Reconstruction and Analysis Methods in Diffusion and Anatomical MRI

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Iman Aganj

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Prof. Guillermo Sapiro

December 2010

Acknowledgements

First, I would like to thank my Master's and PhD supervisor, Prof. Guillermo Sapiro, with whom my Minnesotan years are defined. He has been incredibly patient with me, while always giving me constructive feedback to keep me on the right track. It is my privilege to have him as a friend, and I always feel comfortable turning to him for advice about virtually everything in life. Guillermo has been my inspiration to pursue a life in academia, and he is truly the person whom I am looking forward to becoming like in my future career.

I would also like to acknowledge and thank professors Daniel Boley, Noam Harel, Mostafa Kaveh, and Jim Leger for accepting to be in my PhD committee, to review my thesis, and to provide me with their invaluable feedback.

I am thankful to Prof. Christophe Lenglet, who mentored me on everything I needed in order to advance in most of the projects presented in this thesis. He saved me an immense amount of time by providing me with every implementation detail, evaluating the results with his knowledge of the anatomy, introducing available resources, toolboxes, and datasets, and tutoring me on how to use them. In brief, Christophe has been my live Diffusion MRI encyclopedia, whom I bothered every time I had a question!

I would like to thank our collaborators at the Center for Magnetic Resonance Research, Professors Noam Harel, Essa Yacoub, and Kamil Ugurbil, who provided high-quality datasets for us to test our methods on, and especially Noam who was a great advisor during my summer internship at the Center. I thank Prof. Paul M. Thompson of the University of California – Los Angeles, Neda Jahanshad, Sarah Madsen, and Neelroop Parikshak, for our superb collaboration, which made it possible for us to try our methods on their many datasets. I am grateful to Prof. Rachid Deriche of Institut National de Recherche en Informatique et Automatique, and Emmanuel Caruyer for hosting me in their wonderful Odyssée project team and giving me the opportunity to learn a lot from them during our collaboration. I would also like to thank Prof. Sriram Subramaniam of the National Institutes of Health, Hstau Liao, Rajesh Narasimha, and Alberto Bartesaghi,

with whom I collaborated on my Master's electron microscopy projects.[1] Lastly, I would like to acknowledge Jennifer Campbell of McGill University, Cyril Poupon of Neurospin, Katie McMahon and Greig de Zubicaray of the University of Queensland, and Margaret Wright of Queensland Institute of Medical Research for allowing us to use their valuable datasets.

I would like to thank all my labmates who never hesitated to help me with my work and my life: Mona Mahmoudi, Xue Bai, Diego Rother, Kedar Patwardhan, Hstau Liao, Pablo Sprechmann, Alexey Castrodad, Leah Bar, Ignacio Ramirez, Federico Lecumberry, Adarsh Chandran, Pablo Arias, Anish Mohan, Julien Mairal, Guoshen Yu, Julio Duarte, and Oleg Kuybeda. The countless happy moments we shared together make me never forget this period of my life.

Finally, I thank my parents, Mina and Jalil, and my brother, Ehsan, who always supported me in any way they could, and I am most indebted to. Thank you, Skype, for letting me feel closer to them and helping me see their love and sacrifice, without which, I would not have been able to achieve any of my goals.

The projects in this thesis were partly supported by NIH, NSF, NCRR, NIBIB, ONR, NGA, ARO, DARPA, NHMRC (Australia), the Keck Foundation, the MIND Institute, and the University of Minnesota Institute for Translational Neuroscience. Computing resources were provided by the University of Minnesota Supercomputing Institute, and the Laboratory of Neuro Imaging (UCLA).

Constrained Spherical Deconvolution, Mixture of Wisharts, and the dip test were performed using respectively the MRtrix package by J. D. Tournier, the program "Multi-fiber Reconstruction from DW-MRI" by B. Jian, and a public code by F. Mechler.

[1] I received my Master's degree through a coursework-only track without a thesis, so I never had the chance to acknowledge these wonderful people!

To Mina, Jalil, and Ehsan…

Abstract

Magnetic Resonance Imaging (MRI), along with its extension, Diffusion-Weighted MRI, is a noninvasive imaging modality which has become popular for clinical and research purposes. The raw data obtained from the MRI scanner may not be immediately usable by the health professional, thereby creating the need for additional methods to make more sensible representations of the data and extract the desired information from them. This thesis introduces a few new reconstruction, post-processing, and analysis techniques in diffusion and anatomical MRI, as described next.

Hardware, timing, and SNR considerations restrict the slice-selection and the in-plane resolutions of MRI differently, generally resulting in anisotropic voxels. This non-uniform sampling can be problematic, especially in image segmentation and clinical examination. To alleviate this, the acquisition is divided into (two or) three separate scans, with thick slices yet orthogonal slice-selection directions. In the first part of the thesis, a non-iterative wavelet-based approach to combining the three orthogonal scans is adopted, and its advantages compared to other existing methods, such as Fourier techniques, are mentioned, including the consideration of the actual pulse response of the MRI scanner and lower computational complexity.

Estimating the thickness of the cerebral cortex is a key step in many brain imaging studies, revealing valuable information on development or disease progression. In the second part, a framework for measuring the cortical thickness, based on minimizing line integrals over the probability map of the gray matter in the MRI volume is presented. In contrast to the proposed approach, previous methods often perform a binary-valued hard segmentation of the gray matter before measuring the cortical thickness. Due to image noise and partial volume effects, such a hard classification ignores the underlying tissue class probabilities assigned to each voxel, discarding potentially useful information. The performance of the method is demonstrated on both artificial volumes and real 3D brain MRI data from subjects with Alzheimer's disease and healthy individuals.

Q-ball imaging (QBI) is a Diffusion MRI reconstruction technique which has been proven very successful in resolving multiple intravoxel fiber orientations in MR images. The standard computation of the orientation distribution function (ODF, the probability of diffusion in a given direction) from q-ball data uses linear radial projection, neglecting the change in the volume element along each direction, resulting in artificial blurring of the ODFs. In the third part of this dissertation, a new technique is proposed that, by considering the solid angle factor, uses the

mathematically correct definition of the ODF and results in a dimensionless and normalized ODF expression. In addition, a semi-analytic ODF maxima extraction algorithm is provided, and a measure for fiber crossing is also introduced which is shown to be related to the intelligence quotient. The performances of the proposed techniques are demonstrated on artificial examples and high-resolution real data acquired at 7 Tesla.

In the fourth part, a global probabilistic fiber tracking approach based on the voting process provided by the Hough transform is introduced. The proposed framework tests candidate 3D curves in the volume, assigning to each one a score computed from the diffusion images, and then selects the curves with the highest scores as the potential anatomical connections. The algorithm avoids local minima by performing an exhaustive search at the desired resolution, and is easily extended to multiple subjects. Experimental results are presented on HARDI volumes, ranging from simulated and 1.5T physical phantoms to 7T and 4T human brain and 7T monkey brain datasets.

Table of Contents

List of Figures

1 Thesis outline

Magnetic Resonance Imaging (MRI) is a noninvasive imaging modality which has become popular for clinical and research purposes. An MRI dataset, being a volumetric image, contains massive amount of information about the scanned object – typically the neuronal tissue. In addition, Diffusion-Weighted MRI (DWI) extends this information to the fiber architecture of the tissue by measuring the diffusion of water in three-dimensional space.

The raw data obtained from the MRI scanner may not be immediately usable by, e.g., the health professional. This creates the need for different reconstruction methods to make more sensible representations of the data, and various post-processing algorithms to extract the desired information from the immense amount of available data.

The contribution of this thesis is several new reconstruction and analysis methods in diffusion and anatomical MRI. In this section, I provide a brief overview of each chapter. Since the problems addressed in this thesis are not necessarily of the same nature, I refer the interested reader to the "Introduction" section of each chapter for the basic background about the corresponding problem.

In Chapter 2, I give an overview about the problem of reconstruction of a high-resolution image with isotropic voxels from multiple lower-resolution images with anisotropic-sized voxels, with the low resolutions being in perpendicular directions. I describe a new wavelet-based fusion method and compare it with the existing ones.

Measuring the cortical thickness has long been a topic of interest for neuroscientists. In Chapter 3, I introduce a new definition of tissue thickness which, contrary to the previously existing methods, takes the *soft*-classified image of the tissue, and computes the thickness using line integrals. A longitudinal study on cortical thinning of the Alzheimer's disease patients vs. control subjects is also presented.

Chapter 4 is the starting chapter of the second part of the thesis, dedicated entirely to DWI. It begins with a review of a popular technique for the reconstruction of the orientation distribution function (ODF; the probability of diffusion of water in a given direction), called the q-ball imaging. Then, it mentions and addresses a mathematical inaccuracy in this method. Experimental results are shown and used to compare the proposed method to other existing ones, and the improvement achieved by this correction is demonstrated.

In Chapter 5, I describe a measure of fiber crossing derived by comparing the two ODFs computed using single- and multiple-fiber models. The experimental results in that chapter show

a significant correlation between the intelligence quotient of a person and the value of the proposed measure in specific areas of the brain.

When the principal directions of diffusion are desired, the ODF needs to be further processed. Chapter 6 presents a new semi-analytical approach to the ODF maxima extraction in spherical harmonic (SH) basis. This technique exploits the fact that in the fourth order real and symmetric SH basis, the extrema of the ODF lie on one-dimensional curves, and therefore the two-dimensional surface of the ODF does not need to be searched.

DWI provides, through fiber tracking, a unique *in-vivo* quantitative measurement of the brain's anatomical connectivity. In Chapter 7, I introduce a global and probabilistic Hough-transform based approach to tractography. Being in principal an exhaustive search, this technique avoids entrapment in the local optima. I compare this method to other existing algorithms, both theoretically and by using experimental results on various simulated and real datasets.

Finally, concluding remarks are mentioned in Chapter 8, and further derivations and implementation details are presented in the appendices (Chapter 9).

2 Wavelet Fusion of Anisotropic MR Images

2.1 Introduction

Three-dimensional (3D) MR volumes are typically tomographic sets of two-dimensional (2D) image slices, selected using radio frequency pulses. Since a different strategy is used for the in-plane encoding (i.e. frequency and phase encoding), it is not surprising that the hardware and timing limitations on the slice-selection and the in-plane resolutions are not the same. The trade-off imposed by these limitations, being principally more restrict on the slice-selection resolution, generally results in thicker acquisition slices and therefore anisotropic voxels. This non-uniform sampling can be problematic, especially in image segmentation and clinical examination, since the image will be missing high frequencies in the slice-selection direction. Furthermore, choosing thinner slices while keeping the same in-plane resolution reduces the signal-to-noise ratio (SNR), given that the voxel size is proportional to the SNR.

In the cases where there are no time or SNR constraints, voxels may be chosen to be isotropic by decreasing the slice thickness. However, acquiring such a high-resolution MR volume may require the subject to be motionless for a (clinically) unreasonably long time, otherwise increasing the risk of motion artifacts. This can be alleviated by dividing the acquisition into (two or) three separate scans in different directions, with thicker slices yet complementary resolutions, each containing a considerable proportion of the high frequencies, in two out of the three directions, missing in the other ones. Every scan will then have a shorter acquisition time and a lower chance of undergoing motion-related distortion. Misalignment between the scans can be corrected by employing a variety of available registration techniques, and the high-resolution image should eventually be reconstructed from these multiple scans (three orthogonal scans are considered here).

Different approaches to combining the three orthogonal scans have been proposed in the literature. Simply averaging the volumes, as done in (1)– (2), introduces artificial blurring and a decrease in the contrast (see Sec. 2.2.2 for more details). To avoid this, the authors of (3)– (6) have suggested a selective combination scheme in the Fourier domain, which is principally averaging the information where it comes from multiple scans, and zero padding where no information is available. This method, however, would only be mathematically accurate when assuming that the point spread function (PSF) of the MR scanner in the slice-selection direction is a *sinc* function, which is different from the actual rectangular-window (RW) function often used

in the literature (e.g., in (7)– (10)). A super-resolution based method has been proposed in (9), where the RW (or potentially any other) kernel can be considered as the system response of the MR scanner, and the orthogonal scans are combined by iteratively minimizing a cost function. In addition, the authors of (11) suggested to reconstruct the high-resolution image iteratively by assuming it to be a linear combination of low-resolution images. Finally, a level-set segmentation algorithm for such data has been introduced in (12).

In this work, we adopt a non-iterative and fast wavelet-based approach, which takes into account the actual PSF of the MR scanner. Wavelet fusion techniques, (13)– (28), are commonly used to combine multiple images into a single one, retaining important features from each, and providing a more accurate description of the object. In our case, ideally, we would like to collect all the meaningful information from the input images (i.e. the low frequencies and in-plane high frequencies), while discarding the parts bearing no information (i.e. high frequencies in the slice-selection direction). The wavelet transform, (29)– (30), simplifies this procedure by splitting the image into blocks of low and high frequencies in different directions, enabling us to easily pick the desired (available) blocks from each image. We will show how, assuming the RW-PSF, the proposed wavelet-based approach achieves to divide the image into *all-information* and *no-information* blocks, whereas the Fourier transform fails to accurately carry out such a separation, resulting in loss of information while combining the images. In addition, wavelet transform has a lower computational complexity than the Fourier transform, and is non-iterative as opposed to (9) and (11).

In Sec. 2.2, we justify the use of wavelets and provide technical details. Experimental results are presented in Sec. 2.3, and concluding remarks are mentioned in Sec. 8.1. This work has been submitted to (31)– (32).

2.2 Methods

2.2.1 Wavelet vs. Fourier

Each input image, having a low acquisition resolution in a unique Cartesian direction, lacks some high-frequency information compared to the desired high-resolution image (the image we would obtain by sampling at high resolution in all three directions), which we wish to reconstruct. The main challenge in image fusion lies in how to *selectively* extract information from the three

input images, to avoid relying on seemingly genuine yet information-free parts of the data. To achieve this, we need a tool that transforms every input image into two distinguishable parts: an *all-information* (AI) block to be used and a *no-information* (NI) block to be discarded. Ideally, the AI block of each input image will be identical to the same block of the desired high-resolution image, and the NI block will bear *no* useful information regarding the same block of the ground-truth image. Provided that this transform is invertible, the output image can be reconstructed in the transform domain by combining only the AI blocks of the input images, and then performing the inverse transform. We now show that by considering the standard RW-PSF in the slice-selection direction of the MR scanner (typically used in the literature (7)– (10)), the one-dimensional (1D) wavelet transform with the Haar basis[2] theoretically achieves such an ideal partition, whereas the performance of the 1D Fourier transform is suboptimal. This is of course in addition to the fact that the computational complexity of the Fast Wavelet Transform, $\mathcal{O}(N)$, is also lower than that of the Fast Fourier Transform, $\mathcal{O}(N \log N)$.

To demonstrate the proposed approach, we first simulate a row of 1×1×2-mm³ voxels. We consider a discrete 1D signal (representing isotropic 1-mm³ voxels) which has been blurred by the RW kernel of length 2 and then downsampled by a factor of 2. Generalization of the following to 3D is straightforward.

The Fourier transform decomposes a discrete signal x_n into its frequency elements, representing them as the function $X(\Omega), \Omega \in [-\pi, \pi]$. According to Nyquist-Shannon theorem, sampling the signal with the rate of 1 out of 2 voxels, results in aliasing unless the signal is band-limited, $X(\Omega) = 0, |\Omega| > \pi/2$, which is not guaranteed for the MR images due to the existing sub-voxel structures and edges. Only if the PSF of the MR scanner were an ideal low-pass filter for the Fourier transform (i.e. the *sinc* function), those high frequencies would be eliminated and no aliasing would occur. Then, in this unrealistic scenario, the low-frequency region $|\Omega| < \pi/2$ of the (upsampled) signal would represent exactly the same region of the desired signal (AI block), and its high-frequency region $|\Omega| > \pi/2$ would carry no information regarding the same region of the signal (NI block). However, the RW-PSF of the scanner in the slice-selection direction, although a low-pass filter, is far from an ideal one. Thus, aliasing does occur, distorting the low-frequency region of $|\Omega| < \pi/2$ and making it inappropriate for being an AI block. Likewise, the high-frequency region of $|\Omega| > \pi/2$ contains *some* information about that region of

[2] Extending this approach to the wavelet bases other than Haar to account for other PSF assumptions, such as the Gaussian, is a subject of future research.

the original signal, since the RW-PSF of the MR scanner does not remove those frequencies. Therefore, this region cannot be considered as an NI block either. We hence conclude that the Fourier transform cannot separate an input image into AI and NI blocks. We will now see how this is solved by the wavelets transform.

In its simplest form (first decomposition level), the Haar wavelets are used to transform the sequence (signal) $x_1, x_2, \dots, x_N$ into two separate sequences (L, H), a low-pass $L = l_1, l_2, \dots, l_{N/2}$ one and a high-pass one $H = h_1, h_2, \dots, h_{N/2}$ (for an even signal length N), with $l_i = (x_{2i-1} + x_{2i})/\sqrt{2}$ and $h_i = (x_{2i-1} - x_{2i})/\sqrt{2}$. Since the signal x_i' acquired on the MR scanner is obtained after blurring the ground-truth signal x_i with the two-tap RW-PSF, $\{1/2\,, 1/2\}$, and then downsampling it with the factor of 2, one can easily verify that $x_i' = l_i/\sqrt{2}$. Therefore, an upsampled version of this signal (e.g. by performing an inverse wavelet transform on the zero-padded $(L', H') = \left(\sqrt{2}x_i', 0\right)$) will have an L' exactly equal to L corresponding to the desired high-resolution signal (AI block), and an information-free $H' = 0$ (NI block). Thus, the Haar wavelet transform splits the blurred signal into a pair of AI (L) and NI (H) blocks, as desired.[3]

2.2.2 Wavelet Volumetric Image Fusion

Extended to higher dimensions, the (above described) simplest wavelet transform of a d-dimensional image consists of 2^d different blocks, each containing either the low- or high-frequency information corresponding to a unique Cartesian dimension of the image. The 3D wavelet transform of an MR image (with "XYZ" coordinates) has eight blocks, often denoted as "LLL," "LLH," "LHL," …, "HHH" (see Fig. 2.1), where, for instance, LHH stands for the block containing low-, high-, and high-frequency information in the X, Y, and Z directions, respectively. Now suppose that the three orthogonal scans have voxel sizes proportional to 2×1×1, 1×2×1, and 1×1×2.[4] Then, as a generalization of previous section, we can assume that the wavelet transform of, e.g., the first scan, has useful information in its four L## (AI) blocks, but

[3] In a sense, the RW function may be considered as the ideal low-pass filter in the Haar wavelet domain, as downsampling after applying such a filter leaves the low-frequency region (L) untouched.

[4] We continue assuming the ratio of the slice-selection/in-plane voxel dimension to be 2. For higher powers of 2 (e.g. 4, 8, …), the wavelet transform needs to be performed in further decomposition levels. In case this ratio is not a power of 2, a 1D interpolation in the slice-selection direction will be necessary to decrease the voxel size in that direction to the closest power of 2. Such an acquisition might then introduce extra blurring.

virtually no information in its H## (NI) blocks ("#" stands for both H and L). Similarly, the informative (AI) blocks of the second and third volumes would be the four #L# and the four ##L blocks, respectively.

It now becomes apparent why simply averaging the three scans produces artificial blurring and thus is an inaccurate reconstruction of the object. Given the linearity of the wavelet transform, averaging the scans would be block-wise averaging the three images in the wavelet domain. This would mean including both the signal-containing AI blocks and the zero NI blocks while computing the mean (destructive superposition), resulting in an artificial attenuation of the high-frequency blocks (e.g. by the factors of ⅓ and ⅔). Hence, simply averaging the input images would function as a low-pass filter.

Therefore, we proceed with the reconstruction of the desired high-resolution image – in the wavelet domain – by using as many AI blocks as available from the datasets, as sketched in Fig. 2.2. For example, the reconstructed LLL block will be the average of the LLL blocks of all the three scans, the LLH will be the average of the LLH blocks of the first two scans, and the LHH will be the LHH of the first scan. The HHH block is the only one missing in all the three scans, which we handle by zero-padding, similarly to how the Fourier techniques do (other filling-in can be used if prior information is available). Performing an inverse wavelet transform will be the final step in making the high-resolution image with isotropic voxels, and exploiting the maximal possible information from the three available scans.

Since the wavelet transform and its inverse are performed using the Fast Wavelet Transform algorithm with the complexity of $\mathcal{O}(N)$ (N being the number of voxels) (30), the entire process is done fast and directly (at once), as opposed to the iterative methods proposed in (9) and (11), and also faster than the direct Fast Fourier Transform-based techniques which have the complexity of $\mathcal{O}(N \log N)$.

2.3 Results and Discussions

Three orthogonal Susceptibility-Weighted Imaging (SWI) datasets were acquired on a 7T MRI (Siemens Avanto) with voxel sizes of 0.4×0.4×0.8 (Fig. 2.3a), 0.8×0.4×0.4 (Fig. 2.3b), and 0.4×0.8×0.4 mm³ (Fig. 2.3c). In addition, a high-resolution ground truth dataset with isotropic voxels of the size 0.4×0.4×0.4 mm³ (Fig. 2.3d) was acquired. The three orthogonal datasets were combined using both simple averaging (Fig. 2.3e) and the proposed wavelet fusion method

(Fig. 2.3f). Haar wavelet functions were used to account for the rectangular-window PSF of the MR scanner in the slice-selection direction. Simply averaging the scans (Fig. 2.3e) results in artificial blurring, due to the fact that the same weight has been given to high frequencies of all images. On the other hand, the image reconstructed by the proposed method (Fig. 2.3f) is sharper and visually closer to the ground truth (Fig. 2.3d). To better appreciate the improvement of the proposed approach, we have generated isosurface rendering of vessels within the imaged volume (Fig. 2.4), from isotropic (yellow) and anisotropic (red) voxels. Note the clear deformations of the vessel reconstruction using anisotropic voxels, making isotropic voxels critical for accurate and non-deformed segmentation of brain structures.

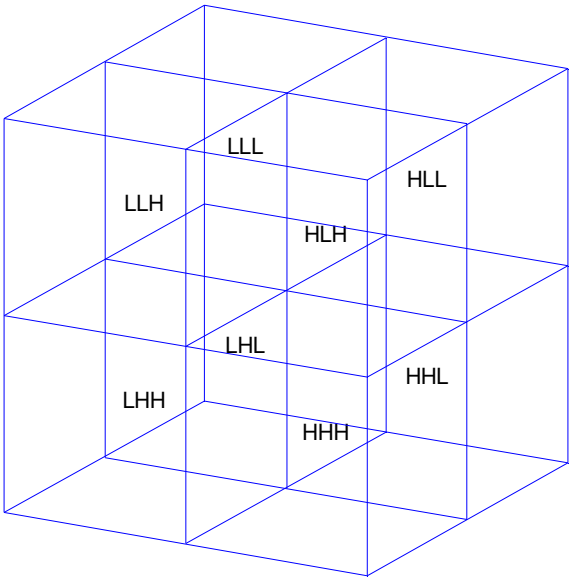


Fig. 2.1. The eight blocks of the 3D wavelet transform in the first decomposition level.

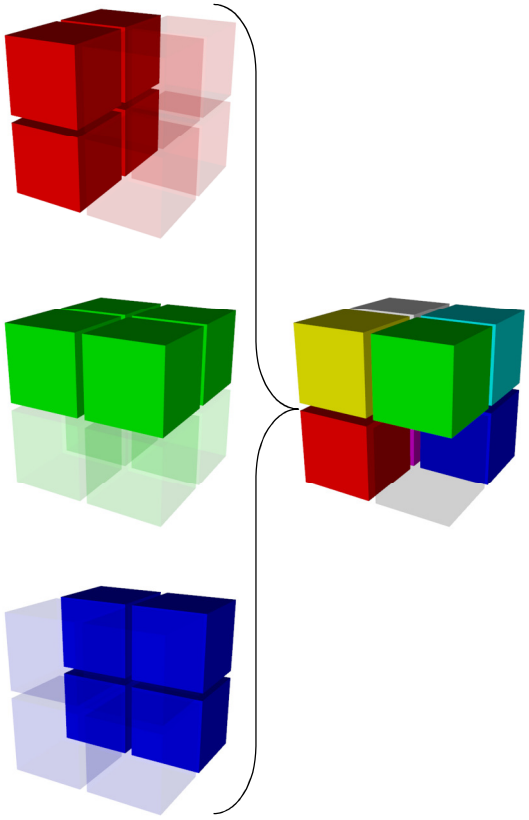


Fig. 2.2. The AI blocks are selected from the wavelet domain of the three anisotropic-voxel scans (left) and combined to make the high-resolution isotropic-voxel image (right).

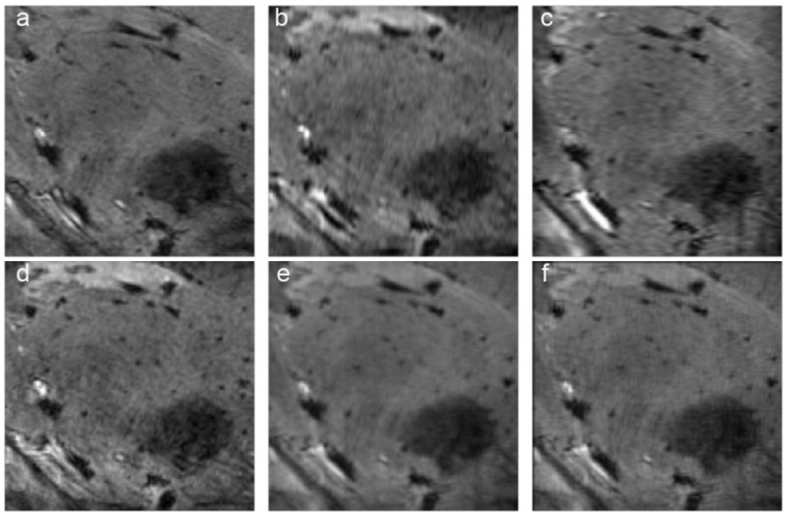


Fig. 2.3. The three anisotropic-voxel scans (a,b,c), the high-resolution ground-truth isotropic-voxel scan (d), and the combination of the three orthogonal scans by simple averaging (e), and the proposed wavelet approach (f).

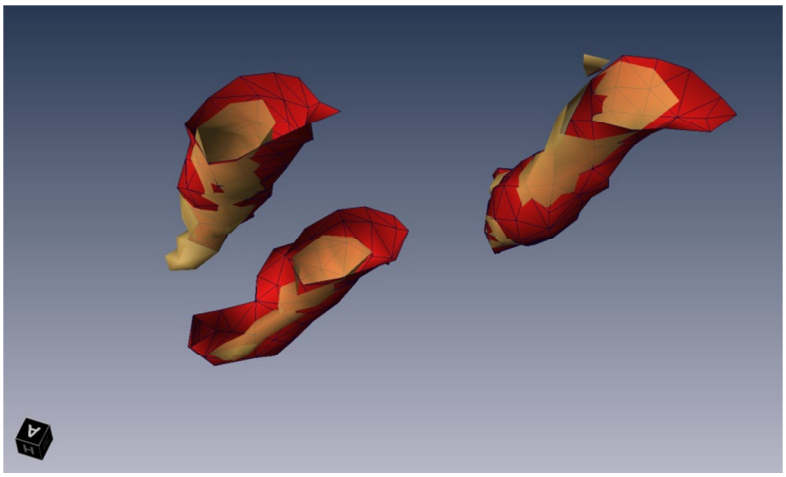


Fig. 2.4. Isosurface rendering of vessels within the imaged volume, from isotropic (yellow) and anisotropic (red) voxels.

3 Measuring Cortical Thickness by Line Integral

3.1 Introduction

Measuring the cortical thickness has long been a topic of interest for neuroscientists. Cortical thickness changes in a characteristic pattern during childhood development and with the progression of neurodegenerative diseases such as Alzheimer's, HIV/AIDS, and epilepsy (33)– (34). Recent studies examining changes in cortical thickness over time have revealed the trajectory of diseases in the living brain, and have been used to quantify treatment effects, identifying regions where cortical thickness correlates with age, cognitive deterioration, genotype, or medication.

Various approaches have recently been proposed to automate this cortical thickness measurement from Magnetic Resonance Imaging (MRI) data, e.g., (35)– (42). The limited spatial resolution of most MRI volumes (typically 1-2 mm) makes it difficult to measure cortical thickness accurately, which varies from 2 to 5 mm in different brain regions and is only a few voxels thick in the images. The neuroscience community has not yet agreed on a unique definition of cortical thickness and so far the various methods proposed measure slightly different quantities. What is common among them is that they virtually all perform a pre-segmentation of the white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), and most extract explicit models of the surfaces between them (i.e., the inner surface between WM and GM and outer surface between GM and CSF). They then use this hard segmentation as the input data for different tissue thickness measurement algorithms (Sec. 3.2 briefly reviews previous work and comments more on this). The disadvantage of this approach is that in the hard segmentation process, information is discarded and decisions are made before measuring the tissue thickness, a significant local error in measured thickness could be introduced by a few misclassified voxels (see Sec. 3.4.1 for an example).

The approach we adopt here uses a soft pre-labeled/classified volume as the input data, keeping valuable information all the way into the step of measuring tissue thickness. Due to the limited resolution of an MRI volume, many voxels contain partial amounts of two or more tissue types (see (43) and the references therein). Their intensity values give us information about the probability or proportion of those voxels belonging to any of the categories of WM, GM, or CSF. Rather than a (hard) pre-classified volume, we use one containing the probability that each voxel

belongs to the GM.[5] These probability values have the same precision as the values in the original MRI volume, and therefore we do not discard any useful information.[6] We compute line integrals of the soft classified data, centered at each voxel and in all possible spatial directions, and then consider their minimum as the local cortical thickness at that voxel.

While hard segmentations are often used as part of the analysis, e.g., to warp surfaces for population studies and/or for visualization, many useful statistics can be performed completely avoiding this hard classification, e.g., region-based statistics. Moreover, volumetric warping avoids hard segmentation. Even if hard segmentation is to be performed for other parts of the analysis, the errors produced by it need not be transferred to the tissue thickness computation. This error transfer is common in the techniques mentioned below and avoided with our proposed framework.

In Sec. 3.2, we review previous work on cortical thickness measurement. Section 3.3 describes our proposed framework, and experimental results are presented in Sec. 3.4. Section 8.2 concludes with a review of the contributions, and finally the implementation is covered in detail in Appendix 9.1. This work was published in (44)– (45), and later extended for longitudinal studies in (46).

3.2 Previous Related Work

We now discuss some of the previously reported work for measuring the cortical thickness. While many additional very important works have been published, those mentioned below provide a good representation of the spectrum of techniques available in the literature. Most methods require a (hard) pre-segmentation of the inner and outer surface, which results in a loss of available information and often inaccuracy in the input for the main thickness measurement algorithm. This loss, while manifested at different levels depending on the sophistication of the algorithm, is intrinsic to all hard segmentation methods.

Coupled-surface methods (35,47), define the cortical thickness as the Euclidean distance between corresponding point pairs on the inner and outer surfaces, often with parametric grids imposed. A displaced surface may result in an overestimation of the thickness (see Fig. 3.1(a)). *Closest point methods* such as (48) compute for each point on one of the two surfaces the closest

[5] When considering partial volume effects, these "probabilities" represent the proportion of GM in the voxel.

[6] The only lost information is that we are not able to distinguish between WM and CSF, which as we will see is not a concern, since we are only interested in the two categories of GM and non-GM. However, we could also preserve that information if needed.

point on the other surface and define the thickness as the Euclidean distance between them. The main drawback with these methods is the absence of symmetry, as seen in Fig. 3.1(b). In another method introduced in (49), the *regional histogram of thickness* is estimated by measuring the length of the line segments connecting the inner and outer surfaces of the GM layer, normal to one of the surfaces. The median of the histogram is then chosen as the local cortical thickness. A detection of the WM-GM and GM-CSF boundaries is however necessary.

Laplace methods (36,39,50) solve Laplace's equation in the GM region with the boundary condition of constant (but different) potentials on each of the two surfaces. The cortical thickness is then defined on each point as the length of the integral curve of the gradient field passing through that point, as illustrated in Fig. 3.1(c). With this approach, the thickness is uniquely defined at every point. Nevertheless, a pre-segmentation of the two surfaces is required, reducing the accuracy of this technique.

Another category of methods defines thickness by making use of a central axis or skeleton (41,51). Thickness is typically estimated as the diameter of the largest enclosed sphere in the GM layer, which is (in some cases only initially) centered on a point on the central axis. As Fig. 3.1(d) demonstrates, a relatively sharp change in the thickness may result in a new branch and affect the topology of the skeleton.

The vast majority of the methods reported in the literature propagate segmentation errors to later steps, and segmentation is still in itself a challenging problem in brain imaging. Considering that the GM layer spans only a few voxels at the commonly used 1-2 mm resolutions, these errors can be significant, and measuring tissue thickness avoiding this hard segmentation step may be very beneficial. This is the approach introduced here and described next.

3.3 Methods

3.3.1 Definition

In its simplest form, we define the thickness of the GM at a given voxel as the minimum line integral of the probability map of the GM over all lines passing through that voxel.[7] Formally:

$$T(\vec{x}) := \min_{l \in L_{\vec{x}}} \int_l P(\vec{r})dl \,, \qquad [1]$$

[7] While here we use lines, the use of other curves of integration is an interesting subject of future research.

where $T(\vec{x})$ is the thickness of the GM at a point $\vec{x} \in \mathbb{R}^3$, $P(\vec{x})$ is the probability of the point $\vec{x}$ belonging to the GM (estimation of this probability is described in Appendix 9.1.1), and $L_{\vec{x}}$ is the set of all lines in three-dimensional space passing through the point $\vec{x}$. In practice, however, $L_{\vec{x}}$ is comprised of all equal-length *line segments* centered at $\vec{x}$, which are sufficiently longer than the expected maximum thickness in the volume. Choosing longer line segments does not greatly affect the integral values since $P(\vec{x})$ decreases significantly on the non-GM regions.[8] Fig. 3.2(a) shows an example of this construction, for a 2D binary probability map, where the probability of belonging to the GM is 1 inside the shape and 0 outside. When computing thickness at the specified point, the line segment marked with oval arrows is selected as the one giving the smallest line integral. The corresponding integral value, which in this case is the length of its overlap with the GM (in bold), is the thickness of the GM at that point. A more realistic situation is shown in Fig. 3.2(b), where the probability map varies between zero and one. A blurred border, which results from the limited resolution of the MRI, includes voxels that partially contain GM. Due to the pre-segmentation, this type of partial volume information is not considered in most prior work in this area.[9]

Our method is based on an intuitive way of measuring the thickness of an object. A simple way to measure the local thickness of an object would be to put two fingers on both edges of the object, and move the finger tips locally (equivalent to varying the angle of the segment connecting them to each other), until the Euclidean distance between them is minimized (Fig. 3.3). For example, among Figs. 3.3(a), 3.3(b), and 3.3(c), we would naturally choose Fig. 3.3(b) as the one that depicts the most accurate local thickness. This distance could then be considered as the local thickness of the object. Thus, we are dealing with a constrained optimization problem: minimizing a distance in a specific region. In our approach, however, this region is identified precisely by the point where we want to define the thickness. Therefore, the

[8] When computing T at a non-GM voxel, the value will be zero due to the stopping criteria explained below.

[9] The proposed framework is independent of how the probability density functions are actually computed. The key is to use the probabilities instead of the hard thresholds. If the soft classifier which provides the probability map is naïve, it will only take into account the intensity values of the MR image, or may in addition use an atlas as prior information. On the contrary, a more intelligent classifier will take into account other factors such as the fact that voxels that are definitely in the gray matter layer must have probability of 1, no matter what can immediately be inferred from the MR intensity values (e.g., in (53), interior voxels are clearly identified). The more accurate the input probabilities are, the better results our thickness measurement algorithm will provide; as the measurement algorithm relies on the accuracy of the input data. The computation of these probabilities for the examples in this paper is detailed in Appendix 9.1.1.

constraint is that the point must be on the line segment connecting the two finger tips. In other words, we consider only the line segments passing through the point where we intend to find the thickness. The minimized distance – or the length of the line segment – is in this case the integral of the probability map on the line containing the segment (Fig. 3.3(e)).

3.3.2 Algorithm

The algorithm basically computes every line integral centered at each point of the volume starting from the point of interest and proceeding in each of the two opposite directions separately (see Appendix 9.1.3 for our discretization method). Once all the line integrals at a point are calculated (meaning in all possible directions), the minimum of them is considered to be the thickness at that point. However, to reduce the effect of noise, an alternative would be for instance to consider the average of some of the smallest integrals.

In practice, a problem may arise, typically in narrow sulci where the outer surface of the folded GM layer has two nearby sides (Fig. 3.4(a)). While computing the thickness on one part of the layer, the GM of the other part may be partially included in some of the line integrals; this will lead to the thickness being overestimated (Fig. 3.4(b)). To avoid this error, we include two stopping criteria, which prevent a line integral from further advancing when it is believed that no more summation is necessary or that we are mistakenly considering a different region of the GM layer. The line integral stops proceeding if the probability map:

1. has been below a specific threshold for a certain number of consecutive voxels, or
2. has been decreasing at least for a certain number of successive voxels and then increasing for an additional number of voxels.

We use the first criterion, since if the probability has been low for a while, we are most likely not in the GM region anymore, and by further summing we would just increase the error. An additional advantage of using this stopping criterion is that summation will be stopped quickly after starting to measure the thickness based on voxels that are not in the GM region. The algorithm will ignore those points and will return near-zero values for the GM thickness at those locations.

The second condition happens when two parts of the GM layer are so close to each other that the probability on the gap between them is not small enough for the first stopping condition to become true, therefore the algorithm stops summing after identifying a valley on the probability

map. As we see in Appendix 9.1.3, the algorithm can be implemented such that gaps as narrow as one voxel are detected by the above stopping criteria.[10]

3.4 Results and Discussion

3.4.1 Artificial Data

To illustrate and validate our approach, we first show results using artificial input data.[11] Fig. 3.5(a) shows the isosurfaces of an artificially created probability map of a parabolic-shaped layer of GM with varying thickness in a volume of 50×50×50 voxels. The two isosurfaces represent the inner and outer gray matter surfaces. Depicted as small circles, a number of sample points have been selected, where the computed thicknesses are illustrated as line segments. The direction of each line segment is the optimal direction that gives the minimum line integral of the probability map. The thickness is indicated in the figure by the length of the line segments. To demonstrate the sensitivity and robustness of our algorithm to noise, we added zero-mean Gaussian noise to the probability map, each time with a different standard deviation. We then computed the mean relative error in the thickness, depicted in Fig. 3.5(b), using the thickness obtained in the noiseless case as the ground truth. As expected from an approach based on integration, a linear behavior is encountered. Of course the standard deviation of the relative error in the computed (noisy) thickness behaves like the standard deviation of the noise scaled by the inverse of the square root of the segment length, meaning that by increasing the resolution, this error is decreasing and the noise in the computation is reduced.

Next, to show the negative consequences of hard segmentation when noise and partial volume effects are present, we reduced the resolution of the volume five times by taking the mean value of every 5×5×5 sub-volume; we also added zero-mean Gaussian noise with standard deviation of 0.2 (Fig. 3.5(d)), and ran our measurement algorithm on it. In addition, we performed hard segmentation on the low-resolution, noisy volume, by substituting the probability values less than 0.5 with 0 and other values with 1 (Fig. 3.5(e)), and re-ran the measurement algorithm. Using the results of the original high-resolution case as ground truth, the experiments on the low resolution and noisy volumes showed an average error in the estimated thickness of 1.9 voxels in the

[10] In case the gap is narrower, both GM layers basically touch each other and the algorithm may overestimate the thickness. Restricting the thickness map to be continuous, which is a part of the future work, may help us overcome this problem. For a further discussion, see e.g. (53).

[11] Due to the lack of ground truth data for tissue thickness (and even a universally accepted definition of tissue thickness), artificial examples as the one here presented are critical to illustrate the importance of the concepts introduced here.

segmentation-free case and 2.2 voxels when hard segmentation was performed as a pre-step for reporting this measurement.

As a further demonstration of the robustness to noise of our soft-classification based algorithm, consider voxels A and B in Fig. 3.5(d). As implied by the symmetry of the original object, the two voxels had equal values (of about 0.5) before the noise was added, which changed the values of A and B to about 0.4 and 0.6, respectively. Since the contributions of these two voxels to line integrals differ from the original value only by a small amount of 0.1, our algorithm results in fairly accurate local thicknesses around A and B. On the contrary, the hard classification wrongly categorizes the two (noisy) voxels differently, as "outside" and "inside" (Fig. 3.5(e)) which affects the measured local thickness noticeably, resulting in significantly different thickness values around A and B.

3.4.2 Real MRI Data

We tested the proposed technique on 44 T1-weighted brain MRI scans, acquired using a standard sagittal 3D MP-RAGE sequence (TR: 2400 ms, minimum TE, inversion time (TI) 1000 ms, flip angle: 8°, 24 cm field of view), with a reconstructed voxel size of 0.9375×0.9375×1.2 mm^3. To adjust for scanner – and session – specific calibration errors, standard corrections were made for gradient nonlinearity, phantom-based scaling, and adjustment of intensity inhomogeneity (52). The actual tissue probabilities $P(\vec{x})$ needed in our method are computed following the procedure explained in Appendix 9.1.1. All computations are done on native space.

A 2D slice from an MRI volume is shown in Fig. 3.6(a) along with its computed thickness map in Fig. 3.6(b). Since we do not extract the GM, the results also contain thickness values for other parts of the head such as the scalp, which may be ignored. As shown in the figure, at this stage the algorithm cannot detect some sulci/gyri where the resolution is too low and two parts of the outer surface touch each other. However, the technique developed by Teo *et al* (53), can be used in this case to *split* the merged GM, and then compute the corresponding tissue thickness for each part. Fig. 3.6(c) illustrates a 3D surface-based mapping of the cortical thickness visualized by the mrGray software, using the steps in (53). As can be seen, although the computed thickness values are volumetric data, it is trivial to map them back onto the mesh (as mrGray was used here to do so), in order to make them usable by other mesh based software. We also note that smoothing data across the cortex – an essential step for voxel/vertex wise statistics – may also be done prior to performing a group statistical analysis of cortical thickness. One approach to do this is by Laplace-Beltrami smoothing of the scalar field on a mesh by using a covariant PDE (54,55), or by

Laplace-Beltrami smoothing using an implicit function whose zero level set is the cortical surface (56).

Our dataset includes pairs of scans over a one-year interval from 22 subjects (total of 44 scans), of whom 9 had been diagnosed with Alzheimer's disease at their first scan, and 13 were age-matched normal subjects. These subjects were included in one of our prior morphometric studies, where the scanning protocol is detailed (57). Each subject was scanned at 1.5 Tesla with a 3D T1-weighted acquisition, with the following parameters: repetition time (TR), 2400 ms; minimum full TE; inversion time (TI), 1000 ms; flip angle, 8°, 24 cm field of view, yielding a reconstructed voxel size of $0.9375 \times 0.9375 \times 1.2$ mm^3. The images were calibrated with phantom-based geometric corrections to ensure stable spatial calibration over time (58).

For comparison,[12] we also analyzed our data using the FreeSurfer thickness computation technique (35). The results of the change in the mean thickness over a year are demonstrated for the individual cases in Fig. 3.7, while the corresponding detailed statistical data can be seen in Table 3.1. From the computational point of view, our approach is 1-2 orders of magnitude faster than the one implemented in FreeSurfer, and it is highly parallelizable (note of course that FreeSurfer computes other characteristics as well, while we concentrate here on the tissue thickness computation step).

Both techniques are able to detect a systematic change in cortical thickness over time, both in the AD (Alzheimer's disease) group and in controls. Our method found thickness declined in AD (t_{paired} = -1.200; p=0.001), and in controls (t_{paired} = -0.274; p=0.009); FreeSurfer found thickness declined in AD (t_{paired} = -1.495; p=0.001), and increased in controls (t_{paired} = +0.175; p=0.001).[13]

First of all, the thickness measures are extremely highly correlated between baseline and follow-up scan, for both measurement methods and both subject groups. Our technique gives measures that correlated highly over the 1 year time-interval in AD (p=0.001) and in controls (p=0.009), and so did FreeSurfer (p=0.001 both for AD and for normal subjects).

Second, although changes over time are more difficult to measure than absolute values of thickness, our technique correlates highly with FreeSurfer in the changes it measures (Pearson's r=0.572; p=0.0028). This agreement between methods was also found in the group of healthy

[12] See Appendix 9.1.4 for a detailed explanation about what voxels are used to perform this comparison (recall that our algorithm produces thickness measurements for all voxels).

[13] This increase, which is anatomically not expected, is statistically significant. No increase is found with our proposed method.

controls, where changes are minimal (r=0.583; p=0.0011), and in the group of AD patients (r=0.558; p=0.0047).

The changes over time that were detected by FreeSurfer tended to be a little lower in general than those detect by our method, and satisfied the following least-squares regression model: $\Delta\text{thickness}_{FS} \sim [0.45*\Delta\text{thickness}_{OurMethod}]-0.04$.[14] Even so, a Student's t test designed to compare thickness values across the two methods did not recover any systematic biases between methods (t_{paired}=0.557, p>0.05). Our method gives a slightly higher SD in the thickness measures at each time point, and this, together with the fact that these errors may be uncorrelated over time, may also lead to a slightly higher estimation of change. Even so, if the sources of errors/deviations are uncorrelated across subjects, there should be no bias incurred by using one method versus the other.

A number of observations can be made from these experiments and the comparison with FreeSurfer. While the actual exact value of cortical thickness decline is not known for the AD subjects, for biological reasons a net increase is not expected for either of the two populations. This is partly because cortical thinning is a natural process that occurs due to neuronal shrinkage in normal aging, and is accelerated in AD due to cell death and neuronal loss in the cortex. Thereby, any apparent increase points to the difficulty and measurement errors in cortical thickness methods. Both our method and FreeSurfer report a few subjects that show increases (less pronounced with our approach), while on average our proposed method shows decline for both populations (much more significant for the AD subjects, as expected). Even so, FreeSurfer shows a net increase for the control subjects, not very large but still statistically significant. A number of reasons might explain this, including registration issues and segmentation errors, which are minimized in our proposed technique. More advanced soft classification techniques will improve the probability assignments and as a result the accuracy of the thickness computed by the proposed approach. Due to the existence of such intrinsic difficulties in measuring cortical thickness, the simultaneous running of different algorithms, with an evaluation of any inconsistencies or consensus, is an important alternative.

[14] We should note that while this is the best fitting straight line, in a least squares sense, the actual fitting is far from ideal, indicating that the relation between the measurements produced by both methods is not linear. Note of course that intrinsically our definition of tissue thickness is fundamentally different from that of FreeSurfer, and should be considered as an alternative.

3.4.3 Correlation with Clinical Data

In addition to classification accuracy, it is desirable that any measure of cortical thickness can be shown to be associated with clinical measures of deteriorating brain function. This is because image-derived measures often serve as a proxy for measures of disease burden that are based on repeated cognitive tests, repeated pathological tests (lumbar puncture). When we modeled factors that affected rates of thinning in AD, our method found a significant sex difference in the rate of thinning (r=0.428; p<0.037), whereas FreeSurfer detected this sex difference at a trend level (r=0.389; p<0.060), with women experiencing a more rapid rate of loss. Overall, the correlations with clinical measures in AD, such as the Mini-Mental State Exam scores at baseline and 6 month follow-up, and the changes in those scores, were around r=0.37-0.5 for FreeSurfer (p=0.013-0.073), and slightly lower for our method (r=0.24-0.30; p>0.1). Even so, this may be due to the slightly higher standard deviation for the changes reported by our method. Conversely, our method detected associations between the cortical thinning rate and Geriatric Depression Scores (r=0.433; p=0.0215), but FreeSurfer did not detect such an association (r=0.106; p>0.1). Clearly, a head-to-head comparison on a larger sample would be useful, including, for example, assessments of the sample sizes needed by both approaches to detect a 25% slowing of the disease with 80-90% power.

Table 3.1. Statistical Results on Real MRI Data

Quantity	Proposed Algorithm	FreeSurfer
Mean change for all subjects	-0.69 %	-0.35 %
Mean change for AD subjects	-1.42 %	-1.15 %
Mean change for normal subjects	-0.19 %	+0.20 %
SD of change for all subjects	3.42	2.68
SD of change for AD subjects	3.52	2.40
SD of change for normal subjects	3.39	2.82
Group separation	0.36	0.50

The relative change in the mean thickness as measured over a one year interval.
Group separation = (Mean for normal - Mean for AD) / (SD for all)

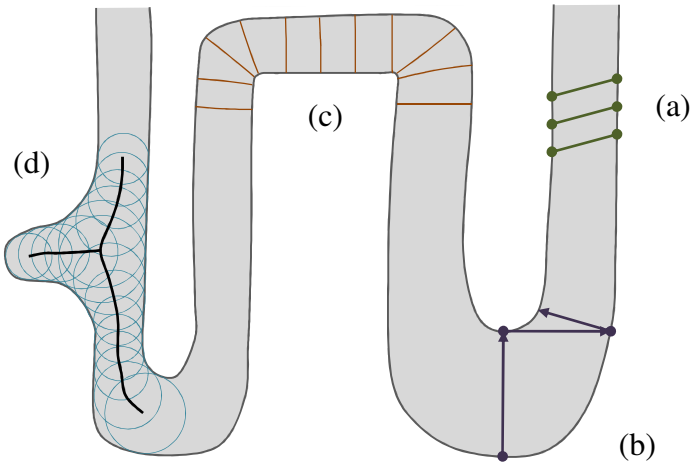
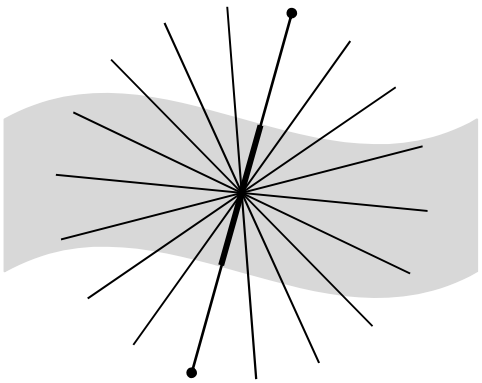
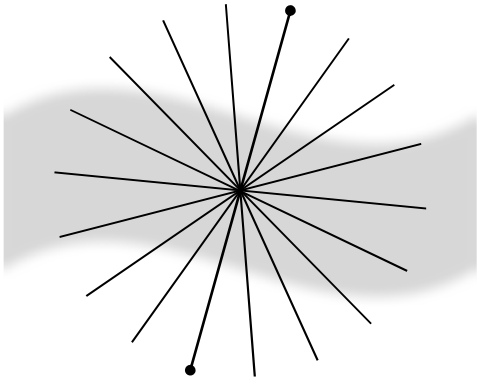


Fig 3.1. Common ways of measuring cortical thickness. (a) Coupled-surface methods. (b) Closest point methods. (c) Laplace ('heat-flow') methods. (d) Largest enclosed sphere methods.



(a)



(b)

Fig 3.2. Computing line integrals passing through a point, and choosing the minimum integral value as the thickness. (a) Binary probability map. (b) Continuous probability map.

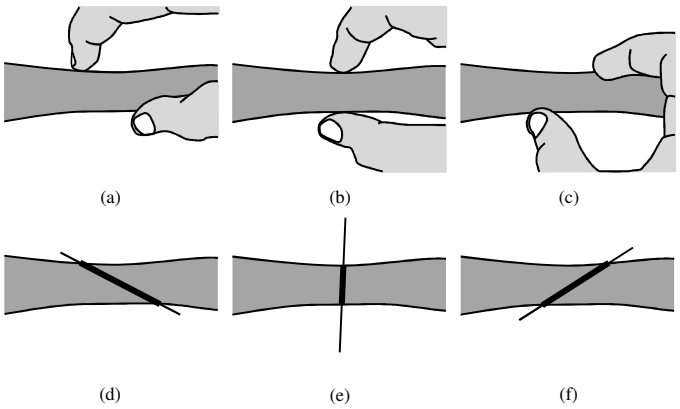
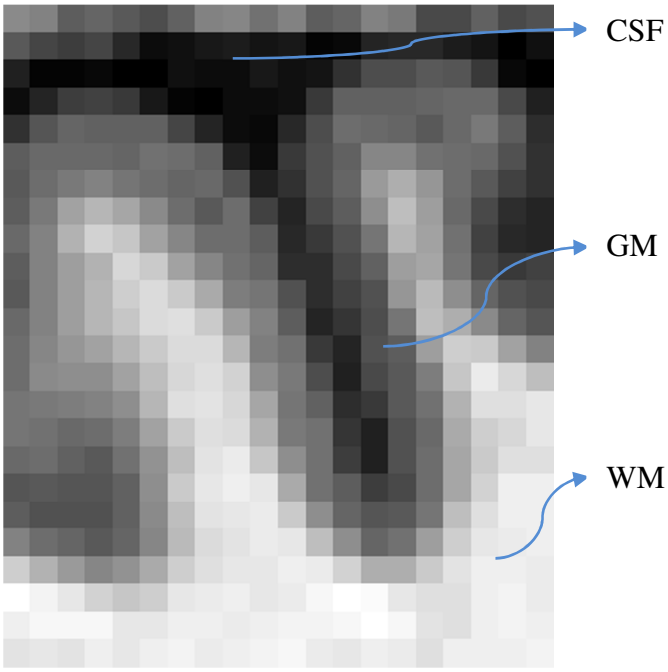
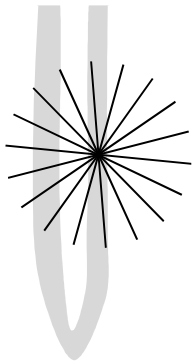


Fig 3.3. (a), (b), (c) An intuitive way of measuring the thickness of an object. (d), (e), (f) Our algorithm produces results similar to this intuitive approach.



(a)



(b)

Fig 3.4. (a) A sulcus in which two sides of the gray matter layer are close to each other. (b) How the algorithm might overestimate the thickness if no stopping criteria were used.

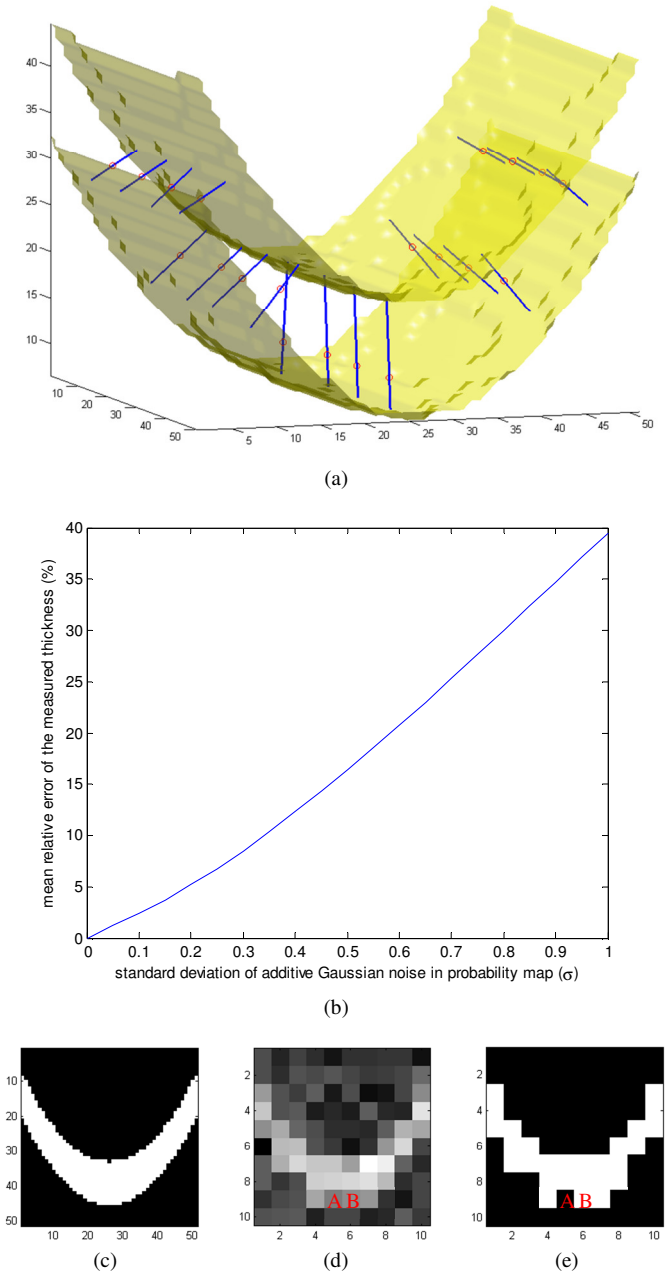


Fig 3.5. Results on an artificial probability map. (a) Inner and outer surfaces of a parabolic-shaped layer of "GM" are depicted. Line segments are chosen by the algorithm such that they give the smallest integrals (of the probability map) among all line segments passing through every selected test point, here shown as small circles. (b) Relative error in measured thickness, introduced by additive Gaussian noise. (c) A 2D slice of the volume. (d) The same slice in a five-times-lower-resolution volume with additive Gaussian noise. (e) Binary classification of the low-resolution volume.

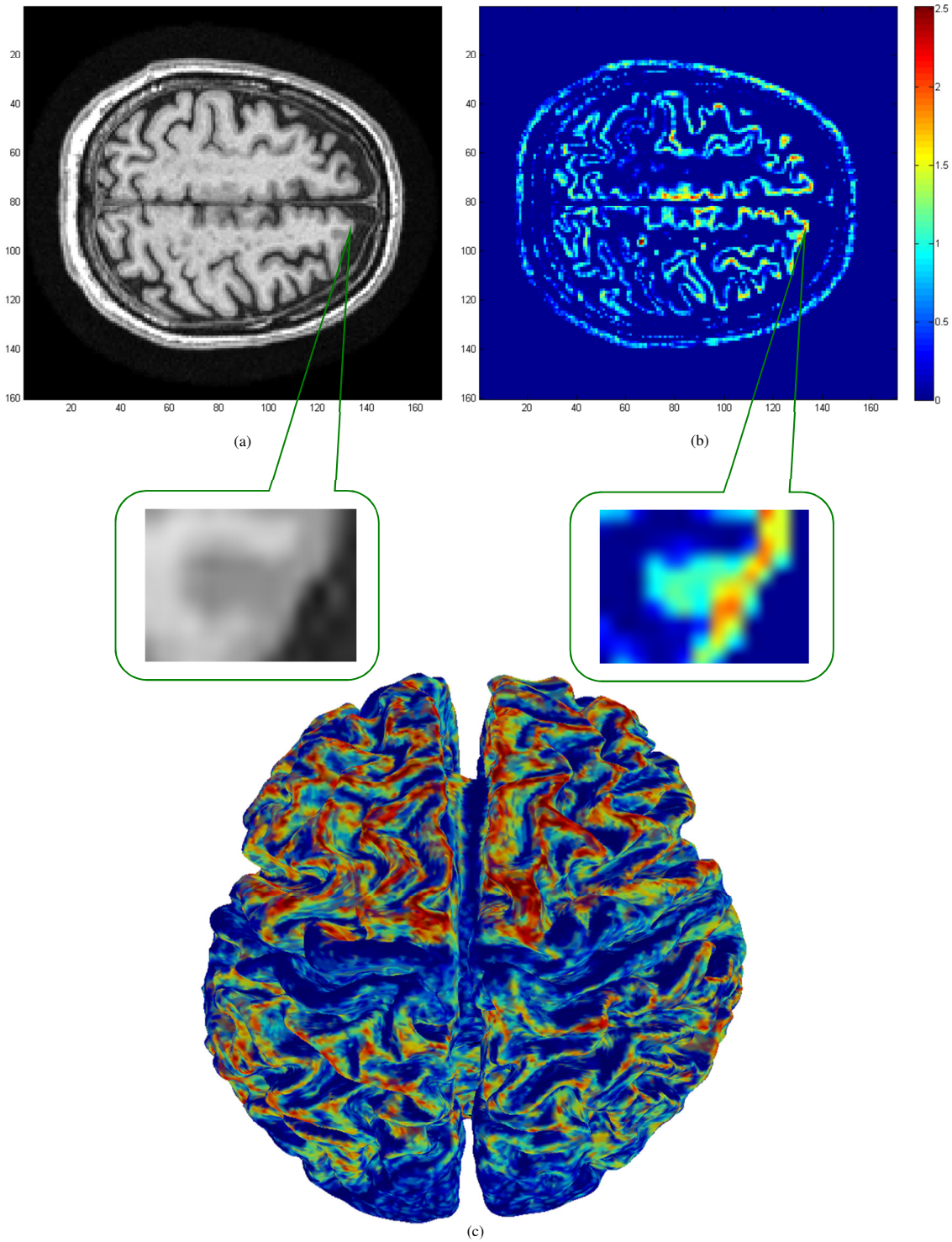
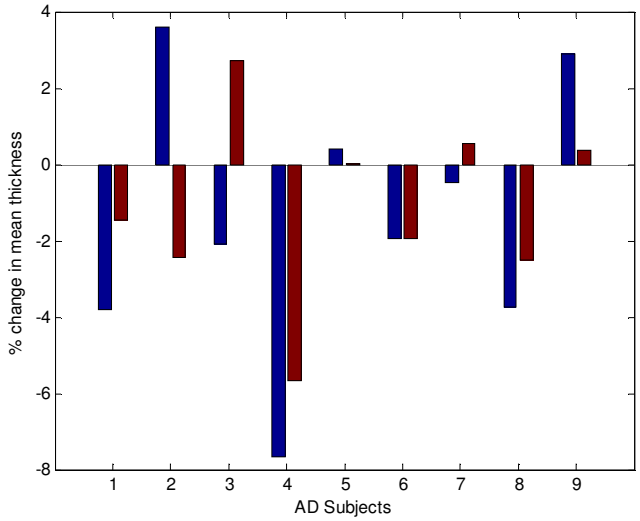
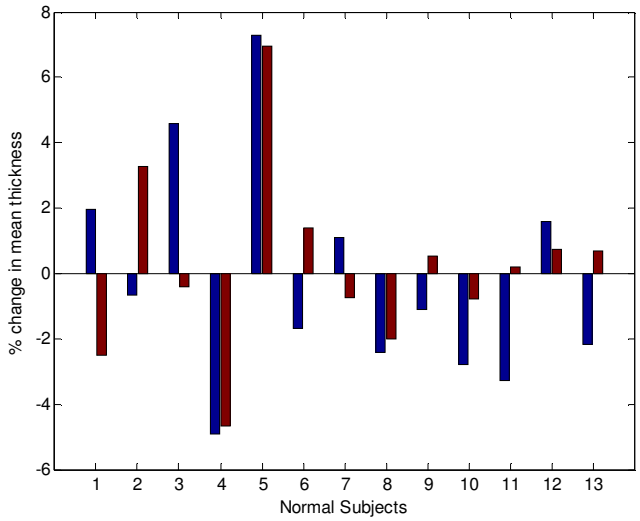


Fig 3.6. Experimental results on MRI data. All computations were done in 3D. A zoomed-in version of a sulcus is shown where low resolution results in overestimation of the cortical thickness. (a) A slice of the original volume. (b) The thickness map of the same slice (blue thinner, red thicker). (c) 3D mapping of the cortical thickness. Note the thinner areas in the pre/post central gyri.



(a)



(b)

Fig 3.7. Relative change in the mean cortical thickness over a one-year interval, comparing our results (left blue bars) with the results from FreeSurfer (right red bars). (a) Subjects diagnosed with Alzheimer's disease (AD). (b) Normal elderly subjects.

4 Orientation Distribution Function in Constant Solid Angle

4.1 Introduction

Diffusion-weighted magnetic resonance imaging provides valuable information about the fiber architecture of tissue by measuring the diffusion of water in three-dimensional (3D) space. The microscopic diffusion may be measured using the model-free diffusion spectrum imaging (DSI) (59), which exploits the direct Fourier inversion of the diffusion signal. This technique is time intensive, as it measures the signal on a 3D (e.g., 11×11×11) Cartesian lattice. Thus, an alternative approach based on sampling only on one or multiple spherical shells in q-space has been proposed, referred to as high angular resolution diffusion imaging (HARDI) (60). The spherical shell, being a 2D manifold, includes a number of measurement points which grows quadratically with the desired angular resolution, as opposed to cubically with the spatial resolution in the entire 3D lattice of DSI.

While the 3D probability density function (PDF) of diffusion is helpful in studying the tissue microstructure, the orientation distribution function (ODF) – the marginal probability of diffusion in a given direction – is the quantity of interest for mapping the orientation architecture of the tissue. Q-ball imaging (QBI), (61), is a widely used reconstruction scheme for HARDI, from which ODFs are approximated through a spherical tomographic inversion called the Funk-Radon transform (62). This technique's simplicity and its ability to resolve intravoxel fiber orientations have made it popular for fiber tracking and characterizing white matter architecture. A number of recently proposed methods have turned QBI into a very efficient and robust technique (63)– (69). Moreover, a few works have suggested exploiting data from multiple q-shells to benefit from the high signal-to-noise ratio (SNR) and high angular contrast-to-noise ratio (CNR) of the data acquired at respectively low and high b-values, (61), (70)– (72). Using multiple q-shells also allows us to employ richer models for the diffusion signal, as discussed in this paper. Nevertheless, QBI only allows us to compute the diffusion ODF, which is a blurred version of the underlying fiber distribution (fiber ODF). To overcome this blurring, spherical deconvolution methods have been proposed, (73)– (75). Other strategies used in HARDI include the Persistent Angular Structure (76) and Diffusion Orientation Transform (DOT) (77), both of which compute non-ODF quantities revealing the orientations of microstructural fibers. (For a comparison between all the mentioned methods, see (78)– (79).) In this work, we attempt to reduce the

blurring in QBI by defining the ODF as the true marginal probability of diffusion, and proposing a suitable computational approach.

The definition of the ODF used in the original QBI is, however, different from the actual marginal PDF of diffusion in constant solid angle. It is computed as a linear radial projection of the PDF, which does not take into account the quadratic growth of the volume element with respect to its distance from the origin (see Sec. 4.2.1 and Fig. 4.1 for more details). This inaccurate formulation generally distorts the ODF, produces non-distribution functions, and has created the need for artificial post-processing such as manual normalization and sharpening.

In this paper, we re-derive the ODF expression for QBI via Fourier analysis, this time starting from the proper definition of the ODF in constant solid angle (CSA). We show that this results in an inherently normalized and dimensionless expression. In addition, we illustrate through our experiments that the new ODFs are naturally sharp and thus multiple fiber orientations are better resolved, potentially improving tractography. We also provide a general formulation for multiple q-shell QBI, and demonstrate the improvement achieved by considering the information from multiple q-shells and using richer multi-exponential models. Furthermore, by making use of the spherical harmonic basis, we demonstrate that the implementation of the new, mathematically correct expression is as straightforward as that of the original formula, or perhaps even simpler, considering that further sharpening (post-processing) is not necessary.

This work was presented in (80)– (83),[15] and later extended to the on-the-fly ODF computation via Kalman filtering (84), motion detection (85), and the ODF computation with spatial regularity (86). It has also been implemented in the *package dti* (87) and the MITK-DI (in MITK 3M3 Image Analysis) (88).

[15] After our conference paper was accepted and its extension to multiple shells was submitted, a parallel and independent work was published (69), where the proper definition of the ODF was considered in *single* q-shell QBI. However, in addition to not considering multiple shells and the richer models as done here, the authors of (69) take the integral of the diffusion signal on a circle and not on the entire plane, and that results in a different formula which is not necessarily normalized and leads to other potential inaccuracies. (See Sec. 4.2.3 for further details.)

4.2 Methods

4.2.1 General ODF Definition

The PDF of the diffusion of water molecules, $P(\vec{r})$, gives the displacement probability $P(\vec{r})dv$ of a molecule, initially placed at the origin, to be in the infinitesimal volume dv located at $\vec{r}$ after a certain amount of time. We make the common assumption that this function is symmetric, i.e. $P(-\vec{r}) = P(\vec{r})$. The PDF can be presented in Cartesian coordinates with $\vec{r} = (x, y, z)^T$ and $dv = dxdydz$. However, for mapping the orientation architecture of the tissue, the representation which mostly interests us is in the standard spherical coordinates, parameterized by (r, θ, ϕ), where $\vec{r} = r\hat{u}$ with $\hat{u}(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^T$ the unit direction vector. The volume element in this case is $dv = r^2 dr d\Omega$ with $d\Omega = \sin\theta d\theta d\phi$ being the infinitesimal solid angle element.

We denote by $ODF(\hat{u})d\Omega$ the probability of diffusion in the direction $\hat{u}$ through the solid angle $d\Omega$, which is computed by integrating the displacement probabilities, i.e. $P(\vec{r})dv = P(r\hat{u})r^2 dr d\Omega$, for all magnitude r, while keeping $\hat{u}$ constant:

$$ODF(\hat{u})d\Omega = \int_{r=0}^{r=\infty} P(r\hat{u})r^2 dr d\Omega ,$$

or simply:

$$ODF(\hat{u}) = \int_0^\infty P(r\hat{u})r^2 dr . \qquad [2]$$

The above definition, which is normalized and dimensionless, is the integral of the probability values in a cone of "very small" constant solid angle (Fig. 4.1, left). This correct definition was used for instance by the authors of (59) in DSI, where $P(\vec{r})$ was first computed from the diffusion data via Fourier inversion and then integrated to calculate the ODF, and also in (89)– (90) for diffusion tensor imaging (DTI), where the ODF was analytically computed. However, the original expression for ODF reconstruction in QBI (61), is different from Eq. [2], in the sense that the integral is not weighted by the important factor r^2 (Fig. 4.1, right). A recent parallel work (69) (published independently after a conference version of our paper (81) had just been accepted), considered this factor in (single shell) QBI and approximated the ODF using Eq. [2]. (See Sec. 4.2.3 for details.)

Computing the ODF without the factor r^2 would be equivalent to assuming the PDF to be $P(\vec{r})/|\vec{r}|^2$, as $\int_0^\infty P(r\hat{u})dr = \int_0^\infty \frac{P(r\hat{u})}{r^2} r^2 dr$. This radial projection gives an artificial weight to $P(\vec{r})$ which is, respectively, too large and too small for points close to and far from the origin, and in fact, the computed quantity would be different just as the zeroth moment of a one-dimensional function $\overline{P}(r) := P(r\hat{u})$ is different from its second moment. For instance, a consequence of not including r^2 is that the computed ODF will not be necessarily normalized, and an artificial normalization factor will be required. Moreover, the ODF will not be dimensionless, since, given that $P(\vec{r})$ has the dimension of L^{-3} (L being the length), the dimensions of $P(\vec{r})r^2 dr$ and $P(\vec{r})dr$ are respectively 1 and L^{-2}.

As an example intended for comparison, we compute the ODFs with r^2 (which we shall call *CSA*-ODF when comparing to the original method) and without r^2 (*original* ODF) in the case of DTI (91), with the following standard Gaussian PDF:

$$P(\vec{r}) = \frac{1}{(2\pi)^{\frac{3}{2}}|D|^{\frac{1}{2}}} e^{-\frac{1}{2}\vec{r}^T D^{-1}\vec{r}}, \quad [3]$$

where D is the covariance matrix (proportional to the diffusion tensor). The computed ODFs are:

$$\underset{CSA}{ODF}(\hat{u}) = \frac{1}{4\pi|D|^{\frac{1}{2}}(\hat{u}^T D^{-1}\hat{u})^{\frac{3}{2}}} \quad [4]$$

$$\underset{orig.}{ODF}(\hat{u}) = \frac{1}{Z}\frac{1}{4\pi|D|^{\frac{1}{2}}(\hat{u}^T D^{-1}\hat{u})^{\frac{1}{2}}}, \quad [5]$$

where Z is the normalization constant that subsequently needs to be computed and considered in the original $ODF(\hat{u})$ (see (61)). An example of this pair of ODFs is illustrated in Fig. 4.2. (No min-max normalization is used in any of the figures.)

Next, we derive a closed-form expression for the ODF in QBI using the correct r^2-weighted integral.

4.2.2 Q-ball Imaging ODF Reconstruction

Let $E(\vec{q})$ be the 3D Fourier transform of $P(\vec{r})$. We have the values of $E(\vec{q})$ measured on a q-ball, i.e., the frequencies with constant norm $|\vec{q}| = q_0$, as $\tilde{E}(\hat{u}) := E(q_0\hat{u}) = S(\hat{u})/S_0$, where $S(\hat{u})$ is

the HARDI signal and S_0 is the non diffusion-weighted (or B0) image. In addition, since the diffusion signal at $\bar{q}=0$ is S_0, one can see that $E(0)=1$. Alternatively, $E(0)$ is the zero frequency of a PDF which is its integral over the entire space, yielding 1.

Our mathematical derivation is based on the following two fundamental facts from Fourier analysis:

- The Fourier transform of $P(\bar{r})|\bar{r}|^2$ is $-\nabla^2 E(\bar{q})$, where ∇^2 is the Laplacian operator (proof presented in Appendix 9.2.1).
- For a symmetric function $f:\mathbb{R}^3\to\mathbb{R}$ with the 3D Fourier transform function $\hat{f}(\bar{q})$, and for the arbitrary unit vector $\hat{u}$, we have that $\int_0^\infty f(r\hat{u})dr=\frac{1}{8\pi^2}\iint_{\hat{u}^\perp}\hat{f}(\bar{q})d^2\bar{q}$, where $\hat{u}^\perp$ is the plane perpendicular to $\hat{u}$ (proof presented in Appendix 9.2.2).

Combining these statements with Eq. [2] leads to

$$ODF(\hat{u})=-\frac{1}{8\pi^2}\iint_{\hat{u}^\perp}\nabla^2 E(\bar{q})d^2\bar{q}\,. \qquad [6]$$

Now, without loss of generality, we choose our coordinates such that $\hat{z}=\hat{u}$, thus making $\hat{u}^\perp$ the q_x-q_y plane. We then use the following expansion for the Laplacian in spherical coordinates (q,θ,ϕ):

$$\nabla^2 E(\bar{q})=\frac{1}{q}\frac{\partial^2}{\partial q^2}(qE)+\frac{1}{q^2}\nabla_b^2 E\,, \qquad [7]$$

where ∇_b^2 is the Laplace-Beltrami operator defined independently of the radial component q, as $\nabla_b^2 E=\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial E}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 E}{\partial\phi^2}$. The surface integral on the q_x-q_y plane is computed by fixing $\theta=\frac{\pi}{2}$ and using the expression $d^2\bar{q}=qdqd\phi$ as the surface element (see also recent work (68)), which yields

$$\begin{aligned} ODF(\hat{z}) &= -\frac{1}{8\pi^2}\int_0^{2\pi}\int_0^\infty \nabla^2 E(\bar{q})\,qdqd\phi \\ &= -\frac{1}{8\pi^2}\int_0^{2\pi}\int_0^\infty\left(\frac{1}{q}\frac{\partial^2}{\partial q^2}(qE)+\frac{1}{q^2}\nabla_b^2 E\right)qdqd\phi \end{aligned} \qquad [8]$$

We can see that the integral of the first term is constant and independent of $E(\bar{q})$ and its derivatives:

$$\int_0^\infty \left(\frac{1}{q}\frac{\partial^2}{\partial q^2}(qE)\right) q dq = \int_0^\infty \frac{\partial^2}{\partial q^2}(qE)\, dq$$
$$= \left[\frac{\partial}{\partial q}(qE)\right]_0^\infty$$
$$= \left[E + qE_q\right]_0^\infty$$
$$= E(\infty) - E(0) + \left[qE_q\right]_{q=\infty} - \left[qE_q\right]_{q=0}$$
$$= -1$$
,

$$\int_0^{2\pi}\int_0^\infty \left(\frac{1}{q}\frac{\partial^2}{\partial q^2}(qE)\right) q dq d\phi = -2\pi\,, \qquad [9]$$

where the subscript indicates the partial derivative. We made the standard assumptions that the diffusion signal and its radial derivative go to zero (sufficiently fast) as $q \to \infty$, and also that the derivative is bounded at the origin. Therefore we have

$$ODF(\hat{z}) = \frac{1}{4\pi} - \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^\infty \frac{1}{q}\nabla_b^2 E(\vec{q})\, dq d\phi\,, \qquad [10]$$

while $\theta = \frac{\pi}{2}$ is kept constant in the integration.

To compute the integral of the second term, the values of $E(\vec{q})$ are required in the entire q-space (as the radial integral is from zero to infinity). The above equation could be used for example in DSI, where direct computation of the ODF from the diffusion images would eliminate the need for 3D Fourier inversion. In QBI, however, the values of $E(\vec{q})$ are available only on the q-ball, from which $E(\vec{q})$ needs to be approximated. In this work, we consider the following radial mono-exponential model:

$$E(\vec{q}) \cong E(q_0\hat{u})^{\frac{q^2}{q_0^2}} = \tilde{E}(\hat{u})^{\frac{q^2}{q_0^2}}\,, \qquad [11]$$

where q_0 is the radius of the q-ball. This type of interpolation has been presented and discussed in (92) and (77) as $E(q\hat{u}) \cong e^{-\tau q^2 \cdot ADC(\hat{u})}$, where the Apparent Diffusion Coefficient (ADC) is assumed to be constant in each direction, and estimated as $ADC(\hat{u}) = -\frac{1}{\tau q_0^{\,2}}\ln E(q_0\hat{u}) = -\frac{1}{\tau q_0^{\,2}}\ln \tilde{E}(\hat{u})$. An advantage of this model over the original QBI model, i.e. $E(q\hat{u}) \cong \tilde{E}(\hat{u})\delta(q - q_0)$ (see (93,68,94)), is the compatibility with $E(0) = 1$.

After applying the mono-exponential assumption and a few more steps of calculations (see Appendix 9.2.3 for details), the following ODF expression is derived:

$$ODF(\hat{z}) = \frac{1}{4\pi} + \frac{1}{16\pi^2}\int_0^{2\pi} \nabla_b^2 \ln\left(-\ln \tilde{E}(\hat{u})\right) d\phi \,. \qquad [12]$$

Finally, rewriting the expression independently of the choice of axes, the following analytical formula can be shown to hold for the ODF:

$$ODF(\hat{u}) = \frac{1}{4\pi} + \frac{1}{16\pi^2} FRT\left\{\nabla_b^2 \ln\left(-\ln \tilde{E}(\hat{u})\right)\right\}, \qquad [13]$$

where FRT is the Funk-Radon transform (62), defined as

$$FRT\left\{f(\hat{u})\right\} := \iint_{\hat{u}^\perp} f(\vec{w})\,\delta(|\vec{w}|-1)\,d^2\vec{w}\,, \qquad [14]$$

with $\delta(\bullet)$ the Dirac delta function. Next, we compare our method from the theoretical aspect with some other approaches.

4.2.3 Theoretical Comparison

The CSA-ODF expression derived in Eq. [13] is dimensionless and intrinsically normalized, since the integrals of the first and second terms over the sphere are respectively 1 and 0. This is in contrast to the ODF formulas used in the original QBI, i.e., $\frac{1}{Z} FRT\left\{\tilde{E}(\hat{u})\right\}$, and also in (69), where an artificial normalization factor Z is needed.

Additional fundamental differences can be observed in the approach presented here, compared to (69). As we demonstrated above, integration of the radial part of the Laplacian on the plane always results in a constant ($1/4\pi$ in Eq. [10]) without assuming any model for the diffusion signal. Yet, (69) uses the Bessel approximation of the Dirac delta function which yields a variable (sometimes negative) term. As for the integral of the tangential term of the Laplacian, we use the exponential model that is in particular consistent with $E(0)=1$, in contrast to (69) that assumes the tangential term of the Laplacian to be zero everywhere except on the q-ball (Bessel approximation again), leading to an expression rather similar to Laplace-Beltrami sharpening (LBS) (95). A major disadvantage of approximating the Dirac delta with a Bessel function while considering the factor r^2 is that, unlike for $P(\vec{r})$ which is typically concentrated near the origin, the projection of $P(\vec{r})|\vec{r}|^2$ may have its highest values at a certain positive radius coinciding with the side lobes of the Bessel function, reducing the accuracy of the approximation (see also (94)).

From Eq. [13], it can be seen that the essential quantity used in computing the ODF from the raw data is $L(\tilde{E}) := \ln(-\ln \tilde{E})$, plotted along with the absolute value of its derivative in Fig. 4.3. Given that the second derivative of this function is zero at the inflection point $e^{-1} \approx 0.368$, $L(\tilde{E})$ is almost linear for the values of the signal close to that point. Therefore, from Eq. [13] it can be observed that the ODFs reconstructed from signals with values close to e^{-1} will be similar to those obtained by the original QBI (61) with LBS (note that FRT commutes with ∇_b^2). This resemblance ends as the range of the signal values approaches 0 or 1 (for example when the diffusion is respectively higher or lower, or the data is acquired at respectively higher or lower b-values), in which case $L(\tilde{E})$ becomes quite nonlinear, amplifying the measured signal. This nonlinearity can be concluded to be the main source of improvement seen in the experimental results obtained using Eq. [13] (predominantly seen in resolving the fiber crossings, see Sec. 4.3).

DOT (77) is a useful method for fiber orientation mapping that computes the PDF of the diffusion of water $P(r_0\hat{u})$ at a fixed distance r_0 from the origin, and is therefore a different quantity from ODF. DOT provides essential information about the fiber microstructure, yet, at the cost of an additional parameter, r_0, which is not always trivial to determine. The optimum r_0 depends on the fiber microstructure and may vary in different regions of the volume. In contrast, ODF sums up the orientation information for all radii. In fact, since CSA-ODF and DOT both use the exponential model for the diffusion signal, we see from Eq. [2] that,

$$ODF(\hat{u}) = \int_0^\infty DOT(r\hat{u})\, r^2 dr \,. \qquad [15]$$

4.2.4 Implementation

Our implementation of the ODF reconstruction makes use of the spherical harmonic (SH) basis, $Y_k^m(\hat{u})$, which is common for the analysis of HARDI data. The steps taken here to numerically compute Eq. [13] are similar to those described in (64). Particularly, we use the real and symmetric *modified SH* basis in (64), where SH functions are indexed by a single parameter $j = k(k+1)/2 + m + 1$, with corresponding k_j and m_j, as follows:

$$Y_j = \begin{cases} \sqrt{2}\,\mathrm{Re}\{Y_{k_j}^{m_j}\}, & -k_j \le m_j < 0 \\ Y_{k_j}^0, & m_j = 0 \\ \sqrt{2}\,\mathrm{Im}\{Y_{k_j}^{m_j}\}, & 0 < m_j \le k_j \end{cases} \qquad [16]$$

We adopt a minimum square error scheme to compute the modified SH coefficients c_j of the double logarithm of the signal, such that

$$\ln\left(-\ln \tilde{E}(\hat{u})\right) \approx \sum_{j=1}^{R} c_j Y_j(\hat{u}), \qquad [17]$$

where $R = (l+1)(l+2)/2$, with l being the order of the SH basis. Next, since the SH elements are eigenfunctions of the Laplace-Beltrami operator, we compute $\nabla_b^2 \ln\left(-\ln \tilde{E}(\hat{u})\right)$ by multiplying the coefficients c_j by their corresponding eigenvalues, $-k_j(k_j+1)$. Then, as suggested in (64), the Funk-Radon transform is computed by multiplying the coefficients by $2\pi P_{k_j}(0)$, where $P_k(\bullet)$ is the Legendre polynomial of degree k, with $P_k(0) = (-1)^{\frac{k}{2}} \frac{1\times 3\times\cdots\times(k-1)}{2\times 4\times\cdots\times k}$ for even k. Finally, given that $Y_1(\hat{u}) = \frac{1}{2\sqrt{\pi}}$, the SH coefficients of the ODF are derived as

$$c'_j = \begin{cases} \dfrac{1}{2\sqrt{\pi}} & j = 1 \\ -\dfrac{1}{8\pi}(-1)^{\frac{k_j}{2}} \dfrac{1\times 3\times\cdots\times(k_j+1)}{2\times 4\times\cdots\times(k_j-2)} c_j & j > 1 \end{cases} \qquad [18]$$

By taking advantage of the SH framework, this implementation of the proposed technique is as straightforward as the one introduced in (64) for the original QBI ODF formula. Additionally, neither normalization, nor sharpening is required.

4.2.5 Regularization

As mentioned before, the essential quantity used in computing the ODF from the raw data is $L(\tilde{E}) := \ln\left(-\ln \tilde{E}\right)$, plotted in Fig. 4.3. Hence, if there is a relatively constant error $\Delta\tilde{E}$ in the diffusion data, the error introduced in the computed ODF will be proportional to the derivative of $L(\tilde{E})$:

$$\Delta L(\tilde{E}) = \left|\frac{dL}{d\tilde{E}}\right| \Delta\tilde{E} = \frac{\Delta\tilde{E}}{\left|\tilde{E}\ln\tilde{E}\right|}. \qquad [19]$$

As Fig. 4.3 suggests, this quantity becomes unstable for values of $\tilde{E}$ very close to 0 and 1, and subsequently amplifies the error in the diffusion data.

To overcome this problem, besides the approach introduced in (86) to impose positivity constraint and spatial regularity, the more robust L_1 error norm may be used to compute the SH coefficients from Eq. [17]. Using a robust error norm, the noise will not contribute to the error function for high values of $L(\tilde{E})$ (when $\tilde{E}$ is very close to 0 or 1) as much as when the L_2 error norm is minimized, therefore the ODF will remain more robust to noise.

We also propose the use of a flexible threshold on the diffusion data to keep their values away from the unstable regions of $[0, \delta_1]$ and $[1-\delta_2, 1]$, where the thresholds δ_1 and δ_2 are manually defined. To perform this operation, we use the following function $f(\tilde{E})$:

$$f(\tilde{E}) = \begin{cases} \frac{\delta_1}{2}, & \tilde{E} < 0 \\ \frac{\delta_1}{2} + \frac{\tilde{E}^2}{2\delta_1}, & 0 \le \tilde{E} < \delta_1 \\ \tilde{E}, & \delta_1 \le \tilde{E} < 1-\delta_2 \\ 1 - \frac{\delta_2}{2} - \frac{(1-\tilde{E})^2}{2\delta_2}, & 1-\delta_2 \le \tilde{E} < 1 \\ 1 - \frac{\delta_2}{2}, & 1 \le \tilde{E} \end{cases}, \qquad [20]$$

which is plotted in Fig. 4.4 for $\delta_1 = \delta_2 = 0.15$. In practice, however, δ_1 and δ_2 do not need to be greater than 0.001 and therefore the distorting effect on the signal is minimal.

Conversely, the ODF is most stable to noise when $\Delta L(\tilde{E})$ is minimum, which is achieved for $\tilde{E} = e^{-1} \approx 0.368$. This gives us a clue on how to choose an optimum b-value in data acquisition. Particularly, in the mono-exponential model, since $\tilde{E} = e^{-b.ADC}$ where the ADC is assumed independent of the b-value, the optimum b-value is obtained as

$$b^* = \frac{1}{\langle ADC \rangle}, \qquad [21]$$

where $\langle ADC \rangle$ is the mean ADC in the region of interest. Note that this result holds only in the simple model which assumes both ADC and $\Delta\tilde{E}$ to be independent of the b-value. For further discussion about the optimal b-value in different techniques, see (79) and (96).

4.2.6 Extension to Multiple q-Shells

4.2.6.1 Multi-Exponential Model

We have so far employed the proposed technique to compute the ODF from a single q-shell. However, if diffusion data are available on multiple q-shells, this technique can be applied to reconstruct the ODF while exploiting the information from all the q-shells. With more available data, richer models become practical and appealing. Here we consider the following radial multi-exponential model (see (77,97)),

$$E(q\hat{u}) \cong \sum_{k=1}^{N} \lambda_k(\hat{u})\, \alpha_k(\hat{u})^{q^2}\,, \qquad [22]$$

with the constraints

$$0 < \alpha_k(\hat{u}), \lambda_k(\hat{u}) < 1\,,$$

$$\sum_{k=1}^{N} \lambda_k(\hat{u}) = 1\,, \qquad [23]$$

where Eq. [23] comes from $E(0)=1$. Once the values of λ_k and α_k are estimated (see Sec. 4.2.6.2), they can be used in the following more general ODF expression, which we have derived in details in Appendix 9.2.4:

$$ODF(\hat{u}) = \frac{1}{4\pi} + \frac{1}{16\pi^2} FRT\left\{ \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\ln \alpha_k(\hat{u})\right) \right\}. \qquad [24]$$

This can be implemented quite similarly to what we explained in Sec. 4.2.4, with Eq. [17] being the only difference, as it now writes

$$\sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\ln \alpha_k(\hat{u})\right) \approx \sum_{j=1}^{R} c_j Y_j(\hat{u})\,. \qquad [25]$$

In addition, the methods suggested in Sec. 4.2.5 can be applied to α_k s, to reduce the effect of noise.

4.2.6.2 Parameter Estimation

To approximate the diffusion signal in a direction $\hat{u}$ by a weighted sum of N exponentials, we need to estimate the $2N$ parameters $\lambda_k(\hat{u})$ and $\alpha_k(\hat{u})$, for $k = 1, \ldots, N$. Therefore, at least $2N-1$ independent equations – besides Eq. [23] – are required, which can be obtained from the HARDI signals measured on M q-balls, for $M \geq 2N-1$, as follows:

$$\sum_{k=1}^{N} \lambda_k(\hat{u})\,\alpha_k(\hat{u})^{q_i^2} = \tilde{E}_i(\hat{u}) \qquad i = 1,\ldots,M \quad , \tag{26}$$

where $\tilde{E}_i(\hat{u}) := E(q_i\hat{u})$ and q_i corresponds to the i^{th} q-ball. Parameterizing the problem in terms of b-values, $b_i = \tau q_i^2$, and choosing the physical units such that the diffusion time $\tau = 1$, we obtain

$$\sum_{k=1}^{N} \lambda_k(\hat{u})\,\alpha_k(\hat{u})^{b_i} = \tilde{E}_i(\hat{u}) \qquad i = 1,\ldots,M \quad . \tag{27}$$

Numerical optimization approaches such as the trust region algorithm, (98), may be employed to solve this non-linear system in the most general case. Here, however, we discuss two special cases (one familiar and one new) with analytical solutions. We continue this subsection considering a fixed direction, and therefore drop the notation $(\hat{u})$.

The mono-exponential assumption (*N=1*) requires measurement on at least *M=1* q-ball. *M=1* leads to $\lambda_1 = 1$ and $\alpha_1 = \tilde{E}_1^{1/b_1}$. As we have shown in Eq. [80] of Appendix 9.2.4, α_1^{γ} can also be a solution with any constant γ. Therefore, choosing $\gamma = b_1$ results in the solution $\alpha_1 = \tilde{E}_1$, which is consistent with what we already derived (Eq. [13]). Furthermore, if measured values are provided on more than one q-shell and the mono-exponential model is still desired, then the assumption in this model (ADC being independent of the b-value) suggests that the best exponential can be fitted by computing the average ADC across all the q-balls.

Another practical case of great interest arises when we consider the richer bi-exponential model (*N=2*, see for example (99)– (100)) to reconstruct the ODFs from (at least) *M=3* q-shells. For *M=3*, the following system of equations holds for each direction:

$$\begin{aligned} \lambda\alpha^{b_1} + (1-\lambda)\beta^{b_1} &= \tilde{E}_1 \\ \lambda\alpha^{b_2} + (1-\lambda)\beta^{b_2} &= \tilde{E}_2 \\ \lambda\alpha^{b_3} + (1-\lambda)\beta^{b_3} &= \tilde{E}_3 \\ 0 < \alpha, \beta, \lambda &< 1 \end{aligned} \quad . \tag{28}$$

In general, this set of equations can be solved numerically. Nevertheless, an analytical solution can be derived for the particular and reasonable case when the sequence $0, b_1, b_2, b_3$ is an arithmetic progress. (The sequence x_i is an arithmetic progress if $x_i - x_{i-1}$ is constant). We describe this solution here, along with some regularization that guarantees the parameters to remain within the correct range.

Without loss of generality, let us assume $\alpha \geq \beta$, and also choose the physical units such that $b_1 = 1$, $b_2 = 2$, and $b_3 = 3$. Then,

$$\begin{aligned} \lambda\alpha + (1-\lambda)\beta &= \tilde{E}_1 \\ \lambda\alpha^2 + (1-\lambda)\beta^2 &= \tilde{E}_2 \\ \lambda\alpha^3 + (1-\lambda)\beta^3 &= \tilde{E}_3 \\ 0 < \alpha, \beta, \lambda &< 1 \end{aligned} \quad . \qquad [29]$$

We first define and calculate the following two quantities:

$$\begin{aligned} A &:= \frac{\alpha+\beta}{2} = \frac{\tilde{E}_3 - \tilde{E}_1\tilde{E}_2}{2\left(\tilde{E}_2 - \tilde{E}_1^{\,2}\right)} \\ B &:= \frac{\alpha-\beta}{2} = \sqrt{\left(\frac{\tilde{E}_3 - \tilde{E}_1\tilde{E}_2}{2\left(\tilde{E}_2 - \tilde{E}_1^{\,2}\right)}\right)^2 - \frac{\tilde{E}_1\tilde{E}_3 - \tilde{E}_2^{\,2}}{\tilde{E}_2 - \tilde{E}_1^{\,2}}} \end{aligned} \quad . \qquad [30]$$

The parameters are afterward computed as:

$$\begin{aligned} \alpha &= A + B \\ \beta &= A - B \\ \lambda &= \frac{1}{2} + \frac{\tilde{E}_1 - A}{2B} \end{aligned} \quad . \qquad [31]$$

However, we still need to ensure that they are real and in the correct ranges. One can verify that these conditions are satisfied by enforcing the following constraints:

$$\begin{aligned} 0 < \tilde{E}_3 < \tilde{E}_2 < \tilde{E}_1 &< 1 \\ \tilde{E}_1^{\,2} &< \tilde{E}_2 \\ \tilde{E}_2^{\,2} &< \tilde{E}_1\tilde{E}_3 \\ \tilde{E}_3 - \tilde{E}_1\tilde{E}_2 &< \tilde{E}_2 - \tilde{E}_1^{\,2} + \tilde{E}_1\tilde{E}_3 - \tilde{E}_2^{\,2} \end{aligned} \quad . \qquad [32]$$

Thus, we can obtain the optimal values of α, β, and λ, by initially projecting $\tilde{E}_i$s onto the subspace defined by the above inequalities, and then computing the parameters. Note that such projection is usually necessary, because the bi-exponential model may not be fully accurate and the data may be noisy. Furthermore, using a small separating margin of $\delta = 0.01 \sim 0.1$ in the inequalities makes the ODFs more stable in practice.

4.3 Results and Discussions

4.3.1 Results from Single q-Shell

To validate our approach, we first show results using artificial data. We simulated fiber crossing by generating diffusion images from the sum of two exponentials,

$\tilde{E}(\hat{u}) = \left(e^{-\hat{u}^T D_1 \hat{u}} + e^{-\hat{u}^T D_2 \hat{u}}\right) \Big/ 2$, where D_1 is a diagonal matrix with diagonal entries (9, 2, 2), and D_2 is D_1 rotated about the *y*-axis by a varying angle. Assuming an ADC of 0.7×10^{-3} mm^2s^{-1} (the mean diffusivity in brain parenchyma), these diffusion values correspond to a b-value of 4800 s/mm^2. We tested the following techniques to reconstruct their corresponding quantities from 76 diffusion directions uniformly sampled on the sphere, in the eighth order SH basis: ODF approaches such as CSA-QBI (with no regularization), the original QBI, and the original QBI followed by LBS, $\left(1-\lambda\nabla_b^2\right)$, with λ=0.15 chosen to produce the optimal results; and also non-ODF approaches such as Constrained Spherical Deconvolution (CSD) (74), Mixture of Wisharts (MoW) (101) using nonnegative least squares (NNLS) with manually optimized parameters, and DOT (77) for two radii. The results are presented in Fig. 4.5. As can be seen, DOT with the larger radius, CSA-QBI, and MoW (respectively) resolve the crossing with the smallest angles. We also verified this using the dip test (102) on sample points generated from the reconstructed distributions. The dip test measures multimodality in a sample by the maximum difference, over all sample points, between the empirical distribution function, and the unimodal distribution function that minimizes that maximum difference. As can be observed in Fig. 4.6, the two modes of the distributions are distinguished in the same order as mentioned above. It should however be noted that DOT and MoW require manual adjustment of parameters to produce the best results, a process that is not so convenient for unknown real data, in contrast to CSA-QBI which is parameter free. In addition, the negative values seem to appear only in the four non-QBI methods in this particular experiment (although they can be corrected, they may indicate potential intrinsic problems with such computations). CSA-QBI (with no regularization) and DOT (for higher radius), however, were less stable to noise compared to the rest of the methods when we repeated the simulation with Rician noise. Nevertheless, it has been shown in (103) that the CSA-ODF produces more consistent generalized fractional anisotropy (GFA) than the original ODF does.

We also tested our method on four real HARDI datasets; first on the physical phantom in (104), which was constructed from excised rat spinal cords and designed to have crossing tracts (90 diffusion images at b=1300 s/mm^2, no regularization; see footnote 29 the for acquisition parameters), and then on a public human brain dataset (105) (200 diffusion images at b=3000 s/mm^2, regularization parameters $\delta_1 = \delta_2 = 0.001$). The ODFs were reconstructed in the fourth order SH using three QBI approaches: our proposed method, the original QBI, and the original QBI followed by LBS with λ=0.5 for the rat data and λ=0.8 for the brain data. Results are

superimposed on the fractional anisotropy (FA) map and presented in Fig. 4.7. (ODFs are shown as they are; no min-max normalization is used in any of the figures.) Our method (left) produces sharper and more accurate ODFs than the original QBI (middle). In addition, although sharpening (right) enhances the original QBI ODFs considerably in anisotropic tissue, it causes significant instability in isotropic regions (e.g. the background of the rat phantom and the cerebrospinal fluid (CSF) in the human brain data), in contrast to our technique which preserves isotropy fairly well (see (103) for more such validations). For the human brain dataset, we focus on the region of the centrum semiovale, where three major fiber bundles intersect: the internal capsule (IC)/corona radiata (CR), the radiations of the corpus callosum (CC), and the superior longitudinal fasciculus (SLF).

We performed additional experiments on a human brain HARDI dataset acquired at 7T with 256 diffusion directions. (For the fourth real dataset, see Sec. 4.3.2.) Similar results as mentioned above were obtained for b=3000 s/mm^2 with no regularization (Fig. 4.8, top row). We then compared the results with those obtained from DOT for three different radii, which are depicted in Fig. 4.8 (middle row). As can be seen, for the smallest radius r_0 (middle row, left), the fiber crossings are not well resolved. As the radius increases (center), the resolution of the crossings improves. Eventually for the largest r_0 (middle row, right), regions with fiber crossings are very well reconstructed. However, negative values (shown in dark red) start to appear in anisotropic regions. (These values can be projected to zero, yet without affecting the eventual fiber orientations.) In all cases, DOT achieves quite isotropic PDFs in the CSF.

Finally, we reconstructed the ODFs in a dataset acquired from the same subject at b=2000 s/mm². In contrast to the previous dataset where no regularization was needed, the noise in this one made the use of regularization necessary, particularly in the highly anisotropic regions (see Sec. 4.2.5). Figure 4.8 (bottom row) shows the corpus callosum in an axial view, where ODFs are reconstructed with the regularization parameters of $\delta_1 = \delta_2 = 0.001$ (bottom, left), $\delta_1 = \delta_2 = 0.01$ (bottom, center), and once without any regularization but using the L_1 error norm (bottom, right). As can be seen, increasing δ_1 and δ_2 more than 0.001 did not result in any significant improvement. However, using the L_1 error norm to compute the SH coefficients considerably reduced the number of negative values.

4.3.2 Results from Multiple q-Shells

To demonstrate the advantages of exploiting multiple q-shells in QBI, we first show experimental results on a synthetic example which consists of large diffusion values in two orthogonal directions. We synthesized diffusion images by sampling the sum of two exponentials, $E(\vec{q}) = \left(|\sin\phi|^{|q^2/2} + |\cos\phi|^{|q^2/2} \right) \Big/ 2$, on seven q-shells ($b = q^2 = 1,2,\ldots,7$) and in 76 directions, uniformly distributed on the sphere. Figure 4.9 illustrates the ODFs reconstructed from single q-shells for different b-values, and from three q-shells with both mono-exponential and bi-exponential models. For the data acquired at low b-values ($b = 1,2,3$), the proposed bi-exponential model using three q-shells is the only method correctly resolving the horizontal and vertical ODF peaks, corresponding to the strong ADC values in the directions $\phi = 0°, 90°, 180°, 270°$. It should be noted, however, that the drawback of such a more general model is its lesser robustness to noise (as low order models are often more robust).

We also tested our method on a 7T monkey brain HARDI dataset introduced in (106). The proposed method was used to reconstruct the ODFs from the three b-values of 1000, 2000, and 3000 s/mm², using both bi-exponential and mono-exponential methods and also from the single q-shells individually. Figure 4.10 depicts the results on a coronal slice through the centrum semiovale area, superimposed on the FA map. (For comparison, one of the sub-figures shows results by the original QBI.) Note how using the bi-exponential method allows for more clear recovery of certain fiber bundles, such as callosal radiations and corticospinal tract, and better resolution of crossing areas (see outlined regions in Fig. 4.10).

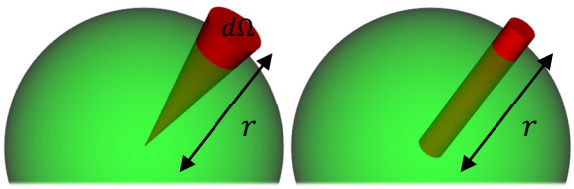


Fig. 4.1. Radial integration of the PDF, (left) in a cone of constant solid angle (i.e., the factor r^2 is considered), and (right) by linear projection (i.e., without the factor r^2 as done in the original QBI).

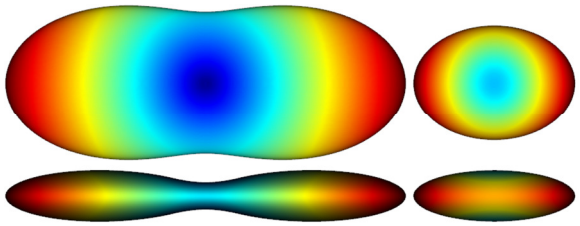


Fig. 4.2. DTI example of ODF reconstruction (with {10, 5, 1} as the diagonal entries of the tensor), shown from two view angles, (left) considering the factor r^2 (CSA) (right) without the factor r^2 and after normalization. Note how less sharp the latter is and how incompletely it represents the true structure of the ODF.

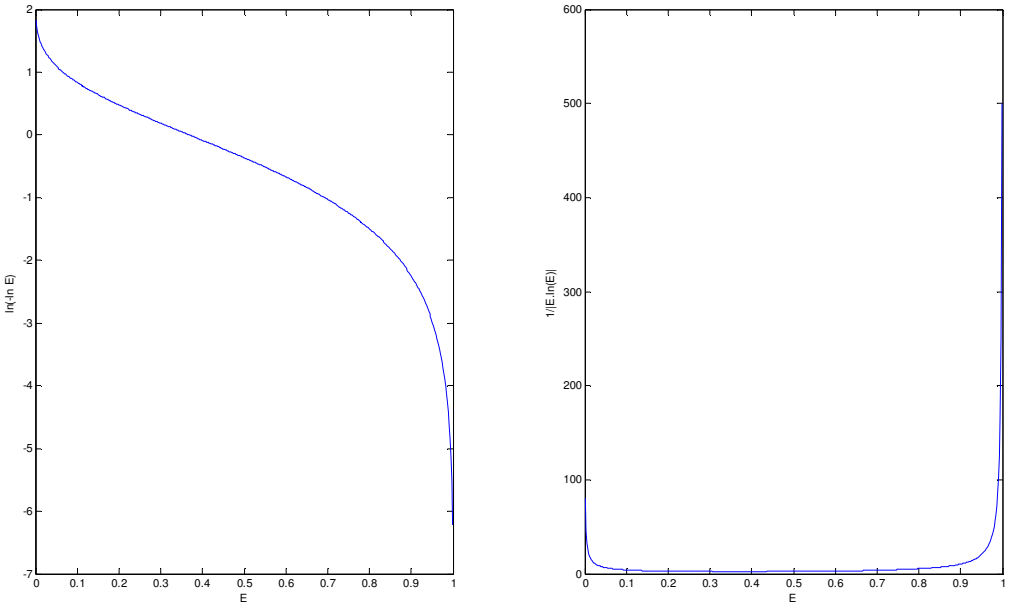


Fig. 4.3. Behavior of (left) ln(–ln *E*) and (right) the absolute value of its derivative with respect to *E*. Note how unstable they are for *E* close to 0 or 1.

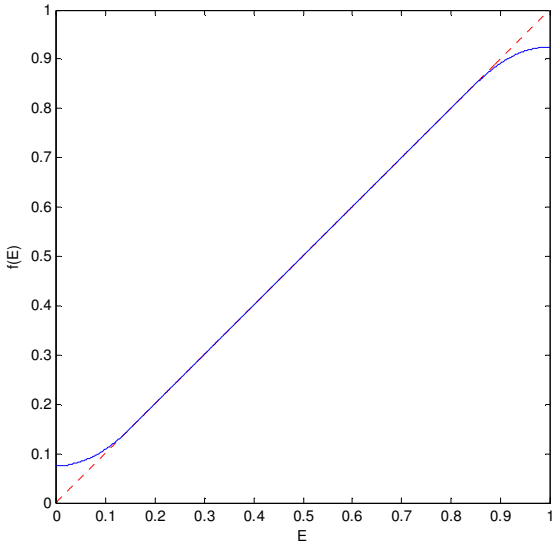


Fig. 4.4. The regularization function used for the diffusion signal to avoid the unstable regions (blue curve). The truncating margins are exaggerated for better visualization.

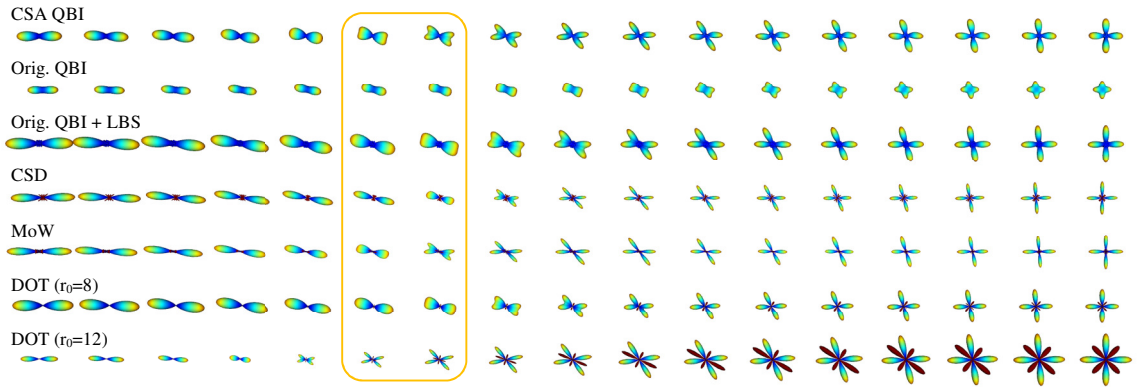


Fig. 4.5. Experimental results on synthetic data with fiber crossing, using: Proposed CSA QBI, original QBI after normalization, original QBI with Laplace-Beltrami sharpening (LBS), Constrained Spherical Deconvolution (CSD), Mixture of Wisharts (MoW) with manually optimized parameters, and Diffusion Orientation Transform (DOT) for two radii. The two columns in the box correspond to the crossing angles of 28.1° and 33.8°. Dark red represents negative values.

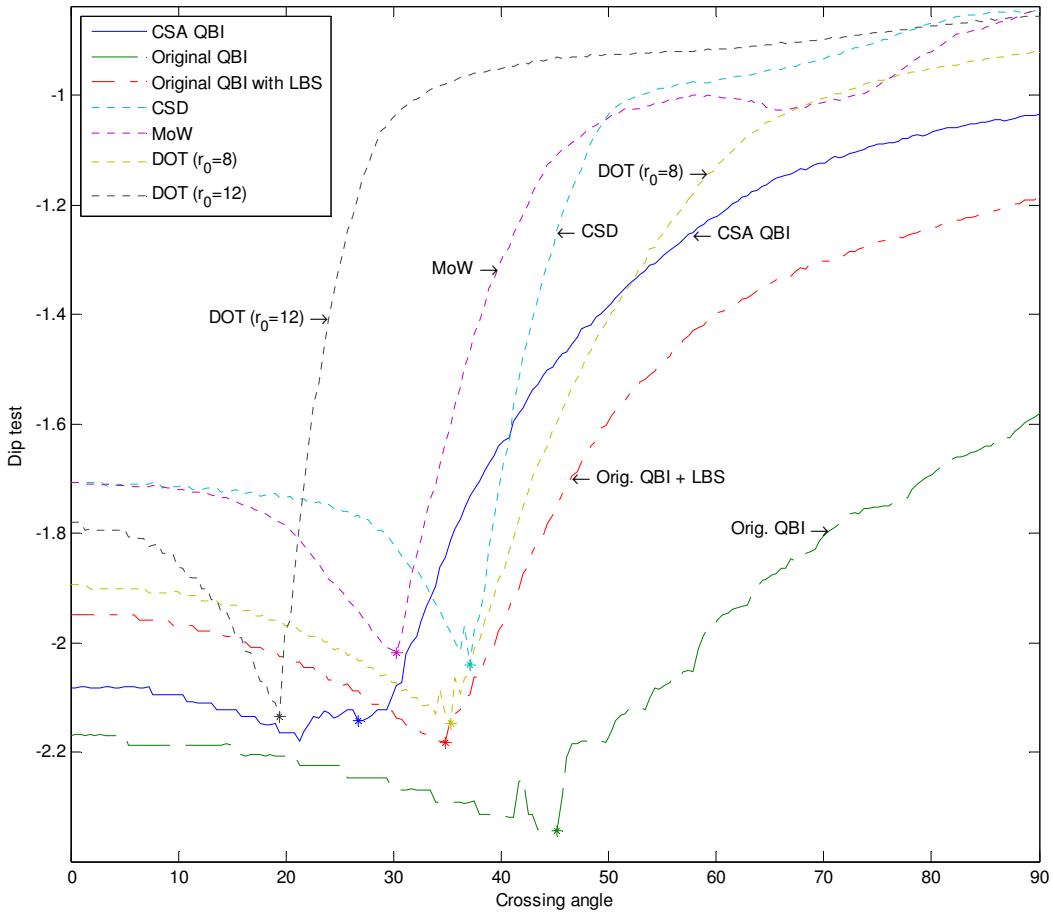


Fig. 4.6. Results of the dip test (a measure of multimodality) using the same distributions as shown in Fig. 4.5. The asterisk (*) on each curve indicates the minimum angle where the bimodality is detected. The *y*-axis is plotted on a logarithmic scale.

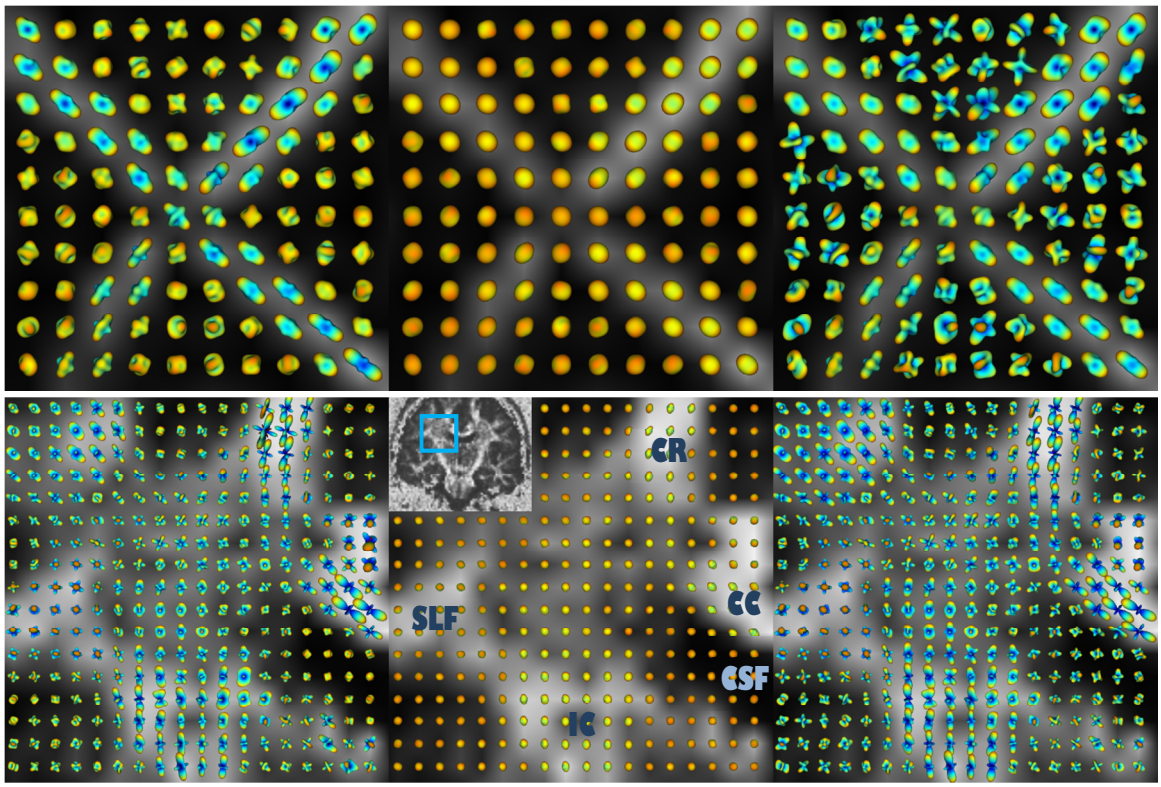


Fig. 4.7. Reconstructed ODFs from (top) rat spinal cord phantom and (bottom) human brain, shown on the FA map, using: (left) CSA QBI, (middle) original QBI after normalization, and (right) original QBI with LBS.

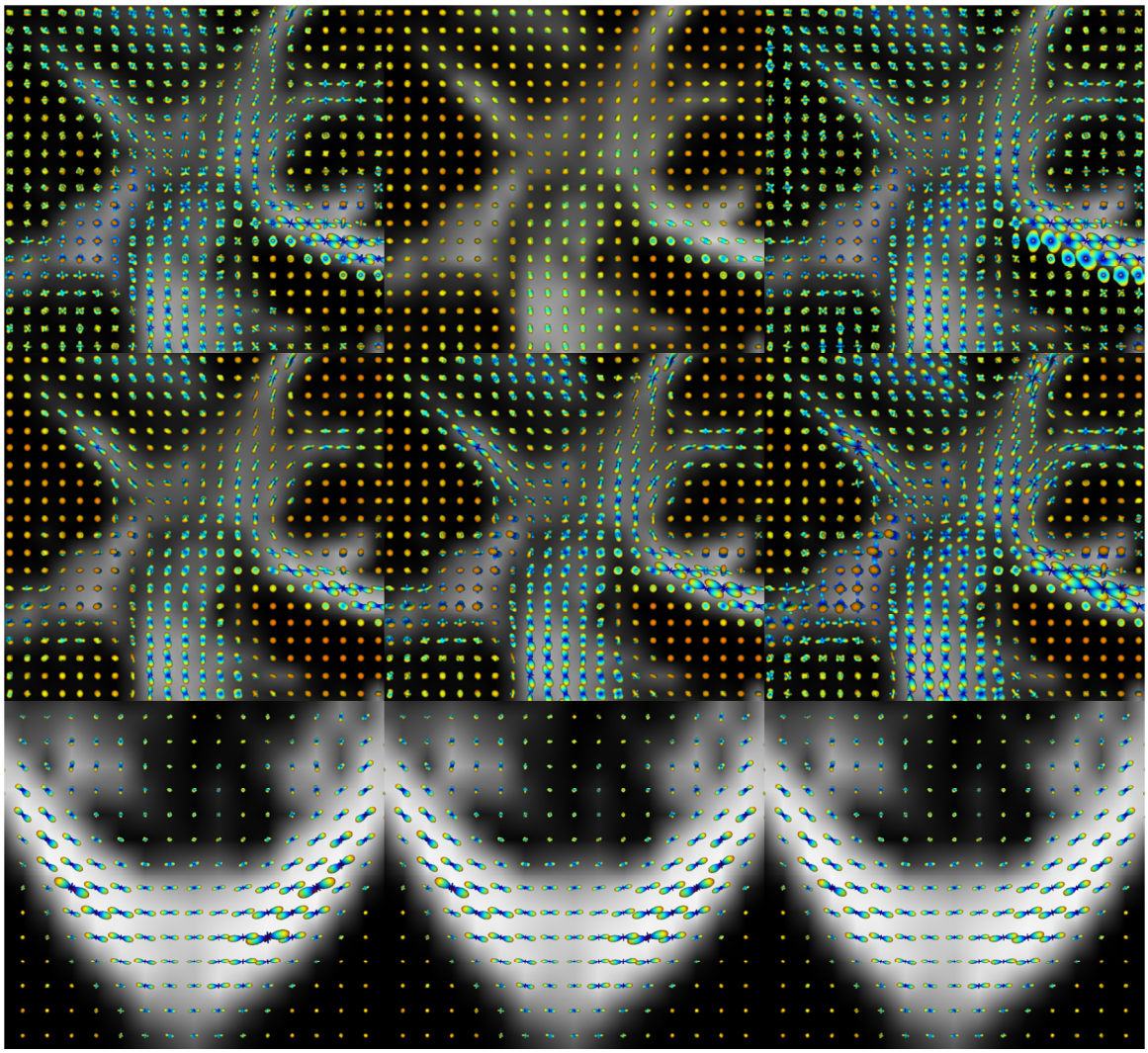


Fig. 4.8. (Top row): Reconstructed ODFs from 7T human brain data shown on the FA map, using: (left) CSA QBI, (middle) original QBI after normalization, and (right) original QBI with LBS. (Middle row): Results of DOT for different radii, ascending from left to right. (Bottom row): ODFs reconstructed using regularization parameters of (left) 0.001, (middle) 0.01, and (right) without regularization using the L_1 error norm. A singly refocused 2D single shot spin echo EPI sequence was used. Image parameters were: FOV: 192×192 mm^2 (matrix: 196×96) to yield a spatial resolution of 2×2×2 mm^3, TR/TE 4800/57 msec., acceleration factor (GRAPPA) of 2 and 6/8 partial Fourier were used along the phase encode direction. Diffusion-weighted images were acquired at three b-values of 1000, 2000 and 3000 s/mm^2 with 256 directions, along with 31 baseline images. EPI echo spacing was 0.57 msec. with a bandwidth of 2895 Hz/Px.

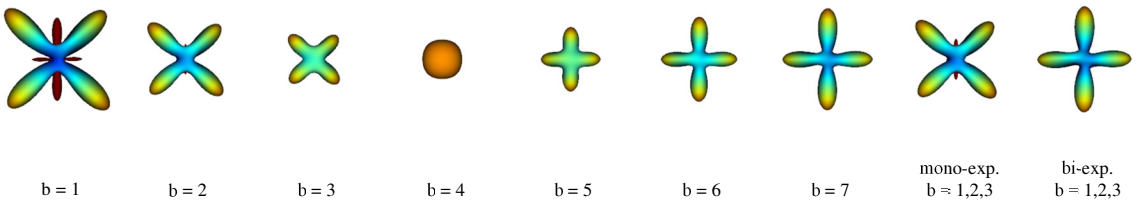


Fig. 4.9. Results of the ODF reconstruction on synthetic data. Note how the bi-exponential model correctly resolves the maxima of the ODF from low b-values. Dark red represents negative values. These values do not appear often in general, nonetheless, a possible formal approach to handle them can be found at (86).

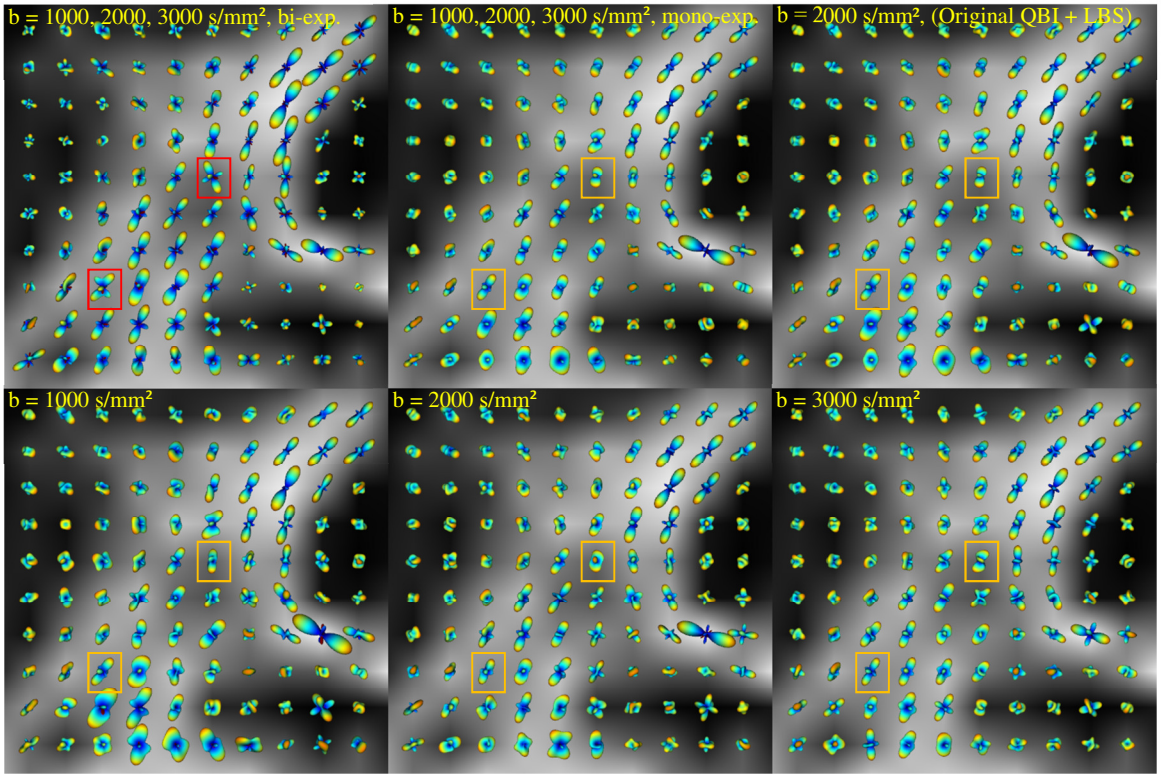


Fig. 4.10. Reconstructed ODFs from the real brain data, shown on the FA map. The bi-exponential model ODFs (top, left) have been scaled down 1.5 times for better comparison. All the ODFs except those in (top, right) are CSA ODFs. Note how the bi-exponential model for diffusion improves the resolution of fiber crossings, compared to the mono-exponential (constant ADC) model. An anesthetized Macaca mulatta monkey was scanned using a 7T MR scanner (Siemens) equipped with a head gradient coil (80mT/m G-maximum, 200mT/m/ms) with a diffusion weighted spin-echo EPI sequence. Diffusion images were acquired (`twice during the same session, and then averaged) over 100 directions uniformly distributed on the sphere. We used three b-values of 1000, 2000, and 3000 s/mm², TR/TE of 4600/65 ms, and a voxel size of 1×1×1 mm³.

5 Detecting Fiber Crossing

5.1 Introduction

Localizing brain structures and pathways associated with intellectual performance and cognition may clarify the networks involved, their development, and their dysfunction in disease. Here we introduce a measure of fiber crossing computed from high angular resolution diffusion images (HARDI), with the goal of better characterizing complex white matter microstructure. We aimed to detect white matter features associated with intellectual performance, previously missed due to limitations of the single tensor diffusion model (work presented in (107)).

5.2 Methods

High magnetic field (4T) diffusion-weighted (94 gradients, 11 non-sensitized) MR images were acquired from 293 right-handed healthy young adult twins and their siblings (23.9 ± 1.9 years old; 112 monozygotic subjects, 153 dizygotic, and 28 siblings). Verbal, performance, and full-scale intelligence quotients were also measured (VIQ: 114.7 ± 15.2; PIQ: 114.2 ± 11.9; FIQ: 111.7 ± 10.4). Images were corrected for motion and eddy current distortions. Each subject's average b_0 image was aligned to a common subject specific template using a nine-parameter affine transformation, and gradient directions were corrected accordingly. Orientation distribution functions (ODFs) in *constant solid angle* were computed as described in Chapter 4 from both the diffusion tensors (Eq. [4]) and q-balls (Eq. [13]). We also created a mean map of the q-ball ODFs, averaged across all subjects. We computed the *geodesic distance* (108,109) between the two ODFs (from the fitted tensor and q-ball) at each voxel for each subject. Both the tensor and q-ball ODF models can correctly reconstruct isotropic and single-fiber geometries. However, only the q-ball ODF can correctly resolve fiber crossings (see Fig. 5.1), so the mentioned tensor–q-ball ODF distance (TQOd) is a reasonable measure of fiber crossing. Each individual's fractional anisotropy (FA) maps were elastically registered to a common FA template including only high FA regions, to improve white matter registration. Individual deformation fields were then applied to the TQOd to identically align each subject's FA and TQOd maps. Using a linear random effects regression analysis to account for familial relations in the twin samples, voxelwise FA and TQOd maps were each regressed against each intelligence score.

5.3 Results and Discussions

Figures 5.2 and 5.3 show the mean q-ball ODF field overlaid on respectively the mean FA and the mean TQOd maps. The mean TQOd clearly identifies known crossings, e.g. where the corpus callosum, corona radiata and internal capsule all intersect. Known isotropic, single-fiber anisotropic, and fiber crossing regions were identifiable, respectively, as regions with low FA/low TQOd, high FA/low TQOd, and low FA/high TQOd. After correcting for multiple comparisons using the false discovery rate (FDR) method, performance IQ – which is thought to be related to neuronal processing speed – was associated with high anisotropy in several regions, as expected (110). TQOd correlations with PIQ were also significant in regions of low FA located in the caudate head, where partial voluming occurs between the internal capsule and striatum. Here, fiber crossing and the presence of multiple tissue types per voxel prevent tensors from fitting correctly, but q-ball reconstruction is more accurate. Figure 5.4 shows voxels that survived FDR correction, with uncorrected p-values of $p < 3.2 \times 10^{-5}$ for TQOd. This correlation highlights the importance of higher-order diffusion modeling to better localize fiber pathways relevant for cognition.

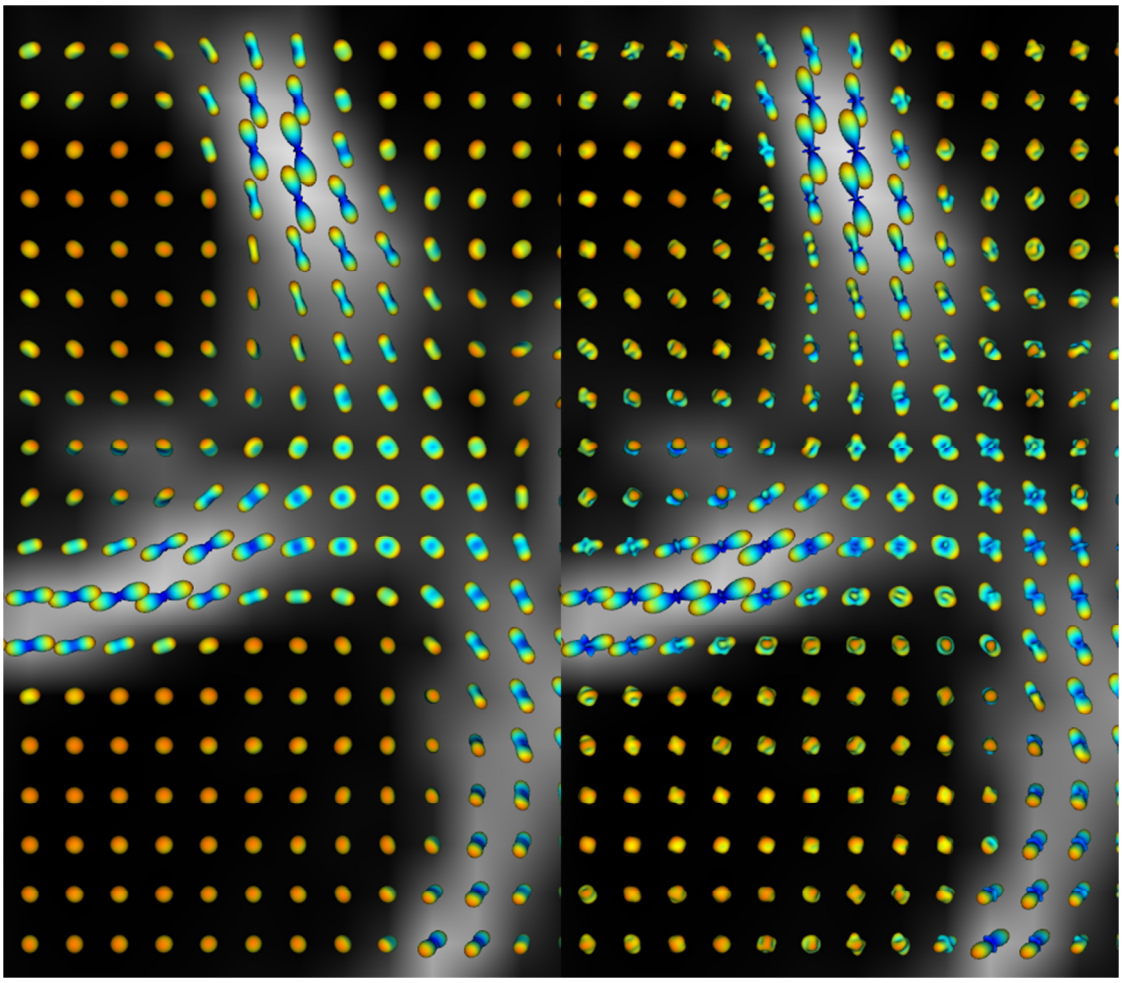


Fig. 5.1. ODFs computed from tensors (left) and q-balls (right). Note the difference in the crossing regions.

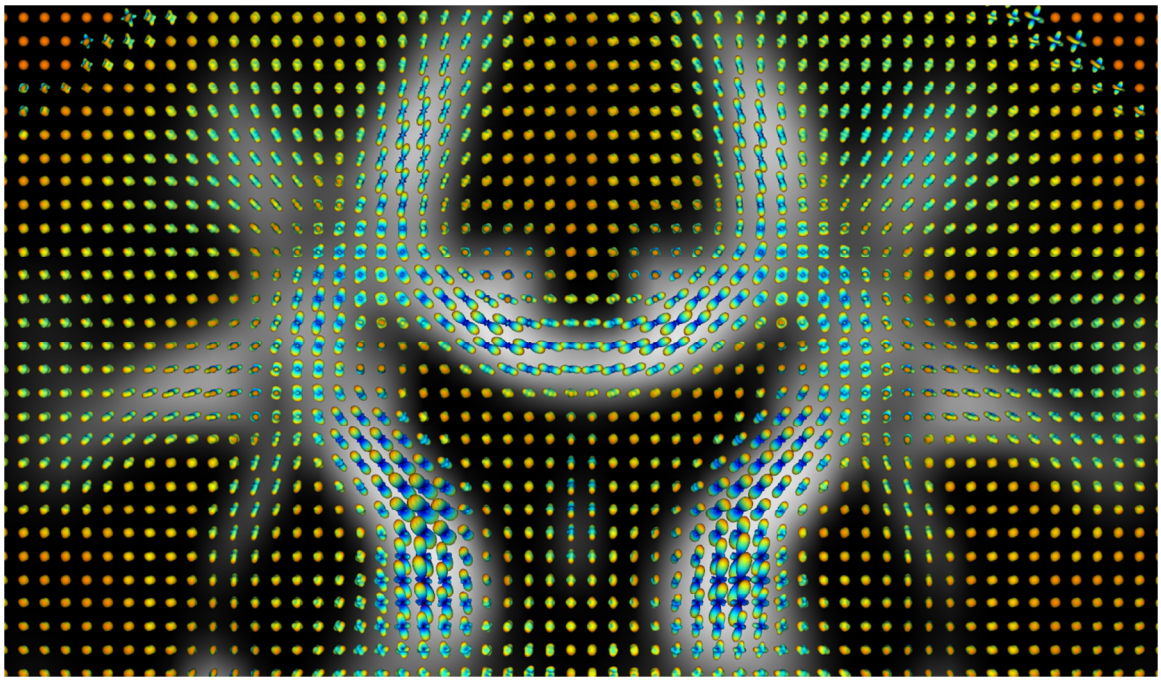


Fig. 5.2. Mean q-ball ODF field overlaid on the mean FA maps.

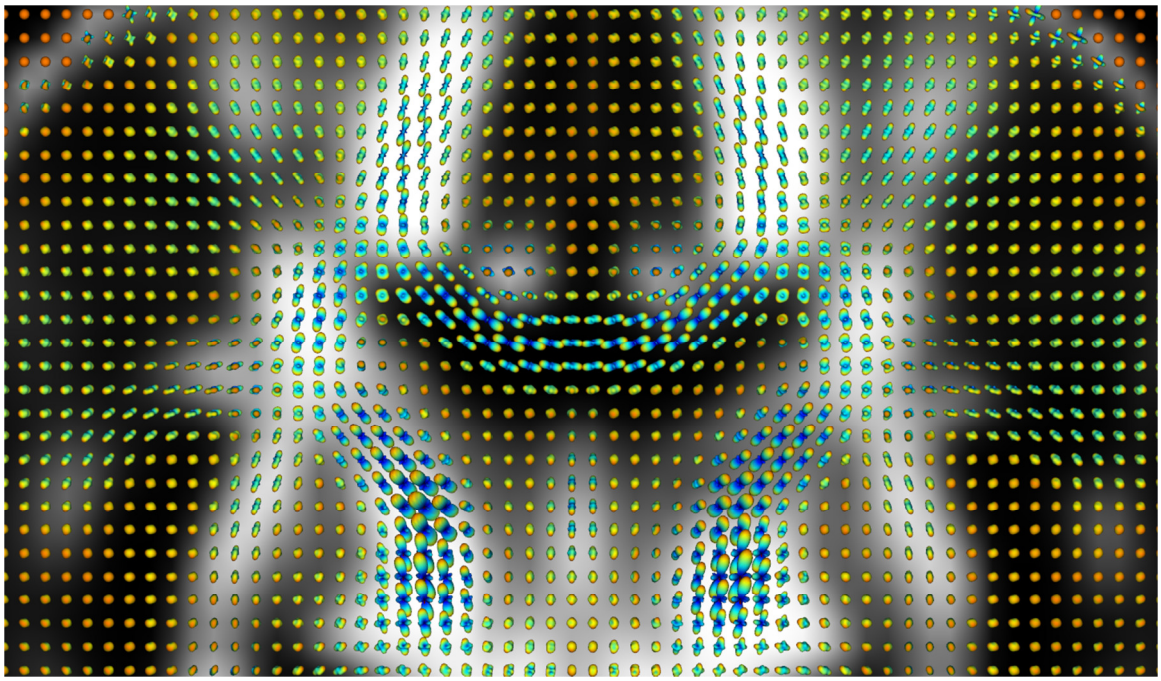


Fig. 5.3. Mean q-ball ODF field overlaid on the mean TQOd map.

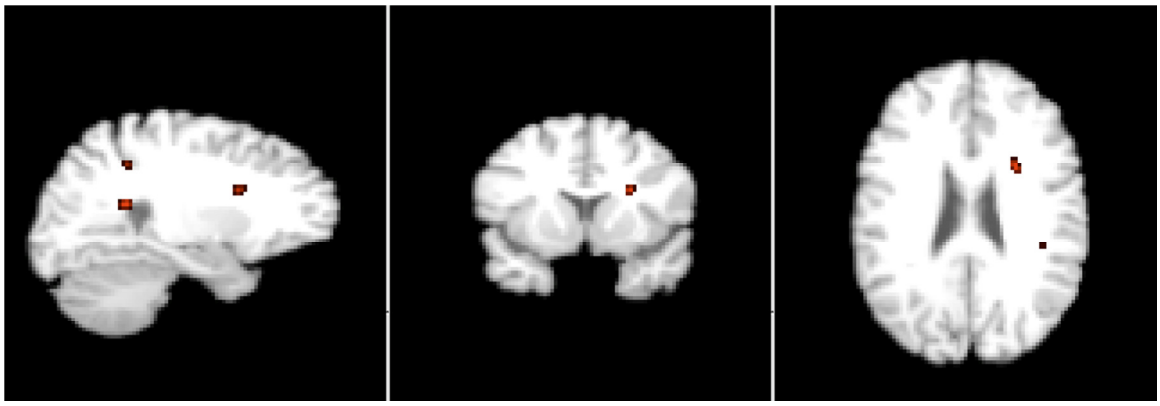


Fig. 5.4. Voxels that survived FDR correction, with uncorrected p-values of $p < 3.2 \times 10^{-5}$ for TQOd.

6 ODF Maxima Extraction

6.1 Introduction

Diffusion-weighted MRI significantly extends the scope of the information obtained from MRI, from being solely spatially dependent to being defined on the spatial-orientational domain. Fiber microstructure and orientation are inferred using this modality from the locally measured diffusion profile of water molecules. Diffusion tensor imaging (DTI) (91) effectively models the diffusion in single-fiber voxels as a Gaussian represented by its covariance tensor. As for more complex fiber architecture, q-ball imaging (QBI) (61), (63)– (65), (83) has been very successful in revealing intravoxel fiber orientations by introducing the orientation distribution function (ODF) as the probability of diffusion in a given direction (see Chapter 4).

Contrary to DTI, where the principal diffusion direction can be readily computed as the major eigenvector of the diffusion tensor, QBI provides a continuous spherical function which, although clearly illustrates the major diffusion orientations as its maxima, does not directly quantify them. Diffusion directions as vectors carry less information than the ODF itself does. On the other hand, their easy interpretation and their application in tractography, e.g., (111)– (113), make the ODF maxima extraction an important post-processing step still to be carefully addressed. The number of peaks can also be interpreted as a measure of white matter complexity. In addition, unlike mixture models that calculate fiber directions by describing the diffusion signal as the sum of finite discrete unidirectional components, ODF maxima are computed without any assumptions about the existence of such components.

Exhaustive search via finite difference method has been exploited in the literature as a straightforward approach to ODF maxima extraction (64,114). This generally requires a two-dimensional (2D) discretization of the unit sphere, resulting in computational complexity that grows quadratically with the desired resolution. Numerical optimization approaches such as gradient ascent (115), Newton-Raphson techniques (73), and Powell's method (76), have also been employed. These techniques require a guarantee of convergence and careful initialization to obtain all the maxima. Lastly, polynomial based approaches, (116)– (118), have been proposed to extract the maxima as a subset of the stationary points of the ODF. These methods exploit a transformation of the real and symmetric spherical harmonic (RSSH) basis (most efficient for ODF reconstruction (64)), to the constrained symmetric tensor or constrained homogenous polynomial bases, resulting in polynomial equations which are solved numerically.

In this paper, we propose a polynomial based approach to reduce the problem of ODF maxima extraction in the fourth order RSSH basis, from a 2D search on the sphere, to a one-dimensional (1D) one on an analytically-derived curve. Compared to the 2D problem, this approach significantly reduces the computational complexity of the search for the maxima of the ODF – or any antipodally symmetric spherical function – without compromising the precision. Contrary to (116)– (118), our method works directly in the RSSH basis and does not require the extra step of transforming the RSSH coefficients to other tensor-based bases. We suggest a discretization scheme for the 1D exhaustive search, and show experimental results on both artificial and human brain data.

We start Sec. 6.2 with a brief review of the RSSH basis, and continue by describing our mathematical derivation. Experimental results are presented in Sec. 6.3. This work was published in (119).

6.2 Methods

6.2.1 ODF in Real and Symmetric Spherical Harmonic Basis

In this work, we use the estimator derived in Chapter 4 to compute the ODF in constant solid angle (CSA). The spherical harmonic basis is commonly used for representing spherical functions such as the ODF, allowing for sampling in any desired direction. Orthonormal spherical harmonic functions are given by

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi}, \qquad [33]$$

where $P_l^m(\cdot)$ is the associated Legendre function, and θ and ϕ are standard spherical coordinates. The assumption of the ODF being real and antipodally symmetric, however, makes the use of the RSSH basis (64) more suitable. RSSH functions are indexed by a single parameter $j = l(l+1)/2 + m + 1$, corresponding to l_j and m_j, as follows (64):

$$Y_j = \begin{cases} (-1)^{m_j}\sqrt{2}\mathrm{Re}\left\{Y_{l_j}^{-m_j}\right\}, & -l_j \le m_j < 0 \\ Y_{l_j}^0, & m_j = 0 \\ \sqrt{2}\mathrm{Im}\left\{Y_{l_j}^{m_j}\right\}, & 0 < m_j \le l_j. \end{cases} \qquad [34]$$

The ODF can be computed in this basis as described in Sec. 4.2.4.

6.2.2 ODF Maxima Extraction

RSSH functions, being smooth, allow us to find all the local maxima of the ODF $\psi(\theta, \phi)$ as points satisfying the following properties (subscripts indicate partial derivatives):

$$\psi_\theta(\theta, \phi) = 0, \quad [35]$$

$$\psi_\phi(\theta, \phi) = 0, \quad [36]$$

$$\det\big(H(\theta, \phi)\big) \geq 0, \quad [37]$$

$$\mathrm{tr}\big(H(\theta, \phi)\big) \leq 0, \quad [38]$$

with the Hessian matrix $H(\theta, \phi)$ defined as

$$H(\theta, \phi) = \begin{pmatrix} \psi_{\theta\theta}(\theta, \phi) & \psi_{\theta\phi}(\theta, \phi) \\ \psi_{\theta\phi}(\theta, \phi) & \psi_{\phi\phi}(\theta, \phi) \end{pmatrix}. \quad [39]$$

Equations [35] and [36] guarantee that (θ, ϕ) is either an extremum or a saddle point of the ODF. Inequalities [37] and [39] filter out, respectively, the saddle points and the local minima (including possible negative lobes), leaving us only with the local maxima of the ODF. The above expressions can all be analytically computed for an ODF expressed in the RSSH basis. However, the main challenge is to find the points that simultaneously satisfy equations [35] and [36]. Once they are identified, applying inequalities [37] and [39] to filter out undesired points is trivial.

Iterative approaches (e.g., Newton method) may be applied to solve equations [35] and [36]. Yet, being quite sensitive to the initialization, they are not guaranteed to converge to all the maximum points. Alternatively, an exhaustive search will result in all the maxima with an accuracy determined by the discretization resolution. Nonetheless, with the ODF being a 2D manifold, the search space, and consequently the computational complexity of the algorithm, grows quadratically with the desired resolution.

We will next show how the fourth order RSSH basis makes it possible to confine the search to a 1D space, thereby creating an efficient method to extract the maxima.

6.2.3 Reducing the Dimension of the Search Space

Let us assume that the ODF has been approximated in the fourth order RSSH basis, as

$$\psi(\theta,\phi)=\sum_{j=1}^{15}a_jY_j(\theta,\phi). \qquad [40]$$

Combining equations [33], [34], and [40], while substituting the values of $P_l^m(\cos\theta)$ from Table 6.1 leads to

$$\begin{aligned}\psi(\theta,\phi)=\frac{a_1}{2\sqrt{\pi}}&+A(3\cos^2\theta-1)+B(\phi)\sin\theta\cos\theta+C(\phi)\sin^2\theta\\&+D(35\cos^4\theta-30\cos^2\theta+3)\\&+E(\phi)(7\cos^2\theta-3)\sin\theta\cos\theta+F(\phi)(7\cos^2\theta-1)\sin^2\theta\\&+G(\phi)\cos\theta\sin^3\theta+H(\phi)\sin^4\theta,\end{aligned} \qquad [41]$$

where $A=\sqrt{\frac{5}{16\pi}}a_4$, $B(\phi)=\sqrt{\frac{15}{8\pi}}(a_3\cos\phi-a_5\sin\phi)$, etc. (We drop the notation (ϕ) in the rest of this subsection.)

We now attempt to solve Eq. [36] by deriving Eq. [41] with respect to ϕ. We then divide it by $\sin\theta\cos^3\theta$ and rearrange it, while using the identity $\sec^2\theta=1+\tan^2\theta$, to obtain

$$\begin{aligned}(H+C-F)_\phi\tan^3\theta+(G+B-3E)_\phi\tan^2\theta+(6F+C)_\phi\tan\theta+(B+4E)_\phi\\=0.\end{aligned} \qquad [42]$$

Equation [42] is a cubic function of $\tan\theta$, and can be analytically solved, leading to a closed-form expression for $\theta(\phi)$.[16] Thus, for each given ϕ, we obtain one, two, or three different real values for θ which satisfy Eq. [36].

The curve characterized by the pair $(\theta(\phi),\phi)$ (Fig. 6.1(b)) is in fact our new 1D search space which contains all the ODF maxima as points satisfying equations [35], [37], and [39] (Fig. 6.1(c&d)). The number of these maxima does not need to be initially specified, since it is

[16] Each solution of $\tan\theta$ corresponds to a unique value of $\theta\in[0,\pi)$. Please note that this approach can also be applied in the RSSH basis of higher orders, with the difference that there may be no analytical solution for $\theta(\phi)$, and numerical methods may need to be applied.

automatically determined by the algorithm and depends on various factors, such as the number of real solutions to Eq. [42]. This is particularly important in practice, as different regions of the white matter naturally exhibit different complexity. The maxima can be found using a 1D exhaustive search (see Sec. 6.2.4), which is considerably faster than exploring the entire 2D manifold of the ODF.[17]

6.2.4 One-Dimensional Exhaustive Search

Here we detail the discretization scheme used to perform the aforementioned 1D exhaustive search for the maxima. We exploit the closed-form description of the curve $(\theta(\phi), \phi)$ provided by Eq. [42] and parameterize the curve with $\phi \in [0, 2\pi)$. To achieve a constant spatial resolution $\Delta s = \sqrt{\Delta\theta^2 + \sin^2\theta\, \Delta\phi^2}$, we need a variable step size $\Delta\phi$:

$$\Delta\phi = \frac{\Delta s}{\sqrt{\theta'^2(\phi) + \sin^2\theta(\phi)}} = \frac{1 + t^2(\phi)}{\sqrt{t'^2(\phi) + t^2(\phi) + t^4(\phi)}}\Delta s, \qquad [43]$$

which is rewritten as a function of $t(\phi) \coloneqq \tan\theta(\phi)$. For every ϕ, Eq. [42] results in one, two or three real values for $t(\phi)$, for each of which $t'(\phi)$ can be computed simply by deriving Eq. [42] with respect to ϕ, and substituting for ϕ and t. Therefore, at each step we choose $\Delta\phi$ to be the minimum of the three (or fewer) values obtained from Eq. [43].

Next, we keep all the *candidate* points satisfying inequalities [37], [39], and the following, which is a relaxation of Eq. [35],

$$|\psi_\theta(\theta, \phi)| < \alpha. \qquad [44]$$

We found an appropriate value of $\alpha = 0.02 \sim 0.03$ for the threshold. Note again that inequalities [37], [39], and [44] can all be computed analytically using equations [33], [34], [40], and Table 6.1. The ODF maxima are then computed as the mean directions corresponding to the *clusters* of points, created by processing all the candidate points, as follows: Each point is added to a previous cluster if its Euclidean distance to the representative (mean) point of that cluster is minimum among all other clusters and is smaller than a threshold (0.4 was used here). If no such

[17] Such 1D exhaustive searches can also be performed using tensor-based approaches (116)– (118).

cluster is found, a new cluster is created, and the algorithm goes on until all the candidate points are processed.

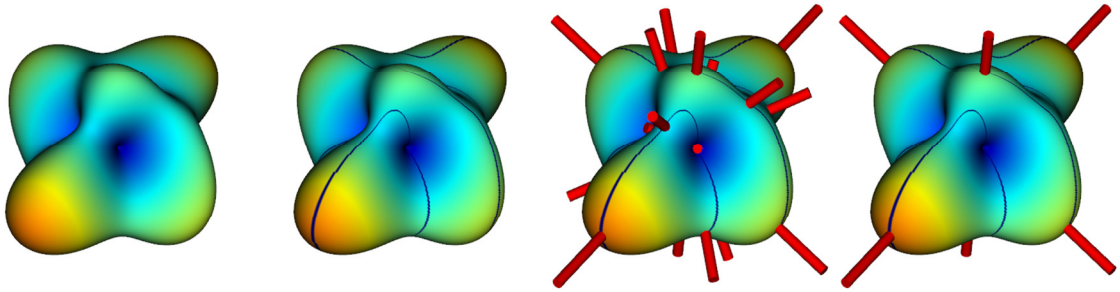
6.3 Results and Discussion

To validate our approach, we first show results on artificial data. We simulated fiber crossing by generating diffusion images from the sum of two exponentials, $S(\hat{u}) = \left(e^{-\hat{u}^T D_1 \hat{u}} + e^{-\hat{u}^T D_2 \hat{u}}\right)/2$, where D_1 is a diagonal matrix with diagonal entries (9, 2, 2), and D_2 is D_1 rotated about the z-axis by a varying angle. CSA-ODFs were reconstructed in the fourth order RSSH basis from 76 diffusion directions, uniformly sampled on the sphere. The maxima were then extracted using the proposed technique, and results are depicted in Fig. 6.2 (top). Increasing the angular precision to 0.5° revealed that multiple fiber orientations are resolved starting at the crossing angle of 37.5°. Choosing a spatial resolution of $\Delta s = 0.001$, required the evaluation of the ODF at 7.7×10^4 points, whereas a 2D search on the sphere with the same resolution would cost 1.6×10^7 operations. When we repeated the experiment by adding Rician noise with a signal-to-noise ratio (SNR) of 40 (Fig. 6.2, bottom), the minimum angle where crossing was detected increased to 48°. Such experiments are commonly employed to evaluate the robustness of the ODF reconstruction algorithm to noise.

We also tested our method on a popular public human brain dataset (105). CSA-ODFs were reconstructed in the fourth order RSSH basis from 200 diffusion images acquired at b=3000 s/mm². Figure 6.3 illustrates the ODFs with their extracted maxima superimposed on the fractional anisotropy (FA) map, in the region of the centrum semiovale, where three major fiber bundles intersect. To demonstrate the performance of the proposed technique, *all* the maxima are shown here, including those corresponding to slight variations in the ODF (for example due to noise). Major ODF peaks corresponding to fiber orientations may however be selected by placing a threshold on the ODF (64) or on its curvature (117).

Table 6.1. The associated Legendre functions required for the proposed algorithm.

Function		Expression
$P_0^0(\cos\theta)$	=	1
$P_2^0(\cos\theta)$	=	$(1/2)(3\cos^2\theta - 1)$
$P_2^1(\cos\theta)$	=	$-3\cos\theta\sin\theta$
$P_2^2(\cos\theta)$	=	$3\sin^2\theta$
$P_4^0(\cos\theta)$	=	$(1/8)(35\cos^4\theta - 30\cos^2\theta + 3)$
$P_4^1(\cos\theta)$	=	$-(5/2)(7\cos^2\theta - 3)\cos\theta\sin\theta$
$P_4^2(\cos\theta)$	=	$(15/2)(7\cos^2\theta - 1)\sin^2\theta$
$P_4^3(\cos\theta)$	=	$-105\cos\theta\sin^3\theta$
$P_4^4(\cos\theta)$	=	$105\sin^4\theta$



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

Fig. 6.1. (a) Reconstructed ODF. (b) Analytically defined 1D space is searched. (c) All the extrema and saddle points are identified. (d) ODF maxima are extracted.

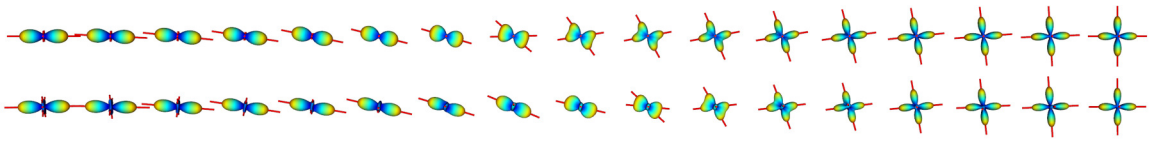


Fig. 6.2. Extracted maxima from synthetic ODFs with fiber crossing, in noise-free case (top), and with SNR=40 (bottom).

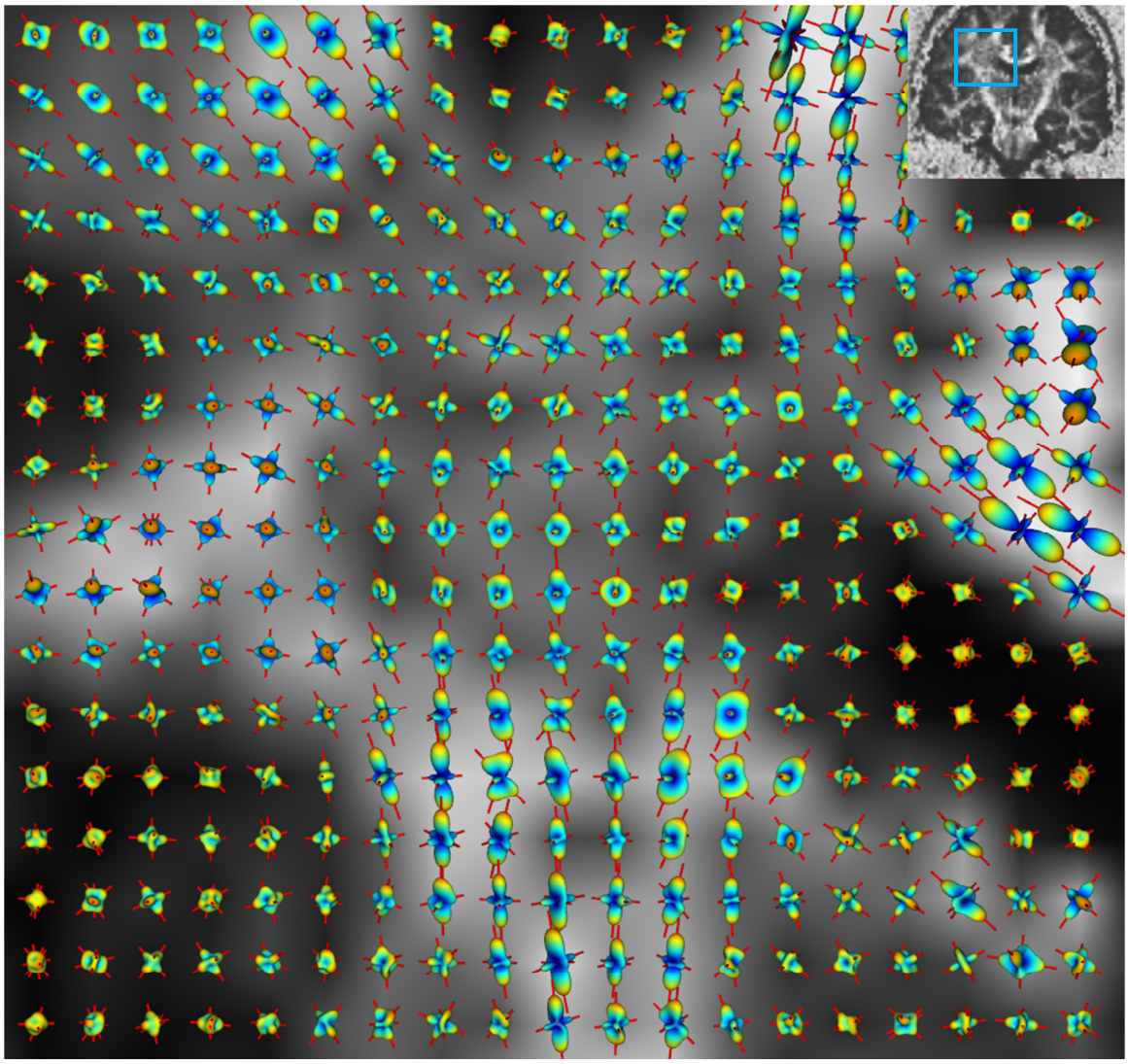


Fig. 6.3. Experimental results of the maxima extraction on human brain data, superimposed on the FA map.

7 Hough-Transform Tractography

7.1 Introduction

Understanding the connectivity between different areas of the brain is essential in studying brain function and development. Diffusion-weighted magnetic resonance imaging (DWI) provides, through tractography, a unique *in-vivo* quantitative measurement of the brain's anatomical connectivity. In addition to its benefits in neurosurgical planning, DWI tractography has considerable clinical importance by noninvasively quantifying changes in the white matter connectivity at different stages of diseases or development. Moreover, it can be used to segment fiber bundles of the central nervous system, or in tract-based statistical analysis of scalars such as the fractional anisotropy (FA). Performing tractography in multiple subjects is invaluable for population studies and creating fiber bundle atlases.

DWI provides local information about the fiber orientation by measuring the diffusion of the tissue water, *in vivo*, assuming a high correlation between the fiber and diffusion orientations. However, there is no unique solution as to how to integrate these voxel-scale local orientations to infer global connectivity. Early fiber tractography algorithms, known as streamline methods, are based on following the principal diffusion orientation (111,112,120–122). Despite their simplicity, these methods are prone to cumulative errors caused by noise, partial volume effects, and discrete integration, and have difficulty in distinguishing fiber crossing and kissing mostly due to the fact that the entire diffusion information is not globally used and integrated. This led to the development of other successful approaches, including probabilistic techniques (113,123–128), global techniques based on front propagation (104,129–133), simulation of the diffusion process or fluid flow (134–139), DWI geodesic computations (140–143), graph theoretical techniques (144–146), spin glass models (147,148), and Gibbs tracking (149). Generally speaking, for virtually every tractography method, a particular putative *subset* of all possible curves is implicitly considered from which the resulting tracts are chosen according to some criteria, which are different depending on the particular selection strategy. The closer the subset is to the universal set of curves, the more accurate we expect the results to be. For a recent thorough discussion on different tractography techniques, see (150).

Prior approaches for multi-subject tractography are typically based on the post processing of tractography results from individual subjects (151–157). These methods generally require aligning the tracts and mapping them into a common fiber coordinate system, which is

challenging due to the large number of high-dimensional fiber trajectories per subject and the lack of clearly defined criteria for aligning curves and particularly tracts.

In this work, we present a global probabilistic approach inspired by the voting procedure provided by the popular *Hough transform* (158,159). Our proposed tractography algorithm essentially tests candidate 3D curves in the volume, assigning a score to each of them, and then returning the curves with the highest scores as the potential anatomical connections. The score is accordingly derived from the DWI data. Being an exhaustive search, this proposed algorithm avoids entrapment in local minima within the discretization resolution of the parameter space.[18] Furthermore, the specific definition of the candidate tract score has the desired effect of attenuating the noise through the integration of the real-valued local votes derived from the diffusion data. We also introduce a simultaneous *multi-subject* tractography technique which takes as input a single representative volume – where the HARDI data from all the (registered) subjects are non-linearly integrated – and generates population-representative tracts. The multi-subject tractography algorithm is run only once, and no tract alignment is necessary. We present experimental results on HARDI volumes such as a simulated phantom, a biological phantom dataset acquired at 1.5T, a monkey brain dataset acquired at 7T, and a number of human brain datasets acquired at 4T and 7T.

In Sec. 7.2 we present the proposed algorithm in detail. Experimental results are presented in Sec. 7.3, and Sec. 8.3 contains a review of the contributions. Additional implementation details are provided in Appendix 9.3. We published this work in (90,160,161), and later used it on the same population as introduced in Sec. 5.2, for human connectome (162), and fiber clustering (163).

7.2 Methods

We first randomly generate a sufficiently high number of initial seed points inside a brain mask or a region of interest. From each initial point, we consider as many passing curves as desired, based on the expected resolution and available computational resources (Fig. 7.1, left). A score is computed for each curve, and the one(s) with the maximum score is (are) then chosen as the best

[18] Please note that many prior approaches may as well be modified to perform efficient exhaustive searches.

curve(s) representing the fiber bundle passing through that seed point (Fig. 7.1, middle & right).[19] This process is detailed in the following subsections.

7.2.1 Curve Parameterization

We parameterize the 3D curves by the arc length s, with $s = 0$ corresponding to the seed point (Fig. 7.2). The unit tangent vector of the curve is identified at each point by standard polar coordinates $\theta(s)$ and $\phi(s)$:

$$\hat{t}(s) = \begin{pmatrix} \sin\theta(s)\cos\phi(s) \\ \sin\theta(s)\sin\phi(s) \\ \cos\theta(s) \end{pmatrix}. \quad [45]$$

In our proposed model, we consider simple polynomial approximations of these two angles with respect to the arc length:

$$\theta(s) = \sum_{k=0}^{N} a_k s^k, \quad [46]$$

$$\phi(s) = \sum_{k=0}^{N} b_k s^k, \quad [47]$$

where N is the polynomial order (different orders for θ and ϕ can be considered if desired). In addition, two extra parameters L_- and L_+ determine the partial (Euclidean) lengths of the curve on each side of the seed point (Fig. 7.2), with $0 \leq L_{-,+} \leq L_{max}$ where the constant L_{max} is chosen as the maximum expected curve (fiber) length, essentially about the largest dimension of the volume. The Hough-inspired process will be used to select the best possible coefficients a_k, b_k, L_-, and L_+, based on the available diffusion data. Each curve initiated from the seed point $\vec{x}_0$ is then represented using $d = 2N + 4$ unique parameters, $\{a_0, \dots, a_N, b_0, \dots, b_N, L_-, L_+\}$, and explicitly computed by integrating the tangent vector:

$$\vec{x}(s) = \vec{x}_0 + \int_0^s \hat{t}(s')ds', \quad s \in [-L_-, L_+]. \quad [48]$$

[19] As customary in probabilistic techniques, several candidate curves may as well be selected per seed point, each carrying a score.

7.2.2 Fiber score computation

A score, intended to estimate the log-probability of the existence of a fiber, is assigned to each possible 3D curve passing through a seed point $\vec{x}_0$. In this work, the score is defined as

$$\begin{aligned} S_{\vec{x}_0}(a_0, \dots, a_N, b_0, \dots, b_N, L_-, L_+) &\coloneqq \int_{-L_-}^{L_+} \left(\log P\left(\vec{x}(s), \hat{t}(s)\right) + \lambda\right) ds \\ &= \int_{-L_-}^{L_+} \log P\left(\vec{x}(s), \hat{t}(s)\right) ds + \lambda(L_- + L_+). \end{aligned} \quad [49]$$

The expression $P(\vec{x}, \hat{t})d\Omega$ represents the probability for the point $\vec{x}$ to be located inside a fiber bundle passing in the direction $\hat{t}$ through the infinitesimal solid angle $d\Omega$.[20] The constant λ, which is used to compensate for the absence of $\log d\Omega$ in the integral, can also be interpreted as a prior on the length of the fiber bundles, as choosing a larger λ favors the score of longer curves.[21] By enforcing the curves to remain in the brain mask, and choosing a value of λ which is large enough, the curves which are oriented towards the cortex are expected to actually reach the gray matter region (e.g., see Fig. 7.8).

$P(\vec{x}, \hat{t})$ can be computed using the conditional probability formula as,

$$P(\vec{x}, \hat{t}) = P(\vec{x})P(\hat{t}|\vec{x}). \quad [50]$$

The prior probability of the existence of a fiber at the point $\vec{x}$, $P(\vec{x})$, is considered to be equal to either the fractional anisotropy (FA) or generalized fractional anisotropy (GFA) inside the brain tissue, and zero outside the brain mask and inside the cerebrospinal fluid. This comes from the assumption that the more anisotropic a region is, the more likely a fiber bundle may be passing through that region. In addition, as long as no further constraints or selections are provided by the user, the initial seed points are chosen randomly with a spatial probability distribution proportional to $P(\vec{x})$.[22] Other choices for $P(\vec{x})$, such as the white matter complexity introduced in (164), are also possible.

---

[20] The curve score is defined by assuming that adjacent voxels are independent, making it only an approximation of the true fiber log-probability. The incorporation of spatial coherence and continuity is the subject of future research.

[21] Not including λ may cause the zero-length curve ($L_- = L_+ = 0$) to gain the maximum score, due to the potentially negative values of the logarithm.

[22] A uniform distribution for the seed points would in fact force the algorithm to pick as many curves in the lesser anisotropic regions of the brain as in the highly anisotropic fiber bundles, sometimes producing unrealistic results.

Next, assuming that a fiber is actually passing through the point $\vec{x}$, the probability for it to be in the direction $\hat{t}$, i.e. $P(\hat{t}|\vec{x})$, is derived from the orientation distribution function (ODF) at each voxel in the volume. Computed from various DWI modalities, the diffusion ODF is defined as,

$$ODF(\hat{u}) \coloneqq \int_0^\infty PDF(r\hat{u})r^2 dr, \qquad [51]$$

which is the integration of the $PDF(\vec{r})$ (the spatial probability density function of the diffusion of water after a certain amount of time), in a cone of constant solid angle (CSA) in the direction of the unit vector $\hat{u}$. In Diffusion Spectrum Imaging (DSI) (59), $PDF(\vec{r})$ is available on a discrete Cartesian grid, and therefore the ODF is directly computed from the above formula. In the case of the Diffusion Tensor Imaging (DTI) (91), the ODF is computed by integrating the 3D normal distribution, as in Eq. [4].

In this work, however, we use Q-ball Imaging (QBI) (61), which is a popular HARDI reconstruction method proven successful in resolving multiple intravoxel fiber orientations. The original ODF expression in QBI does not include the Jacobian factor r^2, creating the need for post-processing such as artificial sharpening. Here we use the normalized and dimensionless ODF estimator in QBI, derived in Chapter 4, which by considering the factor r^2 computes the CSA-ODF (Eq. [13]). This ODF reconstruction scheme has been shown to outperform the original QBI by improving the resolution of multiple fiber orientations (Sec. 4.3), and producing more stable and consistent GFA (103). To allow sampling in any desired direction $\hat{t}$, the ODFs were approximated in the real and symmetric modified spherical harmonic basis, following the method proposed by Descoteaux et al. (64) for the original QBI, and subsequently adapted in Sec. 4.2.4 for the CSA-QBI.

Putting all this together, and using for instance the FA as $P(\vec{x})$, the score in Eq. [49] thus becomes,

$$S_{\vec{x}_0}(a_0, \dots, a_N, b_0, \dots, b_N, L_-, L_+) \coloneqq \int_{-L_-}^{L_+} \left(\log\left[ODF_{\vec{x}(s)}\left(\hat{t}(s)\right)FA\left(\vec{x}(s)\right)\right] + \lambda\right)ds, \qquad [52]$$

where $ODF_{\vec{x}(s)}\left(\hat{t}(s)\right)$ stands for the ODF at the 3D position $\vec{x}(s)$ evaluated in the direction $\hat{t}(s)$, with $\vec{x}(s)$ and $\hat{t}(s)$ represented via the polynomials as specified in equations [45–48]. The score integral in equations [49] and [52] has the additional nice effect of attenuating the additive noise in the data through summation.

7.2.3 Hough Transform

As discussed in Sec. 7.2.1, every curve starting from a particular seed point is presented as a point in a d-dimensional space, with $d = 2N + 4$ being the number of necessary parameters. In theory, we would like to find all possible curves which pass through the seed point while computing their scores, to eventually choose the one(s) with the highest score(s) as the potential fiber tract(s) passing through the seed point.[23] However, we can only perform such an exhaustive search within a finite resolution, by discretizing $\mathbb{R}^d$ and assigning discrete values to the curve parameters within some predefined limits.[24] The resulting d-dimensional array of curve scores is often called the *Hough transform* (158,159) of the data with respect to the curves passing through the chosen seed point. This can be seen as a voting process where the voxels cast real-valued votes for the curves. The overall vote is the integrand of the score integral (Eq. [49] or [52]) if a curve passes through a voxel, and zero otherwise.

The proposed method avoids entrapment in local minima by performing an exhaustive search in the (discretized) high-dimensional space of the curves. Nevertheless, the discretization resolution of the parameter space causes the algorithm to obtain an approximation of the true global optimum (which is improved by increasing the resolution as desired). To alleviate this issue, we choose the best curve in a multi-resolution approach: once the point (in the parameter space $\mathbb{R}^d$) corresponding to the curve with the highest score is found in one resolution level, the neighborhood (in the parameter space) of that point is discretized again with a higher resolution and the search is continued at the next level.[25] We have performed our experiments using three levels of resolution.

This concludes the description of the proposed technique for a single dataset. We now show how this can be efficiently extended to multiple datasets or subjects.

7.2.4 Extension to Multiple Subjects

Here we extend our Hough transform-based global approach to obtain average representative tracts from multiple subjects or datasets. We perform this by first registering the HARDI volumes, using either linear transformation or more sophisticated algorithms such as (165), and

[23] As customary in probabilistic techniques, several candidate curves may be selected per seed point, each carrying a score. In our experiments, however, we used a winner-take-all approach and select/visualize only the maximum-score curve.
[24] This is the standard procedure in the Hough transform, where the accumulator is discretized.
[25] Note that this multi-scale approach is performed to discretize the Hough transform (the parameter space), and not the spatial domain itself.

then running the algorithm described above on a single equivalent volume composed of the voxel-wise mean ODF and mean FA across all the subjects.[26] We may use either the arithmetic or the geometric mean, however, the linearity of the curve score (Eq. [52]) with respect to the logarithms of the ODF and FA makes the use of the geometric mean more appealing (since the arithmetic mean of the logarithms of the ODF and FA values equals the logarithm of their geometric mean). Hence, we reconstruct the effective ODF and FA for each voxel by computing the geometric mean of their values across the subjects,

$$ODF_{\vec{x}}^{eq}(\hat{t}) \coloneqq \left(\prod_{i=1}^{M} ODF_{\vec{x}}^{i}(\hat{t})\right)^{\frac{1}{M}}, \quad [53]$$

$$FA^{eq}(\vec{x}) \coloneqq \left(\prod_{i=1}^{M} FA^{i}(\vec{x})\right)^{\frac{1}{M}}, \quad [54]$$

where the superscripts i and eq indicate respectively the i^{th} subject (out of a total of M) and the *equivalent* subject. We eventually use the equivalent ODF and FA volumes in the single-subject tractography algorithm, thereby running it only once for all the subjects and avoiding complications due to curve (tracts) registration.

7.3 Experimental Results and Discussion

7.3.1 Results for Single Subjects

We tested our method on various HARDI datasets, also using each of them to explain a different aspect of the proposed algorithm. The FA and the CSA-ODFs of each dataset were computed as explained in Sec. 7.2.2. The initial seed points were chosen randomly with a spatial probability distribution proportional to the FA, except for the simulated and the monkey brain datasets where the distribution was uniform.

To validate our approach, we first show results on artificial data and compare our approach to one deterministic and another probabilistic streamline method, both implemented in the Camino Diffusion MRI toolkit (166) and exploiting q-ball ODFs.[27] Using simulated diffusion data, we created a phantom containing two crossing fiber bundles. As Fig. 7.3 demonstrates, this phantom

[26] This could be interpreted as multiple votes per voxel, cast by each corresponding voxel in each volume.
[27] We chose to compare with Camino (and not e.g. FSL) to have a fair comparison between ODF-based methods.

has various challenging fiber configurations, including fanning, crossing, and merging.[28] The phantom was made as a 50×50×1 slice; however, to simulate the partial volume effects, we initially created it ten times larger, then downsampled it after applying a low-pass filter. Figure 7.4 shows tractography results using the proposed algorithm (with uniform $P(x)$) (left), the deterministic streamline method (middle), and the Probabilistic Index of Connectivity (PICo) technique (167) (right). We performed tractography from randomly generated seed points with uniform spatial distribution, once for 500 points in the phantom mask (λ = 0.7; Fig. 7.4, two top rows), and a second time for 80 points in a smaller region identified in Fig. 7.3 (bottom, right) as red (λ = 1.8; Fig. 7.4, two bottom rows). We tested both the noise-free case (rows 1 & 3), and with the signal-to-noise ratio of 0 dB (rows 2 & 4). The streamline methods follow the principal diffusion direction (PDD), which may be misleading in the crossing regions, as the two directions might be mingled in the ODFs (or tensors), resulting in a false PDD different from the original ones. The global nature of our algorithm, however, alleviates this problem. For instance, as can be observed in the two top rows of Fig. 7.4, our method clearly reveals the crossing in the intersection region, resulting in curves going in the directions of the two perpendicular bundles, as opposed to the streamline methods, where curves seem to be going in an "average" direction not corresponding to any of the two bundles. In the case of the smaller seed-point region (two bottom rows), this results in the complete miss of one of the branches by the two streamline methods, in addition to more susceptibility to noise.

Next, we used the biological phantom in (104), constructed from excised rat spinal cords and designed to have crossing tracts (90 diffusion images at b = 1300 s/mm², 1.5T).[29] We computed the tracts from 200 seed points, using three different bias values of λ = 2.0, 2.5, and 3.0 (see Sec. 7.2.2), and polynomials of order $N = 3$, resulting in a total number of $d = 10$ parameters to represent the candidate 3D curves initiated from each seed point. Figure 7.5 (top, left) shows the ODFs superimposed on the FA map, and the rest of the subfigures show the tractography results using different values for λ. Increasing λ results in longer curves being selected. The color and the opacity of each tract (in all the figures) increase with the score, from transparent blue to opaque

[28] Matlab codes to regenerate the phantom can be found at: netfiles.umn.edu/users/iman/www/Synth.zip

[29] Scanning parameters, quoting from (104): The cords were scanned 1 h after the surgeries with a Siemens 1.5T Sonata MR scanner (Siemens Medical Systems, Erlangen, Germany) using a knee coil. A single-shot spin-echo echo planar sequence with twice-refocused balanced gradients, designed for minimization of eddy current artifacts, was used. For diffusion tensor reconstruction, four coregistered datasets were acquired, consisting of 90 diffusion weighted images with isotropically spaced diffusion weighting directions (b = 1300 s/mm², TR = 8s, TE = 110 ms, 2.5 mm isotropic voxels, 40 slices), as well as 10 images with b = 0 s/mm² and otherwise identical imaging parameters.

red. Given the simple structure of this phantom dataset, experiments on more complex real datasets were required to test our algorithm; these are described next.

We performed additional experiments on a human brain HARDI dataset acquired at 7T. A single refocused 2D single shot spin echo EPI sequence was used. Image parameters were: FOV: 192×192 mm^2 (matrix: 196×96) to yield a spatial resolution of 2×2×2 mm^3, TR/TE 4800/57 msec., acceleration factor (GRAPPA) of 2 and 6/8 partial Fourier were used along the phase encode direction. Diffusion-weighted images were acquired at b = 3000 s/mm^2 with 256 directions, along with 31 baseline images. EPI echo spacing was 0.57 msec. with a bandwidth of 2895 Hz/Pixel. Tracts were computed from 1500 seed points in two experiments, using polynomials of orders $N = 2$ and $N = 3$ (Fig. 7.6). High scoring curves are concentrated in major fiber bundles such as corona radiata, corpus callosum, cingulum, superior longitudinal fasciculus, and arcuate fasciculus. Higher polynomial order brings more flexibility to the curves, resulting in them being spread out more continuously in the white matter regions (e.g. in corona radiata). A 3D stereoscopic rendering of the results is shown in Fig. 7.7. Figure 7.8 shows a sagittal slice of the baseline image among the computed tracts. A significant portion of the curves can be seen to reach the gray matter, even though the FA (used as the prior) is lower around this region.

Next, we used the monkey brain HARDI dataset introduced in (106) to test the performance of our method on specific fiber bundles. An anesthetized Macaca mulatta monkey was scanned using a 7T MR scanner (Siemens) equipped with a head gradient coil (80mT/m G-maximum, 200mT/m/ms) with a diffusion weighted spin-echo EPI sequence. Diffusion images were acquired at b = 3000 s/mm^2 (twice during the same session, and then averaged) over 100 directions uniformly distributed on the sphere. We used TR/TE of 4600/65 ms, and a voxel size of 1×1×1 mm^3. We computed the tracts from 1350 seed points uniformly distributed in a mask containing the intersection of the forceps minor and the inferior longitudinal fasciculus, using the polynomials of order $N = 3$. Results are depicted in Fig. 7.9. A fiber density map was created by counting, at each voxel, the number of intersecting curves while taking into account their respective score. A three-dimensional isosurface was then generated by thresholding this map to keep the most relevant connections. It is as well presented in Fig. 7.9, overlaid on a structural MRI. Major tracts including the splenium fibers, posterior corona radiata, tapetum, as well as the inferior fronto-occipital and longitudinal fasciculi – including the optic radiations – are clearly identified. Moreover, fibers of the optic tract are recovered until they reach the optic chiasm.

7.3.2 Results for Multiple Subjects

We used our multi-subject tractography algorithm to compute mean tracts from five HARDI datasets, introduced in (168), each acquired from a different healthy young adult. Images were acquired using a 4T Bruker Medspec MRI scanner. Diffusion-weighted images were acquired using single-shot echo planar imaging with a twice-refocused spin echo sequence to reduce eddy-current induced distortions. Imaging parameters were: 23 cm FOV, TR/TE 6090/91.7 ms, with a 128×100 acquisition matrix. Each 3D volume consisted of 55 2-mm thick axial slices with a 1.8×1.8 mm² in-plane resolution. 105 images were acquired: 11 with no diffusion sensitization (i.e., T2-weighted b_0 images) and 94 diffusion-weighted images (b = 1159 s/mm²) with gradient directions evenly distributed on the sphere. Scan time was approximately 14 minutes. Images were corrected for motion and eddy current distortions. Each subject's average b_0 image was aligned to a group-specific minimal deformation template (MDT) using a nine-parameter affine transformation. This transformation was then applied to each individual DWI and gradient directions were corrected accordingly for ODF calculations.

We combined the individual datasets into two equivalent volumes, using the geometric and the arithmetic means (see Sec. 7.2.4). We tested the tractography algorithm on both equivalent volumes, and also for comparison, on two of the five individual datasets. In each experiment, polynomials of order $N = 3$ were used to represent 3D tracts initiated from 1500 seed points. Figure 7.10 (two top rows) show the mean tracts from the five subjects using respectively the geometric and arithmetic means. The two bottom subfigures show tracts from individual subjects. As Fig. 7.10 demonstrates, combining the volumes improves the results by producing fibers that are less scattered and better concentrated in major fiber bundles. Note particularly how corticospinal tracts are enhanced.

7.3.3 Discussion

Although every possible curve can be represented with an infinite polynomial using the Taylor expansion, keeping only $N + 1$ coefficients gives us an approximated curve which, as N grows, converges to the true curve. Through our experiments, we observed that at least a polynomial order of $N = 2$ is necessary to correctly extract the fiber bundles from the data, since $N = 0$ (basically resulting in straight lines) and $N = 1$ turn out to be models which are too simple. Since we did not observe noticeable improvement in our human brain tractography results by increasing N from 2 to 3 (see, e.g., Fig. 7.6) and beyond that, we deduced that the second or third order

polynomials are most likely sufficient for this type of data.[30] Depending on the available resources, the slight improvements by increasing the polynomial order from $N = 2$ to $N = 3$ may or may not be worth the two added dimensions in the search space (see the Appendix 9.3 for more details about the complexity of the algorithm). The proposed framework is not limited to these low orders, and increasing the order will only result in additional computational cost.

Throughout our experiments, we determined the best values of λ (see Sec. 7.2.2) assessing the results obtained using different values for it. As seen in Fig. 7.5, a value of λ which is too large or too small might result in false positives and false negatives, respectively. We observed that the manually-determined optimal λ remains consistent for different datasets with the same acquisition parameters. This means that the method can be easily used in population studies, without the need to tune λ separately for each subject.

[30] This, however, needs further investigation which is part of the future research.

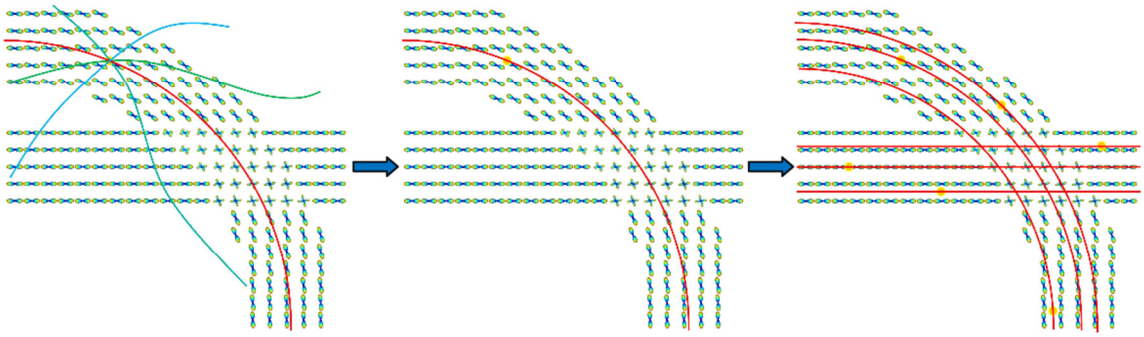


Fig 7.1. (Left) Different possible curves passing through a seed point are tested and their scores are computed. (Middle) The curve with the highest score is selected. (Right) The process is repeated for all the remaining seed points.

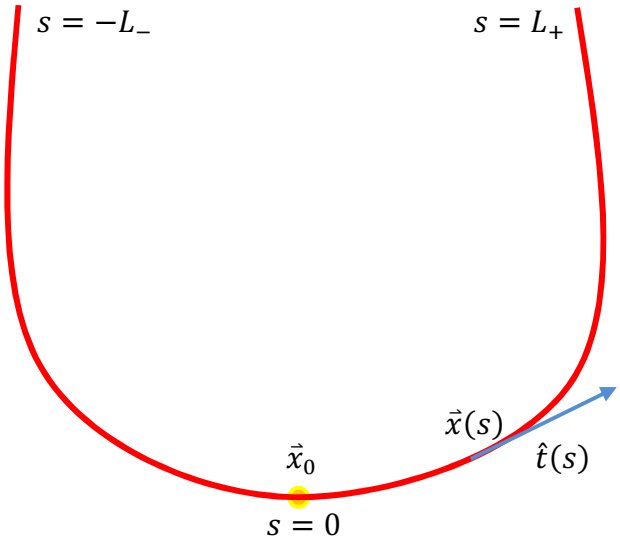


Fig 7.2. Curves starting from the seed point $\vec{x}_0$ are parameterized by the arc length, $s \in [-L_-, L_+]$. The unit tangent vector, $\hat{t}(s)$, is approximated with polynomials.

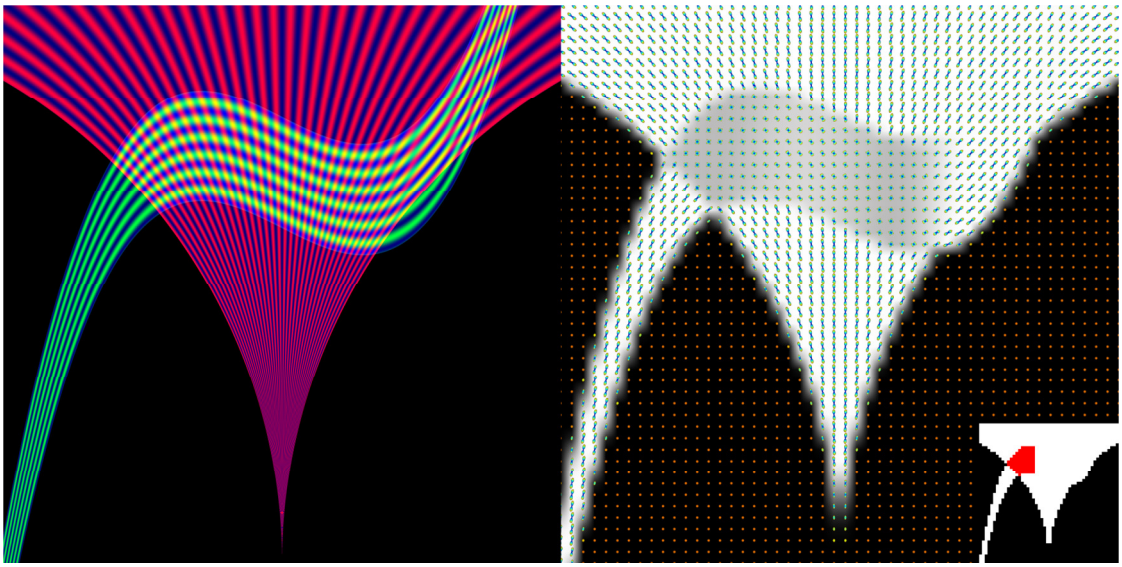


Fig 7.3. Ground truth (left) and the ODFs overlaid on the GFA map (right) of the simulated phantom. The local region for the seed points and the phantom mask are indicated in respectively red and white (bottom, right).

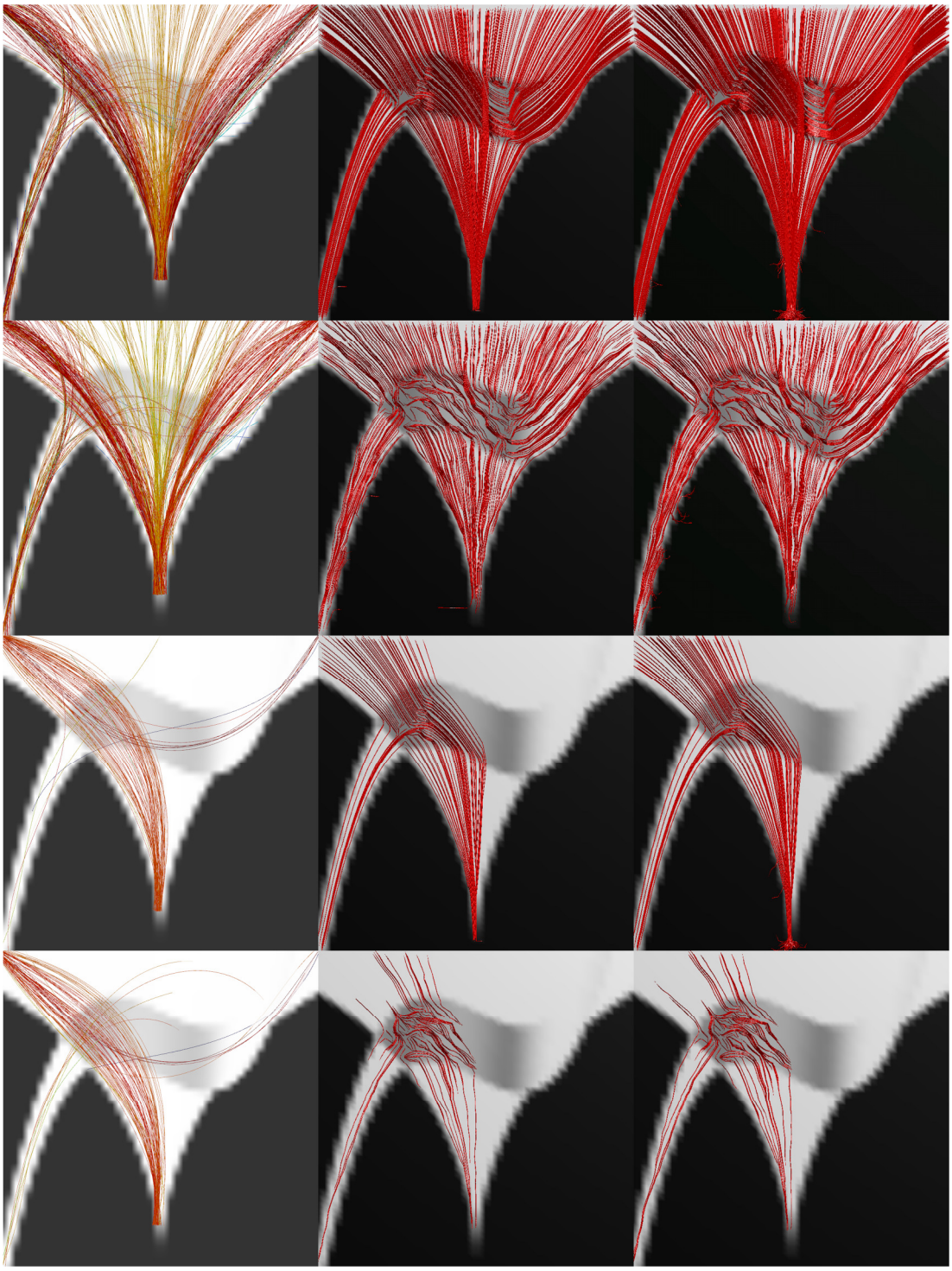


Fig 7.4. Comparison of the proposed method (left) with streamline deterministic (middle) and probabilistic (right) techniques from seed points chosen in the phantom mask (two top rows) and the region identified in *Fig. 7.3 (bottom, right)* as red (two bottom rows), in noiseless (rows 1 & 3) and noisy (rows 2 & 4) cases.

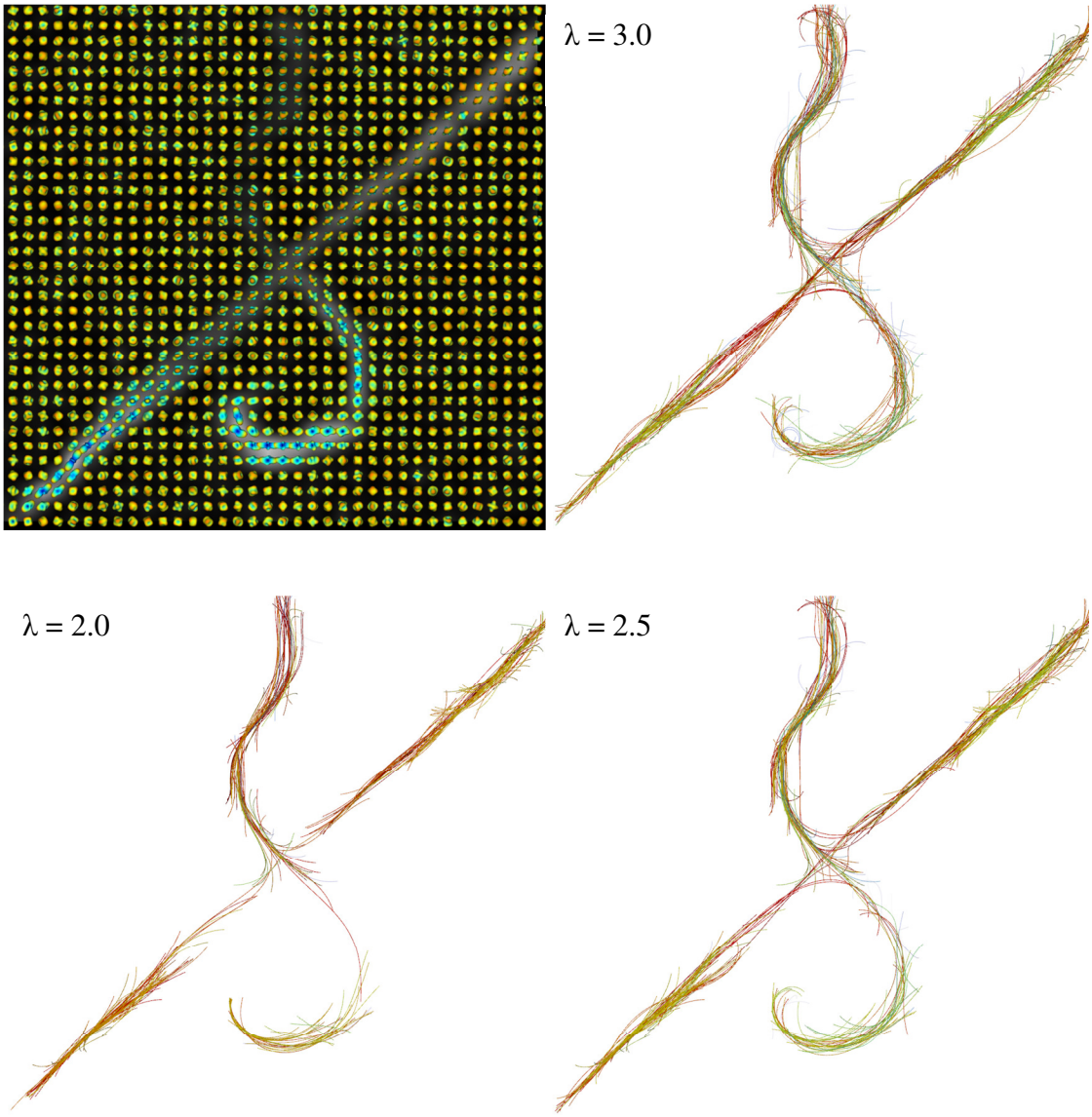


Fig 7.5. Reconstructed ODFs (top left) and the tractography results (rest of the subfigures) on the excised rat spinal cords, using various values for the bias parameter λ.

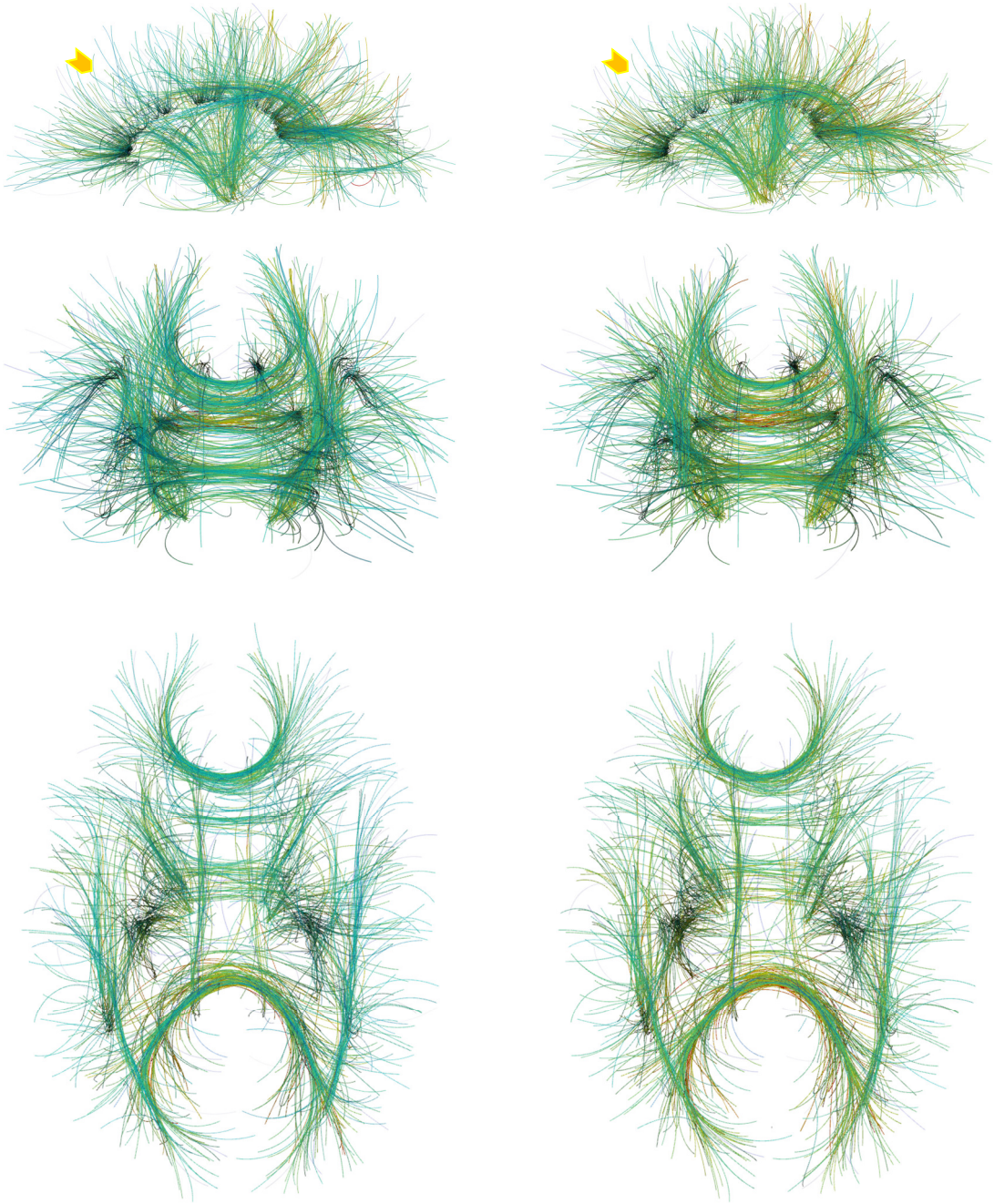


Fig 7.6. Tractography results on a human brain HARDI dataset using polynomial orders of (left) $N = 3$ and (right) $N = 2$, shown in (top) sagittal, (middle) coronal, and (bottom) axial views. The arrows indicate areas where the computed curves are more uniformly spread out when using the higher polynomial order.

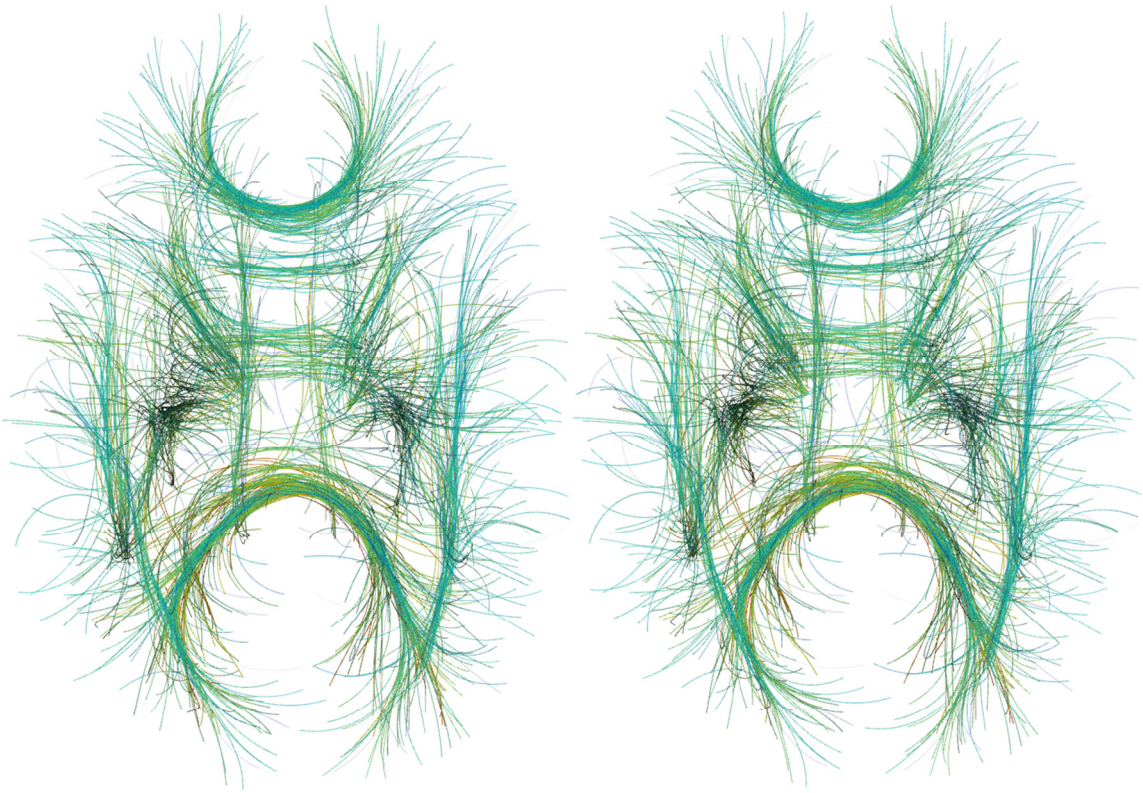


Fig 7.7. Stereoscopic rendering of *Fig. 7.6 (bottom, left)*. To see this image in 3D, please cross your eyes and move the image closer or further away from you until you see what appears to be a third, 3D image in the middle. This figure would be a bonus for those who can perceive 3D with standard eye-crossing techniques.

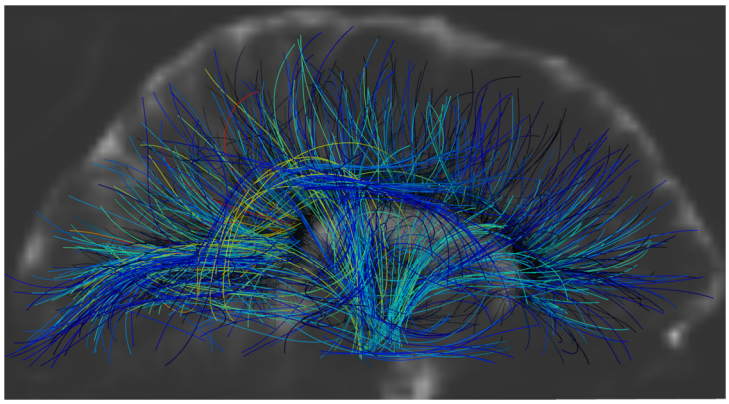


Fig 7.8. A sagittal slice of the human brain baseline image in the middle of the computed tracts. Cerebrospinal fluid is identified as white regions.

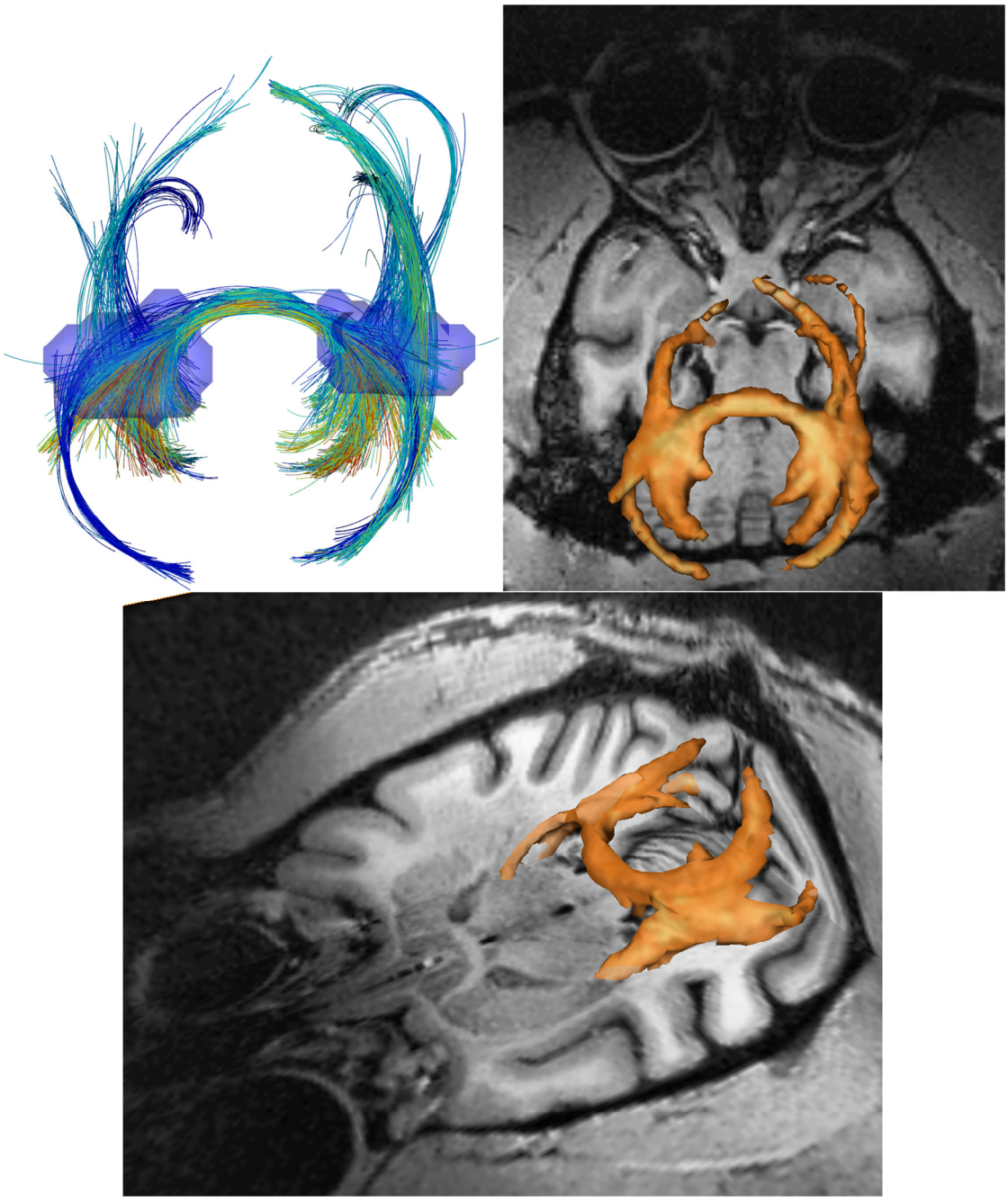


Fig 7.9. Tractography results on a monkey brain HARDI dataset shown in axial (top) and tilted (bottom) views. Seed points were randomly generated inside the transparent blue regions (top, left). A mask of the tracts is shown superimposed on the T1 anatomy image (top right & bottom).

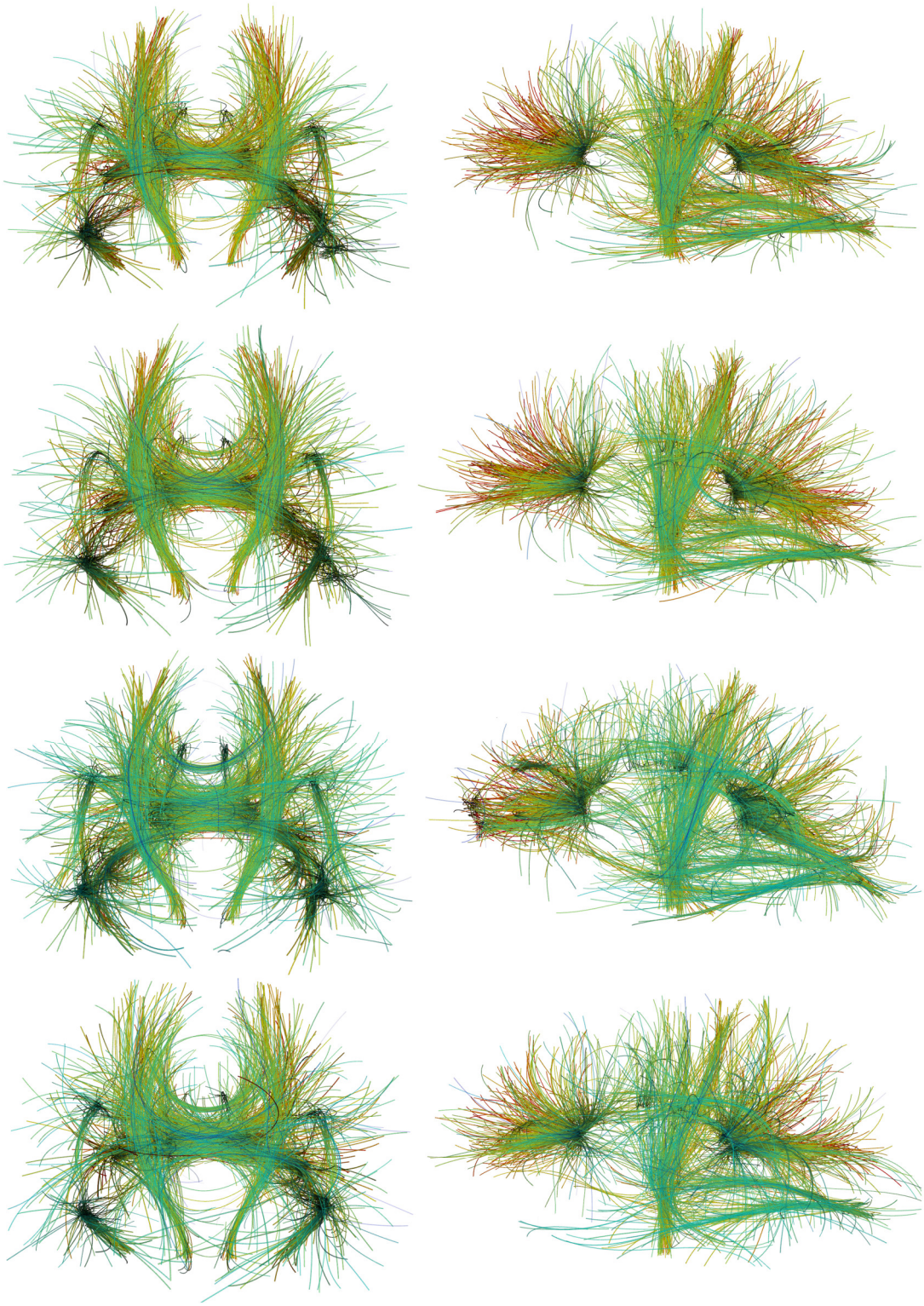


Fig 7.10. Tractography results from five human brain HARDI datasets combined using geometric (top row) and arithmetic (row two) means, and from individual subjects (two bottom rows), shown in coronal (left) and sagittal (right) views.

8 Conclusions

8.1 Wavelet-Based Image Fusion

In Chapter 2, an image fusion technique based on the wavelet transform was proposed to combine MR images with orthogonal (low-resolution) slice-selection directions. The advantages of the wavelet transform to the commonly used Fourier transform were demonstrated, and experimental results on MRI data were shown.

8.2 Cortical Thickness Measurement

In Chapter 3, we presented a new definition of cortical thickness along with an algorithm for computing it. We were motivated by the importance of measuring the thickness of the cerebral cortex for quantifying the progression of various neurodegenerative brain diseases. Our method calculates the thickness at each voxel, by computing all line integrals of the probability map of the gray matter passing through that voxel, and choosing the minimum integral value. Two stopping criteria are taken into consideration to address issues created by narrow sulci. Unlike most prior work, we take into account the probability of each voxel belonging to the gray matter layer and do not carry out a hard segmentation prior to measuring the thickness. The proposed algorithm is significantly faster than popular ones such as the one implemented in FreeSurfer. After such accurate voxel-based computations, the thickness measurements can be mapped into mesh representations if these are needed for other processing steps. This can be done in a number of different ways, e.g., by assigning to each vertex in the mesh the value of the corresponding voxel or a weighted average of nearby voxels. Thereby, the proposed mesh- and segmentation-free tissue thickness computation can easily be integrated into mesh based pipelines, enjoying the best of both representations.

We have validated the technique with artificial data and presented reasonable results for longitudinal MRI scans of Alzheimer's disease and normal subjects. Further improvements are expected when considering, for example, more sophisticated soft-classification techniques as input to the proposed framework.

We are currently investigating the incorporation of smoothness constraints, where the minimum line integral at a given voxel is influenced by those of neighborhood voxels. This can be done, in principle, making the optimization a global one, acting on all voxels of interest at once, and adding a penalization term for large local deviations of the corresponding minimal line. The main challenge here to be addressed is the computational one. Note of course that

smoothness can also be incorporated as part of the pre-processing probability computation and/or as part of a post-processing step after the tissue thickness and minimal lines have been computed. We are also applying the framework introduced here to large population studies.

8.3 Diffusion-Weighted MRI

In Chapter 4, we have proposed a novel framework to correct a mathematical inaccuracy in the original q-ball imaging, and have demonstrated its improved performance through experiments on simulated and real HARDI data. We also extended our technique to the case with multiple q-shells, and observed enhancement in the reconstructed orientation distribution functions (ODFs) which were, however, achieved at the cost of additional (while still moderate) acquisition time. Whether using multiple q-shells and reducing the number of diffusion directions (to maintain the same acquisition time) would still be helpful, is a subject of current research.

In Chapter 5, a measure of fiber-crossing was introduced, which exploited the difference between the ODFs computed from tensors and q-balls in the regions with crossing fibers. This quantity was computed for a large dataset, with some brain regions revealing significant correlation with the performance intelligence quotient.

Chapter 6 describes a semi-analytical approach to ODF maxima extraction in the spherical harmonic basis. This approach uses analytically derived pathways on the ODF to reduce the search space from two-dimensional to one-dimensional, considerably reducing the complexity of the exhaustive search.

In Chapter 7, we have introduced a global approach for single- and multi-subject probabilistic tractography, based on the voting process provided by the Hough transform. We presented experimental results on a physical phantom and brain HARDI datasets, and showed that using this approach, data from multiple subjects can be non-linearly combined and exploited to obtain population statistics and more accurate tractography results. The incorporation of spatial coherence and continuity in curve score computation is the subject of future research.

References

1 Goshtasby AA, Turner DA. Fusion of short-axis and long-axis cardiac MR images. Computerized Medical Imaging and Graphics. 1996;20(2):77–87.

2 Tsougarakis K, Steines D, Timsari B, inventors. Fusion of multiple imaging planes for isotropic imaging in MRI and quantitative image analysis using isotropic or near-isotropic imaging. 2009. Patent #7634119.

3 Hamilton CA, Elster AD, Ulmer JL. "Crisscross" MR imaging: improved resolution by averaging signals with swapped phase-encoding axes. Radiology. 1994;193:276–279.

4 Herment A, Roullot E, Bloch I, Jolivet O, De Cesare A, Frouin F, Bittoun J, Mousseaux E. Local reconstruction of stenosed sections of artery using multiple MRA acquisitions. Magnetic Resonance in Medicine. 2003;49(4):731–742.

5 Hoge WS, Mitsouras D, Rybicki FR, Mulkern RV, Westin CF. Registration of multidimensional image data via subpixel resolution phase correlation. In: IEEE Intl. Conf. on Image Processing ; 2003; Barcelona, Spain. p. 707–710.

6 Roullot E, Herment A, Bloch I, de Cesare A, Nikolova M, Mousseaux E. Modeling anisotropic undersampling of magnetic resonance angiographies and reconstruction of a high-resolution isotropic volume using half-quadratic regularization techniques. Signal Processing. 2004;84(4):743–762.

7 Kwan RKS, Evans AC, Pike GB. MRI simulation-based evaluation of image-processing and classification methods. IEEE Transactions on Medical Imaging. 1999;18(11):1085–1097.

8 Greenspan H, Oz G, Kiryati N, Peled S. MRI inter-slice reconstruction using super-resolution. Magnetic Resonance Imaging. 2002;20:437–446.

9 Bai Y, Han X, Prince JL. Super-resolution reconstruction of MR brain images. In: 38th Annual Conference on Information Sciences and Systems; 2004. p. 1358–1363.

10 Carmi E, Liu S, Alon N, Fiat A, Fiat D. Resolution enhancement in MRI. Magnetic Resonance Imaging. 2006;24(2):133–154.

11 Tamez-Peña JG, Totterman S, Parker KJ. MRI isotropic resolution reconstruction from two orthogonal scans. In: Proc. SPIE; 2001; San Diego, CA.

12 Museth K, Breen DE, Zhukov L, Whitaker RT. Level set segmentation from multiple non-uniform volume datasets. In: Proc. of the IEEE conference on Visualization; 2002; Boston, MA.

13 Huntsberger T, Jawerth B. Wavelet-based sensor fusion. In: Proc. SPIE; 1993; Boston, MA. p. 488–498.

14 Ranchin T, Wald L, Mangolini M. Efficient data fusion using wavelet transform: the case of SPOT satellite images. In: Proc. SPIE; 1993. p. 171–178.

15 Koren I, Laine A, Taylor F. Image fusion using steerable dyadic wavelet transform. In: Proc. IEEE ICIP; 1995; Washington, DC. p. 232–235.

16 Li H, Manjunath S, Mitra SK. Multisensor Image Fusion Using the Wavelet

Transform. Graphical Models and Image Processing. 1995;57(3):235–245.

17 Wilson TA, Rogers SK, Myers LR. Perceptual-based hyperspectral image fusion using multiresolution analysis. Optical Engineering. 1995;34(11):3154—3164.

18 Yocky DA. Image merging and data fusion by means of the discrete two-dimensional wavelet transform. J. Optical Society of America. 1995;12(9):1834–1841.

19 Chipman LJ, Orr TM, Graham LN. Wavelets and image fusion. In: Wavelet Applications in Signal and Image Processing III; 1995; San Diego.

20 Lejeune C. Wavelet transforms for infrared applications. In: Proc. SPIE; 1995; San Diego, CA. p. 313–324.

21 Rockinger O. Pixel-level fusion of image sequences using wavelet frames. In: Proc. of 16th Leeds Applied Shape Research workshop; 1996.

22 Jiang X, Zhou L, Gao Z. Multispectral image fusion using wavelet transform. In: Proc. SPIE; 1996; Beijing, China. p. 35–42.

23 Peytavin L. Cross-sensor resolution enhancement of hyperspectral images using wavelet decomposition. In: Proc. SPIE; 1996; Orlando, FL. p. 193–197.

24 Petrovic VS, Xydeas CS. Cross-band pixel selection in multiresolution image fusion. In: Sensor Fusion: Architectures, Algorithms, and Applications III; 1999; Orlando, FL.

25 Zhang Z, Blum RS. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proc. of the IEEE. 1999;87(8):1315–1326.

26 Shu-Long Z. Image fusion using wavelet transform. In: Symp. Geospatial Theory, Processing and Applications; 2002; Ottawa.

27 Hill P, Canagarajah N, Bull D. Image fusion using complex wavelets. In: Proc. 13th British Machine Vision Conference; 2002; Cardiff, UK.

28 Pajares G, Manuel de la Cruz J. A wavelet-based image fusion tutorial. Pattern Recognition. 2004;37(9):1855–1872.

29 Daubechies I. Ten lectures on wavelets. SIAM; 1992.

30 Mallat S. A wavelet tour of signal processing: the sparse way. 3rd ed. Academic Press; 2008.

31 Aganj I, Lenglet C, Yacoub E, Sapiro G, Harel N. A wavelet fusion approach to the reconstruction of isotropic-resolution MR images from anisotropic orthogonal scans. Submitted to ISMRM; 2011; Montreal.

32 Aganj I, lenglet C, Yacoub E, Sapiro G, Harel N. A wavelet fusion approach to the reconstruction of isotropic-resolution MR images from anisotropic orthogonal scans. To be submitted to Magnetic Resonance in Medicine. 2010.

33 Thompson PM, Hayashi KM, Sowell ER, Gogtay N, Giedd JN, Rapoport JL, de Zubicaray GI, Janke AL, Rose SE, Semple J, et al. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. NeuroImage. 2004;23:S2–18.

34 Thompson PM, Lee AD, Dutton RA, Geaga JA, Hayashi KM, Eckert MA, Bellugi U, Galaburda AM, Korenberg JR, Mills DL, et al. Abnormal cortical complexity and thickness profiles mapped in Williams syndrome. J Neurosci. 2005;25:4146–4158.

35 Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci USA. 2000;97:11050–11055.

36 Jones SE, Buchbinder BR, Aharon I. Three-dimensional mapping of cortical thickness using Laplace's equation. Hum Brain Mapp. 2000;11:12–32.

37 Kabani N, Le GG, MacDonald D, Evans AC. Measurement of cortical thickness using an automated 3-D algorithm: a validation study. NeuroImage. 2001;13:375–380.

38 Lohmann G, Preul C, Hund-Georgiadis M. Morphology-based cortical thickness estimation. In: Inf Process Med Imaging; 2003. p. 89–100.

39 Yezzi AJ, Prince JL. An Eulerian PDE approach for computing tissue thickness. IEEE Trans Med Imag. 2003;22:1332–1339.

40 Lerch JP, Evans AC. Cortical thickness analysis examined through power analysis and a population simulation. NeuroImage. 2005;24:163–173.

41 Thorstensen N, Hofer M, Sapiro G, Pottmann H. Measuring cortical thickness from volumetric MRI data. Unpublished. 2006.

42 Young K, Schuff N. Measuring structural complexity in brain images. NeuroImage. 2008;39:1721–1730.

43 Pham DL, Bazin PL. Simultaneous boundary and partial volume estimation in medical images. In: Proc Intl. Conf. MACCAI; 2004. p. 119–126.

44 Aganj I, Sapiro G, Parikshak N, Madsen SK, Thompson P. Segmentation-free measurement of cortical thickness from MRI. In: Proc. IEEE ISBI; 2008; Paris, France.

45 Aganj I, Sapiro G, Parikshak N, Madsen SK, Thompson P. Measurement of cortical thickness from MRI by minimum line integrals on soft-classified tissue. Human Brain Mapping. 2009;30(10):3188–3199.

46 Li Y, Wang Y, Xue Z, Shi F, Lin W, Shen D, ADNI. Consistent 4D cortical thickness measurement for longitudinal neuroimaging study. In: Proc. MICCAI; 2010; Beijing, China. p. 133–142.

47 MacDonald D, Kabani N, Avis D, Evans AC. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage. 2000;12:340–356.

48 Miller MI, Massie AB, Ratnanather JT, Botteron KN, Csernansky JG. Bayesian construction of geometrically based cortical thickness metrics. NeuroImage. 2000;12:676–687.

49 Scott MLJ, Thacker NA. Cerebral cortical thickness measurements. Tina Memo. 2004 2007–2007.

50 Haidar H, Soul JS. Measurement of cortical thickness in 3D brain MRI data: Validation of the Laplacian method. NeuroImage. 2006;16:146–153.

51 Pizer S, Eberly D, Fritsch D, Morse B. Zoom-invariant vision of figural shape: The

mathematics of cores. Comput Vision Image Understanding. 1998;69:55–71.

52 Leow AD, Klunder AD, Jack CR, Toga AW, Dale AM, Bernstein MA, Britson PJ, Gunter JL, Ward CP, Whitwell JL, et al. Longitudinal stability of MRI for mapping brain change using tensor-based morphometry. NeuroImage. 2005;31:627–640.

53 Teo PC, Sapiro G, Wandell BA. Creating connected representations of cortical gray matter for functional MRI visualization. IEEE Trans Med Imaging. 1997;16:852–863.

54 Chung MK, Worsley KJ, Taylor J, Ramsay JO, Robbins S, Evans AC. Diffusion Smoothing on the Cortical Surface. NeuroImage. 2001;13:95.

55 Lerch JP, Pruessner JC, Zijdenbos A, Hampel H, Teipel SJ, Evans AC. Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy. Cereb Cortex. 2005;15:995–1001.

56 Mémoli F, Sapiro G, Thompson PM. Implicit brain imaging. NeuroImage. 2004;23:S179–S188.

57 Hua X, Leow AD, Lee S, Klunder AD, Toga AW, Lepore N, Chou YY, Brun C, Chiang MC, Barysheva M, et al. 3D characterization of brain atrophy in Alzheimer's disease and mild cognitive impairment using tensor-based morphometry. NeuroImage. 2008;41:19–34.

58 Jack CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, Whitwell J, Ward C, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI Methods. J Magn Reson Imaging. 2008;27:685–91.

59 Wedeen VJ, Hagmann P, Tseng WI, Reese TG, Weisskoff RM. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magnetic Resonance in Medicine. 2005;54(6):1377–1386.

60 Tuch DS, Reese TG, Wiegell MR, Makris N, Belliveau JW, Wedeen VJ. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magnetic Resonance in Medicine. 2002;48(4):577–582.

61 Tuch DS. Q-ball imaging. Magnetic Resonance in Medicine. 2004;52(6):1358–1372.

62 Funk P. Über eine geometrische Anwendung der Abelschen Integralgleichung. Mathematische Annalen. 1916;77:129–135.

63 Anderson AW. Measurement of fiber orientation distributions using high angular resolution diffusion imaging. Magnetic Resonance in Medicine. 2005;54(5):1194–1206.

64 Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Regularized, fast, and robust analytical q-ball imaging. Magnetic Resonance in Medicine. 2007;58(2):497–510.

65 Hess CP, Mukherjee P, Han ET, Xu D, Vigneron DB. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magnetic Resonance in Medicine. 2006;56(1):104–117.

66 Michailovich OV, Rathi Y. On approximation of orientation distributions by means of spherical ridgelets. In: Proc. of 5th IEEE ISBI; 2008; Paris.

67 Rathi Y, Michailovich O, Shenton ME, Bouix S. Directional functions for orientation

distribution estimation. Med Image Anal. 2009;13(3):432–444.

68 Canales-Rodríguez EJ, Melie-García L, Iturria-Medina Y. Mathematical description of q-space in spherical coordinates: Exact q-ball imaging. Magnetic Resonance in Medicine. 2009;61(6):1350–1367.

69 Tristán-Vega A, Westin CF, Aja-Fernández S. Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. NeuroImage. 2009;47(2):638–650.

70 Khachaturian MH, Wisco JJ, Tuch DS. Boosting the sampling efficiency of q-ball imaging using multiple wavevector fusion. Magnetic Resonance in Medicine. 2007;57(2):289–296.

71 Wu YC, Field AS, Alexander AL. Computation of diffusion function measures in q-space using magnetic resonance hybrid diffusion imaging. IEEE Transactions on Medical Imaging. 2008;27(6):858–865.

72 Descoteaux M, Deriche R, Le Bihan D, Mangin JF, Poupon C. Diffusion Propagator Imaging: Using Laplace's equation and multiple shell acquisitions to reconstruct the diffusion propagator. In: Proc. of 21st Intl. Conf. on IPMI; 2009; Williamsburg.

73 Tournier JD, Calamante F, Gadian DG, Connelly A. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage. 2004;23(3):1176–1185.

74 Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. NeuroImage. 2007;35(4):1459–1472.

75 Jian B, Vemuri BC. A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. IEEE Trans Med Imaging. 2007;26(11):1464–71.

76 Jansons KM, Alexander DC. Persistent angular structure: new insights from diffusion magnetic resonance imaging data. Inverse Problems. 2003;19:1031–1046.

77 Özarslan E, Shepherd TM, Vemuri BC, Blackband SJ, Mareci TH. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). NeuroImage. 2006;31:1086–1103.

78 Lenglet C, Campbell JSW, Descoteaux M, Haro G, Savadjiev P, Wassermann D, Anwander A, Deriche R, Pike GB, Sapiro G, et al. Mathematical methods for diffusion MRI processing. NeuroImage. 2009;45(1):S111–S122.

79 Seunarine KK, Alexander DC. Multiple fibers: Beyond the diffusion tensor. In: Johansen-Berg H, Behrens TEJ. Diffusion MRI: From quantitative measurement to in vivo neuroanatomy. 1st ed. Academic Press; 2009.

80 Aganj I, Lenglet C, Sapiro G. Accurate ODF reconstruction in q-ball imaging. In: 15th OHBM; 2009; San Francisco, CA.

81 Aganj I, Lenglet C, Sapiro G. ODF reconstruction in q-ball imaging with solid angle consideration. In: Proc. of 6th IEEE ISBI; 2009; Boston.

82 Aganj I, Lenglet C, Sapiro G, Yacoub E, Ugurbil K, Harel N. Multiple q-shell ODF

reconstruction in q-ball imaging. In: Proc. 12th Intl. Conf. MICCAI; 2009; London.

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

84 Caruyer E, Aganj I, Lenglet C, Sapiro G, Deriche R. Online orientation distribution function reconstruction in constant solid angle and its application to motion detection in HARDI. In: Proc. 7th ISBI; 2010; Rotterdam, Netherlands.

85 Caruyer E, Aganj I, Lenglet C, Sapiro G, Deriche R. Online motion detection in high angular resolution diffusion imaging. Submitted to 8th IEEE ISBI; 2011.

86 Goh A, Lenglet C, Thompson P, Vidal R. Estimating orientation distribution functions with probability density constraints and spatial regularity. In: Proc. 12th Intl. Conf. MICCAI; 2009; London.

87 Polzehl J, Tabelow K. Beyond the diffusion tensor model: The package dti. WIAS Preprint No. 1563; 2010.

88 Fritzsche K, Meinzer HP. MITK-DI - A new diffusion imaging component for MITK. In: Deserno TM, Handels H, Meinzer HP, Tolxdorff T. Bildverarbeitung für die Medizin. Heidelberg: Springer; 2010. p. 246-250.

89 Robinson EC, Valstar M, Hammers A, Ericsson A, Edwards AD, Rueckert D. Multivariate statistical analysis of whole brain structural networks obtained using probabilistic tractography. In: Proc. 11th Intl. Conf. MICCAI; 2008; New York City. p. 486–493.

90 Aganj I, Lenglet C, Keriven R, Sapiro G, Harel N, Thompson PM. A Hough transform global approach to diffusion MRI tractography. In: Proc. of 17th Annual Meeting of ISMRM; 2009; Honolulu.

91 Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. J Magn Reson B. 1994;103(3):247–254.

92 Stejskal EO, Tanner JE. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. Journal of Chemical Physics. 1965;42(1):288–292.

93 Assemlal HE, Tschumperlé D, Brun L. Efficient and robust computation of PDF features from diffusion MR signal. Medical Image Analysis. 2009;13(5):715–729.

94 Barnett A. Theory of q-ball imaging redux: Implications for fiber tracking. Magnetic Resonance in Medicine. 2009;62:910–923.

95 Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. A linear and regularized ODF estimation algorithm to recover multiple fibers in q-ball imaging. Research report 5768. INRIA; 2005.

96 Jones DK. Gaussian modeling of the diffusion signal. In: Johansen-Berg H, Behrens TEJ. Diffusion MRI: From quantitative measurement to in vivo neuroanatomy. 1st ed. Academic Press; 2009.

97 Ronen I, Kim KH, Garwood M, Ugurbil K, Kim DS. Conventional DTI vs. slow and fast diffusion tensors in cat visual cortex. Magnetic Resonance in Medicine. 2003;49(5):785–790.

98 Branch MA, Coleman TF, Li Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. SIAM Journal on Scientific Computing. 1999;21(1):1–23.

99 Niendorf T, Dijkhuizen RM, Norris DG, Campagne ML, Nicolay K. Biexponential diffusion attenuation in various states of brain tissue: Implications for diffusion-weighted imaging. Magnetic Resonance in Medicine. 1996;36(6):847–857.

100 Yeh CH, Cho KH, Lin HC, Wang JJ, Lin CP. Reduced encoding diffusion spectrum imaging implemented with a bi-Gaussian model. IEEE Transactions on Medical Imaging. 2008;27(10):1415–1424.

101 Jian B, Vemuri BC, Özarslan E, Carney PR, Mareci TH. A novel tensor distribution model for the diffusion-weighted MR signal. NeuroImage. 2007;37(1):164–176.

102 Hartigan JA, Hartigan PM. The dip test of unimodality. The Annals of Statistics. 1985;13(1):70-84.

103 Fritzsche KH, Laun FB, Meinzer HP, Stieltjes B. Opportunities and pitfalls in the quantification of fiber integrity: What can we gain from Q-ball imaging? NeuroImage. 2010;51(1):242–251.

104 Campbell JSW, Siddiqi K, Rymar VV, Sadikot AF, Pike GB. Flow-based fiber tracking with diffusion tensor and q-ball data: Validation and comparison to principal diffusion direction techniques. NeuroImage. 2005;27(4):725–736.

105 Poupon C, Poupon F, Allirol L, Mangin JF. A database dedicated to anatomo-functional study of human brain connectivity. In: Proc. of the 12th Annual Meeting of OHBM; 2006.

106 Lenglet C, Yacoub E, Ghose G, Adriany G, Krüger G, Sapiro G, Ugurbil K, Harel N. High resolution diffusion MRI on in-vivo monkey brains at 7T. In: Proc. 14th Annual Meeting of OHBM; 2009; San Francisco.

107 Aganj I, Jahanshad N, Lenglet C, Toga AW, McMahon KL, de Zubicaray GI, Wright MJ, Martin NG, Sapiro G, Thompson P. Relating fiber crossing in HARDI to intellectual function. In: 16th OHBM; 2010; Barcelona, Spain.

108 Goh A, Lenglet C, Thompson P, Vidal R. A nonparametric Riemannian framework for processing High Angular Resolution Diffusion Images (HARDI). In: IEEE Conference on Computer Vision and Pattern Recognition; 2009. p. 2496–2503.

109 Cheng J, Ghosh A, Jiang T, Deriche R. A Riemannian framework for orientation distribution function computing. In: Proc. of 12th International Conference on Medical Image Computing and Computer Assisted Intervention; 2009. p. 911–918.

110 Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C, Klunder AD, Toga AW, McMahon KL, de Zubicaray GI, et al. Genetics of brain fiber architecture and intellectual performance. Journal of Neuroscience. 2009;29(7):2212-2224.

111 Mori S, Crain BJ, Chacko VP, Van Zijl PCM. Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Annals of Neurology. 1999;45(2):265–269.

112 Conturo TE, Lori NF, Cull TS, Akbudak E, Akbudak AZ, Shimony JS, McKinstry RC, Burton H, Raichle ME. Tracking neuronal fiber pathways in the living human brain. Proc. of National Academy of Sciences. 1999;96(18):10422–10427.

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

114 Frey S, Campbell JSW, Pike GB, Siddiqi K. Dissociating the human language pathways with high angular resolution diffusion fiber tractography. Journal of Neuroscince. 2008;28(45):11435–11444.

115 Berman JI, Chung S, Mukherjee P, Hess CP, Han ET, Henry RG. Probabilistic streamline q-ball tractography using the residual bootstrap. NeuroImage. 2008;39(1):215–222.

116 Ghosh A, Tsigaridas E, Descoteaux M, Comon P, Mourrain B, Deriche R. A polynomial based approach to extract the maxima of an antipodally symmetric spherical function and its application to extract fiber directions from the Orientation Distribution Function in Diffusion MRI. In: Proc. of Workshop on Computational Diffusion MRI, MICCAI; 2008.

117 Bloy L, Verma R. On computing the underlying fiber directions from the diffusion orientation distribution function. In: Proc. of 11th MICCAI; 2008; New York. p. 1–8.

118 Qi L, Han D, Wu EX. Principal invariants and inherent parameters of diffusion kurtosis tensors. Journal of Mathematical Analysis and Applications. 2009;349(1):165–180.

119 Aganj I, Lenglet C, Sapiro G. ODF maxima extraction in spherical harmonic representation via analytical search space reduction. In: Proc. 13th MICCAI; 2010; Beijing, China. p. 84–91.

120 Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A. In vivo fiber tractography using DT-MRI data. Magnetic Resonance in Medicine. 2000;44(4):625–632.

121 Jones DK, Simmons A, Williams SCR, Horsfield MA. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. Magnetic Resonance in Medicine. 1999;42(1):37–41.

122 Lazar M, Weinstein DM, Tsuruda JS, Hasan KM, Arfanakis K, Meyerand ME, Badie B, Rowley HA, Haughton V, Field A, et al. White matter tractography using diffusion tensor deflection. Human Brain Mapping. 2003;18(4):306 –321.

123 Björnemo M, Brun A, Kikinis R, Westin CF. Regularized stochastic white matter tractography using diffusion tensor MRI. In: Proc. of 5th MICCAI; 2002; Tokyo.

124 Descoteaux M, Deriche R, Knösche TR, Anwander A. Deterministic and probabilistic tractography based on complex fibre orientation distributions. IEEE Transactions on Medical Imaging. 2009;28(2):269–286.

125 Friman O, Farnebäck G, Westin CF. A Bayesian approach for stochastic white matter tractography. IEEE Transactions on Medical Imaging. 2006;25(8):965–978.

126 Jones DK. Tractography gone wild: Probabilistic fibre tracking using the wild

bootstrap with diffusion tensor MRI. IEEE Transactions on Medical Imaging. 2008;27(9):1268–1274.

127 Lazar M, Alexander AL. Bootstrap white matter tractography (BOOT-TRAC). NeuroImage. 2005;24(2):524–532.

128 Parker GJM, Haroon HA, Wheeler-Kingshott CAM. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. Journal of Magnetic Resonance Imaging. 2003;18(2):242–254.

129 Jackowski M, Kao C, Qiu M, Constable R, Staib L. White matter tractography by anisotropic wavefront evolution and diffusion tensor imaging. Medical Image Analysis. 2005;9(5):427–440.

130 Parker GJM, Wheeler-Kingshott CAM, Barker GJ. Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. IEEE Transactions on Medical Imaging. 2002;21(5):505–512.

131 Pichon E, Westin CF, Tannenbaum AR. A Hamilton-Jacobi-Bellman approach to high angular resolution diffusion tractography. In: Proc. of 8th MICCAI; 2005; Palm Springs.

132 Prados E, Lenglet C, Pons JP, Wotawa N, Deriche R, Faugeras O, Soatto S. Control theory and fast marching methods for brain connectivity mapping. In: Proc. IEEE Conf. CVPR; 2006; New York.

133 Tournier JD, Calamante F, Gadian DG, Connelly A. Diffusion-weighted magnetic resonance imaging fibre tracking using a front evolution algorithm. NeuroImage. 2003;20(1):276–288.

134 Batchelor PG, Hill DLG, Atkinson D, Calamante F. Study of connectivity in the brain using the full diffusion tensor from MRI. In: Proc. of 17th IPMI; 2001; Davis.

135 Hageman NS, Toga AW, Narr KL, Shattuck DW. A diffusion tensor imaging tractography algorithm based on Navier–Stokes fluid mechanics. IEEE Transactions on Medical Imaging. 2009;28(3):348-360.

136 Hagmann P, Thiran JP, Jonasson L, Vandergheynst P, Clarke S, Maeder P, Meuli R. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. NeuroImage. 2003;19(3):545–554.

137 Kang N, Zhang J, Carlson ES, Gembris D. White matter fiber tractography via anisotropic diffusion simulation in the human brain. IEEE Transactions on Medical Imaging. 2005;24(9):1127–1137.

138 O'Donnell L, Haker S, Westin CF. New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space. In: Proc. of 5th MICCAI; 2002; Tokyo.

139 Yörük E, Acar B, Bammer R. A physical model for DT-MRI based connectivity map computation. In: Proc. of 8th MICCAI; 2005; Palm Springs.

140 Jbabdi S, Bellec P, Toro R, Daunizeau J, Pelegrini-Issac M, Benali H. Accurate anisotropic fast marching for diffusion-based geodesic tractography. International

Journal of Biomedical Imaging. 2008;2008.

141 Lenglet C, Prados E, Pons JP, Deriche R, Faugeras O. Brain connectivity mapping using riemannian geometry, control theory, and PDEs. SIAM Journal on Imaging Sciences. 2009;2(2):285–322.

142 Melonakos J, Mohan V, Niethammer M, Smith K, Kubicki M, Tannenbaum A. Finsler tractography for white matter connectivity analysis of the cingulum bundle. In: Proc. of 10th MICCAI; 2007; Brisbane.

143 Pechaud M, Descoteaux M, Keriven R. Brain connectivity using geodesics in HARDI. In: Proc. of 12th MICCAI; 2009; London.

144 Iturria-Medina Y, Canales-Rodríguez EJ, Melie-García L, Valdés-Hernández PA, Martínez-Montes E, Alemán-Gómez Y, Sánchez-Bornot JM. Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. NeuroImage. 2007;36(3):645–660.

145 Sotiropoulos SN, Bai L, Morgan PS, Constantinescu CS, Tench CR. Brain tractography using q-ball imaging and graph theory: Improved connectivities through fibre crossings via a model-based approach. NeuroImage. 2010;49(3).

146 Zalesky A. DT-MRI fiber tracking: a shortest paths approach. IEEE Transactions on Medical Imaging. 2008;27(10):1458–1471.

147 Fillard P, Poupon C, Mangin JF. A novel global tractography algorithm based on an adaptive spin glass model. In: Proc. of 12th MICCAI; 2009; London.

148 Mangin JF, Poupon C, Cointepas Y, Rivière D, Papadopoulos-Orfanos D, Clark CA, Régis J, Le Bihan D. A framework based on spin glass models for the inference of anatomical connectivity from diffusion-weighted MR data - a technical review. NMR in Biomedicine. 2002;15(7–8):481–492.

149 Kreher BW, Mader I, Kiselev VG. Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. Magnetic Resonance in Medicine. 2008;60(4):953–963.

150 Behrens TEJ, Jbabdi S. MR diffusion tractography. In: Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy. 1st ed. Academic Press; 2009.

151 El Kouby V, Cointepas Y, Poupon C, Rivière D, Golestani N, Poline JB, Le Bihan D, Mangin JF. MR diffusion-based inference of a fiber bundle model from a population of subjects. In: Proc. of 8th MICCAI; 2005; Palm Springs.

152 Jbabdi S, Woolrich MW, Behrens TEJ. Multiple-subjects connectivity-based parcellation using hierarchical Dirichlet process mixture models. NeuroImage. 2009;44(2):373–384.

153 Leemans A, Sijbers J, De Backer S, Vandervliet E, Parizel P. Multiscale white matter fiber tract coregistration: a new feature-based approach to align diffusion tensor data. Magnetic Resonance in Medicine. 2006;55(6):1414–1423.

154 Maddah M, Grimson WEL, Warfield SK. Statistical modeling and EM clustering of white matter fiber tracts. In: Proc. of 3rd IEEE ISBI; 2006.

155 O'Donnell LJ, Westin CF. Automatic tractography segmentation using a high-

dimensional white matter atlas. IEEE Transactions on Medical Imaging. 2007;26(11):1562–1575.

156 Voineskos AN, O'Donnell LJ, Lobaugh NJ, Markant D, Ameis SH, Niethammer M, Mulsant BH, Pollock BG, Kennedy JL, Westin CF, et al. Quantitative examination of a novel clustering method using magnetic resonance diffusion tensor tractography. NeuroImage. 2009;45(2):370–376.

157 Wakana S, Jiang H, Nagae-Poetscher LM, van Zijl PCM, Mori S. Fiber tract–based atlas of human white matter anatomy. Radiology. 2004;230:77–87.

158 Duda RO, Hart PE. Use of the Hough transformation to detect lines and curves in pictures. Commun. ACM. 1972;15(1).

159 Gonzalez RC, Woods RE. Digital Image Processing. 3rd ed. Prentice Hall; 2008.

160 Aganj I, Lenglet C, Sapiro G, Chiang MC, Thompson PM. Multi-subject diffusion MRI tractography via a Hough transform global approach. In: Proc. of 15th Annual Meeting of OHBM; 2009; San Francisco.

161 Aganj I, Lenglet C, Jahanshad N, Yacoub E, Harel N, Thompson P, Sapiro G. A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography. Submitted to Medical Image Analysis. 2010.

162 Jahanshad N, Aganj I, Lenglet C, Joshi A, Jin Y, Barysheva M, McMahon KL, de Zubicaray GI, Martin NG, Wright MJ, et al. Sex differences in the human connectome: 4-Tesla high angular resolution diffusion imaging (HARDI) tractography in 234 young adult twins. Submitted to 8th IEEE ISBI; 2011.

163 Prasad G, Jahanshad N, Aganj I, Lenglet C, Sapiro G, Toga AW, Thompson PM. Atlas-based fiber clustering for multi-subject analysis of high angular resolution diffusion imaging tractography. Sumbitted to 8th IEEE ISBI; 2011.

164 Haro G, Lenglet C, Sapiro G, Thompson PM. On the Non-Uniform Complexity of Brain Connectivity. In: Proc. 5th IEEE ISBI; 2008; Paris.

165 Chiang MC, Barysheva M, Lee AD, Madsen S, Klunder AD, Toga AW, McMahon KL, de Zubicaray GI, Meredith M, Wright MJ, et al. Brain fiber architecture, genetics, and intelligence: A high angular resolution diffusion imaging (HARDI) study. In: Proc. of 11th MICCAI; 2008; New York.

166 Cook PA, Bai Y, Nedjati-Gilani S, Seunarine KK, Hall MG, Parker GJ, Alexander DC. Camino: Open-source diffusion-MRI reconstruction and processing. In: 14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine; 2006; Seattle, WA. p. 2759.

167 Seunarine KK, Cook PA, Hall MG, Embleton KV, Parker GJM, Alexander DC. Exploiting peak anisotropy for tracking through complex structures. In: IEEE 11th International Conference on Computer Vision, Workshop on MMBIA; 2007; London.

168 de Zubicaray GI, Chiang MC, McMahon KL, Shattuck DW, Toga AW, Martin NG, Wright MJ, Thompson PM. Meeting the Challenges of Neuroimaging Genetics. Brain Imaging and Behavior. 2008;2(4):258–263.

169 Laidlaw DH, Fleischer KW, Barr AH. Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms. IEEE Trans Med Imaging. 1998;17:74–86.

170 Choi HS, Haynor DR, Kim YM. Multivariate tissue classification of MRI images for 3-D volume reconstruction—A statistical approach. In: Proc SPIE Medical Imaging III: Image Processing; 1989. p. 183–193.

171 Choi HS, Haynor DR, Kim YM. Partial volume tissue classification of multichannel magnetic resonance images—A mixel model. IEEE Trans. Med. Imaging. 1991;10:395–407.

172 Jovicich J, Czanner S, Greve D, Haley E, van der Kouwe A, Gollub R, Kennedy D, Schmitt F, Brown G, Macfall J, et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage. 2006;30:436–443.

173 Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging. 1998;17:87–97.

9 Appendices

9.1 Implementation of the Cortical Thickness Measurement

We now explain in detail how we implemented our proposed algorithm. The input to the algorithm is the raw MRI dataset which is in the form of a three-dimensional matrix, and the result is the thickness map volume with the same size and spatial sampling as the input.

9.1.1 Computing the Probability Map

We consider the probability of each voxel belonging to the GM as a Gaussian distribution on the intensity value of the voxel in the MRI volume. The mean of the Gaussian is the mean value of manually-selected sample voxels in the GM, while the standard deviation is the difference between the manually estimated mean values of GM and WM. We could use more sophisticated soft classification algorithms, such as Partial-Volume Bayesian algorithm (PVB) (169), Probabilistic Partial Volume Classifier (PPVC) (170), and Mixel classifier (171), to further improve the results. Before segmentation, several pre-processing steps were applied to ensure the accurate calibration of the scans over time and across subjects. Specifically: 1) a procedure termed *GradWarp* was applied for correction of geometric distortion due to gradient non-linearity (172), 2) a "B1-correction" was applied, to adjust for image intensity non-uniformity using B1 calibration scans (58), 3) "N3" bias field correction, for reducing intensity inhomogeneity caused by non-uniformities in the radio frequency (RF) receiver coils (173), and 4) geometrical scaling, according to a phantom scan acquired for each subject (58), to adjust for scanner – and session – specific calibration errors. In addition to the original uncorrected image files, images with all of these corrections already applied (GradWarp, B1, phantom scaling, and N3) are available to the general scientific community (www.loni.ucla.edu/ADNI).

9.1.2 Preparing the Masks

In order to compute the line integrals, we made separate cubic mask volumes for all the line segments, which are in different quantized directions in three-dimensional space. The length of each dimension of the mask volumes was chosen to be slightly longer than the actual length of the line segment. However, since binary masks are highly inaccurate to apply in numerical integration, we considered continuous mask volumes. We first built binary masks four times bigger in each dimension, and instead of line segments, we considered cylinders four times longer than the line segments with diameter four. Next, we applied a low-pass filter to it and

downsampled the results to achieve the desired non-binary masks (note that this smoothing acts on the mask, not on the actual MRI data). They were afterwards normalized so that the values in each mask added to the length of the line segment. A two-dimensional example is depicted in Fig. 9.1.1.

The directions were chosen by quantizing the unit hemisphere uniformly, as the complete sphere would have been redundant. In the standard spherical coordinates, the unit hemisphere was sampled every 10° in the latitudinal direction, whereas the longitudinal sampling rate varied as follows:

$$\begin{aligned} \theta &\in [0°, 90°) && \text{Every } 10° \\ \varphi &\in [0°, 360°) && \text{Every } 10°/\cos\theta \end{aligned} \qquad [55]$$

This particular quantization was chosen in order to achieve a relatively constant surface sampling rate, since the solid angle would be

$$\Delta\Omega = \cos\theta \Delta\theta \Delta\phi = \left[\cos\theta . 10° . \frac{10°}{\cos\theta} \right]_{in\ steradians} = const. \qquad [56]$$

In addition, each mask contained only half of its corresponding line segment (extending from the origin to one point on the hemisphere), given that the two halves were identical and also that the line integrals on them would later need to be taken separately.

Next, the mask volumes were stored in a sparse format that is a list of non-zero elements with their coordinates. To facilitate the integration, elements were sorted into ascending order with respect to their Euclidean distance from the origin. The masks were then stored for further use in the numerical integration step, for various input data.

9.1.3 Numerical Integration

We now briefly explain how the line integrals of a probability map are numerically calculated. To compute the thickness measure at each given voxel, we integrate the probability map on line segments centered at that voxel in different directions using the pre-computed mask volumes, and then choose the smallest integral value as the local thickness. Since our stopping criteria make the entire operation nonlinear, we are not able to use convolution methods – which are computationally faster if done in frequency domain, although demanding more memory – to compute the line integrals.

Each line integral is computed in two similar phases. Starting from the main voxel for which we intend to compute the thickness, in each phase we advance in one of the two opposite

directions, using the mask volume in the first phase, and its symmetry about the origin (with negated coordinates) in the second phase. In each phase, we start adding the probability values on the region around the main voxel using as weights the values in the sparse form of the corresponding mask. Since the non-zero elements of the mask are sorted into ascending order with respect to their Euclidean distance to the origin, the values of the closer voxels to the main voxel are added first.

The line integral does not need to be entirely computed, since we have a record of the minimum integral value that has been computed so far for a voxel, and we do not need to exceed that. Therefore, the summation continues until one of the following occurs:

1. All the non-zero elements of the mask are used.
2. The sum reaches or exceeds the minimum integral value that has so far been computed for that voxel.
3. One of the stopping criteria becomes true (see Sec. 3.3.2).

For the first stopping condition (explained in Sec. 3.3.2), a counter is set for the number of successive voxels (with the original sampling size) encountered with probabilities less than the threshold. In this work we used the threshold of 0.3, and summing stopped as the counter reached 10. For the second condition, two states are necessary. In the initial state, a variable counts the number of consecutive decreases in the probability. When it reaches a specific number, the state changes. In the second state, a new variable counts the number of increases in the probability and the summation stops when it reaches a predefined number. In this work 10 was used for both mentioned numbers. Since the masks are made in an interpolative way, they each have many more non-zero voxels than the real length (in voxels) of their corresponding line segments. Therefore, the number of counted iterations is usually about ten times larger than the actual proceeded length, which results in higher precision in counting. For example, a one-voxel thick valley can be detected after about five consecutive decreases and five consecutive increases of the probability.

The process of measuring and comparing the thickness for a pair of scans took about 2 to 3 hours for our technique, in contrast to 2 to 3 days running time of FreeSurfer (note once again that FreeSurfer is computing a number of things during this time, not just tissue thickness). In both cases, a Linux machine with a 2.33GHz Intel CPU was used. One can also parallelize the above procedure, since it is done independently for each voxel in the volume. For instance, we divided the volume into eight sub-volumes and ran eight parallel jobs in different processors to obtain the thickness map in considerably less time.

9.1.4 Finding the Skeleton

Since the output of the proposed algorithm is the set of thickness values for all the voxels in the volume, we need to find the particular voxels that lie inside the GM layer to compute statistical results for group comparisons, and for comparing the results with FreeSurfer.[31] A simple GM mask may not necessarily represent the most appropriate set of voxels to examine. Not only might the thickness be underestimated at the voxels close to the inner/outer surfaces, but the values would also be weighted depending on how thick each segment is, i.e., more voxels with the same thickness value are counted on the thicker parts of the GM. Thus, it is best to make a *skeleton* of the GM, which is a mask with constant thickness positioned in the middle of the layer, far from the inner/outer surfaces. This is of course just one possible way of reporting the rich amount of information produced by the proposed algorithm, where every voxel contains a tissue thickness measurement. The particular way of exploiting this information might depend on the task at hand.

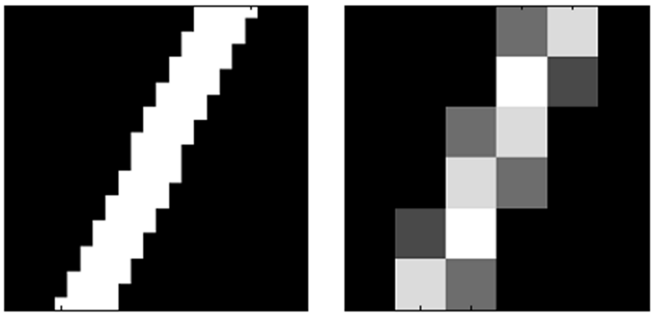
In order to find the skeleton without explicit segmentation, we modified the algorithm such that for each voxel v, in addition to the total thickness $T(v)$, it reports the two different thickness values $t_1(v)$ and $t_2(v)$, the lengths of the two sub-segments on each side of the voxel on the optimal line segment ($t(v) = t_1(v) + t_2(v)$, see Fig. 9.1.2). Next, we considered the skeleton to be the set of the voxels with the following properties:

$$\begin{aligned} |t_1(v) - t_2(v)| &\le 0.5, \\ P(v) &> 0.4, \end{aligned} \qquad [57]$$

where $P(v)$ is the probability of v belonging to the GM. The first condition guarantees that the skeleton is always at most one voxel thick, and the two conditions together guarantee that the skeleton remains in the middle of the GM layer.

In addition, to better remove the skull and non-GM parts of the brain, we applied a rough GM mask obtained from FreeSurfer (FreeSurfer was used just for this task).

[31] Note once again that such segmentation is not an intrinsic component of the proposed framework, it is just one way of reporting the results.



(a) (b)

Fig 9.1.1. Illustration of how line integral masks are generated. (a) A high-resolution binary mask, which is four times larger in each direction, is first generated. (b) The real-size non-binary mask is produced by downsampling the binary mask.

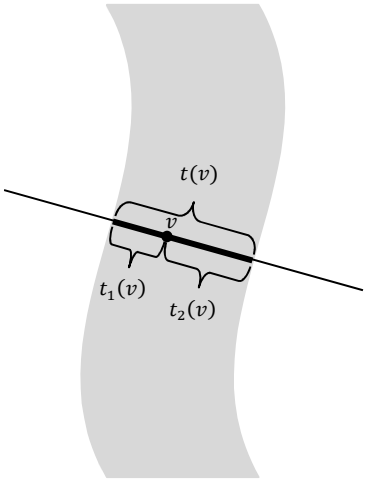


Fig 9.1.2. The lengths, $t_1(v)$ and $t_2(v)$, of the two sub-segments are used in finding the gray matter skeleton for computing the cortical thickness variation.

9.2 Computation of the CSA-ODF

9.2.1 Fourier Transform of $P(\vec{r})|\vec{r}|^2$

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

$$\begin{aligned} F\{xP(\vec{r})\} &= i\frac{\partial}{\partial q_x}E(\vec{q}) \\ F\{x^2P(\vec{r})\} &= -\frac{\partial^2}{\partial q_x{}^2}E(\vec{q}) \end{aligned}, \qquad [58]$$

where $F\{\bullet\}$ is the Fourier transform. By writing the second equation for y and z and summing them all, we will get:

$$F\{r^2P(\vec{r})\} = -\nabla^2 E(\vec{q}). \qquad [59]$$

This has also been exploited in a parallel work, (69).

9.2.2 Radial Projection of a Symmetric Function

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

Without loss of generality, we choose our coordinates such that $\hat{z}=\hat{u}$, thus making $\hat{u}^\perp$ the q_x-q_y plane. We first rewrite the expression as a volume integral over the entire space, with the help of Dirac delta functions:

$$\int_0^\infty f(r\hat{z})dr = \int_0^\infty f(0,0,z)dz = \frac{1}{2}\iiint_{R^3} f(x,y,z)\delta(x)\delta(y)dxdydz. \qquad [60]$$

The factor ½ is required because we need the integral only on the positive half of the z-axis, and the function is symmetric. Let us define $g(x,y,z):=\delta(x)\delta(y)$. For the two functions $f,g:\mathbb{R}^3\to\mathbb{R}$ with Fourier transform functions $\hat{f}(\vec{q})$ and $\hat{g}(\vec{q})$, Parseval's theorem states that

$$\iiint_{R^3} f(x,y,z)g^*(x,y,z)dxdydz = \frac{1}{(2\pi)^3}\iiint_{R^3}\hat{f}(q_x,q_y,q_z)\hat{g}^*(q_x,q_y,q_z)dq_xdq_ydq_z. \qquad [61]$$

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

$$\begin{aligned}\int_0^\infty f(r\hat{z})\,dr &= \frac{1}{2}\iiint_{R^3} f(x,y,z)\,g(x,y,z)\,dxdydz \\ &= \frac{1}{2(2\pi)^3}\iiint_{R^3} \hat{f}(q_x,q_y,q_z)\,2\pi\delta(q_z)\,dq_x dq_y dq_z \\ &= \frac{1}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{f}(q_x,q_y,0)\,dq_x dq_y\end{aligned} \quad [62]$$

.

The integral is taken on the q_x-q_y plane, which is $\hat{u}^\perp$. This completes the proof.

9.2.3 Mono-Exponential Model for the ODF

We will show that by assuming the mono-exponential model, $E(\vec{q}) = \tilde{E}(\hat{u})^{\frac{q^2}{q_0^2}}$, we have:

$$\int_0^{2\pi}\int_0^\infty \frac{1}{q}\nabla_b^2 E(\vec{q})\,dq\,d\phi = -\frac{1}{2}\int_0^{2\pi}\nabla_b^2 \ln\left(-\ln\tilde{E}(\hat{u})\right)d\phi, \quad [63]$$

while $\theta = \frac{\pi}{2}$ is kept constant in the integration. We begin by proving a lemma:

Lemma: For a continuous and differentiable function $f(\theta,\phi): S^2 \to \mathbb{R}$ with S^2 being the unit sphere, we have:

$$\int_0^{2\pi}\left(\nabla_b^2 f\right)\left(\frac{\pi}{2},\phi\right)d\phi = \int_0^{2\pi} f_{\theta\theta}\left(\frac{\pi}{2},\phi\right)d\phi, \quad [64]$$

where the subscript indicates the partial derivative.

Proof: We use the following expansion for Laplace-Beltrami:

$$\begin{aligned}\nabla_b^2 f &= \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2} \\ &= \cot\theta.f_\theta + f_{\theta\theta} + \frac{1}{\sin^2\theta}f_{\phi\phi} \\ \nabla_b^2 f\big|_{\theta=\frac{\pi}{2}} &= f_{\theta\theta} + f_{\phi\phi}\end{aligned} \quad [65]$$

.

Integral of the second term is zero, because of the periodicity of f_ϕ:

$$\int_0^{2\pi} f_{\phi\phi}\left(\frac{\pi}{2},\phi\right)d\phi = f_\phi\left(\frac{\pi}{2},\phi\right)\Big|_0^{2\pi} = 0. \quad [66]$$

Thus, the only remaining term in the integral is $f_{\theta\theta}$, which completes the proof of the lemma.

We now change the order of the integrals twice, while using the lemma in between:

$$\begin{aligned}\int_0^{2\pi}\int_0^{\infty}\frac{1}{q}\nabla_b^2 E(\bar{q})dqd\phi &= \int_0^{\infty}\frac{dq}{q}\int_0^{2\pi}\nabla_b^2 E(\bar{q})d\phi \\ &= \int_0^{\infty}\frac{dq}{q}\int_0^{2\pi}E_{\theta\theta}(\bar{q})d\phi \\ &= \int_0^{\infty}\frac{dq}{q}\frac{\partial}{\partial\theta}\int_0^{2\pi}E_{\theta}(\bar{q})d\phi \\ &= \int_0^{2\pi}\frac{\partial}{\partial\theta}\int_0^{\infty}\frac{1}{q}E_{\theta}(\bar{q})dqd\phi\end{aligned} \quad [67]$$

Next, we compute the radial integral:

$$\begin{aligned}\int_0^{\infty}\frac{1}{q}E_{\theta}(\bar{q})dq &= \int_0^{\infty}\frac{1}{q}\left(\tilde{E}(\hat{u})^{\frac{q^2}{q_0^2}}\right)_{\theta}dq \\ &= \int_0^{\infty}\frac{1}{q}\left(e^{\frac{q^2}{q_0^2}\ln\tilde{E}(\hat{u})}\right)_{\theta}dq \\ &= \left(\ln\tilde{E}(\hat{u})\right)_{\theta}\int_0^{\infty}\frac{q}{q_0^2}e^{\frac{q^2}{q_0^2}\ln\tilde{E}(\hat{u})}dq \\ &= \frac{\left(\ln\tilde{E}(\hat{u})\right)_{\theta}}{2\ln\tilde{E}(\hat{u})}e^{\frac{q^2}{q_0^2}\ln\tilde{E}(\hat{u})}\Bigg|_0^{\infty}\end{aligned} \quad [68]$$

Given that $0<\tilde{E}(\hat{u})<1\rightarrow\ln\tilde{E}(\hat{u})<0$, the above expression vanishes as $q\rightarrow\infty$. Also, since for a negative function $f(\theta)$ we have $\frac{f_{\theta}(\theta)}{f(\theta)}=\left[\ln|f(\theta)|\right]_{\theta}=\left[\ln(-f(\theta))\right]_{\theta}$, the above integral simplifies as:

$$\int_0^{\infty}\frac{1}{q}E_{\theta}(\bar{q})dq=-\frac{1}{2}\left[\ln\left(-\ln\tilde{E}(\hat{u})\right)\right]_{\theta}. \quad [69]$$

Substituting in Eq. [67]:

$$\begin{aligned}\int_0^{2\pi}\int_0^{\infty}\frac{1}{q}\nabla_b^2 E(\bar{q})dqd\phi &= -\frac{1}{2}\int_0^{2\pi}\left[\ln\left(-\ln\tilde{E}(\hat{u})\right)\right]_{\theta\theta}d\phi \\ &= -\frac{1}{2}\int_0^{2\pi}\nabla_b^2\ln\left(-\ln\tilde{E}(\hat{u})\right)d\phi\end{aligned} \quad [70]$$

We completed the proof by reusing the lemma in the last step.

9.2.4 Multi-Exponential Model for the ODF

By assuming the multi-exponential model, $E(q\hat{u})\cong\sum_{k=1}^{N}\lambda_k(\hat{u})\alpha_k(\hat{u})^{q^2}$, we will show that:

$$\int_0^{2\pi}\int_0^{\infty}\frac{1}{q}\nabla_b^2 E(\vec{q})dqd\phi = -\frac{1}{2}\int_0^{2\pi}\nabla_b^2\sum_{k=1}^{N}\lambda_k(\hat{u})\ln(-\ln\alpha_k(\hat{u}))d\phi\,, \qquad [71]$$

while $\theta=\frac{\pi}{2}$ is kept constant in the integration. The ODF will then be derived by replacing the above expression in Eq. [10].

The proof is an extension of Appendix 9.2.3. We first introduce the new non-negative variable $s_k(\hat{u}) := -\ln\alpha_k(\hat{u})$ which yields $E(q\hat{u}) = \sum_{k=1}^{N}\lambda_k(\hat{u})e^{-s_k(\hat{u})q^2}$. For simplicity, here we drop the notation $(\hat{u})$. We then continue from Eq. [67] and compute the radial integral

$$\begin{aligned}\int_0^{\infty}\frac{1}{q}E_\theta dq &= \int_0^{\infty}\frac{1}{q}\left(\sum_{k=1}^{N}\lambda_k e^{-s_k q^2}\right)_\theta dq \\ &= -\int_0^{\infty}\sum_{k=1}^{N}\lambda_k s_{k\theta}e^{-s_k q^2}qdq + \int_0^{\infty}\frac{1}{q}\sum_{k=1}^{N}\lambda_{k\theta}e^{-s_k q^2}dq\end{aligned}\,. \qquad [72]$$

The first integral is computed the same way as in Appendix 9.2.3:

$$\begin{aligned}-\int_0^{\infty}\sum_{k=1}^{N}\lambda_k s_{k\theta}e^{-s_k q^2}qdq &= -\sum_{k=1}^{N}\lambda_k s_{k\theta}\int_0^{\infty}qe^{-s_k q^2}dq \\ &= \sum_{k=1}^{N}\lambda_k\frac{s_{k\theta}}{2s_k}e^{-s_k q^2}\Big|_0^{\infty} \\ &= -\frac{1}{2}\sum_{k=1}^{N}\lambda_k(\ln s_k)_\theta\end{aligned}\,. \qquad [73]$$

Regarding the second integral of Eq. [72], let us define $I(\vec{s}) := \int_0^{\infty}\frac{1}{q}\sum_{k=1}^{N}\lambda_{k\theta}e^{-s_k q^2}dq$, with $\vec{s}$ the vector of s_k s, and derive it with respect to s_k :

$$\begin{aligned}\frac{\partial}{\partial s_k}I(\vec{s}) &= -\int_0^{\infty}q\lambda_{k\theta}e^{-s_k q^2}dq \\ &= \frac{\lambda_{k\theta}}{2s_k}e^{-s_k q^2}\Big|_0^{\infty} \\ &= -\frac{\lambda_{k\theta}}{2s_k} \\ &= \frac{\partial}{\partial s_k}\left(-\frac{\lambda_{k\theta}}{2}\ln s_k\right)\end{aligned}\,. \qquad [74]$$

Then, we can see that:

$$I(\vec{s}) = -\frac{1}{2}\sum_{k=1}^{N} \lambda_{k\theta} \ln s_k + C \quad , \tag{75}$$

where C is independent of $\vec{s}$. By evaluating the function for $\vec{s} = 1$ (vector of all 1s), we obtain $C = I(1)$, which we then compute using the original definition of $I(\vec{s})$:

$$\begin{aligned} I(1) &= \int_0^\infty \frac{1}{q} \sum_{k=1}^{N} \lambda_{k\theta} e^{-q^2} dq \\ &= \int_0^\infty \frac{1}{q} \left(\sum_{k=1}^{N} \lambda_k \right)_\theta e^{-q^2} dq \\ &= \int_0^\infty \frac{1}{q} (1)_\theta e^{-q^2} dq \\ &= 0 \end{aligned} \quad . \tag{76}$$

Therefore,

$$I(\vec{s}) = -\frac{1}{2}\sum_{k=1}^{N} \lambda_{k\theta} \ln s_k \, . \tag{77}$$

We now insert the values of the two integrals in Eq. [72]:

$$\begin{aligned} \int_0^\infty \frac{1}{q} E_\theta dq &= -\frac{1}{2}\sum_{k=1}^{N} \lambda_k (\ln s_k)_\theta - \frac{1}{2}\sum_{k=1}^{N} \lambda_{k\theta} \ln s_k \\ &= -\frac{1}{2}\left(\sum_{k=1}^{N} \lambda_k \ln s_k \right)_\theta \\ &= -\frac{1}{2}\left(\sum_{k=1}^{N} \lambda_k \ln(-\ln \alpha_k) \right)_\theta \end{aligned} \quad . \tag{78}$$

Finally, substituting in Eq. [67]:

$$\begin{aligned} \int_0^{2\pi} \int_0^\infty \frac{1}{q} \nabla_b^2 E(\vec{q}) dq d\phi &= -\frac{1}{2} \int_0^{2\pi} \left[\sum_{k=1}^{N} \lambda_k(\hat{u}) \ln(-\ln \alpha_k(\hat{u})) \right]_{\theta\theta} d\phi \\ &= -\frac{1}{2} \int_0^{2\pi} \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln(-\ln \alpha_k(\hat{u})) d\phi \end{aligned} \quad . \tag{79}$$

We completed the proof by using the lemma introduced in Appendix 9.2.3.

An interesting observation is that if $\alpha_k(\hat{u})$ is a set of estimated parameters in the multi-exponential model, then for any constant γ, the set $\alpha_k(\hat{u})^\gamma$ results in the same ODF:

$$
\begin{aligned}
\nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\ln \alpha_k(\hat{u})^{\gamma}\right) &= \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\gamma \ln \alpha_k(\hat{u})\right) \\
&= \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \left[\ln\left(-\ln \alpha_k(\hat{u})\right) + \ln \gamma\right] \\
&= \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\ln \alpha_k(\hat{u})\right) + \nabla_b^2 \left(\ln \gamma \sum_{k=1}^{N} \lambda_k(\hat{u})\right) \\
&= \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\ln \alpha_k(\hat{u})\right) + \nabla_b^2 \ln \gamma \\
&= \nabla_b^2 \sum_{k=1}^{N} \lambda_k(\hat{u}) \ln\left(-\ln \alpha_k(\hat{u})\right)
\end{aligned}
\qquad [80]
$$

This is expected, since the ODF is dimensionless and should not depend on the physical units of q.

9.3 Implementation of the Hough-Transform Tractography

In this appendix, we provide additional details on the implementation of the proposed tractography technique.

The *Hough transform* is often used in global optimization problems to avoid local optimum solutions. This approach is however characterized by its high computational complexity, given that in a straightforward implementation, all possible solutions must be tested in order to reconstruct the table of scores. This issue could yet be alleviated by parallelizing the exhaustive searches at the seed points, as they can be computed independently of one another. Note that out of the $d = 2N + 4$ parameters (10 for a polynomial of order $N = 3$), two are the partial curve lengths (L_+ and L_-), for different values of which the score is computed on the fly while the integral of Eq. [52] is taken. This integral is in fact computed progressively by discretizing the 3D curve and summing the integrand while advancing on the curve. At each step, the value of the accumulator represents the score for a fixed set of polynomial coefficient values, with a new value for the length parameter. This reduces the space where the score-integral is computed to $d' = 2N + 2$ dimensions (8 for $N = 3$). For $N = 2$, the algorithm (implemented in C) took about one and a half minutes to compute the maximum-score curve among over four billion curves for each seed point. This was increased to one to two hours per seed point in the case of $N = 3$ to test over three hundred billion curves. We ran about 50 to 100 parallel jobs to be able to finish the entire tractography (using about 1500 seed points) in less than a day. Also note that the (potentially) high-dimensional table of scores need *not* be stored in computer memory, since the maximum score can be computed on the fly, thus circumventing any memory exhaustion.

A question which may arise while implementing the proposed algorithm, is whether all the polynomial coefficients should have the same discretization resolution, and if not, how to determine it. From Eq. [46] it can be seen that the small change Δa_k in the k^{th} coefficient (due to its discretization resolution) results in the following change in θ at the arc length s:

$$\Delta\theta(s) = \Delta a_k s^k. \tag{81}$$

Ideally, we would like a uniform resolution for θ, which would mean that $\Delta\theta(s)$ needs to be independent of s and k. Although this dependency cannot be eliminated, it can be minimized by choosing a specific value for Δa_k. Assuming the desired constant value δ for $\Delta\theta(s)$, we minimize the following squared error integral to obtain the optimum value Δa_k^*:

$$\begin{aligned} \Delta a_k^* &= \underset{\Delta a_k}{\operatorname{argmin}} \int_{-L_{max}}^{L_{max}} (|\Delta\theta(s)| - \delta)^2 ds \\ &= \underset{\Delta a_k}{\operatorname{argmin}} \int_{-L_{max}}^{L_{max}} \left(\left|\Delta a_k s^k\right| - \delta\right)^2 ds \\ &= \frac{\delta}{L_{max}^k}\left(2 - \frac{1}{k+1}\right). \end{aligned} \tag{82}$$

By choosing such resolutions for the polynomial coefficients a_k, and similarly for b_k, the curve space is discretized more uniformly, hence increasing the accuracy of the search for the high scoring curves.