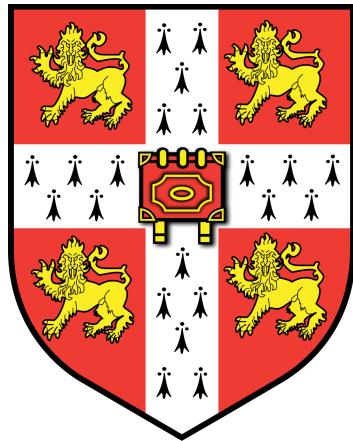


TOWARDS AN ACCURATE BRAIN TRACTOGRAPHY



ELEFTHERIOS GARYFALLIDIS

Wolfson College
University of Cambridge
UK

March 2012

A dissertation submitted to the University of Cambridge for the degree of
Doctor of Philosophy

Supervisors

DR. IAN NIMMO-SMITH

Medical Research Council
Cognition and Brain Sciences Unit
Cambridge, UK

DR. GUY B. WILLIAMS

Wolfson Brain Imaging Centre
Department of Clinical Neurosciences
University of Cambridge
Cambridge, UK

Dedication

To
my parents
Stavros and Amalia

for his creativity and her patience.

I needed both to excell.

To
the giants of OpenSource

who pushed the world forward.

Acknowledgements

Ian
Matthew
Guy
Marta
John + Virginia
Stephan
Frank
++ Daniel, Mina, Periklis, Andreas, Magda,
Rik, Director of Maths, College tutors new and old,
Vassilis, Dimitris, Alexandros
Tsiaras, Argyros, Vassilas

Disclaimer

This dissertation is the result of my own work and contains nothing which is the outcome of work done in collaboration with others, except where stated explicitly.

No part of this dissertation has previously been submitted for any degree or diploma at any institution.

This dissertation does not exceed 60,000 words in length (including tables, footnotes, bibliography and appendices).

Throughout this dissertation the plural pronoun 'we' is used for stylistic reasons and should be taken to refer to either the singular author, the reader and the author or, when stated explicitly, the author and collaborators. The form assumed should be apparent from the context.

Eleftherios Garyfallidis
January 2012

Publications

CONFERENCES

Garyfallidis E, Brett M, Nimmo-Smith I (2010), “*Fast Dimensionality Reduction for Brain Tractography Clustering*”, 16th Annual Meeting of the Organization for Human Brain Mapping.

Garyfallidis E, Brett M, Tsiaras V, Vogiatzis G, Nimmo-Smith I (2010), “*Identification of corresponding tracks in diffusion MRI tractographies*” Proc. Intl. Soc. Mag. Reson. Med. 18

Correia M.M, Williams G.B, Yeh F-C, Nimmo-Smith I, **Garyfallidis E** (2011), “*Robustness of diffusion scalar metrics when estimated with Generalized Q-Sampling Imaging acquisition schemes*”, Proc. Intl. Soc. Mag. Reson. Med. 19

Garyfallidis E, Brett M, Amirbekian B., Nguyen C., Yeh F-C, Olivetti E., Halchenko Y., Nimmo-Smith I (2011), “*Dipy - a novel software library for diffusion MR and tractography*”, 17th Annual Meeting of the Organization for Human Brain Mapping.

Garyfallidis E, Gerhard S, Avesani P, Nguyen T, Tsias V, Nimmo-Smith I, Olivetti E (2012), “*A software application for real-time, clustering-based exploration of tractographies*”, 18th Annual Meeting of the Organization for Human Brain Mapping.

JOURNALS

Garyfallidis E, Brett M, Correia M, Williams G.B, Nimmo-Smith I (2012), “*Effective Real-Time Tractography Clustering with QuickBundles*”, NeuroImage XXX , submitted.

Chamberlain SR, Hampshire A, Menzies LA, **Garyfallidis E**, Grant JE, Odlaug BL, Craig K, Fineberg N, Sahakian BJ (2010), “*Reduced brain white matter integrity in trichotillomania: a diffusion tensor imaging study.*” Arch Gen Psychiatry 67(9):965-71

OUTREACH

Garyfallidis E, Nimmo-Smith I (2010), "Surfing your brain superhighways".
Presented at the 350th celebration of the Royal Society, London, UK. In the presence of Her Majesty Queen Elizabeth II and His Royal Highness the Duke of Edinburgh.

SOFTWARE

Diffusion Imaging in Python – dipy.org
Free On Shades – fos.me

Abstract

The objective of this thesis is to improve on the methods for inferring neural tracts from diffusion weighted magnetic resonance imaging (dMRI). Accordingly, I present improvements to the reconstruction, integration, segmentation and registration modalities of dMRI analysis.

I compare and evaluate different Cartesian-grid q-space dMRI acquisition schemes, using methods based on the Fourier transform of the diffusion signal, with reconstructions by diffusion spectrum imaging or generalised q-ball imaging methods. I propose a new reconstruction method called diffusion nabla imaging (DNI) which works with all these acquisition schemes, using an algorithm that directly approximates the orientation distribution function using the Laplacian of the signal in q-space. DNI has impressive accuracy on low angle crossings.

Most previously published reconstruction methods are closely linked to their own specific track integration method. I have formulated a general, non-inferential, deterministic tractography algorithm (EuDX) which is based on Euler integration and trilinear interpolation, which works with voxel level information about fibre orientations including multiple crossings, and employs a range of stopping criteria. The purpose of this algorithm is to be faithful to the reconstruction results rather than try to correct or enhance them by introducing regional or global considerations.

I have developed an entirely new, fully automatic, linear time, clustering method (QuickBundles) which reduces massive tractographies to just a few bundles. These bundles are characterised by representative tracks which are multi-purpose and can be used for interaction with the data or as the basis for applying higher-complexity clustering methods which would have been impossible or too slow with the full data set. QuickBundles is currently the fastest known tractography clustering algorithm.

After applying QuickBundles to tractographies from different subjects, I show how to use the representative tracks to identify robust landmarks within each subject which I use to directly register the different tractogra-

phies together in a highly efficient way. The resulting correspondences provide important evidence for the anatomical plausibility of the derived bundles. I demonstrate how these methods can be used for group analysis, and for atlas creation.

This thesis contributes to the understanding of the diffusion signal in the context of dMRI acquisitions and builds on this foundation towards a more robust brain tractography which approximates more closely the underlying fibre architecture.

Contents

1	Background	4
1.1	Introduction	4
1.2	Molecular diffusion	5
1.3	Acquisition sequences	5
1.4	Single gradient signal models	9
1.5	Q-space reconstruction	13
1.6	Diffusion Tensor	17
1.7	Orientation Distribution	18
1.8	Tractography	20
1.9	Known problems	27
1.10	Segmentation	29
2	Cartesian Lattice Q-space Reconstructions	33
2.1	Overview	33
2.2	Theory	34
2.3	Equatorial Inversion Transform	39
2.4	Other methods	42
2.5	Implementation	44
2.5.1	Standard EIT	44
2.5.2	Fast EIT	45
2.6	Peak Finding	46
2.7	Spherical Angular Smoothing	47
2.8	Comparisons and Results	49
2.8.1	Multi-fibre Simulations	49
2.8.2	Digital Phantoms	55
2.8.3	Results with digital phantoms	57
2.8.4	Results with real data sets	62
2.9	Anisotropy metrics	67
2.9.1	Non-parametric Anisotropy	67
2.9.2	Quantitative Anisotropy	72
2.9.3	Robustness of QA	73
2.10	Discussion and Conclusion	77

3 Euler Delta Crossings Propagation	80
3.1 Overview	80
3.1.1 Local	81
3.1.2 Global	83
3.1.3 Simulated	85
3.2 Algorithm	86
3.3 Results with digital phantoms	91
3.4 Results with humans	95
3.5 Conclusion	100
4 Highly Efficient Tractography Clustering	102
4.1 Overview	102
4.2 Track distances and preprocessing	102
4.3 Related Work	106
4.4 Data sets	110
4.5 QuickBundles (QB) Clustering	111
4.5.1 The QB Algorithm	111
4.5.2 Powerful simplifications	116
4.5.3 Complexity and timings	118
4.5.4 Virtual tracks, exemplar tracks and other descriptors. .	123
4.5.5 The bi-directionality problem	124
4.6 Comparisons within- and between-subjects	125
4.6.1 Robustness under reordering	125
4.6.2 Measures to compare classifications	125
4.6.3 Tightness comparisons	130
4.7 Parallel version	132
4.7.1 Algorithm	132
4.7.2 Merging two sets of bundles	132
4.8 Direct applications	133
4.8.1 Rapidly detecting erroneous tracks	133
4.8.2 Alignments, landmarks and atlases	135
4.8.3 QB as input to other learning methods	138
4.8.4 Exemplars vs ROIs vs Masks	141
4.9 Affinity Propagation	141

4.10	Direct Tractography Registration	146
4.11	Strategies with Small fibres	150
4.12	Discussion and conclusion	152
5	Discussion	154
5.1	Overview	154
5.2	Clinical Usage	155
5.3	Why speed is so important?	155
5.4	Future Work	155
6	Appendix	180
6.1	Data and sequences	180
6.2	The cosine transform	181
6.3	Fourier transform of $P(\mathbf{r})r^2$	181
6.4	Radial projection of a symmetric function	182
6.5	The Tensor in GQI	183

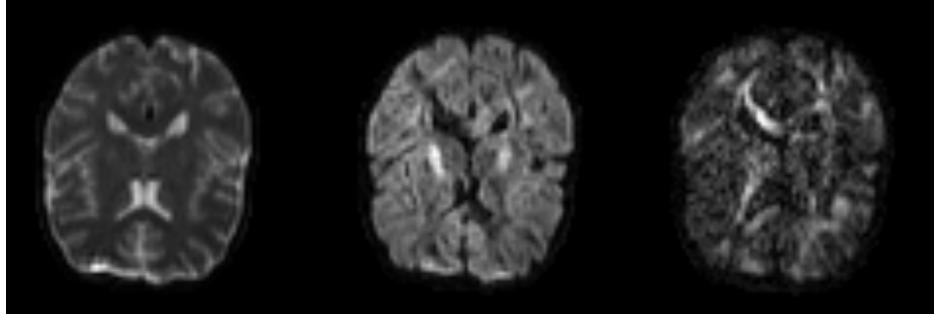


Figure 1.1: Three slices from diffusion data sets gathered with zero gradient strength on the left, medium gradient strength on the middle and high gradient strength on the right.

1 Background

1.1 Introduction

Diffusion MRI (dMRI) is the principal non-invasive method that provides information about the directional structure of neural tracts found in white matter and the cortex. dMRI acquires one or more T_2 -weighted reference images, and a collection of diffusion-weighted images (see Fig.1.1) that attenuate the T_2 signal according to the amount of diffusion along prescribed gradient directions [103]. The information is not complete and the tracts cannot be reconstructed in full detail [48]. However, some spatial structures and patterns can be visualised. These are usually represented as trajectories [197, 143] or connectivity maps [18]. The unique new area of study that aims to reconstruct the neural tracts from diffusion data is called diffusion tractography. Other types of tractography are based in staining using for example luxol-fast blue [134] but these can only be used with *in vitro* brains and they lack ease of repeatability. For non-human brains as for example in macaque there are even *in vivo* methods for tracing down to single axons [189] however these are not available or recommended for human studies because they are highly invasive.

1.2 Molecular diffusion

Molecular diffusion is a process that occurs incessantly in biological materials, fluids in particular, and accounts for a number of interesting phenomena; the dMRI signal measures the history of the random (Brownian) displacements of spin-labelled hydrogen protons (spins) resolved in the direction of a magnetic field gradient. Though the actual probability displacement function of the protons is unaffected by the presence or variation in the magnetic field, the cumulative phase change in the spins reflects the changes in the position-dependent spin frequency induced by the field gradient. Components of the diffusion motion along the direction of the gradient induce such changes. The signal change due to cumulative dephasing is greatest when this coincides with a direction that allows greater random displacements, e.g. because of the orientation of a microstructure within which the proton is moving. It is this link between the directional dependence of the dMRI signal and the orientations of the supposed underlying brain fibres that provides the unique insights of diffusion tractography. In dMRI it is assumed that the protons will move more along the directions of the axons and less perpendicular to that direction.

Anisotropy is one of the terms that are very common in diffusion terminology. Anisotropy means that the average displacement of the particles is greater along some directions than along others. On the other hand, isotropy means that the average displacement is equal in all directions i.e. complete lack of anisotropy. It is this level of anisotropy that is the basis of dMRI as a method of investigation of the structure of biological materials. For a biological interpretation of the signal measured with dMRI see [16], [208] and [103].

1.3 Acquisition sequences

MRI data are collected by changing certain magnetic fields on and off in a prescribed sequence, known as pulse sequence[103, 145]. The pulse sequence determines the content, quality, contrast and resolution of the image. MR images primarily reflect the signal from hydrogen nuclei from water and fat concentrations. The hydrogen nuclei possess a magnetic dipole

which is often referred to as spin. These dipoles can align themselves with an externally applied magnetic field. The MRI scanner generates a strong, static magnetic field B_0 which is typically measured in tesla (T). A second magnetic field is applied for only a brief duration and oscillates at radio frequencies; known as the RF pulse[144].

RF pulses are used primarily for excitation and refocusing. In the excitation phase spins will rotate away from their preferred orientation along B_0 . Excited spins precess about B_0 at a frequency ν given by the Larmor equation $\nu = \gamma B$ where γ is a constant known as the gyromagnetic ratio. The precessing part that is perpendicular to the direction of B_0 decays exponentially with a time constant T_2 and the spins realign themselves exponentially in the direction of B_0 with a time constant T_1 . T_1 and T_2 vary with tissue but $T_2 < T_1$ for the same tissue type [139]. The generated magnetic field from the coherently precessing spins induces a current in the receiver coils; this current is the signal used to generate MR images and corresponds to the image brightness. The more coherent the phase of the precessing spins the higher the brightness in the image pixels. However, with time, spins lose their phase coherence. Signal loss from both T_2 decay and dephasing is called T_2^* signal loss ($T_2^* < T_2$). Often a second RF pulse is applied at some time $TE/2$ after excitation and flips the spins in the plane perpendicular to B_0 . If the conditions stay the same all spins will be back in phase at a time TE after the excitation pulse. The moment of spin coherence is called a spin echo and it realizes photons. Acquisition sequences which use a refocusing pulse are called spin echo pulse sequences; and gradient echo sequence otherwise.

The final magnetic fields generated by an MRI scanner are called magnetic field gradients or simply gradients (G). Including the applied gradients the magnetic field in the scanner is given by $B = B_0 + G_x(t)x + G_y(t)y + G_z(t)z$ where x, y and z the three orthogonal directions. Gradients have a special role in diffusion weighting as we will discuss next.

The best known pulse sequence for generating diffusion-weighted images is called Pulsed Gradient Spin Echo method (PGSE), also known as the Stejskal and Tanner method [182]. This has 90° - 180° spin echo pair of RF pulses with one gradient before the second pulse and one equal gradient af-

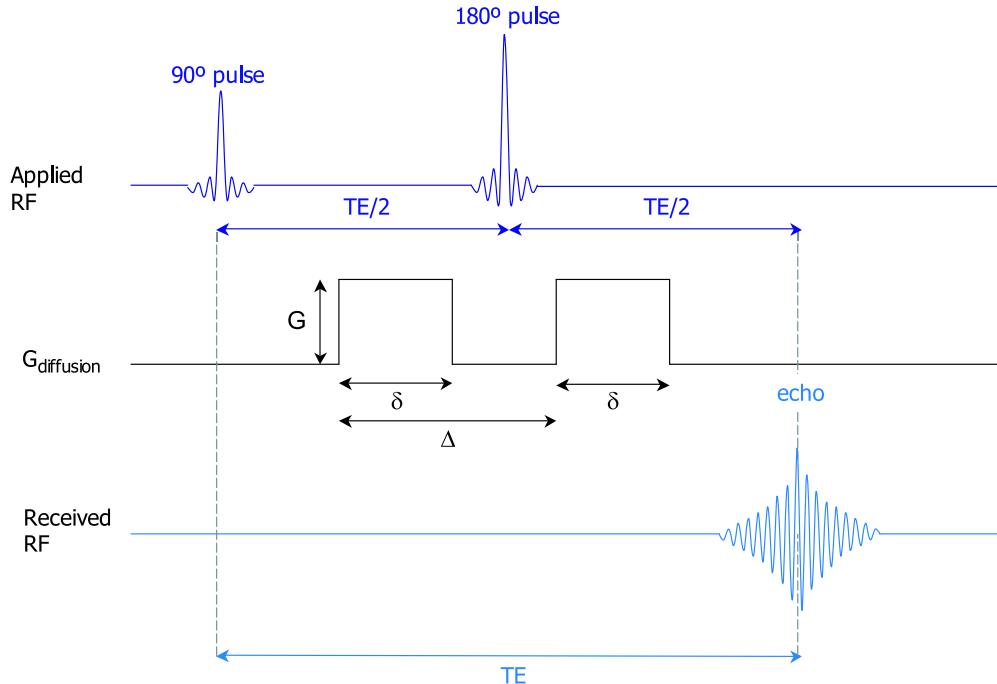


Figure 1.2: Pulsed Gradient Spin Echo (PGSE)

ter the second pulse (see [139], p329), see Fig. 1.2. The refocusing is perfect only when the spins do not move between the two pulses. The diffusion weighted contrast acts as an inverse T_2 weighting i.e. tissues with mobile water molecules give lower signal than more solid tissues with smaller mobility.

Eddy currents caused by the onset and offset of the gradients are a problem with PGSE and most recent systems (including Siemens scanners) use Twice-Refocused Spin Echo (TRSE) sequences [168] to reduce these artefacts. Every time the magnetic field gradients switch they generate currents that produce other smaller magnetic fields which disturb the spins. The TRSE sequence is an improvement on the PGSE. The improvement is made using another refocusing pulse surrounded by the inverse mirror of the previous diffusion gradients (see Fig. 1.3). By adjusting the timing of the diffusion gradients, eddy currents can be nulled or greatly reduced. This sequence improves the image quality without loss of scanning efficiency i.e. TR duration and it is the standard in most modern MRI scanners.

In the experiments described in this thesis we used a recent (2010) Work

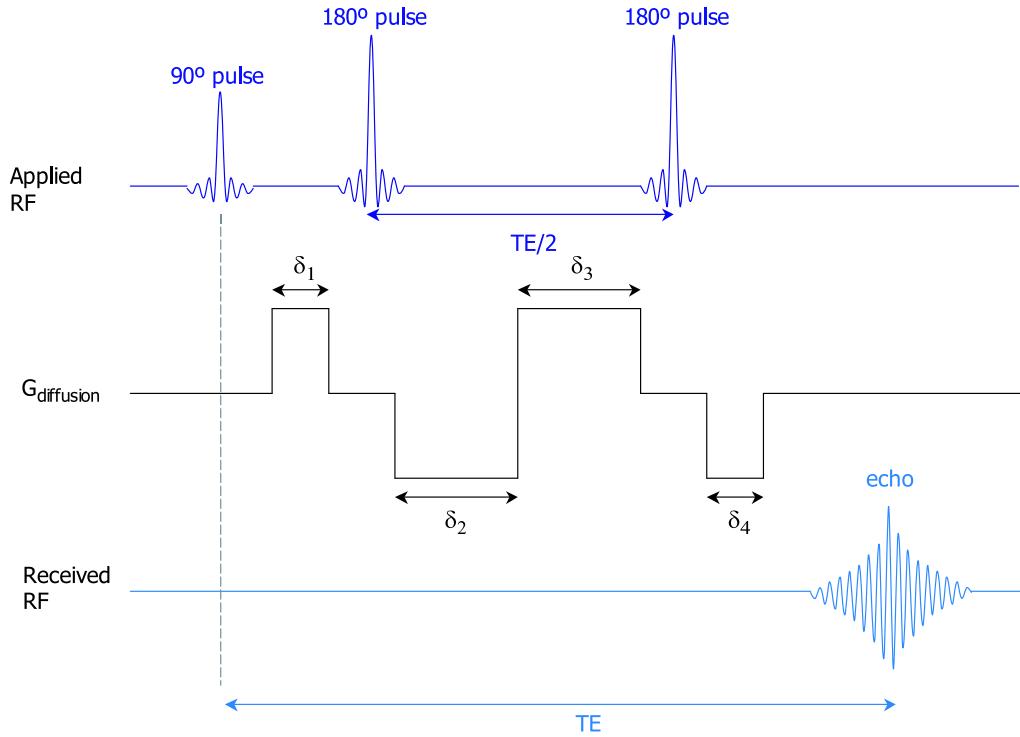


Figure 1.3: Twice-Refocused Spin Echo (TRSE)

In Progress (WIP) protocol from Siemens which uses the STEAM (STimulated Echo Acquisition Mode) sequence [135]. STEAM, is presented in Fig. 1.4 and works in the following way. Three 90° pulses are used to produce a stimulated echo. The first two pulses are separated by a time delay τ . After the same delay τ following the same pulse, a stimulated echo is produced. In order to introduce diffusion weighting into the stimulated echo, two identical diffusion gradient lobes are applied, one during the first and one during the second τ interval. Because the magnetization of the stimulated echo is stored along the longitudinal axis between the second and the third RF pulses, it does not experience any T_2 or T_2^* dephasing during the time interval TM . TM , however, does contribute to the diffusion gradient separation time, Δ . Thus, a high b-value can be obtained without incurring the TE -induced signal loss, as compared to the standard spin echo sequence. The signal amplitude of the stimulated echo is, however, less than that of the corresponding spin-echo sequence with the same TE , because the maximum amplitude of the stimulated echo is one-half of a spin echo.

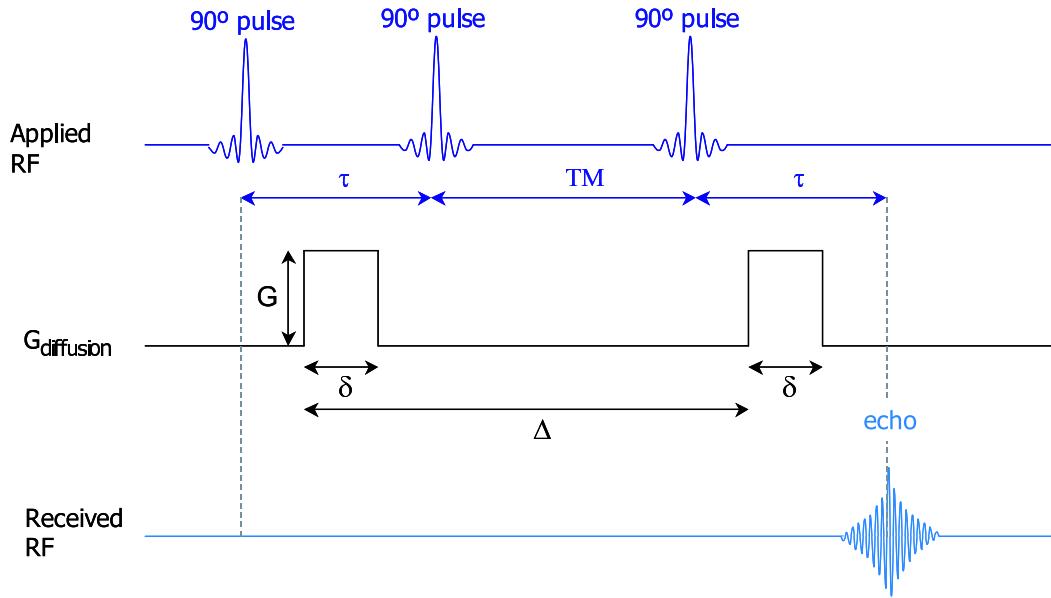


Figure 1.4: STimulated Echo Acquisition Mode (STEAM)

Using STEAM we were able to generate high quality datasets in nearly half the time per direction required with the standard DTI sequence. For example we were able to acquire 102 grid directions with maximum b-value 4000 at about 14 minutes, compared with 9 minutes for 64 directions with $b_{\max} = 1200$.

1.4 Single gradient signal models

Under the Brownian motion assumption the diffusion signal strength is described by the following model known as the Stejskal-Tanner[182] formula

$$S_b = S_0 e^{-bD} \quad (1)$$

where S_0 is the measured signal when no gradient direction is applied, D is the diffusion coefficient that we wish to measure and b the crucial experimental diffusion weighting parameter or b -value which summarises the amount of diffusion sensitising gradient history. D is often referred as the diffusivity value or apparent diffusivity coefficient (ADC). The units of D are mm^2/sec (for water at $37^\circ D \approx 3 \cdot 10^{-3} \text{m}^2/\text{sec}$), and of b are sec/mm^2 , typically in the range $0\text{--}5,000 \text{ sec}/\text{mm}^2$ though some acquisition paradigms

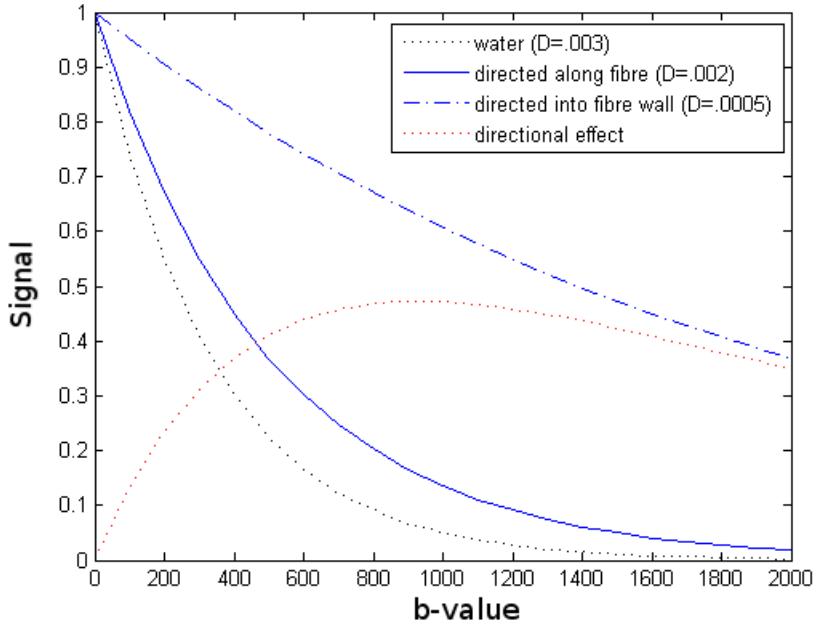


Figure 1.5: Signal as a function of b for various values of D

can call for very much larger values e.g. greater than 10,000 sec/mm²[28],[197]. In Fig. 1.5 we see the signal decay for different b-values and specific diffusivities. The obvious conclusion here is that signal in areas with high diffusivity, as in the corticospinal fluid (CSF) were water persists, will be always lower than the signal from areas of lower diffusivity as those found in fibrous brain structures, as the corpus callosum. Furthermore, there is least diffusivity and least signal loss when the gradient direction is into the wall of a fibre; along the fibre has greater diffusivity and greater signal loss; free water has greatest diffusivity and greatest signal loss. Signal decreases exponentially with b ; Finally we can see that the “directional effect”, i.e. the contrast between ‘along’ and ‘transverse’ signal is greatest when $b \sim 1000$.

The b-value b or *diffusion weighting* is a function of the strength, duration and temporal spacing and timing parameters of the specific paradigm. This function is derived from the Bloch-Torrey equations[24]. In the case of the classical Stejskal-Tanner pulsed gradient spin-echo (PGSE) sequence (described further in 1.3), at the time of readout

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

where γ is the gyromagnetic ratio, δ denotes the pulse width, G is the gradient amplitude and Δ the centre to centre spacing. γ is a physical constant, but we can change the other three parameters and in that way control the b-value.

Although the PGSE is useful for expository clarity, in reality as indicated in 1.3 more complicated but related sequences such as the twice-refocused spin-echo (TRSE) [92, 168] sequence and subsequent refinements such as TRASE [11] are employed as a means of removing the distortion effects from eddy currents resulting from the initial and final ramps of the gradient pulses.

An important point to stress here is that we can control the size of b-values by changing the strength and timings of the gradient pulses, and that depending on the b-value, we can expect different amount of signal loss. In Fig. 1.6 we see the directional dependence of the simulated signal of a single fibre oriented at 0° for two b-values. Note the important features that the signal is lower in the direction of the fibre and at higher b-value.

In eq.1 we assume that the signal can be expressed by a single exponential term (mono-exponential). In fact there is evidence that this assumption may break down at higher b -values and more complicated models have been proposed in order to deal with this issue. One of these, is the multi-exponential model used by Ozarslan et al.[156] which is expressed as

$$S_b = S_0 \sum_i^N f_i e^{-bD_i} \quad (2)$$

where N is the number of exponents or compartments, D_i is the i -th diffusion coefficient and f_i is the volume fraction of the i -th compartment.

Another model of higher order is found in diffusion kurtosis imaging (DKI) proposed by Jensen et al.[101]

$$\ln(S_b) = \ln(S_0) - bD + \frac{1}{6}b^2 D^2 K + O(b^3) \quad (3)$$

where K is the apparent diffusion kurtosis coefficient.

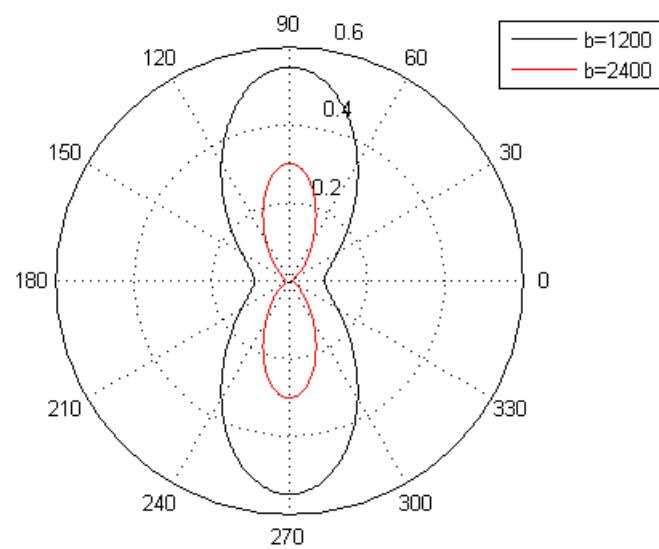


Figure 1.6: Directional dependence of signal for two b-values. Signal is drawn as a function of direction for a Gaussian diffusion function with a horizontal principal direction (0°) and two values of b . D is set to 0.002.

By putting together these single gradient models (Eq. 1,2,3) for every gradient direction we can create systems of equations where we can fit and identify their unknown parameters. In this thesis we will try to avoid fitting in that way but we will try to reconstruct the signal in an non-parametric Fourier-analysis based approach using the combined information from many gradient directions.

1.5 Q-space reconstruction

Bloch and Torrey [183] established differential equations governing MR diffusion in non-isotropic magnetic fields by analogy with Fick's Laws [65] for spontaneous dispersion along concentration gradients of inhomogeneous substances. Callaghan [24] also showed how these bulk properties could be derived by statistical methods from the collective spin histories of individual protons. When a molecule is at position x_0 , we cannot read exactly where it will be after time t , we can only model a distribution of possible locations. This motion is described by a propagator $P(\mathbf{x}; \mathbf{x}_0, t)$ which defines the probability of being in \mathbf{x} after a time t , starting at \mathbf{x}_0 .

Stejskal and Tanner [182] showed that the spin echo magnitude $S(\mathbf{q}, t)$ from a pulsed gradient spin echo (PGSE) experiment (see 1.3) is directly related to the diffusion propagator by the following (inverse) Fourier relation

$$S(\mathbf{q}, t) = S_0 \int P(\mathbf{r}, t) e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} \quad (4)$$

where S_0 is the signal in the absence of the applied magnetic diffusion gradient \mathbf{g} , \mathbf{r} is the relative spin displacement $\mathbf{x} - \mathbf{x}_0$ at diffusion time t , \mathbf{q} is the spin displacement wave vector, related to the applied magnetic gradient \mathbf{g} by the formula $\mathbf{q} = (2\pi)^{-1}\gamma\delta\mathbf{g}$. With the corresponding direct Fourier transform we can reconstruct the diffusion propagator P by measuring the signal in a number of different directions and gradient magnitudes. Q-space imaging (QSI) and Diffusion Spectrum Imaging (DSI) are the best known methods which try to reconstruct the full diffusion propagator in that way. In eq.4 the spin density is implicit in S_0 , in a later chapter we will see a more general interpretation of this relationship (see eq.11).

Q-space is the space defined by the coordinates of the 3D spin displace-

ment wave vectors \mathbf{q} as shown in equation 4. The vector \mathbf{q} parametrises the space of diffusion acquisitions. It is related to the applied magnetic diffusion gradient \mathbf{g} by the formula $\mathbf{q} = (2\pi)^{-1}\gamma\delta\mathbf{g}$ [122]. Every single vector \mathbf{q} has the same orientation as the direction of diffusion gradient \mathbf{g} and length proportional to the strength g of the gradient field, We actually have $\mathbf{q} = k\sqrt{b}\hat{\mathbf{g}}$ where b is the b-value, $\hat{\mathbf{g}}$ is the unit gradient direction $\mathbf{g}/|\mathbf{g}|$, and k is a multiplication constant which is a function of the timing parameters of the acquisition scheme.

Every single point in q-space corresponds to a possible diffusion weighted acquisition i.e. a 3D brain volume of measured signal for a specific gradient direction and strength (see Fig. 1.8). Therefore if for example we have programmed the scanner to apply 60 gradient directions then our data should have 60 diffusion volumes with each volume obtained for a specific gradient. A Diffusion Weighted Image (DWI) is the volume acquired from only one direction gradient. Hence, in the previous example we would gather 60 DWI volumes corresponding to 60 locations (\mathbf{q} -values) in \mathbf{q} -space. An alternative way to think of \mathbf{q} is in mathematical terms as the combination of parameters which produces the inverse Fourier transform relationship between the diffusion signal and the probability displacement distribution. In these terms (see Callaghan [24]) \mathbf{q} is the reciprocal of the probability displacement vector \mathbf{r} , just as in conventional MRI \mathbf{k} space is the reciprocal parametrisation of the brain space vector \mathbf{v} .

One problem in the diffusion imaging literature is that names for techniques often refer both to a particular type of imaging acquisition, and to a particular method to reconstruct the directional organisation of the voxel. All dMRI acquisition methods acquire data in q-space, and the methods can be categorised by their sampling pattern in q-space.

We refer to a method as a *q-space spherical shell* method if it is an acquisition method using a collection of points in q-space that can be thought of as being on a curved surface, typically a sphere. Examples include techniques referred to as HARDI (High Angular Diffusion Imaging), Q-ball Imaging and HYDI (Hybrid Diffusion Imaging). A q-space shell method might involve a single shell (Fig. 1.8B,D), e.g. QBI, or multiple shells (Fig. 1.8C) e.g. EQBI. In clinical settings where we can only use a few directions (less than

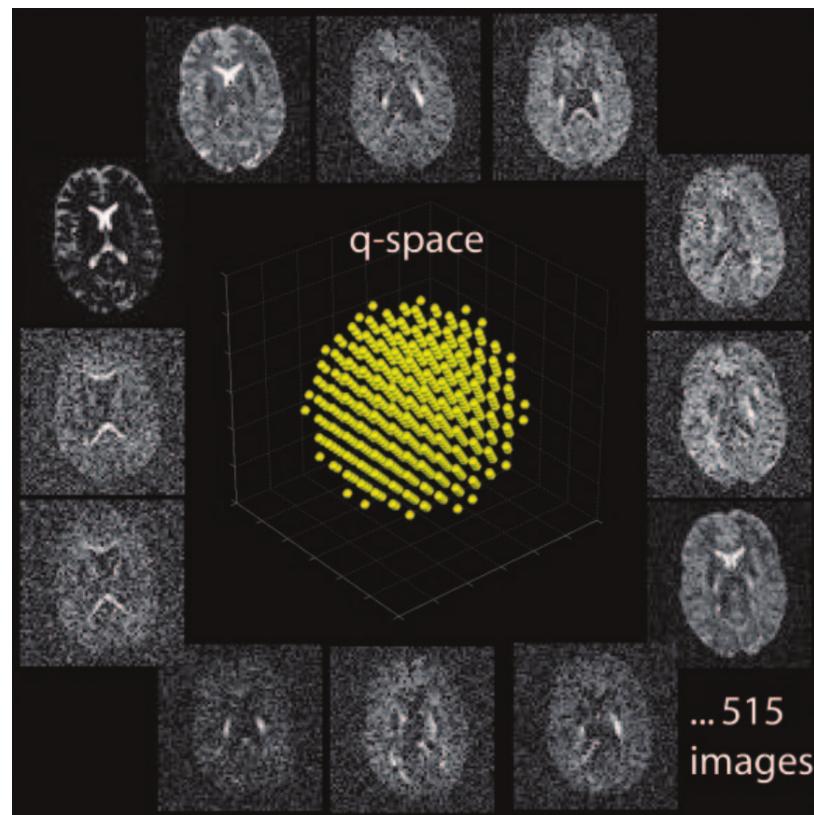


Figure 1.7: One volume is collected for every sampling point in q-space.
Picture from Hagmann et al. [89]

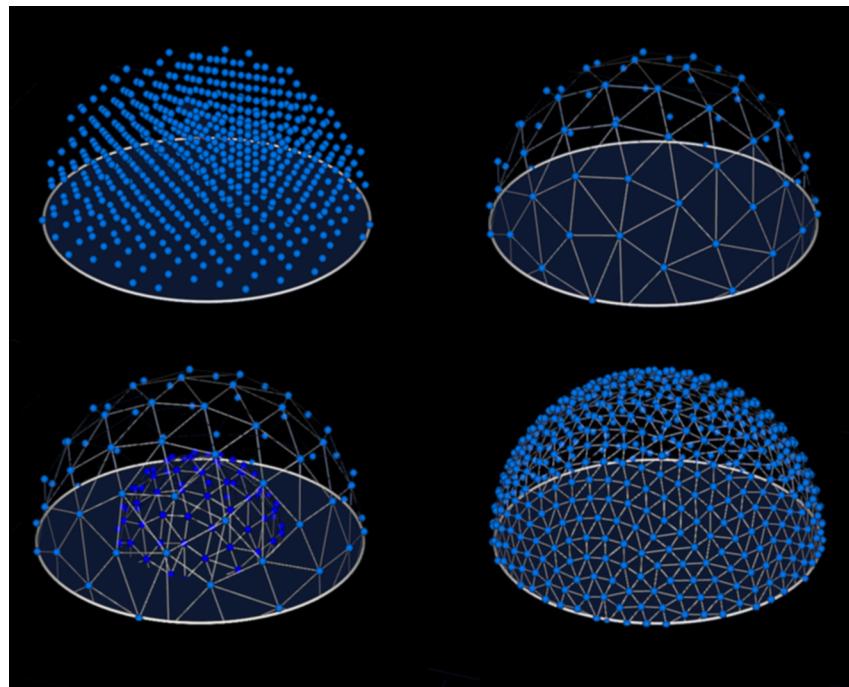


Figure 1.8: A (Upper Left): DSI 604 directions, B (Upper Right): HARDI 65 directions, C (Lower Left): HARDI with 2 shells of 65 directions in each shell, D (Lower Right): QBI 515 directions

60 with minimum 6) it is recommended to use the SDT (simple diffusion tensor) reconstruction model which is relatively easy to fit as it has only a few parameters.

A *q-space 3D grid (volume)* method is an acquisition method that is more easily thought of as a collection of points regularly distributed through a region in q-space. The characteristic example is DSI (diffusion spectrum imaging) (Fig. 1.8A). and QSI (*q*-space imaging)[109, 24]. In chapter ?? we will concentrate more on q-space grid methods and propose a new reconstruction method which we call DNI (Diffusion Nabla Imaging) which also uses q-space grid data.

1.6 Diffusion Tensor

Assuming that the diffusion propagator is given by a 3-dimensional Gaussian distribution from eq. 4we can write

$$P(\mathbf{r}, t) = \frac{1}{\sqrt{4\pi t^3 |\mathbf{D}|}} \exp\left(-\frac{\mathbf{r}^T \mathbf{D}^{-1} \mathbf{r}}{4t}\right) \quad (5)$$

where \mathbf{D} is known as the diffusion tensor. This tensor is a 3×3 positive definite symmetric matrix that can be completely described by a centred ellipsoid with 3 principal axes and associated eigenvalues $\lambda_1, \lambda_2, \lambda_3$. The trace of the diffusion tensor has been found valuable for detecting and evaluating brain ischemia and stroke [146, 195]. Frequently, mean diffusivity (MD) is used instead of the trace defined as

$$MD = \frac{\text{trace}(D)}{3} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \quad (6)$$

Fractional Anisotropy (FA) is the most common scalar metric used in diffusion imaging which is used to characterise the presence or absence of a preferred direction for diffusion. Like the MD it depends only on the eigenvalues.

$$FA = \frac{1}{\sqrt{2}} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (7)$$

If FA is equal to 1 that means very anisotropic (infinitely prolonged el-

lipsoid or a ‘stick’) and if FA is equal to 0 that means completely isotropic (sphere). FA is used in clinical studies to diagnose diseases like stroke and cancer and asses the progress of therapy [127].

Whenever we are using FA volumes we are implicitly assuming that the propagator of the spin displacements in every voxel has a 3D Gaussian distribution. This assumption is used in most of the diffusion related literature where DTI or Diffusion Tensor Imaging is the prevailing term. Unfortunately, in reality things are much more complicated; inside our brain the axons are semi-permeable (restriction), the water molecules interact with many different elements in the complex intra fibre fluid, the fibres might cross, kiss, divert or bend inside a voxel or between voxels. Therefore, assuming a Gaussian distribution is a non-trivial approximation. However, FA is still prevalent as it is easy to calculate and it gives similar values across different acquisitions.

1.7 Orientation Distribution

Since the primary interest of dMRI is the way that the signal depends on the direction of underlying fibre orientations, it is the orientation information in this propagator that is principally of interest.

One possible approach would be to replace the diffusion probability density function with an isosurface, which is a surface that passes through all points of equal value. For instance a 3D Gaussian isosurface is an ellipsoid. A more commonly used technique that is less sensitive to noise involves the computation of the orientation distribution function from the displacement distribution[89, 191, 198]. An orientation distribution function may be considered a sphere polar plot whose radius in a given direction is proportional to the integral of the diffusion probability density function in that direction. For ease of visualization, we color code the surface according to the diffusion direction ($[x,y,z] = [r,b,g]$, where r = red, b = blue, and g = green). An orientation distribution function or isosurface can be plotted for each individual MR imaging voxel in a section (see Fig.1.7).

The orientation distribution function (ODF) expresses the probability of a spin displacing into a differential solid angle about a possible fibre direc-

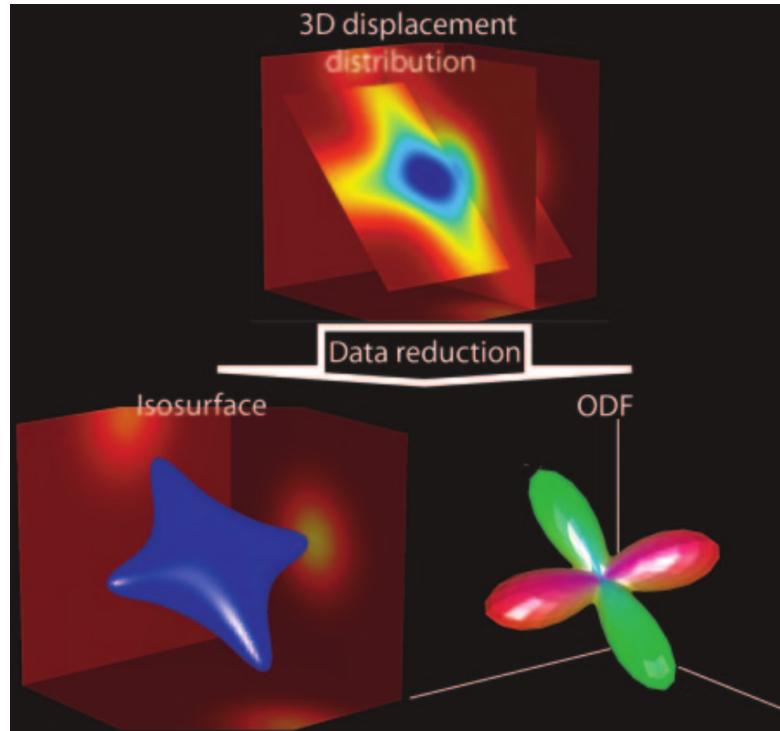


Figure 1.9: On the top picture we see the reconstruction of the 3D displacement probability distribution also known as the diffusion propagator or ensemble average propagator (EAP) or diffusion spectrum produced from the inverse Fourier transform of the diffusion signal. Two approaches that may be used to simplify the visual representation of the EAP is the replacement of the displacement distribution with an isosurface, and computation of the commonly used orientation distribution function (ODF). This displacement distribution resembles a crossing of two fibres, however the ODF is used to identify the primary directions of the fibres. Picture from Hagmann et al.[89].

tion $\hat{\mathbf{u}}$. This is used in order to model and visualise the directional information in diffusion propagator and in simple words it just projects the diffusion function on to the sphere by integrating over the radial coordinate of the diffusion function. The ODF representation symbolised below with f sacrifices all the radial information but retains the relevant directional information:

$$\psi(\hat{\mathbf{u}}) = \int_0^\infty P(r\hat{\mathbf{u}})r^2dr \quad (8)$$

where $\hat{\mathbf{u}}$ is a unit normal vector, and r is the radial coordinate in the diffusion space. By construction $\psi(\hat{\mathbf{u}})$ is a probability distribution over $\hat{\mathbf{u}}$.

The ODF is a function on the sphere. The sphere is usually represented by a discrete spherical grid with evenly distributed points. It is a common procedure to identify the direction of the underlying fibres from the points where the maximum values are found (see Fig.1.7). This procedure is also known as *peak finding*.

1.8 Tractography

Once we know the orientation of fibres at every voxel, we can try to join these directions up to reconstruct complete tracks and hence brain connections. In its simplest form, this consists of starting at a seed location and following the preferred direction until we reach a new voxel. We can then change to this voxel's referred direction and carry on until an entire track is propagated (see Fig. 1.8).

The two best known families of track propagation algorithms are: deterministic [36, 143] and probabilistic [20]. A track propagation algorithm belongs to the probabilistic domain if the fibre model that is being used incorporates uncertainty i.e. errors in estimating the orientation of the fibre at every voxel. If it does not assume any uncertainty along the path of the track then it is deterministic. In chapter 3 we will give a short overview of many other track propagation methods including global tractography[110].

One of the simplest and first deterministic methods is called Fibre Assignment by Continuous Tracking (FACT) [143] (see Fig. 1.12). The FACT algorithm starts through the input of an arbitrary point in the volume and then tracks in both directions e.g. forward and backwards. Perhaps the most

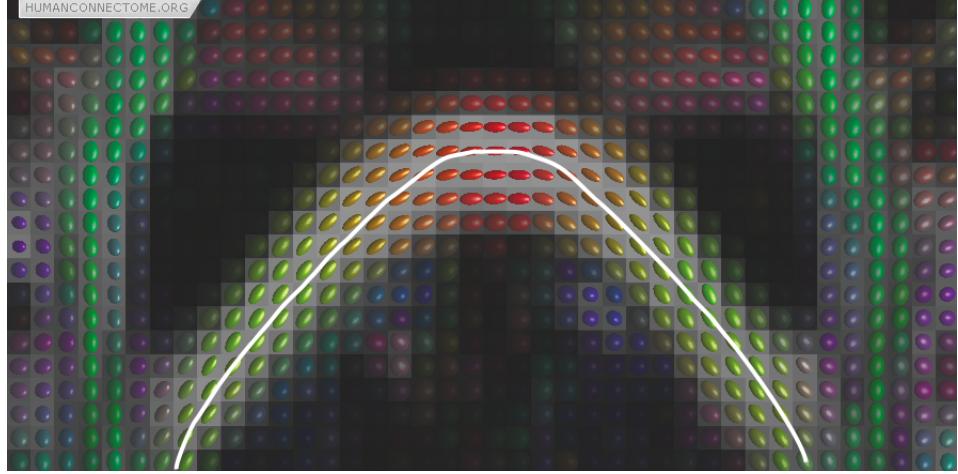


Figure 1.10: The white line shows the track obtained by connecting up a set of voxels based on the direction of the axis of the maximum tensor eigenvalue and is an example of deterministic tractography. The color is a complementary way of coding the tensor ellipsoid direction where red denotes left-right, green denotes back-front and blue up-down.

interesting part with these tracking methods is how they decide when to stop tracking. FACT uses a single threshold variable $R = \sum_i^s \sum_j^s |\mathbf{e}_i \cdot \mathbf{e}_j| / s(s - 1)$ where s is the number of neighbouring voxels and \mathbf{e} is the eigenvector corresponding to the largest eigenvalue in each voxel. The simplest case for defining a neighbourhood of a voxel is to use all 26 other adjacent (touching) voxels. Now, let's think of how R will behave in different neighbours. When adjacent fibres are aligned strongly R will increase as the absolute value of the dot product grows bigger as the normalised vectors become more co-linear. On the other side, R will grow shorter in regions without continuity in fibre direction, where $R = 0$ in perfect isotropic conditions. In voxels with R less than a prespecified threshold e.g. 0.8 the tracks will stop being propagated and FACT will terminate. An important problem with FACT is that it fails to track axons deviating from major pathways where it only tracks one of the major branches.

Another popular deterministic approach commonly used in the field of fluid dynamics for flow simulation was applied to the field of dMRI by [36] and [13] assuming again a continuous vector field but now the track is



Figure 1.11: Deterministic whole brain tractography based on EuDX and generated using Diffusion Imaging in Python (www.dipy.org). The color encodes the orientation of the mid-segment of every track using a colourmap based on Boy's real projective plane immersion [46].

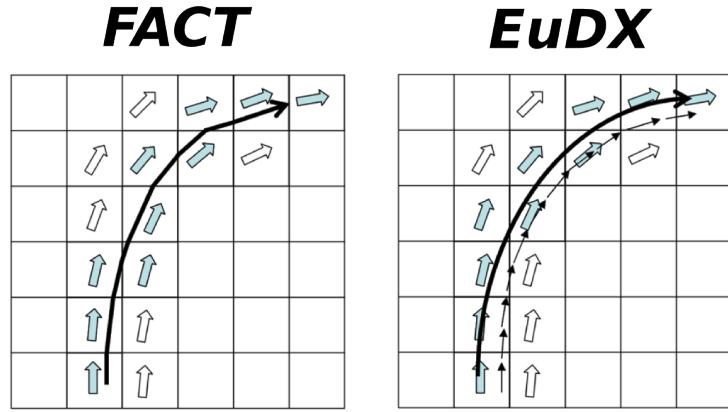


Figure 1.12: FACT (left) uses propagates with different steps indicated by the entering point on each voxel. Other deterministic methods use static steps. In that case interpolation is necessary. For example in EuDX (right) trilinear interpolation is used.

propagated by the solution of a system of differential equations subject to an initial condition, the position of the seed point. Here the authors propose that a fibre can be represented by a 3D curve \mathbf{r} parametrised by the arc length s of the track. The iterative solution of the equation

$$d\mathbf{r}(s)/ds = \mathbf{e}(\mathbf{r}(s)) \quad (9)$$

where \mathbf{e} is the primary direction of the Tensor or a different model. The solution of this system of ordinary differential equations (ODE) in its simplest case is given by iterative methods like Euler integration bound to an initial condition $\mathbf{r}(0) = \mathbf{r}_0$ where \mathbf{r}_0 is the seed point.

$$\mathbf{r}(s_1) \sim \mathbf{r}(s_0) + \alpha \mathbf{e}_1(\mathbf{r}(s_0)) \quad (10)$$

where $0 < \alpha \leq 1$ defines the integration finite step. In Euler integration we are using the first two terms of the Taylor expansion. Conturo et al.[36] used Euler integration and Basser et al. [13] used a higher order approximation called 4th order Runge-Kutta scheme.

Wedgeen et al.[197] showed that one could derive the local orientation field of vectors \mathbf{e} from the local maxima of the ODF calculated in each voxel. In that way we could visualise crossing distributions and depict crossing fi-

bres. The authors suggests that diffusion MRI with sufficient signal to noise ratio (SNR) could make tractography a mathematically well-posed problem. However, we need much more scanning time to reach the necessary resolution which is often impractical. Our novel tractography algorithm EuDX is presented in chapter 3 (see Fig.1.8).

In summary, the deterministic algorithms propagate tracks by making a series of discrete locally optimum decisions. These are fast, simple and easy to interpret. Usually, we depict them using streamlines. A streamline is a curve that is always tangent to the velocity vector of the flow. The disadvantages of deterministic algorithms are that a pathway either exists or not (no uncertainty) and that they do not explore the entire space of possible white matter tracts. In other words, local thresholding makes our tractography vulnerable to small noise aberrations.

Probabilistic tractography is meant to deal with this problem of noise and propagate tracks even in regions where the tracking is unclear. This is made possible by assuming that uncertainty exists concerning the orientation of the fibre at each point of the track.

Imagine a particle in a seed voxel moving in a random manner with a constant speed within the brain whitematter. The transition probability to a neighbouring point depends on the local orientation distribution or underlying model. This yields high transitional probabilities along the main fibre directions. Hence, the particle will move in parallel to the fibre direction with a higher probability than in a perpendicular direction. In this probabilistic method, we start a large number of particles from the same seed point, let the particles move randomly according to the local ODF and count the number of times a voxel is reached by the path of a particle (connectivity values). The random walk is stopped when the particle leaves the white matter volume. The 3-dimensional image of the connectivity values is called a tractogram (see Fig. 1.8).

We will try to illustrate the difference between the two approaches using a simplified 2D example shown in Fig. 1.13 (i,ii). Generalising afterwards in 3D is straightforward.

Let's imagine that we have a 2D slice where in each pixel we have calculated a vector showing the primary direction for that specific pixel. This

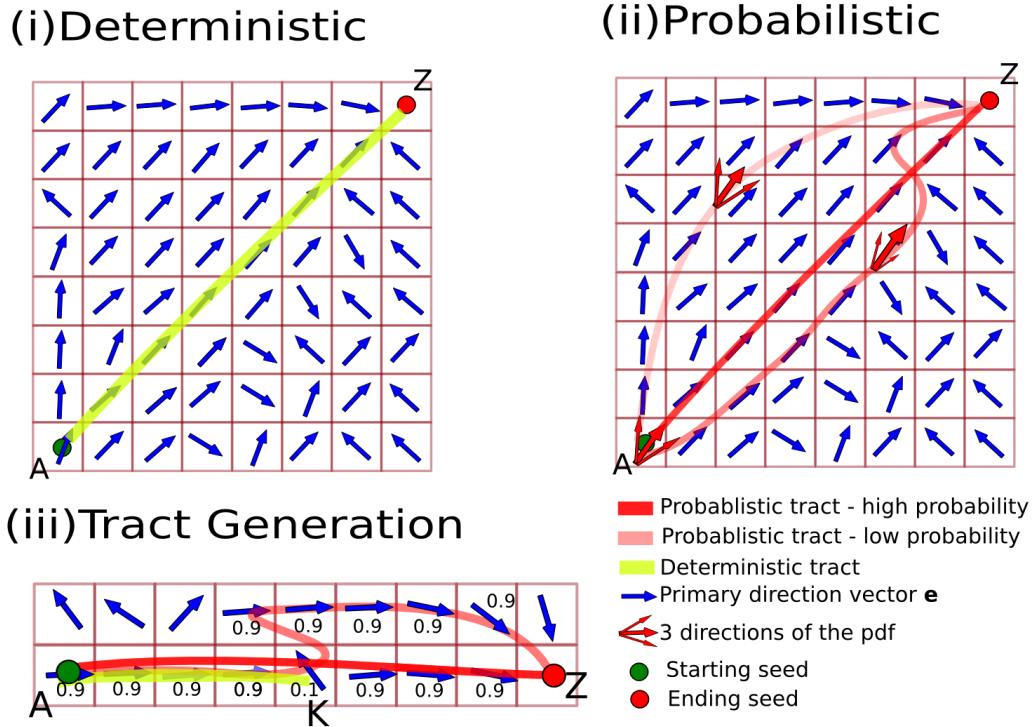


Figure 1.13: A simplified example showing in (i) & (ii) the same dataset but in (i) the yellow line shows the result of deterministic tractography which is given by a single trajectory and in (ii) is given by connectivity matrix depicting with red the probability of different pathways throughout the hole slice. Here only 3 possible pathways are depicted for the ease of understanding. Finally, in (iii) an example is given where it shows that probabilistic tractography weights more closer connections. However it can tract further deep than deterministic tractography.

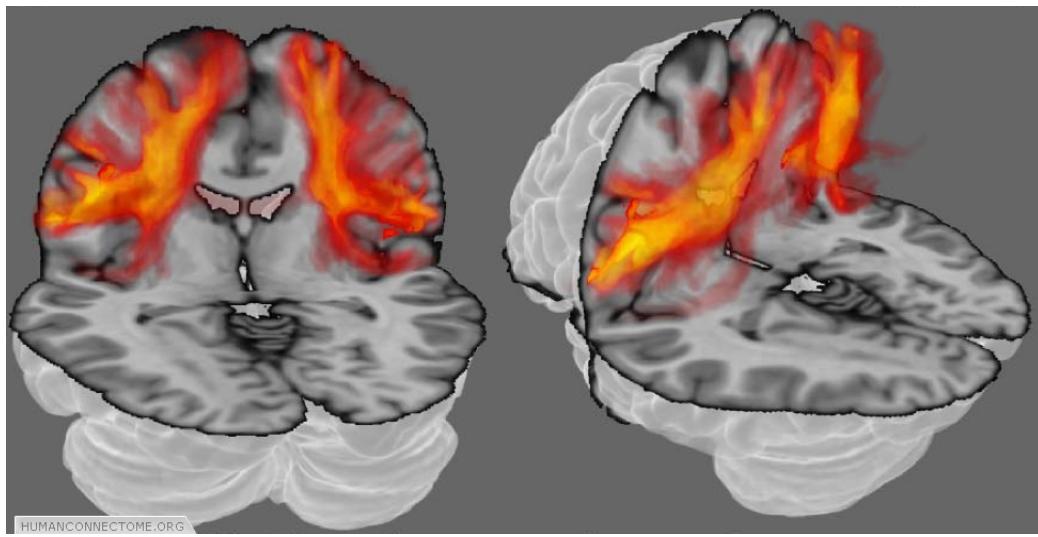


Figure 1.14: The 3D distribution of voxels connected to the seed voxel is called a tractogram.

vector e could have been calculated from the tensor as the principal eigenvector or as a principal direction of a different model e.g. the maximum point of the ODF. Let's now think that we want to find the best track from seed A to seed Z. When using deterministic tractography we are using only the local direction information in every pixel therefore in a direction field as this of Fig. 1.13(i) we will have to use only one track and this is the diagonal pathway with yellow colour. However there are other possible tracks as well in this diagram e.g rather than taking the diagonal we could go first up north from A and then right.

Probabilistic tractography aims to identify all the possible tracks by assigning to each one of them a weight. All the weights of all the tracks together sum to 1. This is possible by generating samples from a probability distribution for every pixel. In this toy example shown in Fig. 1.13(ii) the orientation of the blue vectors can be represented by a single parameter, angle ω . ω here is a random variable that takes values from a probability distribution function (PDF). Now we have many possible directions to move next but with different probabilities. The weights of all directions again sum to 1. After this explanation we can identify in Fig. 1.13(ii) that the most likely track is again the diagonal (with deep red) but there are other possible tracks

(with lighter red) that are less likely. In the same diagram we show with 3 combined red arrows some of the many directions that are possible in each point. However, some are more certain than others.

Let's try now to understand how a track is valued as more probable than others. In Fig. 1.13(iii) we have drawn a very simple image with only two pixel rows and we are assuming that the probability of moving along the primary direction (shown with blue arrow) is 0.9 and there is only a secondary direction given by 0.1 (1-0.9) i.e. we assume for ease of understanding here only 2 directions. We can see that there is a discontinuity in position K. In that point an Euler based deterministic approach (without interpolation) would have to stop at K (see yellow line). The probabilistic method will continue tracking and it will generate two tracks that both reach the target. The probability of each track is calculated by multiplying the probability of a specific direction of each point. Therefore, the shorter and dark red track will have $p_s = 0.9^7 \times 0.1$ and the longer and lighter red track will have $p_l = 0.9^{10}$. It is obvious that $p_s > p_l$ and that the darker red track is more likely to exist according to this method.

This method of multiplying the probabilities at each voxel along the paths has been proposed by many [76],[6], [22], [87],[93],[165] and used in many software packages as well , FSL FDT [19, 17, 18] in FSL, and others. However there are some problems.

1.9 Known problems

Although the probabilistic methods are able to identify many known tracts, they miss several large tracts such as the visual pathways LGN-MT and callosal MT[177]. The visual pathways are useful test cases for algorithmic development and testing because they diverge and bend significantly. Probabilistic methods have the advantages that they expand the track search space beyond deterministic algorithms and that they can easily expand with complex models supporting crossings (usually not more than 2 crossings). However, they don't compute an accurate probability of brain connections. The phraseology "Connection probabilities" or "estimation of global connectivity" or "the probability of the existence of a connection through the data

field, between any two distant points" found in [19] can be very misleading because someone might believe that they represent the actual connectivity profile of the subject. For example we are certain (with probability 1) that LGN (lateral geniculate nucleus) is connected to V1 and V2 (primary visual cortex) in any healthy brain however the estimated connection probability in FDT is much less than 1. In addition, current probabilistic algorithms fail to identify pathways even when they are known to exist or in a few cases they generate pathways even when they do not exist [20, 99, 176]. For example no connections between left MT+ and the posterior portion of Corpus Callosum were found in PiCo [159] or FDT although it is well established that they do exist.

In 2008 Sherbondy[177] in order to try to deal with the problems explained in the previous paragraph introduced an algorithm that separates the pathway sampling and scoring steps. In that way the scoring does not depend any more on whether we are tracking from seed A to seed B or from B to A therefore, it assumes symmetry when the other probabilistic methods do not. At the same time Sherbondy's method assumes independence between different tracts i.e. path $A \rightarrow B$ is independent from $A \rightarrow C$ or $K \rightarrow L$. This does not happen in most other methods where a pathway depends on other pathways starting from the same seed. So, Sherbondy's method was designed to estimate connections that are known to exist. The disadvantage of this method is that it needs a lot of user interaction to add the known tracks and the user needs to be a specialist in white matter anatomy otherwise the results might be biased. Other tools like FSL's FDT uses waypoint masks to reduce the effect of this problem. But these need to be defined by the user defined too.

A small summary of the comparative strengths of deterministic and probabilistic tractography is given in the qualitative Table 1.1 where we can see that although deterministic tractography will most likely stop more often at a noisy voxels it is much faster to calculate than probabilistic tractography. They both can generate non-existing tracts because of propagation errors or errors in the reconstruction step. For noise related problems e.g. motion and eddy correction and possible solutions see [93, 13, 106, 91, 214, 5, 202], and for methodology and ideas comparing across subjects see [107, 144, 127, 49,

	Deterministic	Probabilistic
Voxel Noise Resistance	Less	More
Non-existing Tracts	Yes	Yes
Execution Time	Fast	Slow
Memory Size	Less	More
Biased on Tract Length	Yes	Yes

Table 1.1: Known problems with deterministic and probabilistic tractography

[45, 149, 58, 44]. We showed in the previous section that probabilistic tractography usually gives higher weight to shorter pathways. In the deterministic tractography an opposite weighting is required as longer pathways will have higher representation in the datasets and so it is more likely to have more seeds along a long track rather than along a short track. Therefore, a normalization by length both for probabilistic and deterministic tractography are highly recommended.

1.10 Segmentation

The white matter contains pathways known as fibre tracts that connect functional areas of the brain. A diagram of commonly found fibre tracts is sketched in Fig. 1.10 and real fibre tracts are shown in Fig. 1.10. The white matter contains three types of fibre tracts: commissural, association, and projection[150]. Commissural tracts connect related regions of the two cerebral hemispheres. Association fibres connect regions in the same hemisphere. Association fibres come in various sizes: the smallest fibres are completely within the cortex, the medium ones are called u-fibres and connect one gyrus to the next, and the longest association bundles connect different lobes. Finally, projection fibres connect the cortex and subcortical structures such as the thalamus, basal ganglia, and spinal cord.

A major issue with white matter is that it is particularly dense (see Fig. 1.10) i.e. the boundaries between different tracts are very difficult to distinguish, many bundles cross or touch other bundles (see Fig. 1.10) and most bundles diverge into smaller and smaller tracts as they reach and enter gray matter areas.

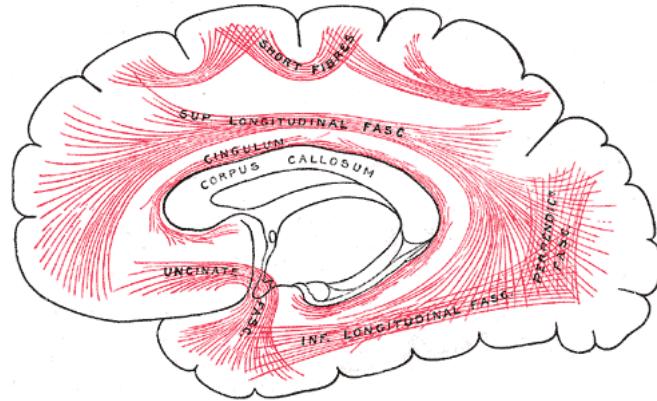


Figure 1.15: Diagram showing principal systems of association fibres in the cerebrum. The white matter fibre tracts are large bundles of axons that interconnect the gray matter processing areas both within and across hemispheres. The association fibres connect fibres from the same hemisphere. Picture from Gray's Anatomy #751 [85]

Previously it was not possible to automatically create white matter models similar to these anatomical atlas diagrams of fibre tracts *in vivo*. However, methods are now available to estimate white matter fibre trajectories using diffusion MRI. In this thesis we present a method for segmentation of the trajectories estimated from diffusion MRI by automatically grouping them into anatomical regions or to be more accurate regions of similar proximity and shape characteristics. This is the main topic of chapter 4 where we propose QuickBundles an extremely efficient algorithm for tractography segmentation.

There are three main goals which should be satisfied by an automatic tractography segmentation algorithm[150]: automatic grouping of like trajectories into regions, region correspondence across subjects, and anatomical labeling of regions. Our ability to perform automatic, subject-specific definition of the white matter fibre tracts has applications in neuroanatomical visualization, neurosurgical planning, and neuroscientific studies of white matter integrity, structure, and variability.

In the next chapters we will use the following definitions. Define a track (or streamline) as a polyline $s = \{x_1, \dots, x_n\}$, where $x \in \mathbb{R}^3$. Then the entire tractography is defined as $T = \{s_1, \dots, s_m\} \sim \mathfrak{T}$, where usually the number

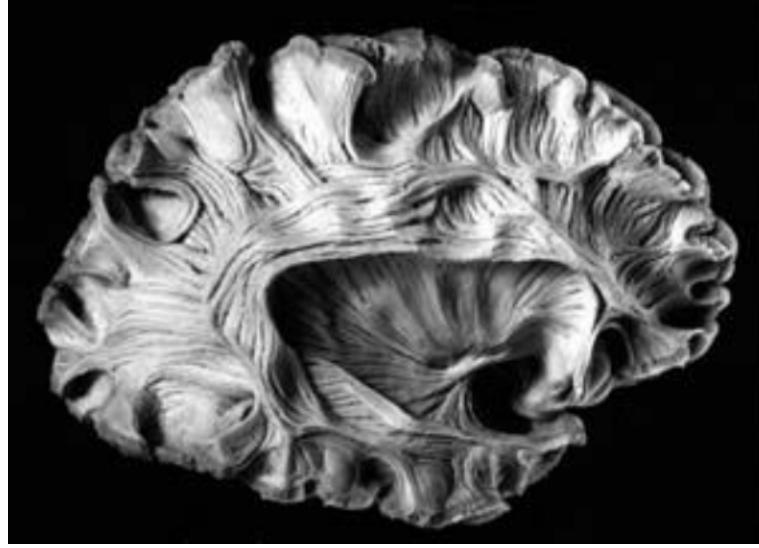


Figure 1.16: Fibre pathways are so densely packed in the real brain that a segmentation algorithm of some kind looks like a possible solution for the neurosurgery planning of the future and further understanding of the brain connectivity. Picture from virtual hospital[201].

of tracks is $|T| \simeq 2 \times 10^5 - 2 \times 10^6$. For an anatomical (physical) bundle (tract) e.g. arcuate fasciculus we use v and for a fibre bundle we use u where $u \subset T$, which approximates v . We think it is very important to always have in mind that we can only approximate a real tract.

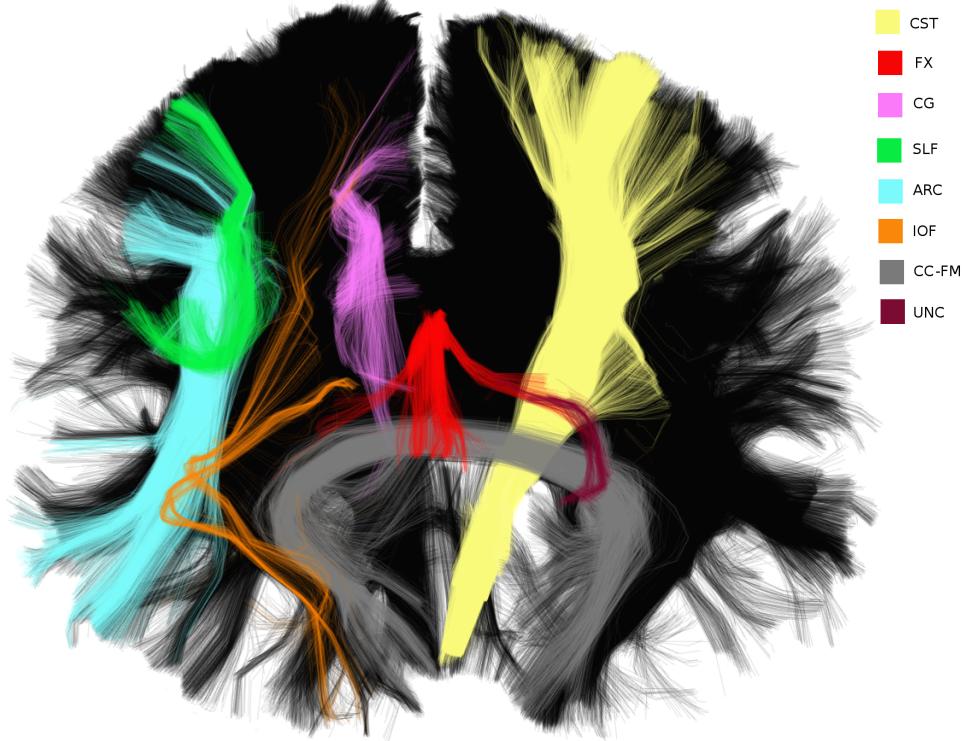


Figure 1.17: An example of a tractography segmentation based on labeling provided by a neuroanatomist. A major project of this thesis is to automatically find the labels or simplify the work of an expert by clustering tractographies. CST: Corticospinal Tract, FX: Fornix, CG: Cingulum Bundle, SLF: Superior Longitudinal Fasciculus, ARC: Arcuate Fasciculus, IOF: Inferior Occipitofrontal Fasciculus, CC-FM: Corpus Callosum Forceps Major, UNC: Uncinate Fasciculus. Data and labels from braincompetition.org

2 Cartesian Lattice Q-space Reconstructions

2.1 Overview

Between one-third to two-thirds of imaging voxels in the human brain’s white matter are thought to contain multiple fibre bundle crossings[20], in which case the Diffusion Tensor model proposed by Basser et al.[14] breaks down. High Angular Resolution Diffusion Imaging (HARDI) [188] such as Diffusion Spectrum Imaging (DSI) [25], [198] or Higher Order Tensors [155], [12] and many more reconstruction methods have been proposed to overcome the limitations of the Diffusion Tensor. These methods can be divided into those who need specific acquisition parameterization for example they have to be sampled in one or more spherical grids like in Qball Imaging[191], or in a Cartesian grid like in Generalized Q-sampling Imaging (GQI) [205] and those which can be used independently of q-space structure like DTI. A further division considers the level of model assumptions for the diffusion process. Although all methods have some underlying assumptions we generally separate them in model-based and model-free. Model-based methods like the Single Tensor or Multi Tensor necessitate the fitting from a few to many parameters. On the opposite side with model-free methods fitting is not necessary and the directionality of the underlying tissue can be approximated by some re-parametrization or re-transformation of the signal. The latter is usually more efficient than fitting models with many parameters.

This document presents and evaluates different model-free methods for the reconstruction of orientation distribution functions using diffusion MRI data sampled on a Cartesian lattice in \mathbf{q} -space. This non-parametric nature of the algorithms described here allows for the identification of multiple fibre crossings. In addition a new method is presented named Diffusion Nabla Imaging (DNI) and a family of methods is defined called the Equatorial Inversion Transform (EIT). The EIT is a new way to represent and reconstruct the diffusion signal. Our results show that it can perform better or as good as the current state-of-the art methods i.e. DSI and GQI.

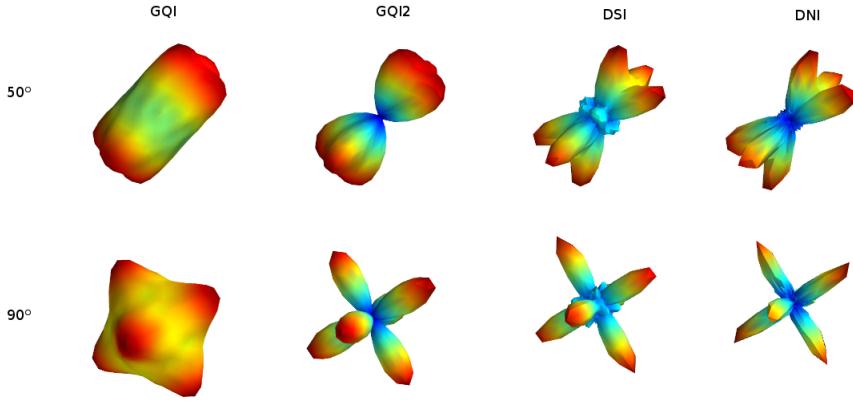


Figure 2.1: Showing the ODFs from two randomly oriented simulated 3-fibre crossings at 50° (top) and 90° angles between each other using different grid based reconstruction methods. The parameters used here are DSI: radial sampling $2.1 - 6$, hanning filter width 36 , GQI: $\lambda=1.2$, GQI2: $\lambda = 3$, DNI: standard.

2.2 Theory

We start from the classical formulation shown in eq.11 of joint \mathbf{k} -space and \mathbf{q} -space imaging (Callaghan 8.3.1 p. 438)[24], [25] using the narrow pulse gradient spin echo (PGSE) sequence of Tanner and Stejskal.

$$RF(\mathbf{k}, \mathbf{q}) = \int \rho(\mathbf{v}) \exp(i2\pi\mathbf{k} \cdot \mathbf{v}) \int P_\Delta(\mathbf{v}, \mathbf{r}) \exp(i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{r} d\mathbf{v} \quad (11)$$

Here RF is the complex RF signal measured at spatial wave number \mathbf{k} and magnetic gradient wave number \mathbf{q} , ρ is the local spin density (number of protons per unit volume contributing to the RF signal), Δ is the diffusion time scale of the sequence, P_Δ is the average diffusion propagator (transition probability distribution), \mathbf{v} is the voxel coordinate and \mathbf{r} is the diffusion displacement.

The \mathbf{k} -space reconstruction gives us diffusion weighted image data S which reveal the average propagator P_Δ of each voxel

$$S(\mathbf{v}, \mathbf{q}) = \int \rho(\mathbf{v}) P_\Delta(\mathbf{v}, \mathbf{r}) \exp(i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{r} \quad (12)$$

For the rest of the chapter we can drop \mathbf{v} as we assume that the formula-

tion is the same for every voxel and Δ because the time scale is the same for the entire sequence. We can also replace the spin density $\rho(\mathbf{v})$ with S_0 i.e. the measured signal without diffusion weighting $\mathbf{q} = \mathbf{0}$. Therefore we can write

$$S(\mathbf{q}) = S_0 \int P(\mathbf{r}) \exp(i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{r} \quad (13)$$

By applying the 3D Fourier transform in eq. 13 we can reconstruct the average propagator also known as the diffusion spectrum [196]

$$P(\mathbf{r}) = S_0^{-1} \int S(\mathbf{q}) \exp(-i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{r} \quad (14)$$

or diffusion propagator. It was shown by Wedeen et al. [196] that the dMRI signal is positive for any type of spin motion without net flux (i.e. spin displacements due to thermal molecular agitation) or other random fluxes such as intravoxel incoherent motion. Under this assumption we can replace in eq. 14 the complex signal S with its modulus $|S|$

$$P(\mathbf{r}) = S_0^{-1} \int |S(\mathbf{q})| \exp(-i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{r} \quad (15)$$

The modulus of the signal coincides with the output of the standard MRI scanners as DWI and that simplifies the acquisition procedure. It represents the density of the average relative spin displacement in a voxel. In other words, $P(\mathbf{r})d\mathbf{r}$ is a measure of the probability for a spin in a considered voxel to make, during the experimental mixing time Δ , a vector displacement \mathbf{r} . We can visualize the propagator for every voxel as a 3D density volume (SEE Figure Introduction).

At each location, diffusion-weighted images are acquired for $N = 515$ or less values of \mathbf{q} -encoding, comprising in \mathbf{q} -space the points of a cubic lattice within the sphere of five lattice units in radius. Therefore,

$$\mathbf{q} = \alpha\mathbf{q}_x + \beta\mathbf{q}_y + \gamma\mathbf{q}_z \quad (16)$$

with $\alpha, \beta, \gamma \in \mathbb{Z}^+$ and $(\alpha^2 + \beta^2 + \gamma^2)^{1/2} \leq 5$. The signal is premultiplied by a Hanning window before Fourier transform in order to ensure a smooth

attenuation of the signal at high q values.

Often to obtain data for the complete grid of 515 \mathbf{q} -vectors (which also means that we need to collect 515 diffusion weighted volumes) is very time consuming and fewer number of unique \mathbf{q} -vectors are employed only at a single hemisphere usually between 102 to 257 points [114]. This is valid because the signal is symmetric therefore the vectors can be mapped on the other hemisphere to create the full \mathbf{q} -space.

Since we are mainly interested in the angular structure of the underlying tissue, we further simplify the data by taking the weighted radial summation of $P(\mathbf{r})$

$$\psi_{DSI}(\hat{\mathbf{u}}) = \int_0^{\infty} P(r\hat{\mathbf{u}})r^2 dr \quad (17)$$

This defines the orientation density function (ODF) for DSI which measures the quantity of diffusion in the direction of the unit vector $\hat{\mathbf{u}}$ where $\mathbf{r} = r\hat{\mathbf{u}}$.

Notice at this point that in order to find the ODF we have to create first the diffusion propagator by applying the Fourier transform on the lattice. Yeh et al.[205] proposed a direct analytical way to calculate a slightly different ODF using the Cosine transform.

In order to represent the average propagator in the scale of spin quantity Yeh et al.[205] introduced the *spin density function* Q which is estimated by scaling the average propagator P_{Δ} with the spin density ρ , i.e. $Q(\mathbf{r}) = \rho P(\mathbf{r}) = S_0 P(\mathbf{r})$. From eq. 13 we obtain

$$S(\mathbf{q}) = \int Q(\mathbf{r}) \exp(i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{r} \quad (18)$$

We can apply the Fourier transform again in eq.18 and obtain

$$Q(\mathbf{r}) = \int S(\mathbf{q}) \exp(-i2\pi\mathbf{q} \cdot \mathbf{r}) d\mathbf{q} \quad (19)$$

Because $Q(\mathbf{r})$ is real and $S(\mathbf{q})$ is symmetric (even), i.e. $S(\mathbf{q}) = S(-\mathbf{q})$, we can use directly the Fourier Cosine transform (see 6.2) to calculate

$$Q(\mathbf{r}) = \int S(\mathbf{q}) \cos(2\pi \mathbf{q} \cdot \mathbf{r}) d\mathbf{q} \quad (20)$$

and obtain the “spin” orientation distribution function (SDF) ψ_{GQI} from an unweighted truncated radial projection

$$\psi_{GQI}(\hat{\mathbf{u}}) = \int_0^\lambda Q(r\hat{\mathbf{u}}) dr \quad (21)$$

$$= \int_0^\lambda \int S(\mathbf{q}) \cos(2\pi r \mathbf{q} \cdot \hat{\mathbf{u}}) d\mathbf{q} dr \quad (22)$$

$$= \int S(\mathbf{q}) \text{sinc}(2\pi r \mathbf{q} \cdot \hat{\mathbf{u}}) d\mathbf{q} \quad (23)$$

where λ is a constant called the diffusion sampling length. This parameter acts like a smoothing factor. The higher λ the more detailed the SDF will be but also more noisy. This ODF was used as the basis of the analysis of the GQI method. It provides a simple direct analytical solution of the ODF which can be written in a simple matrix form

$$\psi_{GQI} = \mathbf{s} \cdot \text{sinc}((6D \cdot G \circ \mathbf{b} \circ \mathbb{1}) \cdot G) \lambda / \pi$$

where \cdot denotes standard matrix or vector dot product, \circ denotes the Hadamard product, \mathbf{y} is ψ_{GQI} as a M -dimensional vector for every direction $\hat{\mathbf{u}}$, \mathbf{s} is a vector with all the signal values, $6D = 0.01506$ where D is a constant known as the free water diffusion coefficient, G is the $N \times 3$ matrix with the gradient vectors, \mathbf{b} is the $N \times 1$ matrix with the b-values and $\mathbb{1}$ is the $N \times 3$ incidence matrix where all values are equal to 1.

For a similar ODF like the one produced using DSI we need to take the weighted truncated radial projection. This will give us a different “spin”

ODF which we symbolize with ψ_{GQI_2}

$$\psi_{GQI2}(\hat{\mathbf{u}}) = \int_0^\lambda Q(r\hat{\mathbf{u}})r^2 dr \quad (24)$$

$$= \lambda^3 \int S(\mathbf{q}) H(2\pi r \mathbf{q} \cdot \hat{\mathbf{u}}) d\mathbf{q} \quad (25)$$

$$\text{where } H(x) = \begin{cases} \frac{2\cos(x)}{x^2} + \frac{(x^2-2)\sin(x)}{x^3}, & x \neq 0 \\ 1/3, & x = 0 \end{cases}.$$

This equation can be similarly implemented with a simple matrix transform

$$\psi_{GQI2} = \mathbf{s} \cdot \mathbf{H}((6D \cdot G \circ \mathbf{b} \circ \mathbf{1}) \cdot G) \lambda^3 / \pi$$

and until today it hasn't been published with real or simulated data sets.

The addition of the spin density plays a very important role on normalizing the ODF and providing scalar or vector metrics for the analysis of dMRI data sets. GQI similarly with DSI expects the q-vectors to sit on a cubic lattice within a sphere. However, because of the direct analytical formulation of the GQI ODFs; the creation of the volumetric grid with the signal values is not necessary. This makes this GQI advantageous on memory and CPU efficiency. Furthermore, no Hanning filter is necessary.

A new method for the calculation of the real ODF is proposed here. This is based on the theoretical work done by Aganj et al. [1],[28] using two important theorems from Fourier Analysis

1. The Fourier transform of $P(\mathbf{r})r^2 = -\nabla^2 E(\mathbf{q})$ where ∇^2 is the Laplacian operator (for proof see 6.3).
2. For a symmetric function $E : \mathbb{R}^3 \rightarrow \mathbb{R}$ and for the arbitrary unit vector $\hat{\mathbf{u}}$ we have $\int_0^\infty E(r\hat{\mathbf{u}})dr = \frac{1}{8\pi^2} \int \int_{\hat{\mathbf{u}}^\perp} E(q)qdq d\phi$ where $\hat{\mathbf{u}}^\perp$ is the plane perpendicular to $\hat{\mathbf{u}}$ (for proof see section 6.4).

From eq. 17 we see that the integration is over $P(r\hat{\mathbf{u}})r^2$, therefore we can write

$$\psi_{DNI}(\hat{\mathbf{u}}) = -\frac{1}{8\pi^2} \int_{\hat{\mathbf{u}}^\perp} \int_0^\infty \nabla^2 E(q) q dqd\phi \quad (26)$$

where ϕ is the angular rotation component operating on the plane perpendicular to $\hat{\mathbf{u}}$, ∇^2 is the Laplacian operator and $E(q) = S(q)/S_0$ is the normalized diffusion signal. Eq. 26 has the advantage that no Fourier transform is necessary. However we need to have a way to calculate the Laplacian of the signal. This can be analytically derived for a spherical grid [1] and we propose here that it can be directly calculated in a cubic grid using the standard 3D discrete Laplacian filter which is given by the 3D kernel defined by the following three planes

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & -6 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This is a function commonly used for image processing. From now on when we use the Laplacian operator in order to measure the directionality of the diffusion signal we will call this reconstruction method Diffusion Nabla Imaging as nabla (∇^2) is the symbol for the Laplacian operator.

2.3 Equatorial Inversion Transform

I propose an important theoretical construction called the Equatorial Inversion Transform (EIT) which creates a general formulation for the interpretation of the directionality of the diffusion signal. This idea is founded on two general properties of the diffusion signal: a) If we visualize the diffusion signal for a single fiber for all gradient directions we see that a shape is generated which is smallest towards the direction of the fiber and biggest on the plane perpendicular to that direction (see fig. 2.2). b) The diffusion signal is additive i.e. $S(\hat{\mathbf{f}}_1) + S(\hat{\mathbf{f}}_2) = S(\hat{\mathbf{f}}_1 + \hat{\mathbf{f}}_2)$, where $\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2$ are the unit directions of the fibres. In simple terms the signal of 2-fibre crossing can be decomposed linearly to the signals of the two fibres that create the crossing. The same holds for any number of fibres in a crossing.

These are two very important geometric properties of the signal that we can try to exploit at the maximum by calculating equatorial integrals in order to identify the directionality of the signal.

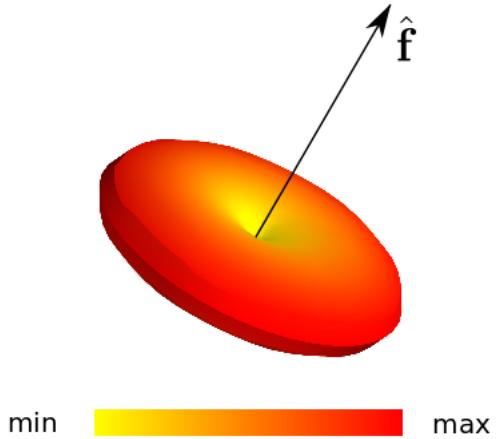


Figure 2.2: The diffusion signal has the beautiful property to be minimum along the direction of a fibre with unit direction \hat{f} and maximum along the equator defined by the plane perpendicular to that fibre direction. This property is the basic inspiration behind the EIT. In this picture the 3D surface plot of a simulated signal for a spherical grid acquisition with b -value 2000 is shown using a yellow-red colourmap.

Apart from the visual confirmation further supporting evidence that equatorial integration is crucial for derivation of directionality can be seen in eq. 26 where an equatorial integral which creates a connection between the real ODF and the signal. The Funk-Radon Transform (FRT) used by [191] is another example where equatorial integration is employed using the reconstruction sphere. We will see next that DNI and FRT are just a subset of the EIT.

With EIT the most important goal is to try to identify the correct directionality of the signal in the most accurate way by generating a spherical density. However it is possible to calculate as well the real ODF as defined by Wedeen et al. [196].

The EIT shown in eq. 27 consists of an integration along the equator and along radial lines. A function F of the signal is multiplied by a radial weighting function O . This construction is a generalization of the previous ODFs and it can support successfully many different function families for F and O which can all more or less accurately identify the directional distribution of the signal. More precisely the EIT is defined as

F	O	Name	Comment
II	1	DNI \equiv EITL	calculates the real ODF without the complications of the FFT
III	1	EITL2	high resolution at low angles
I	1	EITS	impressive resolution without any preprocessing of the signal
I	0	'QBI'-like	similar to the Funk Radon Transform

Table 2.1: The Equatorial inversion transform (EIT) can be used to explain many other reconstructions algorithms.

$$\psi_{EIT}(\hat{\mathbf{u}}) = \int_{\hat{\mathbf{u}}^\perp} \int_0^\infty F(E(q))O(q)dq d\phi \quad (27)$$

where F could be for example any of the following functions

$$F(E(q)) = \begin{cases} E(q) & (I) \\ -\nabla^2(E(q)) & (II) \\ \nabla^4(E(q)) & (III) \\ \dots \end{cases} \quad (28)$$

and O could be for example any of the following functions

$$O(q) = \begin{cases} 1 & (0) \\ q & (1) \\ q^2 & (2) \\ \dots \end{cases} \quad (29)$$

In table 2.1 we see that different functions of F or O can generate older well known or new distribution functions on the sphere. With $F(E(q)) = -\nabla^2(E(q))$ and $O(q) = q$ we can generate ψ_{DNI} which is theoretically identical to the DSI real ODF(ψ_{DSI}). If $F(E(q)) = E(q)$ and $O(q) = 1$ then this is similar with the Funk Radon Transform (used in Qball imaging) but applied to multiple spherical shells. However, we can also try to use different functions like $F(E(q)) = -\nabla^4(E(q))$ and $O(q) = q$ which can potentially

increase the amount of directional information beyond than the standard ODFs. Before starting investigating the realms of EIT we should first give a short overview of other methods commonly found in the literature. These are grid-based, mostly non-grid based and usually parametric.

2.4 Other methods

Pickalov et al.[166] proposed a new method for reconstructing the diffusion propagator by applying an iterative inverse Radon transform on measurements along many radial lines; computing 1D tomographic projections to reconstruct the 3D EAP. This technique measures DW images along a few radial lines of q-space but still requires hundreds of samples to reliably recover the EAP. Currently, to reconstruct the EAP, the state-of-the-art model-free techniques apart from diffusion spectrum imaging are hybrid diffusion imaging (HYDI) [204] and multiple q-shell diffusion propagator imaging (mq-DPI)[52]. HYDI acquires the signal values on 5 concentric spherical q-space shells, then interpolates to a cubic grid and applies the standard Fourier transform in the same way as DSI. In mq-DPI the EAP is calculated by solving Laplace's equation for the diffusion signal using a real and symmetric modified spherical harmonic basis. The EAP can be found analytically by the inversion of a linear system using Laplace-Beltrami regularization. In addition exact q-ball imaging (EQBI)[27] provides a different method to calculate the ODF analytically using multiple spherical q-space shells and similarly Aganj et al.[2], proposed an analytical solution for the multi-shell case by incorporating a mono-exponential or bi-exponential model (CSA-ODF). Another distribution on the sphere was proposed by Özarslan et al.[156] named diffusion orientation transform (DOT). This method calculates a different statistic $P(r_0\hat{\mathbf{u}})$, the probability of finding the particle initially at the origin, at the point $r_0\hat{\mathbf{u}}$ using spherical harmonics. Not surprisingly there is a relationship connecting CSA with DOT which is

$$\psi_{CSA}(\hat{\mathbf{u}}) = \int_0^\infty DOT(r\hat{\mathbf{u}})r^2dr \quad (30)$$

Jansons et al. [97] proposed a different function on the sphere than the ODFs described above, to be used on data sets acquired on a single spherical q-space shell. They called this spherical function persistent angular structure (PAS). This method has very good angular resolution because it uses the principle of maximum entropy however it is rather slow as it uses non-linear fitting in order to identify many parameters. PAS is a statistic on the sphere defined as $PAS(\hat{\mathbf{u}}) = \exp(\lambda_0 + \sum_{j=1}^N \lambda_j \cos(\mathbf{q}_j \cdot k\hat{\mathbf{u}}))$ where λ are the unknown parameters, k is constant and N is the number of DWIs. The relationship $\int PAS(\hat{\mathbf{u}}) \exp(i\mathbf{q}_j \cdot k\hat{\mathbf{u}}) d\hat{\mathbf{u}} = E(\mathbf{q}_j)$ provides the bridge between PAS and the diffusion signal ($E(\mathbf{q})$).

The first reference of using spherical harmonic expansions with diffusivity profiles, which are now very favorable in the literature, was by Alexander et al.[3]. Qball imaging was introduced by Tuch [191] and a new ODF defined as $\psi(\hat{\mathbf{u}}) = \frac{1}{Z} \int_0^\infty P(r\hat{\mathbf{u}}) dr$ where Z is a normalization constant. It was later provided for QBall imaging a fast and analytical solution using spherical harmonics (SH) and Laplace-Beltrami regularization [50]. Tournier et al. [185],[186] introduced a spherical deconvolution method where first the SH coefficients were reconstructed by , then single fiber ODFs are used as a deconvolution kernel estimated from the real data. Then the sharper fODF was obtained by a simple linear transformation [51]. Other deconvolution approaches were proposed in [173] and [206].

On Tensor related methods we have the classical SingleTensor [14], Sticks&Ball[20], Multi-Tensor[163][124] and Higher Rank Tensors [155],[12]. In addition there are also model based methods which try to calculate non-Gaussian properties for example the Kurtosis Tensor [101],[125] which is used in Diffusion Kurtosis Imaging (DKI).

Finally, new model-based methods are emerging which are trying to calculate statistics like the axonal thickness distribution from dMRI data sets. These are usually based on modeling free and restricted components; CHARMED [8][10], AxCaliber [9] and the orientation invariant ActiveAx[4] are some well known methods of this type. Q-space Imaging(QSI) can be used to identify distributions of axon-diameter too[154].

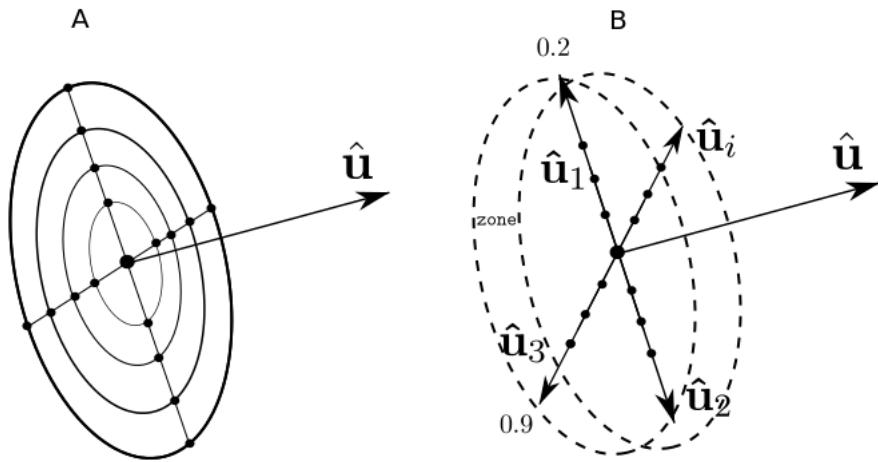


Figure 2.3: (A)Standard EIT vs (B)Fast EIT algorithm. Fast EIT is an order of magnitude faster than standard EIT. The key idea here is that reduce computations by storing the sum of the radial integrals for every vertex in the reconstruction sphere and then we can also precompute the indices of the vertices that are near the equator of every vertex (inside an equatorial zone).

2.5 Implementation

2.5.1 Standard EIT

Eq. 26 and 27 can be implemented in a standard way by evaluating the 3D signal on the grid multiple times for every direction \hat{u} as shown in fig. 2.3(A). This suggests that if for example we use a reconstruction sphere of 642 vertices and the radial integration (q) takes place in 30 steps and the equatorial (ϕ) in 63 steps then we need to interpolate $642 \times 30 \times 30 \simeq 1.2$ million times on the cubic grid. For this reason I invented Fast EIT a new method that needs an order of magnitude less evaluations.

In this document whenever we use the prefix s in front of a method that will mean that this was calculated with the standard EIT algorithm. For example if standard EIT is used for DNI we will write sDNI or sEITL. Of course sDNI and sEITL are equivalent.

Algorithm 1 Fast Equatorial Inversion Transform

Input $U = \{\hat{u}_1 \dots \hat{u}_m\}$, E

Output ψ_{EIT}

Foreach \hat{u}_i **Do**

$$J_i = \{j : |\arccos(\hat{u}_i \cdot \hat{u}_j)| \leq z\}$$

Foreach \hat{u}_i **Do**

$$\mathbb{B}(\hat{u}_i) = \sum_{k=0}^n F(E(q_k \hat{u}_i)) O(q_k \hat{u}_i)$$

where $F(E(q_k \hat{u}_i))$ is interpolated on the lattice.

$$\psi_{EIT}(\hat{u}_i) = \frac{1}{N_i} \sum_{j \in J_i} \mathbb{B}(\hat{u}_j)$$

where N_i is the number of indices in J_i .

2.5.2 Fast EIT

A much faster algorithm than the standard EIT is described here. The main idea is that we can store the sum of the radial integrals for every vertex in the reconstruction sphere and then we can also precompute the indices of the vertices that are near the equator of every vertex (inside an equatorial zone) see fig. 2.3(B). After these calculations the spherical distribution function can be approximated with much less operations. The full algorithm is given in 1. The input is the vertices \hat{u}_i of the reconstruction sphere and the normalized signal E. Then for every point of the reconstruction sphere \hat{u}_i we save the indices of the vertices j of \hat{u}_j , which are inside an equatorial zone, in list J_i . The width of the equatorial zone z is a constant set empirically to 5° . If a very highly dense reconstruction sphere is used with more than 642 vertices which is the one we used then the zone can be smaller. That can potentially increase the angular resolution of the method.

At the next stage we calculate sums along every radius on the direction of \hat{u}_i in the following way $\mathbb{B}(\hat{u}_i) = \sum_{k=0}^n F(E(q_k \hat{u}_i)) O(q_k \hat{u}_i)$ and obtain the final EIT ODF as the average of the sums in the equator $\psi_{EIT}(\hat{u}_i) = \frac{1}{N_i} \sum_{j \in J_i} \mathbb{B}(\hat{u}_j)$ where F is evaluated with trilinear for example interpolation on the lattice and N_i is the number of indices in J_i .

In section 2.8.1 the standard EITL (sEITL) is compared with fast EITL. In fig. 2.8.1 we can see that the fast EIT has very similar results with the standard EIT therefore from now on whenever we see EIT written in this document we will assume that the fast version is used.

Algorithm 2 Peak Finding with a Symmetric Ordered Sphere

Input ODF ψ , faces Φ

Output peaks P and indices I

ForEach face Φ_i **Do**

$f_0, f_1, f_2 = \Phi_i$

$d_0, d_1, d_2 = \psi[f_0], \psi[f_1], \psi[f_2]$

If $d_0 \geq d_1$ **and** d_2 **Do**

$P[f_1] = P[f_2] = 0$

continue

2.6 Peak Finding

After we have generated the ODFs we need to find the peaks (local maxima) from which we can easily approximate the direction of the fibres. Peak finding can be non-trivial if there are many local maxima in the ODFs or the ODFs are noisy. Here we present an algorithm (??) which reduces the amount of small local variations and returns a number of sorted peaks and their indices in the reconstruction sphere. The input of this algorithm is ψ (ODF) and the faces of a symmetric on the z-axis evenly distributed sphere (see fig. 2.6C).

We have constructed a sphere which is symmetric over the z-axis and the faces of the sphere are also ordered along the z-axis. The same sphere was used in [205] for GQI reconstructions. Every face contains the 3 indices which indicate at the 3 different points that create the triangle(face) of the sphere. The idea here is that we can travel from face to face and nullify all points on a face which are lower than the higher value of the face. At the end only local maxima will survive the procedure. The algorithm is presented in detail in alg. 2.

The sphere we use is of course discrete and therefore it adds some constraints on the angular resolution (worst case $\pm 4.96^\circ$) of the peaks found from the ODF. In addition the proposed Peak Finding algorithm can reduce slightly more the angular resolution. For example in Fig.2.6A+B we show that if point **a** was a local maxima then only points **b** could be alternative local maxima for **a** but none of the unlabeled points could be a second peak.

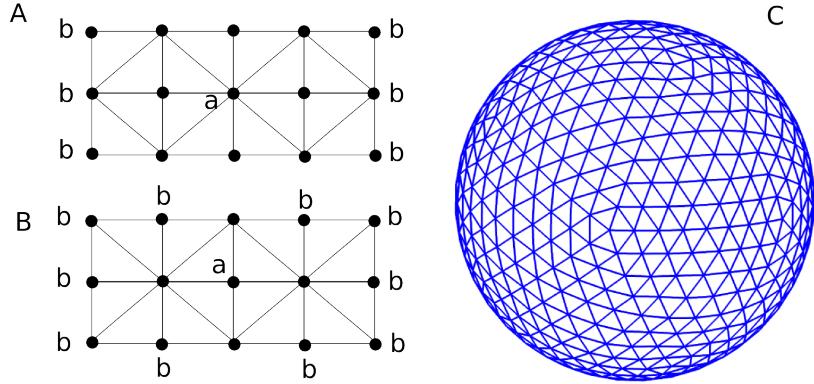


Figure 2.4: A,B: Imagine that point **a** is a local maximum for all its neighboring faces then only at **b** are other possible local maxima. This simple illustration shows that the triangulation of the sphere is important for the determination of closed peaks and that peaks which belong to the same triangle cannot be determined. C: the sphere used for ODF reconstructions consisting of 642 vertices and 1280 faces produced by subdivisions of the icosahedron.

Nevertheless we found alg. 2 to be extremely useful and fast. The same algorithm was used also in [205] but it was not documented as such.

2.7 Spherical Angular Smoothing

All current non-parametric dMRI reconstruction algorithms use some type of “smoothing” to reduce the effect of noise in the real data sets. DSI is using hanning filter and then avoiding sampling from low values in **r**-space. In GQI, smoothing is controlled from a scalar parameter; the diffusion sampling length and in spherical harmonic inversion methods [54],[1] the amount of smoothing is controlled by using only a number of the first components of the SH series.

All these approaches smooth and calculate the ODFs simultaneously. We propose something different. That the ODF can be calculated initially and then smoothed using for example the following operator shown below in matrix form

$$W = \exp\left(\frac{U \cdot U^T}{s}\right)$$

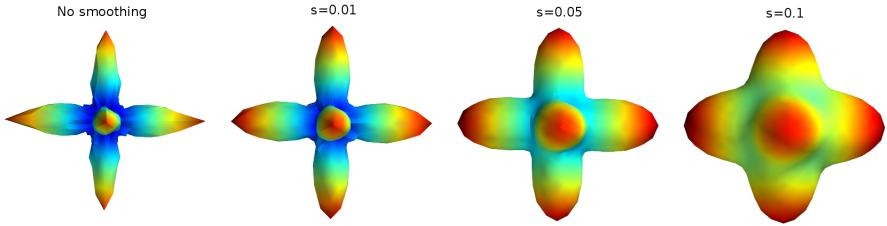


Figure 2.5: An example of spherical angular Gaussian smoothing applied with different smoothing factors on the distribution function of a triple-fibre crossing on the left. The simulation was used using (Sticks&Ball) model with diffusivity value 0.0015, S0=100, and Gaussian noise with SNR 20.

where U is the an $N \times 3$ matrix holding the N points of the ODF reconstruction sphere and s is a smoothing parameter acting like the variance. At the next step we can smooth any $\text{ODF}(\psi)$ creating a new $\text{ODF}(\psi')$ in the following way

$$\psi' = \psi \cdot \frac{W}{\sum_j W_j} \quad (31)$$

where j denotes row indexing, $\sum_j W_j$ acts as a normalization for the angular weighting W , ψ is the initial ODF and ψ' is the smoothed ODF . The advantage of this method is that it is more comprehensive and direct. Also it uses information from all directions simultaneously. Similar operators can be constructed that weight more lower or higher peaks. The operator shown here weighs more peaks that are closer in angular distance. In fig. 2.5 we see the effect of this equation on a simulated triple-fibre crossing; distorted with Gaussian noise with SNR 20 and reconstructed as a EITL density function which we have .

We can easily see in fig. 2.5 that when we increase the smoothing factor s small noisy peaks; as seen in the center of the unsmoothed spherical function can be easily removed. However, with too much smoothing even the longer peaks can lose their definition. This spherical operator can help to set the trade-off between noise and signal and it can also simplify the peak finding process i.e. finding the underlying primary fibre directions as this problem is much easier on smooth surfaces.

Finally, we believe that by decoupling the smoothing from the reconstruction phase we have an important advantage and that is that we can reduce more strongly and independently the effect of the noise to our data. Many spherical operators added as a plugins which is uncontrolled from the reconstruction phase and can work with any function on the sphere. For example Eq. 31 can be used with any function on the sphere.

2.8 Comparisons and Results

Validation of reconstruction and tractography algorithms is not straightforward due to the lack of relevant gold standards. Simulated voxels and digital phantoms is a useful way to overcome this difficulty and test new methods. After the simulation results we also show results with real human data sets.

2.8.1 Multi-fibre Simulations

For single voxel simulations we used the model proposed in Behrens et al.[20]; the multi-compartment model also known as Sticks and Ball which simulates the diffusion signal as

$$S_i = S_0 \left((1 - \sum_{j=1}^N f_j) \exp(-b_i d) + \sum_{j=1}^N f_j \exp(-b_i d \cos(\theta_{ij})^2) \right) \quad (32)$$

where θ_{ij} is the angle between gradient direction $\hat{\mathbf{g}}_i$ and fibre(stick) unit direction $\hat{\mathbf{u}}_j$. The amount of representation for every fibre is given by f and d is the diffusivity value for the entire model. A Multi Tensor [124] approach was also created for digital phantoms using the formula

$$S_i = S_0 \sum_{j=1}^N \exp(-b \hat{\mathbf{g}}^T D_j \hat{\mathbf{g}}) \quad (33)$$

where D_j is the diffusion tensor for every fibre j .

A comparison method/metric is needed in order to evaluate the new/old reconstruction methods discussed in this document. The standard proce-

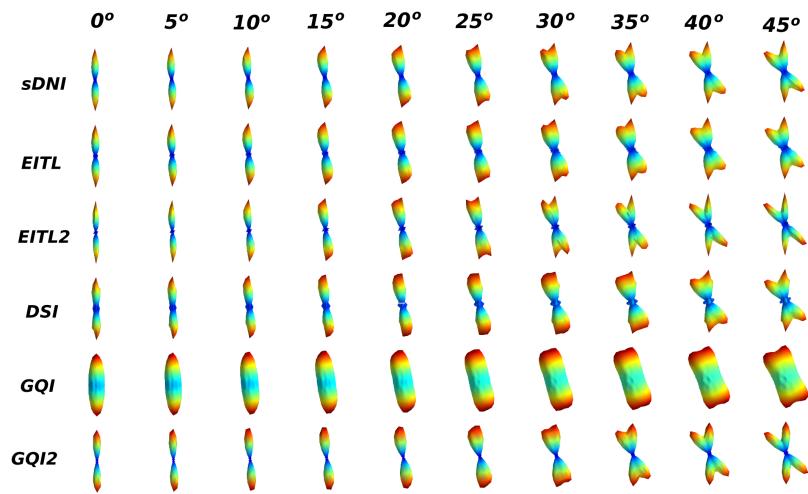


Figure 2.6: Visualizing ODFs created from different reconstruction methods sDNI (sEITL), EITL, EITL2, DSI, GQI, GQI2. These are based on simulations of 2-fibre crossings from 0° to 90° at steps of 5° using eq. 32. We can see that standard DNI (sDNI), EITL and EITL2 can resolve the correct angular fibre directions at lower angles than the other methods. For example see column at angle of 25° .

Known	Measured	AS
$(1, 0, 0), (0, 1, 0)$	$(0, 0, 1)$	0
$(1, 0, 0), (0, 1, 0)$	$(0, 1, 0)$	1
$(1, 0, 0), (0, 1, 0)$	$(0, \sqrt{2}/2, \sqrt{2}/2)$	$\sqrt{2}/2$
$(1, 0, 0), (0, 1, 0), (0, 0, 1)$	$(1, 0, 0), (0, 0, 1)$	2

Table 2.2: Examples of angular similarity (AS) behaviour with simple unit vector sets.

dure is to calculate the similarity between the measured and simulated “golden truth” data sets. We want to calculate the angular precision of the ODFs from simulations derived from eq. 32. We define a new similarity metric called angular similarity (AS) which computes the cosine distance of the best match between the set of measured fibre orientations and the known set of simulated fibres. This metric will be used to compare 2-fibre and 3-fibre crossings. AS is 0 when there is no match i.e. angular distance is maximum (90°), 1 when 1 fibre is matched (0°), 2 when 2 fibres are matched and 3 when 3 fibres are matched. In table 2.2 we show a few examples of AS behaviour with simple unit vector sets.

If our “golden” (known) set consists of $g = [(1, 0, 0), (0, 1, 0)]$ and the measured set consists of $m = [(0, 0, 1)]$ then AS=0. If the measured set was $m = [(0, \sqrt{2}/2, \sqrt{2}/2)]$ then AS is $\sqrt{2}/2$. This is because according to the AS definition we have $AS(g, m) = \max(|g[0] \cdot m[0]|, |g[1] \cdot m[0]|) = \sqrt{2}/2$. If $g = [(1, 0, 0), (0, 1, 0)]$ and $m = g$ then $AS(g, m) = \max(|g[0] \cdot m[0]| + |g[1] \cdot m[1]|, |g[0] \cdot m[1]| + |g[1] \cdot m[0]|) = 2$.

We created an experiment where we set two fibres at an increasing angle of 2.5° from 0° to 90° and then rotate them uniformly around 200 random axes. This operation produces 7400 simulated ODFs and the results are shown in fig. 2.7, 2.8 with different signal to noise ratio. For these simulations noise was normally distributed. What we see in the figures is the average angular similarity where the average is calculated from the 200 random orientations for the same angle.

We can easily observe in fig. 2.7 that EITL2 can resolve more accurately fibre crossings at low angles and continue performing decently well even at

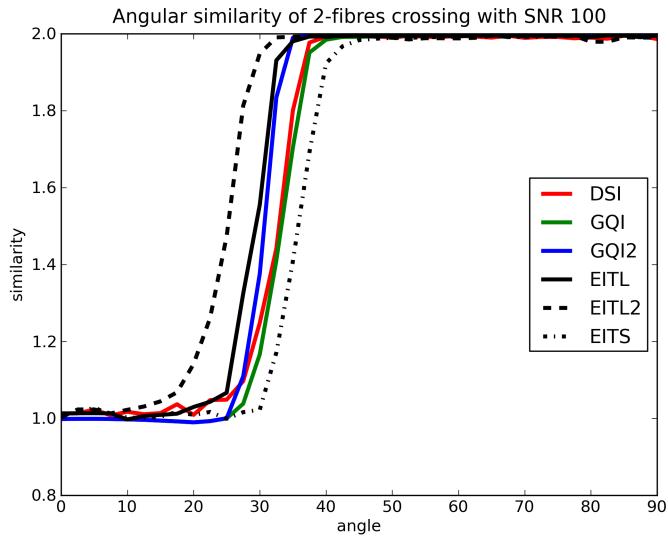


Figure 2.7: Average angular similarity of 2-fibre crossings with SNR 100

higher angles $> 50^\circ$. Then EITL is doing better than DSI, GQI, GQI2 and EITS at low angles and very well at high angles as well. GQI2 performs better than DSI, GQI, and ETS. It is also impressive that EITS can have such a good performance although it is such a simple operation. In summary we say from the that EITL2>EITL>GQI2>DSI>GQI>EITS where $>$ means better average angular similarity. The same pattern takes place even when we increase the noise level see for example fig. 2.8. We will see next that the same pattern will take place even with 3-fibre crossings and high levels of noise.

We also measured the accuracy in 3-fibre crossings. Therefore, we created an experiment where the 3-fibres would always have the same angular distance between each other. That distance would increase from 0° to 90° with steps of 2.3° on average and all 3 fibres would be reoriented 200 times. That gave 8000 simulated crossings.

The results of the 3-fibre crossings shown in fig. 2.9 and 2.10 were very similar with those of the 2-fibre crossings; EITL2 scores higher at low angles with a bit reduced performance at high angles and EITL doing

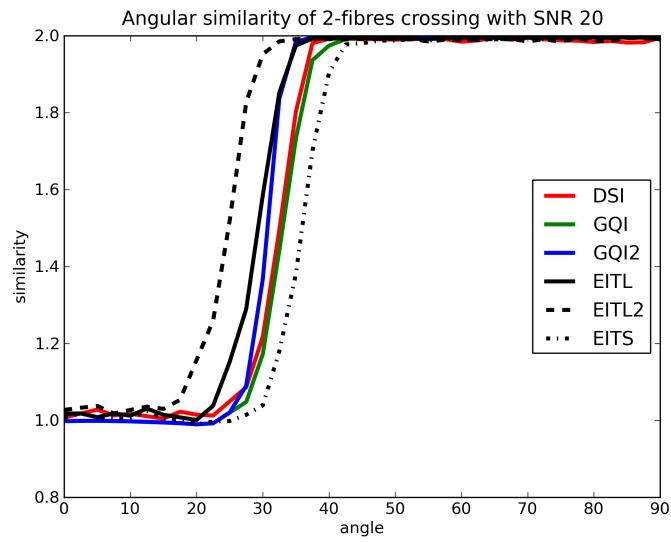


Figure 2.8: Average angular similarity of 2-fibre crossings with SNR 20.

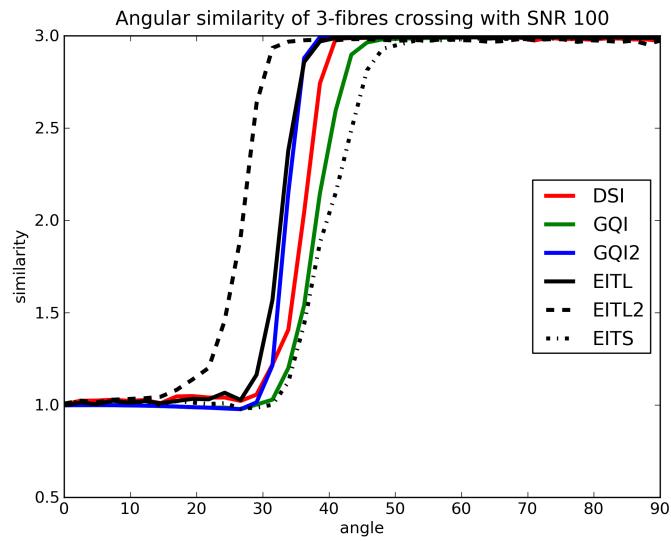


Figure 2.9: Average angular similarity of 3-fibre crossings with SNR 100.

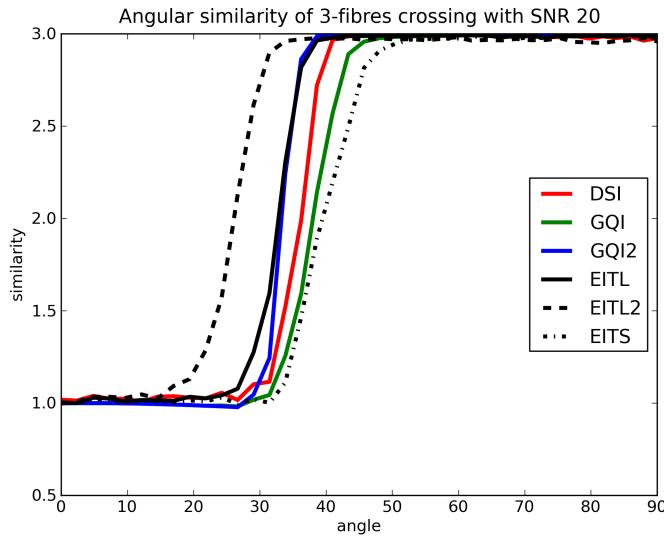


Figure 2.10: Average angular similarity of 3-fibre crossings with SNR 100.

better with low angles than the rest of the methods and also having high accuracy on larger angles.

These concentrative plots give strong evidence that both DNI (EITL) and in general EIT can be used to accurately generate spherical distribution functions for the determination of the directional information of the diffusion signal and that these can do better or similar to the current state-of-the-art grid-based reconstruction methods i.e DSI and GQI. Also the addition of noise didn't manage to affect the determination of the fibre directions considerably.

Furthermore, we can also see that GQI2 can do better than DSI, GQI and that EITS gives results that are very similar to GQI. The parameters used for these simulations were DSI: radial sampling 2.1 – 6, hanning filter width: 36 , GQI: $\lambda=1.2$, GQI2: $\lambda = 3$, and EITS, EITL, EITL2 were all calculated with the standard options zonal width ($z = 5^\circ$), grid size $17 \times 17 \times 17$, radial sampling 0 – 5 with 0.1 steps and no further post-processing or smoothing was used. All methods were using the same reconstruction sphere with 642 vertices and 1280 faces.

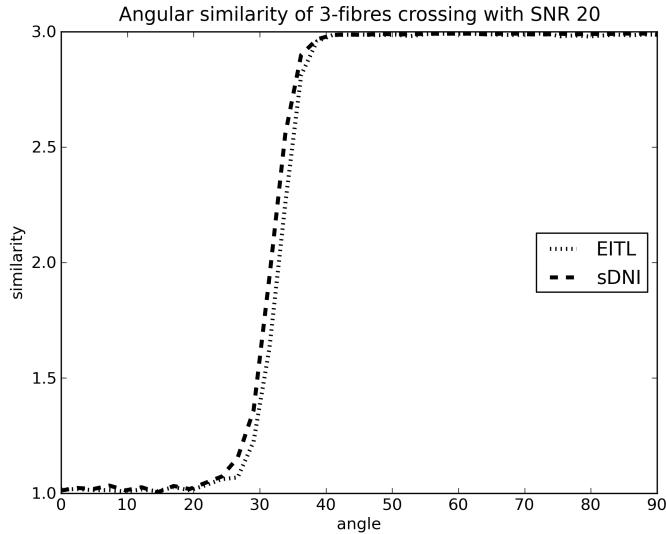


Figure 2.11: This diagram shows that when we compute EITL with the fast or standard method the results are nearly equivalent. We show here that the mean angular similarity for the case of 3-fibres crossings is very similar when using standard DNI or or fast DNI (EITL).

Finally, we should stress that we have never seen any considerable differences between spherical functions created using the standard or fast EIT. For example a simple test for the 3-fibre case as seen in fig. 2.8.1 can show that there is close agreement between the two methods i.e. their results are nearly equivalent. Therefore we can conclude that the fast EIT is great approximation of the standard EIT.

2.8.2 Digital Phantoms

A digital phantom generation tool was developed which can simulate the diffusion weighted signal for one or more fibres represented by different discrete 3D orbital functions. This work is an extension of the phantom developed by Correia et al. [41] who supported only semi-circular functions with analytically calculated derivatives.

The idea here is that we first create any orbital function $f(t) : \mathbb{R} \rightarrow \mathbb{R}^3$ and calculate numerically its derivatives at small steps Δt . Then we can scale it and centre it so that it fits in an image volume of size that we desire.

We expect that many segments of the discrete function f will fall into every voxel in the volume and that more curved parts of f will have higher representation in the voxel than less curved parts. For every segment we can find the main direction of the orbit $\mathbf{v} = \frac{f(t+1) - f(t)}{\Delta t}$ and calculate the rotation matrix \mathbf{R} that rotates $\hat{\mathbf{x}} = (1, 0, 0)$ to \mathbf{v} . Then the signal for each element of the fibre for a given b-value b and a given gradient sampling direction $\hat{\mathbf{g}}$, is given by the following Single Tensor formula

$$\Delta S = S_0 \exp(-b\hat{\mathbf{g}}^T \mathbf{R} \Lambda \mathbf{R}^T \hat{\mathbf{g}}) \quad (34)$$

where S_0 is the unattenuated signal of the fibre, and the diffusion tensor is given by

$$\Lambda = \begin{pmatrix} \lambda_{||} & 0 & 0 \\ 0 & \lambda_{\perp} & 0 \\ 0 & 0 & \lambda_{\perp} \end{pmatrix} \quad (35)$$

Therefore the total signal of the voxel for one gradient direction is given by the summations of all the contributions of the K elements in the voxel

$$S_{vox} = \sum_{i=1}^K \Delta S_i \quad (36)$$

In addition, we can generate simulations of more than one fibres by generating a single volume for every orbit and then add them all together to create complex configurations in the final volume. This is acceptable because the diffusion signal is additive i.e. the signal of a crossing of two fibres is equal to the sum of the signals of the individual fibres. In that way we can simulate phantoms with Multi Tensor based diffusion signals as that described in eq.33. We can increase the thickness of the fibres using a typical smoothing kernel or duplicate the fibres radially. At the end we can add different levels of noise e.g. Rician or Gaussian noise with a prespecified SNR.

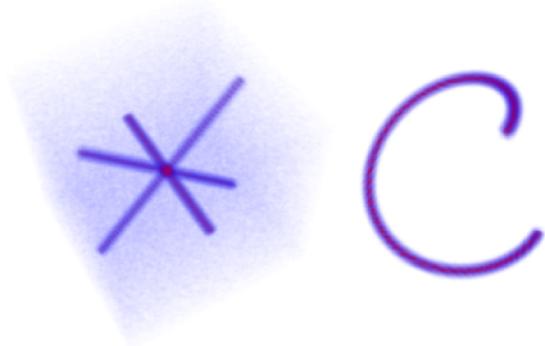


Figure 2.12: Volume renderings of the unattenuated signals of two digital phantoms. On the left 3 fibres intersect on regular angles with Rician noise of SNR=20. On the right a helicoidal fibre is shown clear of noise. For both phantoms $S_0 = 100$ and prolate tensors with eigenvalues $\lambda_{\parallel} = 1.4 \cdot 10^{-3} m^2/sec$ and $\lambda_{\perp} = .35 \cdot 10^{-3} m^2/sec$ were used.

The method we use to create these digital phantoms offers the opportunity to simulate partial volume effects. If partial volume effects are not desired then we need to normalize with the number of fibre elements for each voxel (This function is implemented in `dipy.sims.phantom.orbital_phantom`). In fig. 2.12 we can see the volume renderings of two different phantoms created with the method described here.

2.8.3 Results with digital phantoms

With the purpose of comparing and visualizing the differences between the reconstruction methods described in this document we created a digital phantom of two crossing bundles. The bundles are crossing at an angle of 90° . The digital phantom was generated using the method described in the previous section. Here we describe the basic steps: (a) We first represented the first bundle as a discrete straight orbit starting from point $(-1, -1, 0)$ and ending at point $(1, 1, 0)$ with using 1000 time steps. (b) We scaled, centred and radially expanded this orbit so that it fits a volume of size $64 \times 64 \times 64$. This volume corresponds to the diffusion volume without any weighting. (c) We then applied the weightings for all the following volumes corresponding to non-zero b-values. (d) We replicated the same pro-

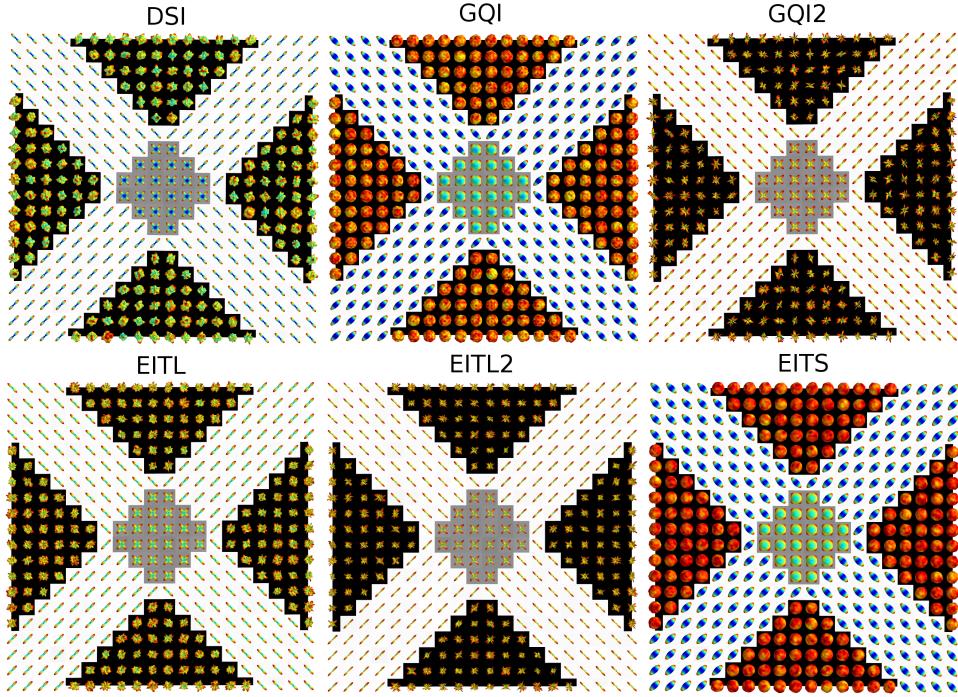


Figure 2.13: Results with an 'x' shape digital phantom. Every single tensor compartment had the following eigenvalues $\lambda_{\parallel} = 1.4 \cdot 10^{-3} m^2/sec$ and $\lambda_{\perp} = .1 \cdot 10^{-3} m^2/sec$. Rician noise was added with SNR = 5. In this figure we can easily perceive that GQI is very similar to EITS, GQI2 is very similar to EITL and DSI is very similar to EITL. In fig. 2.14 the regions at the centers of the phantoms are depicted in higher resolution.

cedure for the other bundle which initially started as an orbit from position $(-1, 1, 0)$ and ended at position $(1, -1, 0)$. (e) We added the two volumes together to create an 'x' shape (see fig.2.13,2.15). (f) We added Rician noise with SNR=5. As in this document we concentrate on Cartesian Lattice Q-space acquisitions we generated b-vectors and b-values by using a keyhole Cartesian sampling grid [190] with 515 q-vectors. The maximum b-value was 11538 and the minimum was 0. Two sets of simulation experiments were performed each using a different type of Tensor.

In the first experiment shown in fig. 2.13,2.14 we used a more anisotropic prolate tensor for the simulation with eigenvalues $\lambda_{\parallel} = 1.4 \cdot 10^{-3} m^2/sec$ and $\lambda_{\perp} = .1 \cdot 10^{-3} m^2/sec$. In the second experiment shown in fig. 2.15,2.16 we used a much less anisotropic prolate tensor with $\lambda_{\parallel} = 1.7 \cdot 10^{-3} m^2/sec$

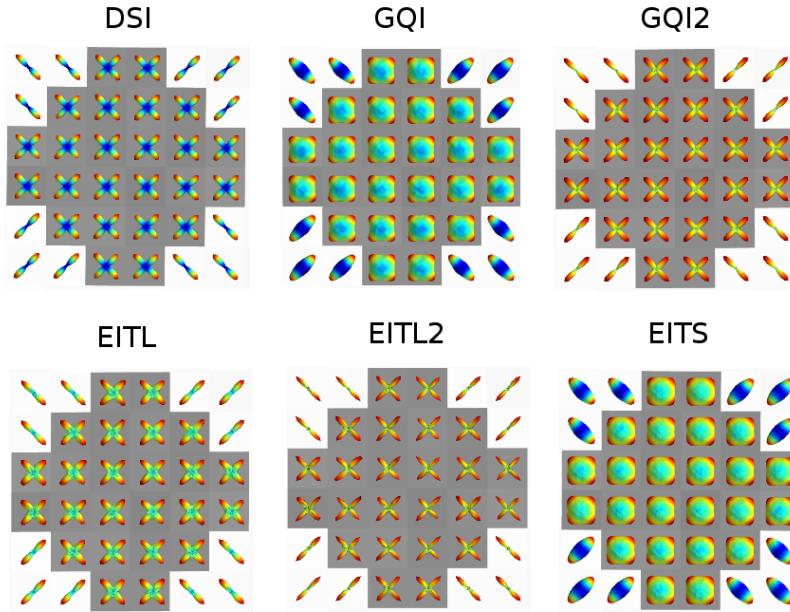


Figure 2.14: Same as previous fig. 2.13 showing the spherical distributions in the centers of the phantoms at higher resolution.

and $\lambda_{\perp} = .3 \cdot 10^{-3} m^2/sec$. It is well known that noise affects more less anisotropic areas. We can see this effect by comparing the overlapped FAs of these two figures (2.13, 2.15). However we can also see that all six methods (DSI, GQI, GQI2, EITL, EITL2, EITS) can resolve correctly the fiber directions by looking at their spherical distribution functions colourmapped with a standard 'jet' colourmap. For visualization purposes all ODFs are shown in relative size as they have been scaled so that their maximum values correspond to 1.

Furthermore, we can easily see that GQI is mostly similar with EITS, GQI2 is very similar with EITL and DSI is mostly similar with EITL. That DSI ODFs are very similar with EITL ODFs is expected as the two methods create theoretically the same real ODFs. Remarkably, EITL can create these ODFs without using the Fourier Transform neither using any filter or thresholds in r-space which are necessary in DSI.

Fig. 2.13, 2.15 shows that all these different grid-based reconstruction methods can reconstruct correctly the underlying fibre directions even when noise is present. However we can see that when tensors are less anisotropic

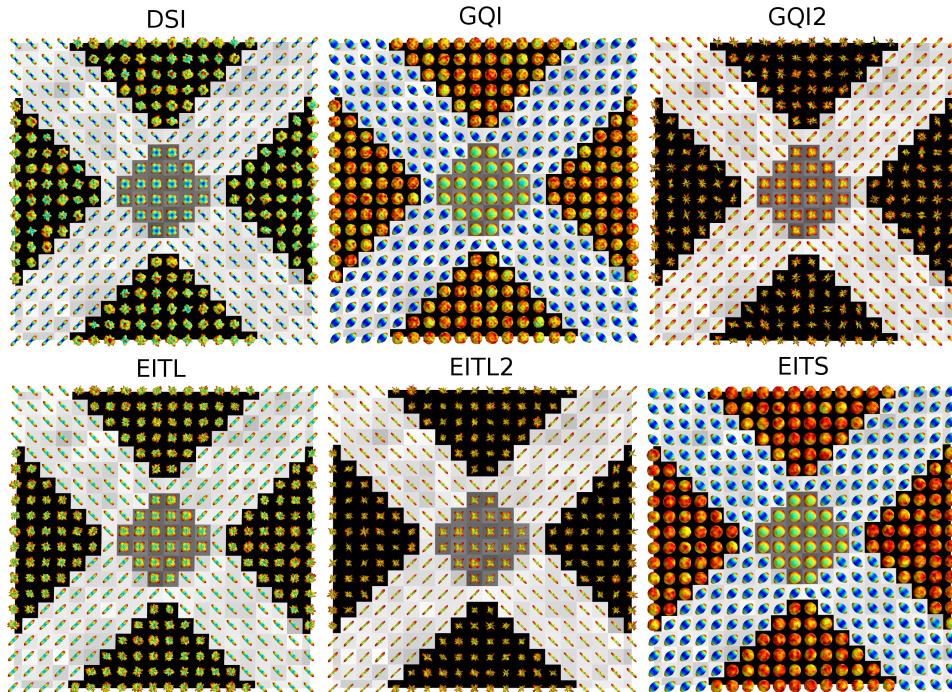


Figure 2.15: Showing the spherical distribution functions (DSI, GQI, GQI2, EITL, EITL2, EITS) of a digital phantom generated by two bundles where each bundle contains single tensors along the direction of the phantom. On the crossing area we see a dual tensor effect in every voxel. Every single tensor compartment had the following eigenvalues $\lambda_{\parallel} = 1.7 \cdot 10^{-3} m^2/sec$ and $\lambda_{\perp} = .3 \cdot 10^{-3} m^2/sec$. Rician noise was added with SNR=5. We also visualize simultaneously the FA for this slice. We can see that in the crossing area (gray background) the FA values drop considerably however the ODFs represent precisely the crossing.

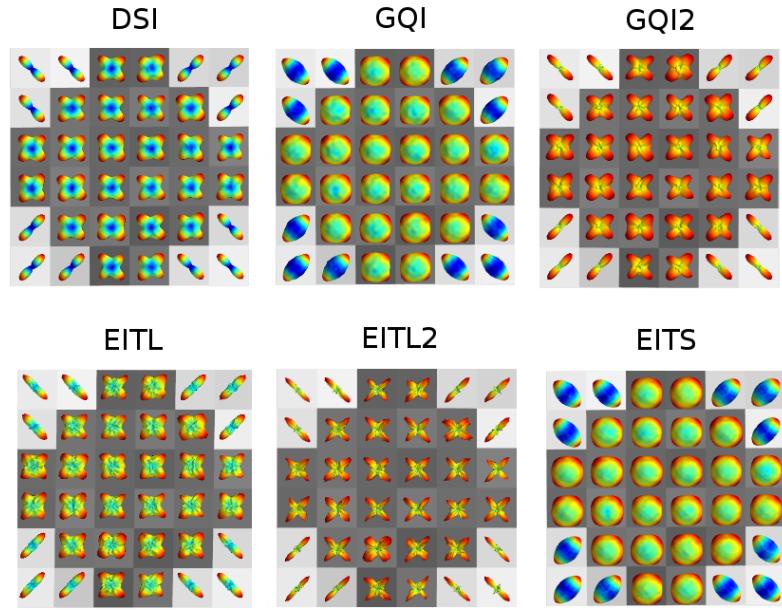


Figure 2.16: A zoomed version of previous fig. 2.15 showing the spherical distributions in the centers of the phantoms at higher resolution.

the noise has a stronger effect in the resulting spherical distributions. We can also see that GQI & ETS are less sharp than DSI & EITL and these are less sharp than GQI2 & EITL2. Also DSI, GQI2, EITL, EITL2 have much lower minima than GQI and EITS.

In the EIT-based reconstruction results shown in fig. 2.14 and 2.16 we do not use any amount of smoothing as used in DSI (through hanning filter), GQI, GQI2 (through sampling length) and it is extraordinary that we obtain so well defined distributions. If we want to apply some weighting/smoothing/denoising in EIT-based methods that is simply possible through the spherical angular smoothing approach described in section 2.7.

The parameters used for these simulations were for DSI: radial sampling 2.1 – 6, hanning filter width: 36 , GQI: $\lambda=1.2$, GQI2: $\lambda = 3$, and EITS, EITL, EITL2 were all calculated with the standard options ($z = \pm 5$) and no further post-processing or smoothing was used. All methods were using the same reconstruction sphere with 642 vertices and 1280 faces.

2.8.4 Results with real data sets

Apparently we want to compare reconstruction methods on Cartesian grid-based acquisitions first with data sets which are rich on directions and commonly used for DSI processing. For this purpose we used a data set which was available online at cmtk.org from the Diffusion Group at Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. So, this data set was brought forth by a 3T scanner (TIM Trio, Siemens) with a 32 channels head coil. The field of view was $210 \times 210 \text{ mm}^2$, matrix size 96×96 , and slice thickness 3 mm . 44 slices were acquired and the voxel resolution was $2.2 \times 2.2 \times 3.0 \text{ mm}^3$. A 258-point half grid acquisition scheme with a maximum b-value of 8011 s/mm^2 also known as DSI515[199] was used. The total acquisition time was 34 min with TR= 8200 ms and TE= 165 ms .

The parameters used for these simulations were for DSI: radial sampling 2.1 – 6, hanning filter width: 36 , GQI: $\lambda=1.2$, GQI2: $\lambda = 3$, and for EITS, EITL, EITL2 were all calculated with the standard options for zonal width ($z = 5^\circ$) and spherical angular smoothing ($s = 0.05$). All methods were using the same reconstruction sphere with 642 vertices and 1280 faces. The results of this experiment are shown on top of an FA slice of a healthy human in fig. 2.17 and in higher resolution in fig.2.18. It is easily observed that EITL, EITL2 and EITS can be used for reconstructing these data sets as their results are clearly as good as the results given by DSI, GQI and GQI2. We can also easily see that EITL and EITL2 are relatively sharp which can be of an advantage for the purpose of recovering correctly the underlying real fibre directions.

We also tested our results with another human brain data set generated at a 3T scanner (TIM Trio, Siemens) in Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK. We used Siemens advanced diffusion work-in-progress sequence, and STEAM [140, 135] as the diffusion preparation method. The field of view was $240 \times 240 \text{ mm}^2$, matrix size 96×96 , and slice thickness 2.5 mm (no gap). 55 slices were acquired to achieve full brain coverage, and the voxel resolution was $2.5 \times 2.5 \times 2.5 \text{ mm}^3$. In this experiment much less gradient vectors were used. A 102-point half grid acquisition with a maximum b-value of 4000 s/mm^2 was used. The total acquisition

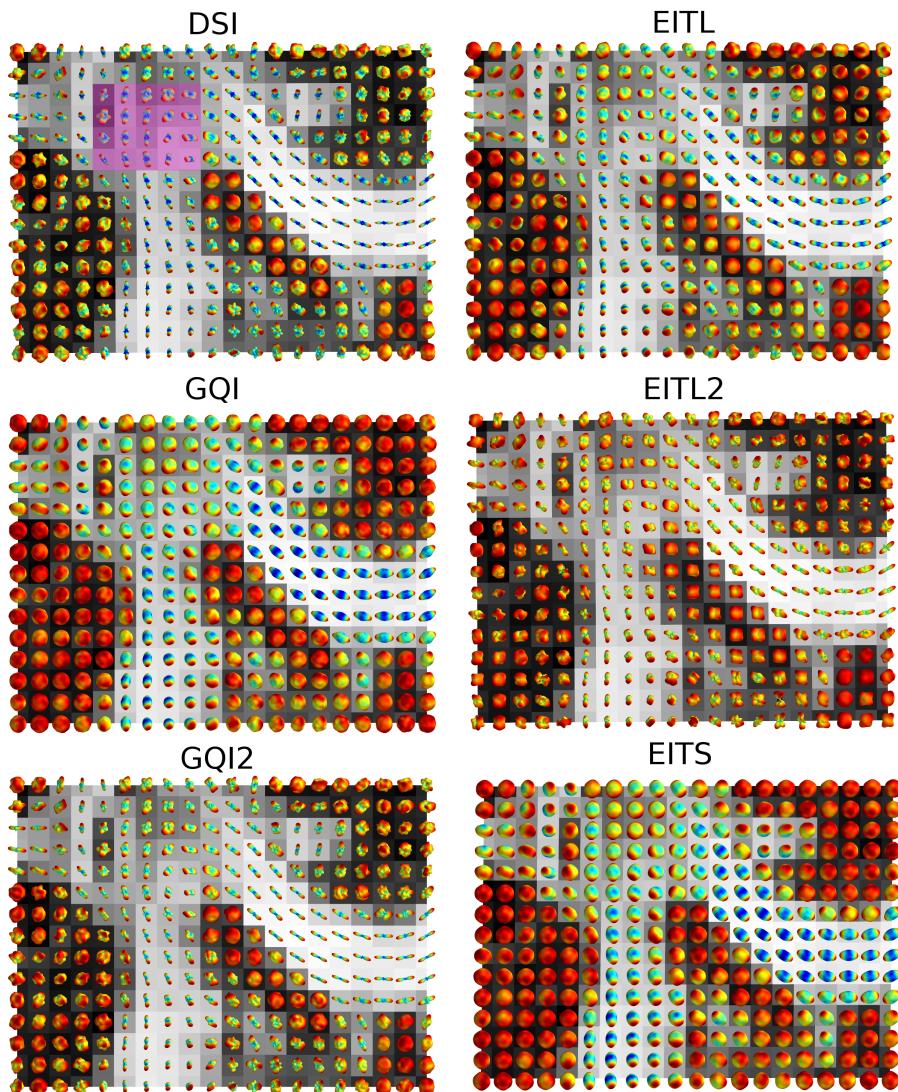


Figure 2.17: Showing the same slice of a human brain reconstructed with 6 different Cartesian grid q-space based methods. We see here that the ODFs are visualized on top of the FA slice. A clearer presentation of a region near the left upper corner (with purple shading) is given in fig.2.18 for all the 6 methods.

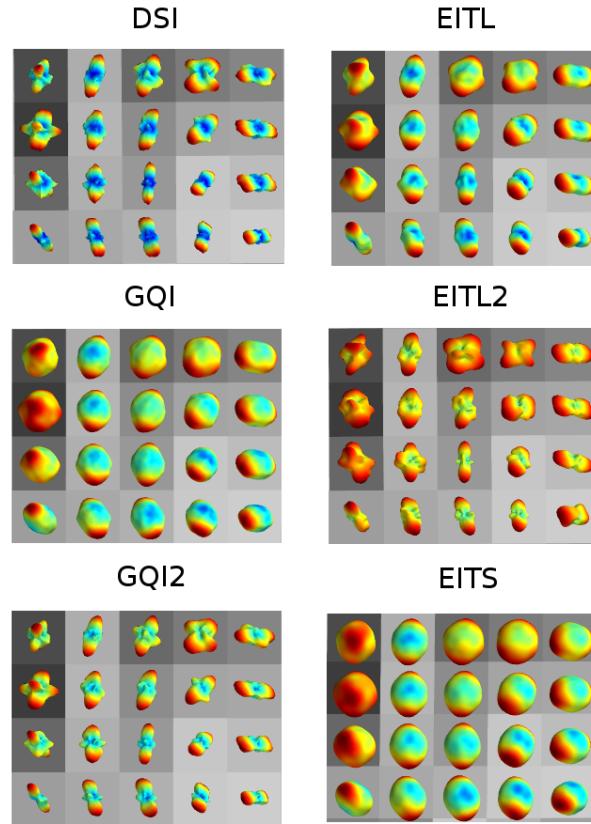


Figure 2.18: The region shown at the upper-left corner (purple shading area) in the panels of fig.2.17 is shown here at higher resolution. The data sets here are from a real human. In contrast with the results shown in simulations 2.14 we applied spherical angular smoothing with $s = 0.05$ for EITL, EITL2 and EITS in order to remove small noisy spikes in the distributions. In addition, similarly with the results of fig.2.14 EITS is very similar with GQI but this time the difference between DSI, GQI2 and EITL, EITL2 is smaller because of the application of angular weighting.

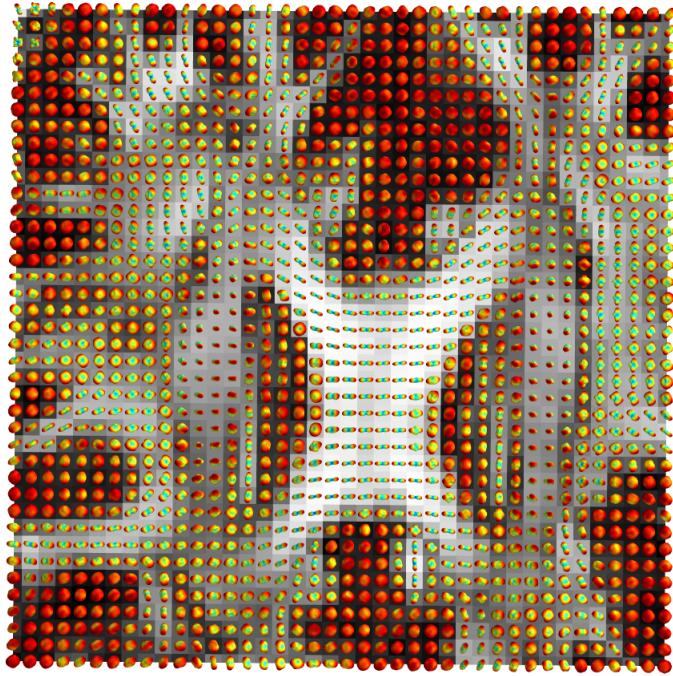


Figure 2.19: EITL ODFs rendered on top of FA of a human brain data set. A small 102-point half grid acquisition with a maximum b-value of 4000 s/mm^2 was used. Figures 2.20, 2.21 are zoomed versions of the same figure. We can see clearly single fibres on the CC and CST areas but also crossing fibres at the Centrum Semiovale and at the areas where big bundles cross. Also the non-white matter areas are evidently more isotropic.

time was only 14 min 21s with TR=8200ms and TE=69ms.

In fig. 2.19 a slice is shown where different parts of white matter are visible with the FA background image. We can clearly see structures like the CC and CST and Centrum Semiovale areas. On top of the FA the ODFs of EITL are shown. The parameters used for EITL were: a standard zonal width $z = 5^\circ$ and spherical angular smoothing $s = 0.05$ with the same reconstruction sphere (642 vertices, 1280 faces) as before.

For illustration purposes the upper part of fig. 2.19 is depicted again in fig. 2.20, and the region with purple shading from fig. 2.20 is given at an even higher resolution in fig. 2.21. We used Mayavi [167] a Python visualization library based on VTK to make the visualizations shown in the figures of this section.

Although much less directions were used in this acquisition scheme we

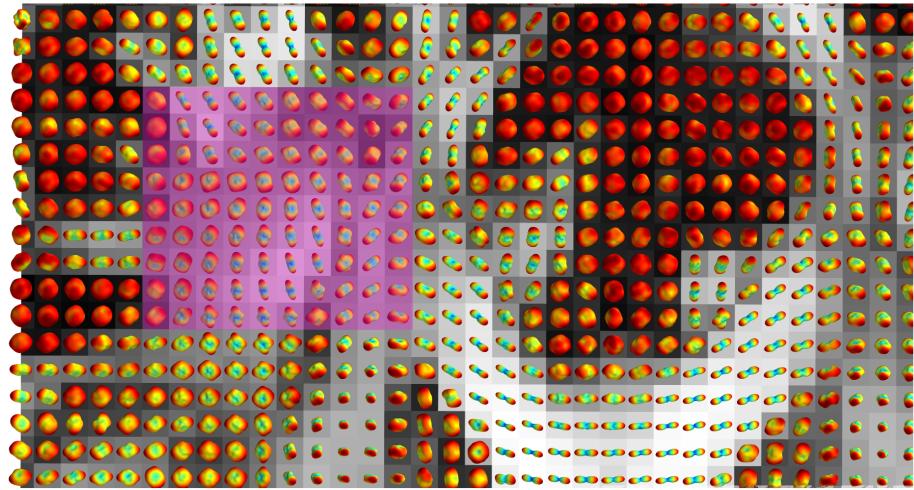


Figure 2.20: The upper part of fig. 2.19 is shown here at higher resolution. The purple shaded part is given in higher resolution in fig. 2.21

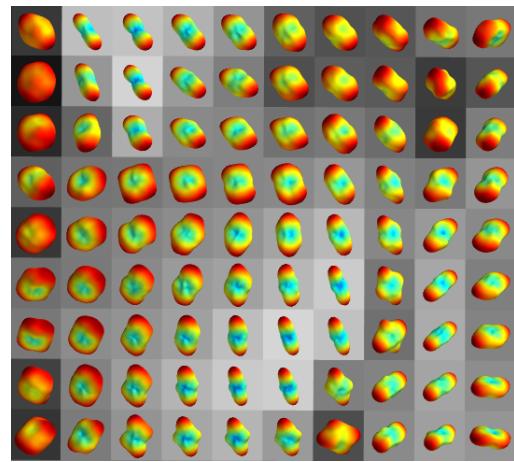


Figure 2.21: EITL ODFs of 1-fibre, 2-fibre and 3-fibre crossings from a real human data set of 101 applied weighted diffusion volume and 1 without weighting (b_0). This picture is a zoomed version of the purple shaded area shown in fig. 2.20.

obtain a similarly accurate depiction of the underlying white matter structure in comparison with that of 258 directions. This gives great hope that we can use grid-based reconstruction methods with half-grid sequences with 100 gradient directions. This was also showed by [114] and [205] who used similar number of directions.

We can see in all the figures with real data sets single fibres as those usually found at the center of CC, and 2 or 3-fibre crossings in the intersection areas of CC with the CST and other bundles.

We can see for example in fig. 2.21 that the effect of spherical angular smoothing can help alleviate the noise effects and focus our concentration on depicting the major directions which are also of highest concern.

2.9 Anisotropy metrics

Until this moment we discussed about density functions on the sphere as a way to represent complex fibre directionality in the voxel. These density functions are represented as multidimensional vectors containing 200 or more dimensions in each voxel and it can be cumbersome to use them directly for subject comparisons or visualization purposes. For this purpose most people use simple scalar summarizing metrics e.g. Tensor-based FA, MD or ODF-based like the Generalized FA (GFA) [188]. In this section we will show that a similar scalar function like FA can be constructed non-parametrically. We call this NPA which stands for non-parametric anisotropy. We will also start experimenting with metrics that have more than one scalar values and can represent more accurately the directionality in each voxel that is lost with FA, GFA and MD. We will investigate and explain here the realms and robustness of Quantitative Anisotropy which was first introduced by Yeh et al. [205].

2.9.1 Non-parametric Anisotropy

Local voxelwise measures such as fractional anisotropy (FA), apparent diffusivity coefficient (ADC), or mean diffusivity (MD) [1,2] have been extensively adopted in clinical and applied research practice based on diffusion

weighted MR imaging (dMRI). This underlines the need for valid and reliable measures which can indicate the degree of local organisation of white matter in the brain. The measures listed above are based on the parametric simple diffusion tensor (SDT) model [14] which works well when there is a single dominant fibre direction but is also known not to give valid information if the local organisation is more complex [205, 190]. We show how model-free, alternatives can yield non-parametric anisotropy (NPA). These are constructed from the GQI ODF. We apply exact analytical results which show the form of the GQI-ODF when the single tensor model is correct, and further indicate how the tensor's parameters may be estimated from this model-free approach. We compare the performance of these parametric and non-parametric measures for simulated data.

Simulations were computed for a 102-point grid sampling scheme, with a maximum b-value of 4000 s/mm^2 . The simulated fibre was aligned with the gradient frame of reference, and the diagonal elements of the diffusion tensor, D , were chosen to match typical values for white matter: $\lambda_1 = 1.4 \times 10^{-3} \text{ mm}^2/\text{s}$, and $\lambda_2 = \lambda_3 = 0.35 \times 10^{-3} \text{ mm}^2/\text{s}$. Variable fibre orientation was realised by spatially rotating the simulated fibres at discrete orientations. 100 orientations were used, which spanned uniformly the space of (θ, ϕ) .

In addition to the SDT a two compartment model with an isotropic component was added with volume fraction 0.5 and diffusivity 0.7×10^{-3} . For each acquisition scheme and fibre type, the “ideal” (noise-free) diffusion weighted signals were calculated according to the SDT model, assuming a constant ideal value of the baseline signal $S_0 = 100$. Complex Gaussian noise was then superimposed upon the ideal signals to provide the complex noise-contaminated signals and their magnitude was then obtained. This results in noisy values with a Rician distribution, which can be scaled in order to set the signal to noise ratio to any desired level. In this study the SNRs were 20,40,60,80 and 100. The GQI ODF and SDT were fitted using dipy (dipy.org).

The GQI ODF was calculated for a tessellated spherical icosahedron with 362 vertices and 720 faces. Two values (1.2 and 3.5) were used for λ , the diffusion sampling length. Non-parametric FA, NPA, was calculated from

the ODF by

1. locating the vertex V_1 with maximum GQI ODF value max_1 ;
2. with V_1 as pole, locating the vertex V_2 on the corresponding equatorial band of width ± 5 degrees with maximum GQI ODF value max_2 ;
3. locating a vertex V_3 in the equatorial band at approximately 90 degrees away from V_2 , denoting the GQI ODF value of max_3 at V_3 .
4. With $npd_1 = max_1^2$, $npd_2 = max_2^2$, and $npd_3 = max_3^2$, non-parametric anisotropy (NPA) was calculated by applying the classical FA [41] formula to the 3 values (npd_1, npd_2, npd_3).

The rationale for the squared ODF values is based on Tuch's formula eq.57 for ODF in the SDT case which implies that the ODF in the 3 principal axes directions of the tensor is proportional to the square root of the corresponding eigenvalue of the tensor. We have further derived an exact formula $max_j \propto \sqrt{\lambda_j}[\Phi(cL_\Delta/\sqrt{\lambda_j}) - .5]$ where c is a constant that depends on the acquisition parameters, and Φ is the cumulative distribution function of the standard Gaussian distribution. NEEDS PROOF?

The average NPA and FA are presented below for 200 simulations for each noise level, and single fibres with or without an isotropic component and with different diffusion sampling length. We can see that NPA gives very similar results with FA and as expected it is modulated by the degree of smoothing controlled by the value of the diffusion sampling length.

We plan to extend this approach with voxels containing multiple peaks where FA would be unable to give an informative result and also extend it to other types of ODFs. In summary, we have shown that an informative new scalar anisotropy function (NPA) can be calculated without fitting just from the GQI ODF which promises to be a model-free proxy for FA. NPA differs from GFA [191] in that it uses just 3 values of the GQI ODF with a geometric relationship instead of the entire ODF.

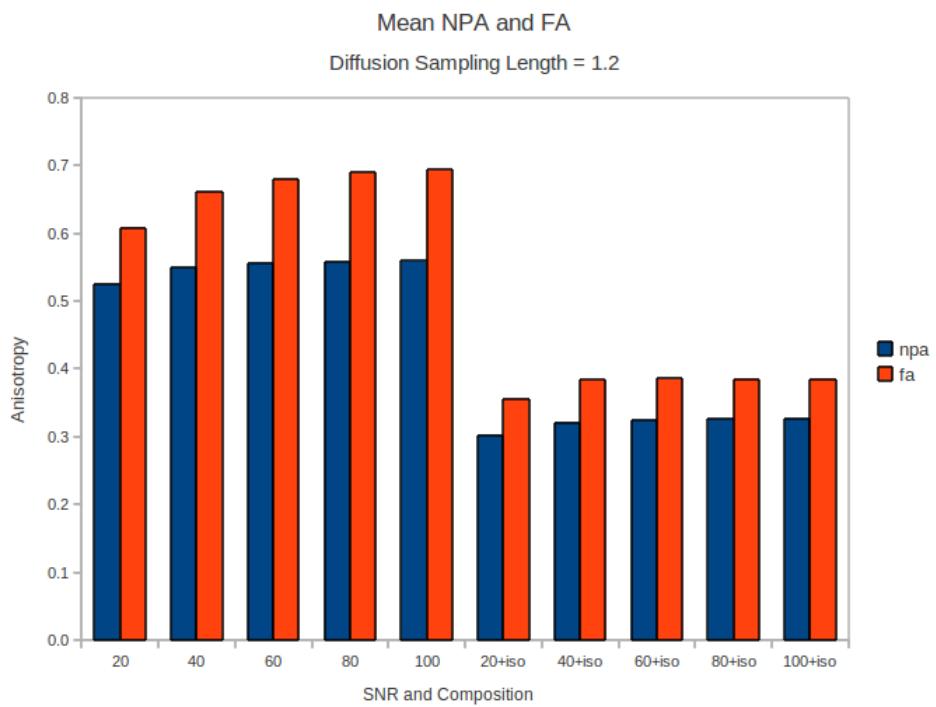


Figure 2.22: Comparison of NPA with FA for single fiber with and without an isotropic compartment at a range of signal to noise ratios.

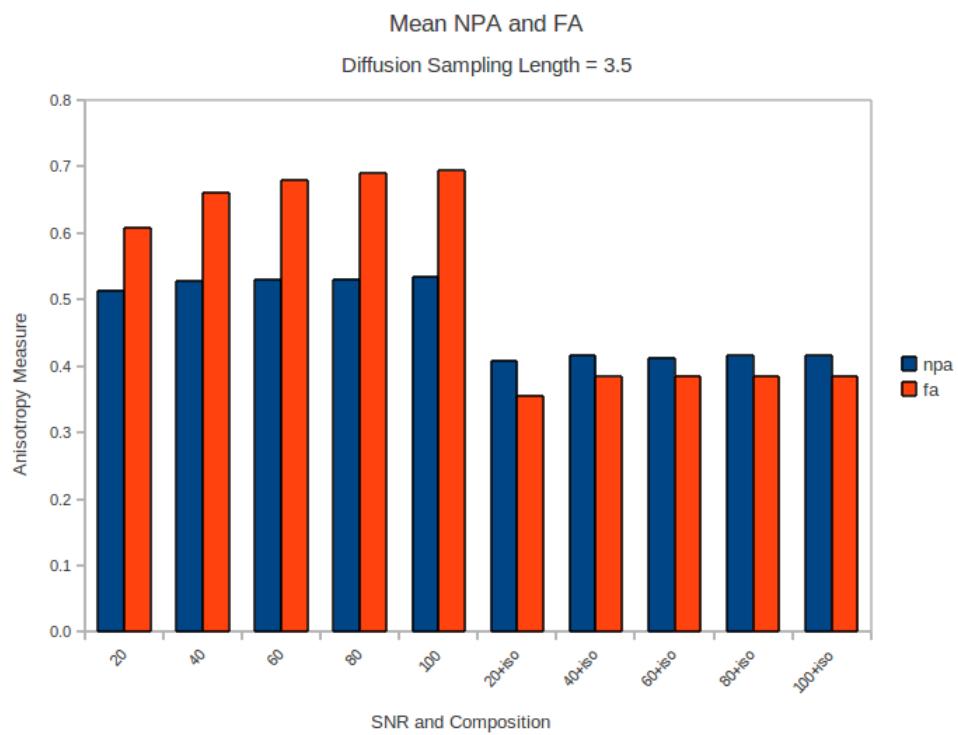


Figure 2.23: As in fig. 2.22 but with higher diffusion sampling length - less smoothing.

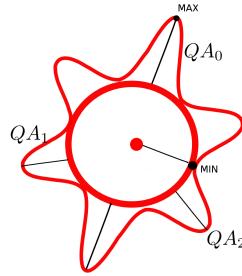


Figure 2.24: This figure shows how QA is calculated from an ODF. The sphere represents the “isotropic” component (minimum value) of a GQI ODF (star) which will be removed from the calculation of QA. QA acts like a differential component with higher values in anisotropic areas and lower in isotropic. Its big advantage over FA is that it can represent crossings.

2.9.2 Quantitative Anisotropy

Quantitative anisotropy (QA) was first used by [205] as a way to represent the peaks of the ODF with as few values as possible. This works in the following way: a) we create the ODF, b) we find the peaks using alg. 2, c) then QA_i is equal to the peak i minus the minimum value for the entire ODF. This is illustrated in fig. 2.24 where we can see a star-shaped ODF with three peaks (symmetric) (PK). This ODF can be represented just with 3 QA values where for example the highest value will be $QA_0 = \max(\psi_{GQI}) - \min(\psi_{GQI})$ where $\max(\psi_{GQI})$ is the value of the first peak PK_0 , $QA_1 = PK_1 - \min(\psi_{GQI})$

$$\text{and } QA_2 = PK_2 - \min(\psi_{GQI}) \text{ with } PK_0 \geq PK_1 \geq PK_2 .$$

QA acts like a differential operator which is higher on anisotropic ODFs and lower on more isotropic. Actually for a purely isotropic ODF $QA = 0$. QA can be also easily normalized by the maximum ODF value of all voxels which is usually at the CSF where there is a great amount of water. If this normalization is in effect then we can very easily remove the background noise i.e. non-white matter areas, scalp, skin, muscles etc. just because these will have very low QA values. We can see this interesting property of QA in fig. 2.25. Of course the most important property of QA is that it can

resolve crossings and assign a weight for every peak. We will make great use of these weightings in Chapter 3. for the creation of tractographies. Fig. 2.25 was created using DSI Studio dsi-studio.labsolver.org and the sequence parametrization is the same with the one provided in the experiments of the next section.

2.9.3 Robustness of QA

GQI was shown to have comparable accuracy to other well established q-space methods when it comes to resolving crossing fibres. In addition, this is achievable with as little as 102 points on a grid sampling scheme, bringing the total acquisition time down to a clinically acceptable level. Another advantage of GQI is that it is also applicable to a shell sampling scheme. Despite their successes in tractography applications, q-space techniques have until now failed to produce scalar metrics that could replace the ones derived from the diffusion tensor model (e.g. mean diffusivity, MD, and fractional anisotropy, FA) in terms of their multi-subject comparability and specificity to pathology. The data acquired with a grid sampling scheme can still be used to estimate a diffusion tensor and respective scalar parameters, but the effects of the high b-values required for q-space imaging ($> 2000 \text{ s/mm}^2$) in the accuracy of the resulting DTI-based (Single Diffusion Tensor SDT) parameters has not been well characterized. The authors of GQI have also proposed a new scalar metric called quantitative anisotropy (QA) which was described in the previous sections, but its properties have not been compared to FA's. In this study we will compare the estimated values of MD, FA and QA_0 (first component of QA) obtained with grid and shell sampling schemes, in terms of their precision and ability to differentiate between different brain fibre populations.

Twelve healthy volunteers aged between 18 and 40 were scanned on a 3T scanner (TIM Trio, Siemens), using Siemens advanced diffusion work-in-progress sequence, and STEAM [140, 135] as the diffusion preparation method. The field of view was $240 \times 240 \text{ mm}^2$, matrix size 96×96 , and slice thickness 2.5 mm (no gap). 55 slices were acquired to achieve full brain coverage, and the voxel resolution was $2.5 \times 2.5 \times 2.5 \text{ mm}^3$. Two sampling

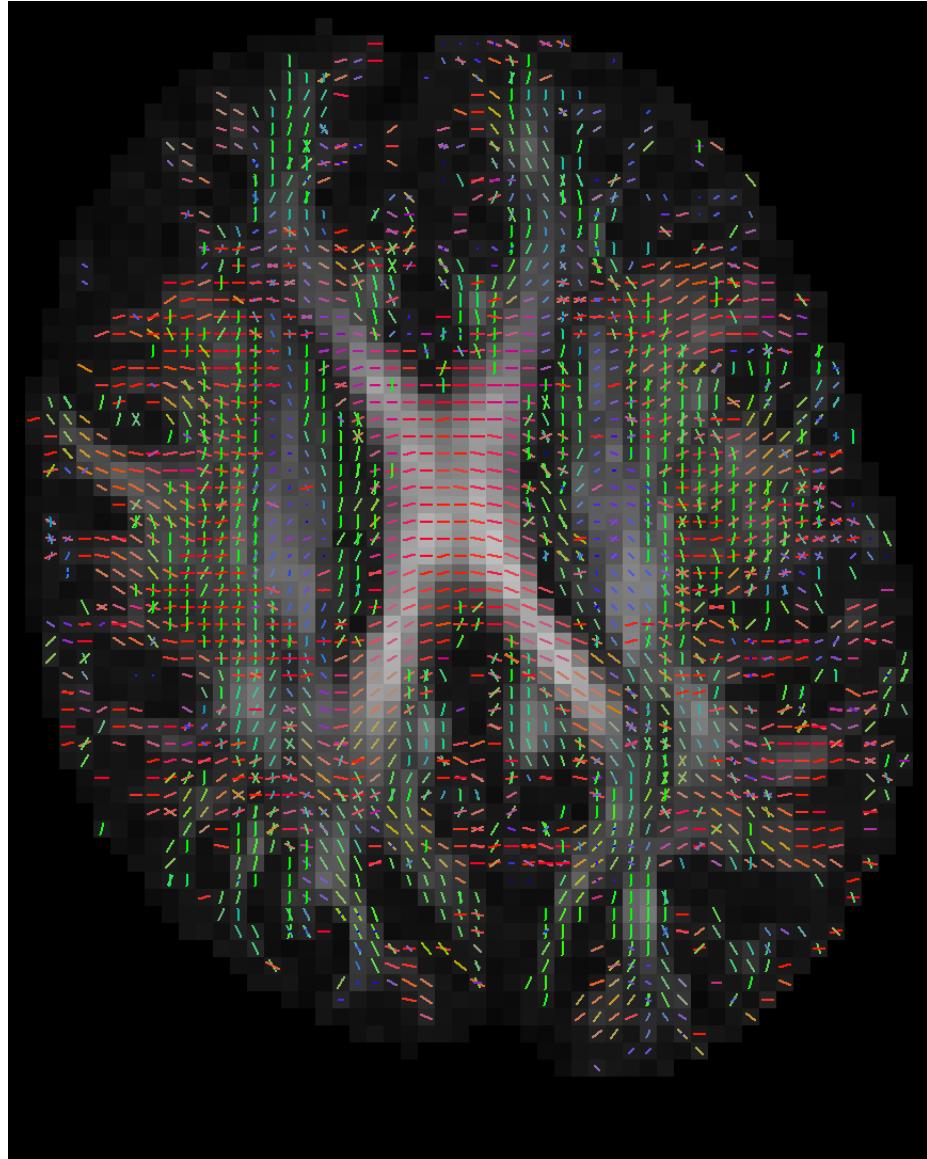


Figure 2.25: Multiple crossings of a real human data set using Quantitative Anisotropy. The first component of QA (QA_0) is also shown in the background.

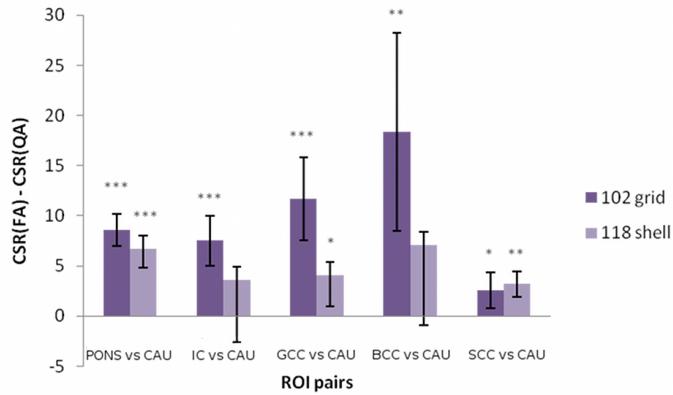


Figure 2.26: Sample results of the paired t-tests comparing CSR(FA) and CSR(QA_0)

schemes were considered: a 102-point grid acquisition with a maximum b-value of 4000 s/mm^2 , and a single shell acquisition using 118 non-collinear gradient directions and a b-value of 1000 s/mm^2 . The two acquisition schemes were matched for total acquisition time ($14\text{ min }37\text{s}$), voxel resolution, and bandwidth. FA, MD and QA_0 maps were then generated for each acquisition scheme and for the 12 volunteers using dipy [77](dipy.org). All the FA datasets were non-linearly registered into MNI space using FSL tools, and the same transformation parameters were applied to MD and QA_0 maps . Fourteen ROIs of different brain regions were drawn in MNI space: putamen (left and right), caudate (left and right), thalamus (left and right), parasagittal white matter (left and right), pons, internal capsule (left and right), and genu, body and splenium of the corpus callosum. Small cubic ROIs were also constructed by finding the centroid of each anatomical ROI and using it as the centre for a $3 \times 3 \times 3$ ROI. For each ROI we calculated the mean value for each metric, and the spatial coefficient of variation (CV) within the ROI (see eq. 37).

$$CV_{ROI} = \frac{\sigma_x}{\langle x \rangle} = \frac{N_{voxels} \sqrt{\sum_{x_i \in ROI} (x_i - \langle x \rangle)^2}}{\sqrt{N_{voxels} - 1} \sum_{x_i \in ROI} x_i} \quad (37)$$

The coefficient of variation of each ROI mean across subjects was also

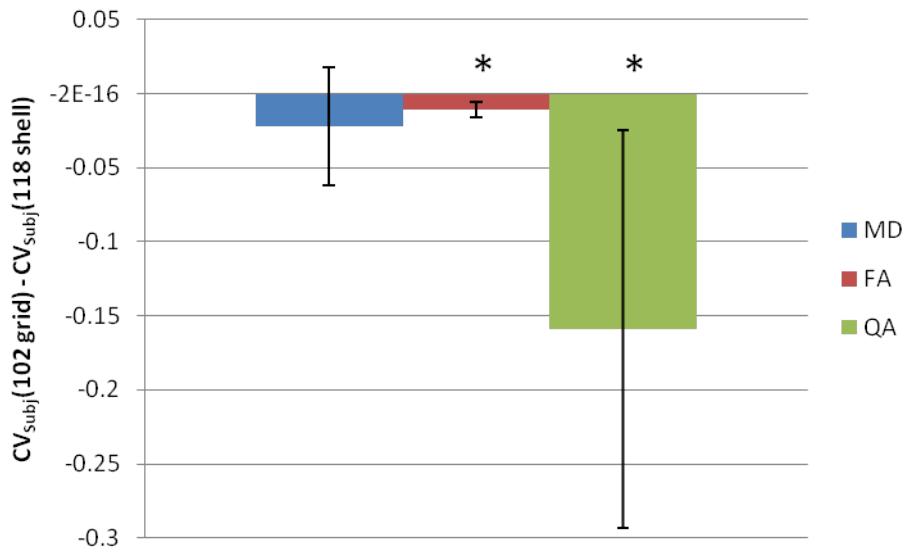


Figure 2.27: Results of the paired t-test comparing the CVs across subjects for MD, FA and QA0.

calculated, as a measure of each metric's comparability between subjects. The contrast-to-scatter ratio (CSR) (calculated for FA in eq. 38) is a good measure of a metric's ability to differentiate between different brain fibre populations [43].

$$CSR(FA) = \frac{mean(FA)_{ROI_1} - mean(FA)_{ROI_2}}{\sqrt{var(FA)_{ROI_1} + var(FA)_{ROI_2}}} \quad (38)$$

Combining the left and right versions of each ROI, we have 9 ROIs of different brain populations, which can be used to define 36 pairs of ROIs, and the CSR of all metrics was calculated for each of these pairs. Paired t-tests were then conducted to compare the performance of each metric with the two acquisition schemes, and also to compare FA and QA_0 directly for each acquisition scheme.

The 102 grid sampling scheme produces significantly higher mean FA and QA_0 values than the ones obtained with the 118 shell scheme, while the opposite was observed for MD. The CSR results for FA and QA_0 were not significantly different between the two acquisition schemes, but the 102

grid scheme produces significantly higher CSRs for MD for 26/36 ROI pairs (fig. 2.26). For MD, no significant difference was found for the CV across subjects, but for FA and QA₀ the 102 scheme produced results more comparable across the different volunteers (fig. 2.27). For FA and MD the 102 scheme showed lower CV within ROIs, especially for white matter, but no difference was found for QA₀. When comparing FA and QA₀ directly, our results show that FA produces higher CSRs than QA₀ for 23/36 ROI pairs for the 102 grid sampling, and for 19/36 ROI pairs for the 118 scheme. FA also shows lower variation across subjects for both acquisition schemes. Finally, FA lower CVs within white matter ROIs, while QA₀ shows less variability for grey matter. The results described and shown above were obtained with the cubic ROIs, but do not differ significantly when the same analysis was applied to larger anatomical ROIs.

Our results indicate that the MD and FA maps generated from a grid sampling scheme designed for GQI are still suitable for analysis, since they do not show poorer performance when compared to a single shell and low b-value acquisition. In fact, the overall results suggest that the 102 grid sampling produces slightly more robust results than the 118 shell acquisition. A previous study [42] has shown that metrics such as MD and FA benefit from the use of multiple b-values, which could explain the better performance of the 102 grid scheme.

2.10 Discussion and Conclusion

Non-parametric methods have the advantage of representing the signal with minimum number of assumptions and without needing any fitting. For many years there has been a trend in science to prefer model-based methods rather than model free (non-parametric) perhaps because model-based can be easier to describe, easier to invent and easier to calculate Bayesian statistics with. However, there are some crucial issues with fitting: (a) Usually the interesting models have many parameters and that makes fitting very slow. (b) Commonly non-linear fitting is needed and accurate fitting is not trivial. (c) Often the model does not represent precisely the complexity of the real problem. (d) The more complex the model, the more difficult to fit

[170], [119],[142].

Non-parametric methods avoid fitting and that makes them very favourable. The concentration of this document was on inventing new non-parametric methods (EIT) or comparing and extending existing ones (GQI2). We showed that a simple, fast and comprehensive transform exists that we call the Equatorial Inversion Transform (EIT). With this transform we showed that we can represent accurately the directional information of the diffusion signal. Furthermore we showed that there are many different functions (F and O see eq. 27,28,29) that can be used in order to create spherical density functions and use these to find the primary fibre directions. With a correct choose of F and O we can create theoretically the same ODF as the real ODF (DSI ODF). This can be done using EITL which is a type of EIT. Nonetheless, other density functions can be created that can identify the leading fibre directions without being real ODFs but they are still different types of spherical densities. EITL2 and EITS where examples of this last case.

Apparently, the EIT concept opens new doors for the investigation of dMRI where many new functionals can be invented in the future that stress more or less different properties of the signal. Notwithstanding, today we have already illustrated and measured that EIT has the best performance with simulations against the other stater-of-the-art methods like DSI and GQI and that empirically, EIT gives as good results with real data sets.

The EIT finds directly the ODF without creating the diffusion propagator. If for some purpose the diffusion propagator is still required then DSI or DPI [52]are favourable. It could be interesting in the future to try and recover the propagator using ideas from the EIT. However, nearly always the propagator is not needed for the analysis. Furthermore, comparing 4D densities like the propagator is a non-trivial problem, also storing the propagator for every voxel is very inefficient.

We discussed that GQI can be used for creating Quantitative Anisotropy (QA). We observed that QA acts like a differential operator but also has some similarities with FA as it is maximum on anisotropic and 0 on isotropic voxels. QA assumes that a fat isotropic part can always be removed from the ODF and that makes it more favourable for spherical functions like those of GQI and EITS. This is in contrast to sharper densities like those of DSI

and GQI2 where QA is not that much useful because the minimum value of these densities will be usually near 0.

GQI needs a manually set parameter; the diffusion sampling length and in contrast the EIT is fully automatic i.e. we always just used the few default parameters for all experiments. The diffusion sampling length can be slightly different from experiment to experiment. The asset of GQI and GQI2 (which was presented together with GQI but not investigated until today) is that they are fast to compute and have simple analytical solutions. GQI2 seems robust and smooth and it has good performance both with simulations and real data.

It is important to stress that there are similarities between all these methods; DSI similar with EITL, GQI with EITS, GQI2 with EITL2. In addition we showed that we can denoise the signal using a Gaussian Spherical Angular method which operates on spherical densities and has a single parameter which is similar with the variance.

Finally, we showed that the first component of QA (highest QA value) can be used for subject comparisons in a similar way with FA. We also showed that NPA could replace FA if we want to calculate anisotropy in a completely geometric way.

The source code for all these methods can be found in dipy (dipy.org in module `dipy.reconst`).

3 Euler Delta Crossings Propagation

3.1 Overview

Tractography methods provide tools to resolve major neuronal fibre bundles non-invasively and in-vivo [32]. From the development of the first tractography algorithms [143, 36] p nearly 12 years ago a great number methods have been published. Most of these algorithms depend strictly on the underlying voxel model or acquisition paradigm making it difficult for other researchers to apply their own reconstruciton methods and evaluate their data sets.

In this work we designed a purely deterministic method which is fast, accurate and all-inclusive. But most importantly it can have as input model-based or model-free reconstruction algorithms of most known algorithms. We call this algorithm EuDX. Eu stands for Euler integration, D stands for Delta function which is a function that checks for many different stopping criteria and X stands for fibre crossings. EuDX can deal with any number of crossing fibres as long as the reconstruction algorithm supports them. The purpose of this algorithm is to be faithful to the reconstruction results rather than try to correct or enhance them by introducing regional or global considerations which is the topic of other methods reviewed below. Therefore, EuDX serves mainly as a robust method for quickly inspecting different reconstruction results using streamlines. EuDX is noise-friendly i.e. if a voxel is too noisy then EuDX will stop tracking on that voxel. This property is often useful when validating underlying reconstruction models. Branching is also supported by a combination of trilinear interpolation and propagation along multiple peaks per voxel. This method is an extension of the method used by Conturo et al.[36] and Yeh et al. [205] with the additional support for multiple fibre-crossings.

In sections 1.8,1.9 we discussed some of the ideas and the problems behind the most popular propagation methods; deterministic and probabilistic. The focus in this section will be to give a more general overview and introduce many more methods.

Most tractography techniques can be grouped in three categories, as pointed

out in Sotropoulos thesis [179]; local, global and simulated. Local approaches propagate a curve from a starting (seed) point using locally greedy criteria, i.e. tracking sequentially through orientation estimates in adjacent voxels. Global approaches identify the best path between two points of interest, according to some optimization criterion, rather than identifying paths arising from a single point. Simulated approaches comprise of algorithms that simulate the diffusion process or solve the diffusion equation to reconstruct white matter tracks. A detailed literature review is given below.

3.1.1 Local

Deterministic tractography was the first to appear. Tracks (also known as streamlines) are created as trajectories in the form of polylines; orthograde and retrograde along an initial direction at a specific point (seed) in the 3D volume. In FACT[143] tracks are propagated in unequal steps governed by the entry point of the streamline in the voxel (see Fig. 1.12). Euler integration with equal steps was used in Conturo et al. [36] and similarly Runge-Kutta integration was used in Basser et al.[13]. Deterministic approaches usually stop propagating when a low anisotropy region (usually $FA < 0.2$) is found. This is useful in order to avoid propagation within CSF where anatomical tracts do not exist or within deep gray matter regions where tracking is uncertain. They usually check also for large angle changes (e.g. larger than 90°) between successive steps to avoid unrealistically sharp turns.

Deterministic methods can be also utilized when multiple orientations are estimated in a single voxel (crossing fibres). These orientations can be obtained for example as the principal eigenvectors of multiple Tensors fitted to the data [190], or from the local peaks of the diffusion ODF estimated using DSI [198] and QBI [191] or from the orientations from the fibre ODFs [185]. There are different approaches for propagating across voxels with more than one fibre directions. One approach is, upon entering a voxel, to choose the orientation that produces the smallest curvature with the incoming path [197]. Another approach follows all orientations that do not exceed a curvature threshold, by initiating a new streamline per orientation [34],[55].

Another interesting point is that most methods which belong in this category utilize only the fibre orientation estimates. Tensor deflection tractography (TEND) proposed by Lazar et al. [118] is a FACT variant that uses the whole DTI Tensor rather than just its principal eigenvector to determine the direction of curve propagation. All the methods presented up to this point provide binary connectivity information i.e. a voxel B can be either connected or not connected to the seed S, depending on whether a streamline from S passes through B.

Probabilistic tractography was introduced by Parker et al. [160] and Behrens et al. [19]. Here the standard procedure is to calculate a spatial distribution of tracks arising from a single seed rather than a single track. In each propagation step of each streamline, a random perturbation of the underlying fibre orientation estimate is followed. Perturbations are generated using functions that characterize the uncertainty in the fibre orientation within each voxel. A probabilistic index of connectivity (PICo) is defined between a seed and an arbitrary point as M/N ; where N is the number of all the tracks that start from the seed and M is the number of tracks that traverse the seed and the arbitrary point.

Probabilistic approaches mainly differ in the way the orientation uncertainty is assessed. Most commonly a Bayesian framework will be used to calculate the posterior probability of the reconstruction model's orientation parameters [19],[20],[94],[75] and [209]. In Behrens et al. [19],[20] Monte Carlo-Markov chain (MCMC) was used to sample the orientation posterior distribution. In Friman et al. [75] the posterior is computed numerically after using Dirac priors. In Zhange et al. [209] particle filtering was used for the same purpose.

Bootstrap tractography is an alternative method to characterize uncertainty in fibre orientation [157],[117]. This is a non-parametric approach where a diffusion acquisition is repeated many times for the same subject and some images from this set are drawn in random with replacement. This process gives a single bootstrap sample. Drawing many samples will give a distribution for the fibre orientation. The advantage of bootstrap tractography is that no ad-hoc assumptions are made on the noise and it is sensitive to all sources of variability that affect the acquired dataset. The disadvantage

is that many repeated acquisitions are required; at least 5 for DTI according to the disadvantage is that many repeated acquisitions are required according to O’Gorman et al. [152].

Model-based residual bootstrap offers an alternative, since it requires only a single data acquisition (Chung et al. [35], Berman et al. [21]; Harroon et al. [90]; Jones et al., [105]). A single bootstrap sample can then be generated by permuting freely the residuals (or just the signs of the residuals using wild bootstrap Jones et al. [105]; Whitcher et al. [200]) between all model predicted values. The bootstrap technique was first introduced by Efron [64] in 1979.

Other probabilistic approaches estimate the orientation uncertainty as an empirically defined function. For example in Parker et al. [160] this is determined using the degree of FA. The higher the FA the higher the confidence on the principal eigenvector of the Tensor. In Parker and Alexander [158] Monte-Carlo simulations are used to predict the orientation uncertainty for multiple Tensors and later in [161] for PAS. In Descoteaux et al. [53] the fODF was used for the same purpose. Cook et al. [37] used a Watson distribution and Seunarine et al. [174] used a Bingham distribution. In the same category we can also classify the work of Bjornemo et al. [22] from 2002 who created a regularized stochastic method for probabilistic tractography. This method utilizes the principles of a statistical Monte Carlo method called Sequential Importance Sampling and Resampling (SISR). This technique is similar with particle filters. The disadvantage of the method is that it has strong assumptions for the single Tensor as the reconstruction model.

3.1.2 Global

A limitation of probabilistic tractography is that the probabilistic index of connectivity decreases with distance from the seed point (see 1.8,1.9). Another limitation is that it is still sensitive to local noise. Global approaches try to overcome these limitations by being distance-independent and by increasing resistance against noise. These are achieved by finding an optimal path between two voxels, according to a global property (see [103]).

Jbabdi et al. [100] developed a **Global Bayesian** model to derive the posterior probability of connections. The path trajectories represented by splines are compatible with the local fibre orientations in regions with low uncertainty estimates. In regions with high uncertainty, the global connectivity information constrains the local parameter estimation and affects the path sampling.

Front evolution techniques often employ fast marching techniques. The front expands from the seed neighbours to the next neighbouring nodes with speeds determined by the local fibre orientations. As the front propagates, a time of front arrival can be associated with each visited voxel. Once all image voxels have been traversed by the front, paths of connection can be obtained going backwards in the map of front arrival times. Starting from an arbitrary voxel, a gradient descent algorithm can find the fastest route back to the seed. A connectivity index can be associated with each path, representing either the weakest link along the path or the agreement between the path tangents and the underlying vector orientation field (see Parker et al. [162], Tournier et al. [184], Cambell et al.[26], Fletcher et al. [71] and Gigandet et al. [80]).

Graph-based tractography utilises weighted networks (graphs). This type of tractography was presented by Iturria-Medina et al.[96], Zalesky et al. [207], Lifshits et al.[123], Fillard et al. [69] and Sotiroopoulos et al. [180]. The common concept of these approaches is that each image voxel becomes a node in the graph where the edges of the graph connect pairs of neighbouring voxels or ROIs. The edges are assigned weights, that can be representative of any type of structural information. Anatomical paths are then defined as chains with successive elements being neighbouring voxels. The weights of the edges are used to determine the path strength. The strongest path between any image voxel and a seed can then be identified using algorithms that search efficiently the image graph.

Energy Minimization [111] methods try to optimize all tracks from the whole brain volume simultaneously. Each tract is represented as a chain of cylinders, whose position and orientation can change. The method tries to find the set of cylinders that best approximate the underlying white matter bundles. This is achieved by minimizing the overall energy of all cylinders

simultaneously, mimicking natural phenomena e.g. the polymerization process which is a process of reacting monomer molecules together in a chemical reaction to form three-dimensional networks. Many standard algorithms e.g. gradient descent are usually employed with this framework; however Gibbs sampling is the most common in tractography. Kreher et al. [110], Reisert et al. [169], Lazar et al. [116] and Fillard et al. [68] showed results using energy minimization. Despite the very promising results shown by Reisert et al. [169], whose team won the Fibre Cup competition[67] the very high computation time is an important disadvantage of this framework.

Microstructure Tracking is a very exciting new family of algorithms which started with MicroTrack[175] that combine global tractography and direct microstructure estimation using diffusion-weighted imaging data. Connectivity via tractography, axon diameter distribution and density estimates are all combined in order to inform one another given the common assumption that microstructural features remain consistent along fibers. This type of algorithms need of course special data sets of acquisitions similar with those employed in ActiveAx invented by Alexander et al.[4].

3.1.3 Simulated

This family of methods take a very different approach from what we have discussed until this moment. They simulate the diffusion of water molecules within the brain tissue or directly solve Fick's second law of diffusion in the entire brain (see [15],[108],[86],[88]).

In Batchelor et al. [15], the diffusion equation is solved using a finite element approach. Successive diffusion simulations over the entire brain, starting from a seed, are performed in Kang et al. [108]. Tractography by simulating fluid flow through a "pressure" Tensor field is performed in Hageman et al. [86]. The Navier-Stokes equation is solved using a finite element approach. However, solving a partial differential equation increases execution time. Furthermore, it is not always easy with these approaches to obtain a connectivity map across the whole brain volume and there is usually a large number of parameters to set.

For further understanding of the current map of all the different trac-

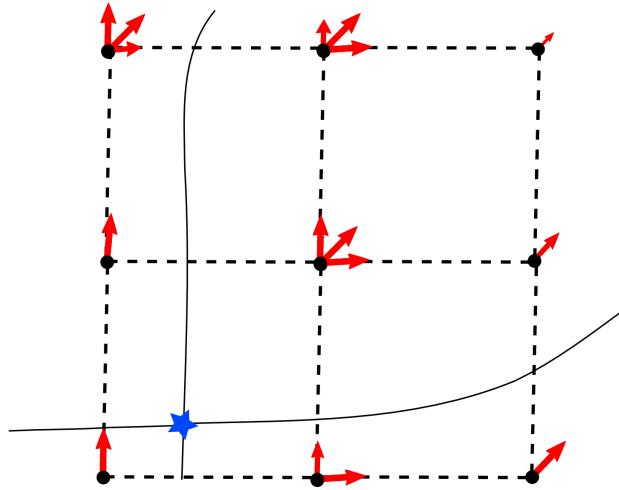


Figure 3.1: In this figure we see that in every voxel centre (black dot) there is one or more vectors. These vectors represent peaks where their length is equal to the anisotropy value and the direction is equal to direction of the peak e.g. calculated in a given ODF. EuDX can track multiple peaks starting from a single seed point (star) if their anisotropy values are higher than a threshold. In order to reduce memory usage we discretize the directions using the vertices of a unit sphere. Then we only need to input to EuDX a 4D volume with anisotropy values \mathcal{A} and an 4D volume \mathcal{I} with the indices of the directionse. The first three dimensions in \mathcal{A} are for the voxel position and the last dimension is for the peak values (anisotropy). Respectively, first three dimensions in \mathcal{I} are for the voxel position and the last dimension contains the indices of the directions(vertices) on a discrete sphere. In that way we track from the same seed towards different directions and support tracking in crossing areas.

tography algorithms see the reviews from [67],[179] and [98]. In Fillard et al. [67] 10 tractographies were simulated using a novel hardware phantom. Comparing all these different methods which all have different parameters and are based on different underlying models is a difficult process and a ground truth from anatomical studies, digital or hardware phantoms is absolutely necessary.

3.2 Algorithm

We created an algorithm that has many similarities with the classical deterministic methods [143, 36, 13] and with more recent as those described in

[55, 205]. Our concentration was to create a more general streamline tractography propagation framework which can be used with very different families of anisotropic functions and work well with fibre crossings. This algorithm that we call EuDX is applied usually in native space image coordinates and it assumes that the voxel dimensions are equal in all three dimensions (isotropic voxel size). In order to create streamlines (tracks) we need to provide initially one or more seed points S . These can be choosed randomly or we could specify them explicitly. However, these seed points need to be constrained by the volume's dimensions. Every seed point \mathbf{p}_0 becomes the starting point for the track propagation.

For the integration we solve for $\mathbf{p}_t = \mathbf{p}_0 + \int_0^t \mathbf{v}(p(\mathbf{s}))d\mathbf{s}$ and we perform the integration numerically using Euler's method

$$\mathbf{p}_{n+1} = \mathbf{p}_n + \mathbf{v}(\mathbf{p}_n)\Delta s \quad (39)$$

where Δs is the propagation step size which should be at least smaller than the voxel size, \mathbf{v} is the propagation direction. Alternatively, Runge-Kutta of 2nd and 4th order can be used. However, in this document we used only Euler's method. For EuDX stopping criteria we can use a standard scalar function like FA but we can also use vector functions like Quantitative Anisotropy (QA) or even the full Orientation Density Function (ODF). The only constraint for these functions is to be greater or equal to zero everywhere in the volume. In most cases all these functions try to measure in some way the anisotropy of diffusion in every single voxel so we decided to use letter \mathcal{A} for the purpose of representing all these fields applied on the image grid. Therefore, when we write $\mathcal{A}(\mathbf{u}_i) = \alpha_i$ this reads for the peak unit direction of \mathbf{u}_i the peak value was α_i . For the simple case of FA, \mathbf{u} is equal with the eigenvector corresponding to the highest eigenvalue, $\mathcal{A}(\mathbf{e}) = FA$. For QA which can allow even for up to 5 peaks or more we use $\mathcal{A}(\mathbf{u}_i) = QA_i$ where $i \in [0, 4]$ denotes the current peak. The peak can be characterised by two things the anisotropy value α_i and the unit direction of the peak \mathbf{u}_i . A visual example of the combination of multiple peaks together with EuDX can be seen in Fig. 3.1.

The EuDX algorithm can be described further in the following way. $\mathcal{A}(\mathbf{u}_i)$

is estimated at every point of the volume. This resembles a composite vector field where every point contains the peak directions superimposed to the anisotropy values. In order to reduce space \mathbf{u} can be replaced by an index to the closest sampling point of an evenly distributed sphere. We create an empty list of tracks $T = \emptyset$ and then we generate random or prespecified seed points \mathbf{p}_{0j} for every track t_j .

We pick a seed point \mathbf{p}_0 and start propagating. We need to remember that the propagation direction can go forward and backward or better said; towards one direction (orthograde) and its opposite direction (retrograde). For the moment we only propagate towards one direction set by

$$\mathbf{v}(\mathbf{p}_0) = \operatorname{argmax}_{u_i} \mathcal{A}(\mathbf{u}) \quad (40)$$

but we need to propagate also towards the opposite direction $-\mathbf{v}(\mathbf{p}_0)$. Furthermore, as \mathcal{A} can have multiple values in each voxel (representing different peaks) we need to remember that when we finish with this track followed by direction \mathbf{v} we will need to propagate towards the direction of second maximum, third maximum etc. This is necessary if and only if \mathcal{A} gives information for multiple peaks (e.g. QA).

Apart from the direction we need to check when to stop tracking. EuDX takes as input a threshold for anisotropy \mathcal{A}_{thr} . If $\mathcal{A}(\mathbf{p}_n) < \mathcal{A}_{thr}$ then EuDX stops propagating. Otherwise it appends the point to the current track t_j . This can be useful for cancelling out any seed points which are in the background or in very low anisotropy areas where tracking is not recommended. One important point here is that \mathcal{A}_{thr} depends on the reconstruction method and it will have a different value for every different metric QA, FA, GFA etc.

In order to generate a smooth tractography it is recommended to use some kind of interpolation. Here we have been using trilinear interpolation which works in the following way. The seed devides the neibouring area (constrained by the centers of the neighbouring voxels) in 8 regions in 3D (it would be 4 in 2D) and the total contribution of the neighbouring points is added according to the weights w of the antipodal side. The weights express subvolumes in 3D (or subareas in 2D). The trilinear interpolation give us the weights w that assign the contribution of the peaks of the neighbouring

Algorithm 3 EuDX Tractography

Input $\mathcal{A}, \mathcal{I}, S, U, \Delta s, VS$
 $\mathcal{A}_{thr}, \theta_{thr}, TW, MNP$ (thresholds)
Output T

$T = \emptyset$
ForEach seed In S
 ForEach peak In $(\mathcal{A}, \mathcal{I})$
 $t = \text{EUDX_Core}(\text{seed}, \text{peak})$
 Append(T, t)

voxels to the seed's next direction.

We find the nearest direction from the seed's initial direction $\mathbf{v}(\mathbf{p}_n)$ to every \mathbf{u}_i of every one of the 8 corners of the neighbourhood of the seed.

If $\arccos(\mathbf{x}_i, \mathbf{p}_n) \leq \theta_{thr}$ we count the corresponding weight otherwise we continue to the next weight. We check simultaneously for the condition $\mathcal{A}(\mathbf{u}) < \mathcal{A}_{thr}$. All the adjacent weights will contribute to the new direction according to the following formula

$$\mathbf{v}'(\mathbf{p}_n) = \sum_m w_m \mathbf{v}(\mathbf{p}_n) \quad (41)$$

where $\mathbf{v}(\mathbf{p}_n) = \frac{\mathbf{v}'(\mathbf{p}_n)}{\|\mathbf{v}'(\mathbf{p}_n)\|}$ is normalized. The next point is calculated with eq. 39. We insert the new point in the track and continue tracking until one of the stopping criteria is met. Finally, we have to repeat the procedure for the opposite direction of the initial $\mathbf{v}(\mathbf{p}_o)$ and for the smaller peaks as in (2).

Many other thresholds apart from the anisotropic threshold \mathcal{A}_{thr} and angular threshold θ_{thr} are also incorporated in EuDX. a) The total weight TW check that there is enough overall neighbourhood weighting to continue tracking (has default value 0.5). This is very useful in edges or corners where you want tracking to stop. b) It is possible for a track to get trapped in a loop and start looping for ever. We can check for that using a threshold for maximum number of points for a track MNP (default 1000). c) Also we need to check that we are always inside the 3D volumetric space. The dimensions of the volume are hold in vector V .

Algorithm 4 EuDX Core

Input seed, peak, ΔS , V thresholds

Output track

```
delta, i_direction = Initial_Direction(seed)
#propagate orthograde
direction=i_direction
While delta Is True
    delta, n_direction = New_Direction(direction)
    If delta Is False
        Break
    point = point +  $\Delta s \times n\_direction$ 
    Append(track,point)
    direction = n_direction
    delta = True
#propagate retrograde
direction = - i_direction
While delta Is True
    #Same as above
```

In alg. 3 we see how EuDX can be called in order to create an entire tractography. The core algorithm which is the same for every seed is given in alg. 4. In alg. 3 the inputs are: \mathcal{A} the 4D volume holding the strength of anisotropy in relation to the indexed directions of sphere U given in \mathcal{I} . U is an array of size $N \times 3$ where N is the number of vertices in the sphere. S is an $M \times 3$ array with the precomputed M seed points (seeds). Δs is the propagation step. We also need the different stopping thresholds: \mathcal{A}_{thr} defines the lowest possible anisotropy value that allows tracking to continue. θ_{thr} is the maximum allowed angle between the current propagation direction and the next direction. TW checks the overall contribution of the neighbourhood for the next propagation direction and MNP checks that a track does not pass from the same point more than a number of times. MNP checks this condition by counting the number of current points appended in a track as the track grows.

The main corpus of EuDX is given in alg. 4. This is called for every seed and peak. The algorithm starts by calling the function Initial_Direction

which finds the closest peak direction to follow from the nearby voxels and if the values of these peaks are lower than A_{thr} then it returns False otherwise it returns True and the vector with the initial direction. The variable delta is used as check point that no stopping criteria are met. As long as delta is True we can continue tracking. The function New_Direction calculates the new direction of the propagation by accumulating the weights created by trilinear interpolation as shown in eq. 41 and finds the nearest direction for the eight neighbouring voxels. It also checks for the total weight threshold TW and if we are inside the volume's boundaries given by V. After this is done we update the points as shown in eq. 39 and repeat the same procedure until one of the stopping criteria is met (by checking variable delta). Finally, we repeat the same for the retrograde direction in order to create the entire track.

3.3 Results with digital phantoms

In order to validate the performance of EuDX we created a 3-dimensional digital phantom with the method described in section 2.8.2-XXX [27] with eigen values $\lambda_{\parallel} = 1.7 \cdot 10^{-3} m^2/sec$ and $\lambda_{\perp} = .1 \cdot 10^{-3} m^2/sec$. The digital phantom consists of two parts: a) a diagonal orbit and b) an elliptical orbit with (with axes ratio $\lambda_2 = 0.6\lambda_1$). Both parts were added together to create a crossing configuration and no partial volume effects were simulated. The average thickness of both parts was 5 voxels. The elliptical orbit did not complete a full ellipsis in order to avoid tracks from looping forever. We created two experiments using the same data set of size $64 \times 64 \times 64 \times 102$ where the 3 first dimensions were giving the size of the volume and the last dimension is the number of diffusion weighted volumes (101) plus the unweighted b0 volume (see Fig. 3.2(i) lower right - jet colourmap). We generated b-vectors and b-values by using a keyhole Cartesian sampling grid[190](see Fig. q-space Introduction 1.5 XXX); therefore, DSI, GQI, EIT or DTI were suitable for the signal reconstruction into orientation distributions or Tensors respectively. For both experiments we used a high SNR of 100 as the main goal was to validate the algorithm on good conditions. We later continue the validation process with human data where SNR is naturally

lower.

In the first experiment (see Fig. 3.2) we generated 200,000 uniformly random seeds in the 3D volume. We used DSI reconstruction with standard parameters: q-space grid size 16, hanning filter width 32, and radial integration range from 2.1 to 6 at steps of .2. The input of EuDX expects the peaks and the directions of the peaks. For every voxel we used DSI to create the corresponding ODF (sampled on an evenly distributed sphere of 642 vertices and 1280 faces) and from the ODF we used the peak finding function introduced in Ch. 2.6 XXX. We further removed the peaks which had values less than 70% of the highest peak and normalized the rest so that the maximum peak equals to 1. We also calculated FA and zeroed the peaks with FA values less than 0.2. In that way we ensured that there will be no tracking in the background area of the phantom.

Finally, we called EuDX with parameters $\mathcal{A}_{thr} = 0.2, \Delta s = 0.5, \theta_{thr} = 60^\circ$ and $TW = 0.5$ and visualized the tracks. The results are shown in Fig. 3.2. The tracks are colourcoded with their orientation defined as the unit vector connecting the first with the last point of each track. Because we generated random seeds that went everywhere in the phantom some of them had to be in the crossing area. In accordance with what we discussed in the previous section we know that if a seed falls in a crossing region EuDX will propagate towards both directions of the crossing if and only if multiple peaks are supported from the underlying reconstruction method. We can confirm this by looking at fig3.2(i) and in (ii), (iii) in higher resolution; where we can see that the crossing area is well represented and that the propagation was successful. We can also observe that not all tracks travel all the distance from end to end. Otherwise you would expect to see the tracks of the entire elliptic part all with one colour and the tracks of the straight orbit all with a second colour. It seems that this had to do with the discrete nature of the phantom especially near the boundaries of the phantom. Therefore, we created a different experiment to evaluate this observation.

In the second experiment we generated 2000 seeds inside specific areas (see Fig. 3.3-left panel). More precisely the end-point areas of the phantom denoted with A,B,C and D. The goal was to count the percentage of tracks which reached any of the other end-point areas. We tried this with 4 differ-

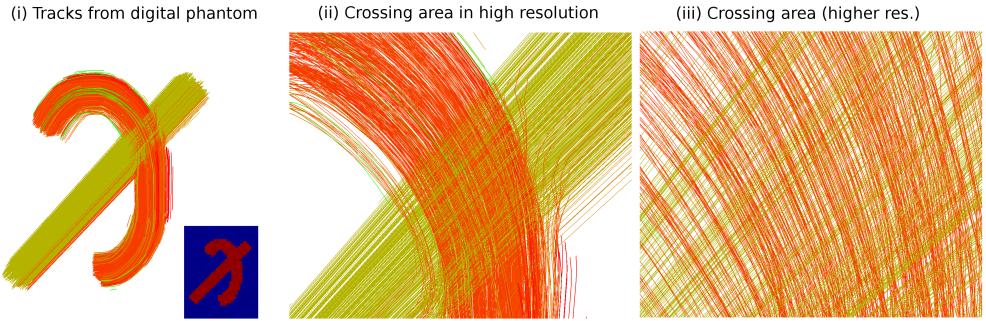


Figure 3.2: Left panel shows the tractography for a software phantom generated by EuDX with DSI as the reconstruction method.

ent reconstruction methods: Single Tensor, EITL, DSI and GQI. The Single Tensor was fitted using weighted least squares. For EITL we used for radial sampling from 0 to 5 with steps of 0.4 and gaussian weight of 0.05. GQI had sampling length of 1.2 and DSI had the same parameters as in the previous experiment. The orientations of the peaks for all the methods were found in a reconstruction sphere of 642 vertices. The same peaking function was also used with the ODFs of EITL, GQI and DSI. For these 3 methods the peaking procedure was the same as described in the previous experiment. The single peak of the DTI reconstruction was found by the corresponding eigenvalues and eigenvectors. EuDX expects the orientation input of the peaks as indices on a unit sphere; therefore, for the eigenvector which corresponds to the maximum eigenvalue we found the vertex in the reconstruction sphere (used above) with the minimum angular distance.

The results of this study are summarized in tab. 3.1. A first general observation is that independently from which end-points the tracks start there will be always a percentage of tracks which will stop before reaching the other end-points. We can see clearly this phenomenon by looking at the last column (\emptyset) of each sub-table; where the best case was 3.7% loss (EITL) and the worst case was 43.5% (DTI).

So, why is there always an amount of tracks which will not reach the other end-points? This happens because most of the area in high-dimensional volumes (3D volumes) is concentrated close to the boundaries of the volumes; therefore many from the random seeds will fall close to the bound-

DTI	A	B	C	D	\emptyset		EITL	A	B	C	D	\emptyset	
A	–	0%	76.4%	0%	23.6%		A	–	63.4%	8.7%	0%	27.9%	
B	0%	–	0%	59.8%	40.2%		B	65.6%	–	0%	5.6%	28.8%	
C	79.9%	0%	–	0%	20.1%		C	14.5%	0%	–	76.8%	8.7%	
D	0%	56.5%	0%	–	43.5%		D	0%	0.5%	95.8%	–	3.7%	

DSI	A	B	C	D	\emptyset		GQI	A	B	C	D	\emptyset	
A	–	65.3%	9.6%	0.0%	25.1%		A	–	57.8%	8.7%	0.0%	33.5%	
B	72.6%	–	0.0%	5.5%	21.9%		B	67.2%	–	0.0%	10.7%	22.1%	
C	14%	0.0%	–	79.9%	6.1%		C	37.5%	0.0%	–	55.5%	7%	
D	0.0%	10.6%	84.8%	–	4.6%		D	0.0%	22.3%	72.3%	–	5.4%	

Table 3.1: Every sub-table shows the percentage of tracks which started from areas A,B,C or D (rows) and reached the other areas A,B,C,D or \emptyset (columns) using EuDX with input from different reconstruction methods DTI, EITL, DSI, GQI. The empty set \emptyset symbolizes here the number of tracks which didn't reach any of the A,B,C,D areas. For example by looking only the first row of each sub-table we easily observe that the crossing area was well represented by EITL, DSI and GQI but not from DTI as it was expected because the Single Tensor cannot resolve crossings. We can also observe by comparing all rows that in the more curved branch of the phantom less tracks reached their target than in the diagonal branch.

aries. In addition, the boundaries are not smooth (see Fig. 3.3-left panel) and for this reason the tracks which will start close to the boundaries will most likely not travel long distances and because the number of seeds which will fall close to the boundaries is not negligible; the number of tracks which will not travel far is not negligible as well. This is something to have in mind when generating tractographies. A naive solution would be to increase the resolution of the volume by interpolation. This has disadvantages as well; the diffusion data sets are already very large and increasing the resolution would increase their already large size. Also interpolation usually introduces smoothing which will reduce the peak resolution in relation to high b-values where the signal is in general low. For this purpose we decided to not do any further preprocessing to the phantom.

Another general observation is that in the curved branch less tracks reached their target than in the diagonal branch. This is sensible because the discretization is more asymmetric in the curved branch (see Fig. 3.3-left panel) therefore more tracks close to the boundaries will stop propagating because

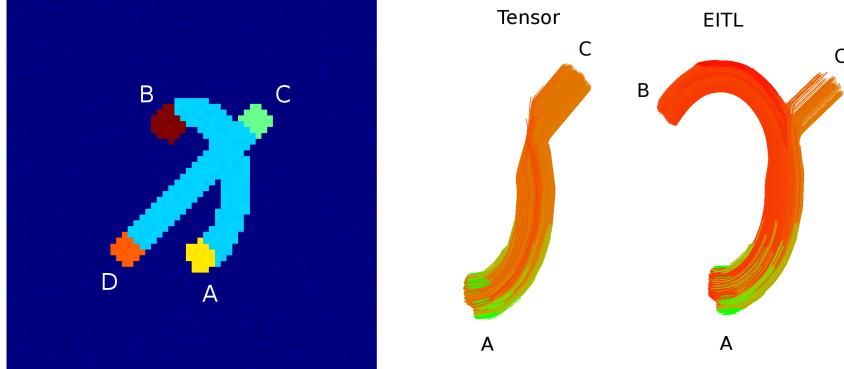


Figure 3.3: On the left panel we see the b0 volume of the digital phantom and on top we see 4 ROIs; A,B,C and D. 2000 seeds were generated inside each ROI and then we measured the amount of tracks which reached the other regions. On the right panel we see that when Single Tensor was used for the reconstruction; the tracking started from A reached only C. However when EITL was used the tracks reached both areas B and C as it was expected.

they reach background voxels where anisotropy is much lower. This is in agreement with EuDX which is designed to stop in low anisotropic regions as defined by the parameter A_{thr} .

In the right panel of Fig. 3.3 we see that the tracks which started from ROI A, reached only ROI C when we used the peaks from the Single Tensor reconstruction method as input to EuDX. However when we used EITL the tracks which started from A, reached both B and C. The tracks didn't reach ROI D because of the angular threshold $\theta_{thr} = 60^\circ$. This was another positive confirmation about the performance of EuDX which identified that even if the seeds are not in the crossing region because of the trilinear interpolation when the tracks will reach the crossing region they will follow the branch to which there are closer to in angular distance.

3.4 Results with humans

We wanted to investigate how EuDX performed with healthy subjects. There are mainly two ways to do validate tractographies with real data. The first is to compare against the results of another tractography method and the second is to ask feedback from specialist medical practitioners who have good

understanding of the underlying anatomy. We investigated first the second approach. In order to do that we generated tractographies (see Fig. 3.4 (right panel)) and asked from expert neuroanatomists Professor Luigi Cataneo and Dr. Nivedita Agarwal from the Center for Brain/Mind Science (CiMeC) in Trento, Italy to validate the quality of the data in relationship to their knowledge from anatomy and manually label bundles of their scientific interest in different subjects. They manually found and labeled on first subject arcuate fasciculus (AC) which is depicted in Fig. 3.4 (left panel). They also identified in a different subject the right Corticospinal bundle depicted in Fig. 3.5 (left panel).

In order to perform the manual labeling we developed an interactive visualization application (fos.me) based on a novel tractography clustering method which is the topic of the next chapter. This application has the capability to create an accessible representation of the initial tractography into bundles of interest (BOIs). Bundles means here a collection of tracks with similar spatial and shape characteristics. After the BOIs are created the medical practitioner can select one or more BOIs interactively and hide the other parts of the tractography which are not interesting to hers/his investigations. This application will not be the topic of this chapter but it is further explained in Chapter 5 (Section? XXXXX).

[74],[171],[81]

The tractographies shown in Fig. 3.4 and left panel of Fig.3.5 were calculated with EuDX parameters $\mathcal{A}_{thr} = 0.02, \Delta s = 0.5, \theta_{thr} = 60^\circ, TW = 0.5$ and with input from QA of GQI reconstruction with sampling length $\lambda = 1.2$. The data sets were generated at a 3T scanner (TIM Trio, Siemens) in Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK. We used Siemens advanced diffusion work-in-progress sequence, and STEAM [140, 135] as the diffusion preparation method. The field of view was $240 \times 240 \text{ mm}^2$, matrix size 96×96 , and slice thickness 2.5 mm (no gap). 55 slices were acquired to achieve full brain coverage, and the voxel resolution was $2.5 \times 2.5 \times 2.5 \text{ mm}^3$. In this experiment much less gradient vectors were used. A 102-point half grid acquisition with a maximum b-value of 4000 s/mm^2 was used. The total acquisition time was only $14 \text{ min } 21\text{s}$ with TR=8200ms and TE=69ms.

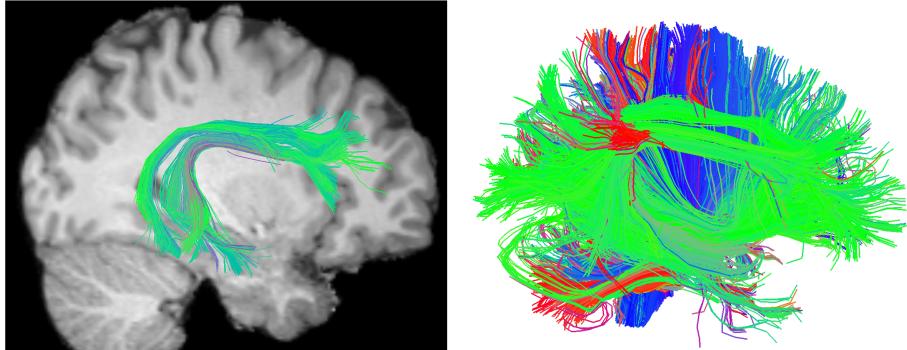


Figure 3.4: Left Panel: right arcuate fasciculus generated by EuDX and segmented by two expert neuroanatomists. The tracks are in MNI coordinates and we visualize simultaneously the T1 slice ($X=29$) for the same subject. Right Panel: the sagittal view of the whole brain tractography of the same subject is shown. For visualization purposes we are depicting only tracks of length from 120mm to 150mm.

The tractographies were generated in diffusion native space and then linearly registered in For that reason FA volumes were generated from the same data sets using Tensor fitting with weighted least squares after skull stripping with FSL BET. These FA volumes were again in diffusion native space therefore we used FSL FLIRT to align them in MNI space. For this purpose a standard template (FMRIB58_FA_1mm) from the FSL toolbox was used as the reference volume. We then applied the affine transformation matrix from the previous step to the initial tractography to align it to MNI space.

In order to help the neuroanatomists to guide themselves with the segmentation of bundles we linearly registered structural MRIs - T1 (MPRAGE) images from the same subjects to the standard template (MNI152_T1_1mm_brain). Therefore, the tractography and the T1 image were properly aligned in the same space. This was very useful for the neuroanatomists because they could find known bundles from the regions they connect in the cortex which are better visible in the T1 rather than in FA.

In the previous section we observed from our simulation experiments that EuDX was able to propagate correctly in crossing areas (see Fig. 3.2) by generating tracks towards both directions of the crossings. We wanted

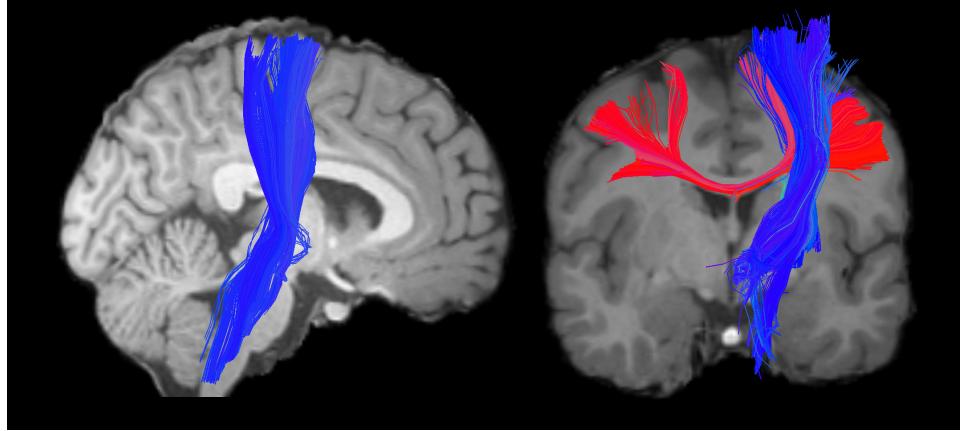


Figure 3.5: Left panel: EuDX tracks from the CST segmented by our expert neuroanatomists. The tracks are linearly registered in MNI standard space and visible is also the T1 slice ($Y=5$) from the same subject. Right panel: Here we see the intersection of the BCC with the right CST. This is a confirmation that EuDX can propagate successfully in crossing areas of human brain data. The T1 slice ($Y=-1$) from the same subject is shown.

to confirm if tracking in crossing areas was also robust with human data where the noise artefacts are less predictable. This was indeed confirmed by looking at the intersection of two well known bundles: the Body of Corpus Callosum (BCC) and the Corticospinal Tract (CST). In Fig. 3.5 (right panel) we can see that a part from the BCC shown with red is intersecting the CST bundle shown with blue without being diverted from the CST. If EuDX was not able to propagate in crossing areas then the tracks from BCC would stop in the intersection area or divert towards the direction of the CST.

We wanted also to compare specific bundles across different healthy subjects. We concentrated at an one not so often studied bundle: the Cingulum; across 12 subjects.

The cingulum (CG) is an association fibre tract that runs within the cingulate gyrus along its entire length. It collects axons from the cingulate gyrus that travel immediately dorsal to the corpus callosum and along the ventral face of the hippocampus, forming a large C-shape trajectory. It carries afferent connections from the cingulate gyrus to the entorhinal cortex. Because of its narrow tubular shape, it is often difficult to reveal its entire length by a single data set [144]. However, using EuDX we see that both left

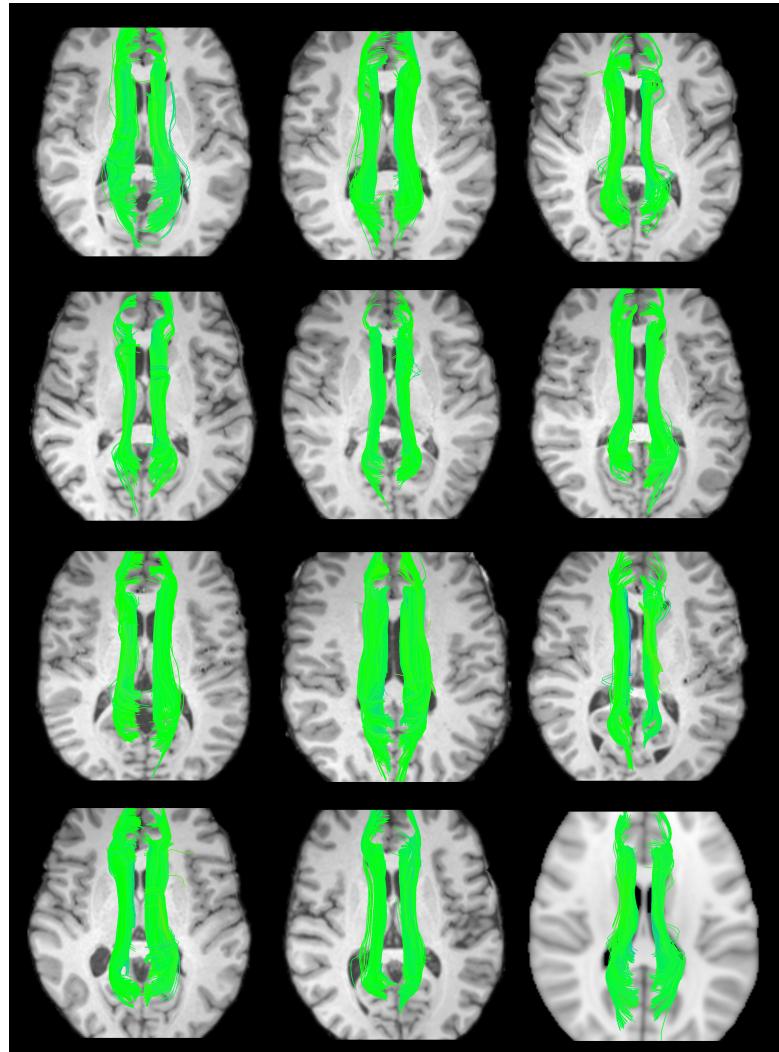


Figure 3.6: The left and right Cingulum (CG) bundles of 12 healthy subjects are presented here. CGs were selected in MNI space from entire EuDX tractophies of 1 million seeds and GQI with sampling length $\lambda = 1.2$. We also show the T1 image slice Z=10 for every subject except from last subject whos T1 was not available. For that subject we use the standard MNI template.

and right CG bundles were consistent across all 12 subjects (see Fig. 3.6). In this atlas, the cingulum was selected using the visualization tool described before from the entire tractography generated with 1 million seeds. The reconstruction parameters and EuDX parameters were the same as before with the exception of the number of initial seeds. Furthermore the tractographies of all subjects were registered together with the structural imagines in MNI space. The only exception was with last subject 12 whos T1 image was not available and it was visualized together with the MNI standard template (MNI152_T1_1mm_brain) (see bottom-right corner of Fig. 3.6).

SLF PROBTRACKX VERSUS EUDX - Try to use with all 3 EITL, DSI and GQI.

ADD Timing experiment.

XXXXX

3.5 Conclusion

We showed that EuDX is a fast deterministic tractography method which can be used to propagate in crossing and non-crossing areas in simulations and human brain data sets.

With the help of expert neuroanatomists we confirmed that EuDX can be used to find bundles which are known from anatomy like AC, CST and we investigated how CG looked between different subjects.

Currently the neuroanatomists are continuing segmenting more and more labels and we hope that in the near future we will have a large collection of segmented bundles which could be used to create better tractography algorithms by understanding better the underlying data sets or exploit supervised learning techniques in order to find the same bundles in other subjects.

The purpose of EuDX is to be faithful to the reconstruction results rather than try to correct or enhance them by introducing regional or global considerations which is the topic of other methods which were reviewed earlier. Therefore, EuDX serves mainly as a robust method for quickly inspecting different reconstruction results using streamlines.

EuDX is will stop tracking on voxel with low anisotropy and will not take into consideration other voxels after it as it would happen with other

probabilistic techniques. This property is often useful when validating underlying reconstruction models.

Furthermore we showed that EuDX can take as input many different anisotropic functions e.g. FA, QA, PK etc. and therefore it has no restrictions for the underlying model neither to the number of peaks per voxel.

The source code for EuDX can be found in `dipy.tracking.eudx` which is available at dipy.org

4 Highly Efficient Tractography Clustering

4.1 Overview

Current tractography propagation algorithms can generate very large tractographies which are difficult to interpret and visualize. A clustering of some kind seems to be a solution to simplify the complexity of these datasets and provide a useful segmentation; however most proposed clustering algorithms are very slow and often need to calculate pairwise distances of size $N \times N$ where N is the number of tracks. This amount of comparisons puts a heavy load on clustering algorithms forcing them to be inefficient and therefore impractical for everyday analysis as it is difficult to compute all these distances or even store them in memory. This adds a further overhead to the use of tractography for clinical applications but also puts a barrier on understanding and interpreting the quality of diffusion datasets. We show in this document that a stable on average linear time clustering algorithm exists and that we can generate meaningful clusters in seconds with minimum memory consumption. In our approach we do not need to calculate all pairwise distances unlike most of the other existing methods. Furthermore we can update our clustering online or in parallel. We show that we can generate meaningful clusters ~ 1000 times faster than any other available method even without parallelism and that it can be used to cluster from a few hundred to many millions of tracks. Moreover our method is multi-purpose; its results can either stand on their own to explore the neuroanatomy directly, or the clustering technique can be used as a precursor tool which reduces the dimensionality of the data, which can then be used as an input to other algorithms of higher order complexity, resulting in their greater efficiency. Beyond the use of this algorithm to simplify tractographies, we show how it can help identify hidden structures, find landmarks, create atlases, compare and register tractographies.

4.2 Track distances and preprocessing

For clarity we first give brief details of various metrics for distances between tracks as they are integral to an understanding of the track clustering liter-

ature. Numerous distance metrics between two trajectories have been proposed in the literature, such as in [56], [130], [212] with the most common being the Hausdorff distance found in [39] and many other studies. We mainly use a very simple symmetric distance proposed in [78] and [192] which we call Minimum average Direct-Flip (MDF) distance $d_{\text{mdf}}(s_A, s_B)$ between track s_A and track s_b , see Eq. (42). This distance can be applied only when both tracks have the same number of points. Therefore we assume that an initial downsampling of tracks has been implemented, where all segments on a track have the same length, and all tracks have the same number of segments. Under that assumption MDF is defined as:

$$\begin{aligned} d_{\text{mdf}} &= \min(d_{\text{d}}, d_{\text{f}}), \text{ where} \\ d_{\text{d}}(s_A, s_B) &= \frac{1}{K} \sum_{i=1}^K \|x_i^A - x_i^B\|_2 \text{ and} \\ d_{\text{f}}(s_A, s_B) &= \frac{1}{K} \sum_{i=1}^K \|x_i^A - x_{K-i}^B\|_2 \end{aligned} \quad (42)$$

where K is the number of points x_i on the two tracks A and B .

In some cases it is still valid to use a family of Hausdorff distances which for simplicity we denote as MAM distances – short for Minimum, or Maximum, or Mean, Average Minimum distance (MAM). We mostly use the Mean version of this family, see Eq. (46) but the others are potentially useful as they can weight different properties of the tracks. These distances are slower to compute but they can work with different number of segments on tracks that is useful for some applications. The equations below show the

formulation of these distances:

$$d_{\text{avg}}(s_A, s_B) = \frac{1}{K_A} \sum_{i=1}^{K_A} d(x_i^A, s_B),$$

$$d_{\min}(s_A, s_B) = \min_{j=1, \dots, K_B} d(x_j^A, s_B), \text{ and} \quad (43)$$

$$d_{\max}(s_A, s_B) = \max_{j=1, \dots, K_B} d(x_j^A, s_B) \text{ where} \quad (44)$$

$$d(x, s_B) = \min_{j=1, \dots, K_B} \|x - x_j^B\|_2.$$

$$\text{MAM}_{\min} = \min(d_{\text{avg}}(s_A, s_B), d_{\text{avg}}(s_B, s_A)) \quad (45)$$

$$\text{MAM}_{\max} = \max(d_{\text{avg}}(s_A, s_B), d_{\text{avg}}(s_B, s_A))$$

$$\text{MAM}_{\text{avg}} = (d_{\text{avg}}(s_A, s_B) + d_{\text{avg}}(s_B, s_A))/2 \quad (46)$$

where the number of points K_A and K_B on the two tracks are not necessarily the same. For the same threshold value MAM_{\min} , MAM_{\max} and MAM_{avg} will give different results. For example, MAM_{\min} will bring together more short tracks with long tracks than MAM_{\max} and MAM_{avg} will have an in between effect. Finally, other distances than the average minimum based on the minimum see Eq. (43) or maximum distance see Eq. (44) can be used. However, we have not investigated them in this work in relation to clustering algorithms.

Coming back to the MDF distance (see Eq.42), its main advantages are that it is fast to compute, it takes account of track direction issues through consideration of both direct and flipped tracks, and that it is easy to understand how it will behave, from the simplest case of parallel equi-length tracks to the most complicated with very divergent tracks. Another advantage is that it will separate short tracks from long tracks and as we will see this will be a good way to find broken or erroneous tracks. An advantage of having tracks with the same number of points is that we can easily do pairwise calculations on them; for example add two or more tracks together to create a new average track . We will see in the next section that track addition is a key property of our clustering algorithm. Some care should be taken into consideration with the number of points allowed in a track (track down-sampling). We always keep the endpoints intact and then downsample in

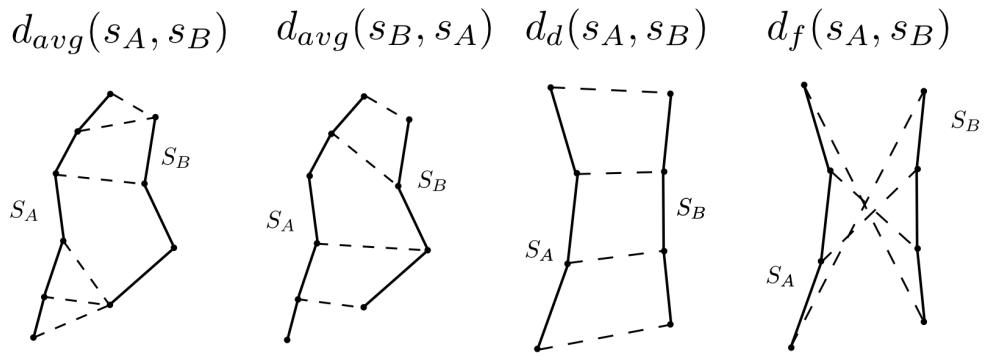


Figure 4.1: Distances used in this work. The main distance used is minimum average direct flip (MDF) distance $d_{df} = \min(d_d, d_f)$ which is a symmetric distance that can deal with the track bi-directionality problem and works on tracks which have the same number of points. Another distance is the mean average distance which is again symmetric but does not need for the tracks to have the same number of points $MAM_{avg} = (d_{avg}(s_A, s_B) + d_{avg}(s_B, s_A))/2$. In this figure the components of both distances are shown; with solid lines we draw the tracks, and then with dashed lines we connect the pairs of points of the two tracks whose distances contribute to the overall metric.

equidistant segments. For example, short tracks having the same number of points as long tracks means that more of the curvature from the long tracks will be lost relative to the short tracks i.e. the short tracks will have higher resolution. We found empirically that this is not an important issue and that for clustering purposes even downsampling to only 3 points in total could be useful [78]. Depending on the application less or more points can be used.

4.3 Related Work

During the last 10 years there have been numerous efforts from many researchers to address the unsupervised and supervised learning problems of brain tractography. As far as we know all these methods suffer from low efficiency, however they provide many useful ideas which we describe in this section.

Tractography clustering algorithms are rarely compared in the literature. The work of Mobergs et al. in [141] was an exception to this. They evaluated different popular hierarchical clustering methods including a less common one, shared nearest neighbor (SNN), against a gold standard segmentation by physicians. The authors concluded that single-link clustering with mean average distance was the method which performed best. Wang et al. [194] proposed a nonparametric Bayesian framework using hierarchical Dirichlet processes mixture model (HDPM). This is one of the very few methods not based on distances. A track is modeled in this work as a discrete distribution over a codebook of discretized orientations and voxel regions. In this paper the authors explain that calculating pairwise distances is very time consuming and therefore they avoid using them. Their approach automatically learns the number of clusters from data with Dirichlet processes priors but it is still not efficient enough for real time operation. Another disadvantage is that their method does not incorporate anatomical knowledge.

Visser et al. [192] used hierarchical clustering and fuzzy c-means together with recombination of subsets of the same tractography to reduce the effect of the large datasets on the distance matrix based on the MDF distance (see section 4.2)[78]. An interesting result with this method was that they could

automatically find the different sub-bundles of the arcuate fasciculus region in accordance with the supervised labeling described in [31]. The algorithm that we present in this chapter also uses the minimum average flip (MDF) metric as a measure of distance between tracks. Gerig et al. [79] also used hierarchical clustering with a symmetrised version of closest point distances, MAM_{avg} and MAM_{max} (Hausdorff). However they tested their method only with two bundles: uncinate fasciculus and the corticospinal tract.

Guevara et al. [84] combined a great number of different algorithms from hierarchical clustering to 3D watershed on track extremities. They first divide the tractography into left-right hemisphere, inter-hemispheric and cerebellum subsets, then create further subsets of different track length, use hierarchical clustering based on the random voxel parcels, use watershed over extremities and finally use hierarchical clustering to merge the different sub-bundles using the Hausdorff distance (see section 4.2). An interesting point with this work is that it stressed the need to divide the data set between shorter and longer tracks. Tsai et al. [187] used a combination of cluster methods based on minimum spanning trees, locally linear embedding and k-means. They showed that they could incorporate both local and global structures by changing a few parameters. The main advantage of this method was that it showed a way to merge a chain of neighbouring structures into one cluster. Zhang and Laidlaw [212] used an agglomerative hierarchical clustering using the same distance as in [211] and later in [210] combined distance-based single linkage hierarchical clustering with expert labeling of specific bundles. Zvitia et al. [217], [218] used adaptive mean shift which is a clustering algorithm that finds automatically the number of clusters for example in contrast with k-means that the user needs to prespecify the number of clusters. They also used this approach for direct registration of tractographies but only from the same subject. El Kouby et al. [66] created a ROI-based connectivity matrix where every i, j point of the matrix holds the number of tracks which connect ROI_i to ROI_j . K-means was used afterwards on the rows of the matrix to cluster the tracks. They showed that their technique can be used for clustering bundles across subjects.

Brun et al. [23] used the mean and covariance of the track as the feature space and normalized cuts based on a graph theoretic approach for

the segmentation. Ding et al. [57] used k-nearest neighbours, another agglomerative approach, applied to corresponding track segments. Corouge et al. [39] used different types of track distances, e.g. Hausdorff distances, and other geometric properties such as torsion and curvature, and in [38] and [40] used Generalized Procrustes Analysis and Principal Components Analysis (PCA) to analyze the shape of bundles.

O'Donnell et al. [148] generated a tractographic atlas using spectral embedding and expert anatomical labeling and then automatically segmented using again spectral clustering and expressed the tracks from as points in the embedded space to the closest existing atlas clusters. The full affinity matrix was too big to compute therefore they used the Nystrom approximation: working on a subset and avoid generating the complete affinity/distance matrix. Later in [151] they tried group analysis on prespecified bundles.

Maddah et al. [129] used B-spline representations of fibre tracts referenced to an atlas, and then the subject's fibre tracts were clustered based on the labeled atlas of the fibre tracts. Later Maddah et al. [128] using a similar track representation (quintic B-splines) calculated a model for each bundle as the average and standard deviation of that parametric representation. In that way created an atlas which is used as a prior for expectation maximization (EM) clustering of the corpus callosal tracks into Witelson subdivisions [203] using population averages. Later in [131] Maddah et al. showed that they could combine spatial priors with metrics for the shape of the tracks to guide tractography clustering.

Jonasson et al. [104] created a large $N \times N$ co-occurrence matrix, where N is the number of the fibers to cluster. The co-occurrence (affinity) matrix contained the number of times two fibers share the same voxel. They then used spectral clustering. Jianu et al. [102] presented a new method for visualizing and navigating through tractography data combining dendograms from hierarchical clustering along with 3d- and 2d-embeddings using the approximation that Chalmers [33] gives for the technique of Eades [62].

Durrleman et al. [61] introduced electrical current models of fibre bundles where a fibre is seen as a set of wires sending information in one direction at constant rate. Currents have good diffeomorphic properties and can be used for registration of bundles as shown in [61] and later in [60].

This methodology does not impose point-to-point or fibre-to-fibre correspondences, however it is sensitive to fibre density and orientation of the bundles and it is computationally expensive.

Leemans and Jones [120] used affinity propagation (section 4.9) to cluster the fronto-occipital fibres, Cingulum and Arcuate Fasciculus after they have reduced the complexity of the data sets using additional frontal and occipital boolean masks on the right cerebrum. However, the authors showed results on a very small part of the entire tractography were clustering is a much easier problem. Later Malcolm et al. [133] used affinity propagation to cluster a full brain tractography created using filtered tractography and suggested that affinity propagation is not suitable for group clustering.

Ziyan et al. [216] introduced a probabilistic registration and clustering algorithm based on expectation maximization (EM) which creates a sharper atlas from a set of subjects on three bundles: Corpus Callosum, Cingulate and Fornix. This work used an initial spectral clustering [148] to label the bundles and then updated these labels iteratively while performing bundle-wise registration combined using polyaffine integration.

Often it is useful to use some protocols in order to add prior information to the automated learning process. Protocols to manually label 11 major white matter tracts were described in Wakana et al. [193] using ROIs to include or exclude tracks generated by deterministic tractography and Hua et al. [95] used regions of interest together with probabilistic tractography in order to create probability maps of known fiber bundles.

From this short review we observe that there are two main trends in the literature. The first which is much more common is to use track distances and calculate distance matrices. In this case deciphering the distance matrix with Hierarchical Clustering or Spectral Embedding are two of the most prevailing approaches. The second case which is much less common is avoid calculating track distances because the computation of the distance matrix is memory intensive. In this case using Dirichlet Processes or Currents or Connectivity based parcelation seem to be some viable solutions. However, if we want to apply clustering in clinical usage or to make practitioners and neuroscientists' analysis less time consuming we need algorithms that can provide useful clusters and cluster descriptors in minimum time and with

low memory usage. None of the papers described in this literature review provide a solution for this issue of efficiency and most of the methods would need from many hours to many days to run on a standard sized data set. The method we propose in this document can provide a solution to this problem and it is an extensive update of our preliminary work found in Garyfallidis et al. [78].

In summary, we can see that most authors agree that unsupervised learning with tractographies is a difficult problem because the data sets are very large, dense, cluttered with noisy tracks which could have no anatomic relevance and bundles which are more than often tangled together in many areas. Furthermore, we can observe that there is a strong disagreement on the number of clusters (from 10 to 60). Because of the difficulty of the problem an international contest was also organized by SchLab in Pittsburgh University (PBC Brain Connectivity Challenge - IEEE ICDM) in 2009. However, the competition didn't conclude to any directly viable solutions. We think that in order to find big clusters a lot of anatomical prior knowledge needs to be introduced in a way that is not yet established. Nevertheless, the clustering that we propose concentrates on reducing the complexity of the data rather than finding bundles with anatomical relevance. We think that this step is more useful at this stage of tractography analysis research.

4.4 Data sets

We experimented with QuickBundles using simulations, 10 human tractographies collected and processed by ourselves and one tractography with segmented bundles which was available online.

Simulated trajectories. We generated 3 different bundles of parametric paths samples at 200 points. The tracks were made from different combinations of sinusoidal and helicoidal functions. In total this data set contained 450 tracks see Fig. 4.5.2 .

Human subjects. We collected data from 10 healthy subjects at the MRC-CBU 3T scanner (TIM Trio, Siemens), using Siemens advanced diffusion work-in-progress sequence, and STEAM [140, 135] as the diffusion preparation method. The field of view was $240 \times 240 mm^2$, matrix size 96×96 ,

and slice thickness 2.5mm (no gap). 55 slices were acquired to achieve full brain coverage, and the voxel resolution was $2.5 \times 2.5 \times 2.5\text{mm}^3$. A 102-point half grid acquisition[205] with a maximum b -value of 4000s/mm^2 was used. The total acquisition time was $14'21''$ with TR=8200ms and TE=69ms. The experiment was approved by the local ethical committee CPREC.

For the reconstruction of the real data sets we used Generalized Q-sampling with diffusion sampling length 1.2 and for the tractography propagation we used EuDX (euler integration with trilinear interpolation) with 1 million random seeds, angular threshold 60° , total weighting 0.5, propagation step size 0.5 and anisotropy stopping threshold 0.0239 (see Fig. 4.8.2,4.11).

PBC human subjects. We also used a few labeled data sets (see Fig.4.5.2,4.5), from the freely available tractography database used in the Pittsburgh Brain Completion Fall 2009 ICDM pbc.lrdc.pitt.edu.

4.5 QuickBundles (QB) Clustering

4.5.1 The QB Algorithm

QuickBundles (QB) is a notably fast algorithm which can simplify tractography representation in an accessible structure in a time that is linear in the number of tracks N . QB is a linear time $O(N)$ (see section 4.5.3) distance based clustering algorithm that we created in order to clarify huge trajectory data sets such as those produced by current state-of-the-art tractography generation algorithms [160, 197]. In general, there are very few linear time clustering algorithms. Just two are well known: CLARANS [147] and BIRCH [213]. QB is different from both of these methods; we will motivate it by describing some aspects of BIRCH as a starting point for the presentation of QB.

BIRCH has two key components: first is relatively simple and involves the use and updating of clusters' descriptors; second is the construction of a tree structure in which the accumulated clusters are held. This second component is aimed at maintaining efficient searchability of the database while balancing what is kept in memory and what is on disc for very large databases. BIRCH uses clustering descriptors which are available for each item in the dataset; these are specific vectors of a fixed dimension of numer-

ical values. Each cluster in turn has a descriptor which is an aggregate of the properties of the items that belong to it (e.g. the sum or mean of the individual clustering feature vectors). Proceeding by a single sweep through the dataset, items are adjoined to clusters on the basis of their proximity to the clusters, subject to a maximum cluster size, or they are added as new leaves into the hierarchical tree structure in which the evolving clusters are held. There then follow updating steps which can involve the merging off previously created clusters in a k-means fashion[181, 126].

It is the linear nature of BIRCH combined with the fixed dimensionality of its cluster descriptors that makes it quite fast. However the further steps involving reorganisation of the accumulated tree do add some major overheads to BIRCH’s performance. QB capitalises on these positive features but does not try to create any kind of hierarchical structure for the clusters. While items in BIRCH are fixed dimension vectors with no additional structure, in QB each item (track) is a fixed-length ordered sequence of points in \mathbb{R}^3 , and uses metrics and amalgamations which take account of and preserve this structure. Moreover each item is either added to an existing cluster on the basis of a distance between the cluster descriptor of the item and the descriptors of the current set of clusters. Clusters are held in a list which is extended according to need.

QB creates an online list of cluster nodes. The cluster node is defined as $c = \{I, \mathbf{h}, n\}$ where I is the list of the integer indices of the tracks in that cluster, \mathbf{h} is an $p \times 3$ matrix, which the most important descriptor of the cluster, and n is the number of tracks on that cluster. \mathbf{h} is a matrix which can be updated online when a track is added to a cluster and is equal to

$$\mathbf{h} = \sum_{i=1}^n s_i \quad (47)$$

where s_i is the $p \times 3$ matrix representing track i , Σ here represents matrix addition, and n is the number of tracks in the cluster. QB assumes that all tracks have the same number of points p , therefore a downsampling of tracks, typically equidistant, is necessary before QB starts. A short summary of the algorithm goes as follows.

Select the first track s_0 and place it in the first cluster $c_0 \leftarrow \{0, s_0, 1\}$.

Then for all remaining tracks (i) goto next track s_i ; (ii) calculate MDF distance between this track and virtual tracks of all existing clusters c_k , where a virtual track is defined on the fly as $\mathbf{v} = \mathbf{h}/n$; (iii) if the minimum MDF distance is smaller than a distance threshold thr add the track to the cluster $c_j = \{I, \mathbf{h}, n\}$ with the minimum distance and update $c_j \leftarrow \{I \cup \{i\}, \mathbf{h} + s, n + 1\}$; otherwise create a new cluster $c_{|C|+1} \leftarrow \{0, s_i, 1\}$, $|C| \leftarrow |C| + 1$ where $|C|$ denotes the current total number of clusters.

Flipping can become an issue when using the MDF distance and adding tracks together, because tracks do not have a preferred direction. A step in QB takes account of the possibility of needing to perform a flip of a track before adding it to a representative track. The complete QB algorithm is described in detail in Alg.5 and a simple step by step visual example is given in Fig.4.5.1. One of the reasons why QB has on average linear time complexity derives from the structure of the cluster node: we only save the sum of current tracks in the cluster and this is achieved cumulatively. By contrast if we were using k-means at every iteration we would have to re-assign tracks to clusters and recalculate averages which is computationally much more intensive. Other reasons are that QB passes through the tracks only once and that a track is assigned to one cluster only.

QB is can be extended for specific applications to contain more information about the clusters. For example we could redefine $c = \{I, \mathbf{h}, n, \mathbf{h}^{(2)}\}$ to obtain second order information and in that way we could calculate the variance of the cluster where

$$\mathbf{h}^{(2)} \leftarrow \{\sum_{i,j} x_{ij}^2, \sum_{i,j} y_{ij}^2, \sum_{i,j} z_{ij}^2, \sum_{i,j} x_{ij}y_{ij}, \sum_{i,j} y_{ij}z_{ij}, \sum_{i,j} x_{ij}z_{ij}\}$$

and x_{ij}, y_{ij}, z_{ij} are the coordinates of the j th point of the i th track in the cluster. Although this alternative would be very useful, as even more refined cluster distances could be used which take into account the additional information, this is not addressed in this document.

One of the disadvantages of most clustering algorithms is that they give different results with different initial conditions; for example this is recognised with k-means, expectation-maximization[47] and k-centers[83] where it is common practice to try a number of different random initial configura-

Algorithm 5 QuickBundles

Input tracks $T = \{s_0, \dots, s_i, \dots, s_{|T|-1}\}$, threshold θ
Output clustering $C = \{c_0, \dots, c_k, \dots, c_{|C|-1}\}$ where cluster $c = \{I, \mathbf{h}, N\}$

```
 $c_0 = \{0, s_0, 1\}$ 
 $C = \{c_0\}$  # the first track becomes the first cluster
 $|C| = 1$  # the total number of clusters is 1
For  $i$  From 1 To  $|T| - 1$  Do # all tracks
     $\mathbf{t} = T_i$ 
    alld = 0 # distance buffer
    flip = 0 # flipping check buffer
    For  $k$  From 0 To  $|C| - 1$  # all clusters
         $\mathbf{v} = C_k.\mathbf{h}/C_k.N$ 
         $d = d_d(\mathbf{t}, \mathbf{v})$ 
         $f = d_f(\mathbf{t}, \mathbf{v})$ 
        If  $f < d$  Then
             $d = f$ 
            flip $_k$  = 1
            alld $_k$  =  $d$ 
         $m = \min(\text{alld})$ 
         $l = \text{argmin}(\text{alld})$ 
        If  $m < \theta$  # append in current cluster
            If  $\text{flip}_l = 1$  Then
                 $C_l.\mathbf{h}+ = t'$ 
            Else
                 $C_l.\mathbf{h}+ = t$ 
                 $C_l.N+ = 1$ 
                 $C_l.I.append(i)$ 
        Else # create new cluster
             $|C|+ = 1$  # total number of clusters increases
             $C_{|C|-1}.I_0 = l$ 
             $C_{|C|-1}.\mathbf{h} = \mathbf{t}$ 
             $C_{|C|-1}.N = 1$ 
```

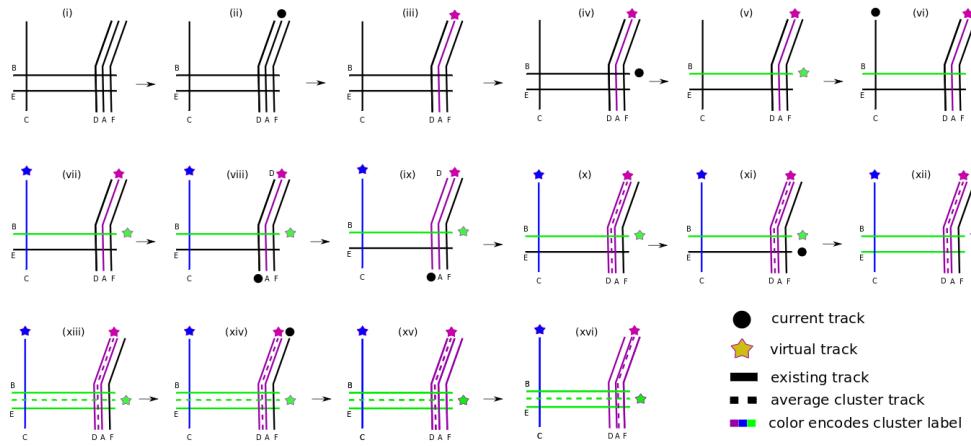


Figure 4.2: QB step-by-step: Initially in panel (i) 6 unclustered tracks (A-F) are presented; imagine that the distance threshold used is the MDF distance (Eq. 42) between B and E. The algorithm starts and in (ii) we see that track A was selected, so no other clusters exist therefore track A becomes the first cluster (labeled with purple color) and the virtual track of that cluster is identical with A as seen in (iii), next in (iv) track B is selected and we calculate the MDF distance between B and the virtual track of the other clusters. For the moment there is only one cluster to compare so QB calculates MDF(B,virtual-purple) and this is obviously bigger than threshold (that being $MDF(B,E)$) therefore a new cluster is assigned for B and B becomes the virtual track of that cluster as shown in (v). In (vi) the next track is selected and this is again far away from both purple and blue virtuals therefore another cluster is created and B is the virtual of the blue cluster as shown in (vii). In (viii) track D is current and after we have calculated $MDF(D,purple)$, $MDF(D,Blue)$ and $MDF(D,green)$ it is obvious that D belongs to the purple cluster as $MDF(D,purple)$ is smaller and lower than threshold as shown in (ix). However we can now see in (x) that things change for the purple cluster because the virtual track is not anymore made by only one track but it is the average of D and A shown with dashline. In (xi) E is the current track and will be assigned at the green cluster as shown in (xii) because $MDF(E,virtual green) = MDF(E,B) = \text{threshold}$, and in (xiii) we see the updated virtual track for the green cluster which is equal to $(B+E)/2$ where + means track addition. In (xiv) the last track is picked and compared with the virtual tracks of the other 3 clusters; obviously $MDF(F,purple)$ is the only with smaller threshold, and so F is assigned to the purple cluster in (xv). Finally, in (xvi) the virtual purple track is updated as $(D+A+F)/3$. As there are no more tracks to select, the algorithm stops. We can see all three clusters have been found and all tracks have been assigned successfully.

tions. The same holds for QB so if there are not distinct clusters such that the distance between any pair of clusters is supra-threshold, then with different permutations of the same tractography we will typically see similar number of clusters but different underlying clusters. We will examine the robustness of QB in this respect in section 4.6.

4.5.2 Powerful simplifications

One of the major benefits of applying QB to tractographies is that it can provide meaningful simplifications and find structures that were previously invisible or difficult to locate because of the high density of the tractography. For example we used QB to cluster the corticospinal tract (CST). This bundle was part of the datasets provided by the Pittsburgh Brain Competition (PBC2009-ICDM) and it was selected by an expert. The result is clearly shown in Fig.4.5.2 where every partition is represented by a virtual track. To generate this clustering we used a tight threshold of 10mm. We observe that only a few virtual tracks travel the full distance from bottom to top and that there are many tracks that are broken (i.e. shorter than what was initially expected) or highly divergent.

Another interesting feature of QB is that it can be used to merge or split different structures by changing the distance threshold. This is shown in Fig. 4.5.2; on the left we see simulated paths made from simple sinusoidal and helicoidal functions packed together. The colour coding is used to distinguish the three different structures. With a lower threshold the three different structures keep remain separated but when we use a higher threshold the red and blue bundles are represented by only one cluster; represented by a purple virtual.

Similarly, with the simulations shown in Fig.4.5.2 we can see the same effect on real tracks, e.g. those of the fornix shown at the left panel of Fig. 4.5 where we can obtain different number of clusters at different thresholds. In that way we can stress thinner or larger sub-bundles inside other bigger bundles.

A full tractography containing 250,000 tracks was clustered using QB with a distance threshold of 10mm (Fig. 4.5). We produced a useful re-

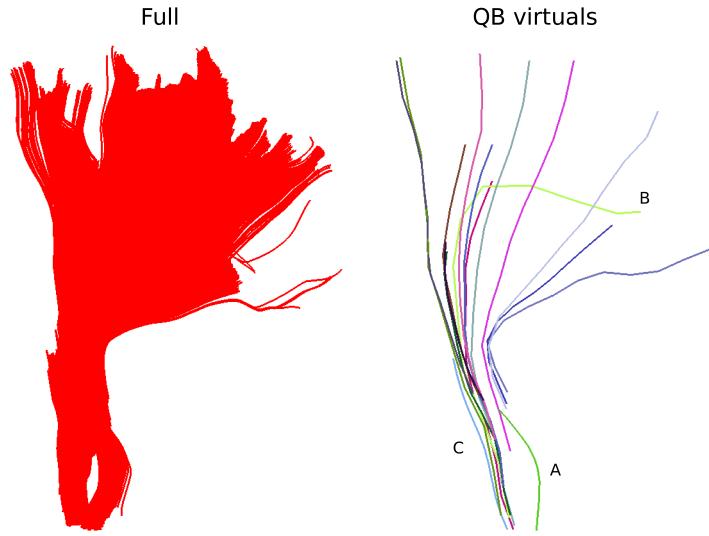


Figure 4.3: This is a part of the CST bundle consisting of 11041 tracks merged by an expert (PBC2009 data) shown with red color. Visually it looks as though all tracks have a similar shape, and possibly merge towards the bottom, and spreading towards the top. However, this is a misreading caused by the opaque density when all the tracks are visualised. QB can help us see the full structure of the bundle and identify its elements. Here on the right hand side we see a simplification (virtual tracks) of the red CST generated by running QB with distance threshold of 10 mm and downsampling to 12 points. We can now clearly see that lots of parts which looked homogeneous are actually broken bundles e.g. dark green (A), light blue (C) or bundles with very different shape e.g. light green virtual track (B). To cluster this bundle took 135 ms $\simeq 0.14$ seconds.

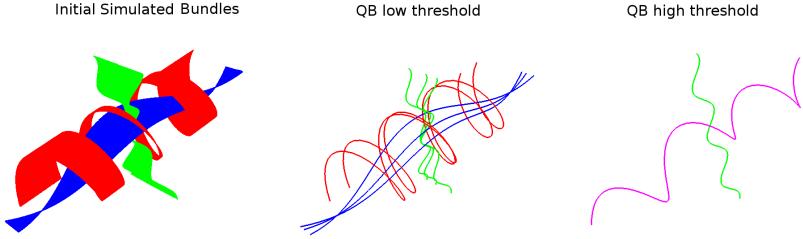


Figure 4.4: On the left we see 3 bundles of simulated trajectories; red, blue and green consisting of 150 tracks each. All 450 tracks are clustered together using QB and the virtual tracks are shown when threshold 1 was used shown in the middle and 8 on the right. We can see that when the threshold is low enough the underlying structure is a more detailed representation of the underlying geometry. However when the distance threshold is higher closer bundles could merge together as seen in the result on the right panel where the red and blue bundle have merged together in one cluster represented by the purple virtual track.

duction of the initial tractography leaving only 763 virtual tracks. Bundles smaller than 10 tracks were removed. Every track shown here represents an entire cluster containing from 10 to 5000 tracks each.

The virtual tracks can be thought as fast access points to explore the entire data set (see Fig. 4.5).

4.5.3 Complexity and timings

To apply QB to a data set we need to specify three key parameters: p , the fixed number of downsampled points per track; θ the distance threshold, which controls the heterogeneity of clusters; and N the size of the subset of the tractography on which the clustering will be performed. When θ is higher, fewer more heterogeneous clusters are assembled, and conversely when θ is low, more clusters of greater homogeneity are created.

The complexity of QB is in the best case linear time $O(N)$ with the number of tracks N and worst case $O(N^2)$ when every cluster contains only one track. The average case is $O(MN)$ where M is the number of clusters however because M is usually much smaller than N ($M \ll N$) we can neglect M and denote it only as $O(N)$ as it is common in complexity theory. We



Figure 4.5: Left - Here we see how QB clustered the fornix bundle with the dataset from the PBC2009 competition. The original fornix is shown in black consists of 1076 tracks. All tracks were equidistantly downsampled at 3 points in this example. With a 5mm threshold our method generates 22 clusters (top right). With 10mm generates 7 (bottom left) and with 20mm the whole fornix is determined by one cluster only (bottom right). Right - an example of a full tractography (250,000 tracks) being clustered using QB with a distance threshold of 10mm. Here we see just 763 virtual tracks depicted which produce a useful simplification of the initial tractography. Every track shown here represents an entire cluster from 10 to 5000 tracks each. These can be thought as fast access points to explore the entire data set. The colour here just encodes track orientation. With an appropriate visualization tool you could click on a track and obtain the entire cluster/bundle that it represents. Visualizing an entire data set of that size is impossible on standard graphic cards and most visualization tools e.g. Trackvis or DSI Studio can only show a small random sample of the full tractography at real time.

Number of tracks (N)	Algorithms	Timings (secs)	QB (secs)	Speedup
1000	Wang et al. [194]	30	0.07	429
60,000	Wang et al. [194]	14400	14.7	980
400,000	Visser et al. [192]	75000	160.1	468

Table 4.1: QB run on $p = 12$ point tracks and distance threshold at $\theta = 10\text{mm}$ compared with some timings reported from other state of the art methods found in the literature. Unfortunately timings were very rarely reported until today as most algorithms were very slow on full data sets. Nonetheless, we can observe in this table that the speedup that QB offers is substantial, holding out the prospect of real-time clustering on data sets containing fewer than $\sim 20,000$ tracks. This experiment was executed on a standard PC using only a single CPU core.

created the following experiment to investigate this claim and we found empirically that the average case is actually $O(N)$ for tractographies (see Fig.4.5.3,4.5.3). In this experiment we timed the duration of QB clustering of tractographies containing from 100 thousand to 1 million tracks, with different initial number of points per track (3, 6, 12 and 18) and different QB thresholds (10.0, 15.0, 20.0, 25.00 mm). The results can be seen in the diagram of Fig.4.5.3 and 4.5.3. You can notice that even when the the threshold value becomes impressively low (10.0 mm) the linearity is only slightly disturbed.

Furthermore, the memory usage of QB is $O(M)$ where M is the number of clusters and because this is usually much smaller than N we consider memory consumption to be negligible. Because in QB we store only the indices of the tracks even for very large tractographies 20 or more clusterings can be stored simultaneously in the RAM of a simple notebook without any problems. Therefore, another feature of QB is memory efficiency.

We compared QB with 12 point tracks and distance threshold at $\theta = 10\text{mm}$ versus some timings reported from other state of the art methods found in the literature (Tab. 4.1). Unfortunately timings were very rarely reported up till now as most algorithms were very slow on full data sets. Nonetheless the speedup that QB offers is obviously of great importance and holds out the prospect of real-time clustering on data sets of fewer than 20,000 tracks (see Tab. 4.1).

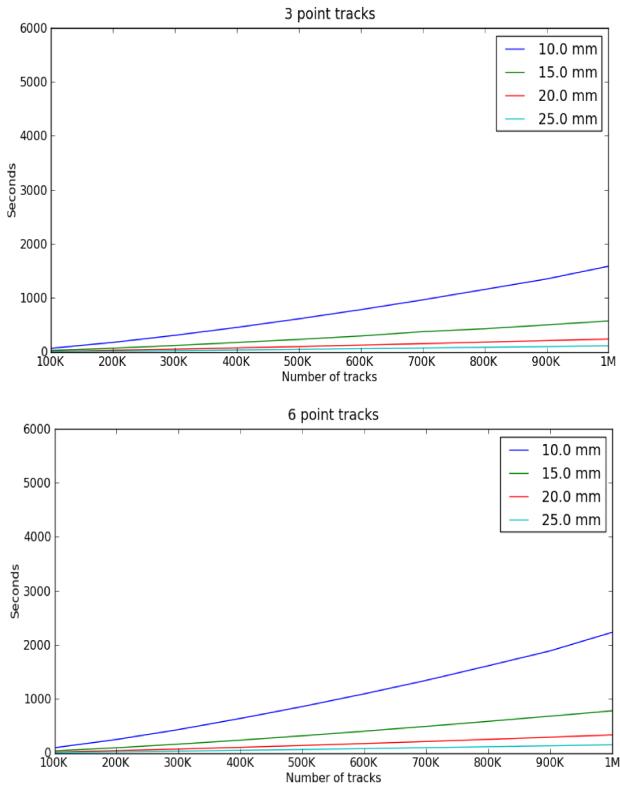


Figure 4.6: QB is a very efficient algorithm whose performance is controlled by just three parameters. The first is the number of tracks, a second is the distance threshold in millimeters - shown with different colours and another is the amount of initial downsampling of the initial trajectories. A last parameter not shown in these diagrams is the underlying structure of the data which is expressed by the number of final clusters. We used a full tractography to generate these figures without removing or preselecting any parts. This results run of a single thread Intel(R) Xeon(R) CPU E5420 at 2.50GHz on a standard PC. (NB This and the following figure use the same vertical scale to assist direct comparison.)

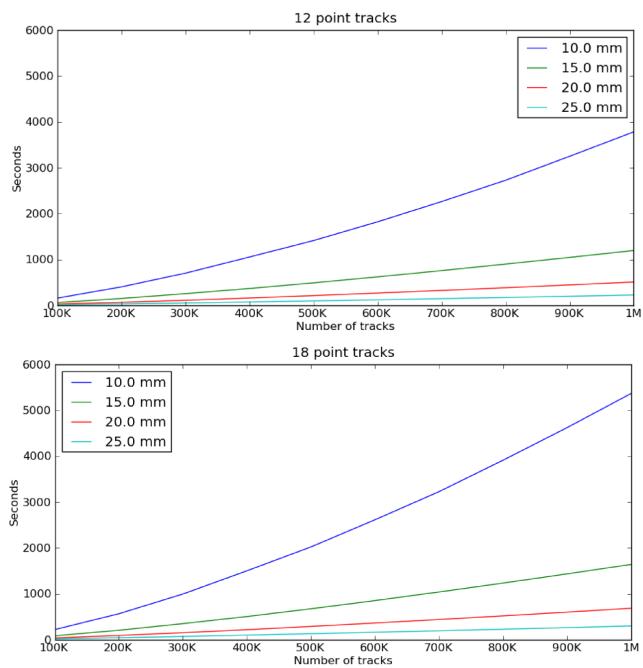


Figure 4.7: Time comparisons of QB using different number of points per track, different distance thresholds and different number of tracks. Same as Fig. 4.5.3. We can observe here how the linearity only reduces slightly even when we use a very low threshold as that of 10mm which can generate many thousand clusters. This experiment concludes that QB is suitable for fast clustering.

4.5.4 Virtual tracks, exemplar tracks and other descriptors.

The virtual tracks created by QB have very nice properties as they represent an average track which can stand as the most important feature of the cluster that they belong to. However, now that we have segmented our tractography into small bundles we can calculate many more potentially important descriptors for the cluster. For instance the Cluster Spread (CS) can be computed for any cluster c as a vector of length p whose j th component is $\sum_{x \in c} |x_j - v_j|^2 / n$. Here x_j is the j -th point in the track x in cluster c , v_j is the corresponding point of the virtual track, and n is the size of the cluster. CS provides a profile of the tightness or looseness of the cluster along the length of the virtual track. Many other similar or higher order statistics can be readily computed in an analogous fashion. One of the most useful features is the calculation of exemplars.

Exemplars. Another fruitful idea relating to the virtual track is to try to identify a corresponding feature for the bundle which actually belongs to the tractography. In other words to find an exemplar or centroid track. Remember that the virtual tracks do necessarily coincide with real tracks as they are just the outcome of large amalgamations. There are many strategies for how to select good exemplars for the bundles. A very fast procedure that we use in this work is to find which real track from the cluster is closest (by MDF distance) to the virtual track. Let's call this exemplar track e_1 such that $e_1 = \arg \min_{x \in C} \text{MDF}(v, x)$. The computational complexity of finding e_1 is still linear in cluster size, and that will be very useful if we have created clusterings with clusters containing more than ~ 5000 tracks (depending on system memory).

A different exemplar can be defined as the most similar track among all tracks in the bundle, which we denote by $e_2 = \arg \min_{x \in C} \sum_{y \in C} \text{MDM}(y, x)$, or if we want to work with tracks with possibly different numbers of points we could instead use $e_3 = \arg \min_{x \in C} \sum_{y \in C} \text{MAM}(y, x)$. Identification of exemplar tracks of type e_2 and e_3 will be efficient only for small bundles of less than 5000 tracks because we need to calculate all pairwise distances in the bundle. We will see next many applications of the exemplars. For example in the

section 4.5.5 we will use them to simplify the bi-directionality problem when merging clusters.

In summary, a virtual (centroid) track is the average of all tracks in the cluster. We call it virtual because it doesn't really exist in the real data set and to distinguish it from exemplar (medoid) tracks which are again descriptors of the cluster but are represented by real tracks.

4.5.5 The bi-directionality problem

Because a track is a sequence of points without a preferred direction, it has two possible orientations when comparing it with another track. Most tractography methods will create tracks with arbitrary directions; meaning that close and similar tracks can have opposite directions. Of course the tracks do not really carry directional information. By direction we mean the encoding of the sequence of points which define the track. Thus a track may be ordered $p_1, p_2 \dots p_{N-1}, p_N$, or $p_N, p_{N-1} \dots p_2, p_1$. We call this the bi-directionality problem. Using the MDF distance we found a way with QB to eliminate this problem. However, if we want to merge clusters together we need to have a way to minimize this problem.

For this purpose we devised the following technique. Choose a fixed point or pole P in the 3D space of the tractography, possibly away from the mid-sagittal plane. Then re-direct all tracks so that the first point of every track is the end closer to P . If the tractography is in native space it suffices to have the origin $(0, 0, 0)$ as the pole point; in MNI space we can use the point $(100, 100, 100)$. It is our empirical experience that this method will redirect correctly most tracks in the sense that similar tracks will have the same direction. However there will still be a small percentage for which the bi-directionality problem persists. We can correct for these by using exemplars rather than virtual tracks as virtual tracks can misrepresent a bundle if the bundle consists of tracks with ambiguous directionality. The exemplars are more preferable than the virtual tracks because of the way the latter can be influenced more by outliers and thus can be less representative in terms of the shape of real tracks in a bundle. The exemplars are similar to the concept of the median and the virtuals more similar to the concept of the mean.

It is well known that the mean is usually influenced more by outliers than the median.

4.6 Comparisons within- and between-subjects

4.6.1 Robustness under reordering

As mentioned above, QB shares the behaviour of most clustering algorithms in that different orderings of the tracks give rise to different clusterings. As a first step towards examining the robustness of QB in this respect we recorded the numbers of QB clusters in 20 different random orderings of the tractographies of 10 human subjects acquired as described in section 4.4. We removed short tracks shorter than 40mm and downsampled the tracks at 12 points. Then we applied QB with threshold at 10mm. The mean number of clusters was 2645.9 (min 1937.6; max 3857.8; s.d. 653.8). There is therefore a considerable between-subject variation in this metric. By contrast the within-subject variability of the number of clusters across random orderings is rather small, with mean standard deviation 12.7 (min 7.3; max 17.4). This suggests an encouraging level of robustness in terms of the numbers of clusters that QB creates. We now consider ways of measuring and comparing the contents of the clusters in a clustering.

4.6.2 Measures to compare classifications

Considerable attention has been paid to measuring the performance of one or more classifiers in the context of supervised learning, see for instance [113]. We now outline some of these metrics before applying them to the comparisons we are interested in. Let $\mathcal{A} = \{A_1, A_2, \dots, A_m\}$ and $\mathcal{B} = \{B_1, B_2, \dots, B_n\}$ be two classifications of N items. Let the number of items in A_i and B_j be a_i and b_j , with t_{ij} items in the intersection $A_i \cap B_j$. There are a number of ways of measuring the similarity or dissimilarity of \mathcal{A} and \mathcal{B} . The first two, Gini Purity and Maximum Likelihood Accuracy, are based on ways we might estimate the \mathcal{A} -labels if we just have the \mathcal{B} -labelling, or *vice versa*.

Purity Suppose we have a probability distribution $P = (p_1, p_2, \dots, p_m)$ such that the probability that any item has label i is p_i . Not knowing what this for any item is we apply ‘probability matching’ and randomly estimate a label from the set $\{1, 2, \dots, m\}$ by random selection using the same distribution P . Then the probability of assigning the correct label is $\sum p_i^2$; this is the Purity of the distribution. The purity of a distribution lies in the range $[\frac{1}{m}, 1]$. The upper limit occurs when P assigns probability 1 to just one label (i.e. a very pure, concentrated distribution); the lower limit occurs when all m labels have equal probability $\frac{1}{m}$. We now extend this to the case when we have some additional information about the item, namely the label that is assigned to it in a different classification \mathcal{B} .

If $P_{\mathcal{A}|B_j}$ is the observed conditional probability distribution ($p_{i|j} = \frac{t_{ij}}{b_j}$, $i = 1, \dots, m$) of \mathcal{A} given B_j , then we define the Purity of \mathcal{A} with respect to \mathcal{B} as $\text{purity}(\mathcal{A}|\mathcal{B}) = \sum_{j=1}^n \frac{b_j}{N} \text{purity}(P_{\mathcal{A}|B_j})$. In terms of the matrix $T = (t_{ij})$ this is the \mathcal{B} -weighted average of the impurities of the rows of T . We similarly define $\text{purity}(\mathcal{B}|\mathcal{A})$ and it is equal to the \mathcal{A} -weighted average of the impurities of the columns of T . In what follows we will use the symmetrised value $\text{purity}(\mathcal{A}, \mathcal{B}) = [\text{purity}(\mathcal{A}|\mathcal{B}) + \text{purity}(\mathcal{B}|\mathcal{A})]/2$.

Maximum probability matching Another way to estimate a label for each item is to assign it the label with maximum probability $i_{\max} = \arg \max p_i$. The Random Accuracy in this case is $p_{i_{\max}} = \max_i p_i$. When we do this conditional on the \mathcal{B} -label and average over those labels, we get the Maximum Probability Matching of \mathcal{A} conditional on \mathcal{B} ,

$$\text{MPM}(\mathcal{A}|\mathcal{B}) = \sum_{j=1}^n \frac{b_j}{N} \max_i p_{i|\mathcal{B}_j}.$$

We define $\text{MPM}(\mathcal{B}|\mathcal{A})$ similarly, $\text{MPM}(\mathcal{B}|\mathcal{A}) = \sum_{i=1}^m \frac{a_i}{N} \max_j p_{j|\mathcal{A}_i}$. A further simplification is to use the symmetrized value

$$\text{MPM}(\mathcal{A}, \mathcal{B}) = [\text{MPM}(\mathcal{A}|\mathcal{B}) + \text{MPM}(\mathcal{B}|\mathcal{A})]/2.$$

Correctness and completeness (splitting and lumping pairs of items) For the next two metrics the focus moves to comparison of the labels assigned by \mathcal{A} and \mathcal{B} to pairs of items. Differences in the partitions \mathcal{A} and \mathcal{B} are reflected in two ways. Items assigned the same label by \mathcal{A} are said to be split by \mathcal{B} if their \mathcal{B} -labels are not equal; alternatively items assigned different \mathcal{A} -labels are said to be lumped by \mathcal{B} if they are assigned the same \mathcal{B} -label. Note that what is lumped (split) by \mathcal{B} will equally be split (lumped) by \mathcal{A} .

The total number of pairs from N items is $\text{pairs}(\mathcal{A}) = \binom{N}{2} = \frac{N(N-1)}{2}$. The number of pairs assigned the same \mathcal{A} -labels is $\text{together}(\mathcal{A}) = \sum_{i=1}^m \binom{a_i}{2}$. The number of pairs assigned different labels is $\text{apart}(\mathcal{A}) = \text{pairs}(\mathcal{A}) - \text{together}(\mathcal{A})$. This can also be written as $\sum_{1 \leq i \neq i' \leq m} a_i a_{i'}$ which in turn can be expressed in terms of the cumulative sum of (a_i) which is an efficient way of programming these calculation of sums of all products with unequal subscripts. The number of \mathcal{A} -pairs split by \mathcal{B} is

$$\text{split}(\mathcal{A}|\mathcal{B}) = \sum_{i=1}^m \left(\sum_{1 \leq j \neq j' \leq n} n_{ij} n_{ij'} \right) = \text{lumped}(\mathcal{B}|\mathcal{A}).$$

Similarly

$$\text{lumped}(\mathcal{A}|\mathcal{B}) = \sum_{j=1}^n \left(\sum_{1 \leq i \neq i' \leq m} n_{ij} n_{i'j} \right) = \text{split}(\mathcal{B}|\mathcal{A}).$$

Completeness and Correctness are defined in terms of these quantities:

$$\text{completeness}(\mathcal{A}|\mathcal{B}) = 1 - \text{split}(\mathcal{A}|\mathcal{B}) / \text{together}(\mathcal{A})$$

and

$$\text{correctness}(\mathcal{A}|\mathcal{B}) = 1 - \text{lumped}(\mathcal{A}|\mathcal{B}) / \text{apart}(\mathcal{A}).$$

Symmetrized measures of completeness and correctness for \mathcal{A} and \mathcal{B} are defined as

$$\text{completeness}(\mathcal{A}, \mathcal{B}) = [\text{completeness}(\mathcal{A}|\mathcal{B}) + \text{completeness}(\mathcal{B}|\mathcal{A})] / 2$$

$$\text{correctness}(\mathcal{A}, \mathcal{B}) = [\text{correctness}(\mathcal{A}|\mathcal{B}) + \text{correctness}(\mathcal{B}|\mathcal{A})]/2.$$

For the clusterings we encounter in tractography typically the number of apart pairs in \mathcal{A} is very high, and only a small percentage (e.g. 0.5%) of these pairs will be lumped by \mathcal{B} . This is because the average cluster size is small by comparison with the number of clusters. As a consequence, the correctness measure is not a particularly useful metric. By contrast the number of together pairs is modest, and the completeness measure is more sensitive.

Maximum Agreement (κ_{\max}) Our fifth metric is Cohen's κ , which is a well-known measure of agreement between raters on the assignment of a set of items to a shared classification scheme. It adjusts the agreements (items on which the raters agree) for the number of agreements that might have occurred by chance:

$$\kappa = \frac{\text{P}_{\text{agreement}} - \text{P}_{\text{chance agreement}}}{1 - \text{P}_{\text{chance agreement}}}.$$

This can be simply represented in terms of the overlap matrix $T = (t_{ij})$ by the formula:

$$\kappa(T) = \frac{\sum_{i=1}^M t_{ii}/N - \sum_{i=1}^M r_i c_i / N^2}{1 - \sum_{i=1}^M r_i c_i / N^2},$$

where r_i and c_j represent the row and column totals of T . We have extended T to a square matrix of size $M = \max(m, n)$ by adding if necessary rows or columns of zeros. When we adapt this measure to the case of comparing two clusterings we further need to take into account the lack of prior correspondence between the two sets of labels. The κ_{\max} statistic is the result of maximising κ over all possible correspondences:

$$\kappa_{\max} = \max_{\pi} \kappa(T_{\pi}) = \frac{\sum_{i=1}^M t_{i\pi(i)}/N - \sum_{i=1}^M r_i c_{\pi(i)}/N^2}{1 - \sum_{i=1}^M r_i c_{\pi(i)}/N^2},$$

where T_{π} is the matrix T with columns reordered by a permutation π . The principal trouble with the κ_{\max} statistic is that its computation is $O(N!)$ if all permutations are tried. One way out of the problem caused by the size of the search set might be to use a randomised search strategy for instance based on a simulated annealing approach.

Matched Agreement via the Hungarian Method An alternative is to look to see if there is a simpler quantity that might be optimised. One obvious choice is the maximized number of agreements $\mu_{\max} = \sum_{i=1}^M t_{i\pi(i)}$ corresponding to the permutation π ; this is the leading term in the numerator of κ_{\max} . Seeking to maximize the number of agreements amongst all permutations π is a classical combinatorial optimization problem (weighted assignment problem on a bipartite graph) that can be reformulated as a linear programming problem whose efficient solution by the Hungarian Method [112] is well known.

We have tried out various published coded implementations of the version by Lawler [115] of the Hungarian Method and have found that the one published by Carpaneto et al. [29] and implemented by them [30] in FORTRAN is both fast and capable of handling assignment problems of unlimited size.

We calculated the average of each of these comparison metrics for QB clusterings of 12 different orderings for each of 10 tractographies (Tab. 4.2). A number of observations are worth making. Matched agreement and matched kappa are essentially the same metric (correlation 0.97). Of these two metric we prefer matched agreement because is both simpler to calculate and also to understand. Correctness, for the reasons discussed above, is too insensitive (near to ceiling) to be of use. We would therefore suggest, and on the basis of the mean of the standard deviations across pairings, that max-

Metric	Purity	MPM	Comp	Corr	MA	MK
Mean	70.8	79.2	65.5	99.9	74.1	74.0
Mean S.D	0.51	0.37	1.11	0.02	0.39	0.39

Table 4.2: Mean and mean standard deviation of six classification comparison metrics for 10 different tractographies: Purity, Maximum Probability Matching (MPM), Completeness (Comp), Correctness (Corr), Matched Agreement (MA) and Matched Kappa (MK). For each of 10 tractographies the 66 pairings of QB clusterings for 12 different orderings were evaluated. All are represented as percentages (%). Matched agreements use the Hungarian Algorithm to create a mapping between each pair of clusters; matched kappa evaluates Cohen’s kappa using this same optimal mapping.

imum probability matching, and matched agreement are suitable metrics for evaluating tractography clusterings. It is also worth noting that maximum probability matching is a simple first approximation to the optimal matching identified by the Hungarian method although it is not necessarily one-to-one.

However we also note that these metrics are all costly to calculate in terms of time and memory requirements. They will not be used further in this study. We instead look at ways to compare clusterings of tractographies that will work when comparing different tractographies either for the same or for different subjects. These need to be based on metrics for distances between tracks, whether virtual tracks, exemplar tracks or raw tracks from the original tractographies. This is the subject of the next section.

4.6.3 Tightness comparisons

We have found rather few systematic ways to compare different clustering results for tractographies in the literature [141]. Being able to compare results of clusterings is crucial for creating stable brain imaging procedures, and therefore it is necessary to develop a way to compare different clusterings of the same subject or different subjects. Although we recognise that this is a difficult problem, we propose the following solution with a metric which we call tight comparison (TC). Tight comparison works as follows. Let us assume that we have gathered the exemplar tracks from clustering A in $E_A = \{e_1, \dots, e_{|E_A|}\}$ and from clustering B in $E_B = \{e'_1, \dots, e'_{|E_B|}\}$ where $|E|$

denotes the number of exemplar tracks of each clustering E . The size of set E_A does not need to be the same as that of E_B (i.e. perhaps $|E_A| \neq |E_B|$ rather than $|E_A| = |E_B|$). Next, we calculate all pairwise MDF distances between the two sets and store them in rectangular matrix D_{AB} . The minima of the rows of D_{AB} provide the distance to the nearest track in B of each track in A ($E_{A \rightarrow B}$) and similarly the minima of the columns of D_{AB} the distance to the nearest track in A of each track in B ($E_{B \rightarrow A}$). From these correspondences we only keep those distances that are smaller than a tight (small) threshold t_{thr} . Then we define TC (Tightness Comparison) to be

$$TC = \frac{1}{2} \left(\frac{|E_{A \rightarrow B} \leq t_{\text{thr}}|}{|E_A|} + \frac{|E_{B \rightarrow A} \leq t_{\text{thr}}|}{|E_B|} \right) \quad (48)$$

where $|E_{A \rightarrow B} \leq t_{\text{thr}}|$ denotes the number of exemplars from A which had a neighbour in B that is closer than t_{thr} and similarly for $|E_{B \rightarrow A} \leq t_{\text{thr}}|$ the number of exemplars from B to A which their distance was smaller than t_{thr} . When $TC = 0$ that means that every exemplar from the one set was further than t_{thr} to all exemplars in the other set. When $TC = 1$ then all exemplars from one set had a close neighbour in the other set. This metric is extremely useful especially when comparing tractographies from different subjects because it does not require $|E_A| = |E_B|$.

We ran an experiment were we compared TC between pairs of 10 subjects with their tractographies warped in MNI space. This generated $\binom{10}{2} = 45$ TC values with $t_{\text{thr}} = 10\text{mm}$. We did this experiment twice; first by keeping only the bundles with more than 10 tracks (TC10) and secondly by keeping only the bundles with more than 100 tracks (TC100). The average value for TC10 was 47% and standard deviation 2.6%. As expected TC100 (bigger landmarks) did better with average value of 53% and standard deviation 4.9%. The difference between TC10 and TC100 is highly significant: Student's $t = 4.692$, $df=88$, $p = 1.97 \times 10^{-5}$, two-sided; and, as a precaution against non-normality of the underlying distributions, Mann-Whitney U = 530., $p = 5.65 \times 10^{-5}$. If we think that the small bundles of size < 100 are more idiosyncratic or possibly more likely to reflect noise in the data, whereas larger bundles are more indicative of substantial structures and landmarks in the tractographies, then we are encouraged to see that on av-

erage the virtual tracks of 50% of larger bundles of each tractography lie within 10mm of those of the other tractographies. This supports the notion that QB can be used to find agreements between different brains by concentrating on the larger (more important) clusters. We will see further evidence for this below (section 4.8.2).

4.7 Parallel version

4.7.1 Algorithm

QB is a very fast algorithm; however we wanted to make it even more efficient so that for example it is trivial to cluster hundreds of subjects together and use many CPUs or computers simultaneously. This could be used to create an atlas of hundreds of subjects in a few hours. Therefore, we have extended QB to a parallel version which we call pQB. This algorithm works as follows. First we redirect and downsample all tracks. Then we put all tracks together and break them into subsets. For every subset we assign a new thread and set QB to run on that thread. Therefore, we have now many QBs running on different CPUs. Then we collect all individual clusterings and start merging them together. We can pair every two results together and merge them in a binary way or just merge all clusterings to the first clustering. We can do merging with many different ways. Here we present the most modest but useful attempt.

4.7.2 Merging two sets of bundles

We can merge bundles using exemplar tracks or virtual tracks. We first set a distance threshold θ usually the same as the one we used for the QBs in the previous step. Let's assume now that we have gathered the virtual tracks from clustering A in $V_A = \{v_0, \dots, v_{|V_A|}\}$ and from clustering B in $V_B = \{v'_0, \dots, v'_{|V_B|}\}$ where $|V|$ denotes the number of virtual tracks of each clustering. $|V_A|$ can be different $|V_B|$. (a) For every v'_i in set V_B we find the closest v_j in set V_A and store the distance between these two tracks. Therefore we now have a set of minimum distances from V_B to V_A . The size of this set is equal to $|V_B|$. (b) Finally, we merge those clusters from B whose vir-

tual tracks have minimum distances smaller than θ into the corresponding clusters of A , and if a virtual track in V_B has no sub-threshold neighbour in V_A then its cluster becomes a new cluster in the merged clustering. In that way clusters from the two sets who have very similar features will merge together and if not new clusters will be created, and we will not have any loss of information from the two sets of clusters.

4.8 Direct applications

We found that QB has numerous applications from detecting erroneous tracks to creating atlases, finding landmarks and guiding registration algorithms. Here we present just a few of the strategies that can be pursued further.

4.8.1 Rapidly detecting erroneous tracks

It is well known that there are different artifacts seen in tractographies caused by subject motion, poor voxel reconstruction, incorrect tracking and many other reasons. However there is no known automatic method to try and detect these tracks and therefore remove them from the data sets. The idea here is to use QB to speed up the search for erroneous tracks. We will concentrate here on tracks that loop one or many times; something that it is considered impossible to happen in nature.

One of the tracks which are most likely erroneous are tracks which wind more than one time, like a spiral. We can detect those with the following algorithm. Lets assume that we have a track s and we want to check if it winds: (a) we perform a singular value decomposition on the centered track $U, \mathbf{d}, V = \text{SVD}(s - \bar{s})$; (b) project the highest singular value \mathbf{d}_0 to the first column of U , U_0 creating the first component of a two dimensional coordinate p_x and the second highest \mathbf{d}_1 to the second column U_1 creating the second coordinate p_y ; and (c) calculate the cumulative winding angle on the 2d plane; d) if the cumulative angle is more than 400° then that would mean that the initial track s is winding and therefore needs to be removed, see Fig. 4.8.1.

Winding tracks can be dangerous when we merge clusters because they could be close to many different clusters. We found that winding tracks

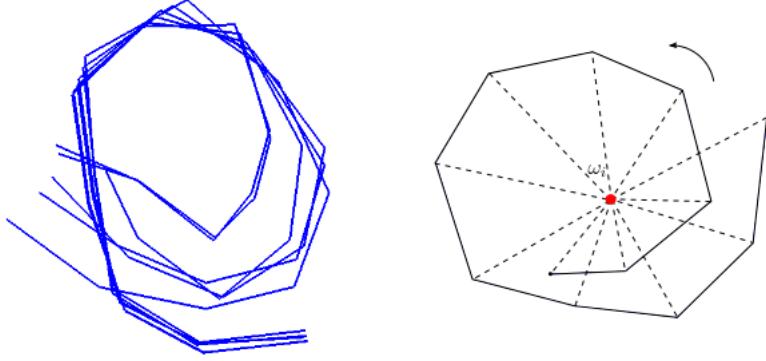


Figure 4.8: Example of detecting a possibly erroneous 3D bundle (on the left) by projecting its exemplar track and counting the winding cumulative angle $\sum_0^N \omega_i$ on the 2D plane as shown on the right, where N is the total number of track segments. Usually bundles with total angle higher than 400° are removed from the data sets as most likely to be erroneous.

often form bundles with many similar tracks. Also, they are usually long tracks so they will not be removed with any filters which remove short tracks. We could use QB with a low threshold to reduce the number of tracks while avoiding embedding winding tracks into otherwise ordinary clusters and then run the winding algorithm just on the exemplar tracks of the bundles rather than the entire tractography.

QB can also simplify detection of tracks which are very dissimilar to others and therefore they are very distant from all other clusters. Usually when we use a QB threshold of about 10mm these tracks will be part of small bundles containing a few tracks (< 10) and the distance of the bundle they belong to from all other bundles will be much higher than average. This can give us another detection method for outliers.

Finally, QB can be used to remove small or broken tracks in an interactive way, for example see Fig. 4.5.2 where the red large bundle has been merged by an expert and then with QB we can extract the skeleton of the bundle and see which parts create that structure. Without QB it would be too difficult to work out that this bundle consists of many small or divergent parts. In this figure both very diverging, small or broken tracks can be identified after the simplification provided by QB.

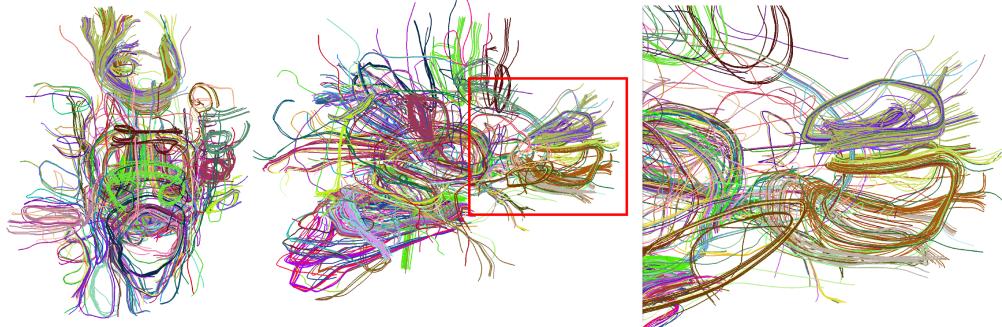


Figure 4.9: 161 most likely erroneous bundles automatically detected by our winding method all having total winding angle higher than 500 degrees and shown with random colours per bundle. On the left panel we see the bundles on their exact position in the data set from the top of the head, on the middle panel we see the same tractography from the side and the third panel we see the part of middle panel on the red box slightly rotated and much zoomed so that some erroneous tracks can be easily shown. To cluster the initial tractography not shown here we used QB with threshold 10mm. This is the first known automatic detection system of outliers and erroneous tracks for tractography data based on more advanced shape characteristics that go beyond simple length. By calculating the number of winding tracks in your data sets over the total number of tracks you can have an indicator of the quality of our data sets.

In summary, we have shown that QB can facilitate a fully automatic, efficient and robust detection system for erroneous tracks in specific bundles or entire tractographies.

4.8.2 Alignments, landmarks and atlases

We have used QB to construct a robust tractographic atlas in MNI space from data for 10 subjects. Here we explain the steps we used to achieve that.

Alignment. Tractographies were created using EuDX as described in section 4.4 and see section 1.3 for acquisition details. The tractographies for all subjects were initially in native space and the goal was to warp them in MNI space, using nonlinear registration.

Because the registration of tractographies is generally considered a difficult problem with a non-unique solution we wanted to make sure that we are using a known, well established and robust method, therefore we

chose to use `fnirt` with the same parameters as used with the first steps of TBSS[178]. For that reason FA volumes were generated from the same data sets using Tensor fitting with weighted least squares after skull stripping with `bet` and parameters `-F -f .2 -g 0`. These FA volumes were again in native space therefore we needed to warp them in MNI space. For this purpose a standard template (`FMRIB58_FA_1mm`) from the FSL toolbox was used as the reference volume. However, we wanted primarily to have the displacements which would do a point wise mapping from native space to MNI space and we found this to be technically very difficult with the FSL tools as they assume that these displacements will be applied only on volumetric data and not with point data as those used in tractographies. Finally, after some considerable effort we found a combination of `flirt`, `invwarp`, `fnirtfileutils` and `fnirtfileutils -withaff` which gave us the correct displacements. As this being very technical we will not describe it further here but the code is available in module (`dipy.external.fsl`). It is also important to say that we didn't use eddy correction with any of this type of data sets because eddy correction is unstable with volumes at high b-values because there is no much signal for guiding a correct registration with the other volumes at lower b-values.

After creating the displacements for every subject; these were applied to all tractographies in the native space so they are mapped in the MNI space of voxel size $1 \times 1 \times 1\text{mm}^3$. Having all tractographies in MNI space is something very useful because we can now compare them against available templates or against each other and calculate different statistics. However this is not where we stop; we proceed to generate a tractographic atlas using QB clusterings.

Tractographic Atlas. For all subjects, (a) load warped tractography (b) re-direct towards a static point (100, 100, 100) as explained in section 4.5.5, (c) downsample the tracks to have only 12 points, (d) calculate and store QB clustering with a 10mm threshold. Then merge all clusterings again with 10mm threshold as explained in section 4.7 (merging). When creating an atlas by merging many different subjects the most important issue is what you remove from the atlas as outliers. QB here provides a possible solution for this problem. If we plot the number of tracks for each cluster sorted in

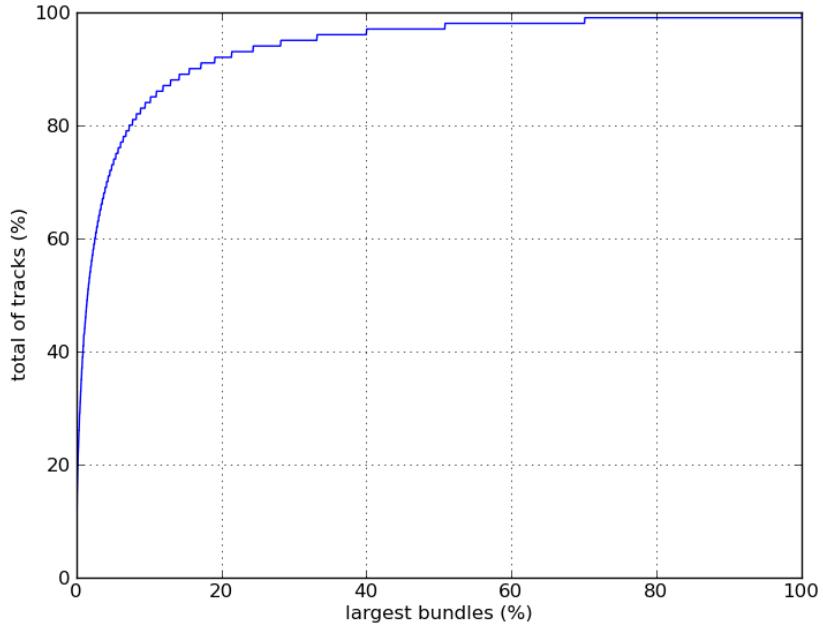


Figure 4.10: 14520 clusters where created by joining the QB clusterings of 10 subjects in MNI space. We found that most of the clusters had a few tracks and only a few clusters had many. In the diagram above we can see 20% of the largest clusters had more than 90% of the total amount of tracks. This result showed that there is much agreement between different subjects which would be useful for a solid atlas with the biggest bundles becoming landmark bundles and the small bundles removed as outliers.

ascending order we can see an interesting pattern see Fig.4.8.2. In this diagram we observe that 20% of the largest clusters had more than 90% of the total amount of tracks. This shows that there is much agreement between the biggest bundles of different subjects. We will use this property to create a solid atlas in which we keep the biggest bundles (landmarks) and remove small bundles (outliers).

Finding and Using Landmarks

One can use this atlas or similar atlases created from more subjects in order to select specific structures and study these structures directly in different subjects without using any of the standard ROI based methods.

A simple example is given in Fig. 4.8.2. In the first row we see a tractographic atlas joined by merging the QB clusterings of 10 healthy subjects as

described in the previous section. Then from these clusters represented by their virtual tracks we keep only 196 biggest clusters i.e. those which contain the highest number of tracks, so that we are sure that there is enough agreement from the different tractographies and from these we just pick by way of an example 19 virtual tracks which correspond to well known bundle structures in the literature: 1 from genu of corpus callosum (GCC), 3 from the body of corpus callosum (BCC), 1 from the splenium (SCC), 1 from the pons cerebellar peduncle (CP), 1 from left arcuate fasciculus (ARC-L), 1 from right arcuate fasciculus (ARC-R), 1 from left inferior occipitofrontal fasciculus (IFO-L) and 1 from right inferior occipitofrontal fasciculus (IFO-R), 1 from right fornix (FX-R), 1 from left fornix (FX-L), 1 optic radiation (OR), 1 left cingulum (CGC-L), 1 from right cingulum (CGC-R), 1 from left corticospinal tract (CST-L), 1 from right corticospinal tract (CST-R), 1 from left uncinate (UNC-L) and 1 from right uncinate (UNC-R). These 19 tracks are coloured randomly. Then on the second row we show, for the first 6 of these selected representative tracks, the tracks closer than 20mm from 3 arbitrarily selected subjects. Similarly, on the third row the tracks closer than 15mm to the next 7 selected tracks. Finally on the last row we bring the tracks from the same 3 subjects which are closer than 18mm. The colours used for the selected tracks are automatically assigned from the colours of tracks picked from the atlas. We can see that there is a significant reliability and continuity both within and between subjects even though we have only selected a very small number of representative tracks. Using a similar procedure we could create a book of bundles for every subject and then compare the subjects at the level of bundles.

4.8.3 QB as input to other learning methods

We found that QB is of great value as an adjunct to many less efficient algorithms e.g. hierarchical clustering, affinity propagation, nearest neighbours, spectral clustering and other unsupervised and supervised learning methods. We present here one example with QB as input to affinity propagation and one with QB as input to hierarchical clustering.

Most clustering algorithms need to calculate all pairwise distances be-

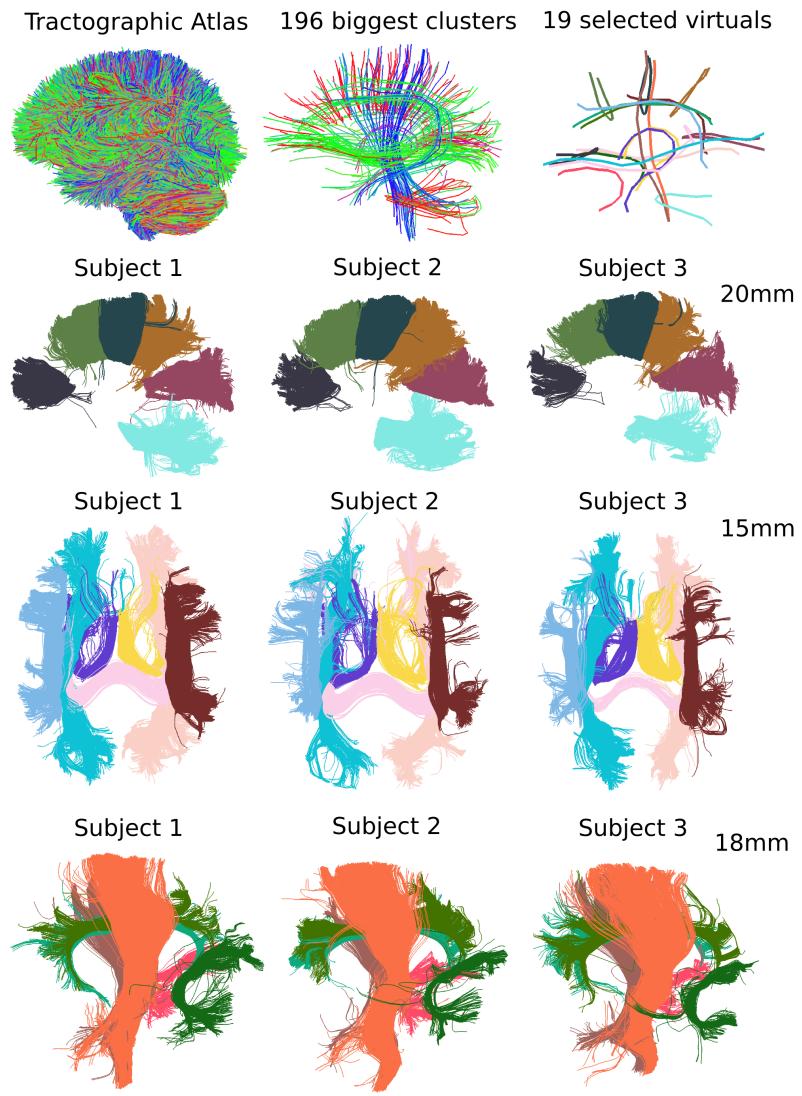


Figure 4.11: A novel way to do comparisons between subjects. Correspondence between different subjects (last 3 rows) and a few landmarks picked from the tractographic atlas generated by merging QB clusterings of 10 subjects (top row). That we can see this amount of agreement and continuity on the last 3 rows from such a few skeletal tracks is a great hope for implementing new robust ways of statistical comparisons using tractographic data sets.

tween tracks; that means that for a medium sized tractography of 250,000 tracks we would need 232 GBytes of RAM with single floating point precision. Something which is not and will not be available soon in personal computers. In those cases some people might hope that sparse matrices could provide a nice approximation; however dense tractographies produce very dense distance matrices. The straightforward solution to this problem is to use QB in order to first segment in small clusters and then use the representatives (i.e. exemplar or virtual tracks) of these clusters with other higher complexity operations and merge the clusters together in bigger clusters.

Procedure:

- (1) Cluster using QB as explained in section 4.8.2
- (2) Gather virtual tracks.
- (3) Calculate MDF distance of virtual tracks with themselves.
- (4) Use any other clustering method to segment this much smaller distance matrix D .

In Fig. 4.8.3 at the left panel we show a result where we used hierarchical clustering with single linkage for step (4) with a threshold of 20mm using the package `hcluster` (see [63]). A known drawback of single linkage is the so-called chaining phenomenon: clusters may be brought together due to single elements being close to each other, even though many of the elements in each cluster may be very distant to each other. Chaining is usually considered as a disadvantage because it is too driven by local neighbours. Nevertheless, we can use this property to cluster the corpus callosum (CC) all together (shown with dark red in left top of Fig. 4.8.3) creating a fully automatic CC detection system. Furthermore, we can use different cutting thresholds on the underlying dendrogram to amalgamate together different structures e.g. see the cingulum bundles in the same panel.

In the right panel of Fig. 4.8.3 we see implementation of step (4) using a more recent algorithm: affinity propagation (AP) [59], which earlier was identified by us (see Fig. 4.8.3) and [133] for being difficult or impossible to be used for group analysis or to cluster entire tractographies of many thousands of tracks. A small outline of how this algorithm works is given in section 4.9. Here we see in the bottom right panel of (see Fig. 4.8.3) how nicely AP, after the simplification provided by QB, has clustered arcuate,

longitudinal occipitofrontal fasciculus and other structures known from the literature. The input of AP was the negative distance matrix— D , the preference weights were set to matrix `median($-D$)` and the hierarchical clustering parameter was set to 20mm.

For hierarchical clustering parts we used the software `hcluster` and for affinity propagation we used the library `scikit-learn`. They are both implemented in Python.

4.8.4 Exemplars vs ROIs vs Masks

Medical practitioners and neuroanatomists often argue that when they use multiple spherical or rectangular masks to select some bundles many tracks are thrown away because they are small and the mask operations cannot get hold of them. Our method provides a solution to this problem as it can identify broken or smaller bundles inside other bigger bundles which are otherwise very difficult or even sometimes impossible to identify visually or with the use of masks. Our method attacks this problem and suggests a very efficient and robust solution which sets the limit for unsupervised clustering of tractographies and facilitates tractography exploration and interpretation. The point here is that one can now use exemplar tracks as access points into the full tractography and with a single click on that exemplar track obtain the entire bundle. Therefore a super-bundle can be created just with a few clicks based on a selection from exemplar tracks.

In order to create this system we implemented a 3D visualization/interaction system for tractographies based on QB in Python and OpenGL. This project is available online at fos.me.

4.9 Affinity Propagation

Affinity propagation (AP) is a very recent $O(N^2)$ clustering method invented by [73], [59] which is inspired by loopy belief propagation [164] and other recent innovations in graphical models and more specifically is an instance of the max-sum algorithm in factor graphs. For the completeness of this thesis and because AP is a relatively new algorithm we give a short description of the AP in this section. AP is an exemplar based clustering method where

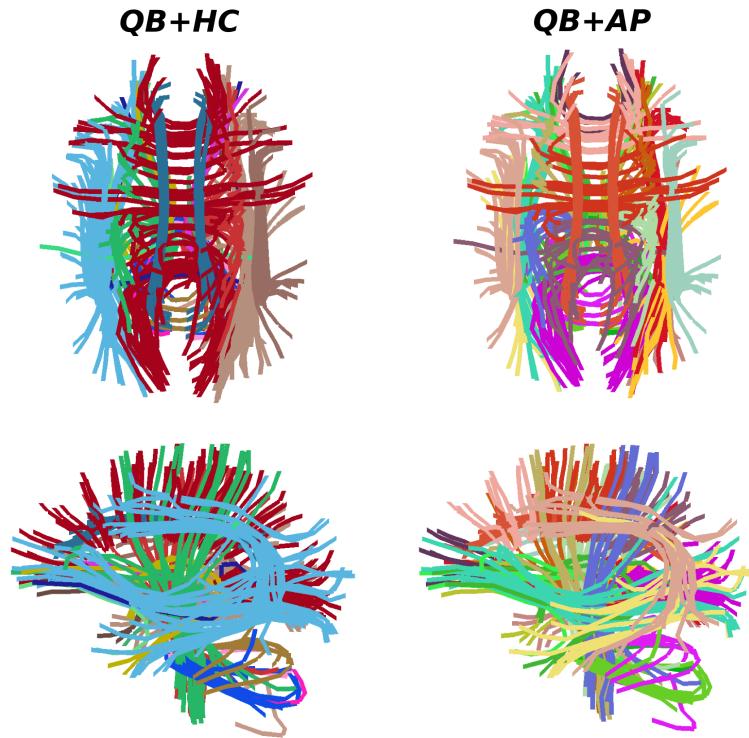


Figure 4.12: Two examples where QB output is used to cluster an entire set of 10 tractographies together and then the result is given as input to hierarchical clustering (HC) using single linkage on the left and to affinity propagation (AP) on the right. Colours encode cluster labels. On the left side we see 19 clusters and on the right 23. QB facilitates significantly the operation of the other two algorithms which would not be able to cluster the entire data sets on current computers. Pay attention at the top left panel where QB+HC have managed to cluster the entire CC as one bundle.

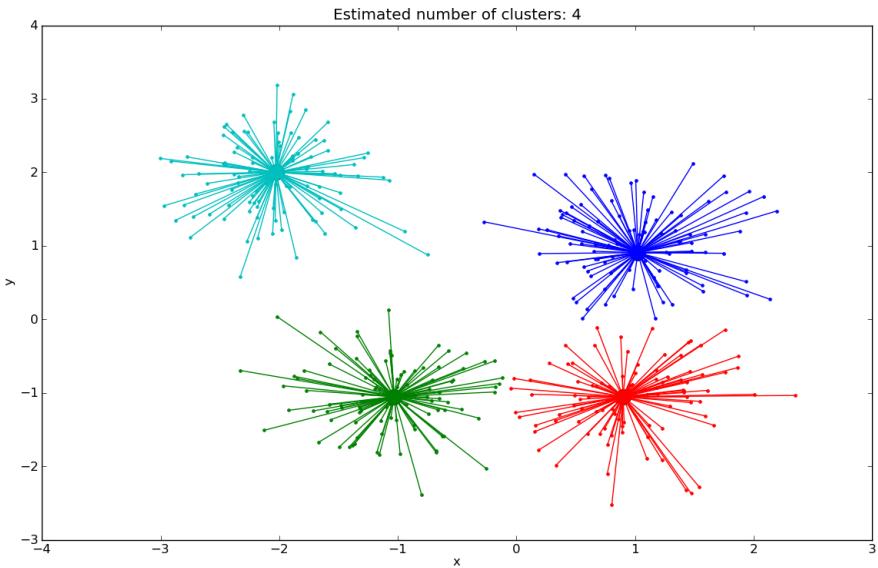


Figure 4.13: Simple example of affinity propagation at work where it can precisely identify 4 different normal distributions with means $(1,1)$, $(-1,-1)$, $(1,-1)$, $(-2,2)$ and standard deviation .5. You can see the exemplars - most representative actual points - with thicker dots perfectly aligned with the means.

the center of a cluster is a real data point (exemplar) as in k-medoids, and k-centres rather than an average virtual point as in k-means. AP starts by simultaneously considering all data points as potential exemplars. Every data point is a node in a network and AP recursively transmits real-valued messages along the edges of the network until a good set of exemplars and corresponding clusters emerges.

AP takes as input a collection of similarities between data points, where the similarity $S(i, k)$ indicates how well the data point with index k is suited to be the exemplar for data point i . In order to understand AP we can think just for the moment that we try to cluster 2D data points and each similarity is expressed as the negative Euclidean distance $S(i, k) = -||x_i - x_k||^2$ see Fig.4.9 therefore S for the moment is the negative complete squared distance matrix. Rather than requiring the number of clusters to be pre-specified, AP adds a real number (preference weights) to the diagonal ele-

ments of S , one for each data point so that larger values of $S(k, k)$ are more likely to become exemplars. For simplicity we can choose the $\text{median}(S)$ as the common preference weight for all points; in this way we don't enforce any *a priori* information for one point to be an exemplar any more than any other point. For some applications this could be an appropriate requirement. There are two different messages exchanged between points (1) responsibilities $R(i, k) = S(i, k) - \max_{k' : k' \neq k} [S(i, k') + A(i, k')]$ and (2) availabilities which are initially $A(i, k) = 0$ and then equal to:

$$\forall i, k : A(i, k) = \begin{cases} \sum_{i' : i' \neq i} \max[0, R(i', k)], & \text{for } k = i \\ \min \left[0, r(k, k) + \sum_{i' : i' \notin \{i, k\}} \max[0, r(i', k)] \right], & \text{for } k \neq i \end{cases} \quad (49)$$

A very interesting fact is the way we get the final exemplars using AP.

After the messages have converged, there are two ways you can identify exemplars:

- 1) For data point i , if $R(i, i) + A(i, i) > 0$, then data point i is an exemplar
- 2) For data point i , if $R(i, i) + A(i, i) > R(i, j) + A(i, j)$, for all i not equal to j , then data point i is an exemplar.

Therefore, the availabilities and responsibilities are added to identify exemplars. For point i , the value of k that maximizes $A(i, k) + R(i, k)$ either identifies i as an exemplar if $k = i$, or identifies the data point that is the exemplar for point i . The message passing procedure is terminated either after a fixed number of iterations, or after changes in the messages stay low, or local decisions stay constant; also the messages are damped - combining previous with current message - to avoid numerical oscillations.

Of course when we need to calculate distances between many points then the distance matrix becomes too big for the available memory. In that case if we are lucky and the data sets are sparse then we can use AP on sparse matrices but if the data sets are not sparse as it is the case with tractographies then we need to reduce the dimensionality of the data sets and this why QB can be very handy. The complete algorithm for AP is given in Alg.4.9.

Algorithm 6 Affinity Propagation

Input Similarity/affinity matrix S where the diagonal elements of $S(k,k)$ indicate the a priori preference for k to be chosen as an exemplar

Output Clustering $CAP = \{c_0, \dots, c_k, \dots, c_{|CAP|-1}\}$, where a cluster $c = \{I, \mathbf{e}, N\}$

$\forall i, k : A(i, k) = R(i, k) = 0$

$S = S + n$ # remove degeneracies

$d = 0.5$ # set damping factor

$last_iter=100$ # last iteration

For iter **From** 1 **To** last_iter **Do**

$R_{old} = R$

$\forall i, k : R(i, k) = S(i, k) - max_{k' : k' \neq k}[S(i, k') + A(i, k')]$

$R = (1 - d)R + d * R_{old}$ # dampen responsibilities

$A_{old} = A$

update availabilities

$$\forall i, k : A(i, k) = \begin{cases} \sum_{i':i' \neq i} \max[0, R(i', k)], & \text{for } k = i \\ \min \left[0, R(k, k) + \sum_{i':i' \neq \{i, k\}} \max[0, R(i', k)] \right], & \text{for } k \neq i \end{cases}$$

$A = (1 - d)A + dA_{old}$ # dampen availabilities

$\forall i, I_e = argmax S(i, I_d)$ # find indices of exemplars

$I_e(I_d) = 1 : size(I_d)$

$L = I_d(I_e)$ # assign labels

$CAP = \{c_0, \dots, c_k, \dots, c_{|CAP|-1}\}$ # clustering output

where a cluster $c = \{I, \mathbf{e}, N\}$ holds the AP exemplars \mathbf{e} ,

the indices I of the cluster elements and N the number of elements

4.10 Direct Tractography Registration

Direct tractography registration is a recently described problem with only a small number of publications, and so far as we know there are no publicly available solutions. By direct registration we mean that no other information apart from the tractographies themselves is used to guide the registration. This is in contrast to the previous sections where we used FA registration mappings applied to tractographies (see section 4.8.2) which is also most commonly used in the literature along with other Tensor based methods [82].

The currently described methodologies on this subject are as follows. Leemans et al. [121] uses the invariance of curvature and torsion under rigid registration along with Procrustes analysis to co-register together different tractographies. Mayer et al. used iterative closest point applied to register pre-selected bundles (bundles of interest - BOI) [136], [138] and extended it using probabilistic boosting tree classifiers for bundle segmentation in[137]. Durrleman et al. [60] reformulated the tracks as currents and implemented a currents based registration. Zvitia et al. [217] [218], used adaptive mean shift clustering to extract a number of representative fibre-modes. Each fibre mode was assigned to a multivariate Gaussian distribution according to its population thereby leading to a Gaussian Mixture model (GMM) representation for the entire set of fibres. The registration between two fibre sets was treated as the alignment of two GMMs and is performed by maximizing their correlation ratio. A further refinement was added using RANSAC[70] to obtain all 12 affine parameters. Ziyan et al.[215] developed a nonlinear registration algorithm based on the log-Euclidean polyaffine framework[7]; however this is not a direct tractography registration algorithm as they first create scalar volumes, therefore they do not try to register the tracks themselves in their space.

We now describe our algorithm and show that it is efficient and simple to use, completely automatic and provides an evidently robust direct rigid tractography registration algorithm available in seconds. This algorithm could be of great use when comparing healthy versus severely diseased brains e.g. stroke or vegetative state patients when non-rigid registration

is not recommended because of severe asymmetries in the diseased brains. The algorithm is based on the robustness of QB to find good representative descriptors.

Procedure. Here we describe a simple algorithm where 2 tractographies T_A, T_B are brought into alignment in native space.

1. All tracks with length smaller than 100mm and longer than 300mm are removed from the data sets. This will reduce the size of tractography to about 1/4 of its initial size (200,000 tracks). (While all the subjects are adults, this filtering may have different effects depending on brain size. We have not investigated this question at present.)
2. Both tractographies are equidistantly downsampled so every track contains only 12 points.
3. We run QB with distance threshold at 10mm for both tractographies.
4. Collect all exemplar tracks from clusters containing more than 0.2% tracks. Let us assume we have these now in E_A and E_B .
5. Calculate all pairwise distances $D = \text{MDF}(E_A, E_B)$ and save them in rectangular matrix D .
6. Create a cost function (optimizer) which we will try to minimize the symmetric minimum distance $\text{SMD} = \sum_i \min_j D(i, j) + \sum_j \min_i D(i, j)$
7. Use modified Powell's method [72] to minimize SMD over rigid rotations of E_B starting with zeroed initial conditions. At each iteration of the optimization, E_B will be transformed by a rigid rotation and SMD will be recalculated. To ensure smooth rotations we use the Rodriguez rotation formula.

In Fig.4.10 A we see the result of this algorithm applied to two tractographies – represented with their exemplar tracks – depicted with orange and purple. We can see in the upper panel that the orange tractography is misaligned with respect to the purple one, and in the lower panel we see their improved alignment after applying our algorithm.

Metric. SMD is proposed here for registration of trajectory data sets, but one could equally use mutual information[132] or the correlation ratio [172] for registration of volumetric data sets. Nonetheless, the advantage of SMD is that it comes from robust landmarks generated by QB which bring together local and global components. Initially, it was not clear if we should use SMD or just the sum of all distances $SD = \sum_{i,j} D(i,j)$. Therefore, we made a small experiment to validate the smoothness and convexity of these two cost functions. We plotted both functions under a single-axis translation or a single-angle rotation of the same tractography as show in Fig.4.10 B and C. From, these two diagrams we can see that although for translations only the SD was entirely convex, with rotations the SD had stronger local minima which is not a good property for registration. Furthermore, the SMD had steeper gradients towards the global minimum which is a positive indicator for faster convergence.

Experiments. The first large scale experiment took place using the same tractography of a single individual copied and transformed 1000 times with range of all three angles from -45 degrees to 45 and range of all x,y,z translations from -113 to 113 mm. Then we registered all transformed tractographies to the static one and calculated all pairwise MDF distances storing them in a square matrix D . We would expect that if the registration was correct then the sum of all diagonals elements of D would be close to 0. This was confirmed with both cost functions used SD and SMD getting close to zero 99.8% of the time however SMD was always closer to perfect alignment than SD, having precision of more than 7 decimals. Consequently we chose SMD as a better cost function for direct tractography registration.

We used GQI-based tractographies from 10 subjects and we registered all combinations of pairs $\binom{10}{2} = 45$. Comparing different tractographies is not a trivial problem however we can use the tightness comparison (TC) metric explained in section 4.6.3. We are happy to report the mean initial TC was $34.8\% \pm 8.0\%$ and the mean final TC after applying our direct registration method was $48.1\% \pm 6.1\%$. This was a statistically highly significant improvement ($t_{\text{paired}}(44) = 11.2, p \leq 10^{-13}$). We are planning in the future to compare this registration method against other standard methods which are common in the literature.

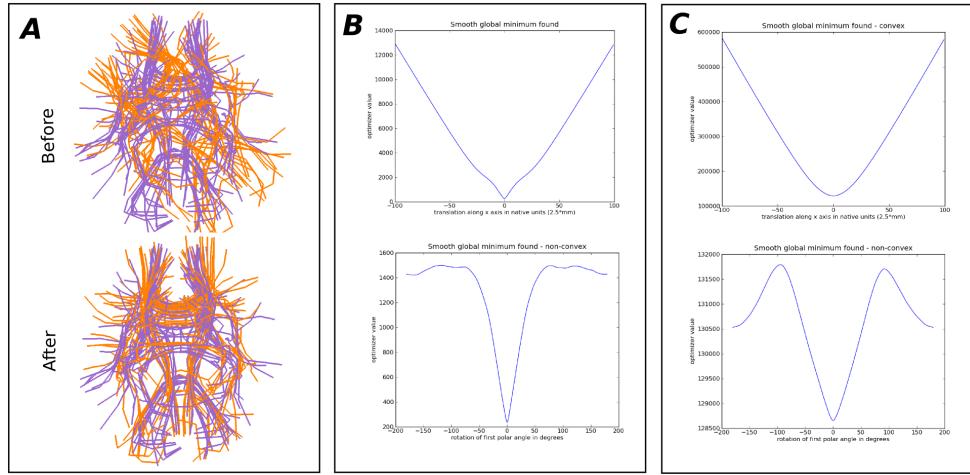


Figure 4.14: In panel A we see two tractographies from different subjects before (top) and after rigid registration (bottom) using our method. In panel B we see the metric SMD that we chose to optimize for two copies of the same tractography with the second copy translated (above) and rotated (below). This metric appears to be smooth with a single global minimum and is only slightly non-convex with small local minima. In panel C another possible candidate metric SD is shown which although more convex on translations it was much more problematic with rotations.

4.11 Strategies with Small fibres

In many parts of this document we did not consider short tracks. That is perfectly valid because (a) the longer tracks are more likely to be used as useful landmarks when comparing or registering different subjects because it is more likely for them to exist in most subjects, (b) removing short tracks facilitates the usage of distance based clustering (no need for manually setting the distance threshold) and interaction with the tractography, (c) someone would first want to see the overall representation of the tractography and later go to the details. Nonetheless, after having clustered the longer tracks there are many ways to assign the smaller bundles to their closest longer bundles. For this purpose we recommend to use a different distance from d_{df} (MDF) for example the minimum version of MAM referred to as MAM_{min} in Eq. 45.

Here are some simple strategies for clustering short fibres. The first is for unsupervised clustering and the second one is for supervised learning.

Strategy 1. Cluster the long tracks using QB with distance threshold at 10mm and then cluster the short tracks ($<100\text{mm}$) to a lower threshold and assign them to their closest long track bundle from the first clustering using the MAM_{min} distance.

Strategy 2. Read the tractography of a single subject, use a tractographic atlas as the one created in section 4.8.2 and pick one or more close skeletal tracks from that atlas and then find the closest tracks from the subject to that selected track using d_{df} , cluster the closest tracks found from the previous step and for each one of these new skeletons find the closest tracks using the m_{in} . We should now have an amalgamation of shorter and longer fibres in one cluster.

An example of this second strategy is shown in Fig. 4.11: (A) a track of interest from the arcuate fasciculus is selected from the tractographic atlas shown in Fig. 4.8.2(top row-middle), (B) the tracks of the subject closer than 15mm (d_{df}) from the selected cluster are shown and clustered with a distance threshold of 6.25mm in (C), (D) from every skeleton track in C we find the closest tracks using the (m_{in}) distance from the entire tractography.

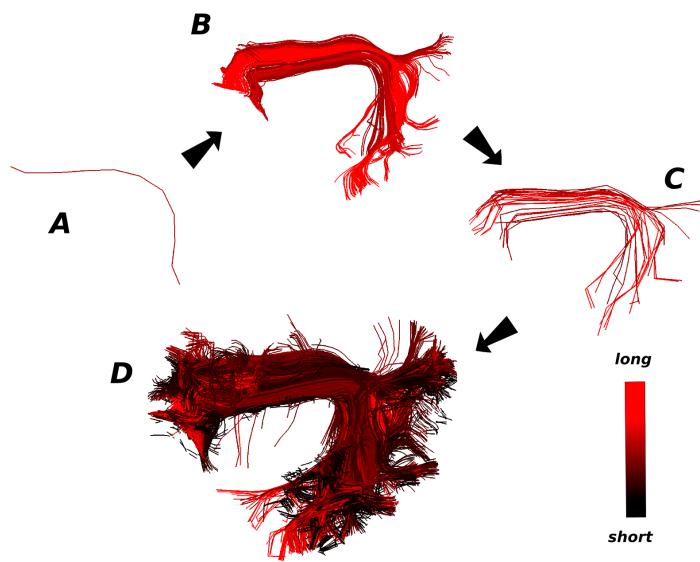


Figure 4.15: A simple and vigorous strategy for handling short and long tracks together by picking a track of interest from one of our atlases. Colourmap here encodes track length. A: one picked selected atlas track, B: 245 subject tracks closer than 15mm (MDF distance), C: B tracks clustered in 23 skeletons, D: 3421 tracks closer than 6mm (MAM distance) from the skeletons of B are shown. We can see that a great number of short tracks have been brought together along with the tracks in B. In that way we managed to bring together an entire bundle consisting both of long and short fibers by just selecting one track.

4.12 Discussion and conclusion

In this document we presented a novel and powerful algorithm – Quick-Bundles (QB). This algorithm provides simplifications to the old problem of white matter anatomy packing which has recently attracted much scientific attention; it can also be used for any trajectory clustering problem and it is recommended when large data sets are involved. QB can be used with all types of diffusion MRI tractographies which generate streamlines (e.g. probabilistic or deterministic) and it is independent of the reconstruction model.

In common with mainstream clustering algorithms such as k-means, k-centers and expectation maximization, QB is not a global clustering method therefore it can give different results under different initial conditions of the data set when there is no obvious distance threshold which can separate the clusters into meaningful bundles; for example we should expect different clusters under different permutations/orderings of the tracks in a densely packed tractography. However, we found that there is enough agreement even between two clusterings of the same tractography with different orderings. If the clusters are truly separable by distances then there is a global solution independent of orderings. This is often perceivable in smaller subsets of the initial tractography. We empirically found that this problem is minimized even with real data sets when a low distance threshold of about 10 – 20 mm is used.

Furthermore the output of QB can become now input for another recent quick algorithm of quadratic time on average $O(M^2)$ called affinity propagation where now $M \ll N$ therefore the overall time stays linear on the number of tracks N . Other algorithms previously too slow to be used on the entire tractography can now be used efficiently too e.g. kNN, hierarchical clustering and many others.

We saw that QB is a linear time clustering method based on track distances, which is on average linear time $O(N)$ where N is the number of tracks and with worst case $O(N^2)$ when every track is a singleton cluster itself. Therefore QB is the fastest known tractography clustering method and even real-time on tractographies with less than 20,000 tracks (depending on

system CPU). We also showed that it uses a negligible amount of memory.

QB is fully automatic and very robust as when we use it we can find good agreements even between different subjects and can be used to create tractography atlases at high speed. Additionally, it can be used to explore multiple tractographies and find correspondences between tractographies, create landmarks used for registration or population comparisons.

QB can be used as well for reducing the dimensionality of the data sets at the time of interaction; providing an alternative way to ROIs using BOIs (bundles of interest) or TOIs (tracks of interest). We also showed that it can be used to find “hidden” tracks not visible to the user at first instance. Therefore QB opens up the road to create a rapid tools for exploring tractographies of any size.

The main concept of this clustering method is that a cluster can be represented by virtual tracks which are used only during cluster comparisons and not updated at every iteration.

A virtual (centroid) track is the average of all tracks in the cluster. We call it virtual because it doesn’t really exist in the real data set and to distinguish it from exemplar (medoid) tracks which are again descriptors of the cluster but are represented by real tracks.

The clustering creates a book of bundles/clusters which have easily obtainable descriptors. When clusters are held in a tree structure this permits upwards amalgamations to form bundles out of clusters, and downwards disaggregation to split clusters into finer sub-clusters corresponding to a lower distance threshold. However, we did not touch this hierarchical extension of this algorithm here and mostly concentrate on one level amalgamations.

We worked mostly with long tracks but strategies for short tracks or bundles are straightforward and documented. We also showed an efficient method where QB can speedup finding erroneous bundles or detecting structures of specific characteristics.

We showed results with simulated, single or multiple real subjects and the code for QuickBundles is freely available at dipy.org.

5 Discussion

5.1 Overview

[?]

[144]

This thesis contains original contributions to many different aspects of diffusion weighted MR imaging (dMRI), from the signal reconstruction, via propagation of tracks, clustering of tracks, detection of bundles of tracks, and creation of atlases of white matter, to visualization tools with high interactivity with data, and for several brains simultaneously.

Our most important new achievements are a powerful way of reconstructing the diffusion signal and a method for simplifying tractographies using a new super-efficient clustering algorithm.

In more detail we give novel and robust solutions for all the following problems

- **How to reconstruct and interpret the directional information in the diffusion signal:**

We compare and evaluate different Cartesian-grid q-space dMRI acquisition schemes, using methods based on the inverse Fourier transform of the diffusion signal, with reconstructions by diffusion spectrum imaging or generalized q-ball imaging methods. These are applied to both software phantom and human data. We propose a new reconstruction method called diffusion nabla imaging (DNI) which works with the same grid type of acquisition schemes, using an algorithm that directly approximates the orientation distribution function using the Laplacian of the signal in q-space.

- **How to integrate tracks from voxel to voxel:**

Most previously published reconstruction methods are closely linked to their own specific track integration method. We have formulated a general, free-standing, data-representative, non-inferential, fast, deterministic tractography algorithm (EuDX) which is based on Euler integration and trilinear interpolation, which works with voxel level information about fibre orientations including multiple crossings, and employs a range of stopping criteria. The purpose of this algorithm is to be faithful to the reconstruction

aaa	bbb
5	6
4	7

Table 5.1:

results rather than try to correct or enhance them by introducing regional or global considerations.

- **How to cluster tracks together into meaningful bundles:**

We have developed an entirely new, fully automatic, linear time, clustering method (QuickBundles) which reduces massive tractographies into just a few bundles. These bundles are characterised by representative tracks which are multi-purpose and can be used for interaction with the data or as the basis for applying higher-complexity clustering methods which would have been impossible or too slow with the full data set. QuickBundles is currently the fastest known tractography clustering algorithm.

- **How to directly register tractographies:**

After applying QuickBundles to tractographies from different subjects, we show how to use the representative tracks to identify robust landmarks within each subject which, with similarity metrics which we have introduced, we use to directly register the different tractographies together in a highly efficient way. We believe the resulting correspondences provide important evidence for the anatomical plausibility of the derived bundles. We demonstrate how these methods can be used for group analysis, and for atlas creation.

5.2 Clinical Usage

5.3 Why speed is so important?

5.4 Future Work

References

- [1] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel. Reconstruction of the orientation distribution function in single-and

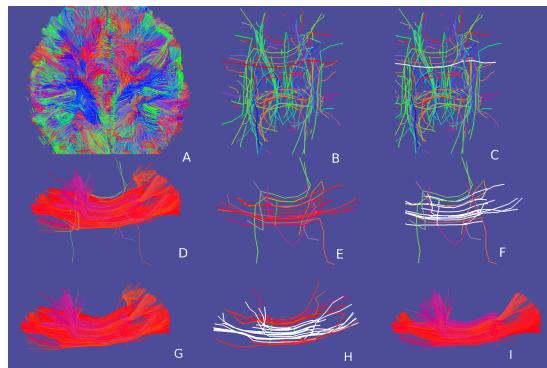


Figure 5.1:

multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2):554–566, 2010.

- [2] Iman Aganj, Christophe Lenglet, Neda Jahanshad, Essa Yacoub, Noam Harel, Paul M Thompson, I M A Preprint Series, and Church Street S E. A HOUGH TRANSFORM GLOBAL PROBABILISTIC APPROACH A Hough Transform Global Probabilistic Approach to Multiple- Subject Diffusion MRI Tractography. pages 612–626, 2010.
- [3] DC Alexander, GJ Barker, and SR Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.
- [4] D.C. Alexander, P.L. Hubbard, M.G. Hall, E.A. Moore, M. Ptito, G.J.M. Parker, and T.B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion mri. *NeuroImage*, 52(4):1374–1389, 2010.
- [5] J.L.R. Andersson and S. Skare. A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *Neuroimage*, 16(1):177–199, 2002.
- [6] A. Anwander, M. Tittgemeyer, D Y von Cramon, A D Friederici, and T R Knosche. Connectivity-Based Parcellation of Broca’s Area. *Cerebral Cortex*, 17(4):816, 2007.

- [7] Vincent Arsigny, Olivier Commowick, Nicholas Ayache, and Xavier Pennec. A Fast and Log-Euclidean Polyaffine Framework for Locally Linear Registration. *Journal of Mathematical Imaging and Vision*, 33(2):222–238, January 2009.
- [8] Y. Assaf and P.J. Basser. Composite hindered and restricted model of diffusion (charmed) mr imaging of the human brain. *Neuroimage*, 27(1):48–58, 2005.
- [9] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P.J. Basser. AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, 59(6):1347–1354, 2008.
- [10] Y. Assaf, R.Z. Freidlin, G.K. Rohde, and P.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.
- [11] E. J. Auerbach and K. Ugurbil. Improvement in diffusion mri at 3t and beyond with the twice-refocused adiabatic spin echo (trase) sequence. 2004.
- [12] A. Barmpoutis, M.S. Hwang, D. Howland, J.R. Forder, and B.C. Vemuri. Regularized positive-definite fourth order tensor field estimation from dw-mri. *NeuroImage*, 45(1):S153–S162, 2009.
- [13] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magn Reson Med*, 44(4):625–32, 2000.
- [14] PJ Basser, J Mattiello, and D LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 1994.
- [15] P. Batchelor, D. Hill, D. Atkinson, and F. Calamante. Study of connectivity in the brain using the full diffusion tensor from mri. In *Information Processing in Medical Imaging*, pages 121–133. Springer, 2001.
- [16] C. Beaulieu. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed*, 15(7-8):435–55, 2002.

- [17] T E Behrens and H. Johansen-Berg. Relating connectional architecture to grey matter function using diffusion imaging. *Philos Trans R Soc Lond B Biol Sci*, 360(1457):903–11, 2005.
- [18] T E J Behrens, H Johansen-Berg, M W Woolrich, C A M Wheeler-Kingshott, P A Boulby, G J Barker, E L Sillery, K Sheehan, O Ciccarelli, A J Thompson, J M Brady, and P M Matthews. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6(7):750–757, 2003.
- [19] T. E. J. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted {MR} imaging. *Magnetic Resonance in Medicine*, 50:1077–1088, 2003.
- [20] T.E.J. Behrens, H. Johansen-Berg, S. Jbabdi, M.F.S. Rushworth, and M.W. Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1):144–155, 2007.
- [21] J.I. Berman, S.W. Chung, P. Mukherjee, C.P. Hess, E.T. Han, and R.G. Henry. Probabilistic streamline q-ball tractography using the residual bootstrap. *Neuroimage*, 39(1):215–222, 2008.
- [22] M. Björnemo, A. Brun, R. Kikinis, and C.-F. Westin. Regularized stochastic white matter tractography using diffusion tensor MRI. In *Fifth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, pages 435–442, Tokyo, Japan, 2002.
- [23] A. Brun, H. Knutsson, H.J. Park, M.E. Shenton, and C.F. Westin. Clustering fiber traces using normalized cuts. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, pages 368–375, 2004.
- [24] Paul T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford University Press, 1991.
- [25] PT Callaghan, CD Eccles, and Y. Xia. Nmr microscopy of dynamic displacements: k-space and q-space imaging. *Journal of Physics E: Scientific Instruments*, 21:820, 1988.

- [26] J.S.W. Campbell, K. Siddiqi, V.V. Rymar, A.F. Sadikot, and G.B. Pike. Flow-based fiber tracking with diffusion tensor and q-ball data: validation and comparison to principal diffusion direction techniques. *NeuroImage*, 27(4):725–736, 2005.
- [27] E.J. Canales-Rodríguez, L. Melie-García, Y. Iturria-Medina, and C.N. Center. Mathematical description of q-space in spherical coordinates: Exact q-ball imaging. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine*, 2009.
- [28] Erick Jorge Canales-Rodríguez, Lester Melie-García, and Yasser Iturria-Medina. Mathematical description of q-space in spherical coordinates: exact q-ball imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 61(6):1350–67, 2009.
- [29] Giorgio Carpaneto, Silvano Martello, and Paolo Toth. Algorithms and codes for the assignment problem. *Annals of Operations Research*, 13:191–223, 1988.
- [30] Giorgio Carpaneto, Silvano Martello, and Paolo Toth. Linear assignment problem freeware. www.assignmentproblems.com/APC_APS.htm, downloaded June 2011.
- [31] M. Catani, D.K. Jones, et al. Perisylvian language networks of the human brain. *Annals of neurology*, 57(1):8–16, 2005.
- [32] Marco Catani, Robert J. Howard, Sinisa Pajevic, and Derek K. Jones. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage*, 17:77–94, 2002.
- [33] M. Chalmers. A linear iteration time layout algorithm for visualising high-dimensional data. In *Visualization'96. Proceedings.*, pages 127–131. IEEE, 1996.

- [34] Y.P. Chao, J.H. Chen, K.H. Cho, C.H. Yeh, K.H. Chou, and C.P. Lin. A multiple streamline approach to high angular resolution diffusion tractography. *Medical engineering & physics*, 30(8):989–996, 2008.
- [35] S.W. Chung, Y. Lu, and R.G. Henry. Comparison of bootstrap approaches for estimation of uncertainties of dti parameters. *Neuroimage*, 33(2):531–541, 2006.
- [36] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci U S A*, 96(18):10422–7, 1999.
- [37] P.A. Cook, D.C. Alexander, and G.J.M. Parker. Modelling noise-induced fibre-orientation error in diffusion-tensor mri. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 332–335. IEEE, 2004.
- [38] I. Corouge, S. Gouttard, and G. Gerig. A statistical shape model of individual fiber tracts extracted from diffusion Tensor MRI. *Medical Image Computing and Computer Assisted Intervention*, pages 671–679, 2004.
- [39] I. Corouge, S. Gouttard, and G. Gerig. Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *International Symposium on Biomedical Imaging*, pages 344–347, 2004.
- [40] Isabelle Corouge, P Thomas Fletcher, Sarang Joshi, Sylvain Gouttard, and Guido Gerig. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Medical image analysis*, 10(5):786–98, 2006.
- [41] Marta Morgado Correia. Development of Methods for the Acquisition and Analysis of Diffusion Weighted MRI Data. *Brain*, (June), 2009.
- [42] M.M. Correia, T.A. Carpenter, and G.B. Williams. Looking for the optimal dti acquisition scheme given a maximum scan time: are more b -values a waste of time? *Magnetic resonance imaging*, 27(2):163–175, 2009.

- [43] Williams G.B. Yeh F-C. Nimmo-Smith I. Correia, M.M. Robustness of diffusion scalar metrics when estimated with generalized q-sampling imaging acquisition schemes. In *19th Annual Meeting of the International Society of Magnetic Resonance in Medicine*, 2011.
- [44] J. Dauguet, S. Peled, V. Berezovskii, T. Delzescaux, S. K. Warfield, R. Born, and C. F. Westin. Comparison of fiber tracts derived from in-vivo DTI tractography with 3D histological neural tract tracer reconstruction on a macaque brain. *Neuroimage*, 37(2):530–8, 2007.
- [45] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. A minimum description length approach to statistical shape modeling. *IEEE Trans Med Imaging*, 21(5):525–37, 2002.
- [46] Cagatay Demiralp, John F. Hughes, and David H. Laidlaw. Coloring 3D line fields using Boy's real projective plane immersion. *IEEE Trans. on Visualization and Computer Graphics (Proc. Visualization '09)*, 15(6):1457–1463, 2009. url`http://vis.cs.brown.edu/docs/pdf/g/Demiralp-2009-CLF.pdf.html`(pdf) url`http://vis.cs.brown.edu/results/bibtex/Demiralp-2009-CLF.bib`(bibtex: Demiralp-2009-CLF).
- [47] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.
- [48] MD Denis Le Bihan, C. Poupon, A. Amadon, and F. Lethimonnier. Artifacts and pitfalls in diffusion MRI. *Journal of Magnetic Resonance Imaging*, 24:478–488, 2006.
- [49] R. Deriche and O. Faugeras. 2-D curve matching using high curvature points: application to stereo vision. In *Pattern Recognition, 1990. Proceedings., 10th International Conference on*, volume 1, 1990.
- [50] M Descoteaux, E Angelino, S Fitzgibbons, and R Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, vol:58no3pp497–510, 2007.

- [51] M. Descoteaux, R. Deriche, and A. Anwander. Deterministic and probabilistic q-ball tractography: from diffusion to sharp fiber distribution. 2007.
- [52] M. Descoteaux, R. Deriche, D.L. Bihan, J.F. Mangin, and C. Poupon. Multiple q-shell diffusion propagator imaging. *Medical Image Analysis*, 2010.
- [53] M. Descoteaux, R. Deriche, T. Knoesche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Trans Med Imaging*, 28(2):269–86, 2009.
- [54] Maxime Descoteaux, Elaine Angelino, Shaun Fitzgibbons, and Rachid Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007.
- [55] Maxime Descoteaux, Rachid Deriche, Thomas R Knösche, and Alfred Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE transactions on medical imaging*, 28(2):269–86, February 2009.
- [56] Z Ding, J Gore, and A Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magn. Reson. Med.*, 49:716–721, 2003.
- [57] Zhaohua Ding, John C Gore, and Adam W Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49(4):716–21, 2003.
- [58] R. F. Dougherty, M. Ben-Shachar, R. Bammer, A. A. Brewer, and B. A. Wandell. Functional organization of human occipital-callosal fiber tracts. *Proc Natl Acad Sci U S A*, 102(20):7350–5, 2005.
- [59] D. Dueck. *Affinity propagation: Clustering data by passing messages*. PhD thesis, University of Toronto, 2009.
- [60] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage*, 2010.

- [61] Stanley Durrleman, Pierre Fillard, Xavier Pennec, Alain Trouvé, and Nicholas Ayache. A statistical model of white matter fiber bundles based on currents. *Proceedings of the 21st Conference on Information Processing in Medical Imaging*, 21:114–25, January 2009.
- [62] P. Eades. A heuristic for graph drawing. *Congressus numerantium*, 42:149–160, 1984.
- [63] Damian Eads. hcluster: Hierarchical clustering for scipy. scipy-cluster.googlecode.com, 2008.
- [64] B. Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 7(1):1–26, 1979.
- [65] A. Einstein. *Investigations on the Theory of the Brownian Movement*. Dover Publications, 1956.
- [66] V El Kouby, Y Cointepas, C Poupon, D Rivière, N Golestani, J B Poline, D Le Bihan, and J F Mangin. MR diffusion-based inference of a fiber bundle model from a population of subjects. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 8(Pt 1):196–204, January 2005.
- [67] P. Fillard, M. Descoteaux, A. Goh, S. Gouttard, B. Jeurissen, J. Malcolm, A. Ramirez-Manzanares, M. Reisert, K. Sakaie, F. Tensaouti, et al. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion mr phantom. *NeuroImage*, 2011.
- [68] P. Fillard, C. Poupon, and J.F. Mangin. Spin Tracking: A Novel Global Tractography Algorithm. *NeuroImage*, 47:S127–S127, 2009.
- [69] Pierre Fillard and Cyril Poupon. A Novel Global Tractography Algorithm based on an Adaptive Spin Glass Model. 2009.
- [70] M.A. Fischler and R.C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

- [71] P. Fletcher, R. Tao, W.K. Jeong, and R. Whitaker. A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor mri. In *Information Processing in Medical Imaging*, pages 346–358. Springer, 2007.
- [72] R. Fletcher. Practical methods of optimization. *John and Sons, Chichester*, 1987.
- [73] B.J. Frey and D. Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972, 2007.
- [74] S Frey, JSW Campbell, and GB Pike. ...human language pathways with high angular resolution diffusion fiber tractography. *Journal of Neuroscience*, 2008.
- [75] O. Friman, G. Farneback, and C.F. Westin. A Bayesian approach for stochastic white matter tractography. *Medical Imaging, IEEE Transactions on*, 25(8):965–978, 2006.
- [76] O. Friman and C. F. Westin. Uncertainty in white matter fiber tractography. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 8(Pt 1):107–14, 2005.
- [77] E. Garyfallidis, M. Brett, B. Amirbekian, C. Nguyen, FC Yeh, Y. Halchenko, and I. Nimmo-Smith. Dipy - a novel software library for diffusion mr and tractography. In *17th Annual Meeting of the Organization for Human Brain Mapping*, 2011.
- [78] Eleftherios Garyfallidis, Matthew Brett, and Ian Nimmo-Smith. Fast Dimensionality Reduction for Brain Tractography Clustering. *16th Annual Meeting of the Organization for Human Brain Mapping*, 2010.
- [79] G. Gerig, S. Gouttard, and I. Corouge. Analysis of brain white matter via fiber tract modeling. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 2, pages 4421–4424. IEEE, 2004.

- [80] X. Gigandet. *Global brain connectivity analysis by diffusion MR tractography: algorithms, validation and applications*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2009.
- [81] M.F. Glasser and J.K. Rilling. Dti tractography of the human brain's language pathways. *Cerebral Cortex*, 18(11):2471–2482, 2008.
- [82] A. Goh and R. Vidal. Algebraic methods for direct and feature based registration of diffusion tensor images. *Computer Vision–ECCV 2006*, pages 514–525, 2006.
- [83] T.F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.
- [84] P Guevara, C Poupon, D Rivière, Y Cointepas, M Descoteaux, B Thirion, and J-F Mangin. Robust clustering of massive tractography datasets. *NeuroImage*, 54(3):1975–1993, October 2010.
- [85] H. Gray. Gray's Anatomy. *Lea and Febiger, 20th u.s. edition edition, 1918. Online edition: bartelby.com*, 2000.
- [86] N.S. Hageman, A.W. Toga, K.L. Narr, and D.W. Shattuck. A diffusion tensor imaging tractography algorithm based on navier–stokes fluid mechanics. *Medical Imaging, IEEE Transactions on*, 28(3):348–360, 2009.
- [87] P. Hagmann, J. P. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder, and R. Meuli. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *Neuroimage*, 19(3):545–54, 2003.
- [88] P. Hagmann, J.P. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder, and R. Meuli. Dt mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *Neuroimage*, 19(3):545–554, 2003.
- [89] Patric Hagmann, Lisa Jonasson, Philippe Maeder, Jean-Philippe Thiran, Van J. Wedeen, and Reto Meuli. Understanding diffusion mr imaging techniques: From scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics*, 26(1):205–223, 2006.

- [90] H.A. Haroon, D.M. Morris, K.V. Embleton, D.C. Alexander, and G. Parker. Using the model-based residual bootstrap to quantify uncertainty in fiber orientations from *q*-ball analysis. *Medical Imaging, IEEE Transactions on*, 28(4):535–550, 2009.
- [91] J. C. Haselgrove and J. R. Moore. Correction for distortion of echo-planar images used to calculate the apparent diffusion coefficient. *Magn Reson Med*, 36(6):960–4, 1996.
- [92] O. Heid. Eddy current-nulled diffusion weighting. In: *Proceedings of the 8th Annual Meeting of ISMRM, Denver*, page 799, 2000.
- [93] T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.
- [94] T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.
- [95] Kegang Hua, Jiangyang Zhang, Setsu Wakana, Hangyi Jiang, Xin Li, Daniel S. Reich, Peter A. Calabresi, James J. Pekar, Peter C.M. van Zijl, and Susumu Mori. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, 39(1):336–347, 2008.
- [96] Y. Iturria-Medina, EJ Canales-Rodríguez, L. Melie-García, PA Valdés-Hernández, E. Martínez-Montes, Y. Alemán-Gómez, and J M Sánchez-Bornot. Characterizing brain anatomical connections using diffusion weighted {MRI} and graph theory. *Neuroimage*, 36(3):645–660, 2007.
- [97] Kalvis M Jansons and Daniel C Alexander. Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems*, 19(5):1031–1046, 2003.
- [98] S Jbabdi, M W Woolrich, J L Andersson, and T E Behrens. A Bayesian framework for global tractography. *NeuroImage*, 37:116–129, 2007.

- [99] S. Jbabdi, MW Woolrich, JLR Andersson, and TEJ Behrens. A Bayesian framework for global tractography. *Neuroimage*, 37(1):116–129, 2007.
- [100] S. Jbabdi, MW Woolrich, JLR Andersson, and TEJ Behrens. A Bayesian framework for global tractography. *Neuroimage*, 37(1):116–129, 2007.
- [101] J.H. Jensen, J.A. Helpern, A. Ramani, H. Lu, and K. 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.
- [102] R. Jianu, C. Demiralp, and D.H. Laidlaw. Exploring 3D DTI fiber tracts with linked 2D representations. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1449–1456, 2009.
- [103] Heidi Johansen-Berg and Timothy E. J. Behrens, editors. *Diffusion MRI*. Academic Press, 2009.
- [104] L. Jonasson, P. Hagmann, JP Thiran, and VJ Wedeen. Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering. In *Int. Soc. Mag. Reson. Med*, page 1310, 2005.
- [105] D.K. Jones. Tractography gone wild: probabilistic fibre tracking using the wild bootstrap with diffusion tensor mri. *Medical Imaging, IEEE Transactions on*, 27(9):1268–1274, 2008.
- [106] DK Jones, MA Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *optimization*, 525, 1999.
- [107] R. A. Kanaan, S. S. Shergill, G. J. Barker, M. Catani, V. W. Ng, R. Howard, P. K. McGuire, and D. K. Jones. Tract-specific anisotropy measurements in diffusion tensor imaging. *Psychiatry Res*, 146(1):73–82, 2006.
- [108] N. Kang, J. Zhang, E.S. Carlson, and D. Gembiris. White matter fiber tractography via anisotropic diffusion simulation in the human brain. *Medical Imaging, IEEE Transactions on*, 24(9):1127–1137, 2005.

- [109] M.D. King, J. Houseman, S.A. Roussel, N. Van Bruggen, S.R. Williams, and D.G. Gadian. q-Space imaging of the brain. *Magnetic Resonance in Medicine*, 32(6):707–713, 1994.
- [110] B. W. Kreher, I. Mader, and V. G. Kiselev. Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. *Proc. Intl. Soc. Mag. Reson. Med.*, 16:425, 2008.
- [111] BW Kreher, I. Mader, and VG Kiselev. Gibbs tracking: a novel approach for the reconstruction of neuronal pathways. *Magnetic Resonance in Medicine*, 60(4):953–963, 2008.
- [112] Harold W Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.
- [113] Ludmila I. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.
- [114] L W Kuo, J H Chen, V J Wedeen, and W Y Tseng. Optimization of diffusion spectrum imaging and q-ball imaging on clinical MRI system. *Neuroimage*, vol:41pp7–18.
- [115] Eugene L. Lawler. *Combinatorial optimization: networks and matroids*. Dover Publications, Mineola, N.Y., 2001.
- [116] M. Lazar and AL Alexander. White matter tractography using random vector (RAVE) perturbation. In *Proceedings of ISMRM Annual Meeting, Honolulu*, 2002.
- [117] M. Lazar and A.L. Alexander. Bootstrap white matter tractography (BOOT-TRAC). *NeuroImage*, 24(2):524–532, 2005.
- [118] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field, and A. L. Alexander. White matter tractography using diffusion tensor deflection. *Hum Brain Mapp*, 18(4):306–21, 2003.
- [119] P.M. Lee. *Bayesian statistics*. Arnold London, UK:, 1997.

- [120] A. Leemans and DK Jones. A new approach to fully automated fiber tract clustering using affinity propagation. In *Int. Symp. on Magnetic Resonance in Medicine (ISMRM)*, volume 17, page 856, 2009.
- [121] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, and P. Parizel. Multiscale white matter fiber tract coregistration: A new feature-based approach to align diffusion tensor data. *Magnetic Resonance in Medicine*, 55(6):1414–1423, 2006.
- [122] C. Lenglet, J.S.W. Campbell, M. Descoteaux, G. Haro, P. Savadjiev, D. Wassermann, A. Anwander, R. Deriche, G.B. Pike, G. Sapiro, et al. Mathematical methods for diffusion mri processing. *Neuroimage*, 45(1):S111–S122, 2009.
- [123] S. Lifshits, A. Tamir, and Y. Assaf. Combinatorial fiber-tracking of the human brain. *NeuroImage*, 48(3):532–540, 2009.
- [124] C. Liu, R. Bammer, B. Acar, and M.E. Moseley. Characterizing non-gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51(5):924–937, 2004.
- [125] H. Lu, J.H. Jensen, A. Ramani, and J.A. Helpern. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, 19(2):236–247, 2006.
- [126] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. California, USA, 1967.
- [127] M. Maddah. *Quantitative Analysis of Cerebral White Matter Anatomy from Diffusion MRI*. PhD thesis, Massachusetts Institute of Technology, 2008.
- [128] M. Maddah, W.E.L. Grimson, and S.K. Warfield. Statistical modeling and EM clustering of white matter fiber tracts. In *IEEE Int. Symp. Biomedical Imaging*, pages 53–56. Citeseer, 2006.

- [129] M. Maddah, A. U. Mewes, S. Haker, W. E. Grimson, and S. K. Warfield. Automated atlas-based clustering of white matter fiber tracts from DT-MRI. *International Conference in Medical Image Computing and Computer Assisted Intervention*, 8(Pt 1):188–95, 2005.
- [130] M. Maddah, W. M. Wells, 3rd, S. K. Warfield, C. F. Westin, and W. E. Grimson. Probabilistic clustering and quantitative analysis of white matter fiber tracts. *Inf Process Med Imaging*, 20:372–83, 2007.
- [131] M. Maddah, L. Zollei, W.E.L. Grimson, C.F. Westin, and W.M. Wells. A mathematical framework for incorporating anatomical knowledge in dt-mri analysis. In *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 105–108. IEEE, 2008.
- [132] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
- [133] J.G. Malcolm, M.E. Shenton, and Y. Rathi. Filtered Tractography: State estimation in a constrained subspace. *Proceedings of MICCAI (Diffusion Modeling and Fiber Cup)*, pages 122–133, 2009.
- [134] G. Margolis and JP Pickett. New applications of the Luxol fast blue myelin stain. *Laboratory investigation; a journal of technical methods and pathology*, 5(6):459.
- [135] Matt A. Bernstein, Kevin F. King, and Xiaohong Joe Zhou. *Handbook of MRI Pulse Sequences*. Elsevier Academic Press, 2004.
- [136] A. Mayer and H. Greenspan. Bundles of interest based registration of White Matter tractographies. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 919–922. IEEE, 2008.
- [137] A. Mayer, G. Zimmerman-Moreno, R. Shadmi, A. Batikoff, and H. Greenspan. A Supervised Framework for the Registration and Segmentation of White Matter Fiber Tracts. *Medical Imaging, IEEE Transactions on*, 30(1):131–145, 2011.

- [138] Arnaldo Mayer and Hayit Greenspan. Direct registration of white matter tractographies with application to atlas construction. In *MICCAI 2007 Workshop Statistical Registration PairWise and GroupWise Alignment and Atlas Formation*, 2007.
- [139] D.W. McRobbie, E.A. Moore, and M.J. Graves. *MRI from Picture to Proton*. Cambridge University Press, 2006.
- [140] K.D. Merboldt, W. Hanicke, H. Bruhn, M.L. Gyngell, and J. Frahm. Diffusion imaging of the human brain *in vivo* using high-speed STEAM MRI. *Magnetic Resonance in Medicine*, 23(1):179–192, 1992.
- [141] B. Mobergs, A. Vilanova, and J.J. van Wijk. Evaluation of fiber clustering methods for diffusion tensor imaging. In *IEEE Visualization*, pages 65–72, 2005.
- [142] D.C. Montgomery, E.A. Peck, G.G. Vining, and J. Vining. *Introduction to linear regression analysis*, volume 3. Wiley New York, 2001.
- [143] S. Mori, B. J. Crain, V. P. Chacko, and P. C. van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol*, 45(2):265–9, 1999.
- [144] S. Mori, S. Wakana, LM Nagae-Poetscher, and PCM Van Zijl. MRI atlas of human white matter. *American Journal of Neuroradiology*.
- [145] Susumu Mori and Jiangyang Zhang. Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*, 51:527–39, 2006.
- [146] M.E. Moseley, K. Butts, M.A. Yenari, M. Marks, and A.D. Crespigny. Clinical aspects of dwi. *NMR in Biomedicine*, 8(7):387–396, 1995.
- [147] R.T. Ng and J. Han. CLARANS: A method for clustering objects for spatial data mining. *IEEE Transactions on Knowledge and Data Engineering*, pages 1003–1016, 2002.

- [148] L. J. O'Donnell and C. F. Westin. Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Trans Med Imaging*, 26(11):1562–75, 2007.
- [149] L. J. O'Donnell, C. F. Westin, and A. J. Golby. Tract-based morphometry for white matter group analysis. *Neuroimage*, 45(3):832–44, 2009.
- [150] L.J. O'Donnell. *Cerebral White Matter Analysis Using Diffusion Imaging*. PhD thesis, 2006.
- [151] L.J. O'Donnell, C.F. Westin, and A.J. Golby. Tract-based morphometry for white matter group analysis. *NeuroImage*, 45(3):832–844, 2009.
- [152] R.L. O'Gorman and D.K. Jones. Just how much data need to be collected for reliable bootstrap dt-mri? *Magnetic resonance in medicine*, 56(4):884–890, 2006.
- [153] F.W. Olver, D.W. Lozier, R.F. Boisvert, and C.W. Clark. NIST handbook of mathematical functions. 2010.
- [154] H.H. Ong and F.W. Wehrli. Quantifying axon diameter and intracellular volume fraction in excised mouse spinal cord with q-space imaging. *Neuroimage*, 51(4):1360–1366, 2010.
- [155] E.
"Ozarslan and T.H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50(5):955–965, 2003.
- [156] E. Ozarslan, T.M. Shepherd, B.C. Vemuri, S.J. Blackband, and T.H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (dot). *NeuroImage*, 31(3):1086–1103, 2006.
- [157] S. Pajevic and P.J. Basser. Parametric and non-parametric statistical analysis of dt-mri data. *Journal of Magnetic Resonance*, 161(1):1–14, 2003.

- [158] G. Parker and D. Alexander. Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *Information Processing in Medical Imaging*, pages 684–695. Springer, 2003.
- [159] G. J. Parker, H. A. Haroon, and C. A. Wheeler-Kingshott. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. *J Magn Reson Imaging*, 18(2):242–54, 2003.
- [160] Geoffrey J M Parker, Hamied a Haroon, and Claudia a M Wheeler-Kingshott. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. *Journal of Magnetic Resonance Imaging*, 18(2):242–54, 2003.
- [161] G.J.M. Parker and D.C. Alexander. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):893, 2005.
- [162] G.J.M. Parker, C.A.M. Wheeler-Kingshott, and G.J. Barker. Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *Medical Imaging, IEEE Transactions on*, 21(5):505–512, 2002.
- [163] O. Pasternak, Y. Assaf, N. Intrator, and N. Sochen. Variational multiple-tensor fitting of fiber-ambiguous diffusion-weighted magnetic resonance imaging voxels. *Magnetic resonance imaging*, 26(8):1133–1144, 2008.
- [164] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [165] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Riviere, A. Constantinesco, D. Le Bihan, and J. F. Mangin. Fiber tracking in q-ball fields using regularized particle trajectories. *Inf Process Med Imaging*, 19:52–63, 2005.

- [166] V. Pickalov and P.J. Basser. 3d tomographic reconstruction of the average propagator from mri data. In *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006*, pages 710–713, 2006.
- [167] P. Ramachandran and G. Varoquaux. Mayavi: 3d visualization of scientific data. *Computing in Science & Engineering*, 13(2):40–51, 2011.
- [168] T.G. Reese, O. Heid, R.M. Weisskoff, and V.J. 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.
- [169] M. Reisert, I. Mader, C. Anastasopoulos, M. Weigel, S. Schnell, and V. Kiselev. Global fiber reconstruction becomes practical. *Neuroimage*, 2010.
- [170] J.A. Rice. *Mathematical statistics and data analysis*. Thomson Learning, 2006.
- [171] J.K. Rilling, M.F. Glasser, T.M. Preuss, X. Ma, T. Zhao, X. Hu, and T.E.J. Behrens. The evolution of the arcuate fasciculus revealed with comparative dti. *Nature Neuroscience*, 11(4):426–428, 2008.
- [172] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. *International Conference in Medical Image Computing and Computer-Assisted Intervention*, page 1115, 1998.
- [173] K.E. Sakaie and M.J. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted mri. *NeuroImage*, 34(1):169–176, 2007.
- [174] K.K. Seunarine, PA Cook, MG Hall, KV Embleton, GJM Parker, and D.C. Alexander. Exploiting peak anisotropy for tracking through complex structures. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.

- [175] A. Sherbondy, M. Rowe, and D. Alexander. Microtrack: An algorithm for concurrent projectome and microstructure estimation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 183–190, 2010.
- [176] AJ Sherbondy, DL Akers, RF Dougherty, M. Ben-Shachar, S. Napel, and BA Wandell. MetroTrac: A metropolis algorithm for probabilistic tractography. *Human Brain Mapping, Florence*, 2006.
- [177] Anthony J. Sherbondy, Robert F. Dougherty, Michal Ben-Shachar, Sandy Napel, and Brian A. Wandell. ConTrack: Finding the most likely pathways between brain regions using diffusion tractography. *J. Vis.*, 8(9):1–16, 7 2008.
- [178] Stephen M. Smith, Mark Jenkinson, Heidi Johansen-Berg, Daniel Rueckert, Thomas E. Nichols, Clare E. Mackay, Kate E. Watkins, Olga Ciccarelli, Zaheer Cader, Paul M. Matthews, and Timothy E.J. Behrens. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31:1487–1505, 2006.
- [179] S.N. Sotiropoulos. Processing of diffusion mr images of the brain: from crossing fibres to distributed tractography. 2010.
- [180] S.N. Sotiropoulos, L. Bai, P.S. Morgan, C.S. Constantinescu, and C.R. Tench. Brain tractography using Q-ball imaging and graph theory: Improved connectivities through fibre crossings via a model-based approach. *NeuroImage*, 49(3):2444–2456, 2010.
- [181] H. Steinhaus. Sur la division des corp materiels en parties. *Bull. Acad. Polon. Sci*, 1:801–804, 1956.
- [182] E. O. Stejskal and J. E. 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.
- [183] H.C. Torrey. Bloch equations with diffusion terms. *Physical Review*, 104(3):563–565, 1956.

- [184] J. Tournier et al. Diffusion-weighted magnetic resonance imaging fibre tracking using a front evolution algorithm. *NeuroImage*, 20(1):276–288, 2003.
- [185] J. Tournier et al. Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.
- [186] 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):1459–72, 2007.
- [187] A. Tsai, C.F. Westin, A.O. Hero, and A.S. Willsky. Fiber tract clustering on manifolds with dual rooted-graphs. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–6. IEEE, 2007.
- [188] David S. Tuch, Timothy G. Reese, Mette R. Wiegell, Nikos Makris, John W. Belliveau, and Van J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582, 2002.
- [189] David S Tuch, Jonathan J Wisco, Mark H Khachaturian, Wim Vanduffel, Leeland B Ekstrom, and Rolf Ko. Q-ball imaging of macaque white matter architecture. (May):869–879, 2005.
- [190] D.S. Tuch. *Diffusion {MRI} of complex tissue structure*. PhD thesis, Massachusetts Institute of Technology, Division of Health Sciences and Technology, 2002.
- [191] DS Tuch. Q-ball imaging. *change*, 2004.
- [192] Eelke Visser, Emil H J Nijhuis, Jan K Buitelaar, and Marcel P Zwiers. Partition-based mass clustering of tractography streamlines. *NeuroImage*, July 2010.
- [193] Setsu Wakana, Arvind Caprihan, Martina M. Panzenboeck, James H. Fallon, Michele Perry, Randy L. Gollub, Kegang Hua, Jiangyang

- Zhang, Hangyi Jiang, Prachi Dubey, Ari Blitz, Peter van Zijl, and Susumu Mori. Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*, 36(1):630–644, 2007.
- [194] X. Wang, W.E.L. Grimson, and C.F. Westin. Tractography segmentation using a hierarchical Dirichlet processes mixture model. *NeuroImage*, 2010.
- [195] S. Warach, J. Gaa, B. Siewert, P. Wielopolski, and R.R. Edelman. Acute human stroke studied by whole brain echo planar diffusion-weighted magnetic resonance imaging. *Annals of neurology*, 37(2):231–241, 1995.
- [196] V Wedeen, R Wang, J Schmahmann, T Benner, W Tseng, G Dai, D Pandya, P Hagmann, H Dâ arceuil, and A. de Crespigny, "Diffusion spectrum magnetic resonance imaging (dsi) tractography of crossing fibers,". *NeuroImage*, vol:41no4pp1267–1277.
- [197] V. J. Wedeen, R. P. Wang, J. D. Schmahmann, T. Benner, W. Y. Tseng, G. Dai, D. N. Pandya, P. Hagmann, H. D'Arceuil, and A. J. de Crespigny. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage*, 41(4):1267–77, 2008.
- [198] V.J. Wedeen, P. Hagmann, W.Y.I. Tseng, T.G. Reese, and R.M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386, 2005.
- [199] VJ Wedeen, RP Wang, JD Schmahmann, and T Benner. ... spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage*, 2008.
- [200] B. Whitcher, D.S. Tuch, J.J. Wisco, A.G. Sorensen, and L. Wang. Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging. *Human brain mapping*, 29(3):346–362, 2008.
- [201] T.H. Williams, N. Gluhbegovic, and J.Y. Jew. *The human brain: dissections of the real brain*. University of Iowa, 1999.

- [202] R. Wirestam, A. Bibic, J. Latt, S. Brockstedt, and F. Stahlberg. Denoising of complex MRI data by wavelet-domain filtering: Application to high-*b*-value diffusion-weighted imaging. *Magnetic Resonance in Medicine*, 56(5), 2006.
- [203] S.F. Witelson. Hand and sex differences in the isthmus and genu of the human corpus callosum. *Brain*, 112(3):799, 1989.
- [204] Y.C. Wu and A.L. Alexander. Hybrid diffusion imaging. *NeuroImage*, 36(3):617–629, 2007.
- [205] F Yeh, V Wedeen, and W Tseng. Generalized Q-sampling imaging. *IEEE Transactions on Medical Imaging*, 29(9):1626–1635, March 2010.
- [206] F.C. Yeh, V.J. Wedeen, and W.Y.I. Tseng. Estimation of fiber orientation and spin density distribution by diffusion deconvolution. *Neuroimage*, 2011.
- [207] A. Zalesky. Dt-mri fiber tracking: a shortest paths approach. *Medical Imaging, IEEE Transactions on*, 27(10):1458–1471, 2008.
- [208] G. Zhai, W. Lin, K. P. Wilber, G. Gerig, and J. H. Gilmore. Comparisons of regional white matter diffusion in healthy neonates and adults performed with a 3.0-T head-only MR imaging unit. *Radiology*, 229(3):673–81, 2003.
- [209] F. Zhang, E.R. Hancock, C. Goodlett, and G. Gerig. Probabilistic white matter fiber tracking using particle filtering and von mises-fisher sampling. *Medical image analysis*, 13(1):5–18, 2009.
- [210] S. Zhang, S. Correia, and D.H. Laidlaw. Identifying White-Matter Fiber Bundles in DTI Data Using an Automated Proximity-Based Fiber Clustering Method. *IEEE transactions on visualization and computer graphics*, 14(5):1044, 2008.
- [211] S. Zhang et al. Visualizing diffusion tensor MR images using streamtubes and streamsurfaces. *IEEE Transactions on Visualization and Computer Graphics*, pages 454–462, 2003.

- [212] S. Zhang and DH Laidlaw. DTI fiber clustering and cross-subject cluster analysis. In *Int. Soc. Magn. Reson. Med*, 2005.
- [213] T. Zhang, R. Ramakrishnan, and M. Livny. BIRCH: A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery*, 1(2):141–182, 1997.
- [214] J. Zhuang, J. Hrabe, A. Kangarlu, D. Xu, R. Bansal, C. A. Branch, and B. S. Peterson. Correction of eddy-current distortions in diffusion tensor images using the known directions and strengths of diffusion gradients. *J Magn Reson Imaging*, 24(5):1188–93, 2006.
- [215] U. Ziyan, M. R. Sabuncu, L. J. O'Donnell, and C. F. Westin. Nonlinear registration of diffusion MR images based on fiber bundles. *International Conference in Medical Image Computing and Computer-Assisted Intervention*, 10(Pt 1):351–8, 2007.
- [216] U. Ziyan, M.R. Sabuncu, W.E.L. Grimson, and C.F. Westin. Consistency clustering: a robust algorithm for group-wise registration, segmentation and automatic atlas construction in diffusion MRI. *International Journal of Computer Vision*, 85(3):279–290, 2009.
- [217] O. Zvitia, A. Mayer, and H. Greenspan. Adaptive mean-shift registration of white matter tractographies. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 692–695. IEEE, 2008.
- [218] Orly Zvitia, Arnaldo Mayer, Ran Shadmi, Shmuel Miron, and Hayit K Greenspan. Co-registration of white matter tractographies by adaptive-mean-shift and Gaussian mixture modeling. *IEEE transactions on medical imaging*, 29(1):132–45, January 2010.

6 Appendix

6.1 Data and sequences

12 Channels 101 and 118 MRC-CBU

3T scanner (TIM Trio, Siemens), using Siemens advanced diffusion work-in-progress sequence, and STEAM [140, 135] as the diffusion preparation method. The field of view was $240 \times 240 \text{ mm}^2$, matrix size 96×96 , and slice thickness 2.5 mm (no gap). 55 slices were acquired to achieve full brain coverage, and the voxel resolution was $2.5 \times 2.5 \times 2.5 \text{ mm}^3$. Two sampling schemes were considered: a 102-point half grid acquisition (TR=8200ms, TE=69ms) with a maximum b-value of 4000 s/mm^2 , and a single shell acquisition using 118 non-collinear gradient directions (TR=7000ms, TE=47ms) and a b-value of 1000 s/mm^2 . The two acquisition schemes were matched for total acquisition time ($14 \text{ min } 37\text{s}$), voxel resolution, and bandwidth.

32 Channels 101 MRC-CBU

3T scanner (TIM Trio, Siemens), using Siemens advanced diffusion work-in-progress sequence, and STEAM [140, 135] as the diffusion preparation method. The field of view was $240 \times 240 \text{ mm}^2$, matrix size 96×96 , and slice thickness 2.5 mm (no gap). 55 slices were acquired to achieve full brain coverage, and the voxel resolution was $2.5 \times 2.5 \times 2.5 \text{ mm}^3$. A 102-point half grid acquisition with a maximum b-value of 4000 s/mm^2 was used. The total acquisition time was $14 \text{ min } 21\text{s}$ with TR=8200ms and TE=69ms.

32 Channels EPFL

3T scanner (TIM Trio, Siemens). The field of view was $210 \times 210 \text{ mm}^2$, matrix size 96×96 , and slice thickness 3 mm . 44 slices were acquired and the voxel resolution was $2.2 \times 2.2 \times 3.0 \text{ mm}^3$. A 258-point half grid acquisition scheme with a maximum b-value of 8011 s/mm^2 (DSI15) was used. The total acquisition time was 34 min with TR=8200ms and TE=165ms.

6.2 The cosine transform

$\int_0^\infty \cos(st)g(t)dt$ where $g(t)$ defined on $[t, \infty)$

Let $f(t)$ be an even function $f(t) = f(-t)$ defined for $-\infty < t < \infty$

$$\begin{aligned}
F(s) &= \int_{-\infty}^{\infty} f(t)e^{its}dt \\
&= \int_0^{\infty} f(t)e^{its}dt + \int_{-\infty}^0 f(t)e^{its}dt \\
&= \int_0^{\infty} f(t)e^{its}dt - \int_{-\infty}^0 f(-t)e^{-its}dt \\
&= \int_0^{\infty} f(t)e^{its}dt + \int_0^{\infty} f(t)e^{-its}dt \\
&= \int_0^{\infty} f(t)[e^{its} + e^{-its}]dt \\
&= \int_0^{\infty} f(t)[\cos(its) + i\sin(its) + \cos(its) - i\sin(its)]dt \\
&= 2 \int_0^{\infty} f(t)\cos(st)dt
\end{aligned}$$

In the third row above we replace $t \rightarrow -t$

If you want to compute the integral in the whole space because of symmetry you have $F(s) = \int_{-\infty}^{\infty} f(t)\cos(st)dt$

6.3 Fourier transform of $P(\mathbf{r})r^2$

From Fourier analysis we know that if $E(\mathbf{q})$ is the Fourier transform function of $P(\mathbf{r})$ then

$$\begin{aligned}
\mathfrak{F}(xP(\mathbf{r})) &= i \frac{\partial E(\mathbf{q})}{\partial \mathbf{q}_x} \\
\mathfrak{F}(x^2P(\mathbf{r})) &= - \frac{\partial^2 E(\mathbf{q})}{\partial \mathbf{q}_x^2}
\end{aligned}$$

where $\mathfrak{F}()$ is the Fourier transform. By writing the second equation for y and z and summing them all together we obtain

$$\mathfrak{F}(r^2 P(\mathbf{r})) = -\frac{\partial^2 E(\mathbf{q})}{\partial \mathbf{q}_x^2} - \frac{\partial^2 E(\mathbf{q})}{\partial \mathbf{q}_y^2} - \frac{\partial^2 E(\mathbf{q})}{\partial \mathbf{q}_z^2} = -\nabla^2 E(\mathbf{q})$$

6.4 Radial projection of a symmetric function

Let $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ be a symmetric function with the 3D Fourier transform function $\hat{f}(\mathbf{q})$ and $\hat{\mathbf{u}}$ be an arbitrary unit vector. We will show that $\int_0^\infty f(r\hat{\mathbf{u}}) dr = \frac{1}{8\pi^2} \int \int_{\hat{\mathbf{u}}^\perp} \hat{f}(\mathbf{q}) q dq d\phi$ where $\hat{\mathbf{u}}^\perp$ is the plane perpendicular to $\hat{\mathbf{u}}$.

Without loss of generality, we align $\hat{\mathbf{u}}$ with the z-axis having $\hat{\mathbf{z}} = \hat{\mathbf{u}}$. Using the Dirac delta function (make use of Lebesgue integral) we can now write

$$\begin{aligned} \int_0^\infty f(r\hat{\mathbf{z}}) dr &= \int_0^\infty f(0, 0, z) dz \\ &= \frac{1}{2} \int \int \int_{\mathbb{R}^3} f(x, y, z) \delta(x) \delta(y) dx dy dz \end{aligned}$$

The factor $1/2$ is required because we need the integral only in the positive half of the z-axis, and the function is symmetric. Let us define $g(x, y, z) \equiv \delta(x)\delta(y)$. For the two functions $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ with Fourier transform functions $\hat{f}(\mathbf{q})$ and $\hat{g}(\mathbf{q})$, Parseval's theorem states that

$$\int \int \int_{\mathbb{R}^3} f(x, y, z) g^*(x, y, z) dx dy dz = (2\pi)^{-3} \int \int \int_{\mathbb{R}^3} f(q_x, q_y, q_z) \hat{g}^*(q_x, q_y, q_z) dq_x dq_y dq_z$$

Furthermore, $\hat{g}(q_x, q_y, q_z) = 2\pi\delta(q_z)$ and replacing it in the above equations leads to

$$\begin{aligned} \int_0^\infty f(r\hat{\mathbf{z}}) dr &= \int \int \int_{\mathbb{R}^3} \frac{1}{2} f(x, y, z) g(x, y, z) dx dy dz \\ &= \frac{1}{2(2\pi)^3} \int \int \int_{\mathbb{R}^3} f(q_x, q_y, q_z) 2\pi\delta(q_z) dq_x dq_y dq_z \\ &= \frac{1}{8\pi^2} \int_{-\infty}^\infty \hat{f}(q_x, q_y, 0) dq_x dq_y \end{aligned}$$

6.5 The Tensor in GQI

We now apply this formulation under the assumption that the diffusion voxel can be represented by a single tensor model. eq. 12 can be written in the form

$$S(\mathbf{q}) = S_0 \exp(-b\mathbf{q}^T D \mathbf{q}) \quad (50)$$

where S_0 is the image when b-value is equal to 0, b is the b-value for a specific direction and D is a 3×3 matrix, known as the diffusion tensor. Then from eq. 19 and 50 the Fourier transform of S is equal to

$$Q(\mathbf{R}) = \int S_0 \exp(-b\mathbf{q}^T D \mathbf{q}) \exp(-j2\pi\mathbf{q} \cdot \mathbf{R}) d\mathbf{q} \quad (51)$$

The same equation in its triple integral form can be written as

$$\begin{aligned} Q(\mathbf{R}) &= S_0 \iiint \exp\left(-b \sum_{i=1}^3 q_i^2 \lambda_i - j2\pi \sum_{i=1}^3 q_i R_i\right) dq_1 dq_2 dq_3 \\ &= S_0 \iiint \prod_{i=1}^3 \exp(-bq_i^2 \lambda_i - j2\pi q_i R_i) dq_1 dq_2 dq_3 \\ &= S_0 \prod_{i=1}^3 \int \exp(-bq_i^2 \lambda_i - j2\pi q_i R_i) dq_i \\ &= S_0 \prod_{i=1}^3 \int \exp\left(-b\lambda_i[q_i^2 + \frac{j2\pi R_i}{b\lambda_i} q_i]\right) dq_i \\ &= S_0 \prod_{i=1}^3 \int \exp\left\{-b\lambda_i[(q_i + \frac{j\pi R_i}{b\lambda_i})^2 + \frac{\pi^2 R_i^2}{b^2 \lambda_i^2}]\right\} \\ &= S_0 \prod_{i=1}^3 \int \exp\left\{-b\lambda_i(q_i + \frac{j\pi R_i}{b\lambda_i})^2\right\} \exp\left\{-\frac{\pi^2 R_i^2}{b\lambda_i}\right\} \end{aligned} \quad (52)$$

In that stage we could try to make use of the formula $\int \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1$. Now we can see that $b\lambda_i = 1/2\sigma^2$ and μ corresponds to $\mu = -jR_i/b\lambda_i$.

Therefore, eq. 52 can now be written as

$$\begin{aligned} Q(\mathbf{R}) &= S_0 \prod_{i=1}^3 \sqrt{\frac{\pi}{b\lambda_i}} \exp\left(-\frac{\pi^2 R_i^2}{b\lambda_i}\right) \\ &= S_0 \left(\frac{\pi}{b}\right)^{3/2} \frac{1}{\sqrt{\prod_{i=1}^3 \lambda_i}} \exp\left(-\frac{\pi^2}{b} \mathbf{R}^T D^{-1} \mathbf{R}\right) \end{aligned} \quad (53)$$

where D is the diffusion tensor. We can replace the displacement vector \mathbf{R} with a scalar value L and a unit vector $\hat{\mathbf{u}}$ i.e. $\mathbf{R} = L\hat{\mathbf{u}}$ and from eq. 53 we can replace $\frac{2\pi^2}{b} \hat{\mathbf{u}}^T D^{-1} \hat{\mathbf{u}}$ with k and $S_0 \left(\frac{\pi}{b}\right)^{3/2} \frac{1}{\lambda_1 \lambda_2 \lambda_3}$ with α . Using that last change of variables we can now write

$$\begin{aligned} \psi_Q(\mathbf{r}, \hat{\mathbf{u}}) &= \int_0^{L_\Delta} Q(\mathbf{r}, L\hat{\mathbf{u}}) dL \\ &= \alpha \int_0^{L_\Delta} \exp\left(-L^2 \frac{k}{2}\right) dL \end{aligned} \quad (54)$$

Setting $m = \sqrt{k}L$ and using the derivation for the error function eq.54 illustrates the remarkable result that we can calculate analytically the spin ODF for Gaussian diffusion using the cumulative distribution function CDF .

$$\begin{aligned} \psi_Q(\hat{\mathbf{u}}) &= \frac{\alpha}{\sqrt{k}} \int_0^{\sqrt{k}L_\Delta} e^{-m^2/2} dm \\ &= \alpha \sqrt{\frac{2\pi}{k}} \left[CDF(\sqrt{k}L_\Delta) - \frac{1}{2} \right] \end{aligned} \quad (55)$$

This can be used as a check to compare the approximated/sampled spin odf that is derived in [205] with eq. 56 for the case of gaussian diffusion.

What is also very interesting is to try to derive what the normalization factor should be for the spin odf in eq.54. Because calculating a spherical intergral from eq.56 seems at the moment very complicated we first work

with the simpler gaussian diffusion ODF derived by Tuch [191]

$$\psi_{p_\Delta} = \frac{1}{Z} \sqrt{\frac{\pi\tau}{\mathbf{u}^T D^{-1} \mathbf{u}}} \quad (57)$$

$$\frac{Z}{\sqrt{\pi\tau}} = \iint_{S^2} (\mathbf{u}^T D^{-1} \mathbf{u})^{-\frac{1}{2}} d\mathbf{u} \quad (58)$$

Lets now define $f_D(\mathbf{u}) = (\mathbf{u}^T D^{-1} \mathbf{u})^{-\frac{1}{2}}$. From [153] (19.31.2) we know that we can calculate the following integral on the entire space (eq. 59). By expanding it in polar form we can find surface integral needed in eq. 58

$$MHG = \iiint f_D(x) e^{-|x|^2} dx \quad (59)$$

$$= \int_0^\infty \left[\iint f_D(r\mathbf{u}) d\mathbf{u} \right] e^{-r^2} r^2 dr, \quad x = r\mathbf{u} \text{ (polar)} \quad (60)$$

However we know that $f_D(r\mathbf{u}) = ((r\mathbf{u})^T D^{-1} (r\mathbf{u}))^{-\frac{1}{2}} = r^{-1} (\mathbf{u}^T D^{-1} \mathbf{u})^{-\frac{1}{2}}$
therefore

$$\begin{aligned} MHG &= \int_0^\infty \left[\iint f_D(u) d\mathbf{u} \right] e^{-r^2} r dr \\ &= \iint f_D(u) du \int_0^\infty r e^{-r^2} dr \\ &= \frac{1}{2} \iint f_D(u) du \end{aligned}$$

Consequently, $\frac{Z}{\sqrt{\pi\tau}} = 2MHG$ where MHG is the multivariate hyper-geometric function with $\mu = -\frac{1}{2}, B = I, n = 3$ and $\lambda_1, \lambda_2, \lambda_3$ the eigenvalues of \mathbf{A} de-

rived from [153] (19.31.2) & (19.16.9). Therefore,

$$\begin{aligned}\frac{Z}{\sqrt{\pi\tau}} &= 2MHG \\ &= \frac{2\pi^{\frac{3}{2}}\Gamma(1)}{\sqrt{\det(I)\Gamma(\frac{3}{2})}} R_{-\frac{1}{2}}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; \lambda_1, \lambda_2, \lambda_3\right)\end{aligned}$$

and $R_{-\frac{1}{2}} = \frac{1}{2} \int_0^\infty t^0(t + \lambda_1)^{-\frac{1}{2}}(t + \lambda_2)^{-\frac{1}{2}}(t + \lambda_3)^{-\frac{1}{2}} dt$ with $\alpha = \frac{1}{2}$ and $\alpha' = 1$.

Given $\lambda_1, \lambda_2, \lambda_3$ we can integrate numerically or even possibly analytically. For the isotropic case the integral simplifies to $\frac{1}{2} \int_0^\infty t^0(t + \lambda)^{-\frac{3}{2}} dt = \frac{1}{\sqrt{\lambda}}$ and for the cylindrical case ($\lambda_2 = \lambda_3$) to $\frac{1}{2} \int_0^\infty (t + \lambda_1)^{-\frac{1}{2}}(t + \lambda_2)^{-1} dt$.

References

- [1] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel. Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2):554–566, 2010.
- [2] Iman Aganj, Christophe Lenglet, Neda Jahanshad, Essa Yacoub, Noam Harel, Paul M Thompson, I M A Preprint Series, and Church Street S E. A HOUGH TRANSFORM GLOBAL PROBABILISTIC APPROACH A Hough Transform Global Probabilistic Approach to Multiple- Subject Diffusion MRI Tractography. pages 612–626, 2010.
- [3] DC Alexander, GJ Barker, and SR Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.
- [4] D.C. Alexander, P.L. Hubbard, M.G. Hall, E.A. Moore, M. Ptito, G.J.M. Parker, and T.B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion mri. *NeuroImage*, 52(4):1374–1389, 2010.
- [5] J.L.R. Andersson and S. Skare. A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *Neuroimage*, 16(1):177–199, 2002.
- [6] A. Anwander, M. Tittgemeyer, D Y von Cramon, A D Friederici, and T R Knosche. Connectivity-Based Parcellation of Broca’s Area. *Cerebral Cortex*, 17(4):816, 2007.
- [7] Vincent Arsigny, Olivier Commowick, Nicholas Ayache, and Xavier Pennec. A Fast and Log-Euclidean Polyaffine Framework for Locally Linear Registration. *Journal of Mathematical Imaging and Vision*, 33(2):222–238, January 2009.
- [8] Y. Assaf and P.J. Basser. Composite hindered and restricted model of diffusion (charmed) mr imaging of the human brain. *Neuroimage*, 27(1):48–58, 2005.

- [9] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P.J. Basser. AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, 59(6):1347–1354, 2008.
- [10] Y. Assaf, R.Z. Freidlin, G.K. Rohde, and P.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.
- [11] E. J. Auerbach and K. Ugurbil. Improvement in diffusion mri at 3t and beyond with the twice-refocused adiabatic spin echo (trase) sequence. 2004.
- [12] A. Barmpoutis, M.S. Hwang, D. Howland, J.R. Forder, and B.C. Vemuri. Regularized positive-definite fourth order tensor field estimation from dw-mri. *NeuroImage*, 45(1):S153–S162, 2009.
- [13] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magn Reson Med*, 44(4):625–32, 2000.
- [14] PJ Basser, J Mattiello, and D LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 1994.
- [15] P. Batchelor, D. Hill, D. Atkinson, and F. Calamante. Study of connectivity in the brain using the full diffusion tensor from mri. In *Information Processing in Medical Imaging*, pages 121–133. Springer, 2001.
- [16] C. Beaulieu. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed*, 15(7-8):435–55, 2002.
- [17] T E Behrens and H. Johansen-Berg. Relating connectional architecture to grey matter function using diffusion imaging. *Philos Trans R Soc Lond B Biol Sci*, 360(1457):903–11, 2005.
- [18] T E J Behrens, H Johansen-Berg, M W Woolrich, C A M Wheeler-Kingshott, P A Boulby, G J Barker, E L Sillery, K Sheehan, O Ciccarelli, A J Thompson, J M Brady, and P M Matthews. Non-invasive mapping

- of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6(7):750–757, 2003.
- [19] T. E. J. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted {MR} imaging. *Magnetic Resonance in Medicine*, 50:1077–1088, 2003.
 - [20] T.E.J. Behrens, H. Johansen-Berg, S. Jbabdi, M.F.S. Rushworth, and M.W. Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1):144–155, 2007.
 - [21] J.I. Berman, S.W. Chung, P. Mukherjee, C.P. Hess, E.T. Han, and R.G. Henry. Probabilistic streamline q-ball tractography using the residual bootstrap. *Neuroimage*, 39(1):215–222, 2008.
 - [22] M. Björnemo, A. Brun, R. Kikinis, and C.-F. Westin. Regularized stochastic white matter tractography using diffusion tensor MRI. In *Fifth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, pages 435–442, Tokyo, Japan, 2002.
 - [23] A. Brun, H. Knutsson, H.J. Park, M.E. Shenton, and C.F. Westin. Clustering fiber traces using normalized cuts. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, pages 368–375, 2004.
 - [24] Paul T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford University Press, 1991.
 - [25] PT Callaghan, CD Eccles, and Y. Xia. Nmr microscopy of dynamic displacements: k-space and q-space imaging. *Journal of Physics E: Scientific Instruments*, 21:820, 1988.
 - [26] J.S.W. Campbell, K. Siddiqi, V.V. Rymar, A.F. Sadikot, and G.B. Pike. Flow-based fiber tracking with diffusion tensor and q-ball data: validation and comparison to principal diffusion direction techniques. *NeuroImage*, 27(4):725–736, 2005.

- [27] E.J. Canales-Rodríguez, L. Melie-García, Y. Iturria-Medina, and C.N. Center. Mathematical description of q-space in spherical coordinates: Exact q-ball imaging. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine*, 2009.
- [28] Erick Jorge Canales-Rodríguez, Lester Melie-García, and Yasser Iturria-Medina. Mathematical description of q-space in spherical coordinates: exact q-ball imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 61(6):1350–67, 2009.
- [29] Giorgio Carpaneto, Silvano Martello, and Paolo Toth. Algorithms and codes for the assignment problem. *Annals of Operations Research*, 13:191–223, 1988.
- [30] Giorgio Carpaneto, Silvano Martello, and Paolo Toth. Linear assignment problem freeware. www.assignmentproblems.com/APC_APS.htm, downloaded June 2011.
- [31] M. Catani, D.K. Jones, et al. Perisylvian language networks of the human brain. *Annals of neurology*, 57(1):8–16, 2005.
- [32] Marco Catani, Robert J. Howard, Sinisa Pajevic, and Derek K. Jones. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage*, 17:77–94, 2002.
- [33] M. Chalmers. A linear iteration time layout algorithm for visualising high-dimensional data. In *Visualization'96. Proceedings.*, pages 127–131. IEEE, 1996.
- [34] Y.P. Chao, J.H. Chen, K.H. Cho, C.H. Yeh, K.H. Chou, and C.P. Lin. A multiple streamline approach to high angular resolution diffusion tractography. *Medical engineering & physics*, 30(8):989–996, 2008.
- [35] S.W. Chung, Y. Lu, and R.G. Henry. Comparison of bootstrap approaches for estimation of uncertainties of dti parameters. *Neuroimage*, 33(2):531–541, 2006.

- [36] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci U S A*, 96(18):10422–7, 1999.
- [37] P.A. Cook, D.C. Alexander, and G.J.M. Parker. Modelling noise-induced fibre-orientation error in diffusion-tensor mri. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 332–335. IEEE, 2004.
- [38] I. Corouge, S. Gouttard, and G. Gerig. A statistical shape model of individual fiber tracts extracted from diffusion Tensor MRI. *Medical Image Computing and Computer Assisted Intervention*, pages 671–679, 2004.
- [39] I. Corouge, S. Gouttard, and G. Gerig. Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *International Symposium on Biomedical Imaging*, pages 344–347, 2004.
- [40] Isabelle Corouge, P Thomas Fletcher, Sarang Joshi, Sylvain Gouttard, and Guido Gerig. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Medical image analysis*, 10(5):786–98, 2006.
- [41] Marta Morgado Correia. Development of Methods for the Acquisition and Analysis of Diffusion Weighted MRI Data. *Brain*, (June), 2009.
- [42] M.M. Correia, T.A. Carpenter, and G.B. Williams. Looking for the optimal dti acquisition scheme given a maximum scan time: are more b -values a waste of time? *Magnetic resonance imaging*, 27(2):163–175, 2009.
- [43] Williams G.B. Yeh F-C. Nimmo-Smith I. Correia, M.M. Robustness of diffusion scalar metrics when estimated with generalized q-sampling imaging acquisition schemes. In *19th Annual Meeting of the International Society of Magnetic Resonance in Medicine*, 2011.
- [44] J. Dauguet, S. Peled, V. Berezovskii, T. Delzescaux, S. K. Warfield, R. Born, and C. F. Westin. Comparison of fiber tracts derived from

- in-vivo DTI tractography with 3D histological neural tract tracer reconstruction on a macaque brain. *Neuroimage*, 37(2):530–8, 2007.
- [45] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. A minimum description length approach to statistical shape modeling. *IEEE Trans Med Imaging*, 21(5):525–37, 2002.
- [46] Cagatay Demiralp, John F. Hughes, and David H. Laidlaw. Coloring 3D line fields using Boy’s real projective plane immersion. *IEEE Trans. on Visualization and Computer Graphics (Proc. Visualization ’09)*, 15(6):1457–1463, 2009. [urlhttp://vis.cs.brown.edu/docs/pdf/g/Demiralp-2009-CLF.pdf.html](http://vis.cs.brown.edu/docs/pdf/g/Demiralp-2009-CLF.pdf.html) [urlhttp://vis.cs.brown.edu/results/bibtex/Demiralp-2009-CLF.bib](http://vis.cs.brown.edu/results/bibtex/Demiralp-2009-CLF.bib) (bibtex: Demiralp-2009-CLF).
- [47] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.
- [48] MD Denis Le Bihan, C. Poupon, A. Amadon, and F. Lethimonnier. Artifacts and pitfalls in diffusion MRI. *Journal of Magnetic Resonance Imaging*, 24:478–488, 2006.
- [49] R. Deriche and O. Faugeras. 2-D curve matching using high curvature points: application to stereo vision. In *Pattern Recognition, 1990. Proceedings., 10th International Conference on*, volume 1, 1990.
- [50] M Descoteaux, E Angelino, S Fitzgibbons, and R Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, vol:58no3pp497–510, 2007.
- [51] M. Descoteaux, R. Deriche, and A. Anwander. Deterministic and probabilistic q-ball tractography: from diffusion to sharp fiber distribution. 2007.
- [52] M. Descoteaux, R. Deriche, D.L. Bihan, J.F. Mangin, and C. Poupon. Multiple q-shell diffusion propagator imaging. *Medical Image Analysis*, 2010.

- [53] M. Descoteaux, R. Deriche, T. Knoesche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Trans Med Imaging*, 28(2):269–86, 2009.
- [54] Maxime Descoteaux, Elaine Angelino, Shaun Fitzgibbons, and Rachid Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007.
- [55] Maxime Descoteaux, Rachid Deriche, Thomas R Knösche, and Alfred Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE transactions on medical imaging*, 28(2):269–86, February 2009.
- [56] Z Ding, J Gore, and A Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magn. Reson. Med.*, 49:716–721, 2003.
- [57] Zhaohua Ding, John C Gore, and Adam W Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49(4):716–21, 2003.
- [58] R. F. Dougherty, M. Ben-Shachar, R. Bammer, A. A. Brewer, and B. A. Wandell. Functional organization of human occipital-callosal fiber tracts. *Proc Natl Acad Sci U S A*, 102(20):7350–5, 2005.
- [59] D. Dueck. *Affinity propagation: Clustering data by passing messages*. PhD thesis, University of Toronto, 2009.
- [60] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage*, 2010.
- [61] Stanley Durrleman, Pierre Fillard, Xavier Pennec, Alain Trouvé, and Nicholas Ayache. A statistical model of white matter fiber bundles based on currents. *Proceedings of the 21st Conference on Information Processing in Medical Imaging*, 21:114–25, January 2009.
- [62] P. Eades. A heuristic for graph drawing. *Congressus numerantium*, 42:149–160, 1984.

- [63] Damian Eads. hcluster: Hierarchical clustering for scipy. scipy-cluster.googlecode.com, 2008.
- [64] B. Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 7(1):1–26, 1979.
- [65] A. Einstein. *Investigations on the Theory of the Brownian Movement*. Dover Publications, 1956.
- [66] V El Kouby, Y Cointepas, C Poupon, D Rivière, N Golestani, J B Poline, D Le Bihan, and J F Mangin. MR diffusion-based inference of a fiber bundle model from a population of subjects. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 8(Pt 1):196–204, January 2005.
- [67] P. Fillard, M. Descoteaux, A. Goh, S. Gouttard, B. Jeurissen, J. Malcolm, A. Ramirez-Manzanares, M. Reisert, K. Sakaie, F. Tensaouti, et al. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion mr phantom. *NeuroImage*, 2011.
- [68] P. Fillard, C. Poupon, and J.F. Mangin. Spin Tracking: A Novel Global Tractography Algorithm. *NeuroImage*, 47:S127–S127, 2009.
- [69] Pierre Fillard and Cyril Poupon. A Novel Global Tractography Algorithm based on an Adaptive Spin Glass Model. 2009.
- [70] M.A. Fischler and R.C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [71] P. Fletcher, R. Tao, W.K. Jeong, and R. Whitaker. A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor mri. In *Information Processing in Medical Imaging*, pages 346–358. Springer, 2007.
- [72] R. Fletcher. Practical methods of optimization. *John and Sons, Chichester*, 1987.

- [73] B.J. Frey and D. Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972, 2007.
- [74] S Frey, JSW Campbell, and GB Pike. ...human language pathways with high angular resolution diffusion fiber tractography. *Journal of Neuroscience*, 2008.
- [75] O. Friman, G. Farneback, and C.F. Westin. A Bayesian approach for stochastic white matter tractography. *Medical Imaging, IEEE Transactions on*, 25(8):965–978, 2006.
- [76] O. Friman and C. F. Westin. Uncertainty in white matter fiber tractography. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 8(Pt 1):107–14, 2005.
- [77] E. Garyfallidis, M. Brett, B. Amirkhanian, C. Nguyen, FC Yeh, Y. Halchenko, and I. Nimmo-Smith. Dipy - a novel software library for diffusion mr and tractography. In *17th Annual Meeting of the Organization for Human Brain Mapping*, 2011.
- [78] Eleftherios Garyfallidis, Matthew Brett, and Ian Nimmo-Smith. Fast Dimensionality Reduction for Brain Tractography Clustering. *16th Annual Meeting of the Organization for Human Brain Mapping*, 2010.
- [79] G. Gerig, S. Gouttard, and I. Corouge. Analysis of brain white matter via fiber tract modeling. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 2, pages 4421–4424. IEEE, 2004.
- [80] X. Gigandet. *Global brain connectivity analysis by diffusion MR tractography: algorithms, validation and applications*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2009.
- [81] M.F. Glasser and J.K. Rilling. Dti tractography of the human brain's language pathways. *Cerebral Cortex*, 18(11):2471–2482, 2008.
- [82] A. Goh and R. Vidal. Algebraic methods for direct and feature based registration of diffusion tensor images. *Computer Vision–ECCV 2006*, pages 514–525, 2006.

- [83] T.F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.
- [84] P Guevara, C Poupon, D Rivière, Y Cointepas, M Descoteaux, B Thirion, and J-F Mangin. Robust clustering of massive tractography datasets. *NeuroImage*, 54(3):1975–1993, October 2010.
- [85] H. Gray. Gray’s Anatomy. *Lea and Febiger, 20th u.s. edition edition, 1918. Online edition: bartelby.com*, 2000.
- [86] N.S. Hageman, A.W. Toga, K.L. Narr, and D.W. Shattuck. A diffusion tensor imaging tractography algorithm based on navier–stokes fluid mechanics. *Medical Imaging, IEEE Transactions on*, 28(3):348–360, 2009.
- [87] P. Hagmann, J. P. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder, and R. Meuli. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *Neuroimage*, 19(3):545–54, 2003.
- [88] P. Hagmann, J.P. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder, and R. Meuli. DtI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. *Neuroimage*, 19(3):545–554, 2003.
- [89] Patric Hagmann, Lisa Jonasson, Philippe Maeder, Jean-Philippe Thiran, Van J. Wedeen, and Reto Meuli. Understanding diffusion mr imaging techniques: From scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics*, 26(1):205–223, 2006.
- [90] H.A. Haroon, D.M. Morris, K.V. Embleton, D.C. Alexander, and G. Parker. Using the model-based residual bootstrap to quantify uncertainty in fiber orientations from q -ball analysis. *Medical Imaging, IEEE Transactions on*, 28(4):535–550, 2009.
- [91] J. C. Haselgrove and J. R. Moore. Correction for distortion of echo-planar images used to calculate the apparent diffusion coefficient. *Magn Reson Med*, 36(6):960–4, 1996.

- [92] O. Heid. Eddy current-nulled diffusion weighting. In: *Proceedings of the 8th Annual Meeting of ISMRM, Denver*, page 799, 2000.
- [93] T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.
- [94] T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.
- [95] Kegang Hua, Jiangyang Zhang, Setsu Wakana, Hangyi Jiang, Xin Li, Daniel S. Reich, Peter A. Calabresi, James J. Pekar, Peter C.M. van Zijl, and Susumu Mori. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, 39(1):336–347, 2008.
- [96] Y. Iturria-Medina, EJ Canales-Rodríguez, L. Melie-García, PA Valdés-Hernández, E. Martínez-Montes, Y. Alemán-Gómez, and J M Sánchez-Bornot. Characterizing brain anatomical connections using diffusion weighted {MRI} and graph theory. *Neuroimage*, 36(3):645–660, 2007.
- [97] Kalvis M Jansons and Daniel C Alexander. Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems*, 19(5):1031–1046, 2003.
- [98] S Jbabdi, M W Woolrich, J L Andersson, and T E Behrens. A Bayesian framework for global tractography. *NeuroImage*, 37:116–129, 2007.
- [99] S. Jbabdi, MW Woolrich, JLR Andersson, and TEJ Behrens. A Bayesian framework for global tractography. *Neuroimage*, 37(1):116–129, 2007.
- [100] S. Jbabdi, MW Woolrich, JLR Andersson, and TEJ Behrens. A Bayesian framework for global tractography. *Neuroimage*, 37(1):116–129, 2007.
- [101] J.H. Jensen, J.A. Helpern, A. Ramani, H. Lu, and K. 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.

- [102] R. Jianu, C. Demiralp, and D.H. Laidlaw. Exploring 3D DTI fiber tracts with linked 2D representations. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1449–1456, 2009.
- [103] Heidi Johansen-Berg and Timothy E. J. Behrens, editors. *Diffusion MRI*. Academic Press, 2009.
- [104] L. Jonasson, P. Hagmann, JP Thiran, and VJ Wedeen. Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering. In *Int. Soc. Mag. Reson. Med*, page 1310, 2005.
- [105] D.K. Jones. Tractography gone wild: probabilistic fibre tracking using the wild bootstrap with diffusion tensor mri. *Medical Imaging, IEEE Transactions on*, 27(9):1268–1274, 2008.
- [106] DK Jones, MA Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *optimization*, 525, 1999.
- [107] R. A. Kanaan, S. S. Shergill, G. J. Barker, M. Catani, V. W. Ng, R. Howard, P. K. McGuire, and D. K. Jones. Tract-specific anisotropy measurements in diffusion tensor imaging. *Psychiatry Res*, 146(1):73–82, 2006.
- [108] N. Kang, J. Zhang, E.S. Carlson, and D. Gembiris. White matter fiber tractography via anisotropic diffusion simulation in the human brain. *Medical Imaging, IEEE Transactions on*, 24(9):1127–1137, 2005.
- [109] M.D. King, J. Houseman, S.A. Roussel, N. Van Bruggen, S.R. Williams, and D.G. Gadian. q-Space imaging of the brain. *Magnetic Resonance in Medicine*, 32(6):707–713, 1994.
- [110] B. W. Kreher, I. Mader, and V. G. Kiselev. Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. *Proc. Intl. Soc. Mag. Reson. Med.*, 16:425, 2008.
- [111] BW Kreher, I. Mader, and VG Kiselev. Gibbs tracking: a novel approach for the reconstruction of neuronal pathways. *Magnetic Resonance in Medicine*, 60(4):953–963, 2008.

- [112] Harold W Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.
- [113] Ludmila I. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.
- [114] L W Kuo, J H Chen, V J Wedeen, and W Y Tseng. Optimization of diffusion spectrum imaging and q-ball imaging on clinical MRI system. *Neuroimage*, vol:41pp7–18.
- [115] Eugene L. Lawler. *Combinatorial optimization: networks and matroids*. Dover Publications, Mineola, N.Y., 2001.
- [116] M. Lazar and AL Alexander. White matter tractography using random vector (RAVE) perturbation. In *Proceedings of ISMRM Annual Meeting, Honolulu*, 2002.
- [117] M. Lazar and A.L. Alexander. Bootstrap white matter tractography (BOOT-TRAC). *NeuroImage*, 24(2):524–532, 2005.
- [118] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field, and A. L. Alexander. White matter tractography using diffusion tensor deflection. *Hum Brain Mapp*, 18(4):306–21, 2003.
- [119] P.M. Lee. *Bayesian statistics*. Arnold London, UK:, 1997.
- [120] A. Leemans and DK Jones. A new approach to fully automated fiber tract clustering using affinity propagation. In *Int. Symp. on Magnetic Resonance in Medicine (ISMRM)*, volume 17, page 856, 2009.
- [121] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, and P. Parizel. Multiscale white matter fiber tract coregistration: A new feature-based approach to align diffusion tensor data. *Magnetic Resonance in Medicine*, 55(6):1414–1423, 2006.
- [122] C. Lenglet, J.S.W. Campbell, M. Descoteaux, G. Haro, P. Savadjiev, D. Wassermann, A. Anwander, R. Deriche, G.B. Pike, G. Sapiro, et al.

Mathematical methods for diffusion mri processing. *Neuroimage*, 45(1):S111–S122, 2009.

- [123] S. Lifshits, A. Tamir, and Y. Assaf. Combinatorial fiber-tracking of the human brain. *NeuroImage*, 48(3):532–540, 2009.
- [124] C. Liu, R. Bammer, B. Acar, and M.E. Moseley. Characterizing non-gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51(5):924–937, 2004.
- [125] H. Lu, J.H. Jensen, A. Ramani, and J.A. Helpern. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, 19(2):236–247, 2006.
- [126] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. California, USA, 1967.
- [127] M. Maddah. *Quantitative Analysis of Cerebral White Matter Anatomy from Diffusion MRI*. PhD thesis, Massachusetts Institute of Technology, 2008.
- [128] M. Maddah, W.E.L. Grimson, and S.K. Warfield. Statistical modeling and EM clustering of white matter fiber tracts. In *IEEE Int. Symp. Biomedical Imaging*, pages 53–56. Citeseer, 2006.
- [129] M. Maddah, A. U. Mewes, S. Haker, W. E. Grimson, and S. K. Warfield. Automated atlas-based clustering of white matter fiber tracts from DT-MRI. *International Conference in Medical Image Computing and Computer Assisted Intervention*, 8(Pt 1):188–95, 2005.
- [130] M. Maddah, W. M. Wells, 3rd, S. K. Warfield, C. F. Westin, and W. E. Grimson. Probabilistic clustering and quantitative analysis of white matter fiber tracts. *Inf Process Med Imaging*, 20:372–83, 2007.
- [131] M. Maddah, L. Zollei, W.E.L. Grimson, C.F. Westin, and W.M. Wells. A mathematical framework for incorporating anatomical knowledge

- in dt-mri analysis. In *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 105–108. IEEE, 2008.
- [132] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
 - [133] J.G. Malcolm, M.E. Shenton, and Y. Rathi. Filtered Tractography: State estimation in a constrained subspace. *Proceedings of MICCAI (Diffusion Modeling and Fiber Cup)*, pages 122–133, 2009.
 - [134] G. Margolis and JP Pickett. New applications of the Luxol fast blue myelin stain. *Laboratory investigation; a journal of technical methods and pathology*, 5(6):459.
 - [135] Matt A. Bernstein, Kevin F. King, and Xiaohong Joe Zhou. *Handbook of MRI Pulse Sequences*. Elsevier Academic Press, 2004.
 - [136] A. Mayer and H. Greenspan. Bundles of interest based registration of White Matter tractographies. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 919–922. IEEE, 2008.
 - [137] A. Mayer, G. Zimmerman-Moreno, R. Shadmi, A. Batikoff, and H. Greenspan. A Supervised Framework for the Registration and Segmentation of White Matter Fiber Tracts. *Medical Imaging, IEEE Transactions on*, 30(1):131–145, 2011.
 - [138] Arnaldo Mayer and Hayit Greenspan. Direct registration of white matter tractographies with application to atlas construction. In *MICCAI 2007 Workshop Statistical Registration PairWise and GroupWise Alignment and Atlas Formation*, 2007.
 - [139] D.W. McRobbie, E.A. Moore, and M.J. Graves. *MRI from Picture to Proton*. Cambridge University Press, 2006.
 - [140] K.D. Merboldt, W. Hanicke, H. Bruhn, M.L. Gyngell, and J. Frahm. Diffusion imaging of the human brain in vivo using high-speed STEAM MRI. *Magnetic Resonance in Medicine*, 23(1):179–192, 1992.

- [141] B. Mloberts, A. Vilanova, and J.J. van Wijk. Evaluation of fiber clustering methods for diffusion tensor imaging. In *IEEE Visualization*, pages 65–72, 2005.
- [142] D.C. Montgomery, E.A. Peck, G.G. Vining, and J. Vining. *Introduction to linear regression analysis*, volume 3. Wiley New York, 2001.
- [143] S. Mori, B. J. Crain, V. P. Chacko, and P. C. van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol*, 45(2):265–9, 1999.
- [144] S. Mori, S. Wakana, LM Nagae-Poetscher, and PCM Van Zijl. MRI atlas of human white matter. *American Journal of Neuroradiology*.
- [145] Susumu Mori and Jiangyang Zhang. Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*, 51:527–39, 2006.
- [146] M.E. Moseley, K. Butts, M.A. Yenari, M. Marks, and A.D. Crespigny. Clinical aspects of dwi. *NMR in Biomedicine*, 8(7):387–396, 1995.
- [147] R.T. Ng and J. Han. CLARANS: A method for clustering objects for spatial data mining. *IEEE Transactions on Knowledge and Data Engineering*, pages 1003–1016, 2002.
- [148] L. J. O'Donnell and C. F. Westin. Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Trans Med Imaging*, 26(11):1562–75, 2007.
- [149] L. J. O'Donnell, C. F. Westin, and A. J. Golby. Tract-based morphometry for white matter group analysis. *Neuroimage*, 45(3):832–44, 2009.
- [150] L.J. O'Donnell. *Cerebral White Matter Analysis Using Diffusion Imaging*. PhD thesis, 2006.
- [151] L.J. O'Donnell, C.F. Westin, and A.J. Golby. Tract-based morphometry for white matter group analysis. *NeuroImage*, 45(3):832–844, 2009.

- [152] R.L. O’Gorman and D.K. Jones. Just how much data need to be collected for reliable bootstrap dt-mri? *Magnetic resonance in medicine*, 56(4):884–890, 2006.
- [153] F.W. Olver, D.W. Lozier, R.F. Boisvert, and C.W. Clark. NIST handbook of mathematical functions. 2010.
- [154] H.H. Ong and F.W. Wehrli. Quantifying axon diameter and intracellular volume fraction in excised mouse spinal cord with q-space imaging. *Neuroimage*, 51(4):1360–1366, 2010.
- [155] E. Ozarslan and T.H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50(5):955–965, 2003.
- [156] E. Ozarslan, T.M. Shepherd, B.C. Vemuri, S.J. Blackband, and T.H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (dot). *NeuroImage*, 31(3):1086–1103, 2006.
- [157] S. Pajevic and P.J. Basser. Parametric and non-parametric statistical analysis of dt-mri data. *Journal of Magnetic Resonance*, 161(1):1–14, 2003.
- [158] G. Parker and D. Alexander. Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *Information Processing in Medical Imaging*, pages 684–695. Springer, 2003.
- [159] G. J. Parker, H. A. Haroon, and C. A. Wheeler-Kingshott. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. *J Magn Reson Imaging*, 18(2):242–54, 2003.
- [160] Geoffrey J M Parker, Hamied a Haroon, and Claudia a M Wheeler-Kingshott. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion

measurements. *Journal of Magnetic Resonance Imaging*, 18(2):242–54, 2003.

- [161] G.J.M. Parker and D.C. Alexander. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):893, 2005.
- [162] G.J.M. Parker, C.A.M. Wheeler-Kingshott, and G.J. Barker. Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *Medical Imaging, IEEE Transactions on*, 21(5):505–512, 2002.
- [163] O. Pasternak, Y. Assaf, N. Intrator, and N. Sochen. Variational multiple-tensor fitting of fiber-ambiguous diffusion-weighted magnetic resonance imaging voxels. *Magnetic resonance imaging*, 26(8):1133–1144, 2008.
- [164] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [165] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Riviere, A. Constantinesco, D. Le Bihan, and J. F. Mangin. Fiber tracking in q-ball fields using regularized particle trajectories. *Inf Process Med Imaging*, 19:52–63, 2005.
- [166] V. Pickalov and P.J. Basser. 3d tomographic reconstruction of the average propagator from mri data. In *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2006, pages 710–713, 2006.
- [167] P. Ramachandran and G. Varoquaux. Mayavi: 3d visualization of scientific data. *Computing in Science & Engineering*, 13(2):40–51, 2011.
- [168] T.G. Reese, O. Heid, R.M. Weisskoff, and V.J. 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.

- [169] M. Reisert, I. Mader, C. Anastasopoulos, M. Weigel, S. Schnell, and V. Kiselev. Global fiber reconstruction becomes practical. *Neuroimage*, 2010.
- [170] J.A. Rice. *Mathematical statistics and data analysis*. Thomson Learning, 2006.
- [171] J.K. Rilling, M.F. Glasser, T.M. Preuss, X. Ma, T. Zhao, X. Hu, and T.E.J. Behrens. The evolution of the arcuate fasciculus revealed with comparative dti. *Nature Neuroscience*, 11(4):426–428, 2008.
- [172] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. *International Conference in Medical Image Computing and Computer-Assisted Intervention*, page 1115, 1998.
- [173] K.E. Sakaie and M.J. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted mri. *NeuroImage*, 34(1):169–176, 2007.
- [174] K.K. Seunarine, PA Cook, MG Hall, KV Embleton, GJM Parker, and D.C. Alexander. Exploiting peak anisotropy for tracking through complex structures. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.
- [175] A. Sherbondy, M. Rowe, and D. Alexander. Microtrack: An algorithm for concurrent projectome and microstructure estimation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 183–190, 2010.
- [176] AJ Sherbondy, DL Akers, RF Dougherty, M. Ben-Shachar, S. Napel, and BA Wandell. MetroTrac: A metropolis algorithm for probabilistic tractography. *Human Brain Mapping, Florence*, 2006.
- [177] Anthony J. Sherbondy, Robert F. Dougherty, Michal Ben-Shachar, Sandy Napel, and Brian A. Wandell. ConTrack: Finding the most likely pathways between brain regions using diffusion tractography. *J. Vis.*, 8(9):1–16, 7 2008.

- [178] Stephen M. Smith, Mark Jenkinson, Heidi Johansen-Berg, Daniel Rueckert, Thomas E. Nichols, Clare E. Mackay, Kate E. Watkins, Olga Ciccarelli, Zaheer Cader, Paul M. Matthews, and Timothy E.J. Behrens. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31:1487–1505, 2006.
- [179] S.N. Sotiropoulos. Processing of diffusion mr images of the brain: from crossing fibres to distributed tractography. 2010.
- [180] S.N. Sotiropoulos, L. Bai, P.S. Morgan, C.S. Constantinescu, and C.R. Tench. Brain tractography using Q-ball imaging and graph theory: Improved connectivities through fibre crossings via a model-based approach. *NeuroImage*, 49(3):2444–2456, 2010.
- [181] H. Steinhaus. Sur la division des corp materiels en parties. *Bull. Acad. Polon. Sci*, 1:801–804, 1956.
- [182] E. O. Stejskal and J. E. 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.
- [183] H.C. Torrey. Bloch equations with diffusion terms. *Physical Review*, 104(3):563–565, 1956.
- [184] J. Tournier et al. Diffusion-weighted magnetic resonance imaging fibre tracking using a front evolution algorithm. *NeuroImage*, 20(1):276–288, 2003.
- [185] J. Tournier et al. Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.
- [186] 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):1459–72, 2007.

- [187] A. Tsai, C.F. Westin, A.O. Hero, and A.S. Willsky. Fiber tract clustering on manifolds with dual rooted-graphs. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–6. IEEE, 2007.
- [188] David S. Tuch, Timothy G. Reese, Mette R. Wiegell, Nikos Makris, John W. Belliveau, and Van J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582, 2002.
- [189] David S Tuch, Jonathan J Wisco, Mark H Khachaturian, Wim Vanduffel, Leeland B Ekstrom, and Rolf Ko. Q-ball imaging of macaque white matter architecture. (May):869–879, 2005.
- [190] D.S. Tuch. *Diffusion {MRI} of complex tissue structure*. PhD thesis, Massachusetts Institute of Technology, Division of Health Sciences and Technology, 2002.
- [191] DS Tuch. Q-ball imaging. *change*, 2004.
- [192] Eelke Visser, Emil H J Nijhuis, Jan K Buitelaar, and Marcel P Zwiers. Partition-based mass clustering of tractography streamlines. *NeuroImage*, July 2010.
- [193] Setsu Wakana, Arvind Caprihan, Martina M. Panzenboeck, James H. Fallon, Michele Perry, Randy L. Gollub, Kegang Hua, Jiangyang Zhang, Hangyi Jiang, Prachi Dubey, Ari Blitz, Peter van Zijl, and Susumu Mori. Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*, 36(1):630–644, 2007.
- [194] X. Wang, W.E.L. Grimson, and C.F. Westin. Tractography segmentation using a hierarchical Dirichlet processes mixture model. *NeuroImage*, 2010.
- [195] S. Warach, J. Gaa, B. Siewert, P. Wielopolski, and R.R. Edelman. Acute human stroke studied by whole brain echo planar diffusion-weighted magnetic resonance imaging. *Annals of neurology*, 37(2):231–241, 1995.

- [196] V Wedeen, R Wang, J Schmahmann, T Benner, W Tseng, G Dai, D Pandya, P Hagmann, H Dâ arceuil, and A. de Crespigny, "Diffusion spectrum magnetic resonance imaging (dsi) tractography of crossing fibers.". *NeuroImage*, vol:41no4pp1267–1277.
- [197] V. J. Wedeen, R. P. Wang, J. D. Schmahmann, T. Benner, W. Y. Tseng, G. Dai, D. N. Pandya, P. Hagmann, H. D'Arceuil, and A. J. de Crespigny. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage*, 41(4):1267–77, 2008.
- [198] V.J. Wedeen, P. Hagmann, W.Y.I. Tseng, T.G. Reese, and R.M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386, 2005.
- [199] VJ Wedeen, RP Wang, JD Schmahmann, and T Benner. ...spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage*, 2008.
- [200] B. Whitcher, D.S. Tuch, J.J. Wisco, A.G. Sorensen, and L. Wang. Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging. *Human brain mapping*, 29(3):346–362, 2008.
- [201] T.H. Williams, N. Gluhbegovic, and J.Y. Jew. *The human brain: dissections of the real brain*. University of Iowa, 1999.
- [202] R. Wirestam, A. Bibic, J. Latt, S. Brockstedt, and F. Stahlberg. Denoising of complex MRI data by wavelet-domain filtering: Application to high-b-value diffusion-weighted imaging. *Magnetic Resonance in Medicine*, 56(5), 2006.
- [203] S.F. Witelson. Hand and sex differences in the isthmus and genu of the human corpus callosum. *Brain*, 112(3):799, 1989.
- [204] Y.C. Wu and A.L. Alexander. Hybrid diffusion imaging. *NeuroImage*, 36(3):617–629, 2007.
- [205] F Yeh, V Wedeen, and W Tseng. Generalized Q-sampling imaging. *IEEE Transactions on Medical Imaging*, 29(9):1626–1635, March 2010.

- [206] F.C. Yeh, V.J. Wedeen, and W.Y.I. Tseng. Estimation of fiber orientation and spin density distribution by diffusion deconvolution. *Neuroimage*, 2011.
- [207] A. Zalesky. Dt-mri fiber tracking: a shortest paths approach. *Medical Imaging, IEEE Transactions on*, 27(10):1458–1471, 2008.
- [208] G. Zhai, W. Lin, K. P. Wilber, G. Gerig, and J. H. Gilmore. Comparisons of regional white matter diffusion in healthy neonates and adults performed with a 3.0-T head-only MR imaging unit. *Radiology*, 229(3):673–81, 2003.
- [209] F. Zhang, E.R. Hancock, C. Goodlett, and G. Gerig. Probabilistic white matter fiber tracking using particle filtering and von mises-fisher sampling. *Medical image analysis*, 13(1):5–18, 2009.
- [210] S. Zhang, S. Correia, and D.H. Laidlaw. Identifying White-Matter Fiber Bundles in DTI Data Using an Automated Proximity-Based Fiber Clustering Method. *IEEE transactions on visualization and computer graphics*, 14(5):1044, 2008.
- [211] S. Zhang et al. Visualizing diffusion tensor MR images using streamtubes and streamsurfaces. *IEEE Transactions on Visualization and Computer Graphics*, pages 454–462, 2003.
- [212] S. Zhang and DH Laidlaw. DTI fiber clustering and cross-subject cluster analysis. In *Int. Soc. Magn. Reson. Med*, 2005.
- [213] T. Zhang, R. Ramakrishnan, and M. Livny. BIRCH: A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery*, 1(2):141–182, 1997.
- [214] J. Zhuang, J. Hrabe, A. Kangarlu, D. Xu, R. Bansal, C. A. Branch, and B. S. Peterson. Correction of eddy-current distortions in diffusion tensor images using the known directions and strengths of diffusion gradients. *J Magn Reson Imaging*, 24(5):1188–93, 2006.

- [215] U. Ziyan, M. R. Sabuncu, L. J. O'Donnell, and C. F. Westin. Nonlinear registration of diffusion MR images based on fiber bundles. *International Conference in Medical Image Computing and Computer-Assisted Intervention*, 10(Pt 1):351–8, 2007.
- [216] U. Ziyan, M.R. Sabuncu, W.E.L. Grimson, and C.F. Westin. Consistency clustering: a robust algorithm for group-wise registration, segmentation and automatic atlas construction in diffusion MRI. *International Journal of Computer Vision*, 85(3):279–290, 2009.
- [217] O. Zvitia, A. Mayer, and H. Greenspan. Adaptive mean-shift registration of white matter tractographies. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 692–695. IEEE, 2008.
- [218] Orly Zvitia, Arnaldo Mayer, Ran Shadmi, Shmuel Miron, and Hayit K Greenspan. Co-registration of white matter tractographies by adaptive-mean-shift and Gaussian mixture modeling. *IEEE transactions on medical imaging*, 29(1):132–45, January 2010.