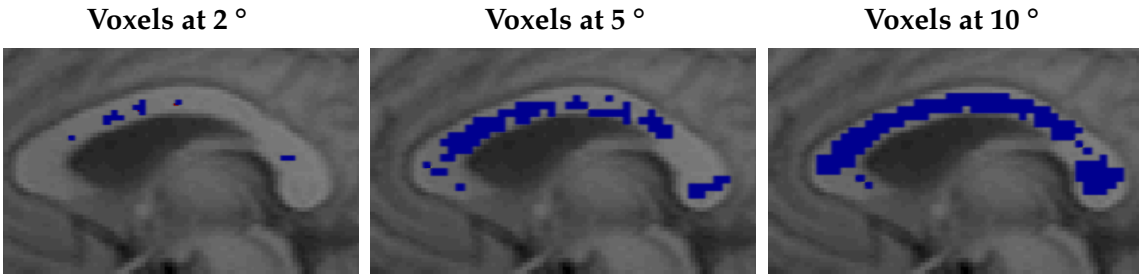


Figure 8.2: The voxels remaining after filtering. The signal of these voxels is then averaged to produce a single signal vector.

DATA PREPROCESSING: To find the best, most representative voxels in the CC, the DT was first fitted to the HARDI data. This is done to find the principal direction and identify those voxels with most coherent fibres. Then, we took voxels with an Fractional Anisotropy (FA) above a threshold, and principal eigenvector direction within a small angle of tolerance from the assumed fibre direction. These voxels were then averaged to give a single data set to which the model could be fitted. We set the FA threshold at 0.5 and repeated the experiment with three different angular thresholds, 2°, 5° and 10°, to establish the effect of orientation dispersion. The signal at each DW was normalised by the corresponding b=0 measurement with the same echo time, to remove the T_2 effects before the fitting. The signal at the noise floor, approximately any normalised signal below 0.1, is deleted from the dataset, and thus disregarded during the fitting. At the lowest TE, 22ms, the voxel-wise SNR was about 20.

MODEL DESCRIPTION: The compartments in Fig.8.3 are used to build the taxonomy of Panagiotaki et al. [2012]. The signal for a model with two or more types of compartments can be expressed as:

$$S = S_0 \left\{ \sum f_{ic}^k S_{ic}^k + f_{rc} S_{rc} + (1 - f_{rc} - \sum f_{ic}^k) S_{ec} \right\} \tag{8.1}$$

where f_{ic} is the weight of the intracellular signal compartment S^{ic}, f_{rc} is the weight of the isotropically restricted signal compartment S_{rc}, S_{ec} is the extracellular signal compartment, and k is the compartment index.

In this collection of models, the extracellular compartment, “hindered" in 3D, can be: a Tensor (full DT), a Zeppelin (cylindrically symmetric DT) or a Ball (isotropic DT). The intracellular compartment, “restricted" in 2D but free in the other direction (anisotropic restriction), can be: a Stick (a spatially oriented line), a Cylinder (a Stick with non-zero radius) or GDRcylinders (Cylinders with a Gamma distribution of radii; the distribution is characterised by shape

parameter κ and scale parameter θ, where $\kappa\theta$ is the distribution's mean, and $\kappa\theta^2$ gives its variance).

There are two special-case models. The first is the one-compartment Tensor, which is simply an ordinary DT. The second is a two-compartment model, the Bizeppelin, which combines two cylindrically symmetric Tensors (a 3D bi-exponential model).

In three-compartment models, the isotropically restricted third compartment can be: a Sphere (where diffusion is restricted to within a sphere of non-zero radius), a Dot (similar to a Sphere, but with a zero radius), Astrosticks (Sticks isotropically distributed in 3D), or Astrocylinders (Cylinders of a single non-zero radius isotropically distributed in 3D).

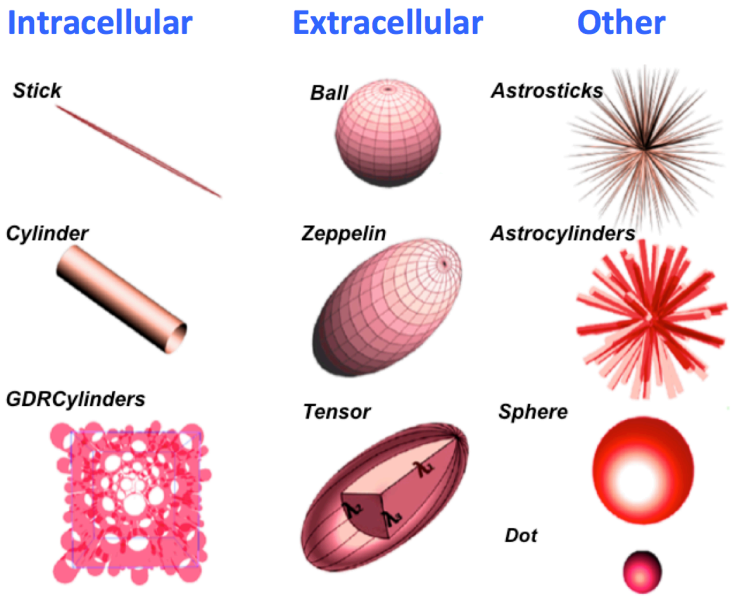


Figure 8.3: The taxonomy's model compartments, designed to capture intracellular diffusion (left), extracellular diffusion (middle) and diffusion in other media (right). Note that diffusion in the extracellular compartments is not restricted by the boundary, as in e.g. the Sphere, but rather hindered. (The figure is adapted from Panagiotaki [2011].)

The collection of models is shown by the network diagram of fig.8.4. Complexity between the families increases left-right, and within each family, there are two strands of three compartment models spun out of the main two-compartment parent: Dot/Sphere models on the one hand, and Astrosticks/Astrocylinders on the other.

MODEL FITTING: 32 models of this taxonomy were fitted to the signal, using the open software tool Camino [Cook et al., 2006]. The fitting uses the Levenberg-Marquardt algorithm with an offset-Gaussian noise model, minimising the objective function:

$$\text{LSE} = \min\left(\sum_{i=1}^{N} \frac{(\tilde{S}_i - \sqrt{S_i^2 + \sigma^2})^2}{\sigma^2}\right) \tag{8.2}$$

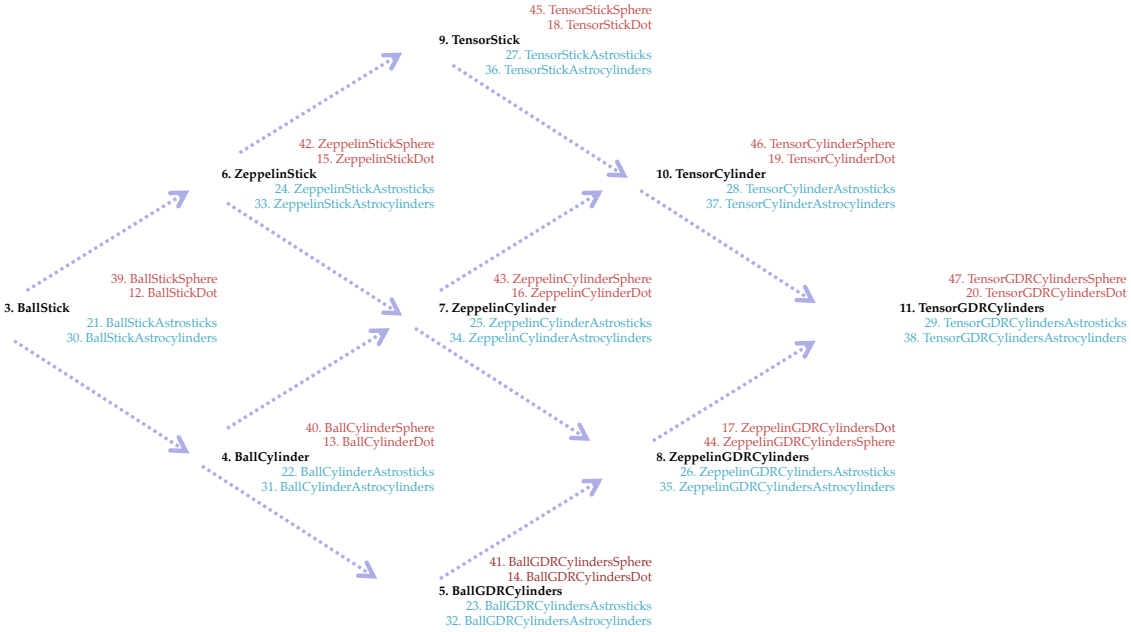


Figure 8.4: The models constructed from the compartments in Fig.8.3. The arrows show the relations between the models, increasing in complexity left-right. The models not shown are the one-compartment Tensor and two-compartment Bizeppelin.

where N is the number of measurements, $\tilde{S}_i$ is the i-th measured signal, and S_i its prediction from the model. σ=0.05 is the noise standard deviation, which we estimate a priori from the b=0 signals. This objective function accounts for bias introduced by the Rician noise inherent in the data in a simplistic way [Jones and Basser, 2004] that is more numerically stable than a full Rician log-likelihood objective function.

We fitted in two stages: after an initial run of 1,000 random starting points, we extract the parameters that produced the minimum objective function. We then execute another 1,000 runs from starting points at small random perturbations from the first minimum. This ensures that local minima are avoided, and the best fit parameters are obtained.

MODEL SELECTION: To compare the models, we use AIC and BIC, as defined in section 6.2.

As the noise level σ is found from the images, thus known *a priori* in the model fit, the first term of BIC (eq.6.1) and AIC (eq.6.3) equates to Least Squares Error (LSE) (eq.8.2), as shown via eq.6.5 (AIC criterion then becomes equivalent to Mallows [1973] C_p.)

8.2 RESULTS

Table 8.1 shows the ranking of the models. The BIC and AIC order and the quality-of-fit score LSE are given for each of the three datasets of three allowances for deviation from the main fibre direction, 2°, 5° and 10°. Most models fit the data similarly well, as seen in the LSE score. An exception is the single compartment model, DT, which is the worst. The most complex

		2 ° deviation			5 ° deviation			10 ° deviation		
Parameters	Models	BIC order	AIC order	LSE	BIC order	AIC order	LSE	BIC order	AIC order	LSE
6	Ball.Stick.Ast	**1**	1	149.05	**1**	1	94.54	**1**	1	74.41
6	Ball.Stick.Dot	**2**	6	151.64	**2**	6	96.44	**2**	6	76.61
7	Ball.Stick.Acl	**3**	2	149.05	**5**	2	94.29	**6**	2	73.73
7	Ball.Cylin.Acl	**4**	3	149.06	**3**	3	94.23	**4**	3	73.52
7	Ball.Cylin.Ast	**5**	4	149.06	**4**	4	94.28	**5**	4	73.61
7	Zepp.Stick.Ast	**6**	5	149.07	**6**	5	94.48	**3**	5	73.48
5	Ball.Stick	**7**	21	162.38	**7**	21	106.82	**9**	21	87.00
7	Zepp.Stick.Dot	**8**	10	151.64	**8**	10	95.97	**7**	10	74.83
7	Ball.Cylin.Dot	**9**	11	151.65	**9**	11	96.16	**8**	11	75.79
7	Ball.Stick.Sph	**10**	12	151.86	**10**	12	96.41	**10**	12	76.51
8	Zepp.Stick.Acl	**11**	7	149.08	**13**	7	94.24	**13**	7	72.99
8	Zepp.Cylin.Acl	**12**	8	149.09	**11**	8	94.19	**12**	8	72.89
8	Zepp.Cylin.Ast	**13**	9	149.14	**12**	9	94.23	**11**	9	72.89
6	Ball.Cylin	**14**	24	162.38	**15**	24	106.60	**17**	24	86.33
6	Zepp.Stick	**15**	25	162.39	**14**	25	105.80	**15**	25	85.44
8	Zepp.Cylin.Dot	**16**	14	151.64	**16**	14	95.72	**14**	14	74.21
8	Ball.Cylin.Sph	**17**	15	151.76	**18**	15	96.16	**18**	15	75.79
8	Zepp.Stick.Sph	**18**	16	151.77	**17**	16	95.92	**16**	16	74.62
9	Tens.Stick.Ast	**19**	13	149.09	**19**	13	94.46	**19**	13	73.43
9	Tens.Stick.Dot	**20**	17	150.96	**23**	17	96.27	**23**	17	76.30
7	Zepp.Cylin	**21**	27	162.39	**20**	27	105.57	**20**	27	84.85
7	Bizepp.	**22**	28	162.39	**21**	28	105.80	**22**	28	85.44
9	Zepp.Cylin.Sph	**23**	20	153.48	**22**	20	95.72	**21**	20	74.20
10	Tens.Cylin.Acl	**24**	18	149.09	**26**	18	94.21	**25**	18	72.97
10	Tens.Stick.Acl	**25**	19	149.14	**25**	19	94.16	**24**	19	72.79
8	Tens.Stick	**26**	30	162.41	**24**	30	105.21	**27**	30	84.68
10	Tens.Cylin.Dot	**27**	22	151.72	**28**	22	95.65	**28**	22	74.17
10	Tens.Stick.Sph	**28**	23	152.30	**29**	23	95.86	**29**	23	74.58
9	Tens.Cylin	**29**	31	162.56	**30**	31	105.09	**30**	31	84.48
10	Tens.Cylin.Ast	**30**	29	157.64	**27**	29	95.40	**26**	29	73.18
11	Tens.Cylin.Sph	**31**	26	152.29	**31**	26	95.69	**31**	26	74.17
7	Tensor	**32**	32	616.71	**32**	32	477.12	**32**	32	449.85

Table 8.1: A ranking of the models for the 2°/5°/10° fibre deviation allowances. Adjacent are the raw scores for their respective Objective Function Residue (*LSE*). The Ball-Stick-Astrosticks does best across both *BIC* and *AIC*, the DT worst. (Abbrev.: Tens.=Tensor; Zepp.=Zeppelin; Cylin.=Cylinder; Ast.=Astrosticks; Acl.=Astrocylinders; Sph.=Sphere).

model, Tensor-Cylinder-Astrocylinders, gives small LSE, but its complexity is penalised under AIC, and even more under BIC.

Ball, Zeppelin and Tensor models fit the data equally well, but the simpler Ball is penalised less by BIC. This is similarly the case for Stick being penalised less compared with Cylinder, and Sphere compared with Dot.

The fitting error LSE decreases markedly as the orientation threshold increases from 2° to 10°. The increased number of voxels being averaged smoothes the signal at the expense of noise in the data, and so the models fit increasingly better. Figure 8.5 illustrates the quality of fit of three selected models from the ranking, including the best and worst models. The model signal is obtained by synthesing signal with the parameters fitted to the 2° threshold data set. The plots reveal the limitations of the over-simplified single-compartment DT model in capturing both signal along and across the fibre. The DT model cannot capture the shape of

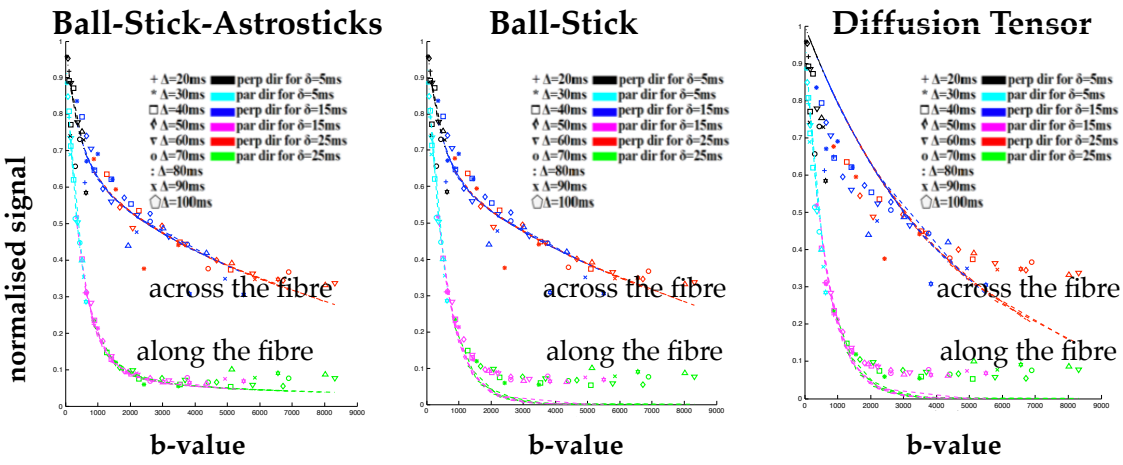


Figure 8.5: These plots illustrate the quality of fit of three models to the data. The model signal is shown as solid line against the signal shown in markers. It is clear that Ball-Stick fits better than DT the signal across the fibre. Even though Ball-Stick-Astrosticks fits the signal along the fibre better than Ball-Stick, the noise floor signal is disregarded in the fitting.

the perpendicular signal decay because it does not model restriction. The three-compartment model, Ball-Stick-Astrosticks, has a slight advantage compared with the two-compartment Ball-Stick fitting both the parallel and perpendicular direction signal. (However, this part of the normalised signal at the noise floor has no weight in the model fitting.)

8.3 DISCUSSION

This experiment was a first step in the translation of the model comparison framework of Panagiotaki et al. [2012] from *ex vivo* rat to *in vivo* human data. We hoped that a long 2.5h protocol would provide data of sufficient richness, and hence provide a reasonable fit of the models.

The increasing fibre incoherence threshold (2°, 5° and 10°) did predictably increase the amount of voxels filtered through which, in turn, raised the SNR of the dataset; this is reflected in the LSE score. However, this made no noticeable difference to the model ranking or differentiation of the models.

The ranking was not very informative as it could not discriminate well between Ball, Zeppelin and Tensor models. One potential and obvious problem could lie with the data used. The previous framework of Panagiotaki et al. [2012] made full use of the strength of animal scanner gradients of up to 1000 mT/m, and an unmatchable length of scanning time of 65h. In the following experiment we look for ways of improving the current *in vivo* human imaging protocol.

9

EXPERIMENT 2: MORE DIFFUSION-SENSITISING DIRECTIONS?

This modest experiment was performed in order to explore the necessity for enriching the protocol of the previous Experiment 1, in chapter 8, with more gradient directions. In addition to the data acquired for that study, we acquired two more HARDI shells; the difference here is that we include the Diffusion Tensor Imaging (DTI)/HARDI measurements in the model fitting analysis. We observe the effect on the model fitting after adding incrementally to the original dataset one half of each HARDI shell. We restrict the number of models used here; from the previous ranking, we chose only three models which had a similar quality of fit but different enough on the complexity spectrum.

We describe below the datasets, and afterwards give the results obtained from the fitting of the models.

9.1 METHODS

The acquisition is an extension to that described in section 8.1. Briefly, in that experiment we had DW measurements contain 12 shells with b_{max}=8,300 s/mm^2, probing the tissue in three mutually orthogonal directions, and a 32-direction HARDI shell with b=711 s/mm^2. Additionally, this scanning includes another 32-direction HARDI shell, with b=2855 s/mm^2.

We construct seven datasets to which the models are fit:

- $0-0$ has the original 2° dataset of DW images, as described in Experiment 1; its preprocessing involved retaining only voxels with principal direction orientation within 2° of the perpendicular to the sagittal slice;
- $0-16$ comprises dataset $0-0$ and half of the first b=711 s/mm^2 HARDI shell (directions uniformly distributed on the sphere);
- $16-0$ as above, but to $0-0$ adds instead half of the second HARDI shell;
- $16-16$ contains $0-0$ and half of the both HARDI shells;
- $16-32$ contains $0-0$, half of the first HARDI shell and the full second HARDI shell;
- $32-16$ contains $0-0$, the full first HARDI shell and half of the second HARDI shell;
- $32-16$ contains $0-0$ and whole HARDI shells.

MODEL FITTING: Here we choose only three models: Ball-Stick-Astrosticks, Bizeppelin and Tensor-Cylinder-Astrocylinders. They were fitted to the data 1000 times, as in the previous experiment, minimising for the objective function LSE of eq.8.2.

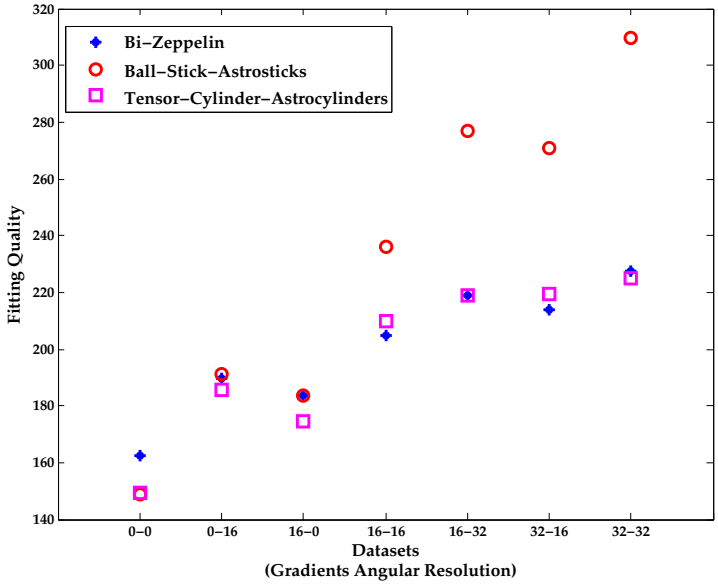


Figure 9.1: The effect on the model fitting of increased angular resolution. The three-direction 2° dataset. The Ball-Stick-Astrosticks model fits the data well; DT fits poorly.

9.2 RESULTS

Figure 9.1 shows the difference between the models in fitting to the data with ever-increasing angular resolution. The first observation is that, for this range of datasets, the models with an anisotropic extracellular compartment, that is Bizeppelin and Tensor-Cylinder-Astrocylinders, show similar behaviour: the initial increase in the fitting error almost reaches a plateau when both halves of the HARDI shells are added. While the fitting error increases as the number of measurements also increases, the rate of change with respect to the number of measurements is considerably reduced beyond this point. The picture is slightly different for Ball-Stick-Astrosticks. Its fitting error has an upward trend as more gradient directions are added. One would expect that this increase would also slow down if even more directions were added.

9.3 DISCUSSION

The point of this experiment was to show a reason for the insufficiency of Experiment 1 in distinguishing between the models. The results from Table 8.1 appeared to cluster most of the models together, even though they were geometrically very different. As an example, we saw an almost identical quality of fit for Tensor-Cylinder-Astrocylinders and Ball-Stick-Astrosticks, followed by Bizeppelin; however, they were ranked at number 1, 24 and 22, respectively, by the BIC criterion. Therefore, for a better comparison of models, it is necessary to amend the original experiment design; we showed that including more gradient directions to the acquisition protocol is potentially beneficial to the testing of the models. We do this in the next chapter, Experiment 3.

10

EXPERIMENT 3: INCREASING ANGULAR RESOLUTION

This study is motivated by Experiment 2. There we saw that, in order to differentiate better between the models, a higher angular resolution of diffusion-sensitising gradients in the signal sampling was beneficial.

We additionally improve the dataset by stretching as much as possible the scanning time, so as to make the protocol richer for even the most complex models. Specifically, this protocol uses a rich, massively multi-shell HARDI protocol, to probe a wide range of gradient orientations, diffusion times, gradient pulse times, and gradient magnitudes. The model comparison framework is extended by using bootstrapping and cross-validation, to provide additional insight into the stability and accuracy of the model ranking. Additionally, we acquire the data twice, to check not only for inter-subject reproducibility of the results, but also for the impact of splitting the scanning into multiple sessions.

The richness of our dataset allows us to assess the model fitting stability through bootstrapping on the dataset.

This chapter starts by describing the acquisition protocol for the data, and the preprocessing done to the measurements. Then follow the details of the fitting procedure, the technique used for comparing the models and, lastly, an evaluation of the robustness of the ranking.

10.1 METHODS

DATA ACQUISITION: The central aim in this acquisition is to cover as large a portion of the measurement space as possible, while retaining a usable signal-to-noise level. The full protocol, henceforth often referred to as the *Achieva+* protocol, has 32 shells of 45-directions each. To enhance overall angular resolution, the set of directions in each shell is a unique random rotation of the 45-direction Camino [Cook et al., 2006] point set (i.e. vector directions, or points, isotropically spread on a sphere, following optimisation for DW imaging). As shown in Table 10.1, each shell has a unique combination of:

- gradient strength $|\mathbf{G}| = \{55, 60\}$ mT/m;
- pulse width $\delta = \{6, 10, 15, 22\}$ ms;
- pulse duration $\Delta = \{30, 50, 70, 90\}$ ms.

Within each shell there are three interwoven b=0 acquisitions. The b-values thus range from 218 to 10,308 s/mm^2, with effective diffusion time (Δ-$\delta/3$) in the range 28 to 82 ms. We use a PGSE sequence on a 3T Philips scanner, with cardiac gating and repetition time TR = 4s. The Sensitivity Encoding (SENSE) factor was 1.10. The echo time TE varies between shells depending

Achieva+ Protocol

δ=6ms					δ=10ms					δ=15ms					δ=22ms				
Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)	Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)	Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)	Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)
1	30	51	55	218	9	30	55	55	577	17	30	60	55	1218	25	30	71	55	2375
2	30	51	60	260	10	30	55	60	687	18	30	60	60	1449	26	30	71	60	2826
3	50	71	55	374	11	50	75	55	1010	19	50	80	55	2192	27	50	87	55	4470
4	50	71	60	445	12	50	75	60	1202	20	50	80	60	2608	28	50	87	60	5320
5	70	91	55	530	13	70	95	55	1443	21	70	100	55	3166	29	70	107	55	6566
6	70	91	60	631	14	70	95	60	1718	22	70	100	60	3768	30	70	107	60	7814
7	90	111	55	686	15	90	115	55	1876	23	90	120	55	4140	31	90	127	55	8661
8	90	111	60	816	16	90	115	60	2233	24	90	120	60	4927	32	90	127	60	10308

Table 10.1: The scanning protocol used. This Achieva+ protocol required two sessions of 4.5h acquisition.

on the values of δ and Δ and is kept to a minimum to maximise signal. There are nine 4 mm thick sagittal slices, acquired with a reduced Field-of-View (FOV) using a ZOnally-magnified Oblique Multi-slice (ZOOM)-EPI technique with outer volume supression [Wilm et al., 2007]. The FOV is centred on the mid-sagittal slice of the CC, where we assume that coherently oriented CC fibres are perpendicular to the image plane. The image size is 64x64 and the in-plane resolution 2x2 mm^2.

The study was approved by the local ethics committee, and written informed consent was obtained from the participant. We acquire the full protocol in a 31-year-old healthy subject in two different ways.

The first full data set is acquired in two separate non-stop sessions, each lasting about 4h 30min; we refer to this as the 2x4h data set. We used the *dynamic stabilisation* facility provided by the scanner, which is designed for long scans to correct for field drifts during the image acquisition. We visually inspected the images and did not observe any obvious shifts from gradient heating. The SNR of $b = 0$ images varied from about 30 at TE_{min}=51ms to 5 for TE_{max}=127ms.

We check for intra-subject reproducibility, we then repeat the protocol in eight sessions, each lasting 1h 15min; we call this the 8x1h data set.

PREPROCESSING / VOXEL SELECTION: We carefully registered the sagittal slices, making in-plane corrections usually in the order of 1 to 3 voxels. The quality of registration is confirmed visually for each individual image. All non-diffusion-weighted images are registered to the first unweighted image of the b=1,202 s/mm^2 shell; the corresponding transformations are then applied to the fifteen DW images that follow each b=0 acquisition, as ordered in the scanning protocol. In this b=1,202s/mm^2 reference shell, we manually segment the subject's image of corpus callosum, and then fit the DT to select a set of voxels with coherently oriented fibres. In particular, all voxels with FA>0.6 and principal eigenvector within η=2° of the assumed fibre direction (perpendicular to the image plane, i.e. left-right in the brain) are retained. In the 2x4h data set, there are 24 voxels that satisfy the imposed criteria, all belonging to the 2 slices closest to the mid-sagittal plane. A similar procedure with η=5° leaves 66

voxels, and η=10° which leaves 99 voxels. In the 8x1h data set, 60, 101 and 166 voxels remain, respectively, sampling the corpus callosum rather more evenly; the same thresholding procedure leads to a slightly different set of voxels because of noise, misalignments, etc. To account for different TE affecting different shells, the signal in each shell is normalised by the average of the three unweighted measurements (b=0) with the same TE.

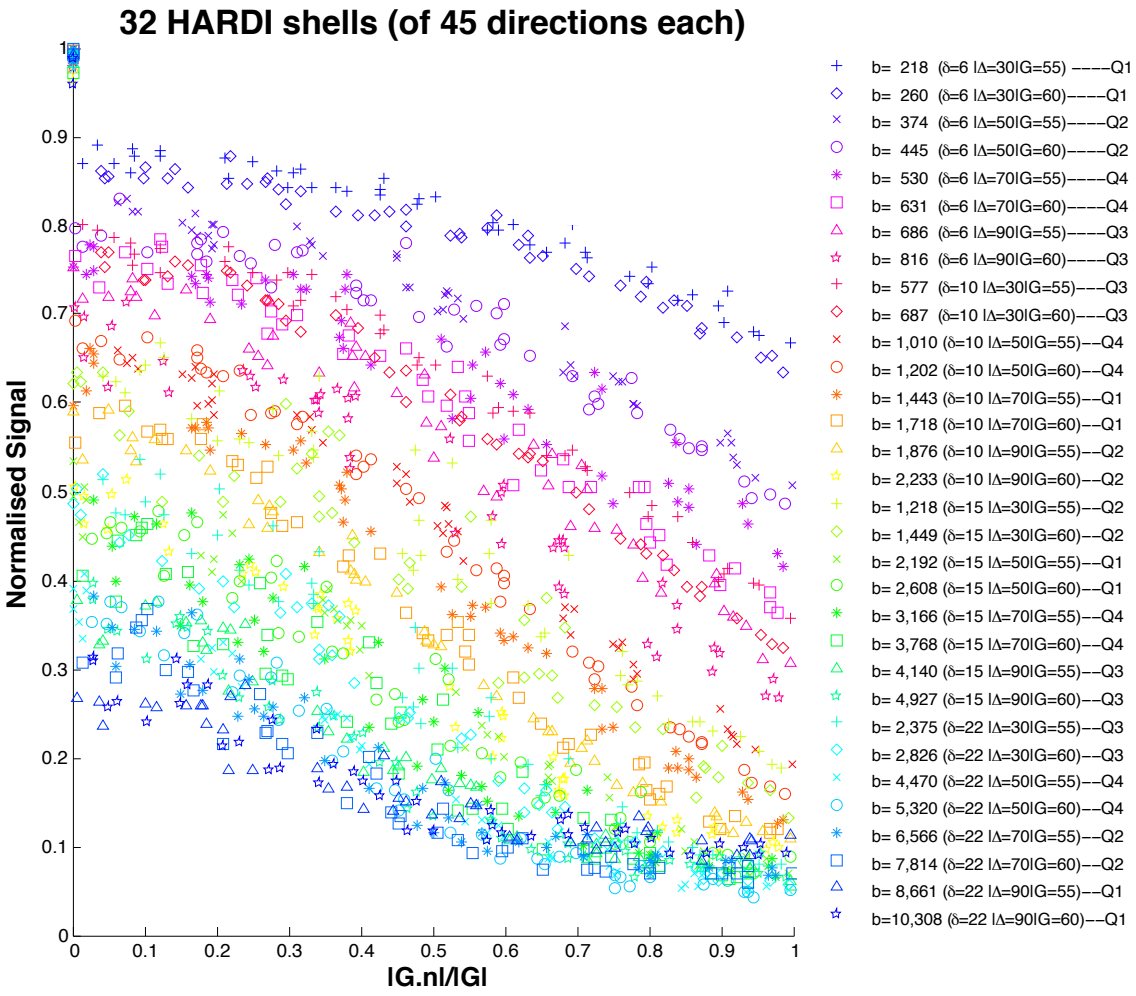


Figure 10.1: The acquired signal for the 2x4h 2° data set. The legend gives b-value ($\delta \mid \Delta \mid |\mathbf{G}|$) in units of s/mm^2(ms|ms|mT/m); Q1-Q4 on the right define the four quarters of the full protocol used in the four-fold cross-validation. **G** is the applied gradient vector and **n** is the fibre direction; the x-axis gives the absolute value of the cosine of the angle between the applied gradient and fibre direction: to the left, the gradient is perpendicular to the fibres; to the right, parallel; the shells' b-value increases going down from the top.

As before, we create a single data set for each η=2°, 5° and 10° by averaging over the voxels selected above. Figure 10.1 shows the signal from the 2x4h data set with η=2° and confirms the rich coverage of the measurement space the protocol provides. The datasets contain 1,356=32*(3+45) measurements each.

MODEL FITTING AND SELECTION: The fitting is as described in the section 8.1. Each model is fitted 250 times, and the final parameters are those that produce the minimum objective function LSE (Eq.8.2).

The criterion BIC (eq.6.1) is chosen to compare the models.

BOOTSTRAPPING: We use classical bootstrap [Efron, 1979] to analyse the stability of the BIC ranking. Each bootstrap data set comes from a random selection in each shell of the same number of data points, with replacement. For each 2x4h and 8x1h data set, we construct 100 bootstrap datasets. We then obtain 100 BIC rankings after fitting the models 50 times to each data set and picking the best parameter estimates. We construct positional variance diagrams, which give the number of times (i.e. bootstrap data sets) out of 100 that each model appears in each position in the ranking.

CROSS-VALIDATION: PREDICTING UNSEEN DATA: Cross-validation provides a complementary model selection to confirm the findings from the BIC. We use four-fold cross-validation and divide the data set into four quarters. Each quarter is constructed by dividing all the shells of each δ into two groups of low Δ (30 and 50 ms) and high Δ (70 and 90 ms). Then, we randomly assign one from each group to each quarter; shells with $|\mathbf{G}| = \{55, 60\}$ mT/m go together.

The cross-validation then proceeds as follows: we divide the data into four quarters, by randomly assigning low and high Δs into four groups. Then, we choose signal coming from three-quarters of the dataset to fit our models to and, from the parameter estimates drawn from these quarters, synthesise signal for the missing part. Next, we evaluate the sum of squared differences LSE compared to that unseen quarter. This provides an alternative model selection routine, to confirm and validate the ranking by BIC.

10.2 RESULTS AND DISCUSSION

Table 10.2 gives the complete model rankings and some parameter estimates across different data sets, 2x4h and 8x1h, and different η. Several distinct groups of models emerge: i) three-compartment models with anisotropic extracellular compartment (Zeppelin/Tensor) and Dot/Sphere third compartment, which produce the best fit (and lowest BIC); ii) three-compartment models with anisotropic extracellular compartment and Astrostick/Astrocylinder third compartment, which are consistently worse than Dot/Sphere equivalents, but better than all other models; iii) three-compartment models with isotropic extracellular compartment and all two-compartment models. The performance boundaries between the groups are very clear. The DT comes below group (iii). The CC voxels selected for averaging are different in each data set, producing some variation. In particular, the axon radius index (shown in the appendix) is higher in the 8x1h data set, which we expect because it has a greater contribu-

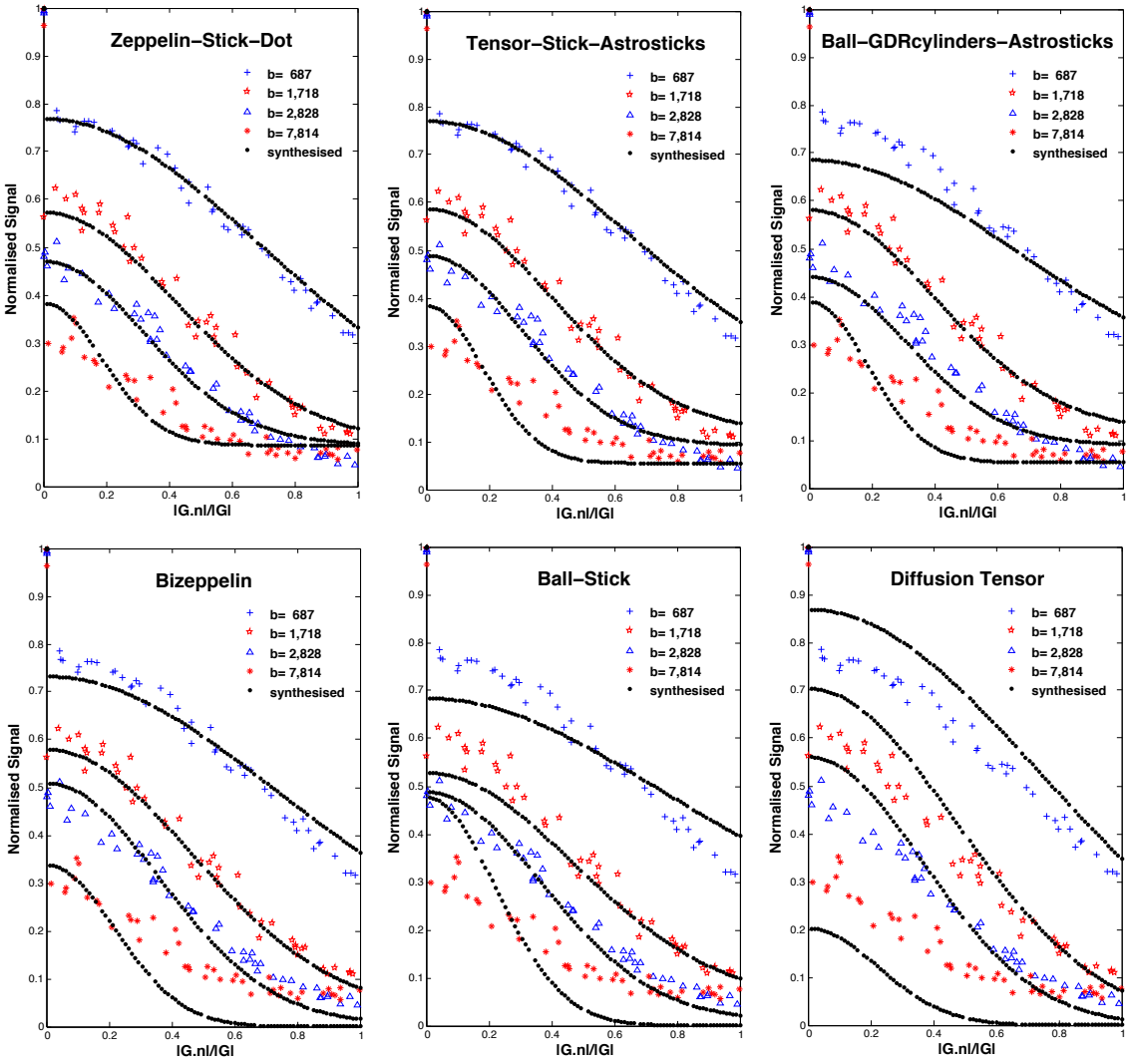


Figure 10.2: The model signal, shown as dotted line, is superimposed on raw data, marked with red/blue colours; for clarity, only 6 representative models are chosen, and only 4 shells are shown across the sampled range of b-values.

tion from the midbody where axons are larger. However, the estimates obtained from multiple sessions (8x1h data) are broadly in line with those of the 2-session data (2x4h data).

As η increases, the LSE would go down because the number of voxels being averaged increases, which increases the SNR. We see a slight increase in the radius estimate and decrease in axial diffusivity as dispersion increases, but the effects are minor.

The parameter estimates show strong consistency within the groups but more variation between groups. In group (i), the intracellular volume fraction is unexpectedly low and about half of the extracellular volume fraction. One possible explanation is a significant free water contribution [Assaf et al., 2008; Barazany et al., 2009] which we do not model explicitly, and so gets absorbed in the extracellular component. Significant within-voxel fibre dispersion [Zhang et al., 2011, 2012; Nilsson et al., 2012] could also cause this observation, as group (ii), which to

some extent model fibre dispersion, show higher intracellular and third compartment volume fractions.

	BIC						Stick/Cyl. Volume Fraction						Axial Diff. ($x10^{-9}m^2/s$)						Cylinder Diameter ($x10^{-6}m$)					
Models	*2x4_2*	*2x4_5*	*2x4_10*	*8x1_2*	*8x1_5*	*8x1_10*	*2x4_2*	*2x4_5*	*2x4_10*	*8x1_2*	*8x1_5*	*8x1_10*	*2x4_2*	*2x4_5*	*2x4_10*	*8x1_2*	*8x1_5*	*8x1_10*	*2x4_2*	*2x4_5*	*2x4_10*	*8x1_2*	*8x1_5*	*8x1_10*
ZeppelinStickDot	813	626	599	877	881	802	0.29	0.30	0.29	0.30	0.29	0.27	1.91	1.85	1.82	1.81	1.77	1.75						
TensorStickDot	814	628	605	853	859	796	0.29	0.29	0.29	0.29	0.28	0.27	1.91	1.85	1.82	1.81	1.78	1.75						
ZeppelinCylinderDot	820	633	605	875	856	788	0.30	0.30	0.30	0.35	0.38	0.35	1.91	1.85	1.82	1.81	1.77	1.75	4.4	5.4	6.9	9.3	11.0	10.5
ZeppelinStickSphere	820	634	606	884	883	808	0.29	0.30	0.29	0.30	0.29	0.28	1.91	1.85	1.82	1.81	1.80	1.77						
TensorCylinderDot	821	635	611	852	836	783	0.29	0.30	0.30	0.34	0.37	0.35	1.91	1.85	1.82	1.81	1.78	1.76	4.4	5.3	6.8	9.2	10.9	10.4
TensorStickSphere	821	635	612	860	861	802	0.29	0.29	0.29	0.29	0.28	0.27	1.91	1.85	1.82	1.81	1.81	1.77						
ZeppelinCylinderSphere	827	641	612	882	863	796	0.30	0.30	0.30	0.35	0.38	0.35	1.91	1.85	1.82	1.81	1.77	1.75	4.4	5.4	6.9	9.3	11.0	10.5
ZeppelinGDRCylindersDot	827	641	613	882	863	796	0.30	0.30	0.30	0.35	0.38	0.35	1.91	1.85	1.82	1.81	1.77	1.75	4.2	5.2	6.8	9.4	11.2	10.7
TensorCylinderSphere	828	642	618	859	843	791	0.29	0.30	0.30	0.34	0.37	0.35	1.91	1.85	1.82	1.81	1.78	1.76	4.3	5.3	6.8	9.2	10.9	10.4
TensorGDRCylindersDot	828	642	618	859	843	791	0.29	0.30	0.30	0.34	0.38	0.35	1.91	1.85	1.82	1.81	1.78	1.76	4.1	5.1	6.7	9.2	11.1	10.6
ZeppelinGDRCylindersSphere	835	648	620	889	870	803	0.30	0.30	0.30	0.35	0.38	0.35	1.91	1.85	1.82	1.81	1.77	1.75	4.3	5.2	6.8	9.4	11.2	10.7
TensorGDRCylindersSphere	836	650	626	867	851	798	0.29	0.30	0.30	0.34	0.38	0.35	1.91	1.85	1.82	1.81	1.78	1.76	4.1	5.1	6.7	9.2	11.1	10.6
ZeppelinStickAstrosticks	961	792	767	1046	1053	987	0.33	0.33	0.33	0.34	0.32	0.31	2.06	1.97	1.94	1.90	1.89	1.87						
TensorStickAstrosticks	961	793	772	1021	1029	981	0.33	0.33	0.32	0.33	0.32	0.31	2.06	1.98	1.94	1.91	1.91	1.88						
ZeppelinCylinderAstrosticks	968	799	773	1045	1029	975	0.33	0.34	0.34	0.39	0.42	0.39	2.06	1.97	1.94	1.89	1.88	1.87	4.0	5.0	6.7	9.1	10.7	10.2
ZeppelinStickAstrocyl.	968	800	775	1054	1060	995	0.33	0.33	0.33	0.34	0.32	0.31	2.06	1.97	1.94	1.90	1.89	1.87						
ZeppelinCylinderAstrocyl.	968	800	775	1054	1052	992	0.33	0.33	0.33	0.34	0.36	0.33	2.06	1.97	1.94	1.89	1.89	1.87	0.2	0.2	0.5	4.5	8.1	6.9
TensorCylinderAstrosticks	968	800	779	1021	1009	970	0.33	0.33	0.34	0.37	0.41	0.38	2.06	1.98	1.94	1.91	1.90	1.87	3.8	4.9	6.6	8.9	10.6	10.1
TensorStickAstrocyl.	969	801	780	1029	1037	988	0.33	0.33	0.32	0.33	0.32	0.31	2.06	1.98	1.94	1.91	1.91	1.88						
TensorCylinderAstrocyl.	969	801	780	1029	1029	986	0.33	0.33	0.32	0.33	0.35	0.33	2.06	1.98	1.94	1.91	1.90	1.88	0.3	0.2	0.4	4.4	8.0	6.9
ZeppelinGDRCylindersAstrost.	975	807	781	1052	1037	982	0.33	0.34	0.34	0.39	0.42	0.39	2.06	1.97	1.94	1.89	1.88	1.87	3.8	5.1	6.6	9.1	10.9	10.3
ZeppelinGDRCylindersAstrocyl.	975	807	782	1061	1060	999	0.33	0.34	0.33	0.34	0.36	0.33	2.06	1.97	1.94	1.89	1.89	1.87	0.6	0.2	0.2	4.5	8.0	6.9
TensorGDRCylindersAstrost.	976	808	786	1029	1016	978	0.33	0.33	0.34	0.37	0.41	0.38	2.06	1.98	1.94	1.91	1.90	1.87	3.6	4.7	6.5	8.9	10.8	10.2
TensorGDRCylindersAstrocyl.	976	808	787	1036	1037	993	0.33	0.33	0.32	0.33	0.35	0.33	2.06	1.98	1.94	1.91	1.90	1.88	0.2	0.2	0.5	4.5	8.0	6.8
Bizeppelin	1079	954	941	1222	1265	1223	0.63	0.60	0.59	0.54	0.54	0.54	1.47	1.42	1.39	1.41	1.35	1.32						
BallGDRCylindersDot	1135	940	894	1162	1063	1002	0.66	0.67	0.68	0.71	0.72	0.70	1.77	1.72	1.68	1.66	1.65	1.63	19.4	19.9	20.2	20.2	20.2	20.2
BallGDRCylindersAstrosticks	1139	965	929	1206	1124	1072	0.56	0.57	0.58	0.64	0.64	0.62	1.99	1.92	1.88	1.86	1.85	1.83	16.3	16.6	17.2	18.4	19.1	19.1
BallGDRCylindersSphere	1142	947	902	1169	1071	1010	0.66	0.67	0.68	0.71	0.72	0.70	1.77	1.71	1.68	1.66	1.65	1.63	19.4	19.8	20.2	20.2	20.2	20.2
BallCylinderAstrosticks	1151	983	948	1232	1143	1101	0.53	0.54	0.54	0.58	0.59	0.56	2.01	1.93	1.89	1.86	1.85	1.83	12.6	12.6	12.8	13.2	13.5	13.4
ZeppelinStick	1177	1016	997	1249	1305	1254	0.40	0.41	0.40	0.41	0.40	0.40	1.49	1.43	1.40	1.41	1.36	1.33						
TensorStick	1179	1019	1004	1229	1286	1250	0.40	0.41	0.40	0.41	0.40	0.40	1.49	1.43	1.40	1.41	1.36	1.33						
BallCylinderDot	1179	995	951	1222	1113	1063	0.60	0.60	0.61	0.64	0.65	0.63	1.75	1.69	1.66	1.65	1.64	1.62	13.4	13.4	13.6	13.9	14.3	14.3
ZeppelinCylinder	1184	1022	1001	1242	1269	1232	0.41	0.42	0.43	0.47	0.51	0.48	1.49	1.43	1.40	1.41	1.36	1.33	4.8	5.6	6.7	8.5	9.8	9.3
TensorCylinder	1186	1025	1008	1223	1253	1229	0.41	0.42	0.43	0.46	0.50	0.48	1.49	1.43	1.40	1.41	1.36	1.33	4.7	5.5	6.6	8.4	9.7	9.2
BallCylinderSphere	1186	1002	958	1229	1120	1070	0.60	0.60	0.61	0.64	0.65	0.63	1.75	1.69	1.66	1.65	1.64	1.62	13.4	13.4	13.6	13.9	14.3	14.3
ZeppelinGDRCylinders	1191	1029	1008	1249	1276	1239	0.41	0.42	0.43	0.47	0.51	0.49	1.49	1.43	1.40	1.41	1.36	1.33	4.7	5.5	6.7	8.6	10.0	9.4
TensorGDRCylinders	1193	1033	1016	1230	1261	1237	0.41	0.42	0.43	0.46	0.50	0.48	1.49	1.43	1.40	1.41	1.36	1.33	4.6	5.4	6.6	8.4	9.8	9.4
BallCylinderAstrocylinders	1231	1064	1036	1319	1239	1197	0.50	0.50	0.51	0.55	0.54	0.52	1.95	1.88	1.83	1.79	1.80	1.77	11.0	11.0	11.2	11.8	12.0	11.8
BallGDRCylindersAstrocyl.	1240	1073	1045	1328	1249	1206	0.50	0.50	0.51	0.55	0.54	0.52	1.95	1.87	1.83	1.78	1.80	1.77	11.1	11.1	11.3	12.0	12.1	11.9
BallGDRCylinders	1362	1190	1153	1396	1357	1319	0.65	0.66	0.67	0.71	0.73	0.70	1.44	1.38	1.36	1.37	1.33	1.30	14.8	14.8	15.3	16.3	17.0	16.7
BallStickAstrocylinders	1387	1225	1217	1568	1504	1429	0.41	0.41	0.41	0.42	0.41	0.40	1.93	1.85	1.81	1.80	1.85	1.79						
BallStickAstrosticks	1388	1228	1227	1613	1580	1483	0.41	0.41	0.41	0.43	0.41	0.40	1.86	1.78	1.72	1.64	1.66	1.62						
BallCylinder	1389	1220	1184	1428	1386	1353	0.61	0.62	0.62	0.66	0.67	0.65	1.43	1.38	1.35	1.36	1.32	1.29	11.3	11.2	11.3	11.8	12.1	11.9
BallStickSphere	1507	1330	1312	1638	1567	1489	0.43	0.43	0.43	0.44	0.43	0.41	1.61	1.55	1.53	1.54	1.58	1.53						
BallStickDot	1536	1360	1351	1719	1710	1594	0.44	0.44	0.43	0.45	0.44	0.43	1.49	1.44	1.40	1.36	1.34	1.33						
BallStick	1584	1415	1404	1743	1761	1662	0.48	0.48	0.47	0.48	0.48	0.47	1.35	1.29	1.26	1.26	1.22	1.19						
DT	2123	2001	1984	2208	2206	2247	0.85	0.84	0.84	0.85	0.84	0.83	1.54	1.48	1.45	1.46	1.40	1.38						

Table 10.2: Various model parameters from different data sets, 2x4h and 8x1h, with different angular thresholds, 2°, 5° and 10°. Regarding the BIC, Raftery [1996] regard the preference for one model against another as 'weak' when their BIC difference is 0-2, 'positive' for 2-6, 'strong' for 6-10, and as 'very strong' for anything above 10.

Table 10.2 gives a more complete set of parameter estimates for the 2° data set of 2x4h. In hindered compartments, the axial diffusivities in groups (i) and (ii) are consistently around $2x10^{-9}m^2/s$, and the radial diffusivities are around $0.7x10^{-9}m^2/s$, in agreement with previous reports [Beaulieu, 2002]. The two radial diffusivities of Tensor models are close, making Tensor and Zeppelin models similar, as we might expect for coherently oriented fibres in the CC, causing the BIC generally to prefer the simpler Zeppelin models.

Looking to the left of table 10.2, within the groups, as expected, the quality-of-fit LSE consistently reduces as the complexity increases, and the BIC ranking rewards simpler compartments, but there is little to choose between the models in group (i).

Cylinder models in group (i) consistently provide axon diameter index values of around 5 μm, which is consistent with axon diameter estimates from the CC in Assaf et al. [2008] and Dyrby et al. [2013]. Other models show more erratic estimates of radius which arise because the models fit the data less well, and so use the parameter to explain effects they do not capture. The GDR Cylinder models' shape parameter κ often hits the upper bound constrained in the fitting to 10. At this value of κ, the Gamma distribution is close to Gaussian shape and is highly peaked about the mean, making the GDR Cylinder model very similar to the Cylinder model. BIC thus prefers the simpler Cylinder model. The Sphere and Astrocylinder radius estimate is usually around 0.1 μm, which makes them very similar to the simpler Dot and Astrosticks models, respectively, which the BIC generally prefers.

Fig.10.2 illustrates the fit of some of the models to the data. The models are fitted to a total of 32 shells, but we select only four to illustrate visually where the models over/under-estimate the signal. While the fitting is not perfect even for the best model of the ranking, the figure reflects clearly the model ranking in the signal prediction.

Fig.10.3 shows on the right the positional variance diagrams of model ranking over 100 bootstrap samples from both the 2x4h and 8x1h η=2° data sets. The group structure of the ranking is very consistent over the bootstraps, although we see some variance of model positions within the groups; the ranking is also consistent between the 2x4h and 8x1h data sets, though some difference is expected, arising from minor imperfections in the registration of images in such large data sets. The group structure is also similar for the η=5° and 10° data sets (results not shown). Differences in the number of voxels averaged in these datasets has little effect on the rankings. To the left of Fig.10.3 we show results from cross-validation. The same group structure emerges with, on average, group (i) performing best, followed by group (ii), and more erratic performance in group (iii). Little distinguishes models within group (i) or group (ii).

LSE	S_0	Nr Parameters	Models	Vol. Fraction	κ (Gamma Dist.)	Diameter / x10^{-6} m	Vol. Fraction	Axial Diff. / x10^{-9} m^2/s	Radial Diff.1 / x10^{-9} m^2/s	Radial Diff.2 / x10^{-9} m^2/s	Theta/ arc degree°	Phi/ arc degree°	Alpha/ arc degree°	Vol. Fraction	Diameter / x10^{-6} m
				Stick/Cylind.			Tensor/Zeppelin/Ball							3rd Compart.	
761	0.96	7	ZeppelinStickDot	0.29			0.62	1.91	0.68		88.9	0.8		0.09	
748	0.96	9	TensorStickDot	0.29			0.62	1.91	0.73	0.63	88.9	0.8	10.8	0.09	
761	0.96	8	ZeppelinCylinderDot	0.30		4.45	0.62	1.91	0.68		88.8	0.9		0.09	
761	0.96	8	ZeppelinStickSphere	0.29			0.62	1.91	0.68		88.8	0.7		0.09	0.20
748	0.96	10	TensorCylinderDot	0.29		4.36	0.62	1.91	0.73	0.63	88.9	0.8	10.8	0.09	
748	0.96	10	TensorStickSphere	0.29			0.62	1.91	0.73	0.63	88.9	0.8	10.8	0.09	0.21
761	0.96	9	ZeppelinCylinderSphere	0.30		4.45	0.62	1.91	0.68		88.9	0.8		0.09	0.20
761	0.96	9	ZeppelinGDRCylindersDot	0.30	7.0	4.19	0.62	1.91	0.68		88.8	0.5		0.09	
748	0.96	11	TensorCylinderSphere	0.29		4.34	0.62	1.91	0.73	0.63	88.9	0.8	11.3	0.09	0.10
748	0.96	11	TensorGDRCylindersDot	0.29	10.0	4.14	0.62	1.91	0.73	0.63	88.9	0.8	10.8	0.09	
761	0.96	10	ZeppelinGDRCylindersSphere	0.30	10.0	4.26	0.62	1.91	0.68		88.8	0.8		0.09	0.20
748	0.96	12	TensorGDRCylindersSphere	0.29	10.0	4.14	0.62	1.91	0.73	0.63	88.9	0.8	10.8	0.09	0.20
909	0.97	7	ZeppelinStickAstrosticks	0.33			0.42	2.06	0.67		88.8	0.6		0.25	
895	0.96	9	TensorStickAstrosticks	0.33			0.43	2.06	0.73	0.58	88.9	0.7	11.2	0.25	
909	0.97	8	ZeppelinCylinderAstrosticks	0.33		3.96	0.42	2.06	0.67		88.9	0.7		0.25	
909	0.97	8	ZeppelinStickAstrocylinders	0.33			0.42	2.06	0.67		88.8	0.6		0.25	0.20
909	0.97	8	ZeppelinCylinderAstrocylinders	0.33		0.20	0.42	2.06	0.67		88.9	0.7		0.25	0.20
895	0.96	10	TensorCylinderAstrosticks	0.33		3.78	0.42	2.06	0.73	0.58	88.9	0.7	11.2	0.25	
895	0.96	10	TensorStickAstrocylinders	0.33			0.43	2.06	0.73	0.58	88.9	0.7	11.3	0.25	0.25
895	0.96	10	TensorCylinderAstrocylinders	0.33		0.25	0.43	2.06	0.73	0.58	88.9	0.7	11.2	0.25	0.25
909	0.97	9	ZeppelinGDRCylindersAstrosticks	0.33	5.2	3.82	0.42	2.06	0.67		88.9	0.7		0.25	
909	0.97	9	ZeppelinGDRCylindersAstrocylinders	0.33	1.4	0.63	0.42	2.06	0.67		88.9	0.9		0.25	0.63
895	0.96	11	TensorGDRCylindersAstrosticks	0.33	10.0	3.60	0.42	2.06	0.73	0.58	88.9	0.7	11.2	0.25	
895	0.96	11	TensorGDRCylindersAstrocylinders	0.33	1.8	0.20	0.43	2.06	0.73	0.58	88.9	0.7	11.2	0.25	0.20
1028	0.97	7	Bizeppelin	0.63			0.37	1.47	1.47		88.8	0.9			
1076	0.99	8	BallGDRCylindersDot	0.66	1.1	19.42	0.27	1.77			88.8	0.7		0.07	
1080	0.99	8	BallGDRCylindersAstrosticks	0.56	1.1	16.33	0.20	1.99			88.9	0.7		0.24	
1076	0.99	9	BallGDRCylindersSphere	0.66	1.1	19.42	0.27	1.77			88.8	0.7		0.07	1.00
1099	0.99	7	BallCylinderAstrosticks	0.53		12.60	0.22	2.01			88.9	0.6		0.24	
1133	0.95	6	ZeppelinStick	0.40			0.60	1.49	0.72		88.9	0.8			
1120	0.95	8	TensorStick	0.40			0.60	1.49	0.76	0.66	88.9	0.8	9.6		
1128	1.00	7	BallCylinderDot	0.60		13.42	0.33	1.75			88.8	0.7		0.07	
1132	0.95	7	ZeppelinCylinder	0.41		4.83	0.59	1.49	0.72		88.9	0.8			
1120	0.95	9	TensorCylinder	0.41		4.71	0.59	1.49	0.76	0.66	88.9	0.8	9.6		
1128	1.00	8	BallCylinderSphere	0.60		13.42	0.33	1.75			88.8	0.7		0.07	0.24
1132	0.95	8	ZeppelinGDRCylinders	0.41	10.0	4.67	0.59	1.49	0.72		88.9	0.8			
1120	0.95	10	TensorGDRCylinders	0.41	10.0	4.57	0.59	1.49	0.76	0.66	88.9	0.8	9.7		
1180	0.99	7	BallCylinderAstrocylinders	0.50		11.04	0.24	1.95			88.9	0.6		0.27	11.04
1181	0.99	8	BallGDRCylindersAstrocylinders	0.50	10.0	11.12	0.24	1.95			88.9	0.6		0.27	11.12
1311	0.98	7	BallGDRCylinders	0.65	1.1	14.82	0.35	1.44			88.7	0.8			
1335	1.00	7	BallStickAstrocylinders	0.41			0.31	1.93			89.0	0.6		0.28	10.38
1344	1.00	6	BallStickAstrosticks	0.41			0.36	1.86			89.0	0.7		0.23	
1345	0.98	6	BallCylinder	0.61		11.26	0.39	1.43			88.8	0.7			
1455	1.00	7	BallStickSphere	0.43			0.46	1.61			88.9	0.7		0.10	17.90
1492	0.99	6	BallStickDot	0.44			0.52	1.49			88.9	0.8		0.04	
1548	0.98	5	BallStick	0.48			0.52	1.35			88.9	0.8			
2071	0.89	7	DT				0.85	1.54	0.19	0.23	88.8	0.9	13.9		

Table 10.3: Parameter estimates obtained after fitting models to the 2x4h 2° data set. The models are ordered top-down by the BIC score. For GDR-Cylinder models, we report the mean of the radius distribution. Angles Theta/Phi/Alpha give the spatial orientation.

Cross-Validation *Bootstrapping*

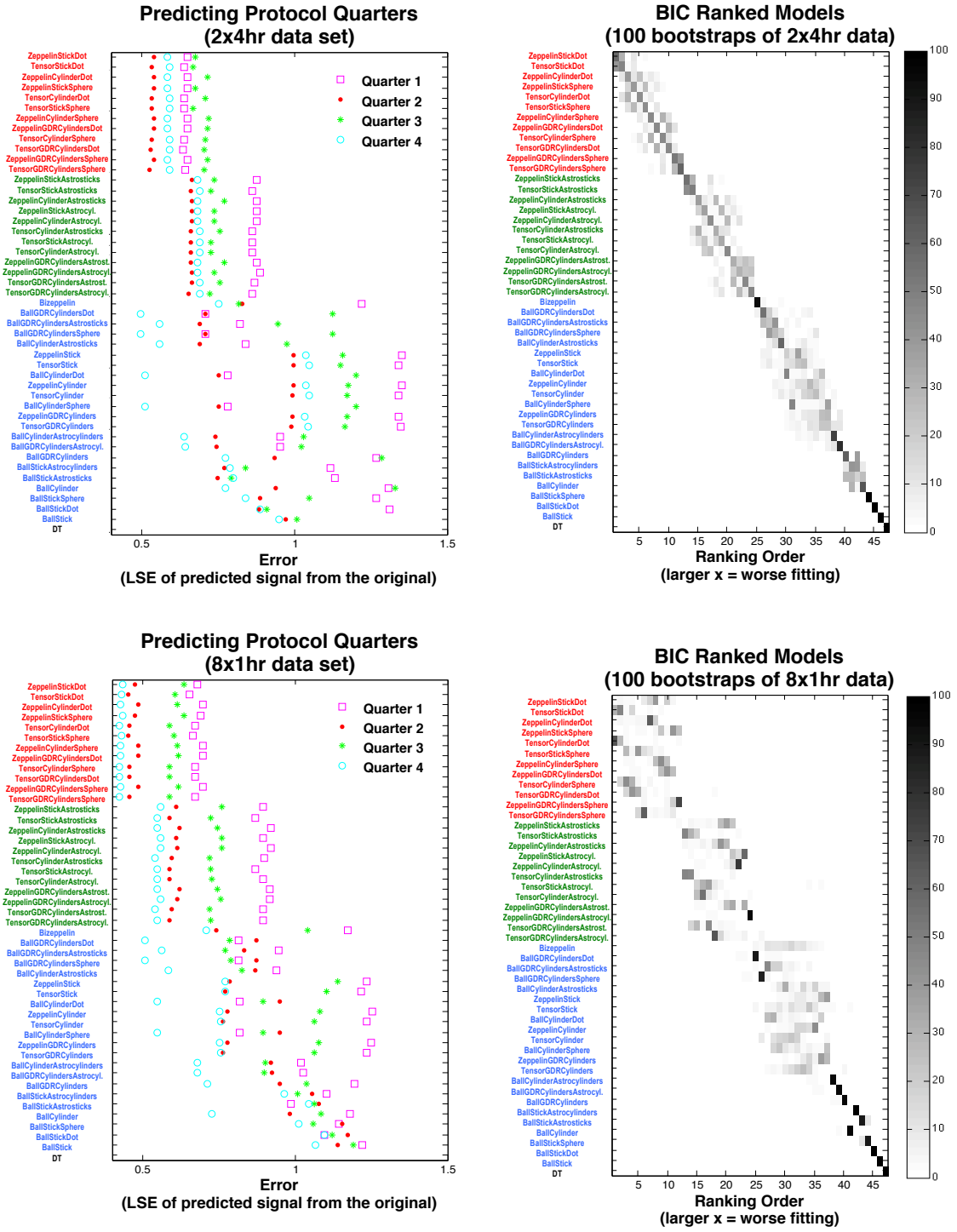


Figure 10.3: Right: Positional variance diagrams over 100 bootstraps from the 2x4h (left matrix) and 8x1h (right) 2° data sets. The frequency of x-axis ranking is given by the shade of grey; e.g. the Tensor comes out last in all 100 bootstrap samples of 8x1h. Left:The accuracy of predicting unseen quarters of the protocol using parameters fitted to data from the remaining three-quarters. Each point is the LSE between the synthesised and measured signal. The ranking is by the BIC score of the 2x4h data set. Dotted lines across the plots indicate the group structure of the ranking

11

SUMMARY

The first experiment follows up on a similar study by Panagiotaki et al. [2012], but using in vivo human data rather than fixed rat tissue. It concluded that the data acquisition protocol used was insufficient to discriminate between the models. Experiment 2 tested, and confirmed as positive, the effect of increased gradient angular resolution on the model performance.

In Experiment 3, we enriched the acquisition protocol: we sampled a wide range of b-values and diffusion times achievable on a clinical system and also used a much higher angular resolution sampling than Panagiotaki et al. [2012]. Additionally, we extended the analysis to determine ranking stability with respect to noise, protocol and model selection technique.

The overall ranking obtained is similar to previous observations from fixed tissue [Panagiotaki et al., 2012], with a few differences. Though there are minor differences due to inter-sessional variability and subsequent image registration, the similarity between 2x4h and 8x1h data sets is important because it means we can construct data sets for this kind of experiment from multiple short sessions, which are much more comfortable for the participant. The additional steps in the analysis reveal a group structure to the model ranking and suggest that the models in group (i) perform similarly well in explaining the full range of PGSE signals acquirable from the human brain on current clinical systems.

The experiments here uses only data from the corpus callosum, which is relatively homogeneous, with little fibre dispersion, crossing or CSF contamination. However, these effects may still influence the measurement to some extent. So, it is useful to explore finer regions of the corpus callosum.

Moreover, the greater angular threshold increases fibre dispersion, which is reflected in the fitting and parameter estimates, and which none of the models we test here is designed to capture. The intention here was to start with the simplest geometry before performing a similar analysis in more complex regions. Even in the corpus callosum, more sophisticated models may outperform the limited set we study here. Models which explicitly cater for fibre features such as dispersion/crossing Zhang et al. [2011, 2012]; Sotiropoulos et al. [2012], CSF pool as in Barazany et al. [2009]; Zhang et al. [2011], will be the focus of the next chapters.

In the next part we compare parametric diffusion MRI models which explicitly seek to explain fibre dispersion in nervous tissue. These models aim at providing more specific biomarkers of disease by disentangling these structural contributions to the signal.

Part IV

MODEL ADVANCES

12

INTRODUCTION

The previous three experiments applied the model comparison framework to the *in vivo* data. One limitation pointed out in the *Summary* of chapter 11 is that the models assume that, within each voxel, the neuronal fibres are straight and coherent. This reasonable assumption can become problematic in the regions where fibres cross/bend/fan or in the estimation of tissue characteristics, e.g. overestimating axon diameter indices [Zhang et al., 2011].

A recent class of parametric models has emerged to describe data better by additionally accounting for fibre directional incoherence, which is abundant in the brain, even at a sub-voxel level. Ball-and-Sticks [Behrens et al., 2003] can have more-than-one intracellular diffusion compartment. Zhang et al. [2012] constructed NODDI to describe fibres with an explicit orientation dispersion index derived from a Watson distribution (an isotropic distribution on the sphere; to be defined in the next experiment) and tested the model with *in vivo* human whole-brain data. Sotiropoulos et al. [2012] design Ball-and-Rackets to describe fibre fanning through a Bingham distribution (an anisotropic distribution on the sphere) by extending the Ball-and-Sticks model [Behrens et al., 2003]. The Bingham distribution extends the Watson distribution to account for asymmetric/anisotropic dispersion. This model is then applied to post-mortem macaque monkey brain data.

In Experiment 4 we incorporate into our taxonomy models similar to NODDI, Ball-and-Rackets, and others which have been constructed from combinations of existing compartments that aim to capture both intracellular and extracellular diffusion. To test these models we use the previous rich data set acquired *in vivo* on the CC of a human brain, and then compare the models via the Bayesian Information Criteria. We test this ranking via bootstrapping on the data sets, and cross-validate across unseen parts of the protocol, as in the previous chapter.

The work in this part has previously been published as:

Ferizi U, Schneider T, Tariq M, Wheeler-Kingshott CAM, Zhang H, Alexander DC: The Importance of Being Dispersed: *A Ranking of Diffusion MRI Models for Fibre Dispersion Using In Vivo Human Brain Data.* Lecture Notes in Computer Science, Springer Berlin Heidelberg, Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2013, vol. 8149, pp.74–81

In addition, using the above data set we organised a challenge, as part of CDMRI workshop at MICCAI'13 conference. Participants were invited to train their models on three-quarters of the data, and they were tested on the missing quarter. The best six entries presented their work at the workshop challenge. There is more information on this in the Appendix chapter B.

13

EXPERIMENT 4: ADDING DISPERSION MODELS

In this experiment we update the taxonomy with various models for not-necessarily-linear intracellular diffusion, represented via multiple sticks, which may be discrete or described via a probabilistic distribution. Other than the models being different, this experiment follows closely the format of Experiment 3: the models are fitted to the same *in vivo* human data, then compared via BIC and, again, both bootstrapping and four-fold cross-validation are used to validate the ranking.

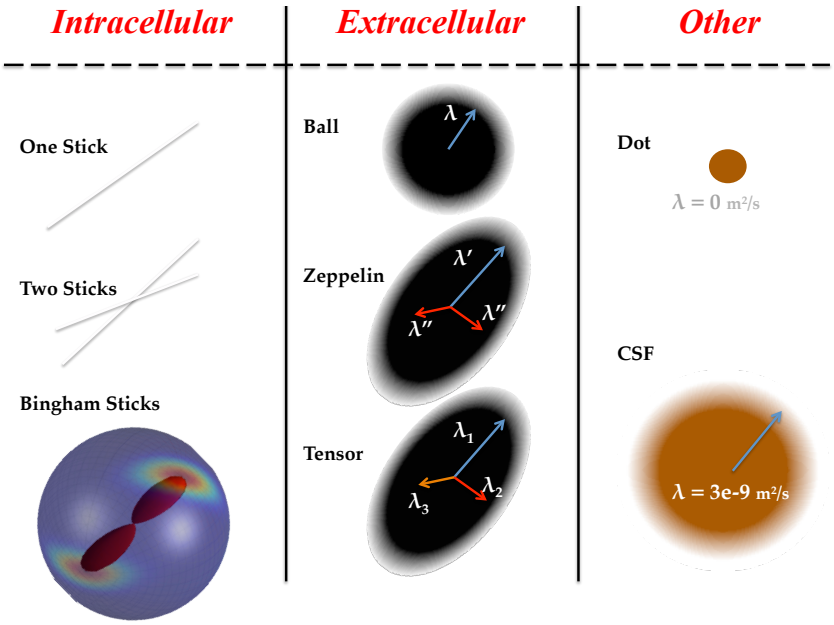


Figure 13.1: The elements of each compartment class designed to capture diffusion through a particular tissue medium: intracellular, extracellular, and the CSF. A model consists of a combination of three compartments, one from each class.

13.1 METHODS

After a description of the models, there follows the data acquisition and the pre-processing done to obtain a set of measurements for fitting the models. Last is the fitting procedure and the criterion applied to compare the models.

EXTRACELLULAR COMPARTMENTS: The compartments used to capture signal outside the axons and the isotropically restricted compartments are the Tensor, the Zeppelin and the Ball, as described in section 8.1. We follow Szafer et al. [1995] to express the *Zeppelin with tortuosity*, where the ratio of the radial vs. axial diffusivities is equal to the ratio of the volume fractions

of the Zeppelin vs. all-bar-CSF/Dot compartments. (The isotropically restricted compartment CSF uses Ball with a fixed diffusivity of $3x10^{-9}$ mm^2/s.)

INTRACELLULAR COMPARTMENTS: Sticks are used to represent the axonal diffusion, via either a discrete set of Sticks [Behrens et al., 2003] (we pick two) or an underlying Bingham/Watson fibre orientation distribution [Zhang et al., 2011; Sotiropoulos et al., 2012]. The *Bingham* distribution is

$$f(\mathbf{n}|\kappa_1, \kappa_2, \mu_1, \mu_2) = [{}_1F_1(\frac{1}{2}, \frac{3}{2}, \kappa_1, \kappa_2)]^{-1} \exp[\kappa_1(\mu_1 \cdot \mathbf{n})^2 + \kappa_2(\mu_2 \cdot \mathbf{n})^2] \tag{13.1}$$

where κ_1 and κ_2 are the concentration parameters, such that $\kappa_1 \geqslant \kappa_2 \geqslant 0$; the mutually orthogonal vectors μ_1 and μ_2 indicate the orientation axes of fibre dispersion. This is similar to a bivariate Gaussian distribution with elliptical contours on the sphere. The denominator, ${}_1F_1$, is a confluent hypergeometric function of first kind [Mardia and Jupp, 2000]. The *Watson* distribution is a special case of the Bingham distribution, where there is only one κ and μ ($\kappa_2 = 0$); this corresponds to circular contours on the sphere.

DATA ACQUISITION AND PRE-PROCESSING Here we use the 2x4h data set from Experiment 3 (Section 10.1). Briefly, the protocol combines many pulse times δ, diffusion times Δ and gradient strengths $|\mathbf{G}|$ to produce 32 shells with a b_{max}=10,308 s/mm^2.

The pre-processing produced three sets of voxels with varying deviation form the main fibre direction: η=2° left 24 voxels, η=5° left 66 voxels, and 10° left 99 voxels. The voxels are located across the genu and mid-body.

The signal is then normalised by the b=0 images with the same TE. A single data set is created by averaging the voxels selected above. Figure 10.1 shows the signal from the 2° data set, containing 1,536=32*(3+45) measurements.

MODEL FITTING AND SELECTION: The fitting is as in the Methods 8.1 of Experiment 1. Each model is fitted 250 times, and the final parameters are those that produce the minimum objective function LSE (eq.8.2).

The criterion BIC (eq.6.1) is then used to compare the models.

BOOTSTRAPPING AND CROSS-VALIDATION As in Methods 10.1 of Experiment 3, we construct 100 bootstrap data sets by sampling with replacement in each shell the same number of data points. In the four-fold cross-validation we divide the data set into four quarters, training the models on three-quarters of the original data and testing the models on the missing/unseen data.

13.2 RESULTS

Table 13.1 ranks some of the models and lists main parameter estimates across all three data sets (with η=2°, 5° and 10°); a more extensive Table C.1, in the Appendix, lists all the models and parameters. Included here are the best model of the previous chapter's ranking of parametric models with no-dispersion, and a similar model with CSF instead of Dot. Four groups can be distinguished:

i) all combinations that include an anisotropic extracellular compartment and a Bingham/Watson intracellular compartment;

ii) models similar to (i) but instead using two-Sticks for their intracellular compartment, excluding models that use tortuosity or those without a spherically restricted compartment;

iii) all models incorporating an isotropic extracellular compartment with a Bingham/Watson intracellular compartment; and

iv) all exceptions to two-Sticks models in (ii).

The models that include a Bingham/Watson distribution outperform two-Sticks ones not simply because of their good quality of fit to the data but also because of their reduced complexity.

LSE	S0	Nr Parameters	MODELS	Stick1/Watson/Bing							Ball/Zep/Ten.				Stick2			CSF/Dot
				Vol. Fraction	Diffusivity (10^{-9} m^2/s)	Kappa (Bin.g/Wats.)	Beta (Bing./Wats.)	Theta/ arc degree°	Phi/ arc degree°	Psi/ arc degree°	Vol. Fraction	Rad. Diff.1 (10^{-9} m^2/s)	Rad. Diff.2 (10^{-9} m^2/s)	Alpha/ arc degree°	Vol. Fraction	Theta/ arc degree°	Phi/ arc degree°	Vol. Fraction
481	1.00	10	Zepp.Bing.CSF.	0.56	2.0	6.94	2.17	74	89	1	0.29	0.5						0.15
511	1.00	7	ZepT.Wat.CSF.	0.59	2.0	5.77		89	1		0.28	0.6						0.13
512	0.98	10	Zepp.Bing.Dot	0.50	2.1	10.33	4.10	71	89	1	0.45	0.9						0.04
526	0.98	11	Tens.Bing.	0.65	2.2	7.18	2.72	67	89	1	0.35	1.2	0.9	55				
550	0.98	8	ZepT.Bing.	0.62	2.2	6.47	1.84	74	89	1	0.38	0.8						
614	0.97	12	Tens.St.St.Dot	0.23	2.0			86	4		0.56	0.8	0.7	20	0.14	75	14	0.07
635	1.00	12	Tens.St.St.CSF.	0.22	1.5			86	5		0.41	0.5	0.3	25	0.17	74	18	0.21
703	1.01	8	Ball.Bing.	0.72	2.2	6.05	1.41	75	88	1	0.28							
703	1.01	9	Ball.Bing.CSF.	0.72	2.2	6.05	1.41	74	89	1	0.28							0.00
761	**0.96**	**7**	**Zepp.St.Dot**	**0.29**	**1.9**			**89**	**1**		**0.62**	**0.7**						**0.09**
801	**1.00**	**10**	**Tens.Cylinder+CSF.**	**0.29**	**1.3**			**89**	**1**		**0.47**	**0.3**	**0.3**	**12**				**0.24**
814	0.98	9	ZepT.St.St.Dot	0.33	1.8			86	2		0.50	1.1			0.12	69	12	0.05
824	0.96	11	Tens.St.St.	0.28	1.7			84	5		0.52	0.9	0.7	29	0.20	70	17	
852	0.99	9	ZepT.St.St.CSF.	0.35	1.5			86	2		0.39	0.9			0.13	66	13	0.12
870	0.97	8	ZepT.St.St.	0.32	1.6			84	4		0.50	1.0			0.18	71	13	
1135	0.99	8	Ball.St.St.	0.28	1.5			80	6		0.46				0.25	75	10	

Table 13.1: Parameter estimates obtained after fitting models to the 2° data set; only a few models are shown, with the full list given in Table C.1. The models are ordered top-down by the BIC score. Here, we also include the estimates (shown in bold) from the best model of the previous chapter's ranking of parametric models with no-dispersion. Angles Theta/Phi/Alpha/Psi give spatial orientation; the number of model parameters includes the parameter S_0, which is the unweighted signal at b=0. [Note: Zepp=Zeppelin; ZepT=Zeppelin with tortuosity;Tens=Tensor; St=Stick; Bing=Bingham; Wat=Watson].

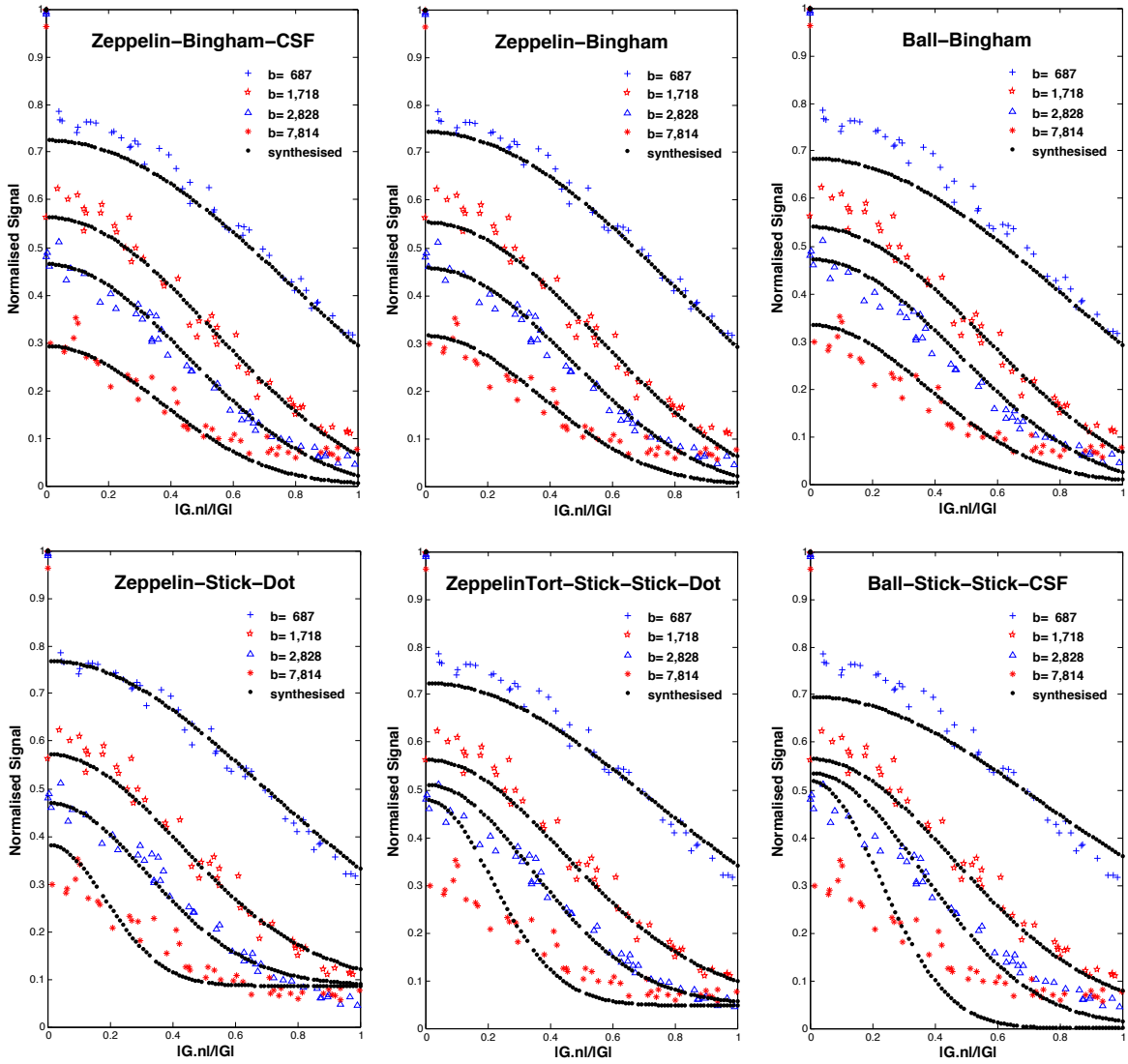


Figure 13.2: A comparison of raw vs. predicted/synthesised signal from six representative models. The models are ordered in decreasing ranking left-right, top-bottom.

Within group (i), CSF models perform best for η=2° but, as η increases, Dot models are best. In this group, models using tortuosity produce similar estimates to those of the unconstrained Zeppelin, suggesting that meaningful constraints on the model parameters, such as the tortuosity assumption, can be used to simplify the problem at little cost to fitting quality.

Across angular thresholds, the axial diffusivity is about $2\text{x}10^{-9}$ mm^2/s, and the radial diffusivity is around one-quarter of this in models with CSF, but one-half in others; this is to be expected as the CSF compartment has a fixed diffusivity of $3\text{x}10^{-9}$ mm^2/s and higher volume fraction than Dot.

As η increases from 2° to 5°, all models reflect the signal improvement from averaging across more voxels (24 vs. 66, resp.) through decreasing BIC and increasing fibre incoherence κ; however, at 10° (with 99 voxels averaged), the fitting improves slightly, but κ reflects the increased fibre coherence through decreasing κ.

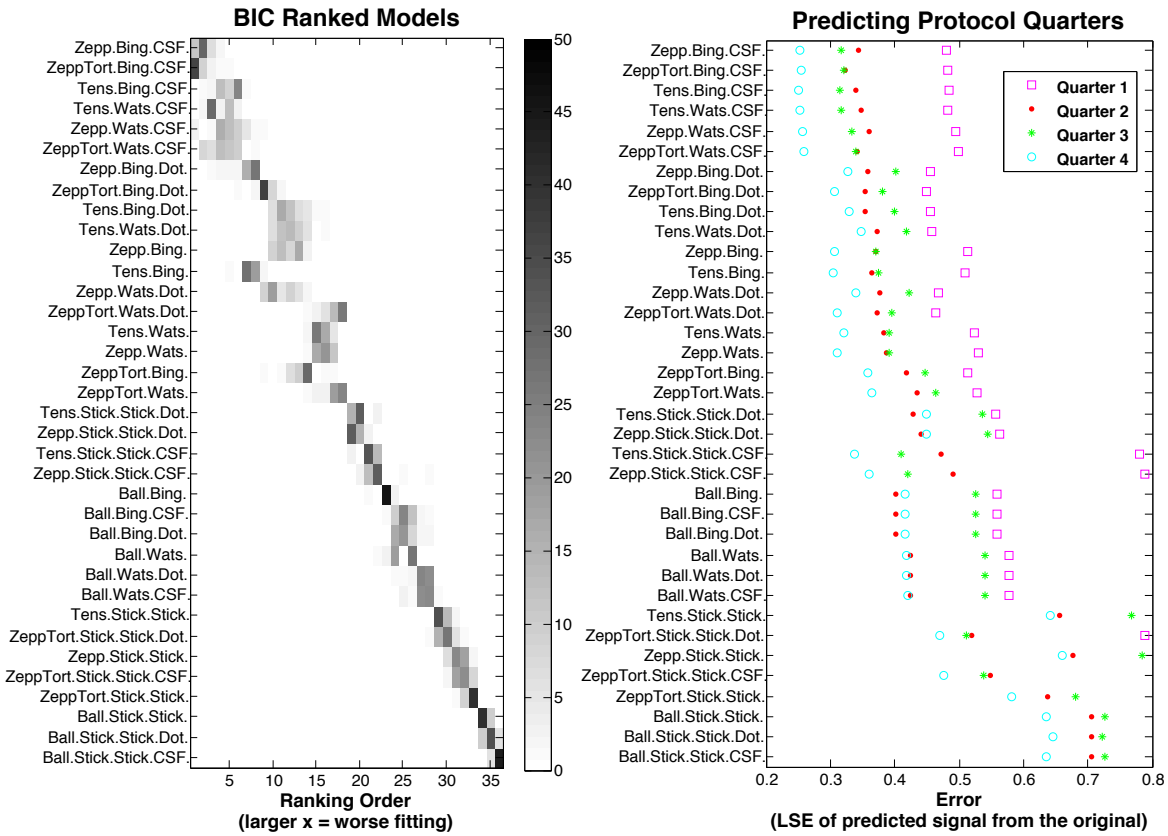


Figure 13.3: LEFT: Positional variance diagrams over 100 bootstraps from the 2° data sets. The frequency of x-axis ranking is given by the shade of grey. RIGHT: The accuracy of predicting unseen quarters of the protocol using parameters fitted to data from the remaining three-quarters. The ranking is as in Table C.1.

Figure 13.2 shows the fit of some representative models to the data, to illustrate the difference between the actual signal and that generated from the model.

Figure 13.3 shows on the left the positional variance diagram for the BIC ranking through classical bootstrap. The ordering of the original ranking remains faithful through these ranking histograms, and group structure remains unchanged, though there are minor variations within each group. On the right of fig.13.3, the relative performance of each model in reproducing unseen parts of the data set is shown. While broadly speaking the trends in both techniques agree, cross-validation is less discriminatory within the groups. This technique also reveals other subgroups within groups, e.g. within the top models of group (i), three-compartment models with CSF do better than those with Dot (i.e. the stationary water compartment), or than other two-compartment models.

13.3 DISCUSSION

This experiment has shown the potential advantage of dispersion models in describing data even in a homogeneous region of the brain such the CC. In such structure, where a multitude

of function specific fibre tracts bundle together, there is inhomogeneity that can produce a dispersion pattern, which is something that these models may reflect.

Because modelling the fibre population through a single-mode Bingham distribution outperformed two-Sticks, the result suggests that the signal arises from small fluctuations in a single fibre orientation rather than a small population of fibres with totally different orientation. On the other hand, even though two Stick models significantly outperform the single Stick models, this does not necessarily make the two-fibre model correct.

The ranking also identified the CSF compartment to be better than Dot. This is expected, as the Dot compartment is designed for stationary water, which would be more appropriate with fixed tissue studies such as that of Panagiotaki et al. [2012], Dyrby et al. [2013], and Richardson et al. [2013].

The bootstrapping on the data sets revealed a ranking order very similar to the original one in Table 13.1. The cross-validation confirmed this ranking, but produced a more stratified group structure, and closer model-similarity within subgroups. This is not unexpected, as bootstrap sampling from such a large data set produces data sets very similar to the original one, hence the ranking is not expected to differ greatly. Four-fold cross-validation data sets, however, are relatively more heterogeneous, as each time we leave out a quarter of the original data set; hence the model ranking stability is somewhat different.

One obvious limitation that arises from our methodology is that averaging voxels across parts of the CC, as well as minor misalignments during image registration, may introduce and/or exaggerate the dispersion. Smaller, ideally voxel-based, analyses would be more appropriate and improve accuracy.

Another limitation in this methodology is that, because of the TE range used in this work, the "normalisation" of the signal in the previous experiments has the disadvantage of making TE-dependent the otherwise constant thermal noise. Though we expect the effect and differences to be minor on the performance of the models relative to one another, accounting for the TE-decay would improve the accuracy on parameter estimation.

The last limitation to list is that the results presented here were obtained from just one subject. It is important to note, therefore, that it is uncertain whether the same results would be obtained in another participant, or from another scanning session.

14

SUMMARY

In this part we experimented with different models; a class of promising dispersion models were fitted to the multi-shell data, and they were found to capture the signal better than previous models for straight and coherent fibres.

The implication of this is that, even in the most coherent structures of the brain, such as the CC, fibre dispersion differs significantly from a delta function. In particular, our analysis showed that the single mode orientation distributions (Watson/Bingham) outperform two discrete orientations (two-Sticks). As in the previous experiments, an anisotropic extracellular compartment benefits the fitting, as does the addition of an isotropically restricted compartment.

We also identified limitations in the methodology: one is that the analysis is confined to a single data set; the other is that this data set is constructed in a way which may inadvertently distort the dispersion models. In the next chapters we will explore ways of correcting for these, and other improvements to the methodology.

Part V

METHODOLOGICAL IMPROVEMENTS

15

INTRODUCTION

In translating the model comparison framework to in-vivo human data, the earlier experiments identified areas of methodology that needed further investigation.

In all of the following experiments we will be using a smaller ROI on which the data will be tested. This is important, especially for models which are designed to capture dispersion and which can be over-sensitive to averaging across the CC. Testing the models voxel-by-voxel will help us see how consistent the model selection is within regions, and help reduce artificially inflated dispersion.

In Experiment 5, we improve the model fitting, by accounting explicitly for the TE decay. Usually, most applications of diffusion MRI models of white matter do not fit relaxation T_2 because the acquisition is usually of a single TE. However, our experiment samples multiple TE values and accounting for TE means that we also need to include the T_2 decay explicitly in our models. Here, we are able to investigate this parameter precisely because our data contains a wide TE range, making T_2 estimation feasible. Work done with multi-spin-echo imaging on living tissue has identified various, usually two or three, T2-specific compartments. The longest, greater than about 1000ms, is attributed to the CSF compartment and the shortest, around 10-20ms to the myelin; the one in between being cellular. Some studies in non-human tissue distinguish between intra- and extra-cellular compartments [Menon et al., 1992; Peled et al., 1999], while others in the normal live human brain [MacKay et al., 1994, 2006] , and therefore more relevant to this study, identify only one peak for the cellular compartment, at around 70-90ms.

On the DW-signal modelling side, we will also fit for compartmentally different diffusivities. The simplification of equal intracellular and extracellular diffusivities is applied in some current white matter models such as ActiveAx [Alexander et al., 2010] or NODDI [Zhang et al., 2012]. The analysis here will show how sensitive the models are to this assumption.

In Experiment 6, we explore another area to be addressed: the inter- and intra-subject reproducibility of model ranking and parameter estimates. Our experiment so far has been confined to one healthy subject, and one acquisition (over two sessions). Therefore, we need to ensure that the results are not affected by spurious imaging effects or particularities of scanning. We include experiments that compare the sensitivity of parameter estimates across similar corresponding regions, across repeated scanning and other healthy subjects, which provides a complementary evaluation of the models.

16

EXPERIMENT 5: ACCOUNTING FOR TE AND T_2

In this experiment, as a first enhancement, we will:

- fit the models voxel-wise over the more homogeneous regions of genu, midbody and splenium, rather than across the whole of the CC;
- investigate how the parameters vary within the defined ROI, as well as across the CC;
- account explicitly for the variable echo time TE among measurements, by also fitting to the data compartmentally different T_2;
- fit for compartmentally different diffusivities.

As before, after describing the Methods, we describe the results.

16.1 METHODS

DATA ACQUISITION The data used here comes from Experiment 3 and 4 (section 10.1). This protocol contains 32 shells with a b_{max}=10,308 s/mm^2.

In contrast with earlier experiments, here we do not normalise the signal, that is we use the raw signal. Nor do we average across voxels, but perform the analysis on a voxel-by-voxel basis, and report the variance across the four ROI voxels. We define three ROI, in the genu, midbody and splenium of the mid-sagittal slice. Each of these regions consist of four voxels. To select the voxels, we first fit the Ball-Stick-CSF model to the whole CC data and then select the four voxels within each area with the highest Stick volume fraction, and with the Stick direction most closely aligned with the mid-sagittal perpendicular. This ensures that the voxels have least CSF contamination and are well embedded in white matter.

COMPARTMENT MODELS All the models considered are a combination of three compartments. The full set of candidate models for each compartment leads to a very large set of three-compartment models. The focus here is on a small subset of models, as identified in the previous chapters, that emerged as strong candidates. In particular, we consider only isotropic free diffusion for the third compartment, modelling CSF contribution, since we consider only *in vivo* data. For the intracellular compartments, we consider only Cylinder, Bingham-Sticks, and one/two Sticks. Similarly, we use only Ball and full Tensor for the extracellular compartment. In the model comparison we include NODDI [Zhang et al., 2012] and MMWMD [Alexander et al., 2010]. Briefly, their extracellular compartment is a cylindrically symmetric tensor/DT, with the radial and axial diffusivity related, as in the tortuosity model of Szafer et al. [1995], with the

axial diffusivity itself fixed. In the two other similar models NODDI+ and MMWMD+, we instead fit for the axial and radial diffusivities.

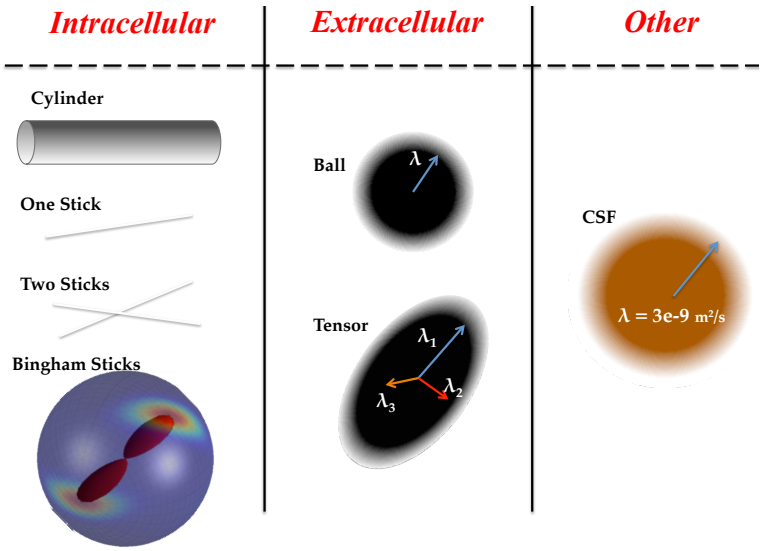


Figure 16.1: The elements of each compartment class designed to capture diffusion through a particular tissue medium: intracellular, extracellular, and the CSF. A model consists of a combination of three compartments, one from each class.

T2 EFFECT Varying TE means that we need to model T_2 effects. The total signal model would then become:

$$S = S_0 \left(f_i \exp\left(-\frac{TE}{T_2^i}\right) S_i + f_e \exp\left(-\frac{TE}{T_2^e}\right) S_e + f_c \exp\left(-\frac{TE}{T_2^c}\right) S_c \right) \tag{16.1}$$

where f_i, f_e and f_c are the weights of the intracellular, extracellular, and third signal compartment S_{intra}, S_{extra} and S_c, respectively; the values of compartmental T_2 are indexed similarly; S_0 is the b=0 signal (and TE-dependent too).

UNEQUAL DIFFUSIVITIES Much previous work [Alexander et al., 2010; Alexander, 2008; Ferizi et al., 2013a,b; Panagiotaki et al., 2012; Zhang et al., 2011, 2012] assumes equal intrinsic diffusivity (i.e. excluding the effects of restriction) of the water in the intracellular and extracellular compartments. For each of our compartment models, we evaluate performance with and without this constraint. To highlight the difference in later results, the models with separate intra/extracellular diffusivities have "-diff" appended to their name.

MODEL FITTING Model fitting has three stages:

1. Estimate voxel-wise the compartmental T_2 from unweighted signals with varying TE. First, we find the T_2 of CSF from a ROI in the ventricle.

2. A preliminary analysis inside homogeneous WM regions of the CC (see Results 16.2) showed that the b=0 signal decays mono-exponentially with TE. Therefore, we assumed

that the intracellular and extracellular T_2 are equal; hereafter, we refer to "T_2 of WM". A bi-exponential model fits WM T_2, S_0 and f_c to the b=0 data.

3. Fit other parameters to all data (i.e. with fixed CSF diffusivity, CSF volume fraction, S_0 signal and each compartmental T_2 of CSF and WM).

Each time we fit a model, we repeat the first two steps 10 times, and the best estimates are fed into the last step. This last step, i.e. fitting to the DW signal, is performed 20 times, with the starting point perturbed from initial estimates (taken from wider literature, including Alexander [2008] and Dyrby et al. [2013]): the volume fractions were equally split across each compartment, the axial diffusivity is 2 $\mu m^2/s$, radial diffusivity is 1 $\mu m^2/s$; the radius is initialised at 2 μm (with an upper limit of 20 μm). The CSF diffusivity was fixed throughout the experiments to 3 $\mu m^2/s$, and all other compartmental diffusivities have this value as an upper bound.

DATASETS We considered various ROIs and subsets of measurements to construct different datasets for fitting and model comparison.

The following four ROI were defined on the Achieva+ acquisition:

- ACH-genu includes four voxels in the centre of the mid-sagittal volume, with no evidence of CSF partial volume;
- ACH-midbody similarly includes four mid-sagittal midbody voxels;
- ACH-splenium includes four mid-sagittal splenium voxels;
- ACH-CSF includes four mid-sagittal ventricular CSF voxels (used for estimating its T_2).

MODEL RANKING For model comparison, instead of the earlier BIC, we return to AIC, as defined in eq.6.3. While the two are not very different in their output, AIC is more often used in cases such as this, where the parameter estimation is through a maximum-likelihood method. The *maximum likelihood* estimation of model parameters means that maximising the posterior distribution in the parameter space is equivalent to maximising the likelihood (the *priors* can be regarded as uninformative in a Bayesian approach).

16.2 RESULTS

T_2 EFFECTS: Figure 16.2 shows the distribution of log-signal at b=0 versus echo time TE for white matter. In both white matter and CSF the data show no significant and meaningful departure from the mono-exponential model assumption, so for each scanner we concluded that, at least for this dataset, the intracellular and extracellular T_2 are the same. In the genu of the CC, T_2 averaged about 59ms. T_2 was higher in the midbody, at 62ms, and splenium at 75ms.

As expected, the T_2 was just-over 1,000 ms on ACH-CSF data. These estimates are in the order of previous estimates [MacKay et al., 1994; Whittall et al., 1997] of about 71 for genu, and over-1000 for CSF.

MODEL RANKING: Figure 16.3 shows the performance of the models fitted to the three datasets. On top of the ranking of ACH-*genu* data are the models containing Tensor and/or Bingham. Combinations of Ball with Stick/s rank last. In ACH-*splenium* the models rank similarly. It is clear that there is more intra-voxel AIC similarity in the more homogeneous regions of genu and splenium. In the midbody, ACH-*midbody*, AIC cannot distinguish well between the models.

Figure 16.4 illustrates the fit of four of the best models to the raw ACH-*genu* signal (Appendix fig.D.1 shows the plots for midbody and splenium). We select only one of the four ROI voxels, after ensuring inter-voxel similarity of plots. From earlier model ranking we pick, from the models with the best ranking performance, two Tensor and two Ball models, as follows: Tensor-Cylinder-CSF-diff, Tensor-Stick-CSF, Ball-Bingham-CSF-diff, and Ball-Stick-CSF-diff. The model signal is shown as solid line, whereas the raw data is shown with markers. We can see that while all four models capture the low b-value data well, Ball-Stick-CSF-diff is visibly worse than other models at capturing the higher (than about 5,000 s/mm^2) b-value shells.

PARAMETER STABILITY Table 16.1 gives the parameter estimates of the models after being fitted to signal from ACH-*genu*. Beside each mean estimate across the five voxels, we provide (in small superscript) the standard deviations of these estimates as a percentage of the mean. Across all models with ACH-*genu*, except Tensor and Stick combinations, the parameters reflect higher intracellular volume fractions, especially in Ball models. The joint intra/extracellular diffusivity is slightly higher in Bingham models than in Cylinder ones. In models with separate intra/extracellular diffusivities, Ball diffusivity is lower than Bingham/Sticks diffusivity, but otherwise in the other models. The Bingham distribution is more dispersed in Ball versus Tensor models because the model has to compensate for the lack of extracellular anisotropy.

Appendix table D.3 concerns the other regions of the CC, with datasets ACH-*midbody* and ACH-*splenium*. Compared with ACH-*genu*, the more heterogenous ACH-*midbody* dataset produces more dispersion (lower κ) and CSF volume, which can be expected in such a thin region prone to CSF contamination. The Tensor-Cylinder models provide higher axon thickness, which follows the expected "low/genu-high/midbody-lowest/splenium" axon-thickness trend in the CC [Aboitiz et al., 1992], unlike the lower estimates provided by the two Ball-Cylinder models. With the exception of most separately-fitted intracellular diffusivities, all other diffusivity estimates are higher compared with ACH-*genu*.

Overall, ACH-splenium produces similar parameter trends to ACH-genu. An obvious difference arises from radii estimation which are slightly under those of ACH-genu; this is consistent with known CC trends (as mentioned above).

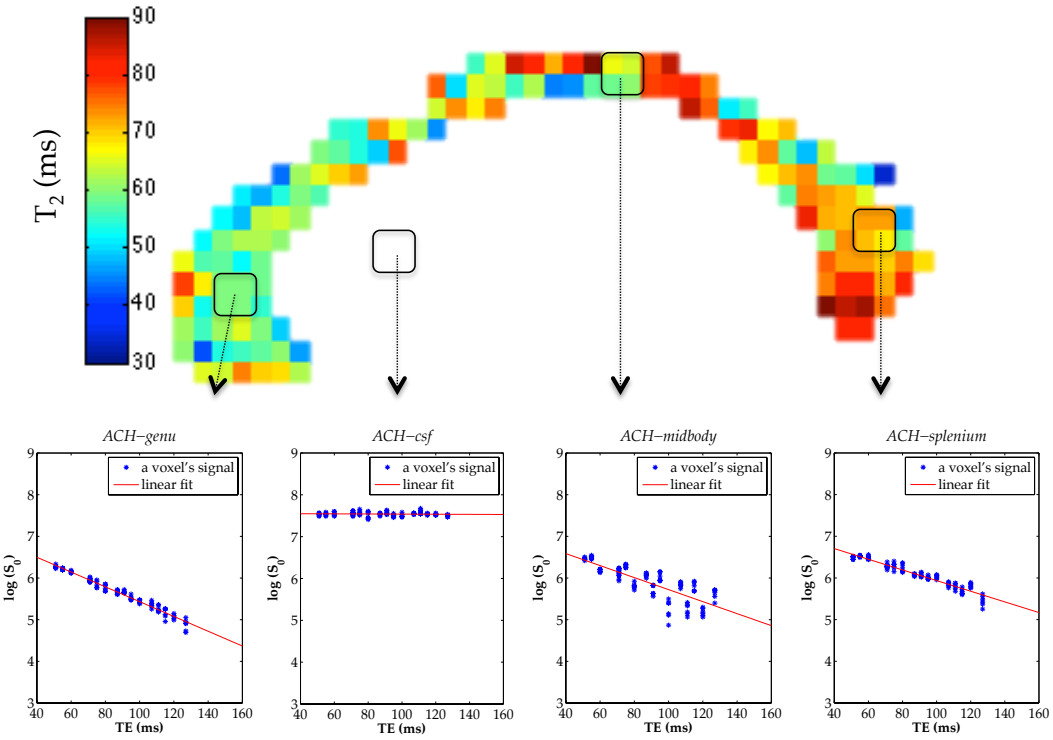


Figure 16.2: A T_2 map of the CC (centre, in colour). The bottom (left-right) the subplots relate to ACH-genu, ACH-CSF, ACH-midbody, and ACH-splenium. Each of the two subplots shows on the left a voxel's $b = 0$ signal decay with TE. There are 6 b=0 signals for each of the 16 TEs, and they are marked in blue; the gradient of the fitted red line gives the negative inverse of the T_2 value for each compartment. For this dataset, the distribution of points in white matter (genu/midbody/splenium) did not suggest two separate rates of T_2 decay, therefore the intra/extracellular T_2 are fixed to be the same.

16.3 DISCUSSION

As mentioned in the Introduction, accounting for varying data TE and fitting for T_2 makes the model fitting more correct. As can be seen from eq.16.1, we expect that fitting for T_2 would have the greatest impact on the relative CSF vs. WM volume fractions. In particular, a much higher (than WM) CSF T_2 decay would affect its volume fraction and, in turn, WM compartmental volume fractions. Therefore, the T_2 is particularly important when fitting to data from regions with partial volume contamination.

In estimating the T_2, there are multiple minima within the range of the fitting runs. This arises as a result of the estimation of T_2 alongside S_0 and volume fraction of CSF: slight perturbations in the values of the latter two cause the T_2 to vary by up to 5ms either side of the

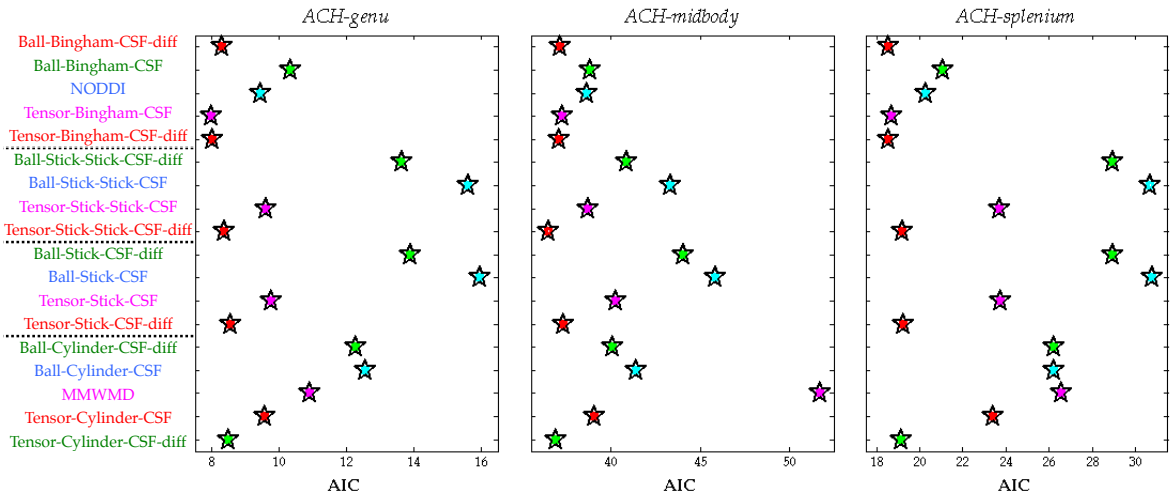


Figure 16.3: Model ranking for the three main Achieva+ datasets: ACH-genu, ACH-midbody and ACH-splenium. A similar trend persists across genu and splenium rankings: combinations of Tensor with Cylinder or Bingham are best, whereas those of Ball with Sticks are worst.

mean. The T_2 reported in the tables of parameter estimates is the one which, along with other estimates produces the minimum objective function for the whole dataset (and not just the b=0 signal). We will report more on the stability of the T_2 in section 23.2.

We regard the results from the genu as more reliable than those from the splenium or midbody. The reason for this is that the genu ROI is best embedded in white matter, surrounded by at least two other extra-ROI voxels of white matter, the splenium ROI is less well "padded", with around one voxel separating it from CSF, whereas the midbody ROI is as thick as the midbody itself and is highly susceptible to CSF contamination. Additionally, the limited imaging FOV is purposely positioned so that the genu is at it centre. So, all subsequent image distortions appeared more pronounced away from the FOV centre and, therefore, image artefact corrections worked less well on the splenium and the midbody than the genu. Another possible source of distortions can be the presence of arteries that supply blood to the CC, one of which is the splenial artery ('posterior pericallosal artery'). Though our scanning was cardiac gated, it is possible that artery pulsations distort one part of the CC more than another. While the results from the midbody and splenium can still be informative, the greater stability of location and homogeneity of the genu ROI through 8h of scanning time render the genu the primary object of study for the CC.

As regards parameters, the apparent volume fraction 'anomaly', i.e. lower intracellular than extracellular volume fractions, has been observed previously, and been given various interpretations: Assaf et al. [2002, 2004] assume "that the majority of the 'slowly diffusing' spins are undergoing restricted diffusion within the neuronal fibers". Assuming that the intracellular space occupies about 70-80% of the total volume, previous bi-tensor model fits [Niendorf et al., 1996; Clark and Le Bihan, 2000], however, attribute the higher volume fraction to 'fast-diffusing' spins.

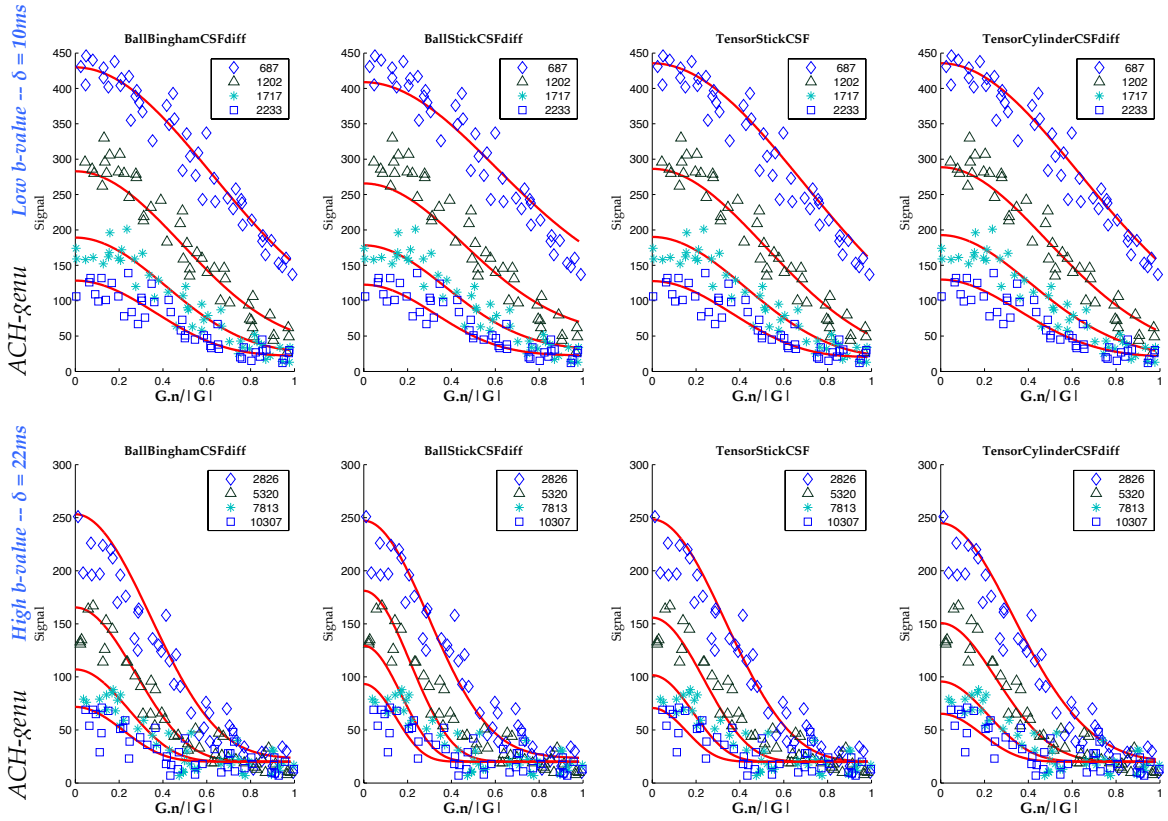


Figure 16.4: Plots illustrating the quality of fit for four selected models fitted to ACH-genu. We select one voxel from the genu ROI, whose raw signal is shown with markers and the model signal shown as solid line. Though the models are fitted to all the data, for clarity, plots show only a few selected high and low b-value shells of $|\mathbf{G}|$=60mT/m. The four Δ-specific shells increase in value from top-bottom.

ACH-genu

Models	LSE (x 10^3)	Cylinder/Bingham/Stick Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	*Tensor/ Ball* Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
Ball-Bingham-CSF-diff	0.81 $^{0.0}$	0.79 $^{0.0}$	*2.56* $^{0.0}$	18 $^{0.0}$		0.21 $^{0.0}$	*0.73* $^{0.0}$		
Ball-Bingham-CSF	1.02 $^{0.0}$	0.92 $^{0.0}$	*2.40* $^{0.0}$	10 $^{0.0}$		0.08 $^{0.0}$	*2.40* $^{0.0}$		
NODDI	0.94 $^{0.0}$	0.66 $^{0.0}$	*1.90* $^{0.0}$	11 $^{0.0}$		0.34 $^{0.0}$	*1.90* $^{0.0}$	*0.64* $^{0.0}$	
Tensor-Bingham-CSF	0.78 $^{0.0}$	0.65 $^{0.0}$	*2.01* $^{0.0}$	28 $^{0.0}$		0.34 $^{0.0}$	*2.01* $^{0.0}$	*0.54* $^{0.0}$	*0.49* $^{0.0}$
Tensor-Bingham-CSF-diff	0.78 $^{0.0}$	0.67 $^{0.0}$	*2.14* $^{0.0}$	25 $^{0.0}$		0.32 $^{0.0}$	*1.80* $^{0.0}$	*0.53* $^{0.0}$	*0.50* $^{0.0}$
Ball-Stick-Stick-CSF-diff	1.35 $^{0.1}$	0.62 $^{0.3}$	*1.98* $^{0.5}$			0.36 $^{0.6}$	*0.89* $^{0.3}$		
Ball-Stick-Stick-CSF	1.55 $^{0.0}$	0.67 $^{0.1}$	*1.53* $^{0.1}$			0.31 $^{0.2}$	*1.53* $^{0.1}$		
Tensor-Stick-Stick-CSF	0.94 $^{0.0}$	0.45 $^{0.3}$	*1.68* $^{0.0}$			0.54 $^{0.3}$	*1.68* $^{0.0}$	*0.59* $^{0.5}$	*0.31* $^{0.4}$
Tensor-Stick-Stick-CSF-diff	0.82 $^{0.0}$	0.39 $^{2.0}$	*0.99* $^{1.2}$			0.59 $^{1.3}$	*2.53* $^{0.7}$	*0.53* $^{2.2}$	*0.27* $^{2.8}$
Ball-Stick-CSF-diff	1.37 $^{0.0}$	0.62 $^{0.0}$	*2.00* $^{0.0}$			0.37 $^{0.0}$	*0.87* $^{0.0}$		
Ball-Stick-CSF	1.58 $^{0.0}$	0.67 $^{0.0}$	*1.51* $^{0.0}$			0.33 $^{0.0}$	*1.51* $^{0.0}$		
Tensor-Stick-CSF	0.95 $^{0.0}$	0.45 $^{0.0}$	*1.68* $^{0.0}$			0.55 $^{0.0}$	*1.68* $^{0.0}$	*0.59* $^{0.0}$	*0.32* $^{0.0}$
Tensor-Stick-CSF-diff	0.83 $^{0.0}$	0.40 $^{0.0}$	*1.00* $^{0.0}$			0.60 $^{0.0}$	*2.51* $^{0.0}$	*0.54* $^{0.0}$	*0.28* $^{0.1}$
Ball-Cylinder-CSF-diff	1.21 $^{0.0}$	0.81 $^{0.0}$	*1.88* $^{0.0}$		6.0 $^{0.0}$	0.19 $^{0.0}$	*0.73* $^{0.0}$		
Ball-Cylinder-CSF	1.25 $^{0.0}$	0.89 $^{0.0}$	*1.63* $^{0.0}$		6.2 $^{0.0}$	0.10 $^{0.0}$	*1.63* $^{0.0}$		
MMWMD	1.08 $^{0.0}$	0.66 $^{0.0}$	*1.70* $^{0.0}$		5.1 $^{0.0}$	0.34 $^{0.0}$	*1.70* $^{0.0}$	*0.57* $^{0.0}$	
Tensor-Cylinder-CSF	0.94 $^{0.0}$	0.53 $^{0.1}$	*1.68* $^{0.0}$		4.7 $^{0.2}$	0.47 $^{0.1}$	*1.68* $^{0.0}$	*0.58* $^{0.0}$	*0.27* $^{0.3}$
Tensor-Cylinder-CSF-diff	0.83 $^{0.0}$	0.43 $^{0.0}$	*1.04* $^{0.0}$		3.4 $^{0.0}$	0.56 $^{0.0}$	*2.57* $^{0.0}$	*0.54* $^{0.0}$	*0.26* $^{0.0}$
NODDI+	0.93 $^{0.0}$	0.65 $^{0.0}$	*1.97* $^{0.0}$	9 $^{0.0}$		0.35 $^{0.0}$	*1.97* $^{0.0}$	*0.53* $^{0.0}$	
MMWMD+	1.06 $^{0.0}$	0.55 $^{0.0}$	*1.67* $^{0.0}$		4.5 $^{0.0}$	0.45 $^{0.0}$	*1.67* $^{0.0}$	*0.46* $^{0.0}$	

Table 16.1: Parameter estimates from model fitting to each full dataset ACH-genu. The volume fraction of CSF is zero, and the estimated ROI mean T_2 is 59ms. Not shown is the "fanning" parameter, κ_2 from Bingham, which is lower than (at about half of) κ_1 making Bingham isotropically dispersed, like Watson. (Appendix fig.D.3 has estimates for ACH-midbody and ACH-splenium.)

17

EXPERIMENT 6: INTER- AND INTRA-SUBJECT REPRODUCIBILITY

To explore the generalisability of our earlier analysis, with regard to model ranking and parameter estimation, we will examine the intra- and inter-subject reproducibility. The inter-subject investigation will be done through scanning other healthy subjects via a reduced protocol; one of these healthy subjects is the main subject of the data used so far. By comparing the results from these acquisitions we can assess the intra-subject reproducibility. In addition, we will see the results from a dataset acquired on the same protocol as the main dataset; we have seen this acquisition in Experiment 3 (Chapter 10), referred to as the "8x1h".

Achieva- Protocol

	δ=8ms					δ=13ms				
Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)	Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)
1	20	54	60	286	7	20	64	60	682
2	40	68	60	616	8	40	73	60	1553
3	60	88	60	945	9	60	93	60	2424
4	80	108	60	1275	10	80	113	60	3294
5	100	128	60	1605	11	100	133	60	4165
6	120	148	60	1935	12	120	153	60	5036

Table 17.1: The reduced 1h scanning protocol, as applied to four different subjects.

17.1 METHODS

THE Achieva+ PROTOCOL (OVER 8 SESSIONS): The protocol used here is as in Experiment 3 (Section 10.1), containing 32 shells with a b_{max}=10,308 s/mm^2. However, rather than covering the protocol over two sessions, here we will use the dataset acquired over eight sessions, to check for reproducibility of the model ranking.

The following three ROI were defined on this 8-session Achieva+ acquisition:

- ACH-8-genu includes four voxels in the centre of the mid-sagittal volume, with no evidence of CSF partial volume;
- ACH-8-midbody similarly includes four mid-sagittal midbody voxels;
- ACH-8-splenium includes four mid-sagittal splenium voxels;

THE Achieva− PROTOCOL: This protocol aimed to cover a subset of the previous measurement space in just 1h, through combinations of δ = {8, 13}ms, Δ = {20, 40, 60, 80, 100, 120}ms, and $|\mathbf{G}|$ = {60}mT/m, with a maximum b-value of ~5,000 s/mm^2. TE varied from 54

to 153ms. There were ten 4mm-thick mid-sagittal slices, with 2mm x 2mm in-plane resolution. Each shell had 45 directions, with five preceding and nine interwoven b=0 acquisitions, giving 708 measurements in total. On this reduced 1h protocol, we scanned two males and two females, aged 25-33 yrs.

These four datasets come from the *Achieva−* acquisitions of the four different healthy subjects. We choose only the genu region, as it is the thickest and least susceptible to artefacts.

- ACH-subject-1 is the signal from four mid-sagittal genu voxels from the first subject;
- ACH-subject-2 is the signal from four mid-sagittal genu voxels from the second subject;
- ACH-subject-3 is the signal from four mid-sagittal genu voxels from the third subject;
- ACH-subject-4 is the signal from four mid-sagittal genu voxels from the fourth subject;

17.2 RESULTS

T_2 EFFECTS: Figure 17.1 shows the T_2 estimation across the CC of the 8x1h *Achieva+* dataset. There is a similar trend as with the 2x4h dataset, shown in fig.16.2, but here the genu T_2 is higher, at about 73ms.

Figure 17.2 shows the genu T_2 estimation in the four different subjects, scanned with the reduced protocol. The average T_2 of the first two subjects (with the least corrupted datasets) were similar, ~54 and 59ms, respectively, in turn being similar to the estimate from ACH-*genu* and ACH8-*genu*. The T_2 extracted from the other two subjects were severely affected by motion artefacts; the wide distribution of T_2 values across the four neighbouring voxels reflects this (compared with the first two subjects, whose T_2 distribution are reasonably tight about the mean).

MODEL RANKING Figure 17.3 gives the ranking of the models to ACH8-*genu*, ACH8-*midbody* and ACH8-*splenium* datasets. Bingham models here are best, but the picture changes a little from fig.16.3; e.g. Tensor with Sticks and Cylinder are no longer top-ranking in the genu data, whereas in the midbody, Bingham models appear increasingly better compared with other models.

Figure 17.4 shows the reproducibility of the ranking obtained above using a reduced scanning protocol on four different healthy subjects. The ranking of models in the first healthy subject is the same as that obtained from ACH-*genu* and ACH8-*genu*. This provides another intra-subject ranking confirmation, as all three datasets correspond to the same person. The data from the second healthy subject produces similar trends: Tensor models are best, especially those with different compartmental diffusivities. As mentioned above, the last two datasets had severe motion artefacts which may have compromised the data, hence there is no clear ranking of the models.

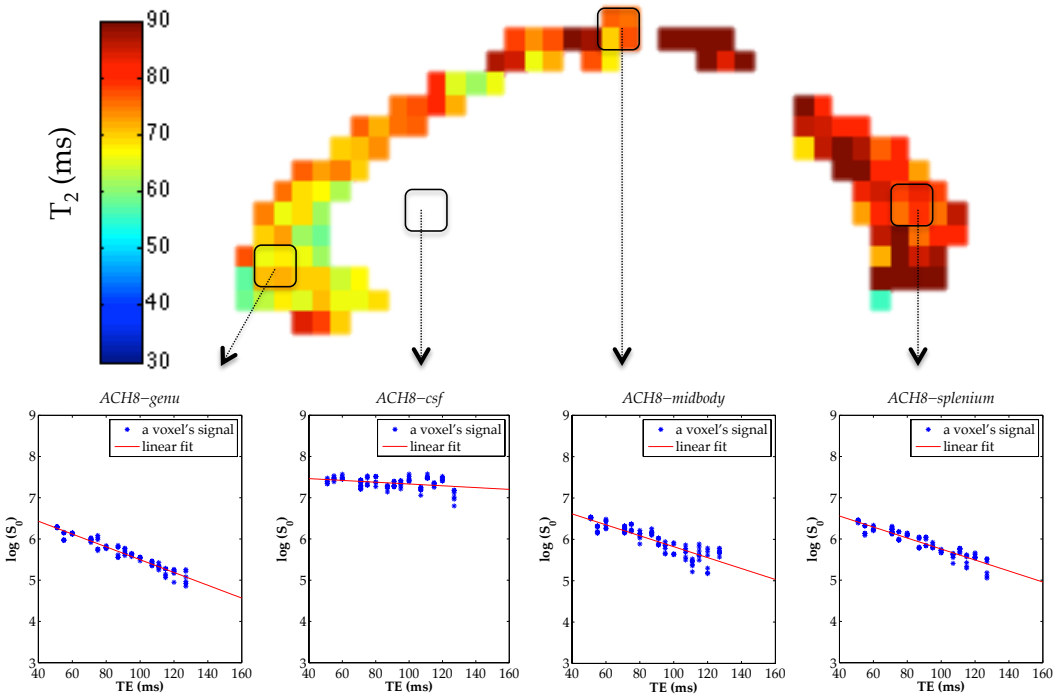


Figure 17.1: Similarly to fig.16.2, this figure shows the T_2 map across the CC for the "2x8h" protocol. The T_2 estimates are visibly higher than in the 2-session scanning of fig.16.2, by up to 14ms in the genu.

PARAMETER STABILITY The estimates from ACH8-genu, as given in table 17.2, show reasonable agreement with the estimates from ACH-genu in table 16.1. As can be seen from Appendix Table D.4, noticeable differences start to appear when using data from the region of midbody, between ACH-midbody and ACH8-midbody, especially in the axial diffusivity. Considerable changes appear in the Tensor volume fraction in the splenium, but Ball compartment estimates are more resilient. These increasing changes are expected as, relative to the genu, the midbody and splenium are more affected by motion artefacts and CSF partial volume. The CSF volume estimate was, on average, 5-10% of the total volume fraction.

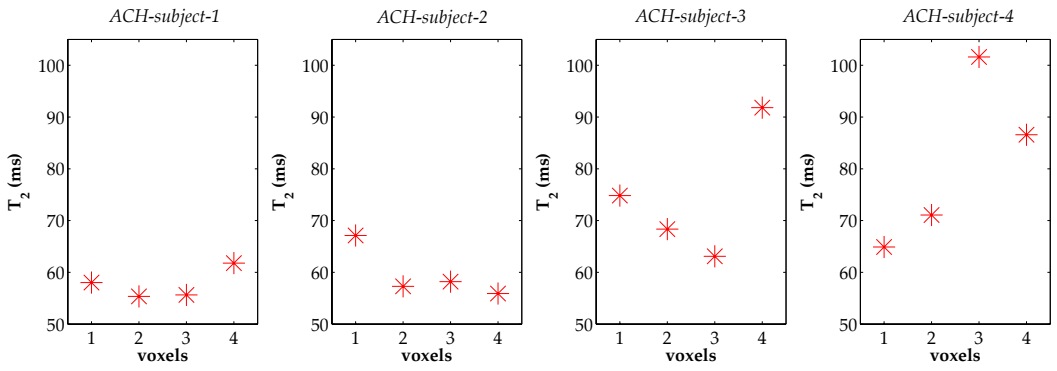


Figure 17.2: The T_2 map across the four different subjects. The wider distribution of estimates in the last two subjects comes as a result of larger motion artefacts.

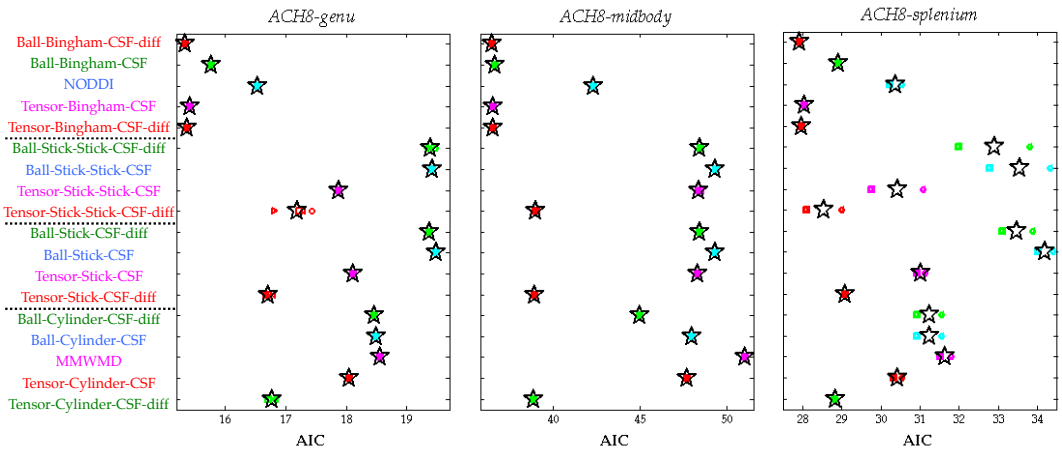


Figure 17.3: Model ranking for the main 8x1hr Achieva+ dataset (fig.16.3 concerns the 2x4hr dataset). Overall, a similar trend persists across all dataset rankings: combinations of Tensor with Cylinder or Bingham are best, whereas those of Ball with Cylinder and Sticks are worst.

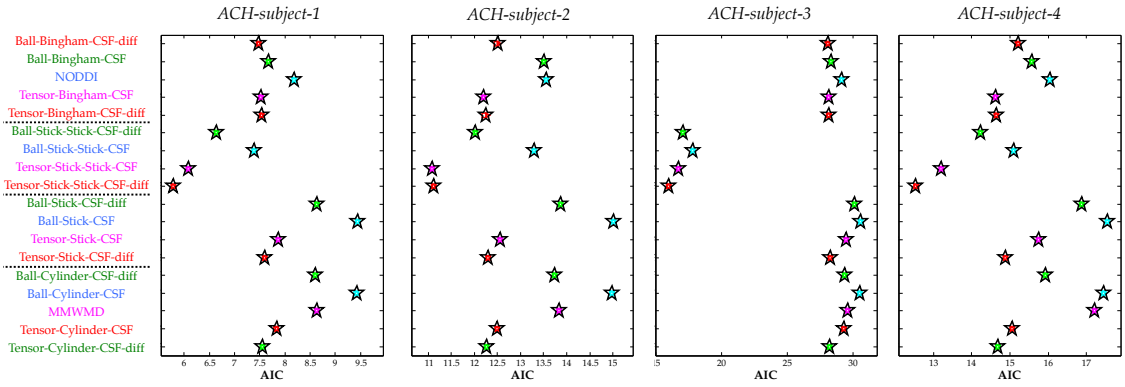


Figure 17.4: Model ranking across four healthy subjects using a reduced protocol. The trends obtained from 16.3 are repeated well across all subjects.

17.3 DISCUSSION

This experiment showed that splitting the 8h scanning protocol into multiple sessions reproduced the T_2 pattern across the CC: there was a lower estimate in the genu than the splenium and midbody. The estimate itself was higher in the eight-session dataset; e.g. by 14ms in the genu. Though these T_2 estimates are indicative of the white matter tissue values, our experiment was not designed for this purpose: limitations include the long acquisition time combined with the relatively large voxel size (2mm x 2mm x 4mm), which in turn increases the possibility to incur motion artefacts, hence more CSF contamination, and more instability in the estimation of T_2. (We discuss these more in sections 16.3 and 23.2.)

The ranking pattern is largely reproduced across scans (8h protocol over 2 sessions, and the same protocol over 8 sessions) and protocols (including scans over a reduced 1h protocol), with some differences. The Bingham models are best for the two large 8h datasets, but models

ACH8-genu

Models	LSE (x 10^3)	Cylinder/Bingham/Stick Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	*Tensor/ Ball* Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
Ball-Bingham-CSF-diff	1.52 $^{0.0}$	0.75 $^{0.0}$	*2.61* $^{0.0}$	15 $^{0.0}$		0.25 $^{0.0}$	*1.75* $^{0.0}$		
Ball-Bingham-CSF	1.56 $^{0.0}$	0.76 $^{0.0}$	*2.38* $^{0.0}$	15 $^{0.0}$		0.24 $^{0.0}$	*2.38* $^{0.0}$		
NODDI	1.65 $^{0.0}$	0.59 $^{0.0}$	*1.90* $^{0.0}$	40 $^{0.0}$		0.41 $^{0.0}$	*1.90* $^{0.0}$	*0.78* $^{0.0}$	
Tensor-Bingham-CSF	1.52 $^{0.0}$	0.72 $^{0.0}$	*2.38* $^{0.0}$	18 $^{0.0}$		0.28 $^{0.0}$	*2.38* $^{0.0}$	*1.81* $^{0.0}$	*1.47* $^{0.0}$
Tensor-Bingham-CSF-diff	1.51 $^{0.0}$	0.75 $^{0.2}$	*2.57* $^{0.6}$	16 $^{1.4}$		0.25 $^{0.6}$	*1.86* $^{1.9}$	*1.86* $^{1.9}$	*1.58* $^{2.4}$
Ball-Stick-Stick-CSF-diff	1.93 $^{0.3}$	0.62 $^{1.0}$	*1.98* $^{0.4}$			0.38 $^{1.6}$	*1.63* $^{1.6}$		
Ball-Stick-Stick-CSF	1.93 $^{0.0}$	0.63 $^{0.1}$	*1.87* $^{0.1}$			0.37 $^{0.2}$	*1.87* $^{0.1}$		
Tensor-Stick-Stick-CSF	1.77 $^{0.0}$	0.42 $^{0.2}$	*1.72* $^{0.0}$			0.55 $^{0.2}$	*1.72* $^{0.0}$	*0.31* $^{0.3}$	*0.22* $^{0.3}$
Tensor-Stick-Stick-CSF-diff	1.70 $^{1.3}$	0.52 $^{4.1}$	*1.47* $^{2.7}$			0.48 $^{4.1}$	*2.99* $^{0.2}$	*0.91* $^{8.0}$	*0.78* $^{12.6}$
Ball-Stick-CSF-diff	1.92 $^{0.0}$	0.63 $^{0.0}$	*1.98* $^{0.0}$			0.37 $^{0.0}$	*1.70* $^{0.0}$		
Ball-Stick-CSF	1.93 $^{0.0}$	0.63 $^{0.0}$	*1.88* $^{0.0}$			0.37 $^{0.0}$	*1.88* $^{0.0}$		
Tensor-Stick-CSF	1.79 $^{0.0}$	0.57 $^{0.0}$	*2.00* $^{0.0}$			0.43 $^{0.0}$	*2.00* $^{0.0}$	*1.25* $^{0.1}$	*0.98* $^{0.1}$
Tensor-Stick-CSF-diff	1.65 $^{0.3}$	0.51 $^{1.6}$	*1.54* $^{1.0}$			0.49 $^{1.6}$	*3.00* $^{0.0}$	*0.97* $^{6.6}$	*0.78* $^{3.2}$
Ball-Cylinder-CSF-diff	1.83 $^{0.0}$	0.75 $^{0.0}$	*1.90* $^{0.0}$		5.4 $^{0.0}$	0.25 $^{0.0}$	*2.19* $^{0.0}$		
Ball-Cylinder-CSF	1.84 $^{0.0}$	0.73 $^{0.0}$	*1.95* $^{0.0}$		5.2 $^{0.0}$	0.27 $^{0.0}$	*1.95* $^{0.0}$		
MMWMD	1.85 $^{0.0}$	0.54 $^{0.1}$	*1.70* $^{0.0}$		0.2 $^{57.1}$	0.46 $^{0.1}$	*1.70* $^{0.0}$	*0.77* $^{0.1}$	
Tensor-Cylinder-CSF	1.79 $^{0.0}$	0.60 $^{0.0}$	*2.00* $^{0.0}$		3.4 $^{0.1}$	0.40 $^{0.0}$	*2.00* $^{0.0}$	*1.30* $^{0.1}$	*1.01* $^{0.1}$
Tensor-Cylinder-CSF-diff	1.66 $^{0.3}$	0.50 $^{2.8}$	*1.51* $^{2.5}$		0.7 $^{121.4}$	0.50 $^{2.8}$	*3.00* $^{0.0}$	*0.88* $^{4.4}$	*0.76* $^{9.2}$
NODDI+	1.55 $^{0.0}$	0.71 $^{0.0}$	*2.19* $^{0.0}$	15 $^{0.0}$		0.29 $^{0.0}$	*2.19* $^{0.0}$	*1.00* $^{0.0}$	
MMWMD+	1.69 $^{0.0}$	0.61 $^{0.0}$	*1.93* $^{0.0}$		3.6 $^{0.0}$	0.39 $^{0.0}$	*1.93* $^{0.0}$	*0.78* $^{0.0}$	

Table 17.2: Parameter estimates from model fitting to the full dataset ACH8-genu. While the CSF volume is almost zero, the estimated ROI T_2 is 73ms. (Estimates for ACH8-midbody and ACH8-splenium are in the Appendix, fig.D.4.)

with two-Sticks come out on top on the reduced datasets for the four subjects. Cylinder and Bingham models seem to benefit most from datasets of a broader b-values range.

As with the ACH-genu dataset, in ACH8-genu the Bingham distribution of fibres appeared isotropic (with the second dispersion concentration parameter close to zero). This suggests that a Watson distribution (equivalent to an isotropic Bingham distribution) may be sufficient for capturing fibre dispersion, at least in the genu.

While the parameters were also largely reproducible, the Bingham distribution reflected a higher dispersion of fibres in the multiple session scanning (higher κ in fig.16.1 than fig.17.2). Indeed, even in the scans of four healthy subjects, the Bingham captures the greater heterogeneity introduced by motion artefacts (higher κ in fig.D.1 than fig.D.2).

18

SUMMARY

In these two experiments, we investigated the compartmental T_2. From our acquired data, we could not justify different T_2 in each intra/extracellular compartment, but found that a separate T_2 for CSF is necessary. The estimate of T_2 varies across the CC, being lower in the genu than in the splenium, and highest in the midbody.

As explained in chapter 16.3, the best representative of white matter in our datasets is the genu, being relatively large and least affected by registration artefacts. The analysis on the genu ROI reveals that, broadly, models with an intracellular Bingham distribution, or an anisotropic extracellular Tensor, rank highest, whereas those with isotropic extracellular Ball and intracellular Sticks rank lowest. The ranking from splenium provided similar results, but midbody data was less discriminatory of model performance.

Beyond the model ranking, we investigated the stability of the parameters. The analysis provides more realistic volume fraction estimates when the Tensor models' intra/extracellular diffusivities are fixed to be the same. Relaxing this assumption confounds the description and assignment of intracellular and extracellular compartment volume fractions and diffusivities. The results from splenium show more variability in Cylinder radius estimation.

Beyond the model ranking, we also investigated the inter- and intra-subject reproducibility of the ranking and parameters. There is good reproducibility of ranking across different scans of the same person, but the results are more robust for the area of genu, whose ROI is less affected by CSF partial volume and motion artefacts. The same applies to the parameters: the reproducibility of the parameters is heavily affected by artefacts, especially in two of the four subjects scanned with the reduced protocol.

Part VI

HIGHER GRADIENTS

19

INTRODUCTION

To restate our aim, we are interested in finding a model which can infer the most information about the microstructure from the tissue signal. We have aimed at maximising the richness of the dataset by sampling densely across the experimental tuneables. To this end, recent improvements in the scanner technology give the potential of learning more about the models, and hence the tissue.

In the work we have done so far, we have used the most recent and modern scanners intended for clinical use. However, as mentioned at the very start of this work, standard clinical scanners, using gradient strengths of 40-60 mT/m, cannot ordinarily afford the gradient strengths found in scanners for fixed tissue or animals that use gradients reaching 300-1,000 mT/m. In diffusion weighted imaging, higher gradients provide higher contrast in the diffusion/dispersion of water molecules; this is very beneficial to the model fitting, and invaluable to this kind of work.

The recent development of human MR systems with 300mT/m gradients, in particular the MGH-UCLA Connectom scanner [Setsompop et al., 2013], aims at mapping through diffusion tractography the structural connections in the live human brain. But this also provides an opportunity for, and is a major step towards, the long-term translation of microstructure imaging techniques.

We also use this scanner to acquire a similarly rich dataset to the one obtained in the previous experiments. Over 8h of scanning, we collect a rich dataset with which to test all the models, and compare the results with those inferred from data acquired with more standard clinical scanners.

20

EXPERIMENT 7: CONNECTOM-SKYRA DATA

Here we explore the generalisability of earlier model comparison results, which use standard human scanners with 60mT/m gradients, to the wider measurement space of human data accessible with the Connectom scanner. We construct a multi-shell HARDI protocol for the Connectom scanner with a wide range of b-values and diffusion times. It is similar in spirit and size to that acquired of Experiment 3 (Section 10.1 and Ferizi et al. [2013a]) but exploits the wider measurement space afforded by the 300mT/m gradients. We concentrate on the set of models used in previous work [Ferizi et al., 2013a,b; Panagiotaki et al., 2012] and Experiments 3/4/5/6 that perform consistently well and compare model rankings between the Connectom datasets and data from Experiments 5 and 6 [Ferizi et al., 2013a] using the AIC.

20.1 METHODS

We describe below the acquisition protocols, which are also summarised in table 20.1. Ethical approval and written consent were obtained prior to scanning for all subjects.

THE Connectom PROTOCOL: This acquisition used the Massachusetts General Hospital Magnetom Skyra Connectom (Siemens Healthcare) scanner, which has a novel AS302 gradient system with a custom-built 64-channel coil, capable of $|\mathbf{G}|$ =300 mT/m and a slew rate of 200 T/m/s. A PGSE [Tanner and Stejskal, 1968] sequence was used, with GeneRalized Autocalibrating Partially Parallel Acquisitions (GRAPPA) parallel imaging, an acceleration factor of 2, cardiac gating and TR=1s. The protocol contains 48 HARDI shells, each with 90 directions (45 unique pairs of opposite directions), and ten interwoven b=0 acquisitions, for a total of 100 measurements per-shell. Each shell has a unique combination of:

- gradient strength $|\mathbf{G}|$ = {60, 100, 200, 300} mT/m;
- pulse width δ = {3, 8} ms;
- pulse duration Δ = {20, 40, 60, 80, 100, 120} ms.

The maximum b-value, therefore, is 46,000 s/mm^2. Each shell uses the minimum TE possible for the combination of δ and Δ; TE thus ranges from 49 to 152ms.

For this protocol, the same healthy subject as in the Achieva+ acquisition was scanned over two 4h non-stop sessions. The imaged volume comprises twenty 4mm-thick whole-brain sagittal slices covering the CC left-right. The image size was 550 x 550 and the in-plane resolution is 2mm x 2mm. The SNR of b=0 images is about 35 at TE=49ms and 6 at TE =152ms. The four shells with δ = 8ms and Δ = 60ms were corrupted so these were omitted from the analysis.

Connectom Protocol

	δ=3ms				δ=8ms				
Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)	Nr	Δ (ms)	TE (ms)	|G| (mT/m)	b (s/mm²)
1	23	49	60	50	25	21	58	60	300
2	17	49	100	100	26	20	58	100	800
3	20	49	200	500	27	20	58	200	3200
4	20	49	300	1100	28	19	58	300	6700
5	44	67	60	100	29	39	72	60	600
6	40	67	100	250	30	40	72	100	1700
7	40	67	200	1000	31	40	72	200	6850
8	37	67	300	2100	32	38	72	300	14550
9	66	87	60	150	*Omitted*	60	92	60	950
10	63	87	100	400		61	92	100	2650
11	59	87	200	1500		60	92	200	10500
12	56	87	300	3200		57	92	300	22350
13	87	107	60	200	37	82	112	60	1300
14	79	107	100	500	38	80	112	100	3550
15	81	107	200	2050	39	80	112	200	14150
16	75	107	300	4300	40	76	112	300	30200
17	109	127	60	250	41	100	132	60	1600
18	102	127	100	650	42	100	132	100	4450
19	100	127	200	2550	43	100	132	200	17850
20	94	127	300	5400	44	95	132	300	38050
21	130	147	60	300	45	121	152	60	1950
22	117	147	100	750	46	119	152	100	5350
23	119	147	200	3050	47	120	152	200	21500
24	113	147	300	6500	48	114	152	300	45900

Table 20.1: The Connectom scanning protocol, which required two 4h sessions. Four corrupted shells, shown as *Omitted*, were excluded from the analysis.

DATA PREPROCESSING All images were corrected for eddy current distortions and co-registered using FSL Flirt [Jenkinson et al., 2002] and *Eddy* module (*www.fmrib.ox.ac.uk/fsl/eddy*). Across the CC of all datasets, and after ensuring our selection was away from partial volume effects and well-embedded in WM, a neighbourhood of five voxels[1], was selected from the middle of the genu, midbody and splenium.

DATASETS Four similar regions of interest from the Connectom acquisition were defined:

- CON-genu has four mid-sagittal genu voxels;
- CON-midbody has four mid-sagittal midbody voxels;
- CON-splenium has four mid-sagittal splenium voxels;
- CON-csf has four mid-sagittal ventricular CSF voxels.

We also consider several subsets of the CON-genu data to explore the influence of measurements with different gradient strengths:

- CON-genu-G60 retains only shells with $|\mathbf{G}|$=60mT/m, with $b_{max} \sim 2,000 s/mm^2$;

1 To do this, we used the Stick volume fraction and direction from the Ball-Stick model fitting.

- CON-genu-G100 retains only shells with $|\mathbf{G}|$=100mT/m, with $b_{max} \sim 5,000 s/mm^2$;
- CON-genu-G200 retains only shells with $|\mathbf{G}|$=200mT/m, with $b_{max} \sim 22,000 s/mm^2$;
- CON-genu-G300 retains only shells with $|\mathbf{G}|$=300mT/m, with $b_{max} \sim 46,000 s/mm^2$.

20.2 RESULTS

Figure 20.1 shows the diffusion-weighted images with gradient in the direction perpendicular to the fibres of the CC. We see that signal persists even at the highest diffusion weighting.

Figure 20.2 shows the CON-genu dataset, split into the four gradient strengths used, together containing 4,411 measurements. The measurements provide good coverage of the signal range. Anisotropy is apparent even at the highest b-value.

T_2 EFFECTS: Figure 20.3 shows the b=0 log-signal against TE for genu, midbody, splenium and CSF. All plots indicate an approximately mono-exponential model. As in Experiment 6, we here assume that the intracellular and extracellular T_2 are equal. In CON-genu, voxel-wise T_2 averaged about 57ms. There was some variation across CC, with T_2 in CON-midbody, at ~ 67ms, and lower for CON-splenium, at ~ 60ms. The T_2 estimate from CON-csf data averaged about 600ms.

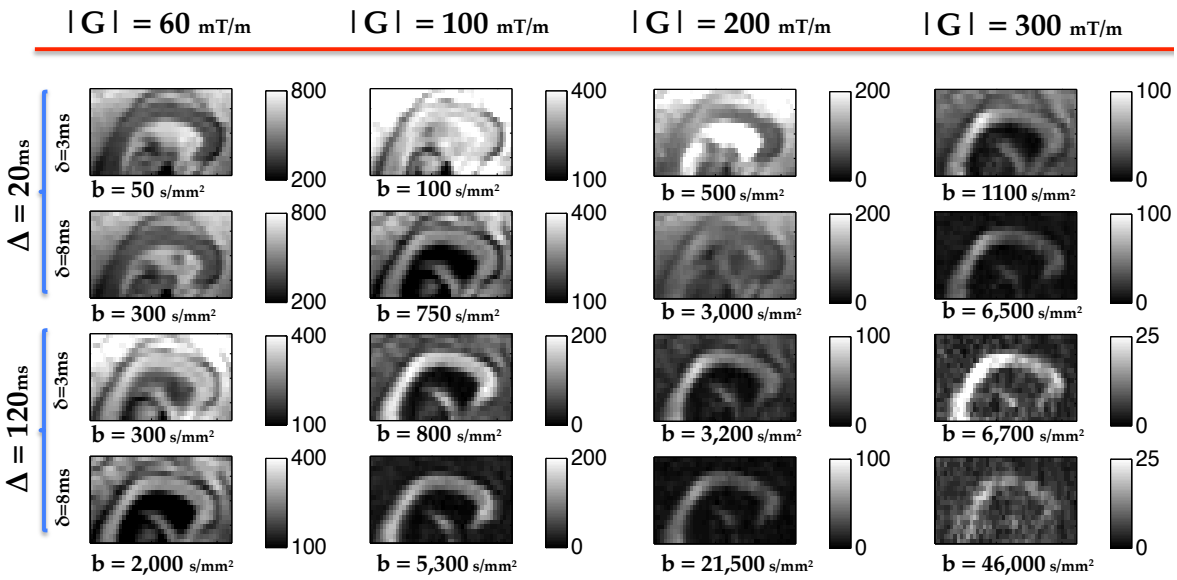


Figure 20.1: Images of the CC showing the signal for gradient direction perpendicular to the fibres, at each $|\mathbf{G}|$, for each pulse time δ, but only for the smallest and largest diffusion times Δ. The grey scale is adjusted in each case so as to give a reasonable contrast between the CC and the background. Signal still persists even at $b \sim 46,000 s/mm^2$.

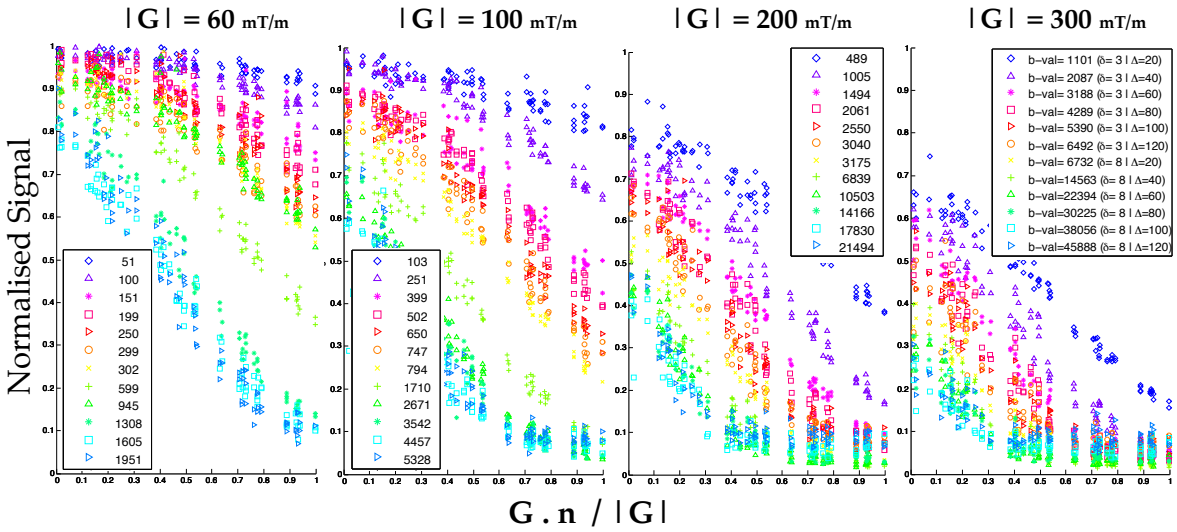


Figure 20.2: The CON-genu dataset, consisting of averaged signal from four voxels in the middle of the genu; the models, however, are fitted to raw data. For clarity, the signal is split across each gradient strength. The legend gives the b-value in units of s/mm^2 and, in the last plot, also diffusion and pulse times ($\delta \mid \Delta$), which are the same across the four plots, in units of ms. **G** is the applied gradient vector and **n** is the fibre direction; the x-axis gives the cosine of the angle between the applied gradient and fibre direction: near 0 to the left, the gradient is perpendicular to the fibres; as it approaches 1 to the right, the gradient direction becomes parallel with the fibre.

MODEL RANKING: Figure 20.4 shows the performance of the models fitted to various datasets. As in Experiment 5 (see Results 16.2), we find that the best models from CON-genu are the Tensor and/or Bingham models, and last are the combinations of Ball with Stick/s.

Compared with ACH-genu (fig.16.3) results, this performance is reproduced in the CON-genu ranking. The differences norrow in the CON-midbody and CON-splenium data. This is expected, as the region is more susceptible to motion artefacts and CSF partial volume.

Figure 20.6 shows that the models become more distinct as gradients increase. The ranking obtained from the first dataset CON-genu-60 is less informative than the ranking obtained from other |**G**|-specific datasets, CON-genu-100 / -200 / -300.

Figure 20.5 shows the fit of four models to CON-genu signal (Appendix fig.E.1 shows similar plots for the midbody and splenium). As with fig.16.4 of Experiment 5, we pick two Tensor and two Ball models: Ball-Bingham-CSF-diff, Tensor-Cylinder-CSF-diff, Tensor-Stick-CSF and Ball-Stick-CSF-diff. The model signal is shown as solid line, and the raw data is shown with markers. The plots reveal that the higher b-value shells of CON-genu data enable a greater differentiation between the models than observed in fig.16.4: Bingham models cannot capture the largest b-value shells, but they capture better than Cylinder the less-restricted signal.

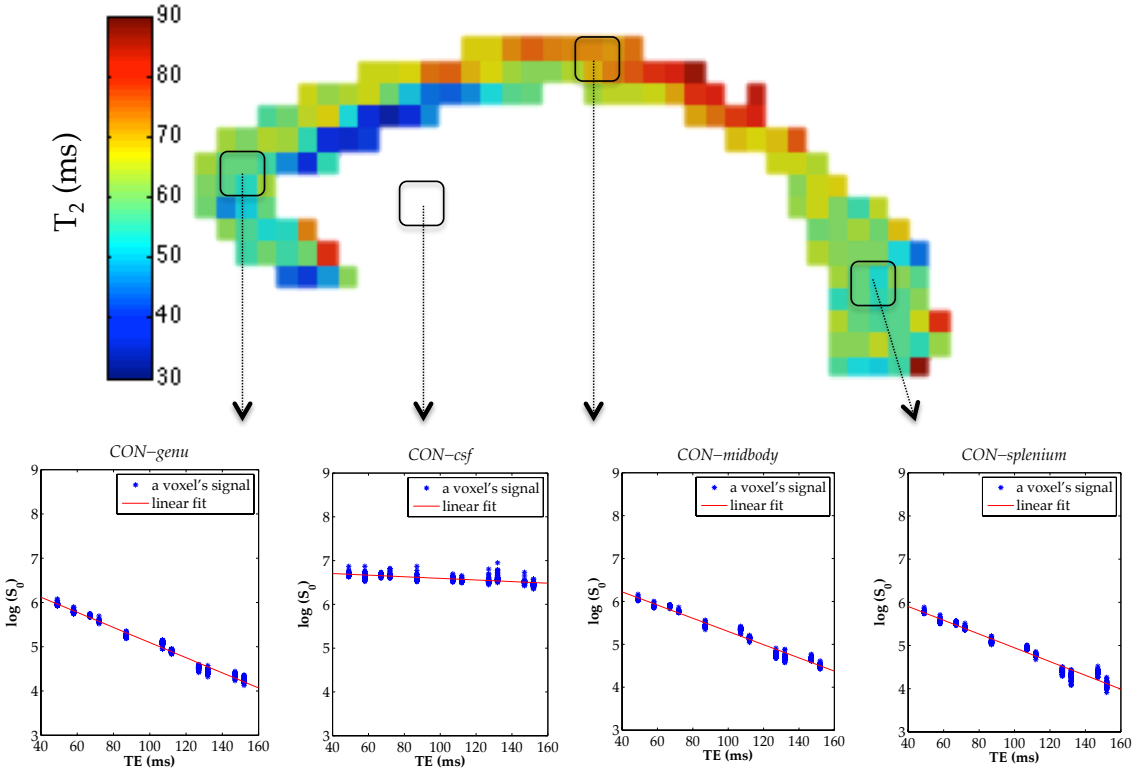


Figure 20.3: As in fig.16.2, here we show the T_2 map of the CC (centre, in colour) for the Connectom data. There are 40 b=0 signals for each of the 12 TEs, marked with blue stars. To the right are all four ROI voxels' T_2 values, starred in red.

PARAMETER STABILITY Table 20.2 shows the mean CON-genu parameter estimates and, in small superscript, their standard deviations gives as the percentage of the mean. The CSF volume fraction is constantly around zero, whereas T_2 calculation shows some variability. The parameters reflect higher intracellular volume fractions, except for Tensor and Stick combinations. The single axial diffusivity is higher in Bingham than in Cylinder models. In models with separate intra/extracellular diffusivities, Ball diffusivity is lower than Bingham/Sticks diffusivity, but the opposite applies to the other models. There is more dispersion in Ball versus Tensor models as the model compensates for of the extracellular isotropy.

Among the best fitting models, separate compartmental diffusivities make little difference to Tensor-Bingham combinations, but a considerable improvement to Tensor-Cylinder and Ball-Bingham models. As in other datasets, separating compartmental diffusivities generally makes the volume fractions closer. Thus, although the separate diffusivity models explain the data better, the parameter estimates, at least the volume fractions, are less consistent with what we might expect.

For the same reasons stated in Discussion 16.3, the genu ROI is a more reliable region than the corresponding ones in the splenium and midbody. So, we leave to Appendix table E.1 parameter estimates for the splenium and midbody. Nevertheless, some parameters in CON-splenium, such as compartmental volume fractions, show slight similarity with those of CON-

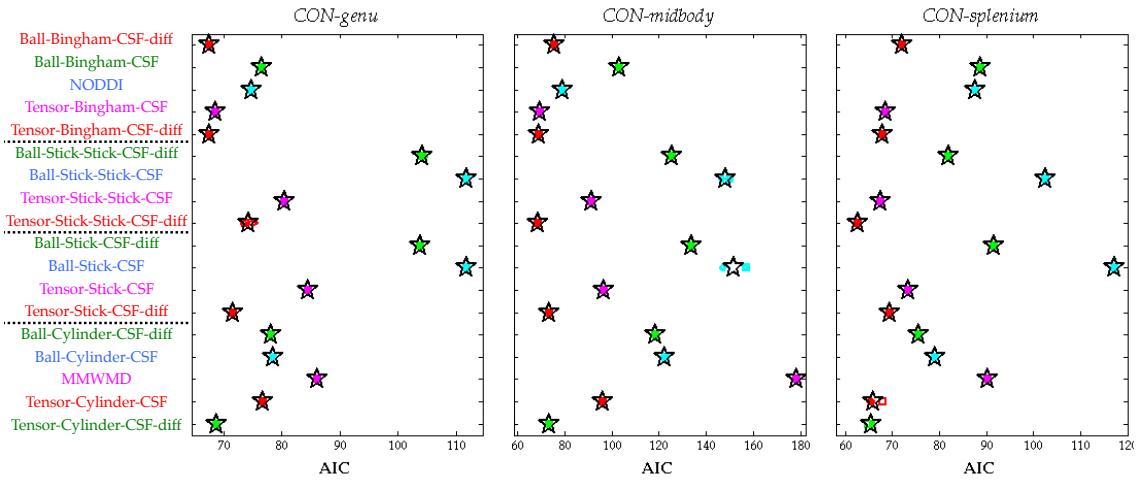


Figure 20.4: Model ranking for each dataset. Overall, a similar trend persists across all dataset rankings: combinations of Tensor with Cylinder or Bingham are best, whereas those of Ball with Cylinder and Sticks are worst.

genu. The diffusivities are more different, especially Tensor radial diffusivities. An unexpected result arises from radii estimation, which are highest in the splenium; however, there is a high variance of the estimate across the ROI voxels. These values are inconsistent with known CC trends.

The more heterogenous CON-midbody dataset produces more dispersion and CSF volume than CON-genu. The Cylinder radii estimates are slightly higher than for genu, which is expected. With the exception of most separately-fitted intracellular diffusivities, all other diffusivity estimates are higher compared with CON-genu. As with ACH-midbody data, this region produces the opposite trend of high-extracellular/low-intracellular diffusivities for Ball models.

Table E.2 shows the dependence of parameter estimates to increasing gradients of the Connectom data. As mentioned earlier, due to differences in the protocol, the model parameters drawn from CON-genu-60 do not provide immediate similarity to the estimates to ACH-genu. In the higher gradient datasets of CON-genu-200 and CON-genu-300 the stability of the T_2 estimates decreased, so we fixed the WM T_2 for these datasets to 57ms. Cylinder models aside, as gradients increase, all Ball compartment volume fractions increase, while the Tensors becomes more anisotropic.

20.3 DISCUSSIONS

Since both scanners used in the study are of same B_0 field strength, 3T, we would expect the T_2 in all datasets CON-genu, ACH-genu, ACH8-genu and ACH-subject-1 to be approximately the same; they respectively averaged 59, 57, 73, and 54ms. The highest estimate of 73ms could

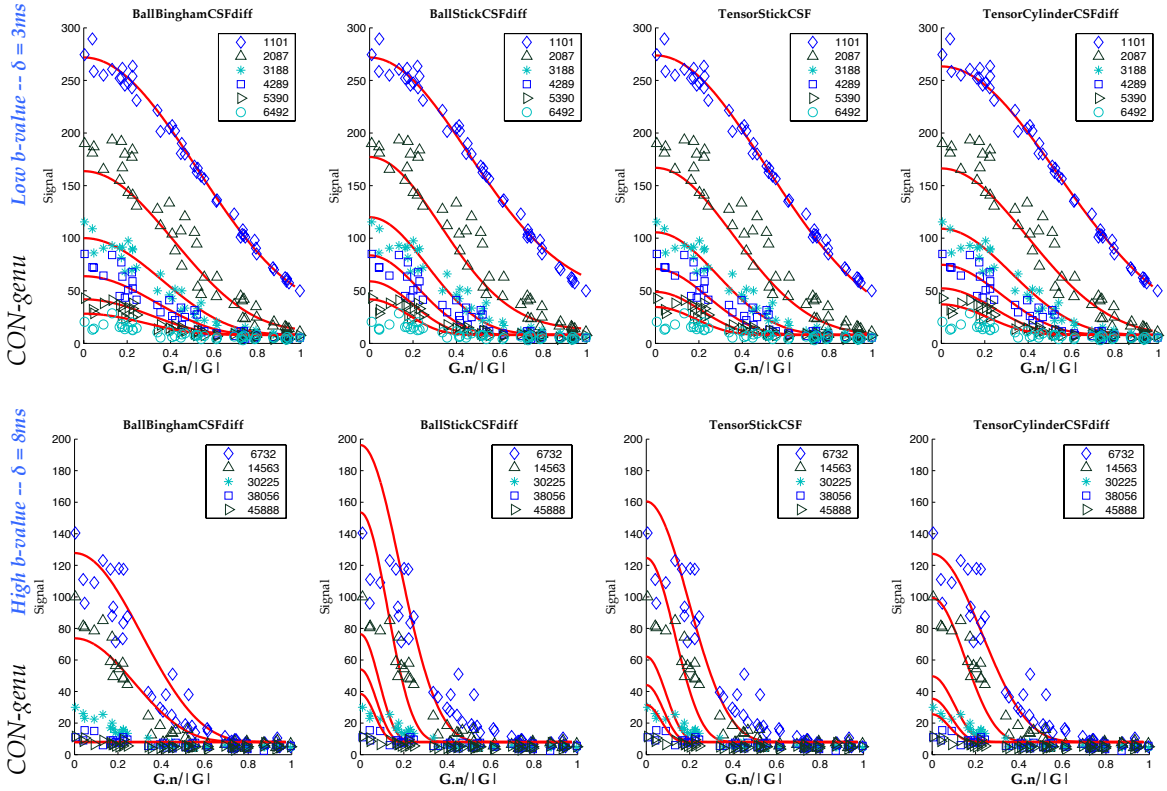


Figure 20.5: Plots illustrating the quality of fit for four selected models fitted to CON-genu. Only one voxel is selected from the genu ROI, whose raw signal is shown with markers and the model signal shown as solid line. For clarity, the plots show only a few selected high/low b-value shells, of $|\mathbf{G}|$=300mT/m. The six Δ-specific shells increase in value from top-bottom. Bingham fails with the largest b-value signal, but is better for less restricted signal.

be attributed to the multi-sessional method of image acquisition, whereas the other estimates are acceptably similar, within 4% deviation from their mean.

The CON-genu-60 ranking, in fig.20.6, is less informative when compared with the ranking from ACH-genu in fig.16.3 and ACH8-genu in fig. 17.3. One possible explanation can be that, though the gradients are similar, the data of ACH-genu and ACH8-genu contain a wider measurement space of b-values, respectively $\sim 10,000 s/mm^2$ vs. $2,000 s/mm^2$ of CON-genu-60, including longer pulse durations δ.

The model fit illustrations to CON-genu data, in fig.20.5, showed that models with a Bingham compartment fail to include the largest b-value shells in their modelling. This result may suggest that the dispersion is microscopic rather than macroscopic [Nilsson et al., 2012], i.e. the signal arises from small fluctuations/undulations in a single fibre orientation rather than a small population of fibres with totally different orientation. Another possible explanation is that the Bingham distribution tails off too slowly, i.e. the true distribution has a stronger peak around the mean.

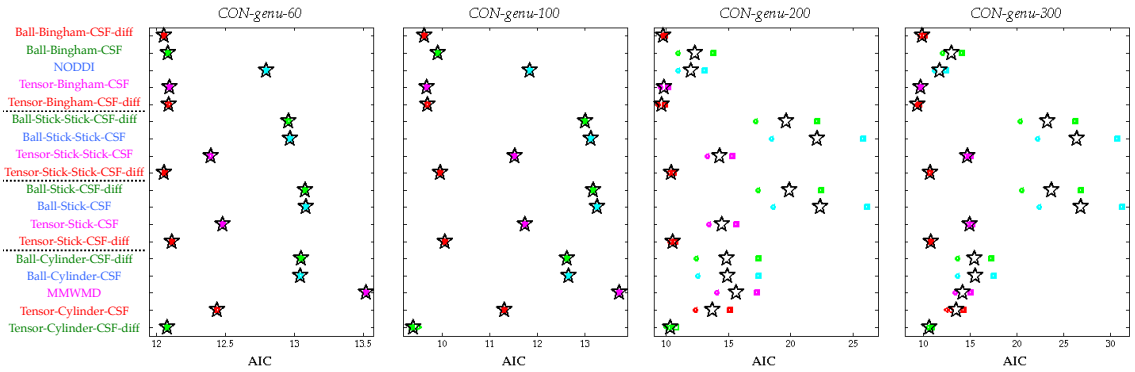


Figure 20.6: Ranking for each |**G**|-specific dataset. Higher |**G**| improves the performance of Tensor and Cylinder models, and increase the difference between Ball with Stick/s models with the rest.

			Cylinder/Bingham/Stick				*Tensor/ Ball*			
	Models	LSE (x 10^3)	Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
CON-genu	Ball-Bingham-CSF-diff	6.74 $^{0.0}$	0.75 $^{0.0}$	*3.00* $^{0.0}$	14 $^{0.1}$		0.25 $^{0.0}$	*0.97* $^{0.0}$		
	Ball-Bingham-CSF	7.64 $^{0.0}$	0.80 $^{0.1}$	*2.41* $^{0.1}$	13 $^{0.4}$		0.20 $^{0.2}$	*2.41* $^{0.1}$		
	NODDI	7.47 $^{0.0}$	0.57 $^{0.0}$	*1.90* $^{0.0}$	24 $^{0.0}$		0.43 $^{0.0}$	*1.90* $^{0.0}$	*0.82* $^{0.0}$	
	Tensor-Bingham-CSF	6.85 $^{0.0}$	0.67 $^{0.0}$	*2.27* $^{0.0}$	20 $^{0.1}$		0.33 $^{0.0}$	*2.27* $^{0.0}$	*0.97* $^{0.0}$	*0.90* $^{0.0}$
	Tensor-Bingham-CSF-diff	6.74 $^{0.0}$	0.75 $^{0.1}$	*3.00* $^{0.1}$	14 $^{1.2}$		0.25 $^{0.3}$	*0.97* $^{0.7}$	*0.97* $^{0.9}$	*0.96* $^{0.3}$
	Ball-Stick-Stick-CSF-diff	10.41 $^{0.2}$	0.60 $^{0.6}$	*2.24* $^{3.2}$			0.40 $^{0.9}$	*1.16* $^{1.9}$		
	Ball-Stick-Stick-CSF	11.17 $^{0.0}$	0.63 $^{0.2}$	*1.75* $^{0.4}$			0.37 $^{0.4}$	*1.75* $^{0.4}$		
	Tensor-Stick-Stick-CSF	8.03 $^{0.0}$	0.31 $^{0.3}$	*1.66* $^{0.0}$			0.63 $^{0.1}$	*1.66* $^{0.0}$	*0.19* $^{0.2}$	*0.11* $^{0.2}$
	Tensor-Stick-Stick-CSF-diff	7.41 $^{1.0}$	0.40 $^{1.9}$	*1.09* $^{3.0}$			0.60 $^{1.1}$	*3.00* $^{0.2}$	*0.53* $^{4.4}$	*0.47* $^{8.3}$
	Ball-Stick-CSF-diff	10.36 $^{0.0}$	0.60 $^{0.0}$	*2.22* $^{0.0}$			0.40 $^{0.0}$	*1.17* $^{0.0}$		
	Ball-Stick-CSF	11.16 $^{0.0}$	0.63 $^{0.0}$	*1.76* $^{0.0}$			0.37 $^{0.0}$	*1.76* $^{0.0}$		
	Tensor-Stick-CSF	8.43 $^{0.0}$	0.49 $^{0.0}$	*1.96* $^{0.0}$			0.51 $^{0.0}$	*1.96* $^{0.0}$	*0.76* $^{0.0}$	*0.58* $^{0.1}$
	Tensor-Stick-CSF-diff	7.14 $^{0.0}$	0.42 $^{2.5}$	*1.21* $^{2.5}$			0.58 $^{1.8}$	*3.00* $^{0.0}$	*0.62* $^{3.6}$	*0.47* $^{3.6}$
	Ball-Cylinder-CSF-diff	7.81 $^{0.0}$	0.77 $^{0.0}$	*1.82* $^{0.0}$		3.9 $^{0.0}$	0.23 $^{0.0}$	*2.19* $^{0.0}$		
	Ball-Cylinder-CSF	7.85 $^{0.0}$	0.75 $^{0.0}$	*1.88* $^{0.0}$		3.9 $^{0.0}$	0.25 $^{0.0}$	*1.88* $^{0.0}$		
	MMWMD	8.61 $^{0.0}$	0.55 $^{0.0}$	*1.70* $^{0.0}$		3.0 $^{0.0}$	0.45 $^{0.0}$	*1.70* $^{0.0}$	*0.77* $^{0.0}$	
	Tensor-Cylinder-CSF	7.66 $^{0.0}$	0.69 $^{0.0}$	*1.92* $^{0.0}$		3.7 $^{0.0}$	0.31 $^{0.0}$	*1.92* $^{0.0}$	*1.42* $^{0.0}$	*1.03* $^{0.0}$
	Tensor-Cylinder-CSF-diff	6.86 $^{0.1}$	0.58 $^{2.2}$	*1.49* $^{1.1}$		3.2 $^{2.4}$	0.42 $^{3.1}$	*3.00* $^{0.0}$	*0.90* $^{4.0}$	*0.67* $^{3.9}$
	NODDI+	6.90 $^{0.0}$	0.65 $^{0.0}$	*2.22* $^{0.0}$	13 $^{0.0}$		0.35 $^{0.0}$	*2.22* $^{0.0}$	*0.86* $^{0.0}$	
	MMWMD+	7.47 $^{0.0}$	0.69 $^{0.0}$	*1.91* $^{0.0}$		3.7 $^{0.0}$	0.31 $^{0.0}$	*1.91* $^{0.0}$	*1.17* $^{0.0}$	

Table 20.2: Main parameter estimates from fitting to CON-genu data.

21

SUMMARY

In this work we sampled a wide measurement space of human data accessible with the Connectom scanner and its 300mT/m gradients with the aim of determining whether similar compartment models for WM explain these unique data as more accessible data from standard systems with 60mT/m.

As in Experiment 5 (chapter 16), we investigated the compartmental T_2. Here too, we found more support for a single intra/extracellular T_2, though separate from a much higher T_2 for CSF. The T_2 estimate was highest in the midbody and splenium, and smallest in the genu.

In the genu, we saw that Tensor combined with Cylinders and Sticks are better with separate intra/extracellular diffusivities; this does not apply to Tensor with Bingham models. In particular, dispersion models generally explain data best, as LSE scores show in table 20.2, but they cannot capture Connectom signal that remains at very high b-values, as shown in fig.20.5.

As regards the parameter estimates, we obtain more realistic volume fractions when Tensor intra/extracelullar diffusivities are equal. From top models, Ball-Bingham-CSF-diff estimates a higher intracellular diffusivity, which could be plausible if, intracellularly, there would be less obstruction inside the axon than in the outside hindered space.

Between Achieva+ and Connectom datasets, the parameters were generally more reproducible in the genu. As expected, the estimation of radii is consistently lower with the Connectom data. This suggests that the higher gradient strengths of CON-genu help ameliorate the overestimation of axon diameter index, consistent with Dyrby et al. [2013].

Part VII

STABILITY OF RANKING; AND AN ALTERNATIVE MODEL SELECTION

22

INTRODUCTION

All the multi-compartment diffusion MRI models we use are non-linear. This necessitates the use of computational methods to search for their best parameters. By 'best' we mean those parameters which minimise the objective function (the error). The optimisation routine searches the parameter space of models until the desired precision is achieved. The models, however, vary in their scale of non-linearity, and so the required computation in finding this solution is not the same in each case. As explained in the *Model Fitting* of Experiment 5, we repeat our optimisation procedure 100 times, each time perturbing randomly the starting estimates. In the next chapter we will start by seeing, for each model, the probability of hitting the minimum in the 100 fitting runs.

Though we do not have ground truth measurements and do not possess many datasets against which to check our model fitting, there are techniques, such as bootstrapping and cross-validation, which investigate the variance and accuracy/bias inherent in the model for a particular dataset. Both methods are used widely in error-prediction and model selection, with some authors preferring one above the other; e.g. Kohavi [1995] identifies cross-validation as a better method for model selection; Efron and Tibshirani [1997] regard cross-validation as a low-bias high-variance method, and recommends bootstrapping instead (specifically, the .632+ method). We use elements from both methods, testing the models to variations in seen and unseen parts of the dataset. This provides a complementary evaluation of the models.

The clinical application of the models we are testing will make use of the parameters as proxy biomarkers for disease development. In using the AIC/BIC criteria we have aimed to evaluate the parsimony of the models in terms of their quality of fit to the data against model complexity. The techniques of bootstrapping and cross-validation test how robust the models are to variations in the dataset, and in predicting unseen data. However, the models need to be applicable across the brain, giving sensible indices, such as fibre density or thickness. So we examine the parameter maps of the models beyond the genu, across the whole mid-sagittal CC.

23

EXPERIMENT 8: STABILITY OF PARAMETERS

We first look at the stability of T_2 estimation in the ACH-*genu* and CON-*genu* datasets. As a reminder, T_2 is estimated in one step with S_0 and CSF volume fraction, and then we estimate the other model parameters. While CSF volume is close to zero in our selected genu ROI, a marginal variation in S_0 estimation does cause fluctuations in the T_2 estimation. In the end, for any particular model, our best solution contains the combination of all parameters which produce the minimum objective function (so two different models may not necessarily produce the same T_2 and S_0).

Assuming that 'ground truth' solution will be found within 100 model fitting runs, we refit the models 100 times to the diffusion-weighted ACH-*genu* and CON-*genu* datasets, and find the minimum number of runs needed to gives us a 99% confidence level that we have found the solution.

Next, we look at the stability of the models to noise variations in the dataset. This involves the construction of training and testing datasets for bootstrapping and cross-validation. We evaluate the relative accuracy of each model in predicting seen and unseen data.

The last part in this chapter involves mapping of the intracellular volume fraction for two of the best models, and two other simpler ones which are already used in previous studies, over the whole mid-sagittal slice of the CC. This will be an indicator of the scale of applicability of the models in the parts of the brain beyond the genu.

23.1 METHODS

DATA: For this experiment we use the ACH-*genu* and CON-*genu* datasets; the first is described in more detail in Experiment 5 (Section 16.1), the second in Experiment 7 (Section 20.1).

MODELS: The selection of models and inherent assumptions about T_2, diffusivity and fitting are as in the previous three experiments.

MODEL RANKING, AND ITS STABILITY: We use bootstrapping [Efron, 1979] and cross-validation [Stone, 1974] as complementary methods for model comparison and to further investigate the stability of the model ranking. We apply them only on the genu data sets. When bootstrapping, we sample with replacement in the original dataset the same number of measurements; we refer to this as the "seen" dataset. The distinct samples left out from the original dataset comprise the "unseen" dataset. The above sampling creates 50 seen and 50

unseen datasets for each of the five voxels in the genu, giving a total of 250 datasets to which we fit, and 250 datasets with which we test, the models. We then measure the model accuracy on the “seen" and the “unseen" datasets. The original .632 method uses a weighted average of the fitting error to the ‘seen’ data set, used to estimate the parameters, and the “unseen" data, left out during fitting, which provides a compromise between bias and variance. We instead look at these linear components separately.

We use leave-one-out cross-validation at the level of HARDI shells, i.e. at each iteration, we leave out a complete HARDI shell, fit the model to the remaining data, and use the fitted parameters to estimate all measurements in the missing shell. The final score is the average fitting error over all shells - 32 for the ACH-genu data and 44 for the CON-genu data.

23.2 RESULTS AND DISCUSSION

Figure 23.1 contains the distributions of T_2 estimates across both ACH-genu and CON-genu datasets; the distributions are cumulative over the four ROI voxels, however, as Appendix fig.F.1 shows, each voxel does exhibit a similar spread. There is a multiplicity of T_2 peaks, which are within the range of white matter T_2 [MacKay et al., 1994; Whittall et al., 1997]. Because the minor instability in S_0 estimation affects the T_2 estimation, our dataset does not make possible distinguishing meaningful components within WM T_2. Hence the assumption of a single WM T_2. In addition, our optimisation has produced approximately the mean of these distributions, 59ms for ACH-genu and 57ms for CON-genu.

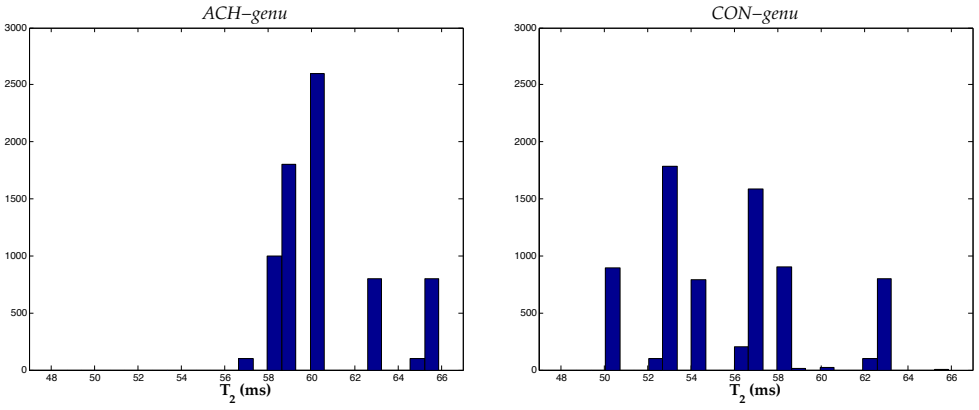


Figure 23.1: Distributions of T_2 estimates, from all the models and voxels, for each dataset.

Appendix table F.1 shows the sensitivity of two models, NODDI+ and MMWMD+, to perturbations in T_2. Prior to fitting the models, we fix T_2 to the mean of the genu region, 57 ms in the case of ACH-genu and 59 ms in the case of CON-genu, perturbing afterwards by 2 ms either side of this mean. The trends in both data sets are noticeable: as T_2 increases, the intracellular volume fraction increases, because the T_2 decay takes some weight off the extracellular compartment; this also explains the increase in the MMWMD+ radius estimate. The axial and radial diffusivities, and dispersion index, on the other hand, are not as sensitive.

Figure 23.2 shows the number of fitting runs needed to capture the best solution in 100 fitting runs, with a 99% confidence. In particular, ACH-*genu* data reveals that the more complex Tensor models require more fitting iterations than simpler Ball models, which is what we expected. The subplot on the right, on the richer CON-*genu* data, reveals that the broader coverage of the measurement space helps simpler models vs. more complex conterparts, e.g. Tensor-Stick-CSF vs. Tensor-Stick-CSF-diff, or the Cylinder equivalent; NODDI and MMWMD are relatively unaffected.

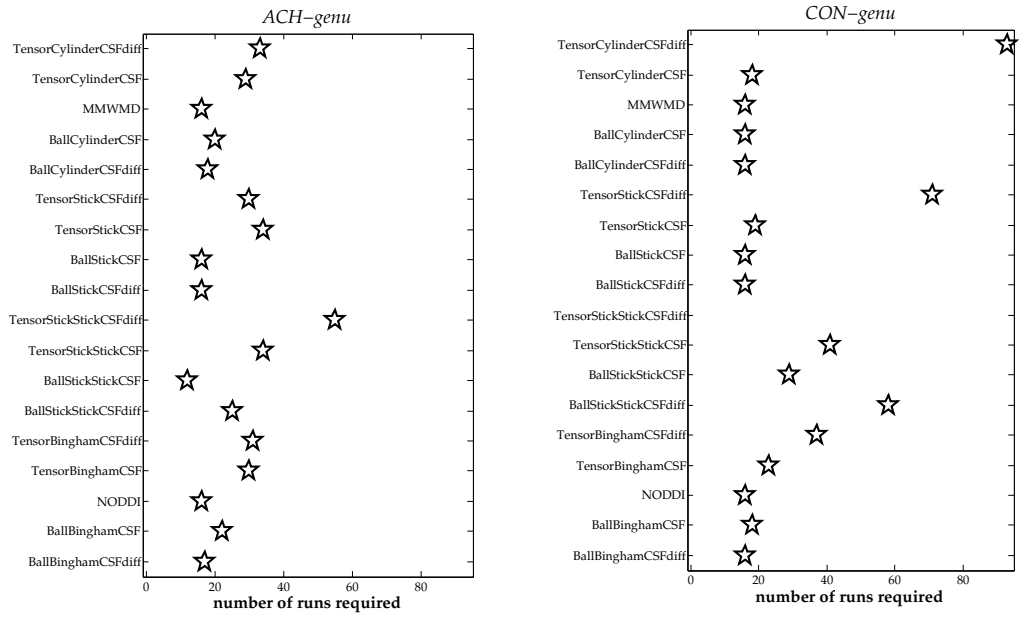


Figure 23.2: The number of runs required to find with high confidence each model's best solution achieved in 100 fitting runs.

Figures 23.3 and fig.23.4 shows how the ranking of models varies over bootstrap iterations and cross-validation folds. For both techniques at each iteration, we compute both the AIC of each model from the 'seen' data and the LSE from the 'unseen' data. Thus we obtain two model rankings from each iteration of each procedure. The positional variance diagrams on the right show a histogram for each model of the position in the ranking over all iterations. The top right figure is the traditional bootstrap result (from the 'seen' data) whereas the bottom left is the traditional cross-validation result (from the 'unseen' data), but we include the other two figures for extra information. (We note that removing the AIC penalisation for complexity in the top diagrams, which does not have a great effect, would leave the two components that add linearly to make the .632 bootstrap result. The .632 method was proposed by Efron and Tibshirani [1997], and used for diffusion model comparison by Scherrer et al. [2013]; it aims to strike a compromise between the under-estimation of variation from the LSE of seen data and the over-estimation from that of unseen data in cross validation.)

The bootstrap and cross-validation results broadly reflect the group structure observed in in Experiment 5 (Results 16.2) and Experiment 7 (Results 20.2), approximately divided into three groups: top-ranking models of Ball-Bingham-CFS-diff, Tensor-Bingham-CFS, Tensor-Bingham-

CFS-diff, Tensor-Cylinder-CFS-diff; the lowest-ranking combinations of Ball with Stick/s; all the models in between. The leave-one-shell-out strategy in our cross-validation creates much greater variation in the model ranking than the random selection of measurements in the bootstrapping experiment that does not consider the shell structure of the acquisition scheme. This shows that the full distribution of b-values is much more influential on the choice of model than the choice of gradient directions.

Using bootstrap datasets, table 23.1 shows the stability of the parameter estimates. The parameter standard deviations, shown as superscript to the mean estimates across the datasets, suggest a greater stability for simpler models. In particular, with the standard scanner data of ACH-*genu*, models combinations of Tensor with Bingham exhibit more sensitivity (i.e. less stability) in the estimation of extracellular Tensor diffusivities and the degree of intracellular Bingham dispersion κ; however, the intracellular Bingham volume fraction and diffusivity are relatively more stable. In Tensor with Cylinder models, it is the Cylinder diameter which is less stable. The sensitivity in the dispersion index κ of Tensor and Bingham models is also reflected with the CON-*genu* data. While other models show greater stability, compared with ACH-*genu* results, model Tensor-Bingham-CSF-diff shows increased sensitivity in compartmental volume fractions and diffusivities too. Across both datasets, the mean estimates are similar to but vary slightly from the parameter estimates obtained from the whole datasets, as in tables 16.1 and 20.2. Again, the bigger differences arise from the more complex models, such as Tensor-Bingham-CSF-diff and Tensor-Bingham-CSF-diff. This could be attributed to the richness of the bootstrap datasets which, though with the same number of measurements as the original dataset, have as distinct elements only about two-thirds of the original measurement space.

Figure 23.5 illustrates the spatial stability of one parameter estimate, the intracellular volume fraction, over the CC. We pick the two most complex models, Tensor-Cylinder-CSF_diff and Tensor-Bingham-CSF_diff, and their simpler counterparts, MMWMD and NODDI. While more complex models may fit the data better at the voxel level, the known volume fraction trends are more recoverable across the CC in the simpler models.

Lastly, Appendix figures F.2 and F.3 give estimates for the case of models fitted to each TE-specific data subset of CON-*genu*. Such subset contains four shells, using one combination of δ and Δ, with $|\mathbf{G}| = \{60, 100, 200, 300\}$ mT/m. Usually, as with CHARMED, NODDI or AxCaliber, one chooses the combination of δ and Δ which gives the best SNR or that is optimised for the experiment. It is not clear what this somewhat speculative investigation informs us about the downward diffusivity pattern in most models in fig. F.2. Previous work has looked at the dependence of intra- and extracellular diffusivity on diffusion time Δ [Novikov et al., 2012; Xu et al., 2014], or the 'fast' and 'slow' pools [Pyatigorskaya et al., 2013]. Though our experiment is not designed to test any of the above, it does inform us that simpler models are more stable across these diffusion regimes, as can be seen through the performance of fibre volume fraction in fig.F.3.

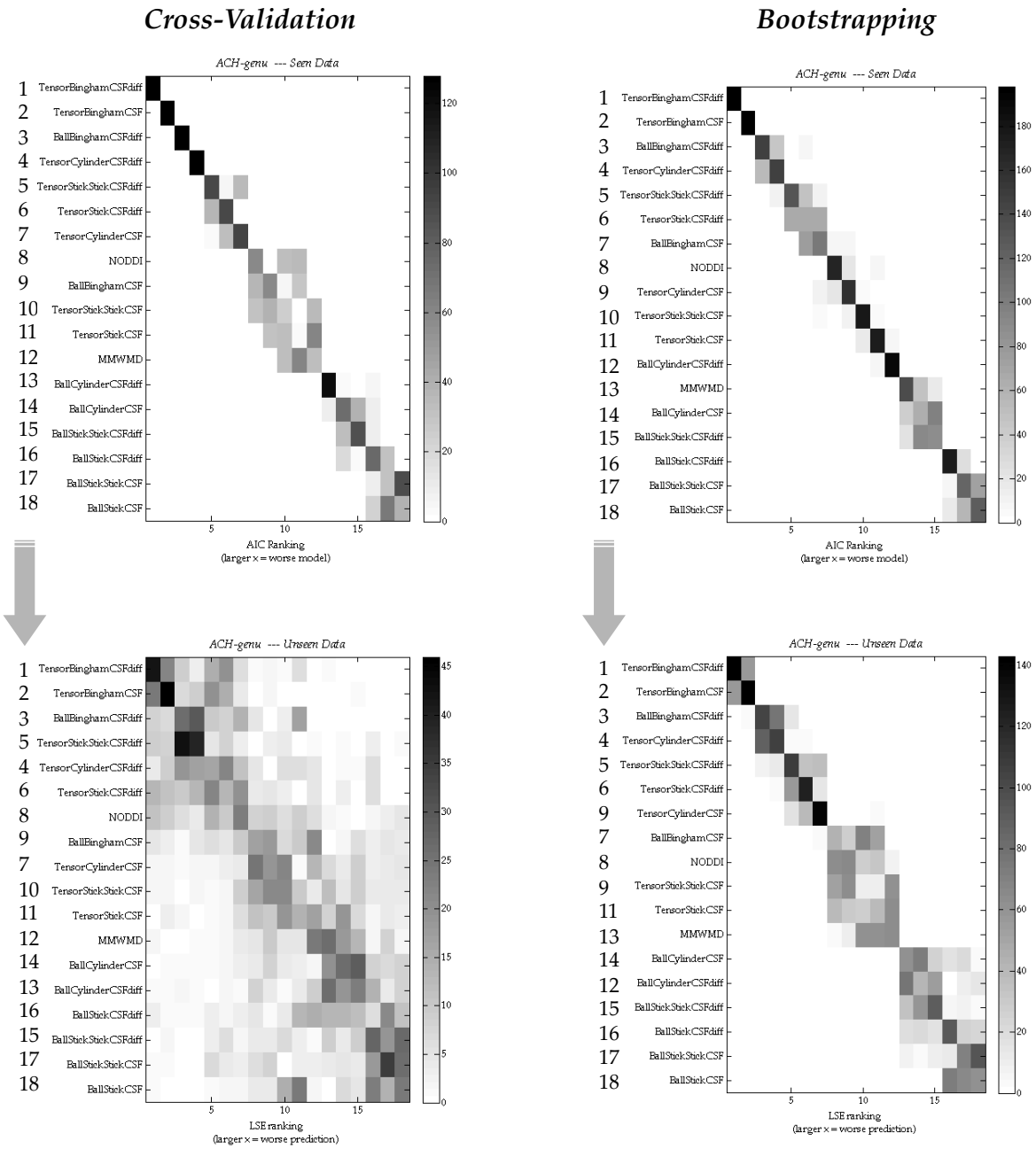


Figure 23.3: In cross-validation plots on the left, each HARDI shell of ACH-genu is left aside at a time; in bootstrapping, for the plots on the right, we construct 50 datasets for each of the four voxels in the ROI. The AIC group performance of fig.16.3 is reflected in both diagrams, but there is greater uncertainty when predicting unseen shells in the cross-validation.

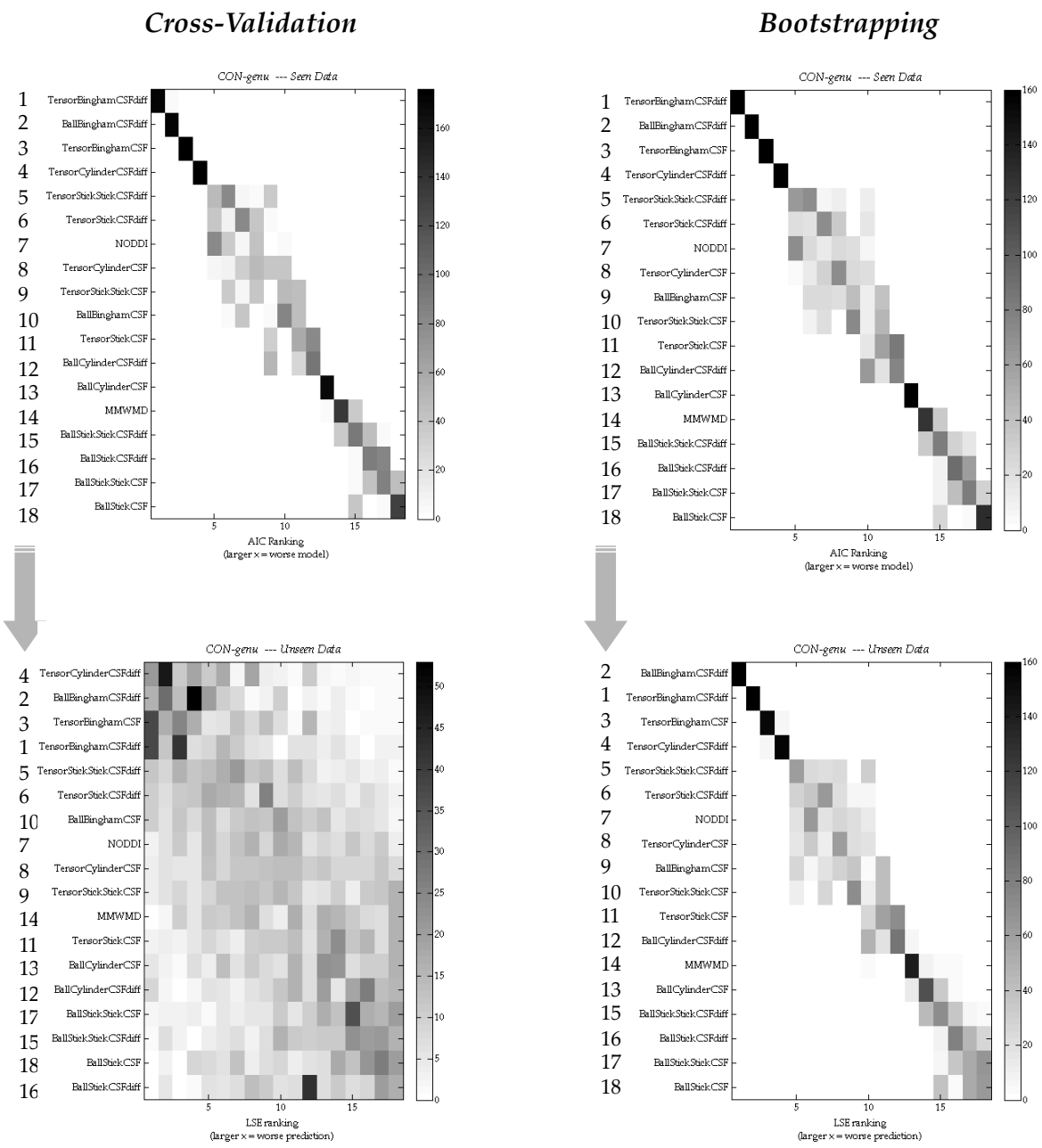


Figure 23.4: Similar to fig.23.3, but here using the CON-genu. As with the previous histogram, the bootstrapping repeats more faithfully on the unseen data the ranking obtained from the 'seen' data (which is also similar to the whole dataset ranking in fig.20.4).

			Cylinder/Bingham/Stick				Tensor/ Ball			
	Models	LSE (x 10^3)	Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
ACH-genu	Ball-Bingham-CSF-diff	6.04 3.7	0.82 1.1	*2.78* *1.9*	13 6.0		0.18 5.3	*0.74* *6.8*		
	Ball-Bingham-CSF	6.85 3.5	0.92 1.8	*2.53* *3.0*	9 8.2		0.08 20.7	*2.53* *3.0*		
	NODDI	7.21 3.7	0.64 0.6	*1.90* *0.0*	13 5.6		0.36 1.0	*1.90* *0.0*	*0.69* *1.0*	
	Tensor-Bingham-CSF	6.01 3.7	0.71 4.8	*2.43* *3.6*	14 35.6		0.29 10.8	*2.43* *3.6*	*0.58* *34.0*	*0.25* *102.4*
	Tensor-Bingham-CSF-diff	5.99 3.6	0.72 9.8	*2.62* *7.1*	13 37.4		0.28 23.2	*1.93* *30.9*	*0.58* *39.2*	*0.32* *80.9*
	Ball-Stick-Stick-CSF-diff	10.22 3.9	0.64 1.4	*1.94* *4.1*			0.36 2.5	*1.07* *7.3*		
	Ball-Stick-Stick-CSF	10.86 4.1	0.67 1.0	*1.60* *1.2*			0.33 2.1	*1.60* *1.2*		
	Tensor-Stick-Stick-CSF	8.05 3.6	0.51 3.1	*1.74* *1.0*			0.49 3.3	*1.74* *1.0*	*0.73* *6.2*	*0.44* *7.8*
	Tensor-Stick-Stick-CSF-diff	6.44 3.8	0.42 3.9	*0.96* *4.1*			0.58 2.8	*2.92* *1.1*	*0.56* *5.9*	*0.32* *6.9*
	Ball-Stick-CSF-diff	10.24 3.9	0.64 1.4	*1.94* *4.2*			0.36 2.5	*1.06* *7.4*		
	Ball-Stick-CSF	10.88 4.1	0.67 1.0	*1.60* *1.2*			0.33 2.1	*1.60* *1.2*		
	Tensor-Stick-CSF	8.06 3.6	0.51 3.0	*1.74* *0.9*			0.49 3.1	*1.74* *0.9*	*0.72* *6.1*	*0.44* *7.4*
	Tensor-Stick-CSF-diff	6.43 3.7	0.42 3.5	*0.96* *3.9*			0.58 2.5	*2.92* *0.9*	*0.56* *5.4*	*0.33* *5.8*
	Ball-Cylinder-CSF-diff	8.74 4.0	0.89 3.3	*1.91* *12.9*		6.7 10.4	0.11 24.4	*1.11* *100.3*		
	Ball-Cylinder-CSF	9.02 3.3	0.86 1.5	*1.70* *1.1*		6.0 2.0	0.14 9.2	*1.70* *1.1*		
	MMWMD	8.57 3.6	0.62 1.7	*1.70* *0.0*		4.7 4.7	0.37 2.8	*1.70* *0.0*	*0.64* *2.8*	
	Tensor-Cylinder-CSF	7.88 3.5	0.61 3.0	*1.75* *0.9*		4.8 3.9	0.39 4.9	*1.75* *0.9*	*0.75* *7.7*	*0.39* *10.7*
	Tensor-Cylinder-CSF-diff	6.43 3.7	0.44 6.2	*0.99* *5.9*		2.9 22.3	0.56 5.1	*2.94* *1.0*	*0.56* *7.7*	*0.31* *7.9*
CON-genu	Ball-Bingham-CSF-diff	6.42 2.8	0.77 0.7	*2.79* *1.4*	16 5.4		0.22 2.5	*0.82* *4.0*		
	Ball-Bingham-CSF	7.47 3.1	0.84 1.2	*2.32* *2.0*	13 10.8		0.15 6.7	*2.32* *2.0*		
	NODDI	7.35 2.5	0.61 0.4	*1.90* *0.0*	17 4.7		0.39 0.6	*1.90* *0.0*	*0.73* *0.6*	
	Tensor-Bingham-CSF	6.42 2.6	0.68 1.5	*2.14* *1.1*	22 10.9		0.31 3.2	*2.14* *1.1*	*0.84* *6.5*	*0.66* *7.3*
	Tensor-Bingham-CSF-diff	6.37 2.6	0.64 10.5	*2.02* *19.0*	23 22.9		0.35 20.1	*2.24* *32.6*	*0.77* *8.9*	*0.55* *13.6*
	Ball-Stick-Stick-CSF-diff	10.14 2.9	0.60 1.3	*2.16* *3.2*			0.38 2.2	*0.98* *5.0*		
	Ball-Stick-Stick-CSF	11.34 3.7	0.63 1.2	*1.64* *1.2*			0.35 2.2	*1.64* *1.2*		
	Tensor-Stick-Stick-CSF	7.94 2.4	0.48 2.4	*1.81* *0.7*			0.51 2.3	*1.81* *0.7*	*0.70* *4.8*	*0.44* *5.5*
	Tensor-Stick-Stick-CSF-diff	6.67 2.5	0.39 2.4	*1.01* *2.6*			0.59 1.6	*2.91* *0.9*	*0.55* *4.3*	*0.34* *5.0*
	Ball-Stick-CSF-diff	10.42 2.9	0.61 1.3	*2.17* *3.3*			0.39 2.1	*0.97* *5.0*		
	Ball-Stick-CSF	11.58 3.8	0.64 1.2	*1.62* *1.2*			0.35 2.2	*1.62* *1.2*		
	Tensor-Stick-CSF	8.11 2.3	0.48 2.5	*1.81* *0.7*			0.51 2.3	*1.81* *0.7*	*0.70* *4.9*	*0.43* *5.6*
	Tensor-Stick-CSF-diff	6.73 2.5	0.40 2.3	*1.00* *2.2*			0.59 1.5	*2.91* *0.8*	*0.56* *3.8*	*0.34* *4.7*
	Ball-Cylinder-CSF-diff	8.38 2.8	0.76 0.9	*1.73* *1.2*		3.8 1.5	0.23 3.1	*1.72* *4.2*		
	Ball-Cylinder-CSF	8.40 2.8	0.77 0.7	*1.73* *0.8*		3.8 1.4	0.23 2.4	*1.73* *0.8*		
	MMWMD	8.45 2.4	0.58 0.6	*1.70* *0.0*		3.0 2.5	0.41 0.8	*1.70* *0.0*	*0.70* *0.8*	
	Tensor-Cylinder-CSF	7.69 2.6	0.63 3.4	*1.79* *0.8*		3.3 8.7	0.36 6.1	*1.79* *0.8*	*1.05* *8.5*	*0.60* *9.5*
	Tensor-Cylinder-CSF-diff	6.56 2.6	0.50 3.7	*1.18* *3.0*		2.7 4.4	0.49 4.0	*3.00* *0.0*	*0.69* *6.2*	*0.40* *7.4*

Table 23.1: The stability of parameter estimates across bootstrap datasets. We report the mean estimate and to their right, as superscript, the standard deviations (as a percentage of the mean).

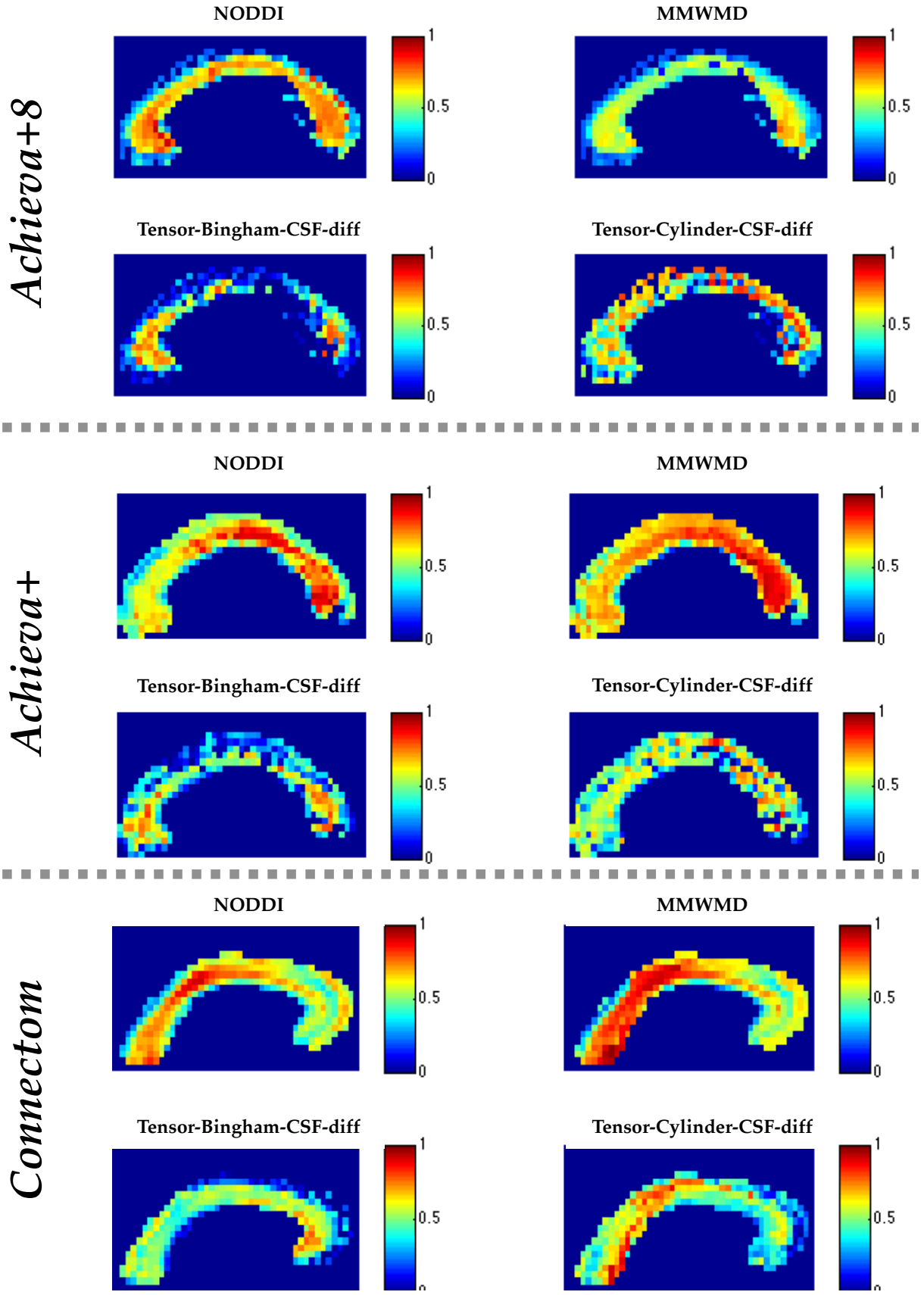


Figure 23.5: Intracellular volume fraction maps of four representative models fitted to the whole CC Connectom data.

24

SUMMARY

A way of evaluating the sensitivity and robustness of parameter estimation is by variations or perturbations in the data. Lacking multiple distinct subjects' datasets, bootstrapping [Efron, 1979] and cross-validation [Stone, 1974] sub-sample the original dataset repeatedly to investigate model stability. Both methods aim at balancing bias with variance. These complementary methods for model comparison largely confirmed the model ranking revealed by the AIC criterion; the results, however were weaker in the k-fold cross-validation, blurring the the distinction between the models in predicting unseen data.

The whole brain maps of the intracellular volume fraction revealed that, though some models may capture the voxel better signal than some others, they do worse at capturing trends beyond the genu ROI we have concentrated on. This will need further investigation, as the fig.23.5 (and the signal plots of fig.E.1 in Experiment 7) suggested that the current models still lack features that enable capturing characteristics across the whole signal spectrum.

Part VIII

CONCLUSIONS

25

THIS WORK

This is the main question posed at the start: *Given so many candidate models for diffusion MRI brain signal, which describes best the signal from the WM of the* in vivo *human brain?*

By using very rich data, sampled for unusually long times, and uniquely strong scanning power, this work has provided a comparison of parametric diffusion MRI models from which we draw some important conclusions. In addition, as with every experiment, several assumptions and limitations need to be born in mind.

The precursor to this work, the study by Panagiotaki et al. [2012], provided us with the initial taxonomy of models and model comparison framework. In our first attempt to translate this study to *in vivo* human data, during Experiment 1, we sampled linearly/evenly across a wide range of b-values and diffusion times on the clinical system, aiming at a sampling space dense enough with measurements for the time available. However, the ranking produced could not distinguish between the three-compartment models. The reasons for the difference could have arisen because of the differences in the type of tissue scanned, which affects water diffusion significantly [Shepherd et al., 2009], or in the imaging protocol, as the previous experiment used a fixed animal on which higher gradient strengths and longer scanning times can be used.

Experiment 2 showed us that increasing the number of measurements, in this case through increasing gradient angular resolution from only three directions, does help in distinguishing the models. [1]

Experiment 3, with data gathered over 8h, used a much richer imaging sequence, to sample many gradient orientations, diffusion times, gradient pulse times, and gradient magnitudes. In common with the *ex vivo* tissue by Panagiotaki et al. [2012] three compartment models are best. With *in vivo* data we obtained a clearer ranking structure.

Experiment 4 showed us that models for fibre dispersion outperformed other models. The results demonstrated the potential benefit of modelling dispersion even in apparently straight and coherent fibres.

However, in using dispersion models, one limitation of Experiment 4 was that averaging voxels across diverse parts of the CC may have exaggerated the dispersion. So Experiment 5 introduced a few improvements to the pipeline by fitting the models voxel-wise, over regions of genu, midbody and splenium, to avoid phantom dispersion in the analysis[2]. In addition, we

1 In hindsight, leaving aside the precise weight of repeated measurements vs. increased angular resolution, measurements with a richer angular resolution would increase even more model specificity in regions where voxel fibres lack coherence and have a more complex structure.

2 Another potential improvement to Experiment 4, so as to avoid introduced phantom dispersion, would have been to re-align (in each voxel, before they were all averaged, and assuming one population of fibres) the deviated principal

also accounted explicitly for the variable echo time in the measurements, as this improved the noise modelling: normalising by the b=0 image can affect the noise in the measurements, making them exponentially dependent on echo time. We also studied compartmentally different T_2 and diffusivity. While we saw that a separate T_2 for the CSF compartment was necessary, we found no support from our data for separate intra/extracellular relaxation rates. As for the different intra/extracellular diffusivities, we saw a variable effect, from a vast improvement in models with extracellular isotropic Ball and intracellular Stick, to hardly any difference in models combining extracellular anisotropic Tensor with intracellular Bingham.

Experiment 6 demonstrated that the results of the model ranking were broadly reproduced across different scanning sessions as well as other healthy subjects. However, the other healthy subjects may have benefited by pre-optimising more carefully the scanning protocol as, at present, it is still long at 2h and prone to motion artefacts, as witnessed in this study.

The data in Experiment 7, collected from the Connectom scanner with gradients of up to 300mT/m gradients, confirmed this ranking, and produced model parameters which were comparable to those acquired with standard scanners. However, this richer dataset also revealed greater differences between the models, and some limitations, e.g. of dispersion Bingham models in capturing the very highest b-value signal.

The bootstrapping and cross-validation techniques in Experiment 8 complemented the AIC and BIC tests for comparing models. While the ranking between model groups remains largely as before, with minor variations, the unseen k-fold cross-validation reveals that the distinction between the models is less clear than what is provided by other techniques.

The methodology applied here can be readily applied to the whole brain. However, there would be some limitations. One immediate drawback comes from the slightly coarser-than-usual resolution of 2mm x 2mm x 4mm used here in the CC, chosen to enhance SNR. Away from the CC and towards the grey matter, thinner fibre bundles would require a finer image resolution, which would lower the SNR; this would particularly affect the high b-value shells and their contrast enhancing ability; the simpler models would then have an advantage. On the other hand, we would expect that models which explicitly capture fibre dispersion would benefit in regions of crossing or fanning fibres. Accounting specifically for WM and CSF T_2 makes this methodology particularly applicable to other areas of the brain with CSF contamination, and increase the model specificity.

eigenvector from the mid-sagittal perpendicular. Experiment 5, however, goes a step further, in giving us voxel-wise specificity, alas at the expense of SNR.

26

FUTURE WORK

Researching for the best model has a goal: to provide a non-invasive histological tool that helps diagnose and monitor human diseases. Before this model can be tested on patients, further improvements can be made on the methodology of acquisition, pre-processing and post-processing of the data.

For a good model comparison, it is necessary to acquire the most representative and high quality data that can be had; as in the present study, this often involves extremely long scans. With the benefit of hindsight, the next scanning acquisition should use a smaller voxel size. Finer resolution would come at the expense of SNR, but this can be improved by taking more signal averages. To make such scanning “do-able", the protocol can be split into three or four sessions, dedicating one-quarter of the session time to some “core" part of the protocol, which can be used to test or measure inter-session scan variability. The dataset would also benefit from a whole brain acquisition, instead of the current time-saving limited *field of view*, as well as acquiring more inter-leaved B_0 images (1 for every 5 DW images vs. 1 for every 10); both these changes would improve image registration and, hence, the quality of the data. (Such adjustments are already in place on the acquisitions with the *Connectom* scanner; it, however, suffers from severe eddy current distortions, which undoubtedly will be addressed in the near future.) A whole-brain dataset would also offer more usability to the data, and offer more anatomical variety. Future work should involve disseminating some of our data to other researchers: we have in the past organised a challenge, for MICCAI 2013, using data from this thesis, and will be organising another for ISBI 2015.

Validating the applicability and appropriateness of models on different types of tissue requires wide sampling of the data, from different sources. Histological samples from electron microscopy or histochemistry staining (as in the study of Assaf et al. [2008] to compare axon diameter measurements), provides one method of validation. The presence of ground truth data against which the performance of the models can be judged can also be provided by synthetic simulations. This simulated data set could be generated for a range of substrates which reflect the brain white matter, and generate data for a broader protocol (to achieve a sort of signal-sampling super-resolution). Comparing this with the *in vivo* data can provide some comparison and an indication of sensitivity and specificity. Another complementary data set could come from emerging phantoms [Hubbard et al., 2014; Zhou et al., 2012], or fixed brain data, which can afford longer scanning times and stronger gradients.

Model selection can benefit from data acquired with other (than PGSE) pulse schemes such as *oscillating-gradient spin-echo* sequences [Callaghan and Stepišnik, 1995; Does et al., 2003], which can be optimised to provide increased sensitivity, e.g. to axon diameters [Drobnjak

et al., 2010], *double-pulsed gradient spin-echo* sequences [Komlosh et al., 2007], which promise deeper microstructural features (such as anisotropy in grey matter), or *twice-refocused spin-echo* sequences which can address specific experimental requirements, such as reducing eddy current artefacts [Reese et al., 2003] and, again, potentially provide added sensitivity to specific parameters such as Clayden et al. [2009].

Further, and broadening the initial question "which model is best?", we need to consider not just how well the models fit the signal, but also how sensible and sensitive the model parameters are. For now, many models appear to do well in some aspects while not so well in others (e.g. the Ball-Stick does provide sensible estimates, and of very little computational cost, but it ranks last in describing the signal). One possibility is to combine all these candidate models into a single framework, through *model averaging* [Burnham and Anderson, 2002]. This would extend the applicability to areas outside the CC. This approach would need careful justification and interpretation as it is unclear what overlap there is between the models (for example, do Stick volume fraction and Cylinder volume fraction represent the same biophysical feature?).

The model selection framework can be extended to include other models. Added complexity (e.g. two or more completely free tensors) could benefit even existing models, which have yet to capture the whole signal. Indeed, the richness of the data may support other non-parametric q-space models [Callaghan et al., 1990; Callaghan, 1991] such as *Diffusion Spectrum Imaging* [Wedeen et al., 2005] or *Diffusion Kurtosis Imaging* [Jensen et al., 2005]). The model comparison will need adjustment, since for this class it is not clear how to penalise for complexity in the AIC and BIC (e.g. spherical deconvolution [Anderson and Ding, 2002; Tournier et al., 2004] has many parameters kept under subject by a regularisation term). To circumvent this problem, one possible fix would be to use *cross-validation* or *bootstrapping*.

Once appropriate models have been identified, experiment design techniques [Alexander, 2008; Caruyer et al., 2013] can determine more economical protocols to replace the current impractical length of the scanning.

The (lengthy) protocol we use here is designed specifically for model selection, rather than large-scale application. In general, reduced data sets will favour the simpler models; larger data sets will support the more complex models. Here we sampled as wide a coverage as possible of the measurement space to get the best idea of what kind of model explains the entire measurement space. Most of the current models in our analysis are not yet appropriate for existing sparse data sets such as off-the-shelf single shell HARDI data, which only support simple models. Rather, these results inform the choice of protocol for future *in vivo* microstructure imaging once we identify the right model.

The work presented on this thesis has been on the very homogeneous part of the CC. Future research, applying these methods and insights elsewhere in the brain, could extend and deepen our knowledge of other white matter structures. Apart from neurological applications, this work can also provide insights for application into other tissue types and diseases, such as cancer.

Part IX

APPENDIX

A

TO PART IV (BACKGROUND)

A.1 RATE OF CHANGE OF MAGNETIC MOMENT

(OR ROTATING COORDINATE FRAMES CALCULUS) Suppose we have two coordinate systems, one rotating at angular velocity $\boldsymbol{\Omega}$ against another that is inertial. Suppose also that their unit vectors are, $r_i = (r^1\ r^2\ r^3)^t$ and $f_j = (f^1\ f^2\ f^3)^t$.

If $\mathbf{P}$ is a rotation matrix (so $\mathbf{PP^T} = \mathbf{I}$), and $\mathbf{r} = \mathbf{Pf}$, then the (i, j)-th member of $\mathbf{P}$ is $\mathbf{P_{(i,j)}} = \mathbf{r_i f_j}$.

Then, by the chain rule, (and commutativity of differentiation and transposition)

$$\mathbf{o} = \frac{\mathbf{d}}{\mathbf{dt}}\mathbf{I} = \frac{\mathbf{d}}{\mathbf{dt}}\left(\mathbf{PP^T}\right) = \left(\frac{\mathbf{d}}{\mathbf{dt}}\mathbf{P}\right)\mathbf{P^T} + \mathbf{P}\left(\frac{\mathbf{d}}{\mathbf{dt}}\mathbf{P^T}\right)$$

$$\Rightarrow -\left(\frac{d}{dt}\mathbf{P}\right)\mathbf{P^T} = \mathbf{P}\frac{d}{dt}\left(\mathbf{P^T}\right) = -\left(\mathbf{P}\frac{d}{dt}\left(\mathbf{P^T}\right)\right)^T$$

$$\Rightarrow \mathbf{P}\frac{d}{dt}\left(\mathbf{P^T}\right) = \begin{pmatrix} 0 & -\Omega^3 & \Omega^2 \\ \Omega^3 & 0 & -\Omega^1 \\ -\Omega^2 & \Omega^1 & 0 \end{pmatrix}$$

with Ω^1, Ω^2,Ω^3 being any three scalars, such that, if $\boldsymbol{\Omega}$=$(\Omega^1\ \Omega^2\ \Omega^3)^t$ and $\mathbf{f} = (\mathbf{f}^1\ \mathbf{f}^2\ \mathbf{f}^3)^t$ any vector, then $\mathbf{P}\frac{d}{dt}\left(\mathbf{P^T}\right)\mathbf{f} = \boldsymbol{\Omega} \times \mathbf{f}$

With $\mathbf{m_r} = \mathbf{Pm_f}$ we get $\mathbf{m_f} = \mathbf{P^T m_r}$ and, by chain differentiation,

$$\frac{d}{dt}\left(\mathbf{m_f}\right) = \frac{d}{dt}\left(\mathbf{P^T m_r}\right) = \frac{d}{dt}\left(\mathbf{P^T}\right)\mathbf{m_r} + \mathbf{P^T}\frac{d}{dt}\left(\mathbf{m_r}\right)$$

$$\Rightarrow \mathbf{P}\frac{d}{dt}\left(\mathbf{m_f}\right) = \mathbf{P}\frac{d}{dt}\left(\mathbf{P^T m_r}\right) = \mathbf{P}\frac{d}{dt}\left(\mathbf{P^T}\right)\mathbf{m}_r + \mathbf{PP^T}\frac{d}{dt}\left(\mathbf{m_r}\right)$$

$$\Rightarrow \mathbf{P}\frac{d}{dt}\left(\mathbf{m_f}\right) = \boldsymbol{\Omega} \times \mathbf{m_r} + \frac{d}{dt}\left(\mathbf{m_r}\right)$$

or

$$\frac{d}{dt}\mathbf{m} = \frac{\partial}{\partial \mathbf{t}}\mathbf{m} + \boldsymbol{\Omega} \times \mathbf{m} \tag{A.1}$$

The left-hand side, $\frac{d}{dt}\mathbf{m}$, gives the rate of change of vector $\mathbf{m}$ with respect to the inertial frame, $\frac{\partial}{\partial t}\mathbf{m}$ with regard to the rotating frame. $\boldsymbol{\Omega}$ can be thought of as the angular velocity vector, with regard to the inertial frame.

If the frame of reference is taken to be the proton's spinning coordinate system, so $\frac{\partial}{\partial t}\mathbf{m} = \mathbf{o}$, i.e. no change of magnitude, the rate of change of the magnetic moment is proportional to the applied torque. This net moment of force comes from the proton's magnetic and gravity force.

A.2 LARMOUR FREQUENCY

(The following is largely drawn from a classic Physics textbook [Goldstein].) We assume that all particles are homogenous, i.e. have the same e/m = charge to mass ratio. If the charged particles move, this will constitute an electric current distribution. This current has its magnetic moment, say $\mathbf{M}$, which can interact with a magnetic field, say $\mathbf{B}$. Then, the rate of change of total angular momentum will equal the applied torque $\frac{d}{dt}\mathbf{L} = \mathbf{M} \times \mathbf{B}$.

With $\mathrm{current\ density}(\mathbf{j}) = \frac{\mathrm{charge}(e)}{\mathrm{mass}(m)} \times \mathrm{density}(\rho) \times \mathrm{velocity}(\mathbf{v})$ and $\mathrm{position}(\mathbf{r})$, the magnetic moment is defined, in Gaussian units, as $\mathbf{M} = \frac{1}{2c}\int(\mathbf{r} \times \mathbf{j})dV = \frac{e}{2mc}\int \mathbf{r} \times (\rho\mathbf{v})dV \quad \Rightarrow \mathbf{M} = \frac{e}{2mc}\mathbf{L}$.

This means that $\frac{d}{dt}\mathbf{L} = \mathbf{L} \times \frac{\mathbf{e}}{\mathbf{2mc}}\mathbf{B}$; and referring to eq.A.1, this is the equation of motion of vector $\mathbf{L}$, of constant magnitude, $\frac{\partial}{\partial t}\mathbf{L} = \mathbf{0}$, rotating in space about the direction of $\mathbf{B}$ with angular velocity $\mathbf{\Omega} = -\frac{e}{2mc}\mathbf{B}$. Therefore, the uniform field $\mathbf{\Omega}$ causes the charged body to precess uniformly with angular velocity $\mathbf{\Omega}$, known as the *Larmour frequency*.

A.3 TENSORS

(The following is also drawn from [Goldstein].) The Angular Momentum $\mathbf{L}$ and Angular Velocity ω of a spinning body are connected by a linear transformation $\mathbf{T}$, such that $\mathbf{L} = \mathbf{T}\boldsymbol{\omega}$, or $\mathbf{T} = \frac{\mathbf{L}}{\boldsymbol{\omega}}$. The right-hand side of this equation expresses a vector dividing by a vector, which is a quantity called tensor (much as in arithmetic, where dividing within the group of integers introduces Real numbers, or taking the square root of negative numbers produces Imaginary numbers).

A TENSOR DEFINITION: In a 3-dimensional space, a tensor $\mathbf{T}$ of rank N (has 3^N components $T_{j_1 j_2 \ldots j_N}$ and) transforms under an orthogonal transformation of coordinates $\mathbf{A}$ according to

$$\tilde{T}_{i_1 i_2 i_3 \ldots} = \sum_{j_1, j_2, j_3} a_{i_1 j_1} a_{i_2 j_2} a_{i_3 j_3} \ldots T_{j_1 j_2 j_3 \ldots}$$

THE MATRIX RELATION When in the field of orthogonal transformations (a subgroup of which are *Rotations*), a tensor is practically the same as the square matrix formed from its components. *The tensor is defined only in terms of its transformation properties under orthogonal coordinate transformations, and the tensor remains unchanged under any choice of coordinate system. But there is no restriction on the type of transformation that can be applied to a matrix.* Examples include:

Scalar = Tensor of Rank 0, transforming very simply as $\tilde{T} = aT$

Vector = Tensor of Rank 1, transforming as $\tilde{T}_i = \sum_j a_{ij} T_j$

Matrix = Tensor of Rank 2, transforming as $\tilde{T}_{i_1 i_2} = \sum_{j_1, j_2} a_{i_1 j_1} a_{i_2 j_2} T_{j_1 j_2}$

A.4 COMPARTMENT MODELS

THE DIFFUSION TENSOR is a symmetric matrix consisting of six free parameters. It is referred to as a *tensor* because, as the matrix undergoes only orthogonal transformations (i.e. rotations of coordinate axes), its shape does not depend on the coordinate system, regardless of what coordinate system you choose to look at it from.

With DT as a model, the signal can be expressed as $S = S_0\, exp(-bd\mathbf{q}^t\mathbf{D}\mathbf{q})$ where

$$\mathbf{D} = d_{\parallel}\mathbf{n}\mathbf{n}^{\mathbf{T}} + d_{\perp_1}\mathbf{n}_{\perp_1}\mathbf{n}_1^{\mathbf{T}} + d_{\perp_2}\mathbf{n}_{\perp_2}\mathbf{n}_{\perp_2}^{\mathbf{T}} \tag{A.2}$$

is the Diffusion Tensor diagonalised, with $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ being the eigenvectors, and $\left\{d_{\parallel}, d_{\perp 1}, d_{\perp 2}\right\}$ the eigenvalues, and $\mathbf{q}$ is the wavevector. These parameters produce two very important and frame-invariant indices: *Mean Diffusivity* $MD = \frac{1}{3}(d_{\parallel} + d_{\perp 1} + d_{\perp 2})$ and *Fractional Anisotropy* $FA = \left(3/2 \cdot (d_{\parallel} - MD)^2 + (d_{\perp 1} - MD)^2 + (d_{\perp 2} - MD)^2/(d_{\parallel}^2 + d_{\perp 1}^2 + d_{\perp 2}^2)\right)^{\frac{1}{2}}$

SUB-CASES of the DT, using eq.A.2, are the *Zeppelin*, which is a cylindrically symmetric DT, i.e. where $d_{\perp_1} = d_{\perp_2}$, the *Ball*, which is an isotropic DT, i.e. where $d_{\parallel} = d_{\perp_1} = d_{\perp_2}$, and the *Stick*, which is a zero-radius (one-dimensional) anisotropic tensor, with $\mathbf{D} = d\mathbf{n}\mathbf{n}^t$.

THE CYLINDER signal expression is a linear combination of the restricted A_r and hindered A_h components (weighted by their respective volume fractions). A_r is a product of the parallel $A_{r\parallel}$ and perpendicular/radial $A_{r\perp}$, both being functions of applied diffusion gradient of magnitude G (with radial component $G_{\perp}$ and axial $G_{\parallel}$, relative to the Cylinder axis) and times Δ and δ [Alexander, 2009]. The axial signal $A_{r|}(G_{\perp}, \Delta, \delta) = exp(-(\Delta - \delta/3)\gamma^2\delta^2 G_{\parallel}^2 d_{\parallel})$, whereas the radial signal $A_{r\perp}(G_{\perp}, \Delta, \delta)$ is [Murday and Cotts, 1968; Vangelderen et al., 1994]:

$$exp\left(-2\gamma^2 G_{\perp}^2 \sum_{m=1}^{\infty} \frac{2d_{\parallel}\alpha_m^2\delta - 2 + 2e^{-d_{\parallel}\alpha_m^2\delta} + 2e^{-d_{\parallel}\alpha_m^2\Delta} - e^{-d_{\parallel}\alpha_m^2(\Delta-\delta)} - e^{-d_{\parallel}\alpha_m^2(\Delta+\delta)}}{d_{\parallel}^2\alpha_m^6(R^2\alpha_m^2 - 1)}\right)$$

where R is the Cylinder radius, $d_{\parallel}$ and $d_{\perp}$ are the apparent diffusion coefficients, and $\alpha_m R$ is the m-th zero of the derivative of the Bessel function of the first kind, order one.

GDRcylinders, instead of a single index R, has its Cylinder radii follow a gamma distribution [Assaf et al., 2008] with *shape* parameter κ and *scale* parameter θ, such that $\kappa\theta$ gives the mean whereas $\kappa\theta^2$ gives the variance of of the distribution.

Astrocylinders model the signal from cylinders with uniform $\mathbf{n}$-vector orientation distribution $p(\mathbf{n})=1/4\pi$, where the above Cylinder signal A_r is integrated over all $\mathbf{n}$ directions so that $A_G = \int A_r p(\mathbf{n}) d\mathbf{n}$ [Panagiotaki et al., 2012]. Analogously, by using Stick signal instead of Cylinder, one can obtain the signal for *AstroSticks*.

The *Sphere* signal is similar to $A_{r\perp}(G_{\perp}, \Delta, \delta)$ above, with the *Dot* compartment as a special case where $R = 0$.

B

MICCAI 2013 - CDMRI'13 CHALLENGE

B.1 DESCRIPTION

As part of the CDMRI workshop, at MICCAI, we organised a model fitting challenge, in which groups were invited to find the best diffusion model to describe the very rich dataset of Fig.B.1. Challenge participants have access to three-quarters of the whole dataset; the winning model was the one that predicted the remaining 'unseen' quarter most closely, the shells being shown as boxed in the figure.

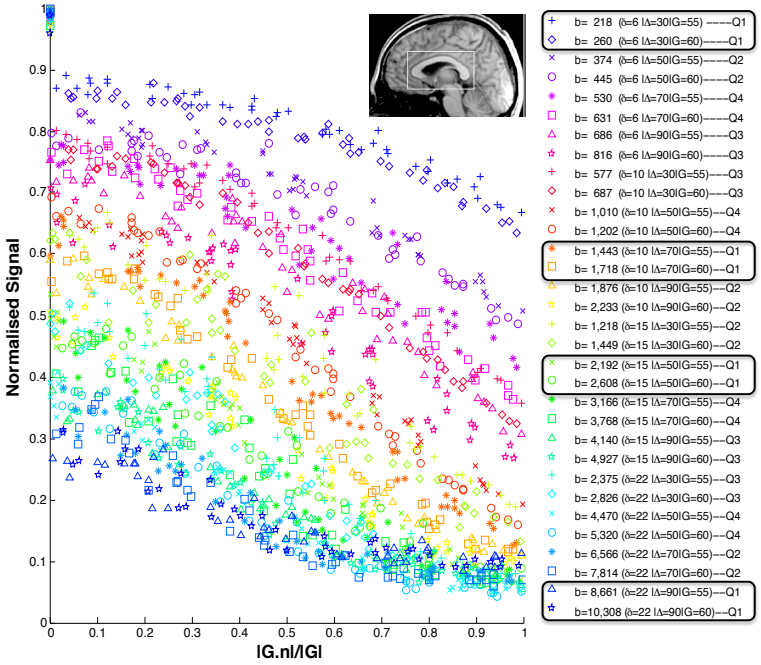


Figure B.1: The signal for training the models.

B.2 CHALLENGE ENTRIES

These were the selected entrants:

- Mohammed Alipoor, University of Chalmers, Sweden: *Weighted LS estimation of 4th order diffusion tensors*

  This proposes a new tensor model and a weighted-least-squares scheme, penalising noisy measurements. Requires at least 15 non-colinear dMRI measurements.

- Benoit Scherrer, University of Harvard, USA: *DIAMOND: a novel diffusion model that characterizes the distribution of anisotropic micro-structural environments with DWI.*

Voxel signal is composed of large scale of compartments, of a continuous distribution of spin packets. Here, there is one isotropic, one anisotropic compartments; 10 parameters.

- Xinghua Zhu, Unversity of Utah, USA, and University of Hong Kong, China: *Predicting Diffusion Weighted MR Signal with Gaussian Process Regression*.

  This model uses a nonparametric Gaussian process regression to estimate the hidden DW signal.

- Lin Mu, University of Zhejiang, China: *A Multicomponent Model For Diffusion Parameters.*

  The model adopts a three-compartment model that consists of 5 tensors, 5 sticks and a constant. 56 parameters.

- Uran Ferizi, UCL, UK: *A diffusion MRI model for fibre dispersion*.

  The 3-compartment model includes a Bingham distribution for fibres, an anisotropic compartment for the extracellular space and a CSF compartment. 10 parameters.

- Torben Schneider, UCL, UK: *A log-normal distribution of axonal diameters.*

  Cylindrically restricted intracellular diffusion across a log-normal distribution of diameters, an anisotropic compartment for the extracellular space and a constant term. 9 parameters.

B.3 RESULTS

While all models performed well, the one with least error was that provided by M.Alipoor of Chalmers. It was interesting that, even though the model of L.Mu of Zhajiang was relatively very complex, it provided the second best score; this showed the ability of the rich data to handle such model complexity.

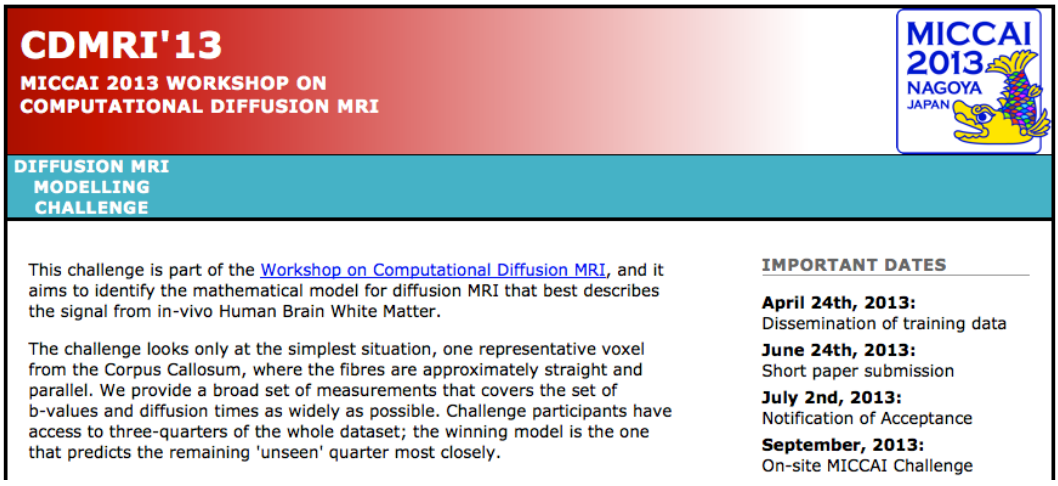


Figure B.2: A cut-out from the challenge website.

C

TO PART IV (MODEL ADVANCES)

Param.		2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10	2x4_2	2x4_5	2x4_10
Nr	**Models**	**BIC**			**Intra.1 Vol.Fr.**			**CSF/Dot Vol.Fr.**			**Axial Diff.**			**Radial Diff.**			**Kappa**		
10	Zepp.Bing.CSF.	513	380	359	0.56	0.59	0.60	0.15	0.11	0.09	2.0	1.9	1.9	0.5	0.6	0.7	6.9	7.1	6.5
9	ZepT.Bing.CSF.	516	377	356	0.59	0.59	0.60	0.13	0.11	0.09	2.0	1.9	1.9	0.6	0.7	0.7	7.0	7.1	6.5
12	Tens.Bing.CSF.	516	383	362	0.56	0.59	0.60	0.15	0.11	0.09	2.0	1.9	1.9	0.6	0.7	0.8	8.1	8.0	7.3
10	Tens.Wat.CSF.	519	392	369	0.55	0.59	0.60	0.16	0.12	0.10	2.0	1.9	1.9	0.5	0.7	0.7	5.4	5.5	5.3
8	Zepp.Wat.CSF.	531	401	373	0.56	0.59	0.60	0.15	0.11	0.09	2.0	1.9	1.9	0.5	0.6	0.7	5.6	5.6	5.4
7	ZepT.Wat.CSF.	533	398	369	0.59	0.60	0.60	0.13	0.11	0.09	2.0	1.9	1.9	0.6	0.6	0.7	5.8	5.6	5.4
9	ZepT.Bing.Dot	542	367	342	0.53	0.52	0.52	0.03	0.04	0.04	2.1	2.0	2.0	0.9	0.9	0.9	9.4	10.4	9.3
10	Zepp.Bing.Dot	544	366	340	0.50	0.48	0.48	0.04	0.05	0.05	2.1	2.0	2.0	0.9	0.8	0.8	10.3	12.0	11.0
12	Tens.Bing.Dot	548	371	345	0.51	0.49	0.48	0.04	0.05	0.05	2.1	2.0	2.0	1.0	0.9	0.9	10.7	12.5	11.5
10	Tens.Wat.Dot	557	385	355	0.49	0.47	0.47	0.04	0.05	0.05	2.1	2.0	2.0	1.0	0.9	0.9	8.1	8.7	8.2
9	Zepp.Bing.	559	398	370	0.64	0.65	0.64				2.2	2.1	2.0	1.0	1.0	1.0	6.6	6.7	6.2
7	ZepT.Wat.Dot	559	390	357	0.53	0.52	0.52	0.03	0.04	0.04	2.1	2.0	2.0	1.0	0.9	0.9	7.4	7.7	7.2
8	Zepp.Wat.Dot	561	389	356	0.50	0.48	0.48	0.04	0.05	0.05	2.1	2.0	2.0	0.9	0.8	0.8	8.0	8.6	8.2
11	Tens.Bing.	561	399	372	0.65	0.65	0.65				2.2	2.1	2.1	1.2	1.1	1.1	7.2	7.4	6.8
9	Tens.Wat.	575	418	384	0.64	0.65	0.65				2.2	2.1	2.1	1.1	1.1	1.0	5.5	5.5	5.3
8	ZepT.Bing.	576	416	385	0.62	0.62	0.62				2.2	2.1	2.1	0.8	0.8	0.8	6.5	6.5	6.0
7	Zepp.Wat.	576	419	383	0.64	0.65	0.65				2.2	2.1	2.1	1.0	1.0	1.0	5.6	5.5	5.3
6	ZepT.Wat.	593	437	398	0.62	0.63	0.63				2.2	2.1	2.1	0.8	0.8	0.8	5.4	5.3	5.1
12	Tens.St.St.Dot	652	464	439	0.23	0.22	0.21	0.07	0.07	0.08	2.0	1.9	1.9	0.8	0.8	0.8			
10	Zepp.St.St.Dot	658	464	437	0.23	0.22	0.21	0.07	0.07	0.08	2.0	1.9	1.9	0.8	0.7	0.7			
12	Tens.St.St.CSF.	674	562	557	0.22	0.24	0.23	0.21	0.18	0.17	1.5	1.5	1.4	0.5	0.5	0.5			
10	Zepp.St.St.CSF.	692	570	565	0.23	0.25	0.24	0.21	0.18	0.17	1.5	1.4	1.4	0.4	0.4	0.5			
8	Ball.Bing.	729	590	583	0.72	0.71	0.71				2.2	2.1	2.1				6.0	6.2	5.9
9	Ball.Bing.Dot	732	593	586	0.72	0.71	0.71	0.00	0.00	0.00	2.2	2.1	2.1				6.0	6.2	5.9
9	Ball.Bing.CSF.	732	593	586	0.72	0.71	0.71	0.00	0.00	0.00	2.2	2.1	2.1				6.0	6.2	5.9
6	Ball.Wat.	745	610	596	0.72	0.72	0.71				2.2	2.1	2.1				5.2	5.3	5.2
7	Ball.Wat.CSF.	748	613	599	0.72	0.72	0.71	0.00	0.00	0.00	2.2	2.1	2.1				5.2	5.3	5.2
7	Ball.Wat.Dot	748	613	599	0.72	0.72	0.71	0.00	0.00	0.00	2.2	2.1	2.1				5.2	5.3	5.2
7	**Zepp.St.Dot**	**784**	**597**	**570**	**0.29**	**0.30**	**0.29**	**0.09**	**0.09**	**0.09**	**1.9**	**1.9**	**1.8**	**0.7**	**0.7**	**0.7**			
10	**Tens.Cyl.CSF.**	**832**	**735**	**739**	**0.29**	**0.31**	**0.31**	**0.24**	**0.22**	**0.21**	**1.3**	**1.3**	**1.3**	**0.3**	**0.4**	**0.4**			
9	ZepT.St.St.Dot	843	652	640	0.33	0.33	0.32	0.05	0.05	0.05	1.8	1.7	1.7	1.1	1.1	1.1			
11	Tens.St.St.	859	687	666	0.28	0.28	0.27				1.7	1.6	1.6	0.9	0.9	0.9			
9	Zepp.St.St.	874	695	674	0.29	0.28	0.27				1.6	1.6	1.6	0.8	0.8	0.8			
9	ZepT.St.St.CSF.	881	718	712	0.35	0.35	0.34	0.12	0.10	0.08	1.5	1.5	1.4	0.9	0.9	0.9			
8	ZepT.St.St.	895	713	696	0.32	0.31	0.30				1.6	1.6	1.5	1.0	1.0	1.0			
8	Ball.St.St.	1,161	978	965	0.28	0.27	0.27				1.5	1.5	1.5						
9	Ball.St.St.Dot	1,162	977	965	0.24	0.26	0.25	0.02	0.02	0.02	1.6	1.5	1.5						
9	Ball.St.St.CSF.	1,164	981	969	0.28	0.27	0.26	0.00	0.00	0.00	1.5	1.5	1.5						

Table C.1: Various model parameters from different data sets of angular thresholds of 2°, 5° and 10°. The models are ordered top-down by the BIC score of 2° data set. [Note: Zepp=Zeppelin; ZepT=Zeppelin with tortuosity;Tens=Tensor; St=Stick; Bing=Bingham; Wat=Watson].

D

TO PART V (METHODOLOGICAL IMPROVEMENTS)

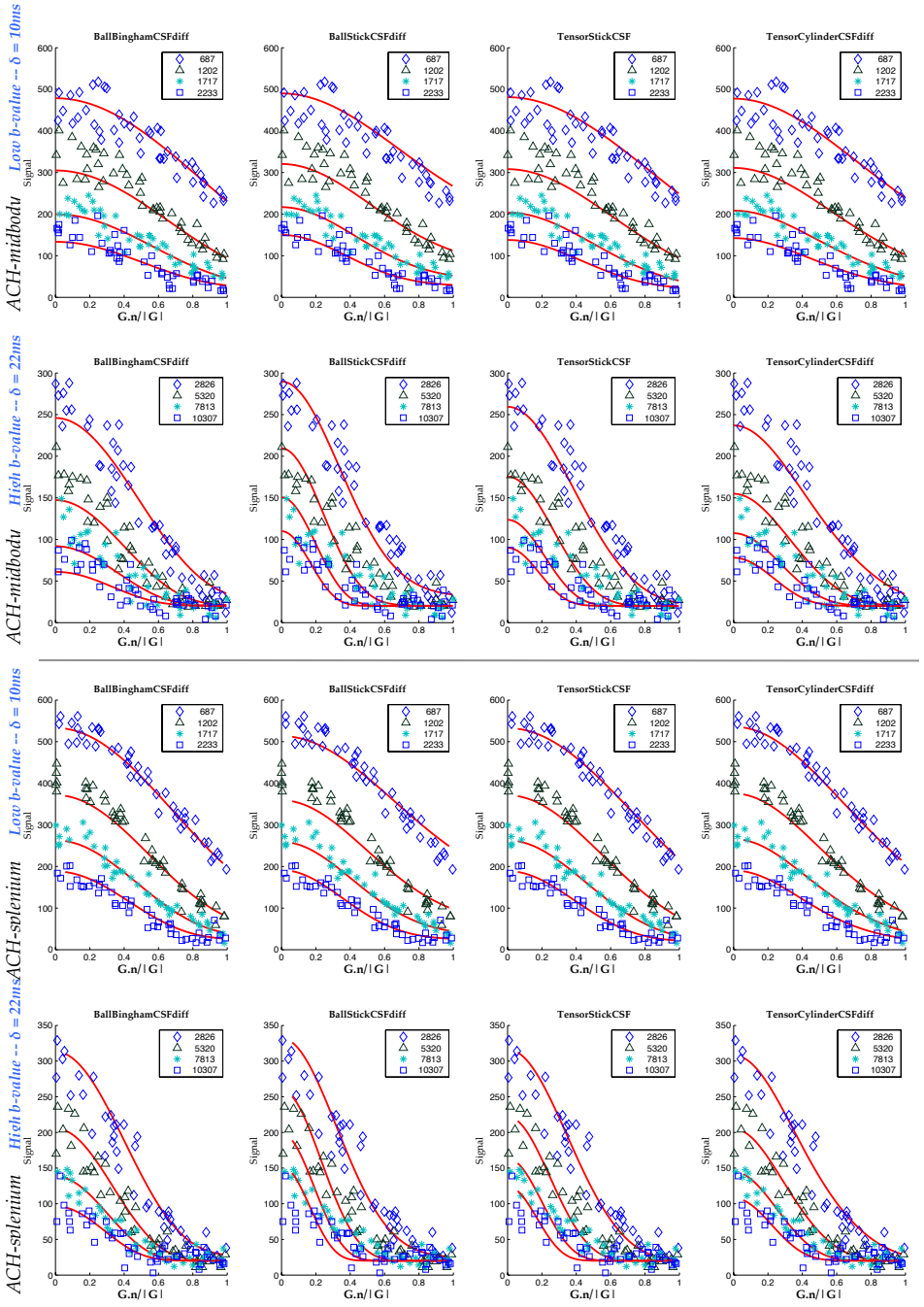


Figure D.1: As in Fig.16.4, the four selected models are fitted to ACH-midbody and ACH-splenium.

	Models	LSE (x 10^3)	Cylinder/Bingham/Stick Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	*Tensor/ Ball* Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
ACH-subject-1	Ball-Bingham-CSF-diff	0.73 $^{0.0}$	0.78 $^{0.1}$	*2.34* $^{0.1}$	52 $^{0.3}$		0.20 $^{0.3}$	*0.51* $^{0.4}$		
	Ball-Bingham-CSF	0.75 $^{0.0}$	0.87 $^{0.0}$	*2.82* $^{0.0}$	10 $^{0.0}$		0.13 $^{0.0}$	*2.82* $^{0.0}$		
	NODDI	0.81 $^{0.0}$	0.74 $^{0.0}$	*1.90* $^{0.0}$	8 $^{0.0}$		0.24 $^{0.0}$	*1.90* $^{0.0}$	*0.45* $^{0.0}$	
	Tensor-Bingham-CSF	0.73 $^{0.0}$	0.60 $^{0.1}$	*1.95* $^{0.1}$	49 $^{1.6}$		0.38 $^{0.2}$	*1.95* $^{0.1}$	*0.33* $^{0.4}$	*0.10* $^{0.7}$
	Tensor-Bingham-CSF-diff	0.73 $^{0.2}$	0.61 $^{6.3}$	*1.96* $^{14.7}$	51 $^{2.9}$		0.37 $^{10.5}$	*1.83* $^{32.4}$	*0.35* $^{6.5}$	*0.15* $^{56.6}$
	Ball-Stick-Stick-CSF-diff	0.65 $^{0.0}$	0.66 $^{0.4}$	*2.12* $^{0.4}$			0.28 $^{1.0}$	*0.58* $^{0.8}$		
	Ball-Stick-Stick-CSF	0.73 $^{0.0}$	0.75 $^{0.0}$	*1.57* $^{0.0}$			0.19 $^{0.0}$	*1.57* $^{0.0}$		
	Tensor-Stick-Stick-CSF	0.59 $^{0.0}$	0.45 $^{0.1}$	*1.63* $^{0.0}$			0.48 $^{0.1}$	*1.63* $^{0.0}$	*0.46* $^{0.1}$	*0.19* $^{0.1}$
	Tensor-Stick-Stick-CSF-diff	0.56 $^{0.0}$	0.37 $^{2.2}$	*0.84* $^{1.7}$			0.57 $^{1.5}$	*2.45* $^{0.6}$	*0.37* $^{2.3}$	*0.16* $^{2.5}$
	Ball-Stick-CSF-diff	0.85 $^{0.0}$	0.66 $^{0.0}$	*2.12* $^{0.0}$			0.32 $^{0.0}$	*0.61* $^{0.0}$		
	Ball-Stick-CSF	0.93 $^{0.0}$	0.76 $^{0.0}$	*1.51* $^{0.0}$			0.22 $^{0.0}$	*1.51* $^{0.0}$		
	Tensor-Stick-CSF	0.77 $^{0.0}$	0.47 $^{0.0}$	*1.58* $^{0.0}$			0.50 $^{0.0}$	*1.58* $^{0.0}$	*0.50* $^{0.0}$	*0.23* $^{0.0}$
	Tensor-Stick-CSF-diff	0.74 $^{0.0}$	0.38 $^{0.0}$	*0.85* $^{0.0}$			0.60 $^{0.0}$	*2.37* $^{0.0}$	*0.41* $^{0.0}$	*0.18* $^{0.0}$
	Ball-Cylinder-CSF-diff	0.85 $^{0.0}$	0.66 $^{0.2}$	*2.11* $^{0.4}$		0.1 $^{2.9}$	0.32 $^{0.3}$	*0.61* $^{0.6}$		
	Ball-Cylinder-CSF	0.93 $^{0.1}$	0.75 $^{0.5}$	*1.50* $^{0.7}$		0.1 $^{26.1}$	0.23 $^{1.7}$	*1.50* $^{0.7}$		
	MMWMD	0.86 $^{0.0}$	0.64 $^{0.1}$	*1.70* $^{0.0}$		0.1 $^{40.0}$	0.34 $^{0.2}$	*1.70* $^{0.0}$	*0.58* $^{0.2}$	
	Tensor-Cylinder-CSF	0.77 $^{0.0}$	0.48 $^{2.4}$	*1.58* $^{0.1}$		0.3 $^{41.9}$	0.50 $^{2.3}$	*1.58* $^{0.1}$	*0.51* $^{3.9}$	*0.24* $^{3.9}$
	Tensor-Cylinder-CSF-diff	0.74 $^{0.2}$	0.37 $^{13.7}$	*0.84* $^{10.6}$		0.1 $^{51.7}$	0.60 $^{8.5}$	*2.36* $^{3.9}$	*0.41* $^{9.7}$	*0.18* $^{8.2}$
	NODDI+	0.76 $^{0.0}$	0.87 $^{0.1}$	*2.80* $^{0.1}$	6 $^{0.3}$		0.13 $^{0.6}$	*2.80* $^{0.1}$	*2.80* $^{0.1}$	
	MMWMD+	0.82 $^{0.0}$	0.51 $^{1.0}$	*1.58* $^{0.0}$		0.1 $^{22.3}$	0.46 $^{1.1}$	*1.58* $^{0.0}$	*0.40* $^{1.5}$	
ACH-subject-2	Ball-Bingham-CSF-diff	1.23 $^{0.2}$	0.70 $^{1.9}$	*3.00* $^{0.0}$	37 $^{20.7}$		0.29 $^{4.6}$	*1.01* $^{2.2}$		
	Ball-Bingham-CSF	1.34 $^{0.0}$	0.89 $^{0.0}$	*3.00* $^{0.0}$	10 $^{0.0}$		0.10 $^{0.0}$	*3.00* $^{0.0}$		
	NODDI	1.35 $^{0.0}$	0.51 $^{0.0}$	*1.90* $^{0.0}$	55 $^{0.0}$		0.49 $^{0.0}$	*1.90* $^{0.0}$	*0.92* $^{0.0}$	
	Tensor-Bingham-CSF	1.20 $^{0.0}$	0.49 $^{1.2}$	*2.70* $^{0.1}$	50 $^{0.6}$		0.51 $^{1.1}$	*2.70* $^{0.1}$	*0.57* $^{1.2}$	*0.17* $^{1.7}$
	Tensor-Bingham-CSF-diff	1.20 $^{0.3}$	0.42 $^{3.7}$	*2.79* $^{13.4}$	49 $^{6.1}$		0.57 $^{2.7}$	*2.57* $^{6.3}$	*0.54* $^{23.3}$	*0.14* $^{16.3}$
	Ball-Stick-Stick-CSF-diff	1.19 $^{0.0}$	0.66 $^{0.6}$	*2.22* $^{0.3}$			0.30 $^{1.4}$	*0.65* $^{2.2}$		
	Ball-Stick-Stick-CSF	1.32 $^{0.0}$	0.76 $^{0.1}$	*1.62* $^{0.1}$			0.20 $^{0.2}$	*1.62* $^{0.1}$		
	Tensor-Stick-Stick-CSF	1.09 $^{0.0}$	0.00 $^{76.7}$	*1.71* $^{0.0}$			0.96 $^{0.0}$	*1.71* $^{0.0}$	*0.22* $^{0.9}$	*0.11* $^{1.9}$
	Tensor-Stick-Stick-CSF-diff	1.09 $^{0.4}$	0.20 $^{79.0}$	*1.58* $^{25.9}$			0.78 $^{17.7}$	*2.27* $^{26.2}$	*0.53* $^{55.2}$	*0.25* $^{53.0}$
	Ball-Stick-CSF-diff	1.37 $^{0.0}$	0.55 $^{0.0}$	*3.00* $^{0.0}$			0.44 $^{0.0}$	*0.92* $^{0.0}$		
	Ball-Stick-CSF	1.49 $^{0.0}$	0.63 $^{0.0}$	*1.86* $^{0.0}$			0.36 $^{0.0}$	*1.86* $^{0.0}$		
	Tensor-Stick-CSF	1.23 $^{0.0}$	0.39 $^{0.0}$	*2.03* $^{0.0}$			0.60 $^{0.0}$	*2.03* $^{0.0}$	*0.86* $^{0.0}$	*0.42* $^{0.0}$
	Tensor-Stick-CSF-diff	1.21 $^{0.0}$	0.32 $^{0.0}$	*1.06* $^{0.0}$			0.67 $^{0.0}$	*2.93* $^{0.0}$	*0.75* $^{0.0}$	*0.36* $^{0.0}$
	Ball-Cylinder-CSF-diff	1.36 $^{0.0}$	0.60 $^{0.0}$	*3.00* $^{0.0}$		5.3 $^{0.1}$	0.40 $^{0.1}$	*0.81* $^{0.2}$		
	Ball-Cylinder-CSF	1.49 $^{0.0}$	0.63 $^{0.2}$	*1.85* $^{0.2}$		0.2 $^{82.2}$	0.36 $^{0.4}$	*1.85* $^{0.2}$		
	MMWMD	1.38 $^{0.0}$	0.62 $^{0.4}$	*1.70* $^{0.0}$		0.2 $^{51.1}$	0.37 $^{0.7}$	*1.70* $^{0.0}$	*0.63* $^{0.7}$	
	Tensor-Cylinder-CSF	1.23 $^{0.0}$	0.39 $^{0.0}$	*2.04* $^{0.0}$		4.2 $^{0.0}$	0.60 $^{0.0}$	*2.04* $^{0.0}$	*0.79* $^{0.0}$	*0.38* $^{0.0}$
	Tensor-Cylinder-CSF-diff	1.21 $^{0.0}$	0.33 $^{0.0}$	*1.06* $^{0.0}$		2.7 $^{0.0}$	0.67 $^{0.0}$	*2.92* $^{0.0}$	*0.74* $^{0.0}$	*0.35* $^{0.0}$
	NODDI+	1.30 $^{0.8}$	0.25 $^{51.6}$	*2.25* $^{23.1}$	32 $^{93.4}$		0.75 $^{17.0}$	*2.25* $^{23.1}$	*0.19* $^{32.3}$	
	MMWMD+	1.32 $^{0.0}$	0.00 $^{135.5}$	*1.69* $^{0.5}$		0.7 $^{86.3}$	0.99 $^{0.0}$	*1.69* $^{0.5}$	*0.18* $^{0.6}$	

Table D.1: Variation of parameter estimates across subjects 1 and 2. The mean T_2 on ACH-subject-1 was 54ms, and 59ms on the ACH-subject-2; the CSF volume fraction is at 2% and 1% respectively.

	Models	LSE (x 10^3)	Cylinder/Bingham/Stick Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	Tensor/ Ball Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
ACH-subject-3	Ball-Bingham-CSF-diff	2.79 $^{0.0}$	0.78 $^{2.2}$	*3.00* $^{0.0}$	8 $^{7.8}$		0.12 $^{14.2}$	*0.47* $^{1.9}$		
	Ball-Bingham-CSF	2.81 $^{0.0}$	0.90 $^{0.0}$	*2.91* $^{0.2}$	6 $^{1.3}$		0.00 $^{165.7}$	*2.91* $^{0.2}$		
	NODDI	2.91 $^{0.0}$	0.56 $^{0.0}$	*1.90* $^{0.0}$	13 $^{0.0}$		0.34 $^{0.0}$	*1.90* $^{0.0}$	*0.65* $^{0.0}$	
	Tensor-Bingham-CSF	2.80 $^{0.0}$	0.66 $^{0.0}$	*2.74* $^{0.0}$	6 $^{0.0}$		0.25 $^{0.0}$	*2.74* $^{0.0}$	*0.18* $^{0.0}$	*0.04* $^{0.0}$
	Tensor-Bingham-CSF-diff	2.79 $^{0.0}$	0.63 $^{1.6}$	*3.00* $^{0.0}$	7 $^{7.2}$		0.27 $^{3.8}$	*2.07* $^{3.9}$	*0.19* $^{12.4}$	*0.05* $^{5.5}$
	Ball-Stick-Stick-CSF-diff	1.70 $^{0.0}$	0.49 $^{1.4}$	*2.83* $^{2.2}$			0.37 $^{1.9}$	*0.70* $^{3.0}$		
	Ball-Stick-Stick-CSF	1.77 $^{0.0}$	0.59 $^{0.0}$	*1.81* $^{0.0}$			0.27 $^{0.0}$	*1.81* $^{0.0}$		
	Tensor-Stick-Stick-CSF	1.66 $^{0.0}$	0.41 $^{0.0}$	*1.92* $^{0.0}$			0.46 $^{0.0}$	*1.92* $^{0.0}$	*0.77* $^{0.0}$	*0.45* $^{0.0}$
	Tensor-Stick-Stick-CSF-diff	1.58 $^{0.2}$	0.32 $^{3.4}$	*0.75* $^{8.5}$			0.55 $^{2.0}$	*2.99* $^{0.5}$	*0.56* $^{4.5}$	*0.43* $^{6.5}$
	Ball-Stick-CSF-diff	2.99 $^{0.0}$	0.53 $^{0.0}$	*3.00* $^{0.0}$			0.37 $^{0.1}$	*0.58* $^{0.1}$		
	Ball-Stick-CSF	3.04 $^{0.0}$	0.64 $^{0.0}$	*1.69* $^{0.0}$			0.26 $^{0.0}$	*1.69* $^{0.0}$		
	Tensor-Stick-CSF	2.93 $^{0.0}$	0.44 $^{0.0}$	*1.80* $^{0.0}$			0.46 $^{0.0}$	*1.80* $^{0.0}$	*0.64* $^{0.0}$	*0.41* $^{0.0}$
	Tensor-Stick-CSF-diff	2.80 $^{0.0}$	0.31 $^{4.0}$	*0.65* $^{6.1}$			0.60 $^{2.0}$	*3.00* $^{0.0}$	*0.43* $^{4.1}$	*0.30* $^{3.1}$
	Ball-Cylinder-CSF-diff	2.92 $^{0.0}$	0.61 $^{0.0}$	*2.99* $^{0.0}$		6.4 $^{0.0}$	0.29 $^{0.0}$	*0.37* $^{0.0}$		
	Ball-Cylinder-CSF	3.04 $^{0.0}$	0.66 $^{0.0}$	*1.70* $^{0.0}$		2.9 $^{0.0}$	0.25 $^{0.0}$	*1.70* $^{0.0}$		
	MMWMD	2.95 $^{0.0}$	0.54 $^{0.0}$	*1.70* $^{0.0}$		4.4 $^{0.0}$	0.37 $^{0.0}$	*1.70* $^{0.0}$	*0.62* $^{0.0}$	
	Tensor-Cylinder-CSF	2.91 $^{0.0}$	0.21 $^{1.3}$	*1.81* $^{0.1}$		10.1 $^{0.0}$	0.69 $^{0.4}$	*1.81* $^{0.1}$	*0.20* $^{1.5}$	*0.12* $^{0.4}$
	Tensor-Cylinder-CSF-diff	2.80 $^{0.0}$	0.30 $^{7.8}$	*0.62* $^{12.9}$		1.9 $^{17.6}$	0.61 $^{3.8}$	*3.00* $^{0.0}$	*0.42* $^{6.6}$	*0.28* $^{4.5}$
	NODDI+	2.82 $^{0.0}$	0.65 $^{0.0}$	*2.73* $^{0.0}$	4 $^{0.0}$		0.26 $^{0.0}$	*2.73* $^{0.0}$	*0.11* $^{0.0}$	
	MMWMD+	2.93 $^{0.0}$	0.22 $^{0.6}$	*1.81* $^{0.0}$		10.1 $^{0.0}$	0.68 $^{0.2}$	*1.81* $^{0.0}$	*0.16* $^{0.3}$	
ACH-subject-4	Ball-Bingham-CSF-diff	1.50 $^{0.2}$	0.75 $^{2.3}$	*3.00* $^{0.0}$	11 $^{14.5}$		0.23 $^{7.6}$	*1.11* $^{6.4}$		
	Ball-Bingham-CSF	1.54 $^{0.0}$	0.89 $^{0.4}$	*3.00* $^{0.0}$	7 $^{3.6}$		0.09 $^{3.7}$	*3.00* $^{0.0}$		
	NODDI	1.60 $^{0.0}$	0.52 $^{0.0}$	*1.90* $^{0.0}$	58 $^{0.0}$		0.46 $^{0.0}$	*1.90* $^{0.0}$	*0.88* $^{0.0}$	
	Tensor-Bingham-CSF	1.44 $^{0.0}$	0.50 $^{0.0}$	*3.00* $^{0.0}$	4 $^{0.1}$		0.48 $^{0.0}$	*3.00* $^{0.0}$	*0.28* $^{0.2}$	*0.15* $^{0.2}$
	Tensor-Bingham-CSF-diff	1.44 $^{0.1}$	0.47 $^{3.3}$	*3.00* $^{0.0}$	3 $^{10.6}$		0.51 $^{3.0}$	*2.90* $^{1.8}$	*0.28* $^{3.9}$	*0.15* $^{6.7}$
	Ball-Stick-Stick-CSF-diff	1.41 $^{0.0}$	0.55 $^{0.7}$	*2.91* $^{1.5}$			0.42 $^{1.0}$	*0.96* $^{1.9}$		
	Ball-Stick-Stick-CSF	1.50 $^{0.0}$	0.63 $^{0.0}$	*1.93* $^{0.0}$			0.34 $^{0.0}$	*1.93* $^{0.0}$		
	Tensor-Stick-Stick-CSF	1.30 $^{0.0}$	0.34 $^{0.1}$	*2.12* $^{0.0}$			0.63 $^{0.1}$	*2.12* $^{0.0}$	*0.66* $^{0.1}$	*0.42* $^{0.1}$
	Tensor-Stick-Stick-CSF-diff	1.24 $^{0.1}$	0.25 $^{4.1}$	*0.80* $^{4.4}$			0.72 $^{1.5}$	*3.00* $^{0.0}$	*0.52* $^{3.5}$	*0.40* $^{2.2}$
	Ball-Stick-CSF-diff	1.67 $^{0.0}$	0.56 $^{0.0}$	*3.00* $^{0.0}$			0.42 $^{0.0}$	*0.90* $^{0.0}$		
	Ball-Stick-CSF	1.74 $^{0.0}$	0.64 $^{0.0}$	*1.90* $^{0.0}$			0.33 $^{0.0}$	*1.90* $^{0.0}$		
	Tensor-Stick-CSF	1.55 $^{0.0}$	0.36 $^{0.0}$	*2.08* $^{0.0}$			0.61 $^{0.0}$	*2.08* $^{0.0}$	*0.65* $^{0.0}$	*0.46* $^{0.0}$
	Tensor-Stick-CSF-diff	1.46 $^{0.0}$	0.25 $^{5.9}$	*0.72* $^{9.5}$			0.72 $^{2.0}$	*3.00* $^{0.0}$	*0.51* $^{3.9}$	*0.37* $^{2.9}$
	Ball-Cylinder-CSF-diff	1.58 $^{0.0}$	0.68 $^{0.2}$	*3.00* $^{0.0}$		7.2 $^{1.5}$	0.30 $^{0.6}$	*0.54* $^{1.4}$		
	Ball-Cylinder-CSF	1.73 $^{0.0}$	0.72 $^{0.0}$	*1.93* $^{0.0}$		5.1 $^{0.0}$	0.26 $^{0.0}$	*1.93* $^{0.0}$		
	MMWMD	1.71 $^{0.0}$	0.55 $^{0.0}$	*1.70* $^{0.0}$		5.9 $^{0.0}$	0.43 $^{0.0}$	*1.70* $^{0.0}$	*0.73* $^{0.0}$	
	Tensor-Cylinder-CSF	1.49 $^{0.0}$	0.32 $^{1.6}$	*2.06* $^{0.1}$		10.1 $^{0.0}$	0.66 $^{0.8}$	*2.06* $^{0.1}$	*0.28* $^{0.6}$	*0.17* $^{0.8}$
	Tensor-Cylinder-CSF-diff	1.45 $^{0.0}$	0.33 $^{4.5}$	*0.91* $^{5.8}$		6.8 $^{2.1}$	0.65 $^{2.2}$	*3.00* $^{0.0}$	*0.41* $^{2.7}$	*0.27* $^{2.3}$
	NODDI+	1.47 $^{0.0}$	0.50 $^{0.1}$	*3.00* $^{0.0}$	3 $^{0.2}$		0.48 $^{0.1}$	*3.00* $^{0.0}$	*0.21* $^{0.2}$	
	MMWMD+	1.51 $^{0.0}$	0.32 $^{0.2}$	*2.06* $^{0.1}$		10.1 $^{0.0}$	0.66 $^{0.1}$	*2.06* $^{0.1}$	*0.22* $^{0.2}$	

Table D.2: Variation of parameter estimates across subjects 3 and 4. These datasets contain considerably more motion artefacts. Notably, the Bingham distribution reflects much higher dispersion, and the Cylinder much higher thickness, than for the first two subjects in Fig.D.1. The ACH-subject-3 mean T_2 is 64ms, ACH-subject-2 T_2 is 59ms; the CSF volume fraction is 10% and 2% respectively.

			Cylinder / Bingham / Stick			Tensor/ Ball			
	Models	LSE (x 10^3)	Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
ACH-midbody	Ball-Bingham-CSF-diff	3.70 0.0	0.75 0.0	*2.11* *0.0*	11 0.0		0.19 0.0	*0.51* *0.0*		
	Ball-Bingham-CSF	3.87 0.0	0.94 0.0	*2.15* *0.1*	6 0.4		0.00 88.0	*2.15* *0.1*		
	NODDI	3.86 0.0	0.76 0.0	*1.90* *0.0*	5 0.0		0.17 0.0	*1.90* *0.0*	*0.33* *0.0*	
	Tensor-Bingham-CSF	3.70 0.0	0.68 0.0	*1.72* *0.0*	15 0.0		0.25 0.0	*1.72* *0.0*	*0.55* *0.0*	*0.21* *0.0*
	Tensor-Bingham-CSF-diff	3.69 0.1	0.57 16.4	*1.29* *38.6*	24 28.0		0.36 25.7	*1.87* *39.9*	*0.52* *6.0*	*0.36* *9.3*
	Ball-Stick-Stick-CSF-diff	4.07 0.0	0.54 0.3	*1.60* *1.0*			0.32 0.4	*0.61* *1.7*		
	Ball-Stick-Stick-CSF	4.32 0.0	0.59 0.0	*1.24* *0.0*			0.27 0.0	*1.24* *0.0*		
	Tensor-Stick-Stick-CSF	3.86 0.0	0.44 0.0	*1.33* *0.0*			0.41 0.0	*1.33* *0.0*	*0.53* *0.0*	*0.39* *0.0*
	Tensor-Stick-Stick-CSF-diff	3.63 0.0	0.36 2.4	*0.66* *4.8*			0.49 1.8	*2.12* *3.8*	*0.43* *1.4*	*0.28* *4.6*
	Ball-Stick-CSF-diff	4.39 0.0	0.57 0.0	*1.54* *0.0*			0.36 0.0	*0.69* *0.0*		
	Ball-Stick-CSF	4.57 0.0	0.62 0.0	*1.15* *0.0*			0.32 0.0	*1.15* *0.0*		
	Tensor-Stick-CSF	4.00 0.0	0.46 0.0	*1.26* *0.0*			0.47 0.0	*1.26* *0.0*	*0.57* *0.0*	*0.35* *0.0*
	Tensor-Stick-CSF-diff	3.71 0.0	0.39 0.0	*0.66* *0.0*			0.54 0.0	*2.18* *0.0*	*0.48* *0.0*	*0.29* *0.0*
	Ball-Cylinder-CSF-diff	4.00 0.0	0.83 0.0	*1.57* *0.0*		6.9 0.0	0.10 0.0	*0.05* *0.0*		
	Ball-Cylinder-CSF	4.13 0.0	0.86 0.0	*1.25* *0.0*		6.0 0.0	0.07 0.0	*1.25* *0.0*		
	MMWMD	5.17 0.0	0.74 0.0	*1.70* *0.0*		5.7 0.0	0.20 0.0	*1.70* *0.0*	*0.33* *0.0*	
	Tensor-Cylinder-CSF	3.89 0.0	0.44 0.4	*1.26* *0.0*		10.1 0.0	0.50 0.3	*1.26* *0.0*	*0.06* *0.5*	*0.00* *79.6*
	Tensor-Cylinder-CSF-diff	3.67 0.0	0.51 0.0	*0.77* *0.0*		4.3 0.0	0.42 0.0	*2.51* *0.0*	*0.49* *0.0*	*0.25* *0.0*
	NODDI+	3.83 0.0	0.69 0.0	*1.66* *0.0*	6 0.0		0.25 0.0	*1.66* *0.0*	*0.47* *0.0*	
	MMWMD+	3.99 0.4	0.51 15.1	*1.26* *0.4*		8.9 23.9	0.43 18.0	*1.26* *0.4*	*0.13* *147.2*	
ACH-splenium	Ball-Bingham-CSF-diff	1.84 0.0	0.81 0.0	*2.49* *0.0*	10 0.0		0.19 0.0	*0.64* *0.0*		
	Ball-Bingham-CSF	2.09 0.0	0.93 0.0	*2.31* *0.0*	7 0.0		0.07 0.0	*2.31* *0.0*		
	NODDI	2.02 0.0	0.75 0.0	*1.90* *0.0*	8 0.0		0.25 0.0	*1.90* *0.0*	*0.47* *0.0*	
	Tensor-Bingham-CSF	1.85 0.0	0.74 0.0	*2.00* *0.0*	13 0.0		0.26 0.0	*2.00* *0.0*	*0.74* *0.0*	*0.41* *0.0*
	Tensor-Bingham-CSF-diff	1.83 0.0	0.61 0.0	*1.39* *0.0*	23 0.0		0.39 0.0	*2.82* *0.0*	*0.58* *0.0*	*0.46* *0.0*
	Ball-Stick-Stick-CSF-diff	2.88 0.0	0.63 0.5	*1.69* *1.8*			0.37 0.9	*0.96* *3.0*		
	Ball-Stick-Stick-CSF	3.06 0.0	0.66 0.1	*1.40* *0.1*			0.34 0.2	*1.40* *0.1*		
	Tensor-Stick-Stick-CSF	2.36 0.0	0.52 0.7	*1.54* *0.3*			0.48 0.7	*1.54* *0.3*	*0.61* *1.0*	*0.47* *1.9*
	Tensor-Stick-Stick-CSF-diff	1.90 0.0	0.44 0.7	*0.86* *1.0*			0.56 0.5	*2.70* *0.4*	*0.49* *1.3*	*0.37* *1.7*
	Ball-Stick-CSF-diff	2.88 0.0	0.63 0.0	*1.71* *0.0*			0.37 0.0	*0.94* *0.0*		
	Ball-Stick-CSF	3.06 0.0	0.66 0.0	*1.40* *0.0*			0.34 0.0	*1.40* *0.0*		
	Tensor-Stick-CSF	2.35 0.0	0.51 0.0	*1.54* *0.0*			0.49 0.0	*1.54* *0.0*	*0.59* *0.0*	*0.46* *0.0*
	Tensor-Stick-CSF-diff	1.90 0.0	0.44 0.0	*0.86* *0.0*			0.56 0.0	*2.72* *0.0*	*0.49* *0.0*	*0.37* *0.0*
	Ball-Cylinder-CSF-diff	2.61 0.0	0.83 0.0	*1.53* *0.0*		5.6 0.0	0.17 0.0	*1.28* *0.0*		
	Ball-Cylinder-CSF	2.61 0.0	0.85 0.0	*1.49* *0.0*		5.7 0.0	0.15 0.0	*1.49* *0.0*		
	MMWMD	2.65 0.0	0.75 0.0	*1.70* *0.0*		5.4 0.0	0.25 0.0	*1.70* *0.0*	*0.42* *0.0*	
	Tensor-Cylinder-CSF	2.33 0.0	0.58 0.0	*1.54* *0.0*		4.3 0.0	0.42 0.0	*1.54* *0.0*	*0.60* *0.0*	*0.44* *0.0*
	Tensor-Cylinder-CSF-diff	1.90 0.0	0.46 0.0	*0.88* *0.0*		2.9 0.0	0.54 0.0	*2.79* *0.0*	*0.49* *0.0*	*0.36* *0.0*
	NODDI+	2.01 0.0	0.71 0.0	*1.84* *0.0*	8 0.0		0.29 0.0	*1.84* *0.0*	*0.43* *0.0*	
	MMWMD+	2.30 0.0	0.61 0.0	*1.49* *0.0*		4.7 0.0	0.39 0.0	*1.49* *0.0*	*0.38* *0.0*	

Table D.3: Similar to Tab.16.1, but using datasets ACH-midbody and ACH-splenium.

	Models	LSE (x 10^3)	Cylinder/Bingham/Stick Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	*Tensor/ Ball* Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
ACH8-midbody	Ball-Bingham-CSF-diff	3.63 0.0	0.79 0.0	*2.53* 0.0	6 0.0		0.21 0.0	*2.01* 0.0		
	Ball-Bingham-CSF	3.65 0.0	0.78 0.0	*2.36* 0.0	6 0.0		0.22 0.0	*2.36* 0.0		
	NODDI	4.22 0.0	0.54 0.0	*1.90* 0.0	7 0.0		0.46 0.0	*1.90* 0.0	*0.88* 0.0	
	Tensor-Bingham-CSF	3.63 0.0	0.76 0.0	*2.37* 0.0	6 0.0		0.24 0.0	*2.37* 0.0	*1.92* 0.0	*1.87* 0.0
	Tensor-Bingham-CSF-diff	3.63 0.0	0.78 0.0	*2.51* 0.0	6 0.0		0.22 0.0	*2.07* 0.0	*2.03* 0.0	*1.94* 0.0
	Ball-Stick-Stick-CSF-diff	4.83 0.0	0.51 0.3	*1.24* 1.2			0.49 0.3	*1.64* 1.1		
	Ball-Stick-Stick-CSF	4.92 0.1	0.50 0.5	*1.43* 0.4			0.50 0.5	*1.43* 0.4		
	Tensor-Stick-Stick-CSF	4.82 0.1	0.48 0.8	*1.50* 0.7			0.52 0.7	*1.50* 0.7	*1.31* 1.7	*1.08* 1.0
	Tensor-Stick-Stick-CSF-diff	3.88 0.1	0.40 3.0	*0.74* 4.1			0.60 2.0	*2.92* 0.8	*1.01* 5.8	*0.81* 4.0
	Ball-Stick-CSF-diff	4.83 0.0	0.51 0.0	*1.24* 0.0			0.49 0.0	*1.66* 0.0		
	Ball-Stick-CSF	4.92 0.0	0.50 0.0	*1.43* 0.0			0.50 0.0	*1.43* 0.0		
	Tensor-Stick-CSF	4.81 0.0	0.47 0.0	*1.49* 0.0			0.53 0.0	*1.49* 0.0	*1.31* 0.0	*1.03* 0.0
	Tensor-Stick-CSF-diff	3.87 0.0	0.40 0.0	*0.75* 0.0			0.60 0.0	*3.00* 0.0	*1.01* 0.1	*0.78* 0.0
	Ball-Cylinder-CSF-diff	4.49 0.0	0.70 0.0	*1.25* 0.0		5.9 0.0	0.30 0.0	*2.49* 0.0		
	Ball-Cylinder-CSF	4.78 0.0	0.60 0.0	*1.49* 0.0		5.1 0.0	0.40 0.0	*1.49* 0.0		
	MMWMD	5.10 0.0	0.46 0.0	*1.70* 0.0		3.3 0.0	0.54 0.0	*1.70* 0.0	*0.92* 0.0	
	Tensor-Cylinder-CSF	4.75 0.0	0.57 0.3	*1.50* 0.1		4.8 0.5	0.43 0.4	*1.50* 0.1	*1.50* 0.1	*1.17* 0.9
	Tensor-Cylinder-CSF-diff	3.87 0.0	0.41 0.1	*0.77* 0.2		2.4 0.6	0.59 0.1	*3.00* 0.0	*1.01* 0.2	*0.78* 0.2
	NODDI+	3.72 0.0	0.76 0.0	*2.38* 0.0	5 0.0		0.24 0.0	*2.38* 0.0	*1.92* 0.0	
	MMWMD+	4.80 0.0	0.60 0.1	*1.49* 0.0		5.1 0.1	0.40 0.1	*1.49* 0.0	*1.49* 0.0	
ACH8-splenium	Ball-Bingham-CSF-diff	2.78 0.0	0.80 0.0	*2.23* 0.0	28 0.0		0.20 0.0	*1.19* 0.0		
	Ball-Bingham-CSF	2.88 0.0	0.82 0.0	*2.01* 0.0	28 0.0		0.18 0.0	*2.01* 0.0		
	NODDI	3.03 0.5	0.68 2.5	*1.90* 0.0	14 12.7		0.32 5.4	*1.90* 0.0	*0.60* 5.4	
	Tensor-Bingham-CSF	2.78 0.0	0.76 0.0	*1.97* 0.1	44 0.0		0.24 0.1	*1.97* 0.1	*1.18* 0.1	*0.98* 0.1
	Tensor-Bingham-CSF-diff	2.77 0.0	0.79 0.1	*2.19* 0.3	30 2.3		0.21 0.4	*1.28* 0.9	*1.28* 0.9	*1.09* 0.2
	Ball-Stick-Stick-CSF-diff	3.28 2.8	0.68 0.2	*1.71* 0.9			0.31 5.1	*1.17* 7.0		
	Ball-Stick-Stick-CSF	3.34 2.3	0.69 0.6	*1.55* 2.2			0.29 6.3	*1.55* 2.2		
	Tensor-Stick-Stick-CSF	3.03 2.1	0.60 0.5	*1.65* 2.1			0.39 2.9	*1.65* 2.1	*0.84* 14.1	*0.52* 1.1
	Tensor-Stick-Stick-CSF-diff	2.84 1.6	0.52 0.3	*1.16* 1.7			0.46 2.7	*2.88* 2.7	*0.65* 13.9	*0.37* 0.9
	Ball-Stick-CSF-diff	3.33 1.1	0.68 0.1	*1.72* 0.8			0.32 0.3	*1.19* 5.9		
	Ball-Stick-CSF	3.40 0.5	0.69 0.5	*1.54* 2.3			0.31 1.1	*1.54* 2.3		
	Tensor-Stick-CSF	3.08 0.4	0.60 0.1	*1.65* 2.2			0.40 0.1	*1.65* 2.2	*0.87* 10.1	*0.56* 4.5
	Tensor-Stick-CSF-diff	2.89 0.0	0.53 0.0	*1.18* 0.0			0.47 0.0	*2.96* 0.0	*0.76* 0.0	*0.39* 0.0
	Ball-Cylinder-CSF-diff	3.11 1.0	0.86 0.4	*1.62* 1.0		5.6 0.2	0.14 2.4	*1.69* 11.0		
	Ball-Cylinder-CSF	3.11 1.0	0.85 1.4	*1.63* 2.1		5.6 0.8	0.15 8.1	*1.63* 2.1		
	MMWMD	3.16 0.4	0.68 4.0	*1.70* 0.0		4.7 6.0	0.32 8.4	*1.70* 0.0	*0.55* 8.4	
	Tensor-Cylinder-CSF	3.02 0.3	0.73 0.7	*1.65* 2.2		4.9 0.0	0.27 1.9	*1.65* 2.2	*1.06* 12.9	*0.54* 9.0
	Tensor-Cylinder-CSF-diff	2.86 0.0	0.61 1.6	*1.28* 1.3		4.2 1.9	0.39 2.4	*3.00* 0.0	*0.79* 3.8	*0.35* 2.7
	NODDI+	2.98 0.1	0.74 0.8	*1.87* 2.5	14 1.0		0.26 2.1	*1.87* 2.5	*0.91* 10.3	
	MMWMD+	3.11 0.6	0.77 0.6	*1.64* 2.0		5.1 0.4	0.23 2.1	*1.64* 2.0	*0.86* 13.5	

Table D.4: Similar to Tab.17.2, but using datasets ACH8-midbody and ACH8-splenium.

E

TO PART VI (HIGHER GRADIENTS)

		Cylinder/Bingham/Stick					Tensor/ Ball			
	Models	LSE (x 10^3)	Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
CON-midbody	Ball-Bingham-CSF-diff	7.58 $^{0.0}$	0.70 $^{0.0}$	*2.21* $^{0.0}$	8 $^{0.0}$		0.28 $^{0.0}$	*0.58* $^{0.0}$		
	Ball-Bingham-CSF	10.30 $^{0.0}$	0.89 $^{0.3}$	*1.98* $^{0.3}$	5 $^{1.0}$		0.09 $^{2.6}$	*1.98* $^{0.3}$		
	NODDI	7.92 $^{0.0}$	0.73 $^{0.0}$	*1.90* $^{0.0}$	4 $^{0.0}$		0.25 $^{0.0}$	*1.90* $^{0.0}$	*0.48* $^{0.0}$	
	Tensor-Bingham-CSF	6.93 $^{0.0}$	0.57 $^{0.1}$	*1.58* $^{0.1}$	7 $^{0.6}$		0.41 $^{0.1}$	*1.58* $^{0.1}$	*0.48* $^{0.4}$	*0.42* $^{0.3}$
	Tensor-Bingham-CSF-diff	6.91 $^{0.0}$	0.56 $^{0.1}$	*1.52* $^{0.1}$	7 $^{0.2}$		0.42 $^{0.1}$	*1.67* $^{0.0}$	*0.48* $^{0.2}$	*0.41* $^{0.1}$
	Ball-Stick-Stick-CSF-diff	12.56 $^{0.0}$	0.46 $^{0.2}$	*1.75* $^{1.0}$			0.49 $^{0.2}$	*0.61* $^{1.1}$		
	Ball-Stick-Stick-CSF	14.82 $^{1.1}$	0.48 $^{4.4}$	*1.05* $^{3.1}$			0.50 $^{8.3}$	*1.05* $^{3.1}$		
	Tensor-Stick-Stick-CSF	9.14 $^{0.0}$	0.32 $^{0.0}$	*1.23* $^{0.0}$			0.64 $^{0.0}$	*1.23* $^{0.0}$	*0.47* $^{0.0}$	*0.34* $^{0.0}$
	Tensor-Stick-Stick-CSF-diff	6.88 $^{0.0}$	0.25 $^{0.0}$	*0.40* $^{0.0}$			0.71 $^{0.0}$	*1.83* $^{0.0}$	*0.40* $^{0.0}$	*0.29* $^{0.0}$
	Ball-Stick-CSF-diff	13.35 $^{0.0}$	0.47 $^{0.0}$	*1.72* $^{0.0}$			0.51 $^{0.0}$	*0.63* $^{0.0}$		
	Ball-Stick-CSF	15.17 $^{3.3}$	0.48 $^{5.0}$	*1.05* $^{3.9}$			0.51 $^{6.2}$	*1.05* $^{3.9}$		
	Tensor-Stick-CSF	9.64 $^{0.0}$	0.33 $^{0.0}$	*1.21* $^{0.0}$			0.66 $^{0.0}$	*1.21* $^{0.0}$	*0.48* $^{0.0}$	*0.34* $^{0.0}$
	Tensor-Stick-CSF-diff	7.35 $^{0.0}$	0.25 $^{0.0}$	*0.40* $^{0.0}$			0.73 $^{0.0}$	*1.81* $^{0.0}$	*0.42* $^{0.0}$	*0.29* $^{0.0}$
	Ball-Cylinder-CSF-diff	11.82 $^{0.0}$	0.61 $^{0.0}$	*1.35* $^{0.0}$		3.6 $^{0.0}$	0.37 $^{0.0}$	*0.78* $^{0.0}$		
	Ball-Cylinder-CSF	12.22 $^{0.6}$	0.62 $^{7.5}$	*1.14* $^{2.4}$		3.6 $^{2.6}$	0.37 $^{14.5}$	*1.14* $^{2.4}$		
	MMWMD	17.84 $^{0.0}$	0.64 $^{0.0}$	*1.70* $^{0.0}$		3.8 $^{0.0}$	0.34 $^{0.0}$	*1.70* $^{0.0}$	*0.58* $^{0.0}$	
	Tensor-Cylinder-CSF	9.59 $^{0.0}$	0.38 $^{0.0}$	*1.21* $^{0.0}$		2.4 $^{0.0}$	0.61 $^{0.0}$	*1.21* $^{0.0}$	*0.52* $^{0.0}$	*0.36* $^{0.0}$
	Tensor-Cylinder-CSF-diff	7.35 $^{0.0}$	0.25 $^{1.3}$	*0.40* $^{1.5}$		0.1 $^{29.6}$	0.73 $^{0.5}$	*1.81* $^{0.3}$	*0.42* $^{0.7}$	*0.29* $^{0.7}$
	NODDI+	7.34 $^{0.0}$	0.57 $^{0.1}$	*1.56* $^{0.1}$	5 $^{0.3}$		0.42 $^{0.1}$	*1.56* $^{0.1}$	*0.45* $^{0.4}$	
	MMWMD+	9.95 $^{0.0}$	0.39 $^{0.0}$	*1.21* $^{0.0}$		2.5 $^{0.0}$	0.60 $^{0.0}$	*1.21* $^{0.0}$	*0.45* $^{0.0}$	
CON-splenium	Ball-Bingham-CSF-diff	7.20 $^{0.0}$	0.75 $^{0.0}$	*2.25* $^{0.3}$	25 $^{0.5}$		0.25 $^{0.1}$	*0.57* $^{0.4}$		
	Ball-Bingham-CSF	8.87 $^{0.0}$	1.00 $^{0.0}$	*2.36* $^{0.1}$	8 $^{0.7}$		0.00 $^{82.4}$	*2.36* $^{0.1}$		
	NODDI	8.77 $^{0.0}$	0.76 $^{0.0}$	*1.90* $^{0.0}$	7 $^{0.0}$		0.24 $^{0.0}$	*1.90* $^{0.0}$	*0.45* $^{0.0}$	
	Tensor-Bingham-CSF	6.83 $^{0.0}$	0.56 $^{0.2}$	*1.65* $^{0.1}$	15 $^{0.4}$		0.44 $^{0.2}$	*1.65* $^{0.1}$	*0.49* $^{0.2}$	*0.26* $^{0.1}$
	Tensor-Bingham-CSF-diff	6.78 $^{0.0}$	0.51 $^{0.5}$	*1.37* $^{0.5}$	23 $^{0.9}$		0.49 $^{0.5}$	*1.93* $^{0.2}$	*0.46* $^{0.2}$	*0.26* $^{0.5}$
	Ball-Stick-Stick-CSF-diff	8.19 $^{0.0}$	0.69 $^{0.0}$	*1.83* $^{0.0}$			0.20 $^{0.0}$	*0.24* $^{0.0}$		
	Ball-Stick-Stick-CSF	10.26 $^{0.0}$	0.82 $^{0.0}$	*1.45* $^{0.0}$			0.06 $^{0.0}$	*1.45* $^{0.0}$		
	Tensor-Stick-Stick-CSF	6.75 $^{0.0}$	0.41 $^{0.2}$	*1.46* $^{0.1}$			0.57 $^{0.1}$	*1.46* $^{0.1}$	*0.47* $^{0.3}$	*0.23* $^{0.3}$
	Tensor-Stick-Stick-CSF-diff	6.25 $^{0.1}$	0.32 $^{1.9}$	*0.73* $^{2.8}$			0.66 $^{0.9}$	*2.12* $^{0.5}$	*0.40* $^{2.6}$	*0.18* $^{1.5}$
	Ball-Stick-CSF-diff	9.14 $^{0.0}$	0.59 $^{0.0}$	*2.07* $^{0.0}$			0.41 $^{0.0}$	*0.59* $^{0.0}$		
	Ball-Stick-CSF	11.71 $^{0.0}$	0.64 $^{0.0}$	*1.26* $^{0.0}$			0.36 $^{0.0}$	*1.26* $^{0.0}$		
	Tensor-Stick-CSF	7.32 $^{0.0}$	0.41 $^{0.0}$	*1.45* $^{0.0}$			0.59 $^{0.0}$	*1.45* $^{0.0}$	*0.48* $^{0.0}$	*0.24* $^{0.0}$
	Tensor-Stick-CSF-diff	6.92 $^{0.0}$	0.33 $^{0.0}$	*0.78* $^{0.0}$			0.67 $^{0.0}$	*2.04* $^{0.0}$	*0.41* $^{0.0}$	*0.20* $^{0.0}$
	Ball-Cylinder-CSF-diff	7.56 $^{0.0}$	0.78 $^{0.0}$	*1.61* $^{0.0}$		3.9 $^{0.0}$	0.22 $^{0.0}$	*0.76* $^{0.0}$		
	Ball-Cylinder-CSF	7.91 $^{0.0}$	0.85 $^{0.0}$	*1.40* $^{0.0}$		4.1 $^{0.0}$	0.15 $^{0.0}$	*1.40* $^{0.0}$		
	MMWMD	9.02 $^{0.0}$	0.74 $^{0.0}$	*1.70* $^{0.0}$		3.9 $^{0.0}$	0.26 $^{0.0}$	*1.70* $^{0.0}$	*0.43* $^{0.0}$	
	Tensor-Cylinder-CSF	6.59 $^{1.7}$	0.50 $^{16.6}$	*1.42* $^{1.2}$		7.7 $^{30.7}$	0.50 $^{16.5}$	*1.42* $^{1.2}$	*0.23* $^{122.8}$	*0.06* $^{169.8}$
	Tensor-Cylinder-CSF-diff	6.53 $^{0.5}$	0.53 $^{10.0}$	*1.21* $^{22.7}$		4.7 $^{51.6}$	0.47 $^{11.2}$	*2.03* $^{21.7}$	*0.46* $^{48.7}$	*0.16* $^{56.1}$
	NODDI+	8.18 $^{0.0}$	0.54 $^{0.0}$	*1.57* $^{0.0}$	12 $^{0.0}$		0.46 $^{0.0}$	*1.57* $^{0.0}$	*0.36* $^{0.0}$	
	MMWMD+	7.69 $^{2.6}$	0.59 $^{16.9}$	*1.41* $^{1.1}$		6.1 $^{40.1}$	0.41 $^{24.0}$	*1.41* $^{1.1}$	*0.26* $^{93.2}$	

Table E.1: Similar to Tab.20.2, but using datasets CON-midbody and CON-splenium.

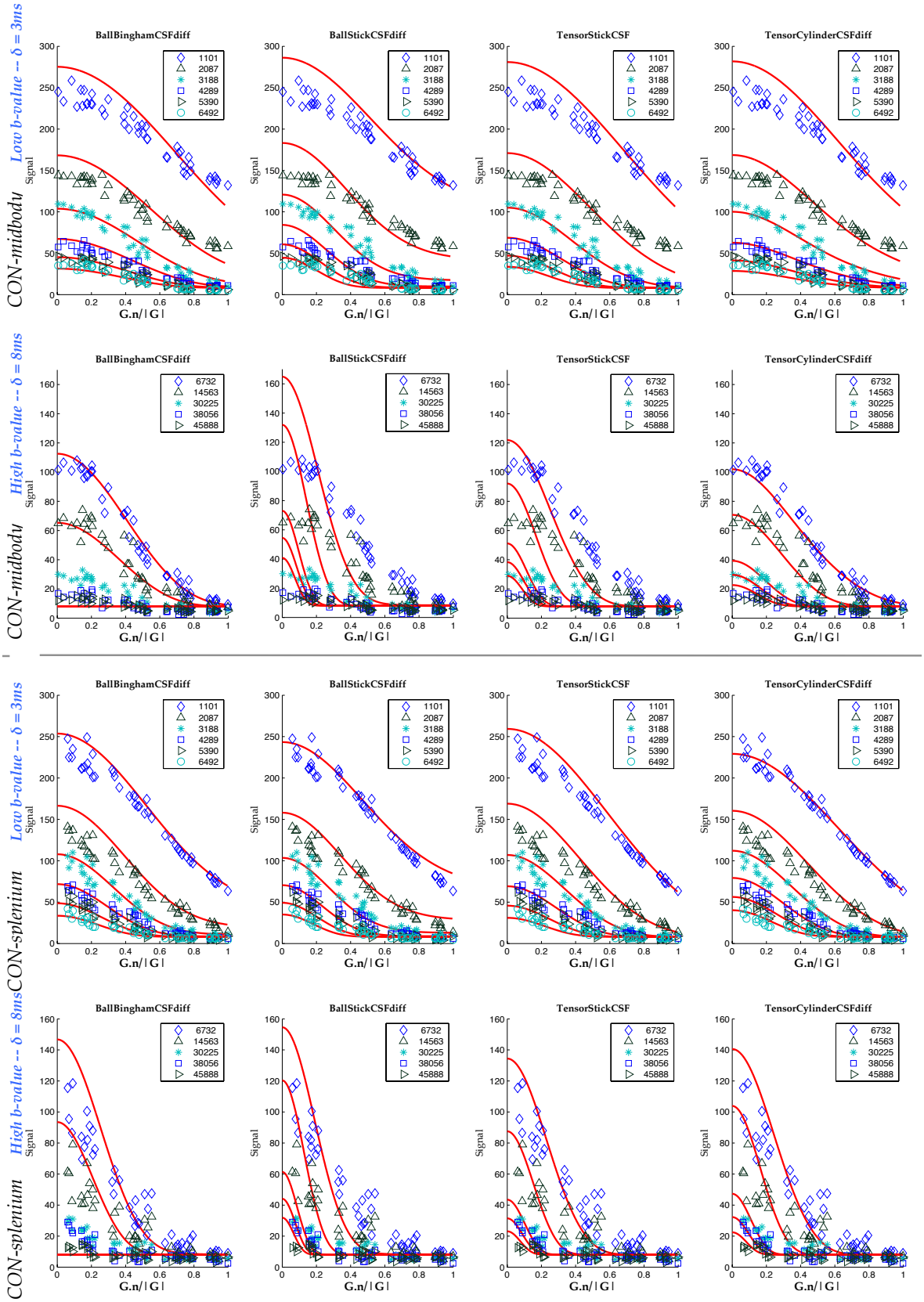


Figure E.1: As in Fig.20.5, on top are the models fitted to CON-midbody and, below, to CON-splenium.

	Models	LSE (x 10^3)	Cylinder/Bingham/Stick Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_i	Radius (μm)	*Tensor/ Ball* Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)	Radial Diff.2 ($\mu m^2/s$)
CON-genu-60	Ball-Bingham-CSF-diff	1.19 $^{0.0}$	0.89 $^{0.0}$	*2.37* $^{0.0}$	13 $^{0.0}$		0.10 $^{0.0}$	*1.34* $^{0.0}$		
	Ball-Bingham-CSF	1.19 $^{0.0}$	0.92 $^{0.0}$	*2.27* $^{0.0}$	12 $^{0.3}$		0.08 $^{0.3}$	*2.27* $^{0.0}$		
	NODDI	1.27 $^{0.0}$	0.64 $^{0.0}$	*1.90* $^{0.0}$	23 $^{0.0}$		0.36 $^{0.0}$	*1.90* $^{0.0}$	*0.68* $^{0.0}$	
	Tensor-Bingham-CSF	1.19 $^{0.0}$	0.87 $^{0.0}$	*2.18* $^{0.0}$	14 $^{0.0}$		0.13 $^{0.0}$	*2.18* $^{0.0}$	*1.67* $^{0.0}$	*1.18* $^{0.0}$
	Tensor-Bingham-CSF-diff	1.19 $^{0.0}$	0.78 $^{3.0}$	*1.82* $^{4.2}$	27 $^{28.5}$		0.22 $^{10.7}$	*3.00* $^{0.0}$	*1.35* $^{4.6}$	*0.85* $^{7.0}$
	Ball-Stick-Stick-CSF-diff	1.28 $^{0.0}$	0.77 $^{0.7}$	*1.90* $^{1.0}$			0.22 $^{2.6}$	*1.51* $^{3.2}$		
	Ball-Stick-Stick-CSF	1.29 $^{0.0}$	0.79 $^{0.0}$	*1.82* $^{0.0}$			0.20 $^{0.0}$	*1.82* $^{0.0}$		
	Tensor-Stick-Stick-CSF	1.22 $^{0.0}$	0.70 $^{0.3}$	*1.83* $^{0.0}$			0.30 $^{0.7}$	*1.83* $^{0.0}$	*1.32* $^{1.3}$	*0.72* $^{0.7}$
	Tensor-Stick-Stick-CSF-diff	1.19 $^{0.0}$	0.64 $^{1.8}$	*1.40* $^{1.4}$			0.35 $^{3.2}$	*3.00* $^{0.1}$	*1.06* $^{5.0}$	*0.57* $^{3.9}$
	Ball-Stick-CSF-diff	1.29 $^{0.0}$	0.77 $^{0.0}$	*1.87* $^{0.0}$			0.22 $^{0.0}$	*1.57* $^{0.0}$		
	Ball-Stick-CSF	1.29 $^{0.0}$	0.79 $^{0.0}$	*1.80* $^{0.0}$			0.21 $^{0.0}$	*1.80* $^{0.0}$		
	Tensor-Stick-CSF	1.23 $^{0.0}$	0.71 $^{0.0}$	*1.83* $^{0.0}$			0.29 $^{0.0}$	*1.83* $^{0.0}$	*1.40* $^{0.0}$	*0.76* $^{0.1}$
	Tensor-Stick-CSF-diff	1.19 $^{0.0}$	0.64 $^{0.8}$	*1.40* $^{0.7}$			0.35 $^{1.5}$	*3.00* $^{0.0}$	*1.09* $^{2.0}$	*0.58* $^{2.2}$
	Ball-Cylinder-CSF-diff	1.29 $^{0.1}$	0.78 $^{0.1}$	*1.83* $^{0.3}$		0.2 $^{58.0}$	0.22 $^{0.5}$	*1.68* $^{1.7}$		
	Ball-Cylinder-CSF	1.29 $^{0.0}$	0.79 $^{0.1}$	*1.80* $^{0.1}$		0.1 $^{32.3}$	0.21 $^{0.5}$	*1.80* $^{0.1}$		
	MMWMD	1.34 $^{0.0}$	0.61 $^{0.0}$	*1.70* $^{0.0}$		4.1 $^{0.0}$	0.38 $^{0.0}$	*1.70* $^{0.0}$	*0.65* $^{0.0}$	
	Tensor-Cylinder-CSF	1.23 $^{0.0}$	0.71 $^{0.2}$	*1.83* $^{0.1}$		0.3 $^{83.7}$	0.29 $^{0.6}$	*1.83* $^{0.1}$	*1.40* $^{0.9}$	*0.76* $^{0.5}$
	Tensor-Cylinder-CSF-diff	1.19 $^{0.0}$	0.64 $^{3.0}$	*1.39* $^{2.3}$		0.1 $^{20.9}$	0.35 $^{5.5}$	*2.99* $^{0.2}$	*1.09* $^{8.9}$	*0.59* $^{7.2}$
	NODDI+	1.24 $^{0.0}$	0.87 $^{0.0}$	*2.18* $^{0.0}$	8 $^{0.0}$		0.13 $^{0.0}$	*2.18* $^{0.0}$	*1.41* $^{0.0}$	
	MMWMD+	1.27 $^{0.0}$	0.72 $^{0.3}$	*1.83* $^{0.1}$		0.2 $^{65.2}$	0.28 $^{0.8}$	*1.83* $^{0.1}$	*1.13* $^{1.1}$	
CON-genu-100	Ball-Bingham-CSF-diff	0.95 $^{0.0}$	0.83 $^{0.0}$	*2.45* $^{0.0}$	12 $^{0.0}$		0.16 $^{0.0}$	*1.57* $^{0.0}$		
	Ball-Bingham-CSF	0.98 $^{0.0}$	0.85 $^{0.0}$	*2.30* $^{0.0}$	12 $^{0.0}$		0.14 $^{0.0}$	*2.30* $^{0.0}$		
	NODDI	1.18 $^{0.0}$	0.59 $^{0.0}$	*1.90* $^{0.0}$	28 $^{0.0}$		0.41 $^{0.0}$	*1.90* $^{0.0}$	*0.77* $^{0.0}$	
	Tensor-Bingham-CSF	0.95 $^{0.0}$	0.80 $^{0.0}$	*2.24* $^{0.0}$	14 $^{0.0}$		0.20 $^{0.0}$	*2.24* $^{0.0}$	*1.56* $^{0.0}$	*1.36* $^{0.0}$
	Tensor-Bingham-CSF-diff	0.95 $^{0.1}$	0.81 $^{4.9}$	*2.29* $^{8.6}$	14 $^{20.8}$		0.19 $^{20.5}$	*2.00* $^{26.4}$	*1.55* $^{7.8}$	*1.40* $^{12.9}$
	Ball-Stick-Stick-CSF-diff	1.29 $^{0.1}$	0.69 $^{0.1}$	*1.89* $^{0.2}$			0.30 $^{0.2}$	*1.53* $^{0.7}$		
	Ball-Stick-Stick-CSF	1.30 $^{0.0}$	0.71 $^{0.1}$	*1.79* $^{0.1}$			0.29 $^{0.2}$	*1.79* $^{0.1}$		
	Tensor-Stick-Stick-CSF	1.14 $^{0.0}$	0.62 $^{0.4}$	*1.85* $^{0.1}$			0.38 $^{0.6}$	*1.85* $^{0.1}$	*1.19* $^{1.5}$	*0.80* $^{1.2}$
	Tensor-Stick-Stick-CSF-diff	0.98 $^{0.1}$	0.56 $^{0.9}$	*1.36* $^{0.5}$			0.44 $^{1.1}$	*3.00* $^{0.0}$	*0.95* $^{1.4}$	*0.64* $^{1.9}$
	Ball-Stick-CSF-diff	1.30 $^{0.0}$	0.70 $^{0.0}$	*1.87* $^{0.0}$			0.30 $^{0.0}$	*1.59* $^{0.0}$		
	Ball-Stick-CSF	1.31 $^{0.0}$	0.71 $^{0.0}$	*1.79* $^{0.0}$			0.29 $^{0.0}$	*1.79* $^{0.0}$		
	Tensor-Stick-CSF	1.15 $^{0.0}$	0.63 $^{0.0}$	*1.85* $^{0.0}$			0.37 $^{0.0}$	*1.85* $^{0.0}$	*1.24* $^{0.1}$	*0.83* $^{0.1}$
	Tensor-Stick-CSF-diff	0.98 $^{0.0}$	0.57 $^{1.0}$	*1.37* $^{0.9}$			0.43 $^{1.3}$	*3.00* $^{0.0}$	*0.98* $^{2.5}$	*0.65* $^{2.1}$
	Ball-Cylinder-CSF-diff	1.25 $^{0.0}$	0.78 $^{0.0}$	*1.87* $^{0.0}$		4.3 $^{0.0}$	0.22 $^{0.0}$	*1.59* $^{0.0}$		
	Ball-Cylinder-CSF	1.25 $^{0.0}$	0.79 $^{0.0}$	*1.82* $^{0.0}$		4.3 $^{0.0}$	0.21 $^{0.0}$	*1.82* $^{0.0}$		
	MMWMD	1.36 $^{0.0}$	0.59 $^{0.0}$	*1.70* $^{0.0}$		4.1 $^{0.0}$	0.40 $^{0.0}$	*1.70* $^{0.0}$	*0.69* $^{0.0}$	
	Tensor-Cylinder-CSF	1.11 $^{0.0}$	0.69 $^{0.0}$	*1.87* $^{0.0}$		4.2 $^{0.0}$	0.31 $^{0.1}$	*1.87* $^{0.0}$	*1.13* $^{0.1}$	*0.68* $^{0.1}$
	Tensor-Cylinder-CSF-diff	0.92 $^{0.8}$	0.56 $^{2.5}$	*1.34* $^{2.5}$		3.7 $^{4.3}$	0.44 $^{3.3}$	*3.00* $^{0.0}$	*0.80* $^{4.7}$	*0.52* $^{3.2}$
	NODDI+	1.02 $^{0.0}$	0.80 $^{0.0}$	*2.23* $^{0.0}$	9 $^{0.0}$		0.20 $^{0.0}$	*2.23* $^{0.0}$	*1.45* $^{0.0}$	
	MMWMD+	1.17 $^{0.0}$	0.69 $^{0.0}$	*1.86* $^{0.0}$		4.1 $^{0.0}$	0.31 $^{0.0}$	*1.86* $^{0.0}$	*0.95* $^{0.0}$	

Table E.2: Variation of parameter estimates for the two lowest gradient strengths, 60 and 100 mT/m. The T_2 is fixed at 56ms, the mean from CON-genu data.

	Models	LSE (x 10^3)		Cylinder / Bingham / Stick Vol.Fraction		Diffusivity ($\mu m^2/s$)		κ_1		Radius (μm)		Tensor/ Ball Vol.Fraction		Axial Diff. ($\mu m^2/s$)		Radial Diff.1 ($\mu m^2/s$)		Radial Diff.2 ($\mu m^2/s$)	
CON-genu-200	Ball-Bingham-CSF-diff	0.96	1.2	0.81	0.9	*2.73*	*8.5*	15	11.7			0.18	1.6	*0.87*	*20.3*				
	Ball-Bingham-CSF	1.22	11.7	0.87	2.1	*2.41*	*6.3*	13	18.8			0.13	17.2	*2.41*	*6.3*				
	NODDI	1.20	9.1	0.64	2.9	*1.90*	*0.0*	21	9.8			0.36	4.0	*1.90*	*0.0*	*0.68*	*4.0*		
	Tensor-Bingham-CSF	0.96	3.4	0.73	4.3	*2.23*	*3.2*	17	8.3			0.27	10.2	*2.23*	*3.2*	*0.77*	*18.6*	*0.65*	*26.3*
	Tensor-Bingham-CSF-diff	0.94	2.8	0.74	6.3	*2.41*	*22.8*	19	31.8			0.26	19.8	*1.92*	*56.0*	*0.71*	*2.7*	*0.60*	*12.2*
	Ball-Stick-Stick-CSF-diff	1.96	12.7	0.66	4.6	*2.07*	*4.9*					0.34	8.2	*1.09*	*8.2*				
	Ball-Stick-Stick-CSF	2.20	16.8	0.68	3.7	*1.71*	*1.1*					0.31	7.2	*1.71*	*1.1*				
	Tensor-Stick-Stick-CSF	1.41	7.3	0.54	10.7	*1.88*	*1.1*					0.45	11.5	*1.88*	*1.1*	*0.62*	*7.2*	*0.46*	*20.2*
	Tensor-Stick-Stick-CSF-diff	1.03	2.0	0.45	17.6	*1.19*	*14.2*					0.54	13.8	*2.99*	*0.2*	*0.48*	*13.1*	*0.36*	*24.9*
	Ball-Stick-CSF-diff	1.97	13.0	0.66	4.3	*2.06*	*4.7*					0.34	7.6	*1.09*	*8.3*				
	Ball-Stick-CSF	2.22	17.0	0.68	3.5	*1.70*	*1.1*					0.32	6.8	*1.70*	*1.1*				
	Tensor-Stick-CSF	1.43	7.9	0.54	10.5	*1.87*	*0.8*					0.46	11.6	*1.87*	*0.8*	*0.63*	*8.4*	*0.47*	*20.0*
	Tensor-Stick-CSF-diff	1.03	2.1	0.46	16.8	*1.18*	*14.6*					0.54	13.6	*3.00*	*0.0*	*0.47*	*10.6*	*0.37*	*24.8*
	Ball-Cylinder-CSF-diff	1.48	17.0	0.85	1.6	*1.79*	*1.2*			4.1	3.2	0.15	6.5	*2.20*	*13.0*				
	Ball-Cylinder-CSF	1.49	16.4	0.84	0.8	*1.82*	*0.0*			4.1	4.0	0.16	1.7	*1.82*	*0.0*				
	MMWMD	1.56	10.2	0.62	1.4	*1.70*	*0.0*			3.0	8.8	0.38	1.3	*1.70*	*0.0*	*0.65*	*1.3*		
	Tensor-Cylinder-CSF	1.36	10.2	0.69	14.6	*1.86*	*1.3*			3.5	5.3	0.31	31.6	*1.86*	*1.3*	*0.98*	*33.3*	*0.66*	*41.6*
	Tensor-Cylinder-CSF-diff	1.02	2.9	0.53	11.0	*1.30*	*8.3*			3.6	19.1	0.47	11.6	*3.00*	*0.0*	*0.47*	*12.8*	*0.36*	*23.7*
	NODDI+	1.03	7.4	0.72	4.2	*2.21*	*2.8*	13	20.1			0.28	9.4	*2.21*	*2.8*	*0.72*	*19.1*		
	MMWMD+	1.40	12.2	0.70	13.4	*1.86*	*1.1*			3.5	5.0	0.30	29.9	*1.86*	*1.1*	*0.83*	*34.2*		
CON-genu-300	Ball-Bingham-CSF-diff	0.97	2.2	0.76	1.0	*2.84*	*5.5*	16	1.0			0.23	2.5	*0.82*	*14.1*				
	Ball-Bingham-CSF	1.29	8.3	0.80	2.6	*2.31*	*2.9*	15	2.9			0.19	13.3	*2.31*	*2.9*				
	NODDI	1.17	5.5	0.61	0.1	*1.90*	*0.0*	18	3.4			0.39	0.2	*1.90*	*0.0*	*0.74*	*0.2*		
	Tensor-Bingham-CSF	0.95	0.0	0.61	0.0	*2.15*	*0.0*	28	0.0			0.39	0.0	*2.15*	*0.0*	*0.56*	*0.0*	*0.49*	*0.0*
	Tensor-Bingham-CSF-diff	0.92	1.4	0.55	4.3	*1.64*	*2.8*	27	43.1			0.44	5.7	*3.00*	*0.0*	*0.62*	*20.8*	*0.45*	*8.7*
	Ball-Stick-Stick-CSF-diff	2.32	12.9	0.60	0.1	*1.87*	*3.5*					0.39	0.6	*0.91*	*15.0*				
	Ball-Stick-Stick-CSF	2.64	16.2	0.62	0.4	*1.46*	*2.8*					0.38	1.3	*1.46*	*2.8*				
	Tensor-Stick-Stick-CSF	1.46	1.9	0.46	8.7	*1.75*	*3.3*					0.53	7.5	*1.75*	*3.3*	*0.55*	*7.3*	*0.40*	*15.3*
	Tensor-Stick-Stick-CSF-diff	1.06	0.3	0.35	5.5	*1.00*	*2.4*					0.64	3.5	*2.99*	*0.4*	*0.42*	*11.1*	*0.27*	*11.1*
	Ball-Stick-CSF-diff	2.35	13.8	0.59	2.4	*1.95*	*0.0*					0.41	3.0	*0.97*	*7.3*				
	Ball-Stick-CSF	2.67	16.7	0.62	0.4	*1.47*	*2.8*					0.38	0.6	*1.47*	*2.8*				
	Tensor-Stick-CSF	1.48	1.7	0.47	8.8	*1.75*	*3.3*					0.53	7.4	*1.75*	*3.3*	*0.56*	*9.5*	*0.41*	*16.6*
	Tensor-Stick-CSF-diff	1.06	0.4	0.37	9.5	*1.01*	*4.9*					0.62	6.0	*3.00*	*0.0*	*0.46*	*17.6*	*0.30*	*18.0*
	Ball-Cylinder-CSF-diff	1.53	12.0	0.76	0.9	*1.69*	*4.9*			3.5	3.6	0.24	1.5	*1.51*	*12.2*				
	Ball-Cylinder-CSF	1.54	12.6	0.77	0.4	*1.66*	*2.2*			3.5	3.9	0.23	2.1	*1.66*	*2.2*				
	MMWMD	1.42	6.0	0.62	0.9	*1.70*	*0.0*			3.0	4.7	0.38	0.9	*1.70*	*0.0*	*0.65*	*0.9*		
	Tensor-Cylinder-CSF	1.33	6.2	0.55	23.1	*1.75*	*2.8*			4.9	62.2	0.45	28.3	*1.75*	*2.8*	*0.62*	*57.3*	*0.41*	*63.1*
	Tensor-Cylinder-CSF-diff	1.05	1.7	0.44	19.3	*1.11*	*11.3*			2.1	23.7	0.56	15.3	*3.00*	*0.0*	*0.53*	*29.0*	*0.34*	*30.6*
	NODDI+	1.09	1.2	0.62	4.5	*2.08*	*1.8*	14	3.5			0.37	8.0	*2.08*	*1.8*	*0.68*	*21.5*		
	MMWMD+	1.35	6.0	0.51	27.5	*1.72*	*2.8*			6.6	52.2	0.48	29.0	*1.72*	*2.8*	*0.39*	*96.6*		

Table E.3: Variation of parameter estimates as gradient strength increases to 200 and 300 mT/m.

F

TO PART VII (STABILITY OF PARAMETERS)

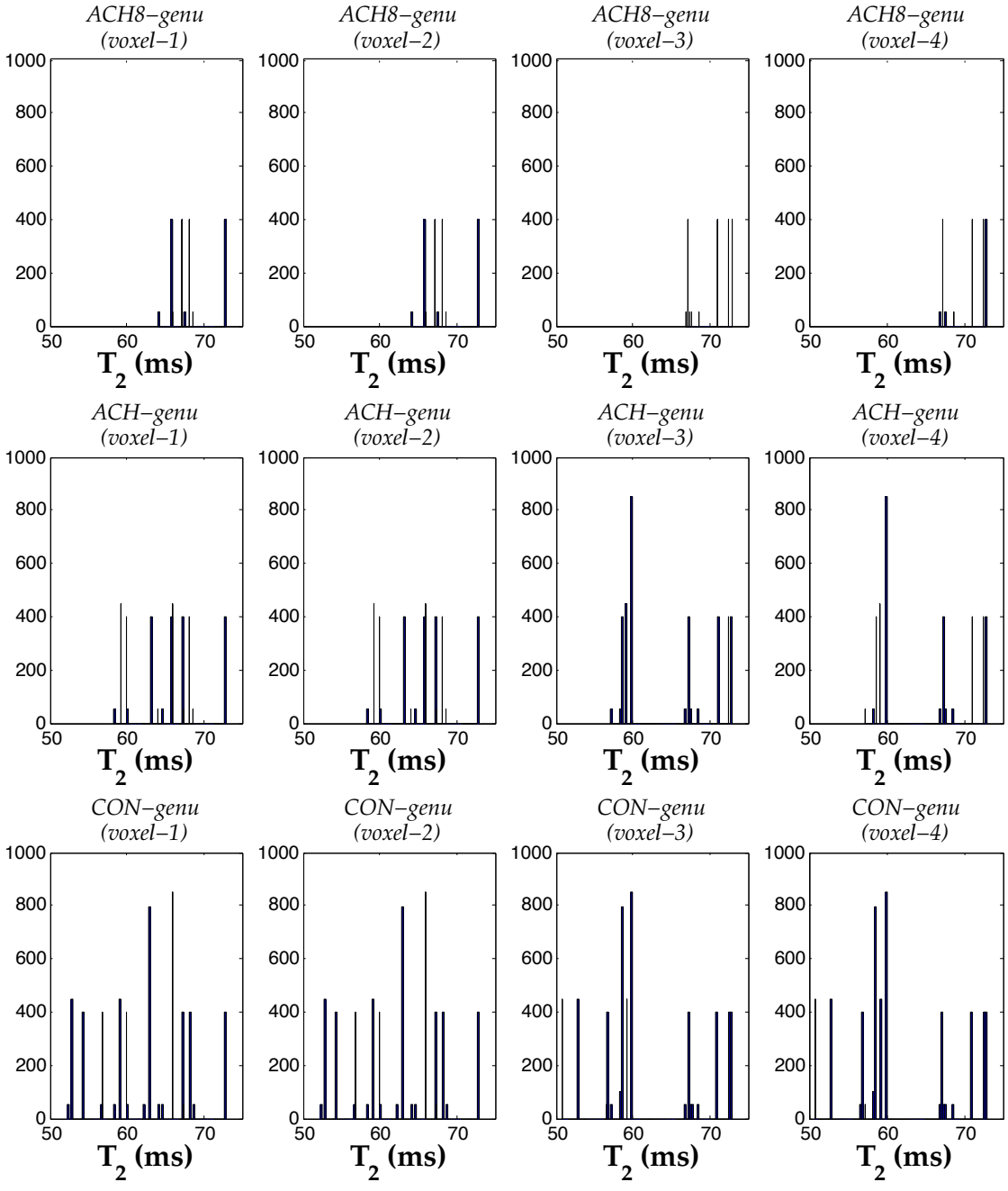


Figure F.1: Distributions of T_2 estimates, from all the models, for each dataset, ACH8-genu, ACH-genu and CON-genu, and each of the four voxels of the ROI.

			Cylinder/Bingham				*Zeppelin*		
		Models	LSE (x 10^3)	Vol.Fraction	Diffusivity ($\mu m^2/s$)	κ_1	Radius (μm)	Vol.Fraction	Axial Diff. ($\mu m^2/s$)	Radial Diff.1 ($\mu m^2/s$)
ACH-genu	T_2^{mean} - 2ms	NODDI+	1.16	0.70	*2.06*	8		0.29	*2.06*	*0.57*
		MMWMD+	1.30	0.65	*1.66*		5.3	0.35	*1.66*	*0.50*
	T_2^{mean} = 59ms	NODDI+	1.12	0.68	*2.04*	8		0.32	*2.04*	*0.58*
		MMWMD+	1.28	0.58	*1.66*		4.8	0.42	*1.66*	*0.51*
	T_2^{mean} + 2ms	NODDI+	1.09	0.64	*1.98*	9		0.36	*1.98*	*0.53*
		MMWMD+	1.25	0.51	*1.65*		4.1	0.49	*1.65*	*0.48*
CON-genu	T_2^{mean} - 2ms	NODDI+	7.12	0.67	*2.10*	12		0.32	*2.10*	*0.69*
		MMWMD+	7.90	0.67	*1.77*		3.4	0.32	*1.77*	*0.87*
	T_2^{mean} = 57ms	NODDI+	7.09	0.66	*2.12*	11		0.33	*2.12*	*0.78*
		MMWMD+	8.16	0.63	*1.79*		3.3	0.36	*1.79*	*0.86*
	T_2^{mean} + 2ms	NODDI+	7.52	0.64	*2.08*	12		0.36	*2.08*	*0.74*
		MMWMD+	8.71	0.57	*1.78*		2.9	0.43	*1.78*	*0.72*

Table F.1: Dependence on T_2 for two models' parameter estimates. We fit NODDI+ and MMWMD+ to the full dataset of ACH-genu, on top, and CON-genu, fixing the T_2 to the mean obtained in the ROI, and 2ms either side of this mean. The estimates for T_2^{mean} are similar to those of Fig.16.1 and Fig.20.2, with slight variation as here we fix the T_2 throughout the ROI voxels, and select the best model fit out of 20 runs (vs. the previous 100).

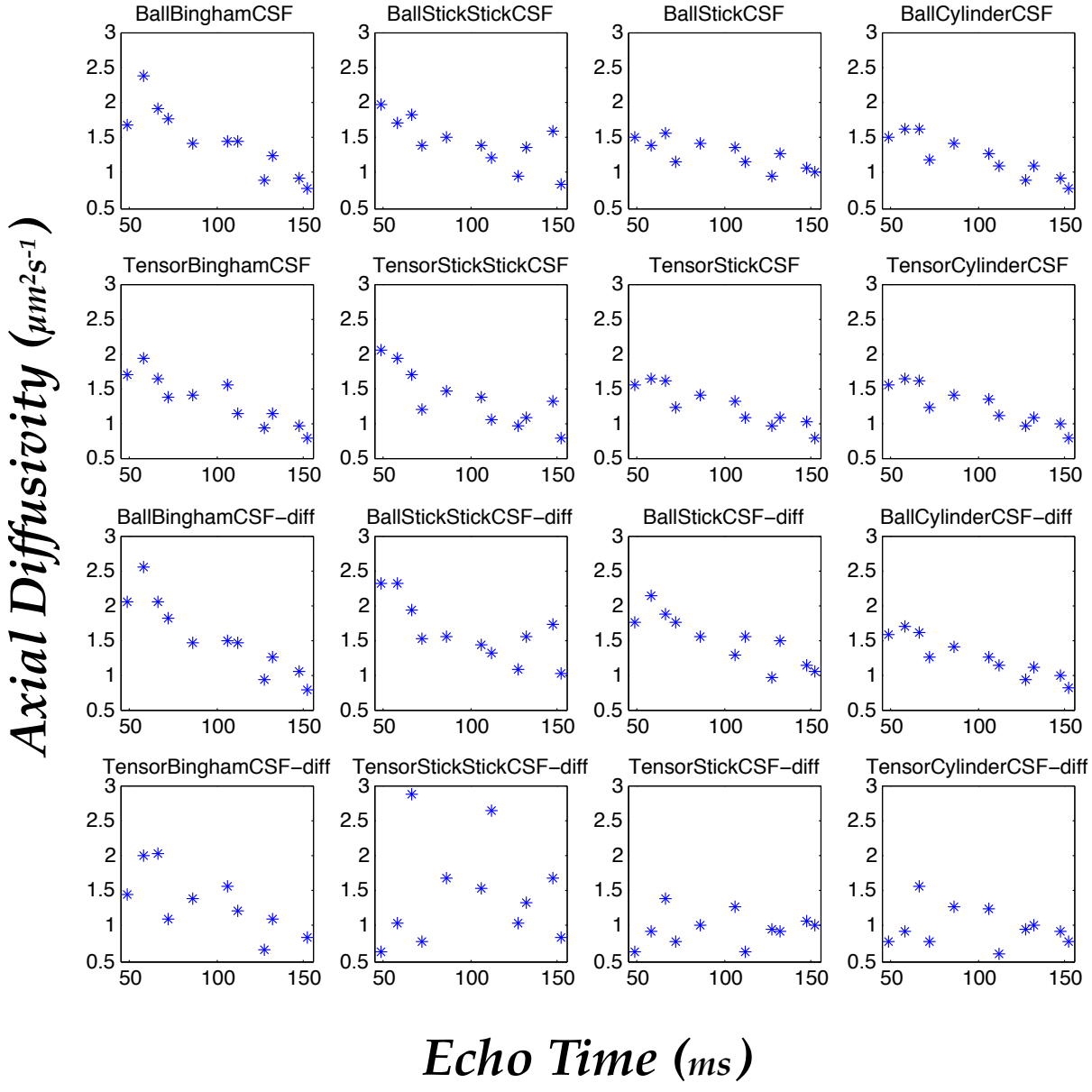


Figure F.2: The estimation of axial diffusivity across TE-specific datasets.

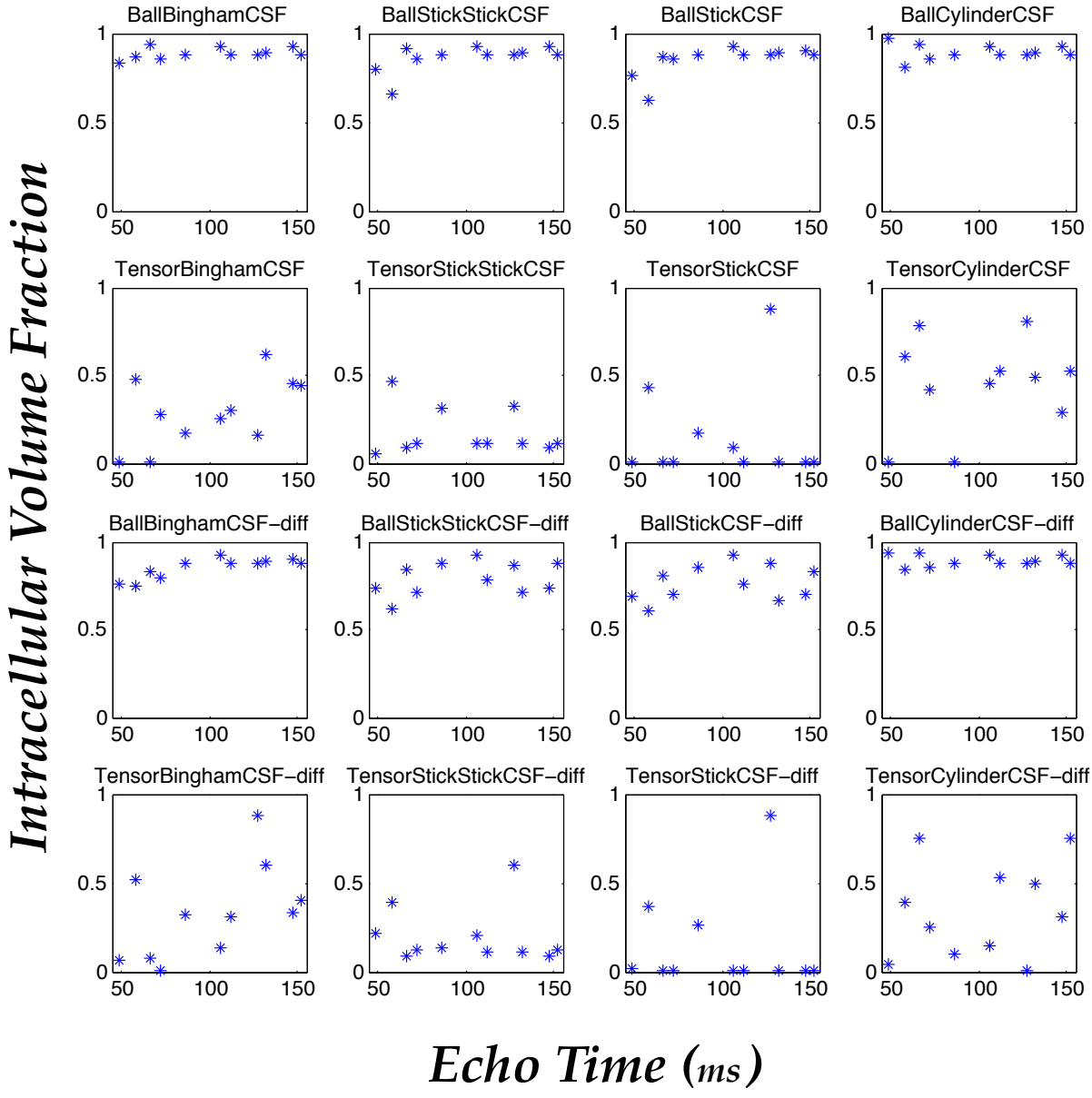


Figure F.3: The estimation of intracellular volume fraction across TE-specific datasets.

BIBLIOGRAPHY

Francisco Aboitiz, Arnold B Scheibel, Robin S Fisher, and Eran Zaidel. Fiber composition of the human corpus callosum. *Brain research*, 598(1):143–153, 1992.

Hirotugu Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

Daniel C Alexander. An introduction to computational diffusion MRI: the diffusion tensor and beyond. In *Visualization and processing of tensor fields*, pages 83–106. Springer, 2006.

Daniel C Alexander. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine*, 60(2):439–448, 2008.

Daniel C Alexander. Modelling, fitting and sampling in diffusion MRI. In *Visualization and processing of tensor fields*, pages 3–20. Springer, 2009.

Daniel C Alexander, Penny L Hubbard, Matt G Hall, Elizabeth A Moore, Maurice Ptito, Geoff JM Parker, and Tim B Dyrby. Orientationally invariant indices of axon diameter and density from diffusion MRI. *Neuroimage*, 52(4):1374–1389, 2010.

A Anderson and Z Ding. Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging. In *Book of abstracts: Tenth Annual Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, CA: ISMRM*, volume 10, page 440, 2002.

Yaniv Assaf and Peter J. Basser. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27(1):48 – 58, 2005.

Yaniv Assaf, Michal Kafri, Hadassah Shinar, Joab Chapman, Amos D Korczyn, Gil Navon, and Yoram Cohen. Changes in axonal morphology in experimental autoimmune neuritis as studied by high b-value q-space 1H and 2H DQF diffusion magnetic resonance spectroscopy. *Magnetic Resonance in Medicine*, 48(1):71–81, 2002.

Yaniv Assaf, Raisa Z Freidlin, Gustavo K Rohde, and Peter J Basser. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52(5):965–978, 2004.

Yaniv Assaf, Tamar Blumenfeld-Katzir, Yossi Yovel, and Peter J Basser. Axcaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, 59(6):1347–1354, 2008.

Daniel Barazany, Peter J Basser, and Yaniv Assaf. In vivo measurement of axon diameter distribution in the corpus callosum of rat brain. *Brain*, 132(5):1210–1220, 2009.

Peter J Basser, James Mattiello, and Denis LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.

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

TEJ Behrens, MW Woolrich, M Jenkinson, H Johansen-Berg, RG Nunes, S Clare, PM Matthews, JM Brady, and SM Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic resonance in medicine*, 50(5):1077–1088, 2003.

TEJ Behrens, H Johansen Berg, Saad Jbabdi, MFS Rushworth, and MW Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage*, 34 (1):144–155, 2007.

Ludwig Boltzmann. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Satzen über das Wärmegleichgewicht. *Wiener Berichte*, 76:373–435, 1877.

Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

Robert Brown. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine, or Annals of Chemistry, Mathematics, Astronomy, Natural History and General Science*, 4(21):161–173, 1828.

Kenneth P Burnham and David R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2002.

Paul T Callaghan. *Principles of nuclear magnetic resonance microscopy*, volume 3. Clarendon Press Oxford, 1991.

Paul T Callaghan and Janez Stepišnik. Frequency-domain analysis of spin motion using modulated-gradient NMR. *Journal of Magnetic Resonance, Series A*, 117(1):118–122, 1995.

Paul T Callaghan, D MacGowan, KJ Packer, and FO Zelaya. High-resolution< i> q</i>-space imaging in porous structures. *Journal of Magnetic Resonance (1969)*, 90(1):177–182, 1990.

Herman Y Carr and Edward M Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 94(3):630, 1954.

Emmanuel Caruyer, Christophe Lenglet, Guillermo Sapiro, and Rachid Deriche. Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance in Medicine*, 69(6):1534–1540, 2013.

A Castriota-Scanderbeg, U Sabatini, F Fasano, R Floris, L Fraracci, Di M Mario, U Nocentini, and C Caltagirone. Diffusion of water in large demyelinating lesions: a follow-up study. *Neuroradiology*, 44(9):764–767, 2002.

Chris A Clark and Denis Le Bihan. Water diffusion compartmentation and anisotropy at high b values in the human brain. *Magnetic Resonance in Medicine*, 44(6):852–859, 2000.

Jonathan D Clayden, Zoltan Nagy, Matt G Hall, Chris A Clark, and Daniel C Alexander. Active imaging with dual spin-echo diffusion MRI. In *Information Processing in Medical Imaging*, pages 264–275. Springer, 2009.

PA Cook, Y Bai, SKKS Nedjati-Gilani, KK Seunarine, MG Hall, GJ Parker, and DC Alexander. Camino: open-source diffusion-MRI reconstruction and processing. page 2759, 2006.

Eleanor F Cox and Penny A Gowland. Simultaneous quantification of T2 and T2* using a combined gradient echo-spin echo sequence at ultrahigh field. *Magnetic Resonance in Medicine*, 64(5):1440–1445, 2010.

NA Diamantidis, D Karlis, and Emmanouel A Giakoumakis. Unsupervised stratification of cross-validation for accuracy estimation. *Artificial Intelligence*, 116(1):1–16, 2000.

Mark D Does, Edward C Parsons, and John C Gore. Oscillating gradient measurements of water diffusion in normal and globally ischemic rat brain. *Magnetic resonance in medicine*, 49 (2):206–215, 2003.

Ivana Drobnjak, Bernard Siow, and Daniel C Alexander. Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. *Journal of Magnetic Resonance*, 206(1): 41–51, 2010.

T Duval, J McNab, K Setsompop, T Witzel, T Schneider, H Susie Yi, B Keil, E Klawitter, LL Wald, and J Cohen-Adad. In vivo estimation of axon diameter in the human spinal cord using 300 mT/m gradients. 2014.

Tim B Dyrby, Matt G Hall, Maurice Ptito, Daniel Alexander, et al. Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magnetic Resonance in Medicine*, 70(3):711–721, 2013.

Bradley Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, pages 1–26, 1979.

Bradley Efron. Estimating the error rate of a prediction rule: improvement on cross-validation. *Journal of the American Statistical Association*, 78(382):316–331, 1983.

Bradley Efron and Gail Gong. A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37(1):36–48, 1983.

Bradley Efron and Robert Tibshirani. Using specially designed exponential families for density estimation. *The Annals of Statistics*, 24(6):2431–2461, 1996.

Bradley Efron and Robert Tibshirani. Improvements on cross-validation: the 632+ bootstrap method. *Journal of the American Statistical Association*, 92(438):548–560, 1997.

Uran Ferizi, Torben Schneider, Eleftheria Panagiotaki, Gemma Nedjati-Gilani, Hui Zhang, Claudia AM Wheeler-Kingshott, and Daniel C Alexander. A ranking of diffusion MRI compartment models with in vivo human brain data. *Magnetic Resonance in Medicine*, 2013a.

Uran Ferizi, Torben Schneider, Maira Tariq, Claudia AM Wheeler-Kingshott, Hui Zhang, and Daniel C Alexander. The importance of being dispersed: A ranking of diffusion MRI models for fibre dispersion using in vivo human brain data. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 74–81. Springer, 2013b.

Jean Baptiste Joseph Fourier. *The analytical theory of heat*. Cambridge University Press, 1878.

Herbert Goldstein. *Classical Mechanics*.

Erwin L Hahn. Spin echoes. *Physical Review*, 80(4):580, 1950.

Tim Hosey, Guy Williams, and Richard Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.

SY Huang, A Nummenmaa, T Witzel, T Duval, J Cohen-Adad, LL Wald, and J McNab. The impact of gradient strength on in vivo diffusion MRI estimates of axon diameter. 2014.

Penny L Hubbard, Feng-Lei Zhou, Stephen J Eichhorn, and Geoffrey JM Parker. Biomimetic phantom for the validation of diffusion magnetic resonance imaging. *Magnetic Resonance in Medicine*, 2014.

Harold Jeffreys. *The theory of probability*. Oxford University Press, 1961.

Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841, 2002.

Jens H Jensen, Joseph A Helpern, Anita Ramani, Hanzhang Lu, and Kyle Kaczynski. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53(6):1432–1440, 2005.

B Jeurissen, A Leemans, JD Tournier, DK Jones, and J Sijbers. Estimating the number of fiber orientations in diffusion MRI voxels: a constrained spherical deconvolution study. *Proceedings of the International Society for Magnetic Resonance in Medicine. Stockholm, Sweden*, page 573, 2010.

Derek K Jones and Peter J Basser. Squashing peanuts and smashing pumpkins: How noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine*, 52(5):979–993, 2004.

Enrico Kaden, Thomas R Knösche, and Alfred Anwander. Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion MR imaging. *NeuroImage*, 37(2):474–488, 2007.

Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, 1995.

ME Komlosh, F Horkay, RZ Freidlin, U Nevo, Y Assaf, and PJ Basser. Detection of microscopic anisotropy in gray matter and in a novel tissue phantom using double pulsed gradient spin echo MR. *Journal of magnetic resonance*, 189(1):38–45, 2007.

Solomon Kullback and Richard A Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, pages 79–86, 1951.

Denis Le Bihan, Eric Breton, Denis Lallemand, Philippe Grenier, Emmanuel Cabanis, and Maurice Laval-Jeantet. Mr imaging of intra-voxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407, 1986.

Zhi-Pei Liang and Paul C Lauterbur. *Principles of magnetic resonance imaging*. SPIE Optical Engineering Press, 2000.

Alex MacKay, Kenneth Whittall, Julian Adler, David Li, Donald Paty, and Douglas Graeb. In vivo visualization of myelin water in brain by magnetic resonance. *Magnetic Resonance in Medicine*, 31(6):673–677, 1994.

Alex MacKay, Cornelia Laule, Irene Vavasour, Thorarin Bjarnason, Shannon Kolind, and Burkhard Mädler. Insights into brain microstructure from the T2 distribution. *Magnetic resonance imaging*, 24(4):515–525, 2006.

Colin L Mallows. Some comments on Cp. *Technometrics*, 15(4):661–675, 1973.

P Mansfield and IL Pykett. Biological and medical imaging by NMR. *Journal of Magnetic Resonance (1969)*, 29(2):355–373, 1978.

Kanti V Mardia and Peter E Jupp. Distributions on spheres. *Directional Statistics*, pages 159–192, 2000.

Jennifer A McNab, Brian L Edlow, Thomas Witzel, Susie Y Huang, Himanshu Bhat, Keith Heberlein, Thorsten Feiweier, Kecheng Liu, Boris Keil, Julien Cohen-Adad, et al. The human Connectome project and beyond: Initial applications of 300mT/m gradients. *NeuroImage*, 80: 234–245, 2013.

RS Menon, MS Rusinko, and PS Allen. Proton relaxation studies of water compartmentalization in a model neurological system. *Magnetic resonance in medicine*, 28(2):264–274, 1992.

Partha P Mitra. Multiple wave-vector extensions of the NMR pulsed-field-gradient spin-echo diffusion measurement. *Physical Review B*, 51(21):15074, 1995.

ME Moseley, Y Cohen, J Mintorovitch, L Chileuitt, H Shimizu, J Kucharczyk, MF Wendland, and PR Weinstein. Early detection of regional cerebral ischemia in cats: comparison of diffusion-and T2-weighted MRI and spectroscopy. *Magnetic Resonance in Medicine*, 14(2): 330–346, 1990.

JS Murday and Robert M Cotts. Self-diffusion coefficient of liquid lithium. *The Journal of Chemical Physics*, 48(11):4938–4945, 1968.

Klaus-Armin Nave and James L Salzer. Axonal regulation of myelination by neuregulin 1. *Current opinion in neurobiology*, 16(5):492–500, 2006.

Thoralf Niendorf, Rick M Dijkhuizen, David G Norris, Menno van Lookeren Campagne, and Klaas Nicolay. Biexponential diffusion attenuation in various states of brain tissue: Implications for diffusion-weighted imaging. *Magnetic Resonance in Medicine*, 36(6):847–857, 1996.

Markus Nilsson, Jimmy Lätt, Freddy Ståhlberg, Danielle Westen, and Håkan Hagslätt. The importance of axonal undulation in diffusion MR measurements: a Monte Carlo simulation study. *NMR in Biomedicine*, 25(5):795–805, 2012.

Dmitry S Novikov, Els Fieremans, Jens H Jensen, and Joseph A Helpern. Characterizing microstructure of living tissues with time-dependent diffusion. *arXiv preprint arXiv:1210.3014*, 2012.

Eleftheria Panagiotaki. *Geometric models of brain white matter for microstructure imaging with diffusion MRI*. PhD thesis, UCL (University College London), 2011.

Eleftheria Panagiotaki, Torben Schneider, Bernard Siow, Matt G Hall, Mark F Lythgoe, and Daniel C Alexander. Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *Neuroimage*, 59(3):2241–2254, 2012.

Sharon Peled, David G Cory, Stephen A Raymond, Daniel A Kirschner, and Ferenc A Jolesz. Water diffusion, T2, and compartmentation in frog sciatic nerve. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine*, 42(5):911, 1999.

Jean Philibert. One and a half century of diffusion: Fick, Einstein, before and beyond. *Diffusion Fundamentals*, 2(1):1–10, 2005.

Nadya Pyatigorskaya, Denis Bihan, Olivier Reynaud, and Luisa Ciobanu. Relationship between the diffusion time and the diffusion MRI signal observed at 17.2 tesla in the healthy rat brain cortex. *Magnetic Resonance in Medicine*, 2013.

Adrian E Raftery. Hypothesis testing and model selection. In *Markov chain Monte Carlo in practice*, pages 163–187. Springer, 1996.

TG Reese, O Heid, RM Weisskoff, and VJ Wedeen. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magnetic Resonance in Medicine*, 49 (1):177–182, 2003.

Simon Richardson, Bernard Siow, Eleftheria Panagiotaki, Torben Schneider, Mark F. Lythgoe, and Daniel C. Alexander. Viable and fixed white matter: Diffusion magnetic resonance comparisons and contrasts at physiological temperature. *Magnetic Resonance in Medicine*, 2013.

Brian D Ripley. Selecting amongst large classes of models. *Methods and models in statistics: In honor of Professor John Nelder, FRS*, pages 155–170, 2004.

Benoit Scherrer, Armin Schwartzman, Maxime Taquet, Sanjay P Prabhu, Mustafa Sahin, Alireza Akhondi-Asl, and Simon K Warfield. Characterizing the distribution of anisotropic micro-structural environments with diffusion-weighted imaging (DIAMOND). In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 518–526. Springer, 2013.

Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2): 461–464, 1978.

Kawin Setsompop, R Kimmlingen, E Eberlein, Thomas Witzel, Julien Cohen-Adad, Jennifer A McNab, Boris Keil, M Dylan Tisdall, P Hoecht, P Dietz, et al. Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *Neuroimage*, 80:220–233, 2013.

Kiran K Seunarine and Daniel C Alexander. Chapter 4. 2009.

Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

Jun Shao. Linear model selection by cross-validation. *Journal of the American statistical Association*, 88(422):486–494, 1993.

Timothy M Shepherd, Peter E Thelwall, Greg J Stanisz, and Stephen J Blackband. Aldehyde fixative solutions alter the water relaxation and diffusion properties of nervous tissue. *Magnetic Resonance in Medicine*, 62(1):26–34, 2009.

Stanford L Smith. Nuclear magnetic resonance imaging. *Analytical chemistry*, 57(4):595A–608A, 1985.

Christopher H Sotak. The role of diffusion tensor imaging in the evaluation of ischemic brain injury–a review. *NMR in Biomedicine*, 15(7-8):561–569, 2002.

Stamatios N Sotiropoulos, Timothy EJ Behrens, and Saad Jbabdi. Ball and rackets: inferring fiber fanning from diffusion-weighted MRI. *Neuroimage*, 60(2):1412–1425, 2012.

Greg J Stanisz, Graham A Wright, R Mark Henkelman, and Aaron Szafer. An analytical model of restricted diffusion in bovine optic nerve. *Magnetic Resonance in Medicine*, 37(1):103–111, 1997.

EO Stejskal. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *The Journal of Chemical Physics*, 43(10):3597–3603, 1965.

EO Stejskal and JE Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, 42(1):288–292, 1965.

Mervyn Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 111–147, 1974.

Mervyn Stone. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 44–47, 1977.

Aaron Szafer, Jianhui Zhong, and John C Gore. Theoretical model for water diffusion in tissues. *Magnetic Resonance in Medicine*, 33(5):697–712, 1995.

JE Tanner and Edward O Stejskal. Restricted self-diffusion of protons in colloidal systems by the pulsed-gradient, spin-echo method. *The Journal of Chemical Physics*, 49(4):1768–1777, 1968.

Maira Tariq, Torben Schneider, Daniel C Alexander, Claudia AM Wheeler-Kingshott, and Hui Zhang. In vivo estimation of dispersion anisotropy of neurites using diffusion MRI. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 241–248. Springer, 2014.

J Tournier, Fernando Calamante, David G Gadian, Alan Connelly, et al. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.

J-Donald Tournier, Fernando Calamante, and Alan Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459 – 1472, 2007.

Jacques-Donald Tournier, Susumu Mori, and Alexander Leemans. Diffusion tensor imaging and beyond. *Magnetic Resonance in Medicine*, 65(6):1532–1556, 2011.

Bruce D Trapp and Grahame J Kidd. Structure of the myelinated axon. *Myelin biology and disorders*, 1:3–27, 2004.

P Vangelderen, D DesPres, PCM Vanzijl, and CTW Moonen. Evaluation of restricted diffusion in cylinders. phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance, Series B*, 103(3):255–260, 1994.

Larry Wasserman. Bayesian model selection and model averaging. *Journal of mathematical psychology*, 44(1):92–107, 2000.

Van J Wedeen, Patric Hagmann, Wen-Yih Isaac Tseng, Timothy G Reese, and Robert M Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386, 2005.

Kenneth P Whittall, Alex L Mackay, Douglas A Graeb, Robert A Nugent, David KB Li, and Donald W Paty. In vivo measurement of T2 distributions and water contents in normal human brain. *Magnetic Resonance in Medicine*, 37(1):34–43, 1997.

BJ Wilm, Jonas Svensson, A Henning, KP Pruessmann, P Boesiger, and SS Kollias. Reduced field-of-view MRI using outer volume suppression for spinal cord diffusion imaging. *Magnetic Resonance in Medicine*, 57(3):625–630, 2007.

Junzhong Xu, Hua Li, Kevin D Harkins, Xiaoyu Jiang, Jingping Xie, Hakmook Kang, Mark D Does, and John C Gore. Mapping mean axon diameter and axonal volume fraction by MRI using temporal diffusion spectroscopy. *NeuroImage*, 2014.

Hui Zhang, Penny L Hubbard, Geoff JM Parker, and Daniel C Alexander. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *Neuroimage*, 56(3): 1301–1315, 2011.

Hui Zhang, Torben Schneider, Claudia A Wheeler-Kingshott, and Daniel C Alexander. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, 61(4):1000–1016, 2012.

Feng-Lei Zhou, Penny L Hubbard, Stephen J Eichhorn, and Geoffrey JM Parker. Coaxially electrospun axon-mimicking fibers for diffusion magnetic resonance imaging. *ACS applied materials & interfaces*, 4(11):6311–6316, 2012.