An Exploratory Approach to Generate Ground Truths of Neural Fiber Bundles

PEHR WESSMARK
AN EXPLORATORY APPROACH TO GENERATE GROUND TRUTHS OF NEURAL FIBER BUNDLES

by

Pehr Wessmark

A thesis submitted to the Department of Applied Physics in partial fulfillment of the requirements for the degree of

Master of Science in Engineering Physics

at

KTH Royal Institute of Technology

Thesis Supervisor: Dr. Rodrigo Moreno, Associate Professor, Medical Image Processing and Visualization, School of Technology and Health

Stockholm, Sweden
July 2017
Abstract

Diffusion magnetic resonance imaging (dMRI) utilizes the propensity of water molecules to diffuse anisotropically along coherent bundles of myelinated nerve fibers, which enables characterization of microstructural tissue changes in vivo. Using tractography coupled with quantitative measures of diffusion, it is possible to reconstruct fiber trajectories throughout the brain and delineate specific white matter fiber tracts that can be further analyzed for research or clinical purposes, including identification of basic mechanisms underlying alterations in neurological conditions. However, the specificity of measurements obtained from the use of tractography is limited due to inferences about how to model diffusion and resolve complex configurations such as curving or crossing fiber populations in individual image voxels. There exists therefore a need to provide a solution for ground truth generation of neural fiber bundles in order to overcome the overarching challenge of validating measurements obtained using diffusion-weighted magnetic resonance imaging techniques. In the present work, an exploratory approach to the issue of validation is described for the purpose of generating realistic ground truths of neural fiber bundles. The aim is to evaluate the feasibility of using synthetic diffusion-weighted datasets as training data for supervised machine learning and to investigate the possibility of constructing viable maps of local neural connections based on models of complex fiber configurations using different types of biophysical compartment models. The method involves the use of Fiberfox, a recently proposed open-source framework for simulating diffusion data from ground truths, an algorithm capable of reducing the number of false-positive fibers in global structural networks through a process entitled linear fascicle evaluation (LiFE), and spherical-deconvolution informed filtering of tractograms (SIFT) for improving the fit between streamline reconstructions and diffusion-weighted images. The results support the argument that collections of single-voxel fiber configurations in Fiberfox are unfit for training datasets. It is shown that the LiFE algorithm does not adequately characterize structural networks in setups involving crossing and curving fiber systems based on stochastic streamline tractography and no statistical evidence is found in favor of using the SIFT method for deterministic tractograms.
Acknowledgements

I am indebted first and foremost to my thesis supervisor Assoc. Prof. Dr. Rodrigo Moreno for sharing his expertise in image processing and medical imaging at all stages of the project and for his continued support and encouragement, without which this work would not have been possible.

My grateful thanks are extended to Dr. Örjan Smedby and the staff at the unit of Medical Image Processing and Visualization at the School of Technology and Health, KTH Royal Institute of Technology. I would especially like to acknowledge the technical support and valuable feedback provided by PhD students Daniel Jörgens and Irene Brusini, which contributed greatly to my research.

I would also like to thank Dr. Peter Neher at the German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ) for his advice on using the MITK software and for describing parameters and settings for partial volume statistics in the diffusion-imaging module.

Finally, I wish to thank Frida for her wonderful support and patience throughout my thesis work. I am lucky and profoundly grateful to have her in my life.
Contents

List of Figures .......................................................... ix
List of Tables ........................................................... xi
Acronyms and Abbreviations ........................................ x

1 Introduction ......................................................... 1
  1.1 Clinical Applications of Diffusion MRI ......................... 1
  1.2 The Challenge of Ground-Truth Validation .................... 2
  1.3 Purpose and goal .............................................. 3

2 Background and General Theory ................................. 4
  2.1 Basic Physics of Nuclear Magnetic Resonance ................. 4
    2.1.1 Signal Acquisition and Pulse Sequence Parameters ...... 5
    2.1.2 Spatial Encoding .......................................... 5
  2.2 Diffusion Magnetic Resonance Imaging ......................... 6
    2.2.1 Molecular Diffusion in Neural Tissue Microstructure ...... 6
    2.2.2 Sources of Diffusion Anisotropy .......................... 7
    2.2.3 Axonal Compartments .................................... 7
    2.2.4 Pulsed Gradient Diffusion-Weighted Imaging ............. 8
    2.2.5 The Diffusion Tensor ..................................... 9
    2.2.6 Deterministic Tractography ............................... 12
    2.2.7 Probabilistic Fiber Tracking and Uncertainty ............. 13
    2.2.8 Diffusion Tensor Imaging Scalar Indices ................ 14
  2.3 Related Work ................................................ 15
    2.3.1 Validation of Tractography ................................ 15
    2.3.2 Standardization and Surgical Dissection Techniques ...... 16
    2.3.3 Physical Phantoms and Ground-Truth Control ............. 16
    2.3.4 Capillary-Based Phantoms ................................ 17
    2.3.5 Software Phantoms ........................................ 17
    2.3.6 Reconstruction of Neural Fibers ........................... 18
    2.3.7 Multi-Compartment Diffusion Models ...................... 18
    2.3.8 Simulation Tools for Generating Synthetic Fibers .......... 19
    2.3.9 Prevalence and Geometry of Sheet Structures in the Brain . 20

3 Methods .......................................................... 22
  3.1 Synthetic Data Acquisition .................................... 22
    3.1.1 Fiber Definition and Configurations ....................... 22
    3.1.2 Diffusion-Weighted Signal Simulation and Preprocessing .. 23
    3.1.3 Compartment Model Selection .............................. 25
  3.2 Real Datasets ................................................ 26
    3.2.1 Ethics Statement .......................................... 26
    3.2.2 Hardware, Protocols, and Preprocessing .................. 26
  3.3 Linear Fascicle Evaluation of Synthetic Data .................. 27
    3.3.1 Stochastic Streamline Tractography ....................... 27
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Time-domain representation of the magnetic moment $\vec{\mu}$ in a magnetic field and the net magnetization following a 90-degree pulse</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Microanatomy of white matter myelinated axons showing the major components of the neuronal cytoskeleton</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>The Stejskal-Tanner prototype pulse sequence diagram showing two gradient pulses with fixed magnitude $G$ and equal duration $\delta$</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>A set of DW images (top row) and three T1-weighted images (bottom row) from a healthy subject using a b-value of $2000 \text{ s mm}^{-2}$</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>The rotationally invariant diffusion values $(\lambda_1, \lambda_2, \lambda_3)$ along a reference coordinate system $(x', y', z')$</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>Principle diagram of a general streamline routinely used in deterministic tractography that indicates the tangent vector at $\mathbf{r}(s_n)$</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>Error estimation and uncertainty $\Delta e_1$ in the fiber orientation around the eigenvector $e_1$ in the principal direction of diffusion</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>Fractional anisotropy maps of three image slices at progressively deeper levels of the brain in the axial, sagittal, and coronal plane</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>Fiducial placement and fiber definition in the Fiberfox framework</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>Schematic depicting the cross-sectional view of the three geometrical configurations of fibers at a macroscopic length scale</td>
<td>23</td>
</tr>
<tr>
<td>11</td>
<td>Applied diffusion-sensitizing gradients in 90 directions at a diffusion sensitivity factor of $1000 \text{ s mm}^{-2}$</td>
<td>27</td>
</tr>
<tr>
<td>12</td>
<td>Schematic diagram of two crossing fiber populations</td>
<td>30</td>
</tr>
<tr>
<td>13</td>
<td>Curving fiber configurations generated by introducing a transverse focus offset of a parabola</td>
<td>31</td>
</tr>
<tr>
<td>14</td>
<td>Comparative examples of noise and artifact simulations in five DW images of the $F_1$ ground truth bundle in Fiberfox</td>
<td>32</td>
</tr>
<tr>
<td>15</td>
<td>Axial view of the noise-free synthetic ground truth DW image with corrected fiber tracks</td>
<td>34</td>
</tr>
</tbody>
</table>
Figure 16  Sequential DW images of the crossing fiber population with respect to the zero-degree ground truth orientation ........................................ 36

Figure 17  Graphs showing the diffusion signals for different compartment models at a diffusion sensitivity factor of 1000 s·mm\(^{-2}\) .................. 37

Figure 18  Synthetic signals registered in central image voxels from the three fiber configurations with a 2-compartment partial volume ball-and-stick model ................................................................. 38

Figure 19  3-dimensional vector plots illustrating the variation between synthetic and real DW datasets ....................................................... 39

Figure 20  Box plot representations of noise-free streamline weighting factors obtained from simulations of crossing fiber systems with a maximum of 20 streamlines per fiber bundle ........................................ 43

Figure 21  Rician noise-corrupted streamline weighting factors from the ground truth and rotated fiber bundle ............................................ 44

Figure 22  Chi-square distributed noise influencing the variability between the ground truth and rotated fiber bundle weighting factors .......... 45

Figure 23  The influence of Gibbs-ringing artifacts on the estimation of streamline weighting factors ............................................................ 46

Figure 24  Aliasing-ridden streamline weighting factors from the ground truth and rotated fiber configurations .......................................... 47

Figure 25  Box plots showing the spread of streamline weighting factors for simulations involving parabolically curved fiber bundles ........ 48

Figure 26  The effect of Gibbs-ringing and aliasing artifacts on weighting factors from the ground truth and curved fiber bundles using single-shell diffusion acquisition ........................................ 49

Figure 27  Distributions of weighting factors for the ground truth and curved fiber configurations using multi-shell diffusion acquisition ......... 50
List of Tables

Table 3.1 Diffusion gradient table describing the diffusion encoding direction for the first 10 out of 90 gradient vectors at 1000 s-mm$^{-2}$ 24

Table 4.1 Invariant diffusion anisotropy indices of simulated datasets 38

Table 4.2 Ratios of crossing bundle fascicle weights to ground truth weights ($w/w_f$) using three diffusion sensitivity factors 40

Table 4.3 Ratios of crossing bundle weights to ground truth weights using a multi-shell acquisition for 1000, 2000, and 3000 s-mm$^{-2}$ 41

Table 4.4 Quotients of the weighting factors of curving fibers to the assigned weights of ground truth fibers 42

Table 4.5 Distribution of right-tailed crossing fiber-based p-values from Welch’s $t$-test for unpaired samples 51

Table 4.6 Distribution of right-tailed curving fiber-based p-values as the parabolic bundle is fitted to different foci $Q - \tilde{P}$ (mm) 52
### Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>dMRI</td>
<td>Diffusion Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>DW</td>
<td>Diffusion-Weighted</td>
</tr>
<tr>
<td>CC</td>
<td>Corpus Callosum</td>
</tr>
<tr>
<td>DTI</td>
<td>Diffusion Tensor Imaging</td>
</tr>
<tr>
<td>DWI</td>
<td>Diffusion-Weighted Imaging</td>
</tr>
<tr>
<td>RF</td>
<td>Radiofrequency</td>
</tr>
<tr>
<td>FID</td>
<td>Free Induction Decay</td>
</tr>
<tr>
<td>TE</td>
<td>Echo Time</td>
</tr>
<tr>
<td>TR</td>
<td>Repetition Time</td>
</tr>
<tr>
<td>PGSE</td>
<td>Pulsed Gradient Spin-Echo</td>
</tr>
<tr>
<td>ADC</td>
<td>Apparent Diffusion Coefficient</td>
</tr>
<tr>
<td>MITK</td>
<td>Medical Imaging Interaction Toolkit</td>
</tr>
<tr>
<td>ROI</td>
<td>Region Of Interest</td>
</tr>
<tr>
<td>FA</td>
<td>Fractional Anisotropy</td>
</tr>
<tr>
<td>RA</td>
<td>Relative Anisotropy</td>
</tr>
<tr>
<td>PTFE</td>
<td>Polytetrafluoroethylene</td>
</tr>
<tr>
<td>ISMRM</td>
<td>International Society for Magnetic Resonance in Medicine</td>
</tr>
<tr>
<td>HCP</td>
<td>Human Connectome Project</td>
</tr>
<tr>
<td>NRRD</td>
<td>Nearly Raw Raster Data</td>
</tr>
<tr>
<td>FOV</td>
<td>Field Of View</td>
</tr>
<tr>
<td>NIfTI</td>
<td>Neuroimaging Informatics Technology Initiative</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>LiFE</td>
<td>Linear Fascicle Evaluation</td>
</tr>
<tr>
<td>SIFT</td>
<td>Spherical-Deconvolution Informed Filtering Of Tractograms</td>
</tr>
<tr>
<td>FOD</td>
<td>Fiber Orientation Distribution</td>
</tr>
<tr>
<td>CSD</td>
<td>Constrained Spherical Deconvolution</td>
</tr>
<tr>
<td>TD</td>
<td>Track Density</td>
</tr>
</tbody>
</table>
1 Introduction

In order to understand the different functions of the human brain, it is necessary to have knowledge of the connection between different anatomical regions. Diffusion magnetic resonance imaging (dMRI) is an imaging technique that can locally detect orientation-dependent anisotropies in the movement of water molecules in the brain, partly because water molecules diffuse more freely along the direction of neural fibers while being influenced by factors such as macromolecules and cell membranes. Mature white matter tissue is highly anisotropic, meaning that the diffusivity largely depends on the tissue orientation. Anisotropies can be used in diffusion-weighted (DW) fiber tractography algorithms for estimating the most likely paths followed by the neural tracts in white matter tissue. Tractography is the only technique for in-vivo visualization of fiber tracts [1] and it offers insight into behavior and properties of white matter, including the level of functional connectivity between different cortical regions in the healthy and diseased brain. Furthermore, the complex microstructural architecture of white matter has led to the development of a wide range of models for diffusion in neural tissue.

The justification for choosing a specific diffusion model with its own set of heuristics is a non-trivial task, since it should perform under a set of idealized assumptions in conjunction with a tractography algorithm, all factors that may affect the measurement accuracy. Extensive efforts have been made to combine DWI tractography with microstructural imaging in order to obtain a representative view of the connectivity between brain regions and the underlying microstructure of neural tissue [2].

1.1 Clinical Applications of Diffusion MRI

Neural tissue undergoes structural changes during pathological neural development. Much of our understanding of physical changes that take place in neural tissue during pathological states is based on observations gathered from in vivo studies of cerebral ischemia in animal models that share physiological characteristics with humans. dMRI can be used in brain maturation studies or to determine the integrity and organization of white matter fiber tracts in patients with schizophrenia, multiple sclerosis, or Alzheimer’s disease [3, 4].

A useful application of dMRI is detection of changes in primary mechanisms, including relative volume changes in segregated fluid compartments and impedance to water diffusion outside cells, that are associated with a restriction in blood supply due to acute ischemia [5]. Recent investigations have been made into Huntington’s disease [6] and its relation to the topological length of white matter connections within the corpus callosum (CC), a neuroanatomically well-defined region with relatively little fiber dispersion and crossing contamination that links the left and right hemispheres of the brain.
Another important aspect is the use of diffusion imaging techniques for neurosurgical planning. Delineation of white matter fiber tracts can provide complementary information about axonal density, myelination, and other microstructural features in perioperative neurological evaluations. Examples include visualization of white matter fiber tracts during epilepsy surgery [7] using a technique called diffusion tensor imaging (DTI) (see section 2.2.5), or employing DTI on the brain in preoperative planning to avoid unaffected areas in the primary motor cortex during tumor resections. The possibility for clinicians to use MR diffusion tractography to segment courses of fiber tracts prior to surgery and invasive procedures is of great importance and cannot be overstated. However, DTI and tractography require careful validation in order to perform accurately in a wide range of circumstances.

1.2 The Challenge of Ground-Truth Validation

Care has to be taken during data acquisition and processing to ensure adequate image integrity. Validation of results constitutes a vital part of assessing the performance evaluation data of dMRI pipelines [8] and the lack of a diffusion phantom that has the appropriate structural characteristics remains a recurrent issue and a major criticism of diffusion imaging. The limited number of validation studies makes it difficult to distinguish true and false positives and other erroneous results, and the validation of results from tractography algorithms need to be improved in order to increase the specificity and accuracy of measurements when compared to known ground truth anatomy.

When assessing anisotropic diffusion, it is important to consider the spatial resolution, which is set by the size of the imaging voxels. In dMRI the voxels are typically a few millimeters in size, which makes it necessary to investigate axons in aggregate instead of single axon fibers. Because of this spatial resolution limit, a primary source of error for tractography methods is inference of the fiber orientation, which may affect the interpretation of the DW signal [9].

As stated, a thorough understanding of the physical connectivity patterns in the brain is required to explain functional networks. Similarly, in order to avoid errors in connectivity mapping, it is necessary to have a working knowledge of the limitations of tractography methods and to be aware of validation issues that may apply to any part of the fiber reconstruction process. These aspects are also important to consider when assessing hypotheses that rely on the strength of anatomical connections about connectivity between brain regions.

Current methods for fiber tracking and tractography based on DTI have high sensitivity for extracting the main neural fiber bundles from dMRI data. However, the methods have very low specificity, which prevents them from being used extensively in research and clinical applications. Also, the problem of automatically generating both nonexistent and redundant fibers has yet to be resolved and fully eliminated from the processing of DW images.

The challenge of specificity lies in the fact that the complexity of neural microstructure requires diffusion models of axonal spaces and that these models
Generating Ground Truths of Neural Fiber Bundles

have inherent biophysical limitations when handling different rates and orientations of molecular diffusion. It is also important to be aware of tradeoffs involved in improving sensitivity to water diffusion in existing diffusion-weighted imaging (DWI) applications, since robust sensitivity is one of the most significant requirements necessary for obtaining adequate image quality.

The choice of algorithms suitable for tractography is not a be-all and end-all solution to the problem of ground truth validation. In vivo data acquisition parameters also need to be considered. Recent studies have demonstrated that no combination of parameter values and algorithms in dMRI are reliable over all possible imaging situations and that quantitative measures depend on a variety of factors such as noise, hardware, and MRI sequence constraints [10]. Suffice to say, approaches to improve steps involved in decision-making processes associated with different tractography methods are subject to a number of potential pitfalls and should be considered when implementing novel solutions.

1.3 Purpose and Goal

In this work, an exploratory approach to generate ground truths of neural fiber bundles is described. The purpose is twofold. First, sets of artificial DW images of non-crossing fiber regions are generated in order to be used as training sets for supervised machine learning algorithms. The ground truth, or imaging phantom, refers to information derived from the set of artificially generated DW images and can be used to develop a reference model for tractography methods. The synthetic fibers should approximate the predicted 3-dimensional shape of real fibers at a local microstructural level, therefore, it is essential to use known neuroanatomy from real image data when assessing the validity of the models. Second, candidate synthetic DWI datasets based on complex fiber systems are used to predict diffusion measurements by forming a family of optimized datasets from probabilistic tractograms as ground truth. The datasets are generated using crossing and curving fiber models along with so-called linear fascicle evaluation resources and informed filtering of tractograms.

The goal is to construct a realistic ground truth by using a collection of distinct synthetic DW datasets with varying user-defined fiber specifications and MRI pulse sequence parameters with the aim of capturing local in vivo fiber architecture. This thesis lays groundwork for direct validation of results from diffusion imaging by investigating the feasibility of generating synthetic data to improve ground truths of neural fiber bundles.
2 Background and General Theory

2.1 Basic Physics of Nuclear Magnetic Resonance

The spin angular momentum ($I$) and magnetic moment ($\mu$) are magnetic properties of atomic nuclei, where $\mu$ is a function of the charge and spin and is defined as the product of the distance between poles and the pole strength. The net magnetic moment is proportional to $I$ with a proportionality constant known as the gyromagnetic ratio ($\gamma$) with SI units radian per second per Tesla ($\text{rad} \cdot \text{s}^{-1} \cdot \text{T}^{-1}$):

$$\mu = \gamma I \quad (2.1)$$

In their resting state, positively charged protons in thermal equilibrium act like randomly oriented magnetic dipoles, with a net magnetic effect of zero. If an external static magnetic field ($B_0$) is applied, the dipoles align with $B_0$ in a spin-up or spin-down polarized configuration. The spinning protons start to precess around the direction of the static field with a frequency known as the Larmor frequency [11], $\omega_0 = -\gamma |B_0|$. The transverse components of the magnetic vectors ($m_{xy}$) cancel out, resulting in a net magnetization vector $M_z$ along the longitudinal magnetic field ($z$-axis). Energy from an applied radiofrequency (RF) pulse, matching the Larmor frequency and perpendicular to $B_0$, can be absorbed by polarized dipoles and reduce the vector $M_z$. Depending on the strength of the RF pulse, the longitudinal components of the dipoles ($m_z$) can be eliminated, and a transverse magnetization vector ($M_{xy}$) may be established in the $xy$-plane (Fig. 1).

Fig. 1. Time-domain representation of the magnetic moment $\mu$ in a magnetic field and the net magnetization following a 90-degree pulse. The net magnetization vector $M_0$ of the entire spin system is shifted toward the transverse plane from an initial flip angle of zero degrees, effectively canceling out the longitudinal spin components and results in $M_z = 0$. Subsequently, the magnitude and direction of the transverse component of the vector ($M_{xy}$) can be described in the $xy$-plane. The cone-shaped dashed line denotes the proton spin precession at frequency $\omega = \omega_0$ and the associated vector $m = m_{xy} + m_z$ in the static field $B_0$ depends on the field strength of the magnet.
2.1.1 Signal Acquisition and Pulse Sequence Parameters

As $\mathbf{M}_{xy}$ rotates in the $xy$-plane it induces an alternating voltage in the receiving RF coil due to changes in the local magnetic field. The envelope of the signal, or the free induction decay (FID), is proportional to the gyromagnetic ratio, proton density, and the magnetic field strength. The vectors $\mathbf{M}_{xy}$ and $\mathbf{M}_z$ combine to form a vector $\mathbf{M}$. During relaxation the transverse magnetization vector will tend to zero as the longitudinal vector increases and $\mathbf{M}$ will spiral to the $z$-axis. The relaxation is caused by two mutually antagonistic methods of energy loss: the spin-lattice relaxation or $T_1$ recovery, and the spin-spin relaxation or $T_2$ decay.

The spin-lattice, or longitudinal relaxation, depends on the propensity of neighboring molecules to absorb thermal energy from excited dipoles, whereas $T_2$ denotes the relaxation of the transverse magnetic components with or without energy transfer. In addition, spin-spin relaxation mechanisms and magnetic field inhomogeneities result in decay of transverse relaxation or $T_2^*$ decay. This relationship is often expressed as $1/T_2^* = \gamma \Delta B + 1/T_2$ where $\Delta B$ is the magnetic field inhomogeneity [7] across a single voxel.

Other important pulse sequence parameters are the operator-controlled echo time (TE), or the time interval between the RF pulse and the reappearing of the signal, and the repetition time (TR), which is the time between consecutive pulses and echoes [12]. It is important to note that the mapping of $T_1$, $T_2$, and the proton density is governed by TE and TR. dMRI methods require a short $T_2$ and therefore necessitates a minimum TE in order to achieve a sufficiently high image contrast. The brightness of a single pixel is determined by the number of protons present in the voxel, the extent of recovery of $\mathbf{M}_{xy}$ from a previous pulse, and the exponential decay of $\mathbf{M}_{xy}$.

2.1.2 Spatial Encoding

The steps involved in the spatial encoding of magnetic resonance images are slice selection, phase encoding, and frequency encoding. Gradient coils surrounding the RF coil are used to select slices by producing a magnetic field gradient, commonly measured in millitesla per meter (mT·m⁻¹), along $\mathbf{B}_0$ with a specific isocentre, and the resonance frequency of the protons along the $z$-axis can be used to determine the location of the signal. Therefore, protons in the region corresponding to the transmit bandwidth of the RF pulse will generate a signal, and the thickness of the slice can be adjusted by varying the gradient field. The phase, frequency, and amplitude of the signals from each individual tissue voxel are collected and stored in the K-space, the Fourier transform of the complex magnetic resonance image.

Following slice selection, phase encoding is applied, in which another gradient pulse orthogonal to the slice introduces a position-dependent phase shift to the precessing dipoles. A third set of coils is used to produce a signal frequency gradient orthogonal to the slice selection gradient. Analogous to phase encoding, frequency encoding is made possible by adjusting the receive frequency bandwidth.
in order to pass frequencies within a certain range. When all spatial frequencies in a slice have been collected, an image can be reconstructed through the distribution of K-space data.

One of the most commonly used pulse sequences in clinical imaging is the spin-echo sequence [13]. A spin-echo can be produced by applying a 180-degree pulse at half the initial echo time after a 90-degree excitation pulse. The 180-degree pulse refocuses the dephased spins due to magnetic field inhomogeneities, and the resonance signal reappears at TE. An alternative to the spin-echo sequence is the gradient echo sequence, in which a single RF pulse together with a gradient reversal is used to generate a signal. The gradient reversal is achieved using a gradient echo that refocuses the spins that are out of phase, which corresponds to a 180-degree pulse. A shorter TR is allowed as a result of using a weaker RF pulse and therefore $M_z$ is not tipped into the transversal plane. A technique called echo planar imaging uses continual reversal of the polarity of the frequency-encoding gradient and switching of the phase-encoding gradient. Echoes of different phases are obtained either by acquiring the entire phase range in a single TR, or by partitioning the range into periods of equally long repetition times.

2.2 Diffusion Magnetic Resonance Imaging

2.2.1 Molecular Diffusion in Neural Tissue Microstructure

Diffusion is the movement of particles from a region of higher concentration to a region of lower concentration and can be described as a 3-dimensional “random walk” where particles follow thermally driven chaotic paths. According to Fick’s first law of diffusion, the net diffusion flux vector $\mathbf{J}$ is proportional to the concentration gradient $C$ and equals the number of particles per unit area per unit time,

$$\mathbf{J} = -D \nabla C$$

(2.2)

The constant of proportionality, or the diffusion coefficient $D$ (m$^2$·s$^{-1}$), is an intrinsic property of the medium determined by the microstructure of the neighboring tissue. Einstein formulated the concept of diffusion in a probabilistic framework and described the displacement of a given particle as a function of $D$ and the diffusion time $t$ as $\langle x^2 \rangle = nDt$ where $\langle x^2 \rangle$ is the mean squared displacement in $n$ dimensions [14, 15]. In particular, restricted diffusion of water molecules forms a basis for probing the underlying microstructure of neural tissue, which is used in dMRI to measure the random motion of water molecules in white matter.
2.2.2 Sources of Diffusion Anisotropy

The central nervous system is composed of grey matter and white matter. Grey matter constitutes the outer layer of the brain and is composed primarily of neuronal cell bodies and unmyelinated axons, whereas white matter lies in the subcortical and central region of the brain and is made up of mostly long myelinated axons that are bundled into larger tracts or fibers. White matter tracts are of three types: projection tracts running to and from the spinal cord, association tracts connecting different regions of the cerebral hemispheres, and commissural tracts that make up the CC, the largest white matter structure in the brain. See Appendix A for details.

Diffusion of water molecules can be used to track and visualize fiber pathways. Isotropic diffusion is equal in all directions and can be described by a single scalar diffusion coefficient and is more prevalent in grey matter than in white matter as the presence of different cells results in random patterns of diffusion directionality [16]. The principal directions of diffusion in white matter are caused by cellular asymmetric structures and barriers within the bundles of myelinated axons and lead to anisotropic diffusion along longitudinal axes of fiber bundles (Fig. 2).

2.2.3 Axonal Compartments

Biophysical compartment models describe the diffusion of water surrounding the axon of a nerve. Accordingly there are three major types of axonal spaces or compartments: intra- and extra-axonal compartments that describe water diffusion inside and outside of the axons, and isotropic restricted compartments that refer to water in non-axonal structures such as glial cells. The compartmental contributions to the total diffusion is still being investigated, however, several models have been developed for the purpose of characterizing diffusion within the immediate vicinity of axonal environments (see section 2.3.7) in order to separate diffusive processes in nervous tissue.

![Fig. 2.](image)

*Fig. 2.* Microanatomy of white matter myelinated axons showing the major components of the neuronal cytoskeleton. Cytoskeletal structures, hydrophobic myelin sheaths, and the axonal membrane constitute barriers to diffusion across the fiber axis. In the intra-axonal space, the movement of water is essentially unhindered which causes free diffusion to dominate along the longitudinal axis of the axon fiber. The diffusivity parallel to the directional orientation of the axon \( D \) is therefore greater than the transverse diffusivity \( D' \).
2.2.4 Pulsed Gradient Diffusion-Weighted Imaging

In addition to a 90- and a 180-degree RF pulse, a T2-weighted spin-echo sequence can be modified to include a pair of diffusion-sensitizing gradients applied before and after the second pulse (Fig. 3). This is known as the Stejskal-Tanner pulsed gradient spin-echo (PGSE) sequence, and it forms a basis for modern DWI sequences, including a combination of echo planar imaging or gradient echo sequences [17]. The short-duration gradients are applied along the directional axis, which allows the pulse duration and the diffusion time to be clearly distinguished. The signal intensity increases if little or no movement of molecules occurs between the gradient pulses. This only holds true if the spins undergo dephasing and rephasing between the two gradients. If the molecules move, the protons will either dephase or rephase, and the signal intensity will decrease.

![Fig. 3. The Stejskal-Tanner prototype pulse sequence diagram showing two gradient pulses with fixed magnitude G and equal duration δ. The strength of the applied magnetic fields will effectively limit the attainable diffusion weighting (b-value) addition to the MRI pulse sequence. Two RF pulses are shown separated by a time TE/2 followed by a free induction decay signal.](image)

The signal intensity $S_i$ or the amplitude in the frequency domain in a voxel for DW images with gradients applied along the longitudinal axis, is calculated with the Stejskal-Tanner equation [18, 19],

$$S_i = S_0 e^{-bADC} = S'_0 \exp \left[ -(\frac{TE}{T2}) - (\gamma G \delta)^2(\frac{\Delta - \delta}{3})ADC \right]$$

(2.3)

where $b = (\gamma G \delta)^2(\Delta - \delta/3)$ is the b-value or the diffusion sensitivity factor in s·mm$^2$. For PGSE measurements, the diffusion time is determined by $\Delta - \delta/3$ and the term $q = \gamma G \delta$ is called the wave vector (see section 3.4.2). $S_0$ is the T2-weighted signal intensity when $b = 0$ s·mm$^2$, whereas $S'_0$ does not include any spin-spin relaxation effects. The b-value is proportional to the magnitude of the gradient pulses $G$, the gyromagnetic ratio, the pulse duration $\delta$, and the time $\Delta$ between the two RF pulses. The measured diffusion cannot be distinguished from
scales of molecular transport such as membrane permeability or mechanisms involving pressure gradients [19]. Therefore, diffusion in anisotropic materials is estimated by using the apparent diffusion coefficient (ADC). The apparent diffusivity for each gradient direction $i$ in a reference coordinate system thus correspond to

$$\text{ADC}_i = -\ln\left(\frac{S_i}{S_0}\right)/b$$

(2.4)

The amount of diffusion affects the exponential decay of the signal intensity and the resulting anisotropic diffusion along the principal fiber axis can be estimated. A large ADC in a specific gradient direction results in a lower signal intensity, and vice versa (cf. Fig. 4).

![Fig. 4. A set of DW images (top row) and three T1-weighted images (bottom row) from a healthy subject using a b-value of 2000 s-mm$^{-2}$. The images are depicted on the axial, coronal, and sagittal plane (Appendix B) and provided by Franco Pestilli at the Department of Psychology at Stanford University and used in the linear fascicle evaluation software (see section 3.3). Post-processing was done using the Medical Imaging Interaction Toolkit (MITK Diffusion, release 2014.10.02, Medical Imaging Computing, German Cancer Research Center, Heidelberg, Germany).](image)

2.2.5 The Diffusion Tensor

A set of $S_0$ and $S_i$ signals with diffusion gradients applied are required for determining the scalar diffusion coefficient $D$ in a specific direction. One set of baseline $S_0$ images and a minimum of six unique diffusion gradient directions

---

corresponding to six diffusion coefficients $D_{xx}$, $D_{yy}$, $D_{zz}$, $D_{xy}$, $D_{xz}$, and $D_{yz}$ need to be applied for a single voxel in order to calculate the three local diffusion vectors that are necessary to visualize diffusion in three dimensions:

$$ S = S_0 \exp \left\{ \sum_{i=x,y,z} \sum_{j=x,y,z} b_{i,j} D_{i,j} \right\} $$

(2.5)

where $b_{i,j} = \gamma^2 G_i G_j [\delta^2(\Delta - \delta/3)]$. After sensitizing the signal $S$ to diffusion along at least six non-collinear directions and calculating the ADCs, it is convenient to model anisotropic diffusion in terms of a symmetric rank-2 tensor (D) containing the voxel-specific displacements:

$$ D = \sum_{j=1}^{3} \sum_{i=1}^{3} D_{ij} \epsilon_i \otimes \epsilon_j = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} $$

(2.6)

where $\otimes$ denotes the dyadic product of two eigenvectors, $\epsilon_1 = (e_{1x} e_{2x} e_{3x})$, $\epsilon_2 = (e_{1y} e_{2y} e_{3y})$, and $\epsilon_3 = (e_{1z} e_{2z} e_{3z})$. $D_{zz}$ is the diffusivity in the $z$-axis, and $D_{xy} = \epsilon_2^T (D \epsilon_1)$ is the directional diffusion coefficient represented by the correlation between the diffusivities in the $x$ and $y$ direction, and so on. The tensor exhibits conjugate symmetry, meaning that the off-diagonal tensor elements above and below the diagonal elements are equal and they represent correlations between diffusivities along the three axes. The diagonal elements are the diffusivity values along each axis in the measurement frame.

The eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \in \mathbb{R}^+$ and eigenvectors of the diffusion tensor can be used to create an orthogonal reference frame $(x', y', z')$ corresponding to the principal directions of a diffusion ellipsoid (Fig. 5),

$$ D = D^T = E^T \Delta E = \begin{pmatrix} e_{1x} & e_{1y} & e_{1z} \\ e_{2x} & e_{2y} & e_{2z} \\ e_{3x} & e_{3y} & e_{3z} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} e_{1x} & e_{2x} & e_{3x} \\ e_{1y} & e_{2y} & e_{3y} \\ e_{1z} & e_{2z} & e_{3z} \end{pmatrix} $$

(2.7)

The off-diagonal elements in the tensor become zero as a result of diagonalizing the $3 \times 3$ vector matrix [20] and the set of elements in $E$ are orthogonal to each other. In DW image acquisition, the signal intensity in a specific gradient direction $i$ can be expressed as

$$ S_i(b, \mathbf{g}_i) = S_0 \exp(-b g_i^T D g_i) $$

(2.8)

where the vector $g_i = (g_{ix} g_{iy} g_{iz})$ denotes the non-collinear gradient directions. This equation constitutes the basis for estimating diffusion with DW imaging techniques. For the diffusion tensor the diffusion sensitivity factor incorporates both the magnitude and direction of $g_i$, and Eq. 2.3 can be rewritten as
\[ S_i / S_0 = \exp(-b_{xx}D_{xx} - b_{yy}D_{yy} - b_{zz}D_{zz} - 2b_{xy}D_{xy} - 2b_{xz}D_{xz} - 2b_{yz}D_{yz}) \] (2.9)

One approach to estimate the diffusion tensor is to define a matrix \( B \) containing the total number of elements that corresponds to each signal measurement [21, 22],

\[
B = \begin{pmatrix}
    b_{xx,1} & 2b_{xy,1} & 2b_{xz,1} & b_{yy,1} & 2b_{yz,1} & b_{zz,1} \\
    b_{xx,2} & 2b_{xy,2} & 2b_{xz,2} & b_{yy,2} & 2b_{yz,2} & b_{zz,2} \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
    b_{xx,n} & 2b_{xy,n} & 2b_{xz,n} & b_{yy,n} & 2b_{yz,n} & b_{zz,n}
\end{pmatrix}
\] (2.10)

The relationship between \( B \) and \( D \) can be expressed in terms of the logarithmic loss of the diffusion signal, or the total number of log-transformed signal intensities \( L \), where \( D = (D_{xx} D_{xy} D_{xz} D_{yy} D_{yz} D_{zz})^T \):

\[
L = \left( \frac{-\ln(S_1 / S_0)}{b} - \frac{-\ln(S_2 / S_0)}{b} \ldots - \frac{-\ln(S_n / S_0)}{b} \right)^T = BD
\] (2.11)

The tensor \( D \) is calculated by taking the pseudo-inverse of \( B \) or by using the method of least squares to determine \( D = (B^T B)^{-1} B^T L \). See Appendix C for details on the properties of tensors and how to solve for the eigenvalues analytically in order determine the principal directions of diffusivity.

**Fig. 5.** a) The rotationally invariant diffusion values \((\lambda_1, \lambda_2, \lambda_3)\) along a reference coordinate system \((x', y', z')\). For isotropic or spherical diffusion as can be observed in the cerebrospinal fluid and astrocytes, the axial diffusivity or diffusion tendency is such that \( \lambda_1 = \lambda_2 = \lambda_3 \), and the corresponding diffusion tensor \( D = dl \) where the scalar diffusivity \( d > 0 \) is the same in all directions. b) Anisotropic diffusion is represented geometrically by a diffusion ellipsoid where \( \lambda_1 \) indicates the longitudinal direction of neural fiber bundles or the principal fiber axis where \( \lambda_1 \geq \lambda_2 \geq \lambda_3 \).
White matter tract directionality in neighboring image voxels can be estimated after modeling the principal diffusion of water. A common technique used to delineate continuous white matter trajectories in the brain is based on DTI, in which the diffusion signal measurements are fitted to a diffusion tensor ellipsoid within each voxel that assumes a discrete fiber architecture, which is extended to include multiple or continuous fiber configurations. It should be noted that the diffusion tensor model estimates the measurement data as a Gaussian diffusion process and does not accurately reflect local directions of crossing fibers or overlaps within a voxel [2, 23]. The fiber orientation within imaging voxels is thus subject to false positive and false negative dispersion errors as a result of potential interleaved and stacked crossings, and the inherent oversimplification of the diffusive behavior often leads to misconstrued measurements for larger b-values \( b > 3000 \text{ s-mm}^{-2} \), which may result in erroneous signal intensities and other partial volume artifacts.

### 2.2.6 Deterministic Tractography

Tractography is a non-invasive technique for visualizing the connectivity between gray matter regions and the organization of white matter fiber pathways by employing deterministic or probabilistic algorithms based on predetermined diffusion models. Deterministic or streamline algorithms operate under the assumption that there exist a linear relationship between the directions of the principal eigenvectors, such that the orientation of the fibers is collinear with the principle fiber axes in each image voxel. The streamline approach to depicting fiber trajectories in regions with limited fiber dispersion is quite intuitive since one expects to find cohesive bundles of fibers distributed in the examined volume (see section 3.4.2).

In the tensor model, tractography streamlines are tangent to the velocity vector of the principal diffusion orientation or the vector field (Fig. 6a). The instantaneous change in position of the streamline as a function of the arc length \( s \) is equal to the first eigenvector of the diffusion tensor, which is a function of the 3-dimensional streamline coordinate \( r \) [24, 25]. The instantaneous displacement at an \( n \)th position is thus expressed as a differential equation, \( dr(s_n)/ds = e_1[r(s_n)] \), where the tangent vector \( t(s_n) = e_1[r(s_n)] \) and the starting position, or seed region, follows from \( r(0) = r_0 \). The direction of the displacement for the entire set of spatial coordinates \( \{r(s_n)\} \) corresponds to \( r(s_{n+1}) - r(s_n) \). Streamline algorithms may produce false negative results when determining trajectories of single bundles by miscalculating the dispersion of fibers. In addition, streamline approaches does not take into account uncertainties in voxel amplitudes and can therefore not be used to produce maps of connectivity.
2.2.7 Probabilistic Fiber Tracking and Uncertainty

Fiber connectivity is commonly visualized by 3-dimensional scalar maps or color-coded orientation distributions of the principal eigenvectors of the diffusion tensors, where each color represents a specific orientation. Distributed connectivity can be visualized with probabilistic, or stochastic tractography, where the most likely fiber orientation is calculated and traced in each voxel. Starting from a seed region of interest (ROI), one can generate path probability maps of white matter fiber measurements that overlap across datasets [2] and that can be used to estimate the spatial distribution of fiber pathways (see section 3.3.1). The technique, which includes stochastic Monte Carlo streamlines and simulated random walks, can be used as an alternative to deterministic streamline approaches to account for uncertainties, errors, and oversimplifications associated with traditional deterministic tractography algorithms (cf. Fig. 7).

Even though stochastic diffusion tractography enables more robust tracking results compared to deterministic algorithms at the expense of higher computational cost, the quality of probabilistic fiber tracking is altered in locations within the brain that are characterized by complex fiber architecture.
2.2.8 Diffusion Tensor Imaging Scalar Indices

The diffusion tensor model is also useful for calculating invariant scalar indices that reflect the microstructural tissue integrity, including the fractional anisotropy (FA) and mean diffusivity \( \langle D \rangle = \text{Tr}(D)/3 \),

\[
\text{FA}(D) = \sqrt{\frac{3}{2} \left( (\lambda_1 - \langle D \rangle)^2 + (\lambda_2 - \langle D \rangle)^2 + (\lambda_3 - \langle D \rangle)^2 \right)} / \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}
\] (2.12)

FA(D) has a real value between zero and one and can be used as an indicator of the extent of intravoxel anisotropy or the homogeneity of the direction of diffusion. Increased radial diffusivity \((\lambda_2 + \lambda_3)/2\) corresponds to a lower FA(D) value (Fig. 5a) while an increased axial diffusivity results in a higher fractional anisotropy (Fig. 5b) and is found in the bulk of the deep parts of the brain (cf. Fig 8).

Another way of interpreting information about isotropic and anisotropic diffusion provided by the tensor is to use the relative anisotropy or RA(D), which is defined as the ratio between the magnitude of the anisotropic and isotropic tensor components in terms of the variance of the eigenvalues to their mean value and vary between zero (isotropic diffusion) and \( \sqrt{2} \) (infinite anisotropy),

\[
\text{RA}(D) = \sqrt{\frac{1}{3} \left( (\lambda_1 - \langle D \rangle)^2 + (\lambda_2 - \langle D \rangle)^2 + (\lambda_3 - \langle D \rangle)^2 \right)} / \langle D \rangle
\] (2.13)
Fig. 8. Fractional anisotropy maps of three image slices at progressively deeper levels of the brain in the axial, sagittal, and coronal plane. The slices were chosen to reflect diffusion, differentiation between anatomic features, and microstructural integrities. The FA is used as a diffusivity measure to indicate organized structures like white matter pathways with a bright appearance, while capitalizing on the fact that isotropic diffusion yields darker image areas. The set of images was obtained from a white-matter atlas provided by the Laboratory of Brain Anatomical MRI at Johns Hopkins University [26].

2.3 Related Work

2.3.1 Validation of Tractography

Addressing difficulties associated with validation of fiber tractography is reflected in both the scientific and professional literature on dMRI. Many methods have been developed in order to identify fiber crossings. So far, however, only up to three crossings in the large centrum semiovale white matter mass in the brain have been identified [27]. Researchers have tried to resolve validation issues associated with confounding factors such as imaging noise and complex fiber configurations and several methods have been developed to directly compare results obtained from a variety of simulation approaches. This includes the use of physical phantoms for evaluating tracking algorithms and methods utilizing known fiber structures to predict algorithm behaviors [28, 29].
2.3.2 Standardization and Surgical Dissection Techniques

Biological phantoms have yet to provide reliable ground truth information for neural architecture in DW imaging applications and even if hardware phantoms can provide ground truths for well-defined fiber geometries, they are not applicable for complex fiber configurations [30, 31]. Several indirect approaches have been suggested as a complement to direct validation by employing functional imaging to reveal physiological activities within tissues in conjunction with results from tractography.

Combined evaluations of probabilistic tractography, functional imaging, and histological studies have demonstrated the validity of brain parcellation approaches. Johansen-Berg and Rushworth [32] found a high correlation between tractography-based parcellation and functional relevance after analyzing functional imaging activations within the thalamus during motor tasks.

There is no standardized method for evaluating DTI registration performance and few of the studies to date shed light on this issue [33]. In 2015 Pujol et al. proposed an approach for devising a more objective way of evaluating DTI tractography by initiating the DTI Challenge [34], an international effort aimed at gaining further understanding into the evaluation of DTI tractography for neurosurgical applications in the absence of ground truth. It was found that the interalgorithm variability was pronounced in the unaffected hemisphere and tumor tissues under the hypothesis that the purpose of the methods presented in the article should yield the same or similar results. Surgical dissection provides ways of comparing results obtained from tractography with fiber tracts in actual brain tissue, which in turn makes it possible to create atlases of white matter pathways and anatomical compartments using DTI streamline tractography [35]. However, experimenter-subjectivity may be of concern and should not be neglected, since the separation of fiber clusters and bundles into coherent fibers relies on careful and meticulous dissection and requires comprehensive understanding of the surrounding tissue structure.

2.3.3 Physical Phantoms and Ground-Truth Control

Many assumptions are made during the fiber reconstruction process, and as a result, errors may degrade the perceived image quality and can adversely affect the tractography results, irrespective of the tractography algorithm being used. Several imaging phantoms have been constructed with the intention to reliably compare results to a known ground truth, and to evaluate different aspects of data acquisition in order to minimize the number of errors present in tractography outputs. Compared to simulated synthetic fiber-based phantoms developed to mimic real fiber trajectories, physical or ex vivo phantoms are able to provide a more realistic ground-truth control by taking advantage of simple geometrical constructions and enabling monitoring of properties that affect the axial diffusivity.
2.3.4 Capillary-Based Phantoms

The characteristic cylindrical shape of axons can be accurately matched by a capillary-based phantom, which is a type of synthetic physical phantom. In 2002 Von dem Hagen et al. [36] compared theoretical results for diffusion in cylindrical fiber bundles with DW measurements in samples with different fiber orientations. This was the first recorded use of water-filled 50 µm polytetrafluoroethylene (PTFE) capillaries to analyze the results of orientational measurements of diffusion by looking at the magnitudes of ADCs as a function of varying b-values [37]. Depending on the orientation, they were also able to investigate the angular orientation-dependence of the ADC.

Despite the fact that the diffusive properties and permeability of PTFE capillaries do not accurately reflect the natural process of intra- and extra-axonal diffusion in neural fiber bundles, Lin and colleagues [38] were able to conduct the first voxel-scale comparison on a physical phantom in order to study sheets of crossing fibers. In addition, trials investigating the use of glass capillaries have been conducted using DTI to study the principal direction of diffusion [37].

Regenerated cellulose has been used to create synthetic fiber-based phantoms with realistic spin-spin relaxation times that can overcome limitations prevalent in PTFE capillaries. Fiber-based phantoms that contain regions of different fiber configurations have been developed, which makes it possible to vary the FA for a selected region of interest [39, 40].

2.3.5 Software Phantoms

The inherent limits of predicting and characterizing real DWI data solely based on simulated datasets makes it a difficult task to develop software phantoms with realistic modeling for validation of tractography, since it is not feasible to fully describe the in vivo fiber architecture [41]. Software phantoms are excellent compared to physical and synthetic fiber-based phantoms for evaluating singled-out characteristics of fiber reconstructions.

The ability to detect both systematic and random errors more easily is but one example of the flexibility of software phantoms. Other adjustable parameters include fiber radii, curvatures, and noise artifacts. Phantoms that omitted fiber crossings at microstructural levels within image voxels was early on proven to be insufficient and were shown in many cases to yield completely invalid results. As demonstrated by Gössl et al. [42], reconstruction techniques should incorporate methods for handling fiber crossings and more complex fiber configurations in order to avoid erroneous results.

Variants of streamline algorithms dominated the early forms of simulated phantoms for tractography. In recent years, researchers have been able find non-streamline approaches to validation of tractography with software phantoms.
2.3.6 Reconstruction of Neural Fibers

96 state-of-the-art tractography pipelines submitted by 20 separate groups of researchers were evaluated during the International Society for Magnetic Resonance in Medicine (ISMRM) 2015 Tractography Challenge [43]. The objective of the challenge was to reconstruct neural fiber pathways in a typical DW image using a ground truth dMRI dataset. The challenge results provided insight into the difficulties related to voxel-aligning streamlines with ground-truth anatomy.

The ground truth dataset was based on 25 segmented fiber bundles acquired from the Human Connectome Project (HCP), an effort funded by the National Institutes of Health that aims to characterize functional and anatomical brain connectivity in healthy adults using dMRI to chart white matter tracts [44]. The HCP diffusion protocol (see section 3.2.2) enabled spatial resolution improvement of the MRI acquisition to a voxel size of 1.25 mm, which opens up the possibility for sub-voxel imaging by using so-called local multi-voxel spatial models.

2.3.7 Multi-Compartment Diffusion Models

Panagiotaki et al. [45] introduced a diffusion-modeling concept that combined compartment-specific models to produce DW images of a rat brain by taking into account the unweighted signal intensity from each of the three axonal spaces to estimate the DW-MRI signal in each voxel. Candidate models for the intra-axonal compartment included the “stick” model, which describes the diffusivity and principal direction of a zero radius cylinder [2], and the “cylinder” model that describes the diffusion as a non-zero radius cylinder. For the extra-axonal compartment, the tensor, “ball”, and “zeppelin” model were analyzed.

In the isotropic ball model, the scalar diffusivity is proportional to the identity tensor, whereas the zeppelin model describes the cylindrically symmetric anisotropic diffusion according to its principal and perpendicular direction [46]. Four models where evaluated for the restricted diffusion of water in other cells, namely the “astrosticks”, “astrocylinders”, “sphere”, and “dot” models (Appendix D). It was shown that intra-axonal restriction models accurately describe the diffusion process and should be prioritized over biexponential or diffusion tensor models. Results obtained with 3-compartment models estimates outperformed 2-compartment measurements of the DW signal, and a biophysical model combining the tensor, cylindrical, and spherical compartment proved most accurate.

In a taxonomy of compartment models using in vivo human brain data, Ferizi et al. [47] demonstrated, with broadly reproducible results, that 3-compartment models for multi-b-value measurements were overall more accurate than 2-compartment models. The authors also concluded that three axonal compartments are necessary for capturing microstructural information about restricted and non-restricted diffusion in the CC.

Because synthetic datasets provide both a reference dataset and DW-MRI image data under a set of specified signal simulation parameters, they are ideal for controlling confounding factors and variables that may otherwise influence in vivo
investigations. Unfortunately, the majority of techniques developed to simulate white matter fiber bundles and trajectories are limited to simple fiber configurations and do not resolve diffusion signals in regions of complex fiber structures. Other techniques neglect effects associated with tissues surrounding the primary source of diffusion, such as partial volume contaminations or changes in the fiber density [48].

As mentioned in the previous section, accurate compartment models are required for assessing and interpreting results obtained with different dMRI techniques. Likewise, it is important to incorporate multi-compartment models in simulations in order to develop techniques that are able to overcome the aforementioned drawbacks, including the ability to realistically depict fiber crossings and regions of white matter tracts that cannot be visualized with crude methods of simulation.

2.3.8 Simulation Tools for Generating Synthetic Fibers

The open-source framework Fiberfox [30] was recently proposed as a simulation tool for generating synthetic white matter fibers and corresponding DW images. These features were added as a software component to MITK, which is a software for medical image analysis that enables processing and visualization of DW images [49]. The Fiberfox framework for signal generation utilizes a combination of the aforementioned compartment models and the 3-compartment model introduced by Panagiotaki and colleagues was extended to include each of the corresponding volume fractions \( f_i \) relative to the user-specific voxel grid and fiber radius. The multi-compartment signal \( S = \sum_{i=1}^{3} f_i S_i \) where \( f_1 + f_2 + f_3 = 1, \, 0 \leq f_i \leq 1 \) was generalized to \( N \) compartments so that the artificial signal \( S_v = \sum_{i=1}^{N} f_{v,i} S_{v,i} \) for a voxel \( v \) in a gradient direction \( i \).

After placing a set of fiducials (Fig. 9a), specifying the fiber bundle properties, and setting the artificial dMRI parameters including voxel size, b-value, and the number of gradient directions, the DW image can be generated by simulating the signal intensity \( S_v \) in each voxel. The result is a complete DW-MRI dataset including metadata and voxel-specific gradient directions. The distribution of individual fibers can either be Gaussian or uniform (Fig. 9b) and it is possible to extract the 3-dimensional organization of individual fibers (see section 3.3.1).

The ground truth described in section 2.3.6 was generated in Fiberfox with a clinical-like dMRI dataset consisting of 32 gradient directions and \( b = 1000 \, \text{s mm}^{-2} \). Motion, distortion, and noise artifacts were also included in the simulated isotropic diffusion data acquisition. A brain connectivity approach was used where the analysis was centered on unexpected relationships between brain regions for the purpose of assessing the validity and accuracy of tractography pipelines, and connectivity criteria were introduced so that true, false, and plausible connections could be clearly distinguished from one another in the evaluation of the reconstruction process [7].
2.3.9 Prevalence and Geometry of Sheet Structures in the Brain

There is no scientific consensus on whether or not sheet-like structures exist in the brain. One of the main reasons why a viable theoretical framework has not been developed is the absence of reliable conditions for sheet-like structures to exist [50]. The characterization of the sheet structure may require assertions that can affect how the sources of measurement uncertainty can be identified and managed. Assuming, for instance, that fiber pathways cross each other at nearly right angles, may drastically affect the consistency of validity and reliability measures across different regions of interest. Even though no formal description of a sheet has yet
been agreed upon by the scientific community, valuable insights and observations have been brought together to provide qualitative results that can be used to indicate the degree to which a sheet structure correlates with fiber orientations in the brain [50].

A multitude of studies have been conducted where comparisons have been made between in vivo measurements of fiber directions and different interpretations of sheet structures. Examples include sheets of single-fiber direction configuration in the CC represented as sheet-like skeletons [51]. By analyzing cerebral fiber pathways in terms of crossings and adjacencies, Wedeen et al. [52] showed that fiber pathways form a distinct 3-dimensional grid. Building on the qualitative results obtained from this investigation, Tax and colleagues reported preliminary research findings at the ISMRM in 2015. They later proposed a theoretical framework that formalized the description of a sheet by introducing a so-called sheet probability index. Using different simulated dMRI datasets together with in vivo measurements and the concept of a sheet tensor, the authors were able to investigate the extent of consistent sheet structures and how they were oriented in different regions in the brain.

If the existence of sheet-like structures was to be properly verified it could impact the way in which both structural and functional connectivity studies and any subsequent microstructural measures are performed. Investigations into the structure and spatial composition of sheets are still subject to validation and assessment issues outlined in the previous sections about ground-truth control.
3 Methods

3.1 Synthetic Data Acquisition

3.1.1 Fiber Definition and Configurations

Three levels of longitudinal fiber configurations were constructed in Fiberfox in order to study synthesized signal attenuation from non-crossing fiber regions superimposed on a baseline T2-weighted dummy image. The configurations were set to increase in geometric complexity while retaining a constant fiducial marker radius \( r \) with respect to a voxel center \( v_{x,y,z} \) across the set of generated images.

The first configuration involved defining the distribution of white matter fibers in a single cylindrical ground truth fiber bundle placed in the center of a representative \( 3 \times 3 \times 3 \) mm volume element (Fig. 10a) in order to reflect local white matter architecture. The 3-dimensional spatial domain of the simulated element was described in a rectilinear Left Posterior Superior (LPS) anatomical coordinate system, which is a patient-based measurement frame defined by basis vectors along the axes of left-right, anterior-superior, and inferior-superior (Appendix B).

Interpolation between consecutive fiducial points was controlled by the tangential properties of the fiber curvature and the bias, or the direction of the tangent vector. The fiducials were displaced parallel to the direction of the fibers, thereby reducing the number of sampling points needed to model individual fibers. The corresponding cylindrical fiber bundle was therefore exclusively characterized by low fiber curvature with consecutive fiducial center points \( O \) aligned to the voxel grid and contained uniformly distributed fibers \( f_{x,y,z} \) over the cylindrical cross-sectional area (Fig. 10d).

The next configuration was constructed as a local level interpretation of the concepts presented by Tax C et al. [50] by copying the initial cylindrical fiber bundle along with its fiber definition parameters. This bundle was repeatedly shifted laterally along the \( x \)-axis, resulting in a sheet of parallel bundles going across the centerline of the representative volume element (Fig. 10b) with a distance of approximately \( 2r \) between neighboring fiducial center points.

The third and final configuration was created by joining the cylindrical fiber bundles in the fiber sheet and adding pairs of copied sheets to the central voxel in the volume element, forming an array of closely packed microscopic fiber bundles (Fig. 10c).
Fig. 10. Schematic depicting the cross-sectional view of the three geometrical configurations of fibers at a macroscopic length scale. The topological skeletonization of a single fiber bundle is portrayed as a straight line going through two consecutive fiducial center points in the geometric center of the representative volume element. The single-bundle ground truth (a) is converted into several adjacent bundles, forming a sheet of contiguous fiber bundles (b) and an additional space-filling array of close-packed bundles (c). d) Uniformly or Gaussian distributed fibers $f_{x,y,z}$ are generated in the disc at a specified fiber density.

3.1.2 Diffusion-Weighted Signal Simulation and Preprocessing

An entire range of phase-encoding steps was registered in a single TR by using the Fiberfox framework to perform a simulated spin-echo echo planar imaging sequence corresponding to a single-coil acquisition with an anterior-posterior phase encoding and a constant coil sensitivity.

No preprocessing of the synthetic data was necessary. Any distributed noise with a variance to the DW signal was not included in the simulations. Ghosting, distortion, motion, or Gibbs-ringing MRI artifacts were not generated during signal acquisitions. Images with eddy-current distortions were discarded from the synthetic datasets to avoid DW images with spatially linear eddy current profiles along the applied diffusion-sensitizing gradients.

The pulse sequence parameters were adapted according to data available from the WU-Minn HCP diffusion protocol [53, 54], which allowed acquisition of high-quality in vivo dMRI data at a voxel size of 1.25 mm. TE and TR were set to 89.5 ms and 5520 ms, respectively. The inhomogenous relaxation was specified by a $T2' = \gamma \Delta B$ relaxation time of 50 ms. Two sets of 90 diffusion directions were used in the simulations: one was randomly generated in Fiberfox, and the other contained uniformly distributed directions generated in the HCP diffusion protocol.
Three different diffusion sensitivity factors \( b \in \{1000, 2000, 3000\} \ \text{s-mm}^{-2} \) were used in the simulations. The signal attenuation in a given gradient direction \( i \in [1, 90] \) corresponded to a matrix \( \mathbf{G} \) of gradient vectors \( \mathbf{g}_i \),
\[
\mathbf{G} = \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_{90} \end{pmatrix} = \begin{pmatrix} g_{1,x} & g_{1,y} & g_{1,z} \\ g_{2,x} & g_{2,y} & g_{2,z} \\ \vdots & \vdots & \vdots \\ g_{90,x} & g_{90,y} & g_{90,z} \end{pmatrix}
\]
(3.1)

Table 3.1. Diffusion gradient table describing the diffusion encoding direction for the first 10 out of 90 gradient vectors at 1000 s-mm\(^{-2}\). Gradients generated in MITK (left) and a partial set of the uniformly distributed non-collinear diffusion directions provided by the WU-Minn HCP diffusion protocol (right) corresponding to a set of unique arrangement of points.

<table>
<thead>
<tr>
<th>( i )</th>
<th>( g_{x,i} )</th>
<th>( g_{y,i} )</th>
<th>( g_{z,i} )</th>
<th>( g_{x,i} )</th>
<th>( g_{y,i} )</th>
<th>( g_{z,i} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-0.939</td>
<td>0.292</td>
<td>0.181</td>
</tr>
<tr>
<td>2</td>
<td>-0.030</td>
<td>0.146</td>
<td>0.989</td>
<td>0.204</td>
<td>0.893</td>
<td>-0.401</td>
</tr>
<tr>
<td>3</td>
<td>-0.209</td>
<td>0.023</td>
<td>0.978</td>
<td>-0.276</td>
<td>-0.340</td>
<td>-0.899</td>
</tr>
<tr>
<td>4</td>
<td>-0.156</td>
<td>-0.204</td>
<td>0.966</td>
<td>-0.219</td>
<td>0.890</td>
<td>0.398</td>
</tr>
<tr>
<td>5</td>
<td>0.070</td>
<td>-0.287</td>
<td>0.955</td>
<td>-0.455</td>
<td>-0.246</td>
<td>0.856</td>
</tr>
<tr>
<td>6</td>
<td>0.284</td>
<td>-0.166</td>
<td>0.944</td>
<td>-0.875</td>
<td>-0.325</td>
<td>-0.357</td>
</tr>
<tr>
<td>7</td>
<td>0.352</td>
<td>0.073</td>
<td>0.933</td>
<td>-0.521</td>
<td>0.568</td>
<td>-0.637</td>
</tr>
<tr>
<td>8</td>
<td>0.019</td>
<td>0.413</td>
<td>0.911</td>
<td>0.796</td>
<td>0.547</td>
<td>-0.259</td>
</tr>
<tr>
<td>9</td>
<td>-0.232</td>
<td>0.370</td>
<td>0.899</td>
<td>-0.494</td>
<td>0.419</td>
<td>0.762</td>
</tr>
<tr>
<td>10</td>
<td>-0.416</td>
<td>0.194</td>
<td>0.888</td>
<td>-0.443</td>
<td>-0.834</td>
<td>-0.322</td>
</tr>
</tbody>
</table>

The length of each gradient vector was confirmed by computing the Euclidean norm \( ||\mathbf{g}_i|| = \sqrt{g_{x,i}^2 + g_{y,i}^2 + g_{z,i}^2} \) = 1. Excluding the 10% automatically generated baseline images where \( b = 0 \ \text{s-mm}^{-2} \), a total number of \( 90 \times 3^3 \) signals in each gradient direction were registered and collected in a set that contained 2430 elements, \( \{S_1, S_2, \ldots, S_{2430}\} \).

No additional linear transformation was added to the voxel signals and any potential scaling scenario was omitted from the list of simulation parameters. By placing two longitudinally oriented waypoints in the voxel grid when given a center point in the middle of the volume element, a set of local level signals \( \{S_1', S_2', \ldots, S_{270}'\} \subseteq \{S_1, S_2, \ldots, S_{2430}\} \ \forall i \) was derived from the full DW dataset. After removing the artificial baseline images, the set was parsed into a column vector \( \mathbf{S} := (S_1, S_2, \ldots, S_{90}) \) in MATLAB (Release 2016b, The MathWorks, Inc., Natick, Massachusetts, United States) that contained all signals in the central image voxel.
The synthetic DW image volume was imported into MATLAB using the *nrrdread* function\(^2\) to read the generated datasets from the specified Nearly Raw Raster Data (NRRD) file format. The function outputs the voxel-wise signal intensities as a multidimensional array together with a structure array containing associated metadata, including fields for the image dimensions, gradient directions, b-value, and the 3-dimensional coordinate system for the simulation. The volume fractions of single compartments were ignored in order to evaluate the effect of fiber density and fiber radius on partial volume effects, or tissue organization heterogeneity. Keeping in mind the consistently occurring patterns of regional differences in fiber size in the CC, the fiber radius was used as a tool to estimate the compartment volume fractions and was set to automatically fill voxels and to realistically model the axonal radii.

Disabling of partial volume effects to mimic signals from voxels embedded inside white matter bundles allowed simulation of effects between distinct tissues: a voxel would be entirely made up of an intra-axonal compartment or no axonal compartment would fill the cross-sectional imaging space, making it possible to study compartment exchanges and the level of anisotropy in the central image voxel for each configuration.

### 3.1.3 Compartment Model Selection

Initial simulations of the synthetic DW images were based on a modified 2-compartment multi-tensor ball-and-stick model that described free diffusion of water and multiple populations of fibers. The simulated T2 relaxation time was set to 110 ms and 80 ms for the stick and ball model, respectively. Both model settings included a scalar diffusivity of 1000 \(\mu \text{m}^2 \cdot \text{s}^{-1}\). For the stick model this refers to the diffusion along the principal axis of the stick, whereas the ball model is described with a single diffusivity parameter and represents free water diffusion as an isotropic tensor \(D\).

In DTI the dispersion of water molecules is modeled with a Gaussian probability distribution \(p\) as a function of the displacement \(x\). In multi-tensor models, however, the Gaussian distribution is replaced with a set of \(N\) Gaussian densities that describe the populations of fibers, and no exchange takes place between the different populations contained in a single image voxel. The contributions \(\{p_1, p_2, \ldots, p_N\}\) corresponding to a specific diffusion tensor for every population in a voxel are expressed as a sum and reflects the probabilistic dispersion,

\[
p(x) = \sum_{i=1}^{N} f_i G(x, D_i, t) \tag{3.2}
\]

where \(f_i\) is the volume fraction of a specific fiber population, \(G\) is the Gaussian function, and \(t\) denotes the diffusion time. The ball-and-stick model aims to

---

characterize separate fiber populations with different tensors and resolve regions of crossing fibers [55]. For our purposes the model was used to estimate and output volume fractions in Fiberfox. However, the model assumes no radial diffusion so the diffusion profile does not change across fiber populations, making it biased towards tissues with anisotropic tendencies.

Following the ball-and-stick simulations, the synthetic fibers were generated with a combination of three diffusion models for the isotropic restricted space and for the intra- and extra-axonal compartments. For diffusion models with two or more compartments, the following general expression for the signal intensity is valid with an implicit assumption of uniform fiber distribution:

\[
S = S_0 \sum_{m=1}^{N} f_m^{IA} S_m^{IA} + f_n^{IR} S_n^{IR} + \left( 1 - f_n^{IR} - \sum_{m=1}^{N} f_m^{IA} \right) S_{EA} 
\] (3.3)

The index \( m \) corresponds to a specific intra-axonal model and \( f_m^{IA} \) is the volume fraction for the same model. \( f_n^{IR} \) denotes the volume fraction relative to the voxel grid and fiber radius for the isotropic restricted model and is determined by one of three restricted models \( n \). Fiberfox only allows for one additional isotropic restricted compartment, therefore Eq. 3.3 was modified to take into account this constraint, which resulted in a simplified form of the expression,

\[
S = S_0 \left[ f_{IA} S_{IA} + f_i^{IR} S_i^{IR} + f_j^{IR} S_j^{IR} + \left( 1 - f_i^{IR} - f_j^{IR} - f_{IA} \right) S_{EA} \right] 
\] (3.4)

where the intra- and extra-axonal compartments are chosen from one out of three options in addition to three possible combinations \( i, j \in [1, 3] \) of restricted compartment models. Refer to the tabulated forms of the compartment models in Appendix D for more details.

### 3.2 Real Datasets

#### 3.2.1 Ethics Statement

The HCP Open Access Data provide imaging datasets from a large population of healthy subjects. The diffusion imaging datasets were used in accordance with the HCP open access data use terms that implemented a two-tiered plan for data sharing to maintain the confidentiality of the data [56]. One can refer to the HCP website for detailed information (www.humanconnectomeproject.org).

#### 3.2.2 Hardware, Protocols, and Preprocessing

Three gradient tables of 90 diffusion directions (cf. Fig. 11) and six baseline acquisitions were obtained in each of the six different runs on a 3T MRI scanner
with 100 mT·m\(^{-1}\) gradients for diffusion encoding. In addition to the WU-Minn HCP diffusion protocol parameters mentioned in section 3.1.2, the HCP MR imaging protocol covered other basic scan parameters, including a flip angle of 78°, a field of view (FOV) set to 2010 \times 180 \text{ mm}, a slice thickness of 1.25 mm and 111 slices, and an echo spacing of 0.78 ms which affected the acquisition time and gradient slew rate. Multiple phase-encoding directions were used by the HCP in order to reduce the level of noise and magnetic susceptibility artifacts. Head-motion and eddy-current corrections were made by simultaneously estimate the artifact effects at different gradient diffusion directions [57].

![Fig. 11. Applied diffusion-sensitizing gradients in 90 directions at a diffusion sensitivity factor of 1000 s·mm\(^{-2}\). a) The collection of generated gradient directions in the half sphere used to simulate diffusion signals across different compartment models in Fiberfox. b) Gradient directions distributed uniformly over the unit sphere obtained from the WU-Minn HCP diffusion protocol.](image)

A single voxel was isolated and evaluated together with the generated DW signals after manually segmenting an atlas-derived ROI within the CC from a HCP subject-specific dataset. This was achieved by processing a large collection of signals emanating from the geometrical center of the in vivo ROI. The collection of signals was assessed in double precision in MATLAB and was intended to function as a source of diffusion with data reminiscent of CC fiber architecture.

### 3.3 Linear Fascicle Evaluation of Synthetic Data

#### 3.3.1 Stochastic Streamline Tractography

As a complementary approach to generate raw and viable artificial fiber bundles, a set of synthetic fibers was created in Fiberfox that served as ground truth fibers for stochastic tractography reconstruction of white matter pathways. A raw DW image was generated from these fibers by using parameter settings from the WU-Minn HCP diffusion protocol and the same method as outlined in section 3.1.2 for constructing cylindrical fiber bundles. The seed polygonal ROI was specified at the
beginning of the fiber bundle, and the number of seeds per voxel was set to a value exceeding the predefined seed density by several orders of magnitude, while limiting the tract length to 100 mm and allowing a maximum memory cache size of 1 Gb. The distance between two consecutive sampling points along a single pathway was kept at a one-to-one relationship with the voxel grid and the spatial coordinates of the pathways in the tractogram were exported from MITK and imported into MATLAB using the readVTR\(^3\) function. The coordinates of the sampling points were stored in a multidimensional matrix and specified by the number of synthetic white matter tracts and sampling points along the longitudinal axis of the ground truth fibers (Appendix E3).

### 3.3.2 Linear Fascicle Structure Setup

The method of evaluating the tractography reconstruction of the synthetic fibers necessitated the use of a framework for evaluation and statistical inference for connectomes created by Pestilli et al. [1] called Linear Fascicle Evaluation (LiFE). The LiFE software was implemented in MATLAB and depended upon the VISTASOFT software repository of the VISTA lab\(^4\) at Stanford University in order to perform analyses on the dMRI data.

The LiFE algorithm used the orientation of the artificially generated white matter neural tracts, or fascicles, and the corresponding raw DWI as inputs and returned an estimation of the diffusion measurements after eliminating unsubstantiated and redundant tractography streamlines. In order to predict the accuracy of the synthetic datasets when compared to the output diffusion data, a set of linear equations were solved in order to obtain the weights \(w_f\) that were used to predict the diffusion signal,

\[
\begin{pmatrix}
S_1(\theta) - I_{1,v} \\
S_2(\theta) - I_{2,v} \\
\vdots \\
S_n(\theta) - I_{n,v}
\end{pmatrix}
= \begin{pmatrix}
O_{1,1}(\theta) & O_{1,2}(\theta) & \cdots & O_{1,m}(\theta) \\
O_{2,1}(\theta) & O_{2,2}(\theta) & \cdots & O_{2,m}(\theta) \\
\vdots & \vdots & \ddots & \vdots \\
O_{n,1}(\theta) & O_{n,2}(\theta) & \cdots & O_{n,m}(\theta)
\end{pmatrix}
\begin{pmatrix}
w_{f,1} \\
w_{f,2} \\
\vdots \\
w_{f,n}
\end{pmatrix}
\]  

(3.5)

where the mean signal \(I_v = (1/N_\theta)\sum_\theta S_v(\theta,b)\) in \(N_\theta\) directions was subtracted from the measured diffusion signal \(S(\theta,b) = S_0 \exp[-bADC_f(\theta)]\) in a voxel filled by a single fiber \(f\). The predicted voxel-specific signal \(S_v(\theta,b)\) was modeled as a weighted sum of the diffusion signals from the set of fibers \(f \in v\), including signals from isotropic compartments:

---


\[ S_v(\theta, b) = I_v + \sum_{f \in v} w_f O_f(\theta) \] (3.6)

\[ O_f(\theta) = S(\theta, b) - \left(1/N_\theta\right) \sum_\theta S_v(\theta, b) \]

is an anisotropic function of the fiber and described fiber-specific modulation around the mean diffusivity. Together with the weight \( w_f \) it constituted the orientation-dependent contributions from the fibers in the compartment. The non-negative weights in Eq. 3.5 were calculated by minimizing the errors for the estimated signals in the ROI that was used to track the white matter fibers:

\[
\arg\min_{w_f} \sum_v \sum_{\theta} \left[ S_v(\theta, b) - I_v - P(\theta, v) \right] \] (3.7)

where \( P(\theta, b) = \sum_{f \in v} w_f O_f(\theta) \) was subtracted from the demeaned signal \( S_v(\theta, b) - I_v \) and describes the predicted modulation of the signal from the fiber.

### 3.3.3 Establishment of Crossing and Curving Fiber Models

A set of candidate DW datasets were constructed in Fiberfox using a single-bundle fiber configuration as ground truth (cf. Fig. 10a) with the default 2-compartment model. The bundle geometry was copied and newly added bundles were rotated about the center point of the original fiber configuration. By default, the geometry of the modified bundles was associative: any changes made to the fiber definition or simulation settings were applied to the entire set of fiber bundles \( \{F_1, F_2, \ldots, F_N\} \) in the image volume.

The sets were generated in order to quantitatively evaluate the frequency of false positive and false negative fibers in the LiFE algorithm that did not contribute to the optimization of the DWI model fitting process. Different configuration complexities were adopted to verify the angular dependence and sensitivity of the methods used in the framework and to identify the threshold for distinguishing separate fiber bundles.

Starting from a simple orthogonal crossing, configurations with discrete crossing angles were used, which imposed an upper limit on the crossing angle equal to 90 degrees. The intersecting fiber system involved two fiber bundles, \( F_1 \) and \( F_2 \), each containing 100 parallel fibers that were uniformly distributed inside the circular fiducial areas. The cross section between the two configurations contained fibers common to both bundles with the second bundle perforating the fixed collection of transverse fibers such that \( F_1 \cap F_2 = \{f : f \in F_1 \land f \in F_2\} \), where \( f \) simply denotes the fibers in the overlapping area. The intersection of \( F_1 \) and \( F_2 \) may be depicted as a 2-dimensional graphical projection onto the plane \( z = 0 \) while leaving out potential interleaved and stacked fiber crossings (Fig. 12).

Increasing the size of the volume element described in section 3.1.1 to \( 21 \times 21 \times 21 \) mm made it possible to account for outliers and diverging fiber tracts generated by the stochastic algorithm in the development of crossing fiber bundles. The local
fiber models were made more interpretable by using one assigned pixel for each fiber bundle as a seed region. The spatial and diffusion information was imported into the MATLAB workspace and further processed in order to be compatible with the LiFE framework. The implementation of the procedures for extracting the fiber coordinates and the corresponding DW measurements is provided in Appendix E3.

![Fig. 12. Schematic diagram of two crossing fiber populations. The first bundle \( F_1 \) remained fixed while the second set of aggregated fibers was rotated relative to the common center of both bundles. The crossing angle \( \theta_c \leq 90^\circ \) was set to vary in a discrete and closed angular interval, here exemplified by rotating \( F_2 \) with respect to the ground truth bundle.](image)

Configurations of crossing fibers were constructed by using a single elemental rotation of the streamline coordinates \( \mathbf{P} = (x_0, y_0, z_0)^T \) through a crossing angle \( \theta_c \) about the center of the fiber bundle in the counterclockwise direction (cf. Fig. 12) and by performing a translation of the coordinates (Appendix E3) such that the rotated coordinate vector \( \mathbf{P}' = R_y(\theta_c)(\mathbf{P} - \bar{\mathbf{P}}) + \bar{\mathbf{P}} \), where \( \bar{\mathbf{P}} \) is the median of the set of coordinates.

Curving fiber models were generated by fitting the spatial coordinates of \( F_2 \) to parabolas of different foci \( Q \) while keeping the vertex of the curves at \( \bar{\mathbf{P}} \) (Fig. 13), such that the transverse coordinates of the fiber pathways in \( F_2 \) were proportional to the square of the coordinate points along the longitudinal ground truth bundle. Five curved bundles of increasingly higher curvature were simulated in MATLAB and evaluated by displacing the focus point \( Q - \bar{\mathbf{P}} \) (mm). The curvature of each fiber bundle was given by \( \kappa = 4(Q - \bar{\mathbf{P}})^2[x^2 + 4(Q - \bar{\mathbf{P}})^2]^{-3/2} \) where \( Q - \bar{\mathbf{P}} \in \{5, 10, 15, 20, 25\} \) mm. The methodology for calculating the parabolic coordinates and the data used for a given focus point are given in Appendix E3.
Fig. 13. Curving fiber configurations generated by introducing a transverse focus offset of a parabola. The foci $Q$ of the positive parabolic bundles lie on the directrix emanating from the negative parabolas, and vice versa. The focus is displaced toward the vertex $P$ as the constant of proportionality increases, which is pictorially denoted by dashed lines.

### 3.3.4 Simulation of Noise and Artifacts

In addition to establishing crossing fiber models for the purpose of generating DW images without imprecisions, Rician and chi-square ($\chi^2$) distributed noise with a variance of 50 to the signal and combinations of artifacts were added to the parameters in the image processing in Fiberfox (Fig. 14). Provided that image artifacts were enabled, the limited size of the K-space produced ringing artifacts around edges due to low frequency filtering across a single image slice. By shrinking the FOV with 5%, aliasing effects could be included in the simulations, making images overlap each other in the process. Because of inherent inhomogeneities in external magnetic fields in different types of MRI scans, distortions were simulated in the processing of DW images in the gradient direction domain by including an extra phase during readout.
Fig. 14. Comparative examples of noise and artifact simulations in five DW images of the $F_1$ ground truth bundle in Fiberfox. The noise-free image in the top-left corner shows the original $21 \times 21$ mm lateral (sagittal) view of the fiber bundle. The simulations were generated with 90 diffusion directions using $b = 1000$ s·mm$^{-2}$.

Existing T1-weighted images and DW volumes provided in the LiFE framework with the file extension Neuroimaging Informatics Technology Initiative (NIfTI) were replaced with generated synthetic datasets. The `load_nii` and `save_nii` functions$^5$ were used to evaluate the preset volumes and corresponding structure array fields in order to assign variable data to the file header specified by the return values and output data created by the `load_nii` function (Appendix E2).

The output data was used to export the DW datasets in NIfTI format using the `save_nii` function, and the content of the data files was verified in both MITK and in ITK-SNAP (version 3.6.0-rc1 2016, Paul A. Yushkevich, Guido Gerig, www.itksnap.org), a neuroimaging viewer capable of interactive visualizations of the multidimensional arrays stored in the NIfTI format.

3.4 Spherical-Deconvolution Informed Filtering

3.4.1 Fiber Orientation Distribution Estimation

Spherical-deconvolution informed filtering of tractograms (SIFT) [58], was used to estimate a cross-section multiplier, or weighting factor, for each streamline from synthetic raw DW images using the MRtrix3 software package [59]. Instead of employing methods of removing false positive fibers as already described, an expectation-maximization framework was used to find an appropriate weighting factor for each streamline. The method required basic DWI processing, including fiber orientation distribution (FOD) estimation and streamline tractography. The fiber orientation was captured by a FOD function ($f_{\text{FOD}}$) that described the probabilistic distribution of fibers within each voxel,

$$f_{\text{FOD}}(\hat{x}) = \sum_{i=1}^{N} w_i \theta_i(\hat{x})$$

(3.8)

where the independent variable $\hat{x}$ represents the fiber direction in $N$ voxels and $\theta_i$ denotes the linear basis functions with corresponding weights $w_i$ [57, 62].

3.4.2 Constrained Spherical Deconvolution Tractography

In order to model the expected signal in a voxel containing a coherently oriented fiber bundle, a response function $R$ was estimated and used as a kernel when performing constrained spherical deconvolution (CSD) [60]. $R$ was derived directly from the synthetic DW image using the Tournier algorithm [61] in MRtrix3. The observed signal $S$ was calculated as the sum of measurements from the ground truth fiber bundle with the fiber directions being weighted by the fraction of fibers having the same orientation,

$$S(q) = \sum_{i=1}^{N} w_i \int \theta_i(\hat{x}) R(q, \hat{x}) d\hat{x}$$

(3.9)

where $q = \sqrt{\frac{b}{(\Delta - \delta)}}$, thereby combining a nonparametric and model-based approach, as exemplified by Eq. 3.2, to reconstruct the diffusion signal.

Deterministic streamline tractography was performed using the FOD estimation for each type of simulation (as explained in section 3.3.4), and the original DW image as input. The `read_mrtrix_tracks` and `write_mrtrix_tracks` functions were used to read and write the MRtrix3 tractography streamline coordinates in MATLAB. To simulate local fiber dispersion, the coordinates of the two generated endpoints of each streamline (cf. Fig. 6a) were used to join the streamline paths into unidirectional Cartesian trajectories throughout the length of the ground truth fiber bundle (Fig. 15).

The FOD estimation and the generated streamlines were used as input to the SIFT algorithm and the resulting weighting factor vectors were analyzed in Python 2.7.10\footnote{http://www.python.org. Accessed 2017-05-25.} and MATLAB. The streamline weighting factors $e^F$ for each fiber configuration were estimated by determining the cross-sectional area for each streamline, and the underlying track density (TD) in each FOD lobe was based on contributions from individual streamline lengths $|s|$ through the central voxels across $N$ streamlines attributed to the lobes in the FOD image [62]:

$$TD = \sum_{i=1}^{N} |s_i| e^F$$

(3.10)

### 3.5 Evaluation Methodology

#### 3.5.1 Quantitative Evaluation of Synthetic Datasets

While systematically varying parameters in Fiberfox, eigendecomposition was performed on the datasets in MATLAB (Appendix E1) and the diffusion tensor in each image voxel was approximated with a local estimation technique on the basis of the gradient vectors and corresponding signal intensities,

$$S_{gg}^T = \begin{pmatrix}
S_1 & g_{1.x} & g_{1.y} & g_{1.z} \\
S_2 & g_{2.x} & g_{2.y} & g_{2.z} \\
... & ... & ... & ...
\end{pmatrix}
\begin{pmatrix}
g_{1.x} & g_{2.x} & ... & g_{90.x} \\
g_{1.y} & g_{2.y} & ... & g_{90.y} \\
g_{1.z} & g_{2.z} & ... & g_{90.z}
\end{pmatrix}$$

(3.11)
The sum of all signals was taken over the set of gradient vectors in the simulation, analogous to principles of orientation distributions and orientation tensors laid down by Moakher [63]:

\[ D \sim \sum_{i=1}^{90} S_i g_i g_i^T \]  

(3.12)

The anisotropic diffusivity was evaluated with emphasis on relative diffusivity by weighting the elements of the tensor based on the maximum value of the leading diagonal, \( D_{ij} \) \( \forall i, j \in \{1, 2, 3\} \), \( i = j \). The root mean square error (RMSE) was used as a measure of the spread between signals generated with different fiber configurations, with the single-bundle configuration dataset as a reference point in place of a predicted signal value. The dot product of the weighted sum of two sets of signals was used as a quantitative metric to characterize \( S \) and to evaluate the similarity between the generated datasets,

\[ \cos \theta \sim \frac{\langle S_n / \sum_{i=1}^{90} S_{ni} \rangle \langle S_m / \sum_{i=1}^{90} S_{mi} \rangle}{\| S_n / \sum_{i=1}^{90} S_{ni} \| \| S_m / \sum_{i=1}^{90} S_{mi} \|} \]  

(3.13)

where \( n \) and \( m \) denotes any two of the three geometrical configurations. In order for the datasets to be regarded as viable candidates and stored as training sets for further implementation, the orientation of the weighted terms should not be parallel to each other.

### 3.5.2 Fascicle Weight Estimation Procedure

The single-bundle ground truth fascicle weights \((w_f)\) and the weights corresponding to the complex fiber configurations \((w)\) were computed as described in section 3.3.2 from Eq. 3.5 by establishing initial conditions in support of the synthetic partial connectome. First, the fiber pathway coordinates were accounted for in MATLAB by defining the track length and the maximum number of seeds per voxel generated in the candidate DW image in order to run the coordinate extraction script across any number of synthetic fiber tracts. In order to evaluate the models in terms of their fascicle weights, the fiber coordinates in the ground truth bundle and complex configurations were concatenated. Second, it was necessary to perform affine correction on the imported datasets by incorporating a 0.625 mm translation with respect to the homogenous coordinate transformation matrix. Third, the algorithm was stopped when the fixed number of fascicle weights was computed in the optimized partial connectome by inserting a breakpoint in the LiFE simulation script after having extracted the coordinates of the white matter voxels. See Appendices E2 and E3 for details on how the transformed coordinates of the complex fiber configurations were implemented with the NIfTI file. In addition to running single-shell acquisition simulations where \( b \in \{1000, 2000, 3000\} \text{ s-mm}^{-2} \), the DW images were obtained by combining shells such that the matrix of gradient vectors was extended to include two additional sets of diffusion-sensitizing gradients, \( G = (g_1, g_2 \cdots g_{270})^T \).
3.5.3 Unpaired Sample Analysis of Streamline Weighting Factors

The same single-bundle ground truth $F_1$ was used in the evaluation of the different crossing and curving fiber models. The streamline weighting factors obtained from the ground truth and modified bundle for each of the different curving and crossing fiber models (Fig. 16) were estimated with single- and multi-shell diffusion acquisitions using the default 2-compartment model in Fiberfox. Differences in contributions from streamlines to the ground truth fiber track density were evaluated using a two-sample unpaired $t$-test (Welch’s $t$-test). The general expression for a single test statistic $t$ in terms of $N$ streamlines simplified to

$$
t = \frac{1}{\sqrt{N(\sigma_1^2 + \sigma_2^2)}} \sum_{n=1}^{N} (e_{n,1}^F - e_{n,2}^F) \tag{3.14}
$$

where the first sample $\{e_{n,1}^F\}_{n=1}^{N}$ was the zero-degree ground truth dataset and where $\{e_{n,2}^F\}_{n=1}^{N}$ was the set of observed weighting factors obtained from the modified bundle with the corresponding standard deviations $\sigma_1^2$ and $\sigma_2^2$ for the ground truth and modified dataset, respectively. The value of the test statistic was used to return a decision for the null hypothesis that the ground truth and modified tracks should be reliably distinguishable from one another with a significance threshold set at $p < 0.05$.

\begin{align*}
\theta_c &= 0^\circ & \theta_c &= 10^\circ & \theta_c &= 20^\circ & \theta_c &= 30^\circ & \theta_c &= 40^\circ \\
\theta_c &= 50^\circ & \theta_c &= 60^\circ & \theta_c &= 70^\circ & \theta_c &= 80^\circ & \theta_c &= 90^\circ 
\end{align*}

Fig. 16. Sequential DW images of the crossing fiber population with respect to the zero-degree ground truth orientation. The figure demonstrates the presumptive diffusion signal contributions that arise from including synthetic crossing fascicles at 10-degree increments of rotation.
4 Results

4.1 Simulations of Synthetic Diffusion MRI Data

4.1.1 The Influence of Compartment Models on Signals

Fig. 17. Graphs showing the diffusion signals for different compartment models at a diffusion sensitivity factor of 1000 s·mm\(^{-2}\). The signals and diffusion-sensitizing gradient directions were generated in Fiberfox for each fiber configuration using a 2-compartment ball-and-stick model (a), and a 3-compartment tensor-stick-ball model (b). The fiber bundles in each configuration contained 100 uniformly distributed fibers. With respect to the single-bundle configuration, the RMSE between the two configurations with the 2-compartment model was 4.69 and 5.99 units of amplitude.
4.1.2 Effect of Diffusion Weighting

Fig. 18. Synthetic signals registered in central image voxels from the three fiber configurations with a 2-compartment partial volume ball-and-stick model. The graphs show that the amplitude of the diffusion signal decreases as the b-value increases.

4.1.3 Diffusion Tensor Calculations and Scalar Indices

The diffusion tensor estimation in the central image voxel for each fiber configuration is given in Table 4.1, in which the relative diffusivity is weighted based on the maximum value of the leading diagonal of D.

Table 4.1. Invariant diffusion anisotropy indices of simulated datasets. Scalar indices are obtained from estimating the diffusion tensor in the central voxel based on 100 fibers with the default 2-compartment model in Fiberfox. The axial diffusivity is given by $\lambda_1$ and the units for the diffusivities are in mm$^2$·s$^{-1}$.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>FA(D)</th>
<th>RA(D)</th>
<th>$D_{xy}$</th>
<th>$D_{yz}$</th>
<th>$D_{xz}$</th>
<th>$\langle D \rangle$</th>
<th>$\lambda_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cylinder</td>
<td>0.194</td>
<td>0.161</td>
<td>0.685</td>
<td>0.972</td>
<td>1.00</td>
<td>0.00523</td>
<td>0.00197</td>
</tr>
<tr>
<td>Sheet</td>
<td>0.203</td>
<td>0.168</td>
<td>0.673</td>
<td>0.971</td>
<td>1.00</td>
<td>0.00562</td>
<td>0.00218</td>
</tr>
<tr>
<td>Array</td>
<td>0.203</td>
<td>0.168</td>
<td>0.673</td>
<td>0.971</td>
<td>1.00</td>
<td>0.00562</td>
<td>0.00218</td>
</tr>
</tbody>
</table>
Generating Ground Truths of Neural Fiber Bundles

Diffusivity is determined by approximating the diffusion tensor as described in Eq. 3.12. The oblate appearance of the simulated distribution (a) is indicative of low fractional anisotropy, compared to the characteristics that define the HCP data sample (b).

Fig. 19. 3-dimensional vector plots illustrating the variation between synthetic and real DW datasets.
4.2 Optimized Synthetic Datasets

4.2.1 Fascicle Weight Estimation With Crossing Configurations

From the results presented in Table 4.2, 4.3, and 4.4, it is clear that aliasing artifacts have a profound impact on the partial connectome evaluation, which is indicated by false positive ratios above an angle of 50 degrees. A ratio equal to zero corresponds to the removal of non-existent fascicles that do not contribute to the ground truth DW image, whereas a value of 1 represents equal fascicle orientation.

Table 4.2. Ratios of crossing bundle fascicle weights to ground truth weights ($w/w_f$) using three diffusion sensitivity factors. The weights were estimated under noise-free conditions, with ringing and aliasing artifacts, and with Rician and chi-squared distributed noise.

<table>
<thead>
<tr>
<th>$\theta_c$</th>
<th>$b = 1000 \text{ s} \cdot \text{mm}^{-2}$</th>
<th>$b = 2000 \text{ s} \cdot \text{mm}^{-2}$</th>
<th>$b = 3000 \text{ s} \cdot \text{mm}^{-2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0^\circ$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10$^\circ$</td>
<td>0.155</td>
<td>0.308</td>
<td>0.559</td>
</tr>
<tr>
<td>20$^\circ$</td>
<td>0.438</td>
<td>0.675</td>
<td>0.781</td>
</tr>
<tr>
<td>30$^\circ$</td>
<td>0.264</td>
<td>0.504</td>
<td>0.195</td>
</tr>
<tr>
<td>40$^\circ$</td>
<td>0.248</td>
<td>0.225</td>
<td>0</td>
</tr>
<tr>
<td>50$^\circ$</td>
<td>0.087</td>
<td>0.316</td>
<td>0</td>
</tr>
<tr>
<td>60$^\circ$</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>70$^\circ$</td>
<td>0.049</td>
<td>0</td>
<td>0.222</td>
</tr>
<tr>
<td>80$^\circ$</td>
<td>0.012</td>
<td>0.118</td>
<td>0</td>
</tr>
<tr>
<td>90$^\circ$</td>
<td>0</td>
<td>0.227</td>
<td>0</td>
</tr>
</tbody>
</table>
Measurements at 90 degrees produced a value of zero for the crossing fascicle weights, except for when Rician noise was present at a diffusion sensitivity factor of 1000 and 2000 s·mm⁻², which corresponded to a ratio of 0.227 and 0.055, respectively.

**Table 4.3.** Ratios of crossing bundle fascicle weights to ground truth weights using multi-shell diffusion acquisition for 1000, 2000, and 3000 s·mm⁻².

<table>
<thead>
<tr>
<th>θ,°</th>
<th>Noise-free</th>
<th>Rician</th>
<th>χ²</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>N/A</td>
</tr>
<tr>
<td>10</td>
<td>N/A</td>
<td>2.810</td>
<td>7.670</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>20</td>
<td>1.350</td>
<td>0.516</td>
<td>0.109</td>
<td>5.450</td>
<td>1.350</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>1.810</td>
<td>0</td>
<td>0.574</td>
<td>2.460</td>
<td>1.810</td>
</tr>
<tr>
<td>50</td>
<td>0.014</td>
<td>0.129</td>
<td>0.549</td>
<td>0</td>
<td>0.014</td>
</tr>
<tr>
<td>60</td>
<td>1.250</td>
<td>0.675</td>
<td>0.298</td>
<td>1.620</td>
<td>0</td>
</tr>
<tr>
<td>70</td>
<td>0</td>
<td>0.231</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>0.025</td>
<td>0.812</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0.461</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

N/A is used to represent the fascicle weight fraction when \( w_f \) tend to zero.
4.2.2 Fascicle Weight Estimation With Curving Configurations

Table 4.4. Quotients of the weighting factors of curving fibers to the assigned weights of ground truth fibers. The displacement of five focus points \( Q - \bar{P} \) (\( \text{mm} \)) is used to generate sets of curving fiber configurations.

<table>
<thead>
<tr>
<th>( Q - \bar{P} )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0.190</td>
<td>0</td>
<td>0.047</td>
<td>1.130</td>
</tr>
<tr>
<td>10</td>
<td>0.379</td>
<td>0.615</td>
<td>0.326</td>
<td>0.473</td>
<td>1.220</td>
</tr>
<tr>
<td>15</td>
<td>1.060</td>
<td>0.912</td>
<td>0.940</td>
<td>1.130</td>
<td>1.370</td>
</tr>
<tr>
<td>20</td>
<td>1.330</td>
<td>0.844</td>
<td>1.120</td>
<td>1.360</td>
<td>1.350</td>
</tr>
<tr>
<td>25</td>
<td>1.520</td>
<td>0.740</td>
<td>1.050</td>
<td>1.350</td>
<td>1.300</td>
</tr>
</tbody>
</table>

\( b = 2000 \text{ s mm}^{-2} \)

<table>
<thead>
<tr>
<th>( Q - \bar{P} )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.019</td>
<td>0.756</td>
</tr>
<tr>
<td>10</td>
<td>0.592</td>
<td>0.621</td>
<td>0.636</td>
<td>0.680</td>
<td>1.290</td>
</tr>
<tr>
<td>15</td>
<td>1.370</td>
<td>1.280</td>
<td>1.220</td>
<td>1.420</td>
<td>1.800</td>
</tr>
<tr>
<td>20</td>
<td>1.670</td>
<td>1.770</td>
<td>1.270</td>
<td>1.670</td>
<td>1.900</td>
</tr>
<tr>
<td>25</td>
<td>1.560</td>
<td>1.810</td>
<td>1.060</td>
<td>1.520</td>
<td>1.760</td>
</tr>
</tbody>
</table>

\( b = 3000 \text{ s mm}^{-2} \)

<table>
<thead>
<tr>
<th>( Q - \bar{P} )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.002</td>
<td>0.444</td>
</tr>
<tr>
<td>10</td>
<td>0.925</td>
<td>0.342</td>
<td>1.020</td>
<td>1.000</td>
<td>1.540</td>
</tr>
<tr>
<td>15</td>
<td>1.780</td>
<td>0.796</td>
<td>1.500</td>
<td>1.840</td>
<td>2.420</td>
</tr>
<tr>
<td>20</td>
<td>2.100</td>
<td>0.953</td>
<td>1.380</td>
<td>2.140</td>
<td>2.670</td>
</tr>
<tr>
<td>25</td>
<td>1.850</td>
<td>1.070</td>
<td>1.020</td>
<td>1.890</td>
<td>2.350</td>
</tr>
</tbody>
</table>

Multi-shell acquisition for 1000, 2000, and 3000 s mm\(^{-2}\)

<table>
<thead>
<tr>
<th>( Q - \bar{P} )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.194</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>3.340</td>
<td>0.875</td>
<td>0.638</td>
<td>16.600</td>
<td>3.340</td>
</tr>
<tr>
<td>20</td>
<td>4.920</td>
<td>0.916</td>
<td>0.911</td>
<td>28.400</td>
<td>4.920</td>
</tr>
<tr>
<td>25</td>
<td>2.650</td>
<td>0.520</td>
<td>0.465</td>
<td>12.100</td>
<td>2.660</td>
</tr>
</tbody>
</table>

4.3 Informed Filtering of Tractograms

4.3.1 Distribution of Streamline Weighting Factors

The results reported in Fig. 20-24 for the ground truth fiber-based weighting factors shows an approximately equal interquartile range across \( \theta_c \), whereas the rotated fiber bundle produces more pronounced ranges of variation. Each type of simulation is presented on a separate page and underscores the influence of noisy datasets on the SIFT algorithm for crossing and curving configurations (Fig. 25-27), and the statistical significance of the findings is made evident in Tables 4.5 and 4.6.
Fig. 20. Box plot representations of noise-free streamline weighting factors obtained from simulations of crossing fiber systems with a maximum of 20 streamlines per fiber bundle. The single-shell acquisition results from the rotated fiber bundle are presented separately from the zero-degree ground truth.
Fig. 21. Rician noise-corrupted streamline weighting factors from the ground truth and rotated fiber bundle.
Fig. 22. Chi-square distributed noise influencing the variability between the ground truth and rotated fiber bundle weighting factors.
Fig. 23. The influence of Gibbs-ringing artifacts on the estimation of streamline weighting factors. The effects are pronounced in the single-shell acquisition simulations and persistent across all crossing angles.
Fig. 24. Aliasing-ridden streamline weighting factors from the ground truth and rotated fiber configurations.
Fig. 25. Box plots showing the spread of streamline weighting factors for simulations involving parabolically curved fiber bundles. The weighting factors produced using single-shell acquisition reveals the streamline contribution to the underlying track density as the focus offset of the parabolic bundles increases for noise-free fiber systems and datasets with Rician and chi-square distributed noise.
Fig. 26. The effect of Gibbs-ringing and aliasing artifacts on weighting factors from the ground truth and curved fiber bundles using single-shell diffusion acquisition.
Fig. 27. Distributions of weighting factors for the ground truth and curved fiber configurations using multi-shell diffusion acquisition.
Table 4.5. Distribution of right-tailed crossing fiber-based p-values from Welch’s t-test for unpaired samples. The p-values of the observed results are deemed statistically significant if \( p < 0.05 \).

\[
b = 1000 \text{ s} \cdot \text{mm}^{-2}
\]

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10°</td>
<td>0.867</td>
<td>0.955</td>
<td>0.942</td>
<td>0.599</td>
<td>0.086</td>
</tr>
<tr>
<td>20°</td>
<td>0.391</td>
<td>0.590</td>
<td>0.610</td>
<td>0.908</td>
<td>0.093</td>
</tr>
<tr>
<td>30°</td>
<td>0.212</td>
<td>0.464</td>
<td>0.417</td>
<td>0.583</td>
<td>0.048*</td>
</tr>
<tr>
<td>40°</td>
<td>0.131</td>
<td>0.381</td>
<td>0.364</td>
<td>0.446</td>
<td>0.011*</td>
</tr>
<tr>
<td>50°</td>
<td>0.163</td>
<td>0.357</td>
<td>0.315</td>
<td>0.493</td>
<td>0.011*</td>
</tr>
<tr>
<td>60°</td>
<td>0.145</td>
<td>0.363</td>
<td>0.300</td>
<td>0.476</td>
<td>0.011*</td>
</tr>
<tr>
<td>70°</td>
<td>0.145</td>
<td>0.428</td>
<td>0.376</td>
<td>0.434</td>
<td>0.011*</td>
</tr>
<tr>
<td>80°</td>
<td>0.118</td>
<td>0.441</td>
<td>0.394</td>
<td>0.345</td>
<td>0.010*</td>
</tr>
<tr>
<td>90°</td>
<td>0.107</td>
<td>0.337</td>
<td>0.299</td>
<td>0.272</td>
<td>0.024*</td>
</tr>
</tbody>
</table>

\[
b = 2000 \text{ s} \cdot \text{mm}^{-2}
\]

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10°</td>
<td>0.818</td>
<td>0.948</td>
<td>0.871</td>
<td>0.871</td>
<td>0.984</td>
</tr>
<tr>
<td>20°</td>
<td>0.486</td>
<td>0.545</td>
<td>0.539</td>
<td>0.539</td>
<td>0.919</td>
</tr>
<tr>
<td>30°</td>
<td>0.307</td>
<td>0.348</td>
<td>0.331</td>
<td>0.331</td>
<td>0.708</td>
</tr>
<tr>
<td>40°</td>
<td>0.224</td>
<td>0.302</td>
<td>0.268</td>
<td>0.268</td>
<td>0.594</td>
</tr>
<tr>
<td>50°</td>
<td>0.246</td>
<td>0.235</td>
<td>0.234</td>
<td>0.234</td>
<td>0.604</td>
</tr>
<tr>
<td>60°</td>
<td>0.218</td>
<td>0.235</td>
<td>0.195</td>
<td>0.195</td>
<td>0.493</td>
</tr>
<tr>
<td>70°</td>
<td>0.224</td>
<td>0.263</td>
<td>0.243</td>
<td>0.243</td>
<td>0.450</td>
</tr>
<tr>
<td>80°</td>
<td>0.185</td>
<td>0.285</td>
<td>0.252</td>
<td>0.252</td>
<td>0.352</td>
</tr>
<tr>
<td>90°</td>
<td>0.183</td>
<td>0.221</td>
<td>0.190</td>
<td>0.190</td>
<td>0.383</td>
</tr>
</tbody>
</table>

\[
b = 3000 \text{ s} \cdot \text{mm}^{-2}
\]

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>Noise-free</th>
<th>Rician</th>
<th>( \chi^2 )</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10°</td>
<td>0.951</td>
<td>0.973</td>
<td>0.925</td>
<td>0.925</td>
<td>0.849</td>
</tr>
<tr>
<td>20°</td>
<td>0.728</td>
<td>0.515</td>
<td>0.474</td>
<td>0.474</td>
<td>0.513</td>
</tr>
<tr>
<td>30°</td>
<td>0.576</td>
<td>0.338</td>
<td>0.285</td>
<td>0.285</td>
<td>0.328</td>
</tr>
<tr>
<td>40°</td>
<td>0.401</td>
<td>0.278</td>
<td>0.215</td>
<td>0.215</td>
<td>0.247</td>
</tr>
<tr>
<td>50°</td>
<td>0.441</td>
<td>0.238</td>
<td>0.179</td>
<td>0.179</td>
<td>0.268</td>
</tr>
<tr>
<td>60°</td>
<td>0.411</td>
<td>0.196</td>
<td>0.152</td>
<td>0.152</td>
<td>0.243</td>
</tr>
<tr>
<td>70°</td>
<td>0.420</td>
<td>0.243</td>
<td>0.188</td>
<td>0.188</td>
<td>0.248</td>
</tr>
<tr>
<td>80°</td>
<td>0.365</td>
<td>0.256</td>
<td>0.189</td>
<td>0.189</td>
<td>0.208</td>
</tr>
<tr>
<td>90°</td>
<td>0.362</td>
<td>0.196</td>
<td>0.144</td>
<td>0.144</td>
<td>0.208</td>
</tr>
</tbody>
</table>

| Multi-shell acquisition for 1000, 2000, and 3000 s \cdot \text{mm}^{-2} |
|---|---|---|---|---|---|
| 10° | 0.951 | 0.921 | 0.901 | 0.901 | 0.951 |
| 20° | 0.728 | 0.562 | 0.445 | 0.445 | 0.729 |
| 30° | 0.576 | 0.334 | 0.271 | 0.271 | 0.576 |
| 40° | 0.401 | 0.271 | 0.202 | 0.202 | 0.401 |
| 50° | 0.441 | 0.229 | 0.168 | 0.168 | 0.441 |
| 60° | 0.411 | 0.179 | 0.149 | 0.149 | 0.411 |
| 70° | 0.420 | 0.239 | 0.181 | 0.181 | 0.420 |
| 80° | 0.365 | 0.240 | 0.165 | 0.165 | 0.365 |
| 90° | 0.362 | 0.184 | 0.128 | 0.128 | 0.362 |

* Significant difference between the crossing and ground truth configurations (\( p < 0.05 \)).
Table 4.6. Distribution of right-tailed curving fiber-based p-values as the parabolic bundle is fitted to different foci $Q - \bar{P}$ (mm).

$$b = 1000 \text{ s} \cdot \text{mm}^{-2}$$

<table>
<thead>
<tr>
<th>$Q - \bar{P}$</th>
<th>Noise-free</th>
<th>Rician</th>
<th>$\chi^2$</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.002*</td>
<td>0.077</td>
<td>0.059</td>
<td>0.001*</td>
<td>0.324</td>
</tr>
<tr>
<td>10</td>
<td>0.297</td>
<td>0.191</td>
<td>0.082</td>
<td>0.059</td>
<td>0.324</td>
</tr>
<tr>
<td>15</td>
<td>0.795</td>
<td>0.402</td>
<td>0.367</td>
<td>0.338</td>
<td>0.045*</td>
</tr>
<tr>
<td>20</td>
<td>0.753</td>
<td>0.816</td>
<td>0.825</td>
<td>0.309</td>
<td>0.055</td>
</tr>
<tr>
<td>25</td>
<td>0.426</td>
<td>0.627</td>
<td>0.674</td>
<td>0.179</td>
<td>0.041*</td>
</tr>
</tbody>
</table>

$$b = 2000 \text{ s} \cdot \text{mm}^{-2}$$

<table>
<thead>
<tr>
<th>$Q - \bar{P}$</th>
<th>Noise-free</th>
<th>Rician</th>
<th>$\chi^2$</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.019*</td>
<td>0.048*</td>
<td>0.045*</td>
<td>0.002*</td>
<td>0.006*</td>
</tr>
<tr>
<td>10</td>
<td>0.511</td>
<td>0.077</td>
<td>0.060</td>
<td>0.064</td>
<td>0.138</td>
</tr>
<tr>
<td>15</td>
<td>0.894</td>
<td>0.324</td>
<td>0.285</td>
<td>0.222</td>
<td>0.353</td>
</tr>
<tr>
<td>20</td>
<td>0.795</td>
<td>0.722</td>
<td>0.714</td>
<td>0.170</td>
<td>0.545</td>
</tr>
<tr>
<td>25</td>
<td>0.538</td>
<td>0.664</td>
<td>0.734</td>
<td>0.080</td>
<td>0.545</td>
</tr>
</tbody>
</table>

$$b = 3000 \text{ s} \cdot \text{mm}^{-2}$$

<table>
<thead>
<tr>
<th>$Q - \bar{P}$</th>
<th>Noise-free</th>
<th>Rician</th>
<th>$\chi^2$</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.009*</td>
<td>0.057</td>
<td>0.042*</td>
<td>0.001*</td>
<td>0.023*</td>
</tr>
<tr>
<td>10</td>
<td>0.331</td>
<td>0.076</td>
<td>0.052</td>
<td>0.122</td>
<td>0.538</td>
</tr>
<tr>
<td>15</td>
<td>0.671</td>
<td>0.304</td>
<td>0.252</td>
<td>0.257</td>
<td>0.905</td>
</tr>
<tr>
<td>20</td>
<td>0.543</td>
<td>0.688</td>
<td>0.649</td>
<td>0.147</td>
<td>0.799</td>
</tr>
<tr>
<td>25</td>
<td>0.307</td>
<td>0.751</td>
<td>0.802</td>
<td>0.064</td>
<td>0.538</td>
</tr>
</tbody>
</table>

Multi-shell acquisition for 1000, 2000, and 3000 s·mm$^{-2}$

<table>
<thead>
<tr>
<th>$Q - \bar{P}$</th>
<th>Noise-free</th>
<th>Rician</th>
<th>$\chi^2$</th>
<th>Gibbs</th>
<th>Aliasing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.009*</td>
<td>0.069</td>
<td>0.039*</td>
<td>0.051</td>
<td>0.009*</td>
</tr>
<tr>
<td>10</td>
<td>0.331</td>
<td>0.080</td>
<td>0.060</td>
<td>0.643</td>
<td>0.331</td>
</tr>
<tr>
<td>15</td>
<td>0.671</td>
<td>0.342</td>
<td>0.255</td>
<td>0.701</td>
<td>0.671</td>
</tr>
<tr>
<td>20</td>
<td>0.543</td>
<td>0.741</td>
<td>0.645</td>
<td>0.753</td>
<td>0.543</td>
</tr>
<tr>
<td>25</td>
<td>0.307</td>
<td>0.658</td>
<td>0.746</td>
<td>0.999</td>
<td>0.307</td>
</tr>
</tbody>
</table>

* Significant difference between the curving and ground truth configurations ($p < 0.05$).
5 Discussion

There is no formal definition of a coherent fiber structure that is characteristic of white matter in the brain. This is one of the reasons why reconciling constraints of fiber reconstruction methods with limitations inherent to validation of tractography is a major undertaking. Moreover, it is known that extrapolation of agreements on particular fiber tracts is not possible due to the physiological interindividual variability and the varying degree to which neural pathways are differentiated, which is also a principal indicator of the complexity of cerebral connectivity.

For simplicity, regions that contain mainly high curvature fibers were avoided in the modeling process. Instead, the initial approaches were geared towards modeling situations where the fiber orientation was unique and did not involve problematic fiber configurations. For this purpose the synthetic fibers were patterned after the main commissural tract, the corpus callosum. It was therefore necessary to properly motivate the use of idealized local configurations (Fig. 10) of fiber bundles in simulations that were aimed at capturing the microstructural characteristics of white matter in order to assess the validity and reliability of tractography methods.

5.1 Viability of Simulated Neural Fiber Bundles

The signals in the central image voxel in the representative volume element were comparable across different fiber configurations and their comparability was made clear by the relative difference in signal intensities with respect to the signals generated with the single-bundle configuration.

Increasing the amount of diffusion weighting resulted in a decreased signal amplitude (Fig. 18), which was to be expected according to the Stejskal-Tanner relationship (Eq. 2.5) in terms of the gradient direction and diffusion sensitivity factor. Similarly, using biophysical models with three or even four axonal compartment models resulted in lower-amplitude signals acquired in the central voxel. For the sake of consistency, a default 2-compartment model was used in the simulations shown in Fig. 17a, and the data was compared with the signals generated with a 3-compartment tensor-stick-ball model (Fig. 17b) that shared characteristics with a tensor-stick-sphere model. The tensor-stick-sphere model was taken as a reference because of its documented ability to accurately predict diffusion measurements, being the second highest rated model in the taxonomy and comparison of multi-compartment models by Panagiotaki et al. [45]. The only available 3-compartment combination of models in Fiberfox that was evaluated by the authors was the zeppelin-stick-dot model, which ranked 15th out of 47 analytic models. As mentioned in section 2.3.7 and demonstrated by Ferizi et al. [47] with high reproducibility, 3-compartment models outperform biophysical models that use fewer axonal spaces to separate diffusive processes in nervous tissue. The same
reasoning was applied to the data obtained with different fiber configurations in the simulation framework and it was concluded that the results, which were in agreement with combinations of different compartment models, could be taken as representative of the synthetic datasets produced using MITK and Fiberfox. This made it possible to avoid exhaustive evaluations of software parameter settings in relation to changes in the artificial diffusion signal.

Since the parallel fibers were simulated in a single voxel, the resulting signal amplitude produced by a specific signal model can be motivated in terms of the total number of bundles in the fiber configuration and the automatic signal scaling implemented in the Fiberfox framework. Specifically, if the single-bundle model (Fig. 10a) produces a signal amplitude $S$, the fiber sheet configuration would produce an amplitude of $5S$ and the array of packed sheets would generate an amplitude of $25S$. Even if the fiber radius was set manually to mimic realistic size distributions in the CC or automatically selected by the software, the resulting signal was scaled in all voxels to obtain at least one full fiber voxel in the volume element, meaning that the voxel with the highest fiber density was interpreted as a local region consisting exclusively of synthetic fiber tracts. The values $S$, $5S$, and $25S$ resulted in approximately the same signal amplitude (Fig. 17), irrespective of the number of axonal compartment models being used. The resulting synthetic signal could therefore be expressed generically as being proportional to $mS/N$ where $m$ is the modifier to obtain a full fiber voxel, $N$ is the number of cylindrical fiber bundles, and $S$ denotes the diffusion signal for a fiber configuration in a specific gradient direction.

As a consequence of obtaining highly similar synthetic datasets with the Fiberfox plug-in to MITK (RMSE of 4.69 and 5.99 units of amplitude between the single-bundle signal reference point and the sheet and array configuration, respectively), no compartment-dependent classification of the DWI signal intensity change was possible. Therefore, the artificial images were not seen as viable options for supervised learning algorithms to map predictions to instances in the open-access preprocessed HCP datasets. To further substantiate the results presented in Fig. 17 and in Table 4.1, the scalar product of the set of all combinations of signal collections was calculated (Eq. 3.13).

Estimates of diffusion tensors from synthetically generated datasets pointed to diffusion more characteristic of isotropic morphology than of anisotropic white matter tissue (Fig. 19a). There were regrettably few instances when the data tended towards true anisotropy. As shown in Table 4.1, the reported approximation of each diagonal tensor component indicated little or no difference along the $x$, $y$, and $z$-axes, and was further validated by the accompanying invariant anisotropy indices.

The diffusion anisotropy values, which were calculated with the MATLAB toolbox for numerous fiber configurations, were found to be within the range of diffusion measurements in the literature. These verifications were of importance since they relate to the issue of validation as discussed in some detail in section 2.3. Ideally, the simulated datasets should exhibit distinct differences in terms of anisotropy estimates and signal intensity characteristics (as exemplified in Fig. 18).
5.2 Estimation of Fascicle Weights

The purpose of the LiFE software was to evaluate the evidence supporting fiber tracts and connections based on a large collection of white matter fascicles. The initial model created by the algorithm contained a prediction of the diffusion measurements in each image voxel given by the fiber pathways in the probabilistic tractography solution, or the connectome.

In this work, the approach was aimed at modifying the algorithm to be able to evaluate and classify customized dMRI datasets. The data supporting the existence of fiber tracts could then be used to generate collections of distinct DW measurement data in order to capture different aspects of simple fiber configurations and crossings. The diffusion direction across all voxels traversed by a single fascicle was predicted from the orientation and position of the fascicle inside a voxel, and contributions from multiple fibers in the voxel were used to generate a global map of the large-scale structural brain network. One concern was, therefore, with the computation of the individual fascicle weights as they were computed taking into account the entire connectome by running a whole-brain voxel-wise analysis of white matter tracts. The objective was thus to determine the relative difference between estimated fascicle weights in the optimized connectome (Table 4.2). Instead of using a candidate tractography connectome as intended, a simple tractogram based on coherently oriented fiber bundles was used as input data in order to evaluate the retained fascicles that contributed to the prediction of the accompanying synthetic DW image. These initial simulations were carried out in order to determine an appropriate course of action and computational procedure for the linear fascicle evaluation approach, and to ensure that the algorithm was rigorously defined and able to run with customized datasets and complex fiber configurations.

One can infer, on the basis of the results presented in Table 4.2, that the LiFE algorithm is highly sensitive to discrete changes in angular orientation of a single set of crossing fibers in relation to a fixed synthetic neural fiber bundle. This was confirmed by straightforward descriptive statistics and exemplified, in some instances, by the fact that two disparate crossing angles \( \theta_c \in \{0^\circ, 10^\circ, \ldots, 90^\circ\} \) could yield approximately the same value. It is important to recognize that this is quite different from ratios being equivalent since the ground truth bundle \( F_1 \) was held fixed while \( \theta_c \) increased. In the simulated cases, it also appeared that the algorithm inaccurately predicted the number of false positive fascicles in the optimized connectome, which led to some results being misconstrued. This was particularly evident when simulating noisy and artifact-ridden datasets. Distortion artifacts (cf. Fig. 14) were not included in the final evaluation of the LiFE algorithm since the DW data produced for each crossing angle was zero.

The results show that the LiFE algorithm tend to not exclusively retain the ground truth fiber bundle and that it is corrupted by sensitivity inhomogeneities at higher \( b \)-values, which is further attested to by contextual outliers such as 3.130, 5.210, and 3.090 at a crossing angle of 10 degrees using \( b = 3000 \text{ s-mm}^2 \). The same holds for acquisitions strategies that combine shells of single \( b \)-values (Table 4.3).
As documented by the fascicle weight fractions in Table 4.4, the algorithm was not able to adequately characterize optimized networks in setups involving curving fiber configurations. With increasing b-values, similar to instances involving crossing fibers, it also exhibited bias towards discarding the ground truth dataset, the most obvious example being the computed ratios from using the 3000 s-mm$^{-2}$ DW image obtained by simulating an insufficient sampling rate. As the distance between the vertex and focus point of the parabola increased (cf. Fig. 13), the noise-free single-shell simulations returned an ever-increasing value, with the exception of the maximum values 1.330, 1.670, and 2.100 at $Q - \bar{P} = 10$ mm. The $F_2$ bundle received zero weights when the focus point was positioned at 5 mm with respect to $\bar{P}$, which was plausible since this was the narrowest parabola used in the investigation with an idealized curvature of $\kappa = 0.100$ at the vertex (section 3.3.3).

5.3 Informed Filtering and Streamline Reconstructions

Contrary to the method outlined for the LiFE algorithm, the weights assigned to each streamline varied and could not be represented as quotients between weighting factors. The modified MRtrix3 SIFT algorithm provided a comparison of the overall $e^F$ distribution range and the variability observed amongst the streamline weighting factors facilitated the assessment of individual streamline contributions to the underlying track density in terms of descriptive statistics. The outliers outside the upper or lower adjacent values were most likely due to the inherent variability of the synthetic DW data. The noisy datasets yielded high numbers of skewed distributions as many of the estimated weighting factors were classified as outliers. The same trends were also found between the zero-degree ground truth and the curving fiber datasets (Fig. 25-27) where the same argument can be made for common ranges, typical values, and the appearance of outliers.

Based on the two-sample test statistic, the ability of the algorithm to reliably distinguish ground truth fiber populations in both noise-free and artifact-ridden simulations (Table 4.5) proved highly improbable ($p > 0.05$). However, much like the calculated distribution of fascicle weights for curved fiber bundles, the algorithm was statistically likely to discern the ground truth fiber tracks from the fiber systems when $Q - \bar{P} = 5$ mm (Table 4.6).

5.4 Incorporation of Parameter Constraints

The tension, continuity, and bias of the real-time generated fibers were constrained to be zero across all diffusion modeling methods, which effectively restrained the sharpness and direction of curvature between consecutive fiducials as illustrated in Fig. 9. These constrictions were also in line with the expected outcome of using longitudinally oriented fibers as simulation ground truth data, which made it easier to focus exclusively on those aspects that were relevant when constructing configurations of parallel fibers. In order to keep the exponential decay of the DW
signal amplitude and resulting diffusion within clinical range, the diffusion sensitivity factor was controlled and did not exceed 3000 s-mm$^{-2}$, a choice based primarily on the information provided in the HCP diffusion protocol (section 3.1.2). This granted the simulations a realistic facet, considering that the amount of diffusion weighting is controlled by the b-value, and thus indirectly affects both TE and the gradient magnitudes.

Due to the sensitivity of the LiFE algorithm, 10-degree increments were used in calculating the ratio between fascicle weights in the optimized connectome. In order to generate noisy DW images, either a Rician or chi-square distributed noise was added to the diffusion signal, in addition to aliasing and ringing artifacts with default settings.

### 5.5 Limitations

In using the `readVTK` function for importing stochastic streamline coordinates from MITK into MATLAB, the spatial data was effectively limited to the number of sampling points along the longest streamline generated for a group of synthetic fiber tracts, irrespective of its position and orientation in 3-dimensional space. This inability of the function to cope with conflicting fiber geometries made any simulated fiber crossings inaccessible to the LiFE algorithm and prohibited the use of certain noise and artifact simulations that altered the geometry of simulated fiber tracks.

When assessing constraints on generalizability, it is necessary to comment on the fact that no additional modifications to the algorithm were implemented except those needed for calculating fascicle weights. This exploratory work was limited partly by not thoroughly investigating the VISTASOFT software repository for potential options available for expediting the quantification of prediction errors. One example, albeit a minor one in terms of it affecting the outcome of evidence supporting properties of candidate connectomes, was the ability to modify the default 2-compartment model used by the LiFE framework.
6 Conclusion

The exploratory approach set out in this work aimed at discovering regularity properties in diffusion anisotropy measurements across three visually distinct fiber configurations and to classify the findings into either regression or classification machine learning systems. Considering the self-similar space-filling sheet and array configurations, one might have expected that the signal contributions from neighboring voxels containing parallel fibers would have drastically altered the diffusion signal in the central image voxel. However, this was not observed in the simulations, and inferences about diffusion characteristics could not be made on the basis of geometrical features. Information derived from synthetically generated datasets in Fiberfox cannot therefore be used to construct viable models for predicting objective ground-truths with the simulation methods outlined in this work.

Even though the synthetic datasets proved to be compatible with the LiFE software, the information derived from the optimized partial connectome was deemed unreliable in predicting probabilistic tractograms of complex intravoxel fiber systems. The same conclusion was reached by using the extended SIFT method. However, this work suggests that SIFT is suitable for general-purpose quantitative dMRI methods and that the algorithm delivers adequate usability.

State-of-the-art tools for resolving complex fiber systems still have some way to go to achieve ground-truth validation. Further research is required to provide a reliable framework for correctly categorizing local fiber configurations such as the ones described in the preceding sections, which would in turn provide a basis for developing validated models for predicting diffusion in complex neural tissue.

6.1 Recommendations for Future Work

Future investigations into the use of synthetically generated sample data for the purpose of creating training sets for supervised machine learning or sparse dictionary learning algorithms would benefit from managing conflicting fiber geometries in a more comprehensive manner by adopting techniques that circumvent the problem of having to adhere to orientation-invariant dependence of fiber positions. An option would be to handle the synthetic fiber coordinate data using the previously mentioned MRtrix3 MATLAB functions and adjust for negative values before calling the data from the LiFE algorithm. Issues related to deriving an optimized connectome with LiFE and complexities associated with determining prediction accuracies should be accorded the same level of attention. It is also advisable to evaluate the results with various non-default fascicle-weighting schemes in order to capture subtle changes in fiber dispersion as a consequence of including complex configurations in the candidate connectome.
As previously outlined, the accompanying repository should be inspected and tested with the aim of characterizing local fiber architecture and confirming the scientific consensus that no combination of parameter values is reliable across all situations and that quantitative measures of diffusion depend on sequence constraints, noise, and artifacts.

It is also advisable to evaluate the fit between streamline reconstructions and DW images with public data from the HCP and to use the LiFE software to investigate the possibility of reducing the number of false positive fibers at a local level within the CC.
References


Appendices

Appendix A  Neuroanatomy of White Matter Pathways

![Diagram of the brain](image)

**Fig. A1.** Frontal section of the brain with a cross section through the cerebral hemispheres. The hemispheres are separated by the longitudinal fissure, showing their simplified structure in relation to the white matter commissural tracts, projection tracts, and short association tracts.

The fiber tracts or fascicles in the brain consists of myelinated bundles of nerve fibers. There are three main types of white matter fibers that transmit nerve impulses in the brain, namely, commissural tracts, projection tracts, and association tracts (Fig. A1). Commisural tracts form connections between regions of grey matter in opposite hemispheres, whereas the association fibers connects areas between gyri and transmit impulses between nuclei in the same cerebral hemisphere. Transmission of nerve impulses between lower (inferior) parts of the brain and along the spinal cord is made possible by projection tracts, which also connects various nuclei and gyri to distinct regions in the brain.

Appendix B  Image Coordinate System Conversion

In this appendix, complementary information to the official NRRD file format semantics and documentation\(^2\) is given. For purposes of this section, the coordinate systems described generally refer to a 3-dimensional space defined by a set of three linearly independent vectors (Fig. B1) relative to the orientation of the surrounding raster grid, which is located in the so-called world space. For further information, refer to the self-contained explanation of image orientation by Gordon L. Kindlmann at the University of Chicago\(^3\).

![Fig. B1. A 3-dimensional coordinate system with a set of orthonormal vectors \(\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}\) depicted before and after affine transformation with origin \(O\) and \(O'\), respectively.](image)

Affine transformations are used to convert between 3-dimensional spaces and include reflections, rotations, shearings, dilations, and translations. The net effect of applying an affine transformation can, in general, be expressed in terms of a matrix multiplication and a 3-dimensional translation \(\mathbf{T} = (T_x \ T_y \ T_z)^T\):

\[
\mathbf{Ax} + \mathbf{T} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}
\]

where the matrix \(\mathbf{A}\) holds the non-translational components for the generalized affine transformation and \(\mathbf{x}\) describes the physical coordinates being mapped onto the new coordinate space \((x \ y \ z)^T\).

The LPS anatomical coordinate system is fixed to the object being scanned in the MRI machine and is defined within the world-space coordinate system by an orthonormal basis \(B_{\text{LPS}}\) that orients the axes of the scanned volume (cf. Fig. B2a) in \(\mathbb{R}^3\). A unique index space is used in neuroimaging software, which is a 3-dimensional array of indices \((i, j, k)\) (Fig. B2c) that correspond to the different anatomical planes defined by \(B_{\text{LPS}}\) (Fig. B2b).


Fig. B2. The axial or transverse plane passes through the body and divides it into two parts: a bottom or inferior (I) section and an upper or superior (S) section. The coronal plane runs through the body and divides it into an anterior (A) or front section and a posterior (P) or back section, whereas the sagittal plane divides the body into right (R) and left (L) portions. a) The LPS anatomical coordinate system and the corresponding index space (b). c) The shaded unit cell represents a pixel $P_{ijk}$ with integral coordinates used for defining its position in the grid.

For coordinate systems in projective spaces, the index space and LPS basis are used to orient images and requires an affine transformation,

$$
\begin{pmatrix}
  x \\
  y \\
  z
\end{pmatrix}
= \begin{pmatrix}
  A_{11} & A_{12} & A_{13} \\
  A_{21} & A_{22} & A_{23} \\
  A_{31} & A_{32} & A_{33}
\end{pmatrix}
\begin{pmatrix}
  T_x \\
  T_y \\
  T_z
\end{pmatrix}
\begin{pmatrix}
  i \\
  j \\
  k
\end{pmatrix}
\begin{pmatrix}
  1
\end{pmatrix}
$$

Homogenous coordinates of an image point in the projective space should not be confused with 4-dimensional quaternions, which cannot be applied to the concepts described in this section.
In general, the geometric relationship between spatial locations described in index space are described by the basis independent vectors \((A_{11} A_{21} A_{31})^T\), \((A_{12} A_{22} A_{32})^T\), and \((A_{13} A_{23} A_{33})^T\). The world-space location \(P(i, j, k)\) can be described as

\[
P(i, j, k) = i \begin{pmatrix} A_{11} \\ A_{21} \\ A_{31} \end{pmatrix} + j \begin{pmatrix} A_{12} \\ A_{22} \\ A_{32} \end{pmatrix} + k \begin{pmatrix} A_{13} \\ A_{23} \\ A_{33} \end{pmatrix} + \mathbf{T}
\]

**Appendix C  Diffusion Tensor Imaging Mathematics**

This appendix serves to give the reader insight into relevant concepts on tensor mathematics, specifically, the rationale behind tensor rotation and solving analytically for the eigenvalues and eigenvectors of a 3-by-3 tensor. The material derives heavily from Joseph C. Kolecki’s and Peter B. Kingsley’s excellent introduction to tensor analysis.""
$$D'_{xy} = abD_{xx} + deD_{yy} + ghD_{zz} + (ae + bd)D_{xy} + (ah + bg)D_{yz} + (dh + eg)D_{yx}$$

$$D'_{xz} = acD_{xx} + dfD_{yy} + giD_{zz} + (af + cd)D_{xz} + (ai + cg)D_{yz} + (di + fj)D_{zx}$$

$$D'_{yz} = bcD_{xx} + efD_{yy} + hiD_{zz} + (bf + ce)D_{yz} + (bi + ch)D_{zx} + (ei + fh)D_{xee}$$

The elements in the $x$, $y$, and $z$-axis in $D$ are specified in the laboratory frame of reference. Note that the same rotation matrices applied for vector rotations can be used for tensor rotations in $\mathbb{R}^3$. There are, however, properties of tensors that are rotationally invariant, they do not change under rotation through an arbitrary angle. Tensors of rank 2 may not just map $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, but may also map between dual spaces under a particular transformation law. If a diagonal tensor with elements $\lambda_x$, $\lambda_y$, and $\lambda_z$ is introduced in place of $D$, the general formula for rotation becomes

$$\begin{pmatrix}
D'_{xz} & D'_{xy} & D'_{xz} \\
D'_{zy} & D'_{yy} & D'_{yz} \\
D'_{az} & D'_{az} & D'_{zz}
\end{pmatrix} = R^T \begin{pmatrix}
\lambda_x & 0 & 0 \\
0 & \lambda_y & 0 \\
0 & 0 & \lambda_z
\end{pmatrix} \times R$$

where the resulting elements are

$$D'_{xz} = a^2\lambda_x + d^2\lambda_y + g^2\lambda_z$$
$$D'_{xy} = b^2\lambda_x + e^2\lambda_y + h^2\lambda_z$$
$$D'_{xz} = c^2\lambda_x + f^2\lambda_y + i^2\lambda_z$$
$$D'_{yz} = a^2\lambda_x + d^2\lambda_y + g^2\lambda_z$$
$$D'_{yy} = b^2\lambda_x + e^2\lambda_y + h^2\lambda_z$$
$$D'_{yz} = c^2\lambda_x + f^2\lambda_y + i^2\lambda_z$$
$$D'_{yz} = ab\lambda_x + de\lambda_y + gh\lambda_z$$
$$D'_{xz} = ac\lambda_x + df\lambda_y + gi\lambda_z$$
$$D'_{yz} = bc\lambda_x + ef\lambda_y + hi\lambda_z$$

The rotation is specified so that the resulting diffusion ellipsoid is aligned with a desired orientation. If the major axis of the ellipse is along the $z$-axis, the rotation matrix and resulting diffusion tensor is

$$R = \begin{pmatrix}
\cos \varphi & -\sin \varphi & 0 \\
\sin \varphi & \cos \varphi & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
\cos \theta & 0 & \sin \theta \\
0 & 1 & 0 \\
-\sin \theta & 0 & \cos \theta
\end{pmatrix} \begin{pmatrix}
\cos \psi & -\sin \psi & 0 \\
\sin \psi & \cos \psi & 0 \\
0 & 0 & 1
\end{pmatrix}$$

$$D = RAR^T$$
C2 Calculation of Eigenvectors and Eigenvalues

When the diffusion tensor has been aligned with the reference frame, the following general equation holds for \( i \in \{1, 2, 3\} \):

\[
\mathbf{D} \mathbf{e}_i = \lambda_i \mathbf{e}_i = \lambda_i \mathbf{I} \mathbf{e}_i
\]

\[
\begin{pmatrix}
D_{xx} & D_{xy} & D_{xz} \\
D_{yx} & D_{yy} & D_{yz} \\
D_{zx} & D_{zy} & D_{zz}
\end{pmatrix}
\begin{pmatrix}
e_{1x} \\
e_{2x} \\
e_{3x}
\end{pmatrix}
\begin{pmatrix}
e_{iz} \\
e_{iy} \\
e_{iz}
\end{pmatrix}
= \lambda_i
\begin{pmatrix}
e_{iz} \\
e_{iy} \\
e_{iz}
\end{pmatrix}
\]

Note that \( \mathbf{D} \mathbf{E} = \mathbf{E} \Lambda \) where \( \mathbf{E} \) is the eigenvector matrix and \( \Lambda \) is the eigenvalue matrix. Therefore,

\[
\begin{pmatrix}
D_{xx} & D_{xy} & D_{xz} \\
D_{yx} & D_{yy} & D_{yz} \\
D_{zx} & D_{zy} & D_{zz}
\end{pmatrix}
\begin{pmatrix}
e_{iz} \\
e_{iy} \\
e_{iz}
\end{pmatrix}
= \lambda_i
\begin{pmatrix}
e_{iz} \\
e_{iy} \\
e_{iz}
\end{pmatrix}
\]

where the matrix \( \mathbf{E}^{-1} = \mathbf{E}^T \). The tensor \( \mathbf{D} \) can now be calculated from the eigenvalue matrix

\[
\mathbf{D} = \mathbf{D} \mathbf{E} \mathbf{E}^T = \mathbf{E} \Lambda \mathbf{E}^T
\]

\[
\mathbf{E}^T \mathbf{E} \Lambda = \Lambda = \mathbf{E}^T \mathbf{D} \mathbf{E}
\]

\( \mathbf{D} \mathbf{e}_i = \lambda_i \mathbf{I} \mathbf{e}_i \Rightarrow (\mathbf{D} - \lambda_i \mathbf{I}) \mathbf{e}_i = 0 \) is the first step to solve for the eigenvalues of a 3-by-3 matrix and equates to

\[
\begin{pmatrix}
D_{xx} - \lambda_i & D_{xy} & D_{xz} \\
D_{yx} & D_{yy} - \lambda_i & D_{yz} \\
D_{zx} & D_{zy} & D_{zz} - \lambda_i
\end{pmatrix}
\begin{pmatrix}
e_{iz} \\
e_{iy} \\
e_{iz}
\end{pmatrix}
= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\]

\[
(D_{xz} - \lambda_i) e_{iz} + D_{xy} e_{iy} + D_{zx} e_{iz} = 0
\]

\[
D_{zy} e_{ix} + (D_{yy} - \lambda_i) e_{iy} + D_{yz} e_{iz} = 0
\]

\[
D_{xz} e_{iz} + D_{zy} e_{iy} + (D_{zz} - \lambda_i) e_{iz} = 0
\]

This set of homogeneous equations has a solution if and only if \( \det(\mathbf{D} - \lambda_i \mathbf{I}) = 0 \) where \( \lambda_1 \geq \lambda_2 \geq \lambda_3 \). The analytic solution to the 3-dimensional cubic equation is given below. For simplicity one can express

\[
\alpha := D_{xx} D_{yy} + D_{yy} D_{zz} + D_{zz} D_{xx} - (D_{xx}^2 + D_{yy}^2 + D_{zz}^2)
\]

\[
\beta := D_{xx} D_{yy} D_{zz} + 2D_{xy} D_{yx} D_{yz} - (D_{xx} D_{yy}^2 + D_{yy} D_{zz}^2 + D_{xx} D_{zz}^2)
\]

\[
\gamma := (D_{xx} D_{yy} - D_{xx} D_{yy} + D_{xx} D_{yy})^2
\]

The analytic solution to the 3-dimensional cubic equation is given below. For simplicity one can express

\[
\alpha := D_{xx} D_{yy} + D_{yy} D_{zz} + D_{zz} D_{xx} - (D_{xx}^2 + D_{yy}^2 + D_{zz}^2)
\]

\[
\beta := D_{xx} D_{yy} D_{zz} + 2D_{xy} D_{yx} D_{yz} - (D_{xx} D_{yy}^2 + D_{yy} D_{zz}^2 + D_{xx} D_{zz}^2)
\]

\[
\gamma := (D_{xx} D_{yy} - D_{xx} D_{yy} + D_{xx} D_{yy})^2
\]
\[
\det(D - \lambda I) = 0 \Rightarrow \lambda^3 - \operatorname{Tr}(D)\lambda^2 + \alpha\lambda - \beta = 0
\]

\[
\lambda_1 = \frac{1}{3} \operatorname{Tr}(D) + \frac{2}{3}[(\operatorname{Tr}(D)/3)^2 - \alpha/3]^{1/2} \cos \theta
\]

\[
\lambda_2 = \frac{1}{3} \operatorname{Tr}(D) - \frac{2}{3}[(\operatorname{Tr}(D)/3)^2 - \alpha/3]^{1/2} \cos(\pi/3 + \theta)
\]

\[
\lambda_3 = \frac{1}{3} \operatorname{Tr}(D) - 2[(\operatorname{Tr}(D)/3)^2 - \alpha/3]^{1/2} \cos(\pi/3 - \theta) = \operatorname{Tr}(D) - \lambda_1 - \lambda_2
\]

where \( \theta = \arccos(\varepsilon/[\operatorname{Tr}(D)/3]^2 - \alpha/3]^{3/2})/3 \) and \( \varepsilon := [\operatorname{Tr}(D)/3]^3 - [\alpha \operatorname{Tr}(D)/6] + \beta/2 \).

The eigenvectors are determined from the following system of homogenous equations,

\[
\begin{pmatrix}
(D_{xx} - \lambda_1)e_{ix} & D_{xy}e_{iy} & D_{xz}e_{iz} \\
D_{yx}e_{ix} & (D_{yy} - \lambda_1)e_{iy} & D_{yz}e_{iz} \\
D_{zx}e_{ix} & D_{zy}e_{iy} & (D_{zz} - \lambda_1)e_{iz}
\end{pmatrix} = 0
\]

where \((D_{xx} - \lambda_1)e_{ix} + D_{xy}e_{iy} = -D_{ix}e_{iz}\) and \(D_{yx}e_{ix} + (D_{yy} - \lambda_1)e_{iy} = -D_{yx}e_{iz}\). Setting the eigenvector \(e_{ix}\) to an arbitrary value, it follows from these set of equations that

\[
e_{ix} = e_{ix}\frac{D_{xy}D_{yx} - (D_{yy} - \lambda_1)D_{xx}}{(D_{xx} - \lambda_1)(D_{yy} - \lambda_1) - D_{yy}^2}
\]

\[
e_{iy} = e_{iy}\frac{D_{xy}D_{yx} - (D_{xx} - \lambda_1)D_{yy}}{(D_{xx} - \lambda_1)(D_{yy} - \lambda_1) - D_{xx}^2}
\]

\[
e_{iz}/\varepsilon_{iz} = e_{iz}/\varepsilon_{iz}\frac{D_{xy}D_{yx} - (D_{yy} - \lambda_1)D_{xx}}{D_{yy}D_{xx} - (D_{xx} - \lambda_1)D_{yy}}
\]

Similar estimations for the other eigenvectors can be made so that

\[
e_{ix} = [D_{xy}D_{yx} - (D_{yy} - \lambda_1)D_{xx}][D_{xx}D_{yy} - (D_{xx} - \lambda_1)D_{yy}]
\]

\[
e_{iy} = [D_{xx}D_{yx} - (D_{xx} - \lambda_1)D_{yy}][D_{xx}D_{yy} - (D_{xx} - \lambda_1)D_{yy}]
\]

\[
e_{iz} = [D_{xy}D_{yx} - (D_{xx} - \lambda_1)D_{yy}][D_{xy}D_{yx} - (D_{xx} - \lambda_1)D_{yy}]
\]

Normalizing the eigenvectors is then a matter of dividing each vector with its corresponding magnitude and then reiterate this process for all eigenvalues:

\[
\hat{e}_i = e_i/\sqrt{e_{ix}^2 + e_{iy}^2 + e_{iz}^2} = e_i/\sqrt{e_i^T e_i}
\]
Appendix D  Axonal Compartment Models

Panagiotaki and colleagues\(^6\) presented the following concepts in a review and taxonomy of axonal compartment models. These models (Table D1) are capable of describing the morphology of white matter fiber tracts while taking the non-Gaussian characteristics of the signal into account. However, they cannot be used to identify precise origins of fibers in the brain.

Fig. D1. Two examples showing the representation of diffusive processes of water in nervous tissue as an isotropic restricted ball model (left) and a diffusion ellipsoid (right) for intra- and extra-axonal water diffusion.

Table D1. List of axonal compartment models tested in Fiberfox. The fiber radius and orientation is denoted by \( r \) and \( n \), respectively. The axial diffusivity \( d_\| \) describes the parallel diffusion along the orientation of the fiber, whereas \( d_\perp \) is the perpendicular diffusivity. Specifically, intra- and extra-axonal models (cf. Fig. D1) are described in spherical coordinates \((\theta, \phi, \psi)\) along the fiber direction in a voxel.

<table>
<thead>
<tr>
<th>Model</th>
<th>Form</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stick</td>
<td>( S = \exp[-bd(n^2)^2] )</td>
<td>( d, \theta, \phi )</td>
</tr>
<tr>
<td>Zeppelin</td>
<td>( \mathbf{D} = \alpha \mathbf{n} \mathbf{n}^T + \beta \mathbf{I} )</td>
<td>( d_|, \beta, \theta, \phi )</td>
</tr>
<tr>
<td>Tensor</td>
<td>( \mathbf{D} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^T + \lambda_3 \mathbf{e}_3 \mathbf{e}_3^T )</td>
<td>( \lambda_1, \lambda_2, \lambda_3, \theta, \phi, \alpha )</td>
</tr>
</tbody>
</table>

Isotropic restricted compartment models

<table>
<thead>
<tr>
<th>Model</th>
<th>Form</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ball</td>
<td>( \mathbf{D} = \alpha \mathbf{I} )</td>
<td>( \alpha )</td>
</tr>
<tr>
<td>Astrosticks</td>
<td>( S_G = \int S, p(n) , dn, r = 0 )</td>
<td>( \alpha )</td>
</tr>
<tr>
<td>Dot</td>
<td>( S = 1, r = 0 )</td>
<td>( \alpha )</td>
</tr>
</tbody>
</table>

Appendix E  MATLAB Scripts

E1  Diffusion Tensor Estimates

%% Import Synthetic Data generated in Fiberfox
% Single-bundle (cylinder) configuration
[S1Cyl, datatbl] = nrrdread('S1Cyl.dwi');
S1Cyl(1:99,:,:,:) = []; % Omit the baseline (b0) image
S1Cent = double(S1Cyl(1:end, 2, 2, 2));

%% Diffusion tensor estimation and eigendecomposition
D = zeros(3,3);
for i = 1:90
    D = D + S1Cent(i)*g(i,:).*transp(g(i,:));
end
[E, lambda] = eig(D);

E2  Load and Save NIfTI Data

%% Read the VTK and DWI file
% The same procedure is used for the baseline image
[S1Cyl, nrrdData] = nrrdread('probDWI.dwi');
gradRows = size(S1Cyl, 1);
b0 = floor(0.10*gradRows); % First 10 percent are due to b0

%% Gradient vector g
bvec1000Art = dlmread('bvec1000_art.txt', '%f');
g = zeros(99, 3); n = 1; m = 3; i = 10;
while n < 270
    subvect = bvec1000Art(n:m);
g(i, :) = subvect';
    n = m+1; m = n + 2; i = i + 1;
end

%% Establish candidate DWI NIfTI file
fiberfox_nii = load_nii('probDWI.nii'); % NIfTI header (3D)
p = load_nii('LiFE_demo_scan1.nii'); % For data comparison

%% Create structure array fields
q.img = SCyl;
[nGrads, xPos, yPos, zPos] = size(SCyl);

% hdr.dime
q.hdr.dime.intent_p1 = 0; % Intent parameters
q.hdr.dime.intent_p2 = 0;
q.hdr.dime.intent_p3 = 0;
q.hdr.dime.intent_code = 0;
q.hdr.dime.datatype = 4;
q.hdr.dime.bitpix = 16; % No. of bits/voxel
q.hdr.dime.slice_start = 0;
q.hdr.dime.slice_end = 0;
q.hdr.dime.scl_inter = 0;
q.hdr.dime.scl_slope = 1; % Data scaling
q.hdr.dime.cal_min = 0;
q.hdr.dime.cal_max = 24731;
q.hdr.dime.glmax = 24731;
q.hdr.dime.glmin = 0;
q.hdr.dime.dim = [4 21 21 21 99 1 1 1];

% hdr.hk
q.hdr.hk = p.hdr.hk;
% hdr.hist
q.hdr.hist.intent_name = '';
q.hdr.hist.magic = 'n+1';
q.hdr.hist.rot_orient = [1 2 3];
q.hdr.hist.flip_orient = [3 3 0];
q.hdr.hist.descr = 'Fiberfox DWI Test';
q.hdr.hist.aux_file = '';
q.hdr.hist.qform_code = 0;
q.hdr.hist.sform_code = 0;
q.hdr.hist.quatern_b = 0;
q.hdr.hist.quatern_c = 0;
q.hdr.hist.quatern_d = 0;
q.hdr.hist.offset_x = -0.6250;
q.hdr.hist.offset_y = -0.6250;
q.hdr.hist.offset_z = 0.6250;
q.hdr.hist.srow_x = [-1.2500 0 0 -0.6250];
q.hdr.hist.srow_y = [0 -1.2500 0 -0.6250];
q.hdr.hist.srow_z = [0 0 1.2500 0.6250];
q.hdr.hist.originator = [21.5000 21.5000 0.5000 0 0];

% original defined after the header is set
q_original = struct('hdr', []);
q_original.hdr = q.hdr;
q_original.hdr = q_original.hdr;

% Filetype, prefix, and machine
q.filetype = 2;
q.fileprefix = 'Fiberfox_DWI_Test';
q.machine = 'ieee-le';

%% Save as NIfTI using save_nii(nii, filename, [old_RGB])
% Save as scan 1
save_nii(q, 'single_fiber_DWI_scan1.nii');
% Save as scan 2
save_nii(q, 'single_fiber_DWI_scan2.nii');
% Save as baseline image (T1-weighted)
save_nii(qb0, 'single_fiber_b0.nii');

E3 Rotations and Concatenation of Coordinates

%% Synthetic fiber coordinates
% Single-bundle configuration (no tractogram)
bndCoord = readVTK('bnd.fib');
% size(bndCoord) % nxmxp, multidimensional array (p dims)
% n = the number of points of the longest fiber, which is proportional to the
% sampling point distance (mm), m = 3
% p is the number of lines, i.e., number of fibers
% Each fiber is represented as 2D n x m matrix along p
coord = bndCoord(1:end,:,1:100); % (x, z, y)
coord(:,2,1) = bndCoord(:,3,1); % y to z
coord(:,3,1) = bndCoord(:,2,1); % z to y

% Spatial evaluation of coordinates
r2 = coord(2,:);'
r3 = coord(3,:);
zcoord = r2(2:3:end);
ycoord = r3(3:3:end);

%% Coordinates derived from stochastic streamline tractography
% Image size: 21x21x21 mm, voxel size: 1.25 mm
tract = readVTK('probDWI.fib');
nrFibers = size(tract, 3);
nrRows = size(tract, 1);
tractCoord = tract(1:end,:,1:1nrFibers);
tractCoord(:,2,1) = tract(:,3,1);
tractCoord(:,3,1) = tract(:,2,1);
```matlab
tempArr = []; 
probTractCoord = cell(nrFibers, 1);
for fiber = 1:nrFibers
    for cols = 1:nrRows
        % Go through every tract and remove NaN entries
        tempArr = tractCoord(:,cols,fiber)';
        tempArr(isnan(tempArr)) = []; 
    end
    % Insert transposed fiber coordinates in the cell matrix
    probTractCoord{fiber, 1} = tempArr;
end
probTract = load('test_life_demo_mrtrix_csd_lmax10_probabilistic');
probTract.coordinateSpace = 'acpc';
probTract.fg.fibers = probTractCoord;
probTract.fg.pathwayInfo = struct('algo_type', zeros(nrFibers,1),
    'seed_point_index', ones(nrFibers,1));

% Save workspace variables
fg = probTract.fg;
versionNum = probTract.versionNum;
coordinateSpace = probTract.coordinateSpace;
save('probabilistic_tract','fg','versionNum','coordinateSpace');

%% Rotating streamline coordinates about the y-axis
theta_c = 90; % Crossing angle
colleft = 13; 
colright = 14;
medVal = 0.5*(probTractCoord{1}(:,colleft) + ... 
    probTractCoord{1}(:,colright));
for fiber = 1:nrFibers
    for row = 1:nrRows 
        tempRotArr = roty(theta_c)*(tractCoord(row,:,fiber)' - ... 
            medVal) + medVal; 
        roty_tractCoord(row,:,fiber) = tempRotArr';
    end
end

%% Fit coordinates to parabola, p(x)
% p(x) = x^2/n + medVal; curvature, kappa = (2*n^2)((4*x^2)+n^2)^-1.5 for 
% each parabola where n is (20,40,60,80,100) and the focal distance is n/4
n = 20; 
tractCoordPar = zeros(nrRows, 3, nrFibers);
for fiber = 1:nrFibers
    y_curve = tractCoord(:,1,fiber);
    a = 1/n; y_curve = a * y_curve.^2 + medVal;
    colTemp = y_curve(1:size(y_curve,1)/2); colTemp = colTemp';
    colTemp = fliplr(colTemp)'; colTemp_flip = fliplr(colTemp');
    colTemp_flip = colTemp_flip';
    % Concatenate with respect to the parabola's axis of symmetry
    catCol = cat(1,colTemp, colTemp_flip); catCol(13:14) = 13.1250;
    tractCoordPar(:,1,fiber) = tractCoord(:,1,fiber);
    tractCoordPar(:,2,fiber) = catCol; 
    tractCoordPar(:,3,fiber) = tractCoord(:,3,fiber);
end

%% Concatenation of the cell arrays probTractCoord and probTractCoordRot
probTractCoordAppend = cat(1, probTractCoord, probTractCoordRot);
```