New Algorithms for Local and Global Fiber Tractography in Diffusion-Weighted Magnetic Resonance Imaging

Dissertation
zur Erlangung des
mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen
im Promotionsprogramm Mathematik
der Georg-August University School of Science (GAUSS)

vorgelegt von
Helen Schomburg
aus Göttingen

Göttingen, 2017

Betreuungsausschuss

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Jens Frahm
Biomedizinische NMR Forschungs GmbH
Max-Planck-Institut für biophysikalische Chemie, Göttingen

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Thorsten Hohage

Koreferent: Prof. Dr. Jens Frahm

Weitere Mitglieder der Prüfungskommission:

Jun.-Prof. Dr. Andrea Krajina
Institut für Mathematische Stochastik
Georg-August-Universität Göttingen

Prof. Dr. Russell Luke
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Martin Uecker
Institut für Diagnostische und Interventionelle Radiologie
Universitätsmedizin Göttingen

Tag der mündlichen Prüfung: 29. September 2017

Contents

1. Introduction

Fiber tractography based on *diffusion-weighted magnetic resonance imaging* (DW-MRI) is to date the only method for the three-dimensional visualization of nerve fiber bundles in the living human brain noninvasively. Therefore, it is an invaluable tool for both clinical practice and medical research. Applications include neurosurgical planning as well as research on human brain anatomy, connectomics, and disorders of the nervous system, such as Alzheimer's disease and multiple sclerosis. Since its introduction at the end of the 20th century, various strategies for fiber tractography have been developed. Early streamline tractography methods, for instance the approach described in the article by Basser et al. (2000), reconstruct in each image voxel a symmetric positive definite *diffusion tensor* (DT) from the diffusion-weighted magnetic resonance measurements and derive fiber tracks by iteratively stepping from predefined starting points and with a certain step length in the direction of the tensor's principle eigenvector. Tensor-based methods, unable to describe more than one fiber population in a voxel, were succeeded by techniques using *orientation distribution functions* (ODFs) of an angular resolution high enough to resolve multiple fiber populations per voxel. Similarly, streamline tractography follows the direction maximizing the local ODF, as for instance presented in the work by Descoteaux et al. (2009). In a previous thesis by Rügge (2015), an ODF reconstruction method was developed. Taking one step further, the present thesis aims at reconstructing nerve fiber bundles from ODF data.

Due to simplicity and computational efficiency, streamline tractography is widely used. However, a drawback inherent in such line propagation methods operating on DW-MR data with low signal-to-noise ratio is the stepwise accumulation of measurement errors. These inaccuracies often result in the reconstruction of fiber tracks which are anatomically incorrect or implausible. The main objective of this thesis is the development of improved tractography strategies which yield an increased proportion of anatomically correct tracks. We present a semi-local approach based on the fast streamline method which includes information about the local neighborhood to increase robustness against noise. For this purpose, candidate tracks are generated and assigned an a-posteriori probability by means of Bayesian statistics. We model the likelihood to represent plausibility of a candidate track using the underlying ODF data. The prior distribution includes

information about track curvature. Moreover, the prior can be used to incorporate anatomical information such as tissue type. From this Bayesian framework we derive both a deterministic tractography algorithm by maximum a-posteriori estimation, and a probabilistic algorithm by drawing from the marginalized posterior distribution. Both algorithms are tested and compared to state-of-the-art methods on data obtained from computer simulations and on MR scans of diffusion phantoms and of a human brain *in vivo*.

The question that now arises is how well fiber tracks which were reconstructed by a local or semi-local method describe the underlying diffusion data globally. A global tractography approach which seeks to filter out invalid tracks in a postprocessing step is introduced in the article by Daducci et al. (2015). There, a large number of fiber tracks is generated from the data and each track contributes individually to a multi-compartment model of the DW-MR signal. The modeled signal is fitted to the measured DW-MR signal by solving a large-scale linear least squares problem

$$\min_{x \geq 0} \|Ax - y\|_2^2$$

with an ill-conditioned dictionary matrix $A \in \mathbb{R}^{m \times n}$, weight vector $x \in \mathbb{R}^n$, and measured data $y \in \mathbb{R}^m$. In a simplified scheme, the vector $x = [x_1|x_2]^\top$ contains fiber weights x_1 and voxel weights x_2. The authors suggest the use of ℓ_1-norm regularization to promote sparseness of the weight vector x. Thus, fiber tracks which do not support the measured signal globally can be discarded. In consideration of the used multi-compartment model, ℓ_1-norm regularization is appropriate for x_1, but regularization promoting spatial smoothness is better suited for the voxel weights x_2. In this thesis, we derive an improved version of the method stated in Daducci et al. (2015) by adding H^s-Sobolev-norm regularization terms. We solve the resulting minimization problem of the form

$$\min_{x \geq 0} \{\|Ax - y\|_2^2 + \alpha_1 \|x_1\|_1 + \alpha_2 \|x_2\|_{H^s}^2\},$$

$\alpha_1, \alpha_2 > 0$, using the *alternating direction method of multipliers* (ADMM) described by Boyd et al. (2011). The matrix A, though relatively sparse, is very large in practice. Exploiting the special structure of A, we develop strategies for the efficient solution of large systems of equations arising in the ADMM method based on dimension reduction using truncated singular value decomposition of certain matrix blocks. The algorithm is tested and compared on data obtained from computer simulations, MR scans of a diffusion phantom, and *in vivo* MR scans.

The thesis is structured as follows: In **Chapter 2**, we summarize the basic principles of MRI and diffusion-weighted MRI, and motivate the use of DW-MRI as

a reconstruction tool for nerve fiber bundles by outlining the diffusion properties of nerve tissue in the brain. Then, we briefly consider the estimation of diffusion tensors and orientation distribution functions from the DW-MRI data, and the computation of different image contrasts that can be derived from MRI and DW-MRI data and will be used in the course of this work. Furthermore, we give a detailed introduction of fiber tractography where we group and describe different classes of algorithms that can be found in literature. Finally, we summarize the approach of modeling white matter microstructure using multi-compartment diffusion models.

Chapter 3 gives an overview of the different data sets used to test the algorithms developed in the following chapters.

In **Chapter 4**, we introduce both a deterministic and a probabilistic neighborhood-informed tractography algorithm derived from a Bayesian framework. Experiments and results on simulated, phantom, and *in vivo* data are illustrated and discussed. Moreover, a literature review examines related work.

Chapter 5 starts by explaining the multi-compartment model and global model-fitting problem proposed in Daducci et al. (2015). Then, we motivate the use of H^s-norm regularization and state the resulting minimization problem. We describe the theory of ADMM based on the article by Boyd et al. (2011), and derive an algorithm that solves our minimization problem. Furthermore, we propose strategies for efficient computation. Finally, we present and discuss numerical experiments and results using simulated, phantom and *in vivo* data.

We conclude with a summary in **Chapter 6**.

2. Diffusion Magnetic Resonance Imaging and Fiber Tractography

The human brain contains a dense network of nerve fibers responsible for transmission and processing of information. Knowledge about the location of these neuronal pathways gives information about connectivity between different brain parts, for instance by deriving a map of connections. This information can then be used to better understand and treat neurological disorders, or can be helpful for planning a surgery. In the brain, diffusion of water is mostly directional, or *anisotropic*, along fiber bundles, and free, or *isotropic*, elsewhere. *Diffusion-weighted magnetic resonance imaging* (DW-MRI) is the only technique that can noninvasively detect neuronal pathways in the living human brain by measuring these diffusion properties. In Sections 2.1 and 2.2, we briefly explain the idea behind magnetic resonance imaging (MRI) and diffusion-weighted MRI, respectively. Section 2.3 reviews diffusion properties in the brain and motivates the use of diffusion MRI as a tool to infer the pathways of nerve fiber bundles. *Diffusion tensor imaging* (DTI), a technique to obtain the principle diffusion direction from the diffusion-weighted measurements, is briefly summarized in Section 2.4. Since DTI requires relatively few diffusion-weighted measurements it is widely used in clinical applications, yet at the cost of inaccuracies due to imprecise model assumptions and the inability of resolving more than one fiber population per voxel. *High angular resolution diffusion imaging* (HARDI) is used to overcome the drawbacks of DTI. In Section 2.5, we consider HARDI-based fiber *orientation distribution function* (ODF) reconstruction techniques employing spherical deconvolution. Important image contrasts obtained from anatomical MR images, DTI, and ODF reconstructions, which will be useful in Section 2.7 and in following chapters, are summarized in Section 2.6. In Section 2.7, we review basic ideas and recent developments of fiber tractography, with the main focus on ODF-based deterministic and probabilistic streamline methods. Finally, in Section 2.8, we regard multi-compartment models which, in contrast to DT and fiber ODF studied in Sections 2.4 and 2.5 that only represent diffusion or fiber density averaged over a whole voxel, seek to explain the underlying tissue microstructure.

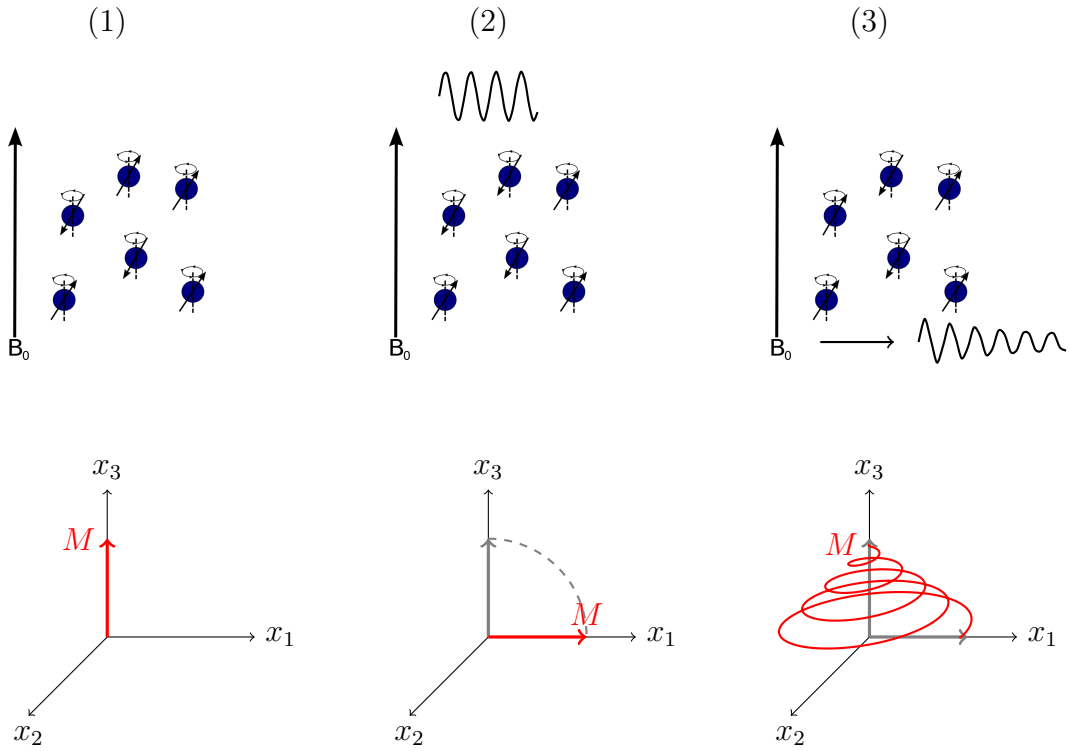


Figure 2.1.: *Above*: Behaviour of hydrogen nuclei, represented as blue dots, in an MR scanner with static magnetic field B_0. The arrows pointing up and down indicate the protons' spin axes at different energy states (here, up: low, down: high), the circular arrows indicate precession around the axes aligned with the direction of B_0. (1) The protons in the MR scanner are in equilibrium state. They are aligned with the magnetic field and in precession. (2) An RF-pulse results in the excitation of spins from low to high energy state. (3) Protons move back from high to low energy state while emitting electromagnetic energy. *Below*: Corresponding net magnetization.

2.1. Basic Principles of MRI

Human tissue consists of about 60% water, and the hydrogen nuclei thereof constitute more than 90% of the overall hydrogen nuclei in the human body. Hydrogen nuclei consist of single protons, which are *spinning* around their own axis. Due to its spin, each proton has an *angular momentum* and a *magnetic moment*. These two properties lay the foundation for MR measurements. Though not solely inherent in hydrogen but in all isotopes with an odd number of protons and/or neutrons, the behaviour of hydrogen nuclei is mainly what is measured in MRI due to the abundance of water in the body. In the following, we discuss the basic ideas of MRI. For a more detailed description, we refer to the books by Slichter (1990) and Haacke et al. (1999).

In an MR scanner where a static magnetic field with strength B_0 is applied, protons can be viewed as being in *precession*, which means that they rotate - in addition to their spin - with their rotation axis around the direction of the B-field. The precession has a certain frequency called *Larmor frequency*, $\omega = \gamma B_0$, which depends only on the field strength and the gyromagnetic ratio γ. In the presence of a static magnetic field, the spins are aligned with the magnetic field and are in one of two different states, the parallel (or low-energy) or the antiparallel (or high-energy) state. In fact, a little more than half of the protons are in the low-energy state, whereas the remaining protons are in the high-energy state. This so called *initial* or *equilibrium state* is exemplarily illustrated in Figure 2.1 (1). The space and time dependent *net magnetization* $M : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ is proportional to the difference of the overall amount of parallel and antiparallel spins in the scanner. Without a magnetic field, the spins are oriented randomly resulting in $M = 0$. In the equilibrium state where the spins are aligned with the magnetic field, spins with opposite alignment cancel out, but $|M| > 0$ as more spins are in the low-energy state. We define the MR scanner coordinate system such that the x_3-axis is aligned with the direction of the static magnetic field. M initially points in x_3-direction, also called *longitudinal* direction. Then, a short radio frequency pulse (RF-pulse) with Larmor frequency is applied to the static magnetic field. As shown in Figure 2.1 (2), the electromagnetic energy emitted from the RF coils results in excitations of the protons from low to high energy states. Correspondingly, the direction of the net magnetization precesses with Larmor frequency around the x_3-axis while spiraling out into the x_1-x_2-plane perpendicular to the direction of the magnetic field, also called the *transverse* plane.

After the RF-pulse has been switched off again, the protons move back to their initial states while emitting electromagnetic energy, as illustrated in Figure 2.1 (3). This restoration of the equilibrium is called *spin relaxation*. A distinction is made between longitudinal and transverse relaxation. Longitudinal relaxation, or T_1 *recovery*, is the restoration process of the net magnetization in direction of the magnetic field due to spins returning to the low-energy state. At the same time, transverse relaxation, or T_2 *decay* proceeds, which can be described as a decrease of transverse magnetization due to phase decoherence of spins. In the following, time constants T_1 and T_2 indicate the time required for T_1 recovery and T_2 decay, respectively. Let $B : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ denote the sum of the static magnetic field with field strength B_0 and a time-dependent, spatially varying gradient field

$$B_G : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3, \quad B_G(x,t) = (G(t)^\top x)e_3$$

with $G : \mathbb{R} \to \mathbb{R}^3$ and $e_3 = (0,0,1)^\top$, that is necessary to encode spatial information in the MR signal. The change in net magnetization with time explained

above is summarized by the *Bloch equation*

$$\frac{dM(x,t)}{dt} = \gamma M(x,t) \times B(x,t) + \begin{pmatrix} -M_1(x,t)/T_2 \\ -M_2(x,t)/T_2 \\ \left(M^0(x) - M_3(x,t)\right)/T_1 \end{pmatrix} \tag{2.1}$$

that was derived in the article by Bloch (1946). Here, $\times$ denotes the cross product, M^0 denotes the initial magnetization where all spins are at the equilibrium state, and M_1, M_2, M_3 are the three components of M.

Basically, the signal at time t that is measured in MRI is given by the total transverse magnetization,

$$S(t) = \int_{\mathbb{R}^3} M_\perp(x,t)\mathrm{d}x,$$

where the transverse magnetization $M_\perp$ is defined by

$$M_\perp(x,t) := M_1(x,t) + iM_2(x,t).$$

Solving the partial differential equation (2.1) yields

$$M_\perp(x,t) = M_\perp(x,0)e^{-t/T_2}e^{-\gamma B_0 t}e^{-i\gamma\int_0^t B_G(x,t)\mathrm{d}t}$$

for the transverse magnetization. Note, that $M_\perp(x,0)$ is proportional to the density of hydrogen atoms. Omitting the scaling parameters e^{-t/T_2} and $e^{-\gamma B_0 t}$ we obtain for the MR signal the commonly used representation

$$S(t) = \int_{\mathbb{R}^3} M_\perp(x,0)e^{-i\gamma\int_0^t G(t)^\top x\ \mathrm{d}t}\mathrm{d}x.$$

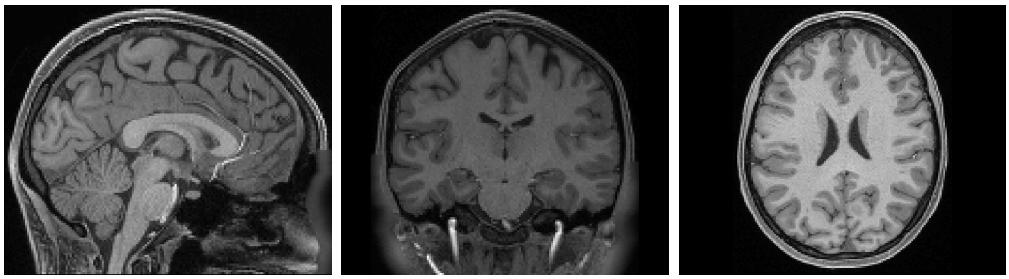
In practice, the MR signal is usually measured as 2D slices, and in *k-space* (Twieg, 1983). For instance, let z denote the x_3-coordinate of the scanned slice. Then, the MR signal of a 2D slice in the transverse plane can be represented as

$$S(z,t) = \int_{\mathbb{R}^2} M_\perp(x,0)\exp\left(-i2\pi k(t)^\top x\right)\Big|_{x_3=z}\mathrm{d}x_1\mathrm{d}x_2, \tag{2.2}$$

with

$$k : \mathbb{R} \to \mathbb{R}^3, \quad k(t) = \frac{\gamma}{2\pi}\int_0^t G(\tau)\mathrm{d}\tau,$$

where G is chosen such that k is sampled on a uniform grid. A 2D inverse Fourier transform of (2.2) yields the respective MR image slice. Different MR



(a) sagittal plane **(b)** coronal plane **(c)** transversal plane

Figure 2.2.: T1w acquisition of the HCP data set (see Sec. 3.3).

contrasts can be obtained by applying suitable sequences of excitation pulses. An important example for a structural contrast that can be generated from MRI by exploiting relaxation time properties of different tissue types is the T_1-weighted image, which can be used to differentiate between different nerve tissue types. In Figure 2.2, different slices through a T_1-weighted image of a human brain are shown. Note that anatomical terms of location are summarized in Appendix A. Apart from identifying static properties such as brain anatomy, MRI can also be used to measure the *movement* of atomic nuclei. In the next section, we discuss the basic idea of *diffusion* MRI, where the concepts summarized above are extended to measure the *diffusion* of water in tissue.

2.2. Diffusion MRI

Diffusion, or *Brownian motion*, describes the process of random motion of particles due to thermal energy. Diffusion MRI is a variant of MRI that provides the possibility of measuring the diffusion of water molecules in biological tissue. As for conventional MRI, we regard the effects on protons in a magnetic field. The measurement sequence can be summarized as follows. First, an RF-pulse results in the excitation of spins. All spins are in precession, aligned, and at the same frequency. In the following, let the 2-sphere be defined by $\mathcal{S}^2 = \{x \in \mathbb{R}^3 : \|x\|_2 = 1\}$. At time t_0, a diffusion sensitizing gradient pulse in direction $g \in \mathcal{S}^2$ is applied with duration δ. This so called *dephasing* gradient introduces a magnetic field inhomogeneity, inducing spin precession at different frequencies. While the magnetic field regains homogeneity, some of the protons are in motion due to diffusion. At time t_1, a second, invertive gradient pulse of same strength and duration is applied to the B-field. This gradient pulse is called *rephasing*, since it results in phases of static protons being the same as in the beginning, and in different phases of protons which have moved. The overall signal attenuation compared to the initial signal is measured, providing

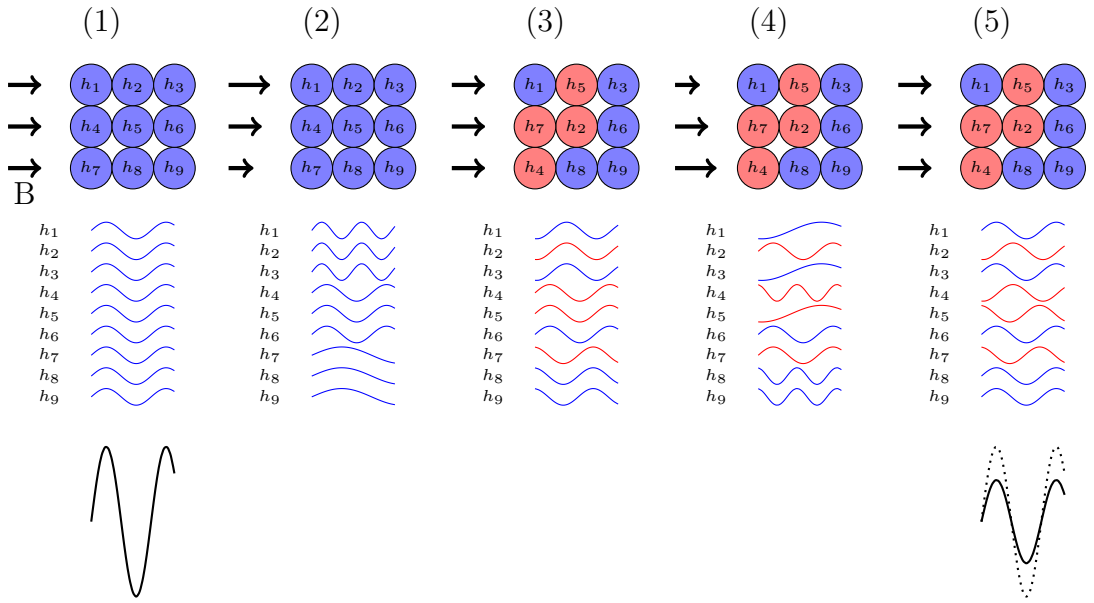


Figure 2.3.: Basic concept of diffusion MRI signal attenuation. (1) All protons h_i are in precession at the same frequency in a homogeneous magnetic field. (2) A diffusion sensitizing gradient pulse is applied resulting in the precession at different frequencies. (3) The magnetic field regains homogeneity. Some protons (h_2, h_4, h_5, h_7) have moved due to self-diffusion. (4) The invertive gradient pulse is applied. (5) Signal attenuation since the protons in motion are in precession at a different phase.

insight into the diffusion process. A simplified illustration of the basic process of diffusion MRI is given in Figure 2.3.

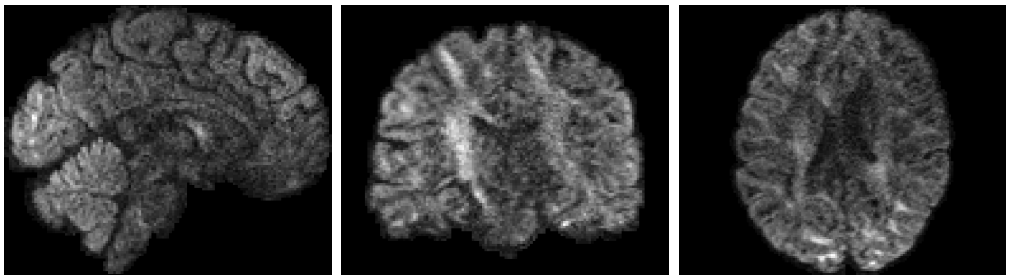
If the duration δ of the gradient pulses is assumed to be infinitely narrow, that is, Dirac delta functions, a transversal slice of the measured DW-MR signal in k-space can be expressed as

$$\hat{S}(q,t,z) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^2} M_\perp(x,0)p(x+r|x,t)e^{-2\pi ik(t)^\top x}e^{-2\pi iq^\top r}\Big|_{x_3=z}\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}r,$$

where $M_\perp$ and k are defined as in (2.2), and $q = \gamma\delta g$. The *diffusion propagator* $p(x+r|x,t)$ describes the probability that a particle that is at position $x \in \mathbb{R}^3$ at time t_0 has at time t_1 drifted to position $x+r$, $r \in \mathbb{R}^3$. A preprocessing step yields the more commonly used representation of the signal as 3D Fourier transform of the diffusion propagator

$$S(x,q) = S_0(x)\int_{\mathbb{R}^3} p(x+r|x,t)e^{-2\pi iq^\top r}dr, \tag{2.3}$$

where S_0 is the MR signal without applied gradient. In the following, we are



(a) sagittal plane **(b)** coronal plane **(c)** transversal plane

Figure 2.4.: Diffusion MRI acquisition of the HCP data set (for one single gradient direction; see Sec. 3.3).

interested in $p(x + r|x, t)$.

In an unrestricted homogeneous medium, the diffusion propagator can be represented by a Gaussian distribution

$$p(x + r|x, t) = \frac{1}{\sqrt{(2\pi\sigma^2)^3}} e^{-\frac{r^\top r}{2\sigma^2}}. \tag{2.4}$$

The average distance σ a particle travels in a period of time t is related to the *diffusion coefficient* d by Einstein's equation $\sigma^2 = 2 \cdot d \cdot t$. Hence, (2.4) can be written as

$$p(x + r|x, t) = \frac{1}{\sqrt{(4\pi d(t_1 - t_0))^3}} e^{-\frac{r^\top r}{4d(t_1 - t_0)}}. \tag{2.5}$$

Substituting (2.5) in Equation (2.3), we obtain

$$S(x, q) = S_0(x) \exp(-bd)$$

with $b = (t_1 - t_0)|q|^2$. The *b-value* b is an important acquisition parameter in diffusion MRI that indicates the amount of diffusion weighting.

The first diffusion-weighted imaging experiments based upon the concepts described above were performed in the mid-1980s by Le Bihan and Breton (1985), Merboldt et al. (1985), and Taylor and Bushell (1985). Since then, Diffusion MRI has been widely used for imaging the movement of water molecules in biological tissue *in vivo*. Figure 2.4 shows sample slices of a diffusion MRI acquisition of a human brain.

The diffusion-weighted MRI signal is in the spatial domain acquired on a three-

(a) Neuron. **(b)** Brain tissue. **(c)** Axon bundles.

Figure 2.5.: Different levels of brain organization. 2.5a: Schematic drawing of a neuron, adapted from commons.wikimedia.org. 2.5b: Coronal slice of a human brain. Figure taken from www.neuroanatomy.ca. 2.5c: Electron micrographs of transverse (above) and longitudinal (below) sections of axons in the monkey's brain. Figures taken from www.bu.edu[1].

dimensional Cartesian grid which we define by

$$\begin{aligned}\Omega_h = &\{x_i = ih_x : i = 1, \dots, N_x\} \\ &\times \{x_j = jh_y : j = 1, \dots, N_y\} \\ &\times \{x_k = ih_z : k = 1, \dots, N_z\} \subset \mathbb{R}^3\end{aligned} \tag{2.6}$$

with voxel-size h_x, h_y, h_z in $x-$, $y-$, and $z-$direction, respectively. The resolution is usually of the order of mm. In the following, a *voxel* denotes a three-dimensional grid cell

$$\left[x_i - \frac{h_x}{2}, x_i + \frac{h_x}{2}\right] \times \left[x_j - \frac{h_y}{2}, x_j + \frac{h_y}{2}\right] \times \left[x_k - \frac{h_z}{2}, x_k + \frac{h_z}{2}\right].$$

Furthermore, the diffusion-weighted signal is usually acquired for approximately 32 to 90 different gradient directions.

2.3. Diffusion in the Brain

Nerve cells, or neurons, are the structural and functional core components of the brain. They connect to networks where information is transmitted and processed via electrical and chemical excitation. The basic parts of a neuron - cell body, dendrites, axon, and synapses - are exemplarily illustrated in Figure 2.5a. Dendrites and synapses are in contact with other neurons to receive and

[1] **The Fine Structure of the Aging Brain**, **Authors: Alan Peters and Claire Folger Sethares**, Boston University School of Medicine, 72 East Newton Street, Boston, MA 02118, www.bu.edu/agingbrain, Supported by the Institute on Aging of the National Institute of Health, grant number P 01-AG 000001

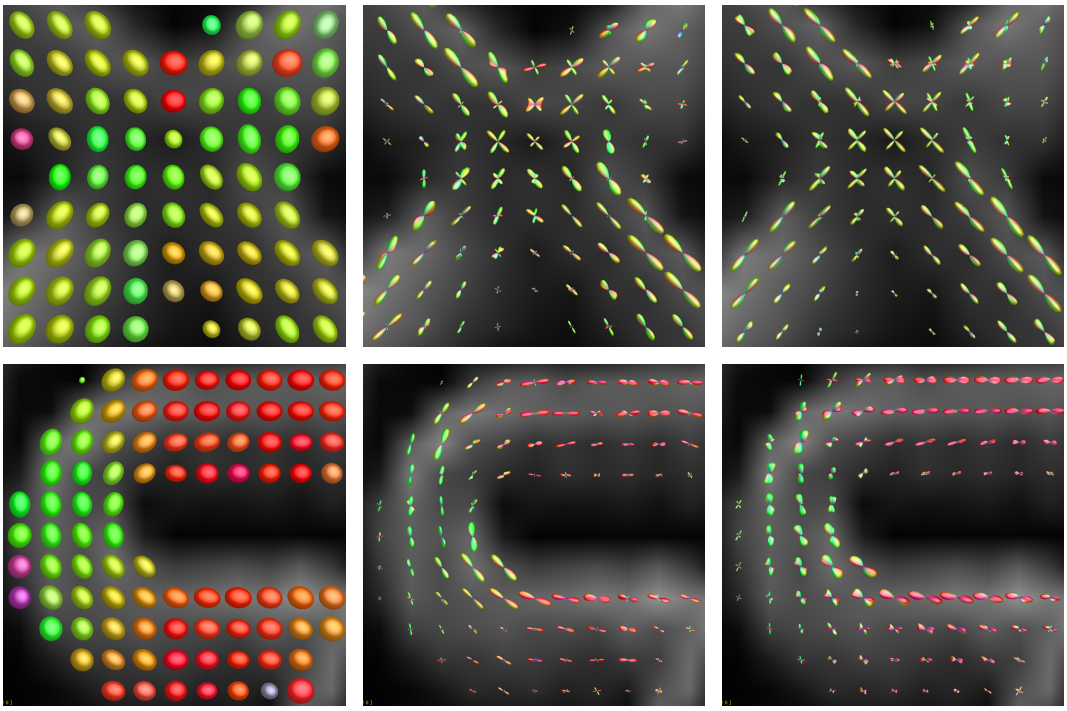
transmit information, respectively. The axon conducts received signals from the dendrites towards the synapses. Axons are often *myelinated* which accelerates signal transmission. Overall, the human brain contains approximately 86 billion nerve cells. The cells are structured and densely packed, forming *gray matter* (GM) and *white matter* (WM) regions. Gray matter contains mainly cell bodies and dendrites, whereas white matter is primarily composed of axons arranged in bundles connecting GM regions. In Figure 2.5b, a coronal slice of a human brain is shown where these two tissue types can be distinguished as lighter (WM) and darker (GM) regions. The surrounding areas and hollow structures of nerve tissue are occupied by *cerebrospinal fluid* (CSF).

Figure 2.5c shows electron microscopy images of transversal and longitudinal sections of myelinated axons in the white matter of a monkey's brain. It can be observed that fibre bundles in white matter consist of coherently organized axons whose cell membranes form longitudinal boundaries. Experiments verified that anisotropic diffusion in the brain is mainly due to these barriers resulting from densely packed axon membranes (Beaulieu, 2002). Water can diffuse relatively fast along the aligned axons, both in intracellular and extracellular space, whereas diffusion perpendicular to the fiber bundles is highly restricted. Hence, diffusion MRI measurements of anisotropic diffusion in white matter can be used to probe the pathways of nerve fiber bundles *in vivo*.

The measurement time of diffusion along a certain gradient direction is usually selected in the range from 10 ms to 100 ms. In the brain, the diffusion constant of water is given by $d \approx 1 \cdot 10^{-3} \frac{\text{mm}^2}{\text{s}}$ (Mori, 2007, Chapter 3). Hence, during measurement time, the average distance water molecules diffuse can be approximated by 4-14 μm using Einstein's equation. Due to the special cellular architecture of white matter, where most axons have a diameter of approximately 1 μm and the length of axon bundles is larger than 2 mm, measurement time is sufficient to capture diffusion anisotropy. However, as the spatial resolution of diffusion MRI is 2-3 mm, the diffusion measurement yields no microscopic diffusion information but the average diffusion information within a voxel.

2.4. Diffusion Tensor Imaging

Various diffusion models were introduced to derive the diffusion properties of tissues from the diffusion-weighted measurements. A pioneering role in diffusion MRI has the widely used *diffusion tensor* (DT) model (Basser et al., 1994). Here,



(a) DT **(b)** CSD-ODF **(c)** spat. reg. ODF

Figure 2.6.: Comparison of different voxel-wise diffusion imaging techniques.

diffusion propagator $p(x+r|x,t)$ is modeled as a Gaussian distribution

$$p(x+r|x,t) = \frac{1}{\sqrt{(4\pi t)^3 \det(D(x))}} \exp\left(-\frac{r^\top D^{-1}(x) r}{4t}\right) \tag{2.7}$$

with symmetric positive definite diffusion tensor $D : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$. Substituting (2.7) in Equation (2.3), the signal simplifies to

$$S(x,q) = S_0(x) e^{-bq^\top D(x) q}.$$

To estimate the diffusion tensor $D(x)$ in each voxel, which has six degrees of freedom due to symmetry, at least six diffusion MR measurements with different gradient directions $(S(x,q_1),\ldots,S(x,q_6))$, and the unweighted signal $S_0(x)$ are necessary. The eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and respective eigenvectors v_1, v_2, v_3 of $D(x)$ are computed to obtain diffusion properties of the respective voxel. If $\lambda_1 > \lambda_2, \lambda_3$ the tensor indicates an anisotropic voxel, where the eigenvector corresponding to λ_1 points in the main diffusion direction. On the other hand, similar eigenvalues ($\lambda_1 \approx \lambda_2 \approx \lambda_3$) imply an isotropic region. The diffusion tensor is usually represented by an ellipsoid, whose axes coincide with the eigenvectors scaled by their corresponding eigenvalues. Computing the tensor for each voxel

yields a tensor field as shown in Figure 2.6a.

An important limitation to DT imaging is, that only one main diffusion direction can be represented in a voxel. This is problematic since due to the coarse resolution of diffusion MRI, the diffusion signal in about one third of all voxels in white matter is caused by crossing fiber bundles (Behrens et al., 2007), but the anisotropic diffusion occurring along individual bundles cannot be resolved by DTI. Figure 2.6a illustrates this problem by means of a DT field computed from physical phantom data with regions of crossing of fiber bundles.

2.5. High Angular Resolution Diffusion Imaging

Various methods have been developed to overcome the limitations of DTI. These techniques are based upon high angular resolution diffusion imaging (HARDI), where diffusion MR data are acquired for a much larger number of, usually 32 to 90, different gradient directions to resolve crossing fibers. From this data, an orientation distribution function (ODF), which describes the diffusion quantity (Jansons and Alexander, 2003; Tuch, 2004; Wedeen et al., 2005) or fiber density (Tournier et al., 2004) in each direction, is reconstructed for each voxel. In the following we consider a reconstruction method using constrained spherical deconvolution (CSD) that was introduced in (Tournier et al., 2004), and a spatially regularized variant of this method described in (Hohage and Rügge, 2015). Both methods are used as input data to test our newly developed tractography algorithms presented in this thesis.

2.5.1. Fiber ODF Estimation with CSD

In the following, let the diffusion data be given on a three-dimensional domain $\Omega \subset \mathbb{R}^3$. We consider a single voxel $x \in \Omega$ in white matter where N different fiber populations are located. Assuming that the diffusion characteristics of all these fiber populations are identical, Tournier et al. (2004) model the DW-MRI signal $S : \mathcal{S}^2 \to \mathbb{R}$ in that voxel as the sum of signal contributions of N individual fiber populations

$$S(q) = \sum_{i=1}^{N} f_i R_{(i)}(q) \tag{2.8}$$

where the axially-symmetric *response function* $R_{(i)}\colon \mathcal{S}^2 \to \mathbb{R}$, whose z-axis is aligned with fiber population i, describes the attenuation of the diffusion weighted signal originating from fiber population i, and $f_i \geq 0$ denotes the volume fraction of fiber population i. Representation (2.8) can be expressed for an arbitrary voxel $x \in \Omega$ as the spherical convolution operation

$$T\psi(x,q) := \int_{\mathcal{S}^2} K(q \cdot u)\psi(x,u)\, du \tag{2.9}$$

of kernel function $K\colon [-1,1] \to \mathbb{R}$ and fiber orientation distribution function (fODF) $\psi\colon \Omega \times \mathcal{S}^2 \to \mathbb{R}$, where T denotes the convolution operator. The fODF $\psi\colon \Omega \times \mathcal{S}^2 \to \mathbb{R}$ characterizes at a point $x \in \Omega$ the fiber density in each direction $q \in \mathcal{S}^2$. Given the response function K the fODF can be computed voxel-wise using constrained spherical deconvolution techniques, as for instance described in Tournier et al. (2007). Various heuristics for approximating the response function exist. For example, K can be selected as the Gaussian diffusion propagator defined in (2.7) with anisotropic diffusion tensor, or it can be estimated from the DW-MRI data, as suggested in Tournier et al. (2004). Figure 2.6 illustrates that, as opposed to DTI, the fiber ODF can represent multiple fiber orientations per voxel.

Spatial Regularization

Due to the characteristically low SNR of DW-MRI images, employing some regularization strategy is beneficial for the reconstruction of the fODF. The spatial regularization method described in (Hohage and Rügge, 2015) supports smoothness along fibers assuming that the point $(x,u) \in \mathbb{R}^3 \times \mathcal{S}^2$ belongs to a structure locally oriented along the horizontal line

$$(-\epsilon, \epsilon) \ni t \mapsto (x + tu, u)$$

for ϵ sufficiently small. The authors define the *horizontal derivative* by

$$D_{\mathrm{hor}}\psi(x,u) := u \cdot \nabla\psi(x,u),$$

and the *horizontal norm* by

$$\|\psi\|^2_{\mathrm{hor}} := \|\psi\|^2 + \|\mathrm{grad}_x\, \psi\|^2 + \|D_{\mathrm{hor}}\psi\|^2,$$

where grad is the surface gradient on $\mathcal{S}^2$. Then the regularized fODF reconstruction given the diffusion-weighted MR signal S is the solution to the constrained

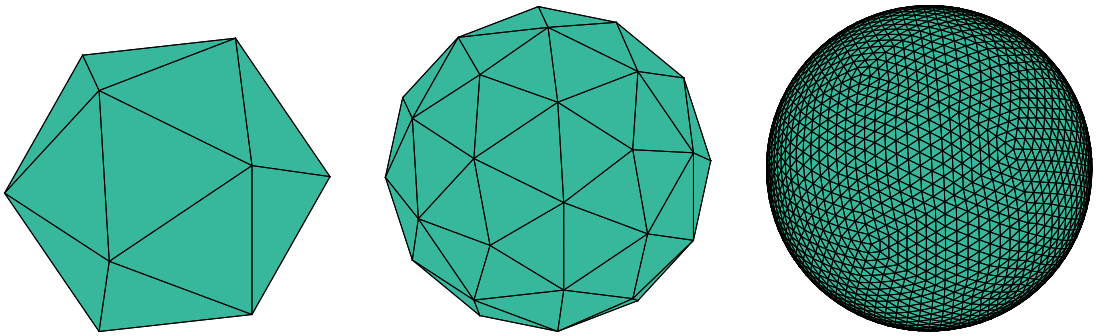


Figure 2.7.: Triangulations of the unit sphere derived after 0, 1 and 4 (*from left to right*) refinement steps.

Tikhonov optimization problem

$$\psi_\alpha := \underset{\psi \geq 0}{\operatorname{argmin}} \left\{ \|T\psi - S\|^2 + \alpha \|\psi\|_{\text{hor}}^2 \right\}.$$

with convolution operator T as defined in (2.9), and regularization parameter $\alpha > 0$. Comparing Figures 2.6b and 2.6c, the benefits of spatially regularized ODFs can be observed, especially in the fiber crossing structure. On the other hand, artifacts may occur in structures of high curvature, as can be discovered in the picture below in Figure 2.6c.

2.5.2. Discrete Representation

Let a set of N_u sampling points distributed on the unit sphere be given by the subset $\mathcal{U} \subset \mathcal{S}^2$. We obtain the sampling points by an iterative refinement of a uniform icosahedron which is centered at the origin and whose circumscribed sphere has a radius of one. In each iteration step, each triangle is subdivided into four smaller, congruent triangles and the coordinates of the resulting new vertices are normalized such that all vertices are located on the unit sphere. In Figure 2.7, a uniform icosahedron and the derived triangular tessellations obtained after one and four iterations are shown. A discrete version of the ODF ψ at each spatial position $x \in \Omega$ is obtained by evaluating $\psi \colon \Omega \times \mathcal{S}^2 \to \mathbb{R}$ at the sampling points contained in $\mathcal{U}$.

Spatial Interpolation

In practice, diffusion data and thus reconstructed ODFs are obtained on a regular grid $\Omega_h \subset \Omega$ as defined in (2.6). Here, we work in dimensionless coordinates which are chosen such that $x_1 = y_1 = z_1 = 1$, $x_{N_x} = N_x$, $y_{N_y} = N_y$, $z_{N_z} = N_z$, and $h_x = h_y = h_z = 1$. For the reconstruction of fiber tracks from ODF data (see Section 2.7) it is useful to have a continuous ODF field on Ω. To achieve this, we estimate ODF-values between grid points by trilinear interpolation.

2.6. Image Contrasts Derived from MRI and DW-MRI

In this section, we define apparent diffusion coefficient and fractional anisotropy which are derived from DW-MRI data and describe diffusion in each voxel by a scalar. Furthermore, we briefly discuss maps obtained from anatomical images.

2.6.1. Scalar Diffusion Maps

Scalar diffusion maps are usually visualized as gray scale images. In each voxel, the diffusion properties derived from DT or ODF are reduced to a single scalar value. In the following, we describe the calculation of the most widely used scalar diffusion values, the *apparent diffusion coefficient* (ADC) and the *fractional anisotropy* (FA).

Apparent Diffusion Coefficient (ADC)

The ADC, or *mean diffusivity*, represents the amount of diffusion present at a voxel $x \in \Omega$. For a DT $D(x)$ as introduced in Section 2.4, it is given by

$$\mathrm{ADC}_{D(x)} := \langle \lambda \rangle := \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3), \tag{2.10}$$

an average of the DT's eigenvalues. Similarly, we define

$$\mathrm{ADC}_{\psi(x,\cdot)} := \langle \psi \rangle := \frac{1}{N_u} \sum_{i=1}^{N_u} \psi(x, u_i) \tag{2.11}$$

for ODF ψ sampled at points $u_i \in \mathcal{U}$, $i = 1, \ldots, N_u$, as specified in Section 2.5.2.

Fractional Anisotropy (FA/GFA)

The FA value in each voxel $x \in \Omega$ is a measure for the degree of anisotropy. It is defined as the standard deviation of the eigenvalues of the DT divided by the root mean squared, that is

$$\mathrm{FA}_{D(x)} := \sqrt{\frac{3\left((\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2\right)}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{2.12}$$

where $\langle\lambda\rangle$ is the mean defined in (2.10). The FA value derived from an ODF, also known as *general* FA (GFA), was introduced in the article by (Tuch, 2004). Similar to (2.12), it is given by

$$\mathrm{FA}_{\psi(x,\cdot)} := \sqrt{\frac{N_u \sum_{i=1}^{N_u} \left(\psi(x, u_i) - \langle\psi\rangle\right)^2}{(N_u - 1)\sum_{i=1}^{N_u} \psi(x, u_i)^2}}$$

with $\langle\psi\rangle$ as defined in (2.11). FA values are in the range from zero to one. A value close to zero indicates isotropic diffusion, whereas a value close to one indicates anisotropic diffusion along a single direction.

2.6.2. Anatomical Maps

Knowledge about the location of different tissue types in a data set can provide very useful additional information for tractography. In the following, we consider binary and tissue partial volume maps.

Binary Tissue Masks

Let A denote an anatomical MR image of size $N_x \times N_y \times N_z$. A binary tissue mask $M_T \in \{0,1\}^{N_x \times N_y \times N_z}$ for each voxel (i, j, k), $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$, $k = 1, \ldots, N_z$, is defined by

$$M_T(i,j,k) = \begin{cases} 1 \text{ if } A(i,j,k) \text{ predominantly contains tissue type } T \\ 0 \text{ else} \end{cases} \tag{2.13}$$

Figure 2.8.: Tissue partial volume maps of a coronal slice generated with FAST from FSL. *Left:* White matter, *middle:* gray matter, *right:* CSF.

where $T \in \{\mathrm{WM}, \mathrm{GM}, \mathrm{CSF}\}$ (see Section 2.3 for the a definition of the different tissue types). White matter masks, for instance, are often used as *tracking masks*. Note that the classification of different tissue types requires a preprocessing step which is addressed at the end of this section.

Tissue Partial Volume Maps

We define WM, GM, and CSF tissue partial volume maps as images I_{WM}, I_{GM}, $I_{\mathrm{CSF}} \in [0,1]^{N_x \times N_y \times N_z}$ such that $I_{\mathrm{WM}} + I_{\mathrm{GM}} + I_{\mathrm{CSF}} = 1^{N_x \times N_y \times N_z}$. Each voxel value represents the proportion of the respective tissue type in that voxel, or could be interpreted as the probability that a random (uniformly distributed) point in that voxel belongs to one of the three categories. An example for tissue partial volume maps is shown in Figure 2.8.

In practice, automated tissue segmentation, both binary and partial volume, is difficult. The MRI analysis software FSL provides the segmentation tool FAST based upon an algorithm described in (Zhang et al., 2001). FAST generates binary segmented masks, as well as tissue partial volume maps, using for instance a T_1-weighted image (see Equation (2.1)) as input. Nonetheless, it is always advisable to check such automatically segmented images, and manually modify them if necessary.

2.7. Fundamentals of Fiber Tractography

Let $\mathcal{D}$ represent the measured diffusion MRI data and $\mathcal{F}$ a set of three-dimensional white matter trajectories corresponding to $\mathcal{D}$. The relation between $\mathcal{D}$

and $\mathcal{F}$ can be described by

$$\mathcal{D} = \mathcal{A}(\mathcal{F}) + \varepsilon,$$

where $\mathcal{A}$ denotes an operator assigning the fiber tracts to their individual signal contributions, and ε denotes noise in the measured data. Given data $\mathcal{D}$ and a strategy of retrieving information about tract location from the data, that is, in a sense, inverting $\mathcal{A}$, *fiber tractography* denotes the inverse problem of reconstructing the set of fiber bundles $\mathcal{F}$.

In the following, the terms *fiber tractography* and *fiber tracking* are used interchangeably. Furthermore, let the output of a fiber tractography algorithm be given by *fiber tracks*. We define a fiber track T by a discrete representation in the form of an array

$$T = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^{n\times 3} \tag{2.14}$$

of n points $x_i \in \Omega$, $i = 1, 2, \ldots, n$. The points will usually be chosen equidistantly, i.e. $\|x_j - x_{j-1}\|_2 = \|x_k - x_{k-1}\|_2$, $1 < j, k \leq n$. A set of fiber tracks of the same data set is called a *tractogram*.

Basically, there are two different categories of tractography algorithms, *streamline* and *global* methods. Streamline tractography reconstructs fiber tracks based on the information obtained from the underlying DT or ODF field by line propagation starting from given *seed points*. On the other hand, global tractography yields a tractogram from optimizing a global objective function. We review streamline tractography methods in Section 2.7.1, and briefly summarize global strategies in Section 2.7.2. Regardless of the method used, one has to deal with the question if a reconstructed track really reflects the underlying brain anatomy. In Section 2.7.3, we regard approaches that seek to evaluate a given tractogram.

2.7.1. Streamline Tractography

In this section, we introduce streamline tractography and state a basic deterministic and a basic probabilistic streamline tractography algorithm. Moreover, we discuss drawbacks and advantages of different streamline approaches that can be found in literature and summarize common strategies for seed point selection, choice of the step length, and stopping criteria.

Deterministic

In the early days of fiber tractography, track propagation was usually based on diffusion tensors. Given a continuous representation of a tensor field, the principle diffusion direction at each point $x \in \Omega$ can be obtained as the eigenvector corresponding to the largest eigenvalue of the diffusion tensor at that point. Assuming its uniqueness at each $x \in \Omega$, we define the principle diffusion direction by the vector field $\varepsilon_1 : \Omega \to \mathbb{R}^3$. In Basser et al. (2000), the evolution of a fiber tract trajectory $\mathbb{R} \to \mathbb{R}^3,\ t \mapsto \chi(t)$ from a seed point $x^{(0)} \in \Omega$ is formulated as an initial value problem:

$$\frac{d\chi(t)}{dt} = \varepsilon_1(\chi(t)), \quad \chi(0) = x^{(0)}. \tag{2.15}$$

Approximating the solution to (2.15) using Euler's method results in the well-known streamline tractography rule for calculating in each iteration step $k = 0, 1, 2, \ldots$ a new fiber path point

$$x^{(k+1)} = x^{(k)} + \lambda d^{(k)} \tag{2.16}$$

with starting point $x^{(0)}$, direction $d^{(k)} = \varepsilon_1\left(x^{(k)}\right)$ and step length $\lambda > 0$. Similar approaches can be found in Conturo et al. (1999) and Mori et al. (1999).

Replacing Euler's method by higher order methods, such as Runge-Kutta schemes, a more accurate approximation to the solution of ODF system (2.15) can be obtained. For example, applying standard fourth order Runge-Kutta integration yields the tractography iteration rule

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 \\ &\text{where} \\ k_1 &= \lambda\varepsilon_1\left(x^{(k)}\right) = \lambda d^{(k)} \\ k_2 &= \lambda\varepsilon_1\left(x^{(k)} + \frac{1}{2}k_1\right) \\ k_3 &= \lambda\varepsilon_1\left(x^{(k)} + \frac{1}{2}k_2\right) \\ k_4 &= \lambda\varepsilon_1\left(x^{(k)} + k_3\right), \end{aligned} \tag{2.17}$$

as described, for instance, in McGraw et al. (2004). So far, we have assumed uniqueness of the principle diffusion direction. In fact, as described in Section 2.4, the sign of the principle diffusion direction is indefinite and other ambiguities occur if the largest eigenvalue is not unique ($\lambda_1 = \lambda_2$, $\lambda_1 = \lambda_2 = \lambda_3$). These

issues are resolved by, for instance, selecting the direction in coherence with the previous tracking direction.

The strategies introduced above essentially reconstruct fibers by successively stepping in the direction of principle diffusion of the current spatial position. More refined methods (e.g. Weinstein et al. (1999); Westin et al. (2002)) try to cope with tensor orientation ambiguities in voxels with multiple fiber orientations by using the full tensor information to increase robustness of tract propagation.

Eventually, the development of HARDI-based methods with the ability to resolve crossing fibers enables fiber tracking in regions of complex tissue architecture, outperforming DT-based tractography. Several methods (e.g. Descoteaux et al. (2009); Wedeen et al. (2008)) have extended the basic concept of tensor methods to ODF data substituting the principle eigenvector of the DT for the direction maximizing the local ODF $\psi\left(x^{(k)}, d\right)$ while producing the smallest angle with the incoming direction $d^{(k-1)}$. In the following, we denote this direction by $d_\psi^{(k)}$. As described in Section 2.5.2, trilinear interpolation is used to obtain values of ψ between grid points. Selecting $d^{(0)} \in \operatorname{argmax}_{d\in\mathcal{U}} \psi\left(x^{(0)}, d\right)$ and $d^{(k)} = d_\psi^{(k)}$ for $k > 0$ in the discrete fiber track evolution equation (2.16) results in a straightforward deterministic tracking method:

Algorithm 2.1. (Basic deterministic tractography)
Input:
$\psi : \Omega \times \mathcal{U} \to \mathbb{R}$ (field of ODFs)
dir $\in \{+, -\}$ (sign of initial tracking direction)
$x^{(0)} \in \Omega$ (initial point)
$d_\psi^{(0)} \in \mathcal{U}$ (initial direction)
$\varphi \in (0, \pi]$ (max. deviation angle from previous direction)

if dir=='−'
$\quad d_\psi^{(0)} = -d_\psi^{(0)}$
end

for $k = 1, 2, \ldots$
$\quad x^{(k)} = x^{(k-1)} + \lambda d_\psi^{(k-1)}$
$\quad$ **if** $x^{(k)} \in \Omega$
$\quad\quad T^{\mathrm{dir}}(k) = x^{(k)}$

$$\text{compute } D = \left\{ d^* \in \mathcal{U} \middle| d^* = \operatorname*{argmax}_{\substack{d\in\mathcal{U} \\ \sphericalangle\left(d, d^{(k-1)}\right) < \varphi}} \psi\left(x^{(k)}, d\right) \right\}$$

$$\text{select } d_\psi^{(k)} \in \operatorname*{argmin}_{d\in D} \sphericalangle\left(d, d_\psi^{(k-1)}\right)$$

```
  else
      break
  end
end
```

Output: T^{dir} (fiber track)

Note that since the ODF is symmetric, the tracking algorithm can proceed from a seed point $x^{(0)}$ both in positive and negative direction. Hence, the complete track T obtained from streamline tractography can be represented by

$$T = \begin{pmatrix} PT^{-} \\ T^{+} \end{pmatrix} \quad \text{with} \quad P = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} \tag{2.18}$$

with tracks T^{+} and T^{-} resulting from invoking the algorithm with dir $= +$ and dir $= -$, respectively, and permutation matrix P.

A method generalizing the DT-based tracking rule derived from Runge-Kutta fourth order integration stated in (2.17) to ODF input data is the state-of-the-art *SD_Stream* algorithm included in the widely-used MRI software tool *MRtrix*. Yet another interesting strategy described in Chao et al. (2008) uses branching of fiber tracks in regions with multiple prominent diffusion directions.

On the whole, deterministic streamline tractography is computationally very efficient. However, a drawback inherent in line propagation methods operating on data with low SNR, like DW-MRI data, is the stepwise accumulation of measurement errors. Moreover, streamline fiber tracking suffers from partial volume effects due to the rather coarse spatial resolution (1 - 4 mm^3) of DW-MRI acquisitions.

Probabilistic

Probabilistic tractography seeks to tackle the problems of deterministic methods by taking into account the uncertainty inherent in reconstructed fiber orientations. For instance, statistical bootstrapping strategies generate a multitude of individual DTI or HARDI data sets by using one of various schemes to resample the corresponding diffusion-weighted MRI acquisitions. Multiple probabilistic fiber tracks are then produced for any seed point by calculating tensor or ODF volumes and applying a deterministic streamline tractography algorithm

to each data set (Lazar and Alexander, 2005; Berman et al., 2008; Jones, 2008; Jeurissen et al., 2011). Another category of probabilistic streamline tractography algorithms is based on Monte Carlo simulations or random walks. Unlike deterministic streamline methods that follow in each iteration a uniquely identified direction of principle diffusion, these techniques determine the stepping direction by sampling at the current spatial position a probability distribution representing local uncertainty of the fiber orientation. The probability distribution is derived from the shape of the DT (Björnemo et al., 2002; Lazar and Alexander, 2002; Koch et al., 2002; Hagmann et al., 2003; Parker et al., 2003) or ODF (Parker and Alexander, 2005), or the direction is directly sampled from the local ODF (Perrin et al., 2005; Descoteaux et al., 2009).

Referring to the latter case and adopting the notion from Algorithm 2.1, we can regard the ODF ψ as a probability distribution

$$P(x^{(k)}, d) = \frac{\psi(x^{(k)}, d)}{\sum_{d \in \mathcal{U}} \psi(x^{(k)}, d)}$$

and deduce the straightforward probabilistic tracking rule:

Algorithm 2.2. (Basic Probabilistic Tractography)
Input:
$\psi : \Omega \times \mathcal{U} \to \mathbb{R}$ (field of ODFs)
dir $\in \{+, -\}$ (sign of initial tracking direction)
$x^{(0)} \in \Omega$ (initial point)
$d_\psi^{(0)} \in \mathcal{U}$ (initial direction)

if dir=='−'
  $d_\psi^{(0)} = -d_\psi^{(0)}$
end

for $k = 1, 2, \ldots$
  $x^{(k)} = x^{(k-1)} + \lambda d_\psi^{(k-1)}$
  if $x^{(k)} \in \Omega$
    $T^{\text{dir}}(k) = x^{(k)}$
    sample $d_\psi^{(k)} \in \mathcal{U}$ from $P(x^{(k)}, d)$
    if $\sphericalangle\left(d_\psi^{(k)}, d_\psi^{(k-1)}\right) > \frac{\pi}{2}$
      $d_\psi^{(k)} = -d_\psi^{(k)}$
    end **else**
    break
  end
end

Output: T^{dir} (fiber track)

Thus, the probability to step in the direction of principle diffusion is high, but other directions are possible as well. Repeating this procedure, a multitude of possible fiber tracks emanating from each seed point is obtained. Note that the complete track is obtained as in (2.18). A state-of-the-art *second* order probabilistic streamline tractography strategy, included in the MRtrix software package, is the algorithm *IFOD-2* described in Tournier et al. (2010).

Similarly, *Bayesian* methods (e.g. Behrens et al. (2003); Kaden et al. (2007); Friman et al. (2006); Ramirez-Manzanares and Rivera (2006)) generate fiber tracks by sampling from a local posterior probability density function on the fiber direction given the measured data and diffusion model. The posterior is calculated as the normalized product of a likelihood or observation density, derived from a parametric model of the diffusion signal, and a prior density modeling knowledge about the model parameters (Behrens et al., 2003; Kaden et al., 2007) or promoting the continuation of the track in the previous stepping direction (Friman et al., 2006; Ramirez-Manzanares and Rivera, 2006). Furthermore, Zhang et al. (2009) combine this approach with particle filtering (sequential Monte Carlo methods). Here, a finite number of particles is propagated from a seed point by drawing samples from an importance function and assigning weights to the evolving path according to a recursively defined posterior distribution. The path with the highest weight is selected as the optimal fiber path estimate. All these probabilistic techniques allow the assignment of a degree of confidence to a reconstructed fiber path. Furthermore, anatomical connectivity indices between different brain regions can be derived.

A different but somewhat related strategy for tractography is referred to as *front evolution*. As probabilistic tractography, *front evolution* methods can provide information about the likeliness of a path. Usually, these methods are based on level set theory. From a seed point, a surface propagates guided by either the field of principle diffusion directions (Parker et al., 2002; Jbabdi et al., 2008), i.e. the eigenvectors corresponding to the largest eigenvalue of the DT, or the full DT or ODF (Campbell et al., 2005). From the resulting map of arrival times of the front surface at each voxel, fiber tracts are reconstructed as the minimal distance path, or *geodesic*, from any spatial point of the volume to the seed point. Though being robust in the presence of noise and partial volume effects due to optimization of a global criterion, the minimal distance approach appears to be problematic in regions of high curvature fibers where a higher connectivity index may be assigned to false positive but smooth tracks than to sharp bending fibers. Apart from using level set theory, similar front evolution methods exist. For instance, Tournier et al. (2003) consider a front whose surface consists of a

finite number of points. In each iteration, the front is evolving by sampling from an ODF derived from the DT at each point of the surface of the front. Number and orientation of the sampled vectors are restricted by a curvature constraint and an ODF-value threshold.

Seed Point Selection

Seed points are selected in a *region of interest* (ROI) in white matter either at random (e.g. Hagmann et al. (2003)), on a regular grid (e.g. Conturo et al. (1999)), or manually (e.g. Parker et al. (2003)). According to the different aims of tractography experiments, the size of the ROI ranges from only few voxels for targeted tractography of individual fiber bundles (Descoteaux et al., 2009; Berman et al., 2008; Parker et al., 2003), to a segmentation of the whole brain, e.g. for the generation of connectivity maps (Hagmann et al., 2003). Some methods also perform a filtering step after tractography where tracks that do not start and end in predefined ROIs are discarded (Conturo et al., 1999; Hagmann et al., 2003).

Choice of step length λ

The step length λ is usually set to a fixed value between 0.1 and 0.5 $\times$ voxel-size (e.g. Basser et al. (2000); Parker et al. (2003); Descoteaux et al. (2009)). Besides, some methods (e.g. Mori et al. (1999); Chao et al. (2008)) employ an adaptive step length scheme. Here, the stepping direction is constant until a boundary of the current voxel is reached and a new direction corresponding to the subsequent voxel is computed.

Stopping Criteria

Apart from termination due to leaving the image boundaries or executing a preselected number of iteration steps, other stopping criteria have been proposed. Most commonly, tracking terminates if the local FA value, defined in Section 2.6.1, falls below a threshold (e.g. Basser et al. (2000); Berman et al. (2008)). Moreover, ODF-based tractography algorithm often use an ODF amplitude threshold (e.g. Tournier et al. (2010); Jeurissen et al. (2011)). In addition to FA and ODF amplitude, tracking masks may contain information such as brain boundaries or tissue properties (Koch et al., 2002; Jeurissen et al., 2011). For

instance, a WM tracking mask M_{WM}, as stated in (2.13) can be included as stopping criterion in the following way: Tractography continues while $M_{\mathrm{WM}}(x^{(k)}) = 1$ but track generation terminates if $M_{\mathrm{WM}}(x^{(k)}) = 0$. Furthermore, a curvature threshold is often selected as termination criterion. For instance, the algorithm described in Wedeen et al. (2008) breaks if the curvature is larger than 0.5 rad in one voxel, while Descoteaux et al. (2009) terminate tracking if the angle between two consecutive steps is larger than 75° where $\lambda = 0.5$.

2.7.2. Global Tractography

Yet another class of methods, usually referred to as *global tractography*, seek to simultaneously determine fiber tracks across the whole data set by optimizing a global objective function. In Kreher et al. (2008), Fillard et al. (2009), and Reisert et al. (2011), for example, each fiber track comprises small fiber segments defined by spatial position and orientation. An initial configuration of fiber segments distributed in white matter regions is iteratively adjusted using stochastic optimization techniques to best fit an overall energy function controlling the difference between measured and predicted signal, as well as anatomical knowledge like low curvature of fibers. Compared to local fiber tracking algorithms, these methods are less prone to errors and image artifacts, but rather expensive regarding computation time. Besides, stochastic optimization techniques do not ensure convergence to a global optimal solution.

A different approach based on the solution of a convex optimization problem is introduced in Daducci et al. (2013). The framework uses a tractogram as input, approximates the diffusion-weighted MR signal by a linear combination of signal contributions of individual fibers, and fits the modeled signal to the measured one by solving a least-squares problem with ℓ_1-regularization. By this means, fibers that do not well support the global diffusion signal can be eliminated from the tractogram. Due to the convexity of the problem, a unique solution is guaranteed. Moreover, depending on the method used to obtain the initial tractogram, the overall technique can be much faster regarding computation time than the stochastic procedure stated above.

2.7.3. Quantitative Validation

In recent years, a variety of both numerical and physical diffusion phantoms have been designed to quantitatively validate the performance of fiber reconstruction methods. Côté et al. (2013) introduced the *Tractometer* evaluation tool for

data sets with known ground truth bundles, comprising several characteristics to quantify a given tractogram. Let n_{VC}, n_{IC}, and n_{NC} denote the numbers of valid connections, invalid connections, and no connections in the regarded tractogram. The number of overall tracks is defined by $n = n_{\mathrm{VC}} + n_{\mathrm{IC}} + n_{\mathrm{NC}}$. The regarded characteristics are:

VC $= \frac{n_{\mathrm{VC}}}{n}$ (proportion of valid connections to total tracks)

IC $= \frac{n_{\mathrm{IC}}}{n}$ (proportion of invalid connections to total tracks)

NC $= \frac{n_{\mathrm{NC}}}{n}$ (proportion of no connections to total tracks)

VB (number of valid bundles)

IB (number of invalid bundles)

Furthermore, Girard et al. (2014) introduce

VCCR $= \frac{n_{\mathrm{VC}}}{n_{\mathrm{VC}}+n_{\mathrm{IC}}}$ (proportion of valid connections to total connections, or the *valid connection to connection ratio*)

CSR $= \frac{n_{\mathrm{VC}}+n_{\mathrm{IC}}}{\#\mathrm{seeds}}$ (proportion of connections to number of seed points, or the *connection to seed ratio*)

In the literature, the Tractometer results are usually stated in percent.

Let the end regions connected by ground truth bundle j be indicated by $E_{j,1}$ and $E_{j,2}$. We regard a fiber track T of length N as defined in (2.14). Furthermore, let successive points of each track in the ground truth bundles have the same Euclidean distance as for track T. We assign track T to one of the three categories, valid, invalid, and no connection, in the following way:

Algorithm 2.3. (Track categorization)
Input:
threshold $\theta \geq 0$,
track T,
GT bundles $j = 1, 2, \ldots$ with end-regions $E_{j,1}, E_{j,2}$

if $(x_1 \in E_{j,1} \wedge x_N \in E_{j,2}) \vee (x_1 \in E_{j,2} \wedge x_N \in E_{j,1})$
  if there exists a track T_{GT} in bundle j such that $\|T - T_{\mathrm{GT}}\|_2 \leq \theta$
    T is *valid connection*
  else
    T is *invalid connection*

end
elseif $(x_1 \in E_{j,\cdot} \wedge x_N \in E_{k,\cdot}), j \neq k$
T is *invalid connection*
else
T is *no connection*
end

In other words, if a track connects two end regions of the same bundle and the difference to one of the ground truth tracks (measured in Euclidean norm) is below a threshold, the track is categorized as VC. On the other hand, if the difference is above a threshold, or if the track connects two end regions of different bundles, the track is assigned to the IC tracks. A track is counted as NC, if it does not connect two regions.

2.8. Modeling WM Microstructure

Compartment models are used to probe white matter microstructure from diffusion MRI measurements. In the article by Panagiotaki et al. (2012), compartment models are reviewed and compared. Three types of compartments are considered. First, the intra-axonal compartment describing diffusion of water inside the axons. Second, models that seek to explain hindered diffusion in the space between axons of fiber bundles, the extra-axonal space. And last, other cellular structures of white matter apart from axons where diffusion is modeled as restricted but isotropic. The myelin coating the axons is assumed to be impermeable to water, thus there is no exchange between intra- and extra-axonal space and in the intra-axonal compartments diffusion is considered restricted. In the following, we denote the signals modeling intra-axonal, extra-axonal, and isotropic diffusion in a voxel by R^{IC}, R^{EC}, and $R^{\mathrm{ISO}} : \mathcal{S}^2 \to \mathbb{R}$, respectively. The overall estimated voxel-wise signal attenuation is defined by

$$\frac{S(q)}{S_0} = f_{\mathrm{IC}} R^{\mathrm{IC}}(q) + f_{\mathrm{EC}} R^{\mathrm{EC}}(q) + f_{\mathrm{ISO}} R^{\mathrm{ISO}}(q) \tag{2.19}$$

where $f_{\mathrm{IC}}, f_{\mathrm{EC}}, f_{\mathrm{ISO}} \in [0, 1]$ denote the proportions of water molecules in the respective compartment, and $f_{\mathrm{IC}} + f_{\mathrm{EC}} + f_{\mathrm{ISO}} = 1$.

The intra-axonal diffusion signal is, for instance, represented as diffusion in a cylinder with radius zero by the *stick* model (Behrens et al., 2003)

$$R^{\mathrm{IC}}(q) = \exp\left(-bd(v_1^\top q)^2\right) \tag{2.20}$$

with b-value as defined in Section 2.2, axonal diffusivity $d > 0$, and fiber orientation $v_1 \in \mathcal{S}^2$. More complex models also incorporate axon diameter and the distribution of multiple axons.

The *zeppelin* model (Alexander, 2008) is often used to describe diffusion in extra-axonal space. Here, diffusion is modeled using an anisotropic, cylindrical symmetric diffusion tensor D with eigenvalues $\lambda_1 > \lambda_2 = \lambda_3$ and corresponding eigenvectors $v_1, v_2, v_3 \in \mathbb{R}^3$. Defining the isotropic tensor

$$I_\lambda := \lambda[v_1, v_2, v_3] \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{bmatrix} v_1^\top \\ v_2^\top \\ v_3^\top \end{bmatrix} \tag{2.21}$$

with identical eigenvalues $\lambda \geq 0$, we can write the diffusion signal as

$$R^{\mathrm{EC}}(q) = \exp\left(-bq^\top \left(\alpha v_1 v_1^\top + I_\beta\right) q\right) \tag{2.22}$$

where $\lambda_1 = \alpha + \beta$ and $\lambda_2 = \lambda_3 = \beta$. Information about diffusivity parallel to the axons is obtained by $d_\parallel = \lambda_1$, and perpendicular diffusivity is given by $d_\perp = \lambda_2 = \lambda_3$. Simpler models use only the isotropic tensor, while for a more accurate approximation the full diffusion tensor is taken into account.

Restricted isotropic diffusion can, for example, be explained by the *ball* model (Behrens et al., 2003) as an isotropic tensor. The resulting diffusion signal is given by

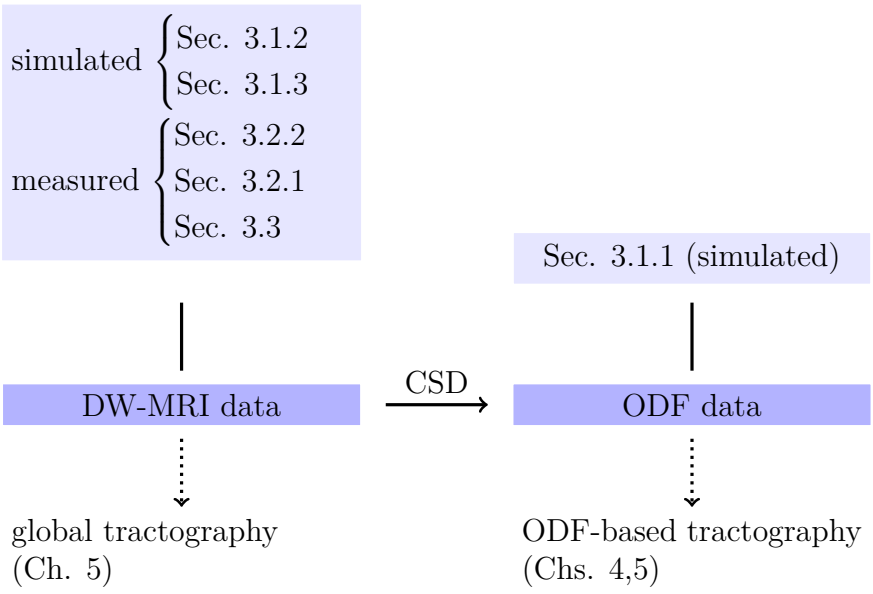
$$R^{\mathrm{ISO}}(q) = \exp\left(-bq^\top I_d q\right) \tag{2.23}$$

with diffusivity d and I_d as defined in Equation (2.21). More models for isotropic diffusion are summarized in Panagiotaki et al. (2012).

In Panagiotaki et al. (2012), estimating the signal attenuation stated in Equation (2.19) substituting R^{IC}, R^{EC}, and R^{ISO} as defined in Equations (2.20), (2.22), and (2.23), respectively, is termed the *StickZeppelinBall* model. Information such as axonal- and extra-axonal diffusivity, fiber orientations, and axon diameter are computed by parameter estimation, where the modeled signal (2.19) is fitted to the measured diffusion-weighted MR data S specified in (2.3).

3. Data

This chapter is intended to give an overview of the different numerically simulated, diffusion phantom and an *in vivo* data set that we used as input data for the methods described in Chapters 4 and 5. We owe the opportunity to test our algorithms performance on various different data sets to the Biomedizinische NMR Forschungs GmbH, who provided several diffusion data sets, as well as to other groups from the diffusion MRI community, who made their data freely available. In the following, we briefly present the generation strategies for the simulated data sets, development and structure of the physical phantoms, and acquisition parameters used to obtain the MRI measurements. Moreover, we state ODF reconstruction details. The diagram below gives an outline of the different data sets discussed in the following sections:



3.1. Simulated Data

The data considered in this section was not acquired by MRI measurements, but was computationally generated by different strategies based on models for the orientation distribution function or the diffusion MRI signal. In Section 3.1.1, we describe a method to generate ODF testing data. In Sections 3.1.2 and 3.1.3, we summarize the properties of two simulated diffusion MR data sets that were obtained by other groups and are publicly available.

3.1.1. Numerical ODF Data

To generate artificial ODF data, we consider a fiber track as a curve

$$\mathbb{R} \supset [a,b] \to \mathbb{R}^3, \quad t \mapsto \chi(t) = \begin{pmatrix} \chi_1(t) \\ \chi_2(t) \\ \chi_3(t) \end{pmatrix}$$

and model the orientation distribution function along x by a function

$$f(u,t) = \left(\frac{u^\top \frac{d}{dt}\chi(t)}{\left\| \frac{d}{dt}\chi(t) \right\|} \right)^p$$

with $u \in S^2$, $t \in [a,b]$ and $p \in \mathbb{N}$. Here, we choose $p = 4$. The function f is then multiplied with a Gaussian function

$$g(\omega,t) = \frac{1}{2\pi\delta} \exp\left(-\frac{(\omega_1 - \chi_1(t))^2 + (\omega_2 - \chi_2(t))^2 + (\omega_3 - \chi_3(t))^2}{2\delta} \right)$$

where $\delta^2 > 0$, $\omega \in \Omega$ and $t \in [a,b]$. We obtain the ODF field by numerical integration of the integral

$$\psi(\omega,u) = \int_a^b f(u,t) g(\omega,t) dt$$

on a uniform grid. In the following, we consider ODF data generated in that way which constitutes a circle intersected by a secant line. The ODFs, each sampled at 2562 points on the sphere, are equidistantly distributed on a Cartesian grid of size $25 \times 25 \times 3$. The three slices in the third dimension are identical and equal to the ODF field shown in Figure 3.1. In Section 4.3.1, tractography results for this data set are shown.

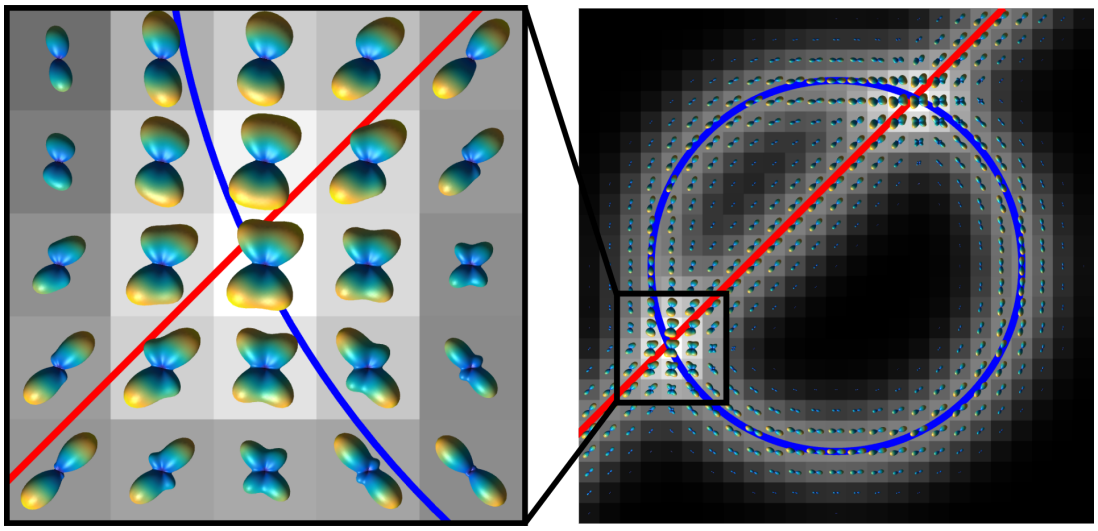
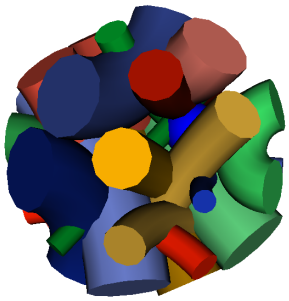


Figure 3.1.: *Right:* Computationally generated ODF field of size 25×25 where the two ground truth fibers are plotted in blue (circle) and red (line), respectively. *Left:* Enlarged section of a crossing region.

3.1.2. ISBI 2013 HARDI Reconstruction Challenge Data

The ISBI 2013 HARDI reconstruction challenge data[1] was provided for the IEEE International Symposium on Biomedical Imaging 2013 in order to compare different schemes for ODF reconstruction and tractography. It contains the simulated MR signal of complex three-dimensional fiber configurations distributed in a spherical domain. The MR signal was generated similar to the method described in the article by Close et al. (2009), but using a multi-compartment model. We chose the training data sets with 64 directions and $b = 3000$ s/mm^2, and both SNR $= 10$ and SNR $= 30$, respectively. The spatial dimensions are $50 \times 50 \times 50$. Furthermore, a brain mask, 40 seeding regions of different size, and a fiber geometry ground truth are provided for the data. Fiber ODF reconstruction was performed by standard CSD in MRtrix. Figure 3.2 illustrates the ground truth fiber configuration and the diffusion signal. Tractography results using this data set are discussed in Section 4.3.3.

[1] http://hardi.epfl.ch/static/events/2013_ISBI/



(a) Ground truth[1].

(b) Sagittal, coronal and transversal slices of the simulated MR signal for a single gradient direction.

Figure 3.2.: Simulated data from ISBI 2013 HARDI reconstruction challenge.

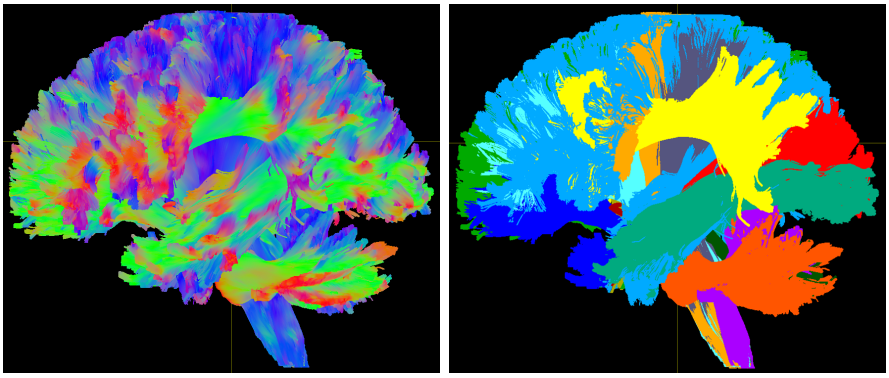
3.1.3. ISMRM 2015 Tractography Challenge Data

The ISMRM 2015 Tractography challenge data set[2] was generated for the comparison of different tractography pipelines. Data simulation consists of the following steps: First, a whole-brain tractogram is computed for an acquisition of the *human connectome project* (HCP) data set (see Sec. 3.3) using a global tractography method. Then, 25 bundles are manually segmented and processed, and the diffusion-weighted MR signal is simulated using the method described in Neher et al. (2014). The procedure results in a realistic simulation of a whole-brain DW-MR image with 32 diffusion-weighted images at $b = 1000$ s/mm^2, and one image without diffusion-weighting. The spatial dimensions are $90 \times 108 \times 90$, and isotropic voxel-size is $2 \times 2 \times 2$ mm^3. We chose the DW image without additional artifacts. Furthermore, a ground truth is provided for the 25 fiber bundles. The ground truth bundles and the simulated diffusion-weighted signal are illustrated in Figures 3.3 and 3.4, respectively. Fiber ODF reconstructions were computed using standard CSD in MRtrix. Results for the ISMRM 2015 data set are discussed in Section 4.3.4.

3.2. Diffusion Phantom Data

The physical phantoms considered in this section are built from synthetic materials to model configurations of nerve fiber bundles in the brain. The phantoms are constructed in such a way that diffusion MR scans yield diffusion anisotropy properties similar to the living brain. Since the fiber configuration is known, the

[1] Figure taken from http://hardi.epfl.ch/static/events/2013_ISBI/
[2] http://www.tractometer.org/ismrm_2015_challenge/



(a) Color corresponding to direction.

(b) Color corresponding to bundle classification.

Figure 3.3.: Ground truth fibers of ISMRM 2015 data (sagittal view).

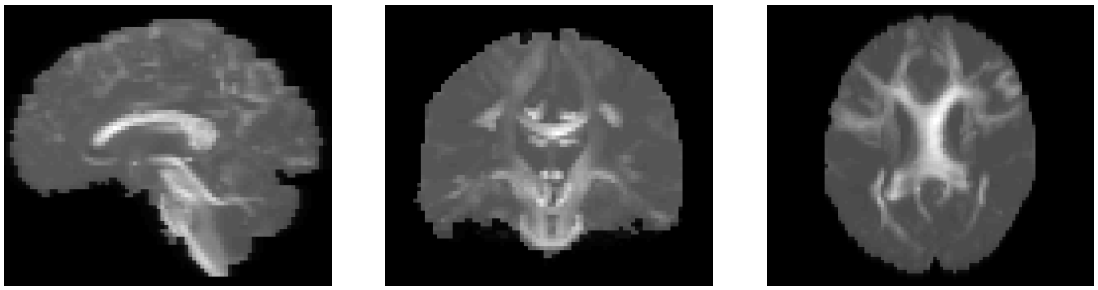


Figure 3.4.: Sagittal, coronal and transversal slices of the simulated data from the ISMRM 2015 tractography challenge for a single gradient direction.

diffusion MR acquisitions obtained from the phantoms can be used to validate diffusion imaging methods, such as ODF reconstruction or tractography.

3.2.1. *Fiber Cup* Phantom

The *Fiber Cup* diffusion phantom[1] was designed for a tractography challenge where ten algorithms were quantitatively evaluated and compared (Fillard et al., 2011). Modeling a coronal slice of a human brain, the phantom contains seven bundles of different curvature, crossing and kissing structures. The bundles, consisting of acrylic fibers of a diameter of 20 µm, are fixed and compressed between two polyurethane disks which have the fiber configuration engraved in. Each fiber is oriented in a plane perpendicular to the z-direction, that is, the phantom does not contain complicated three-dimensional fiber configurations. Fibers are arranged such that there are approximately 100 fibers per bundle and the fiber density is approximately the same in all areas of the fibrous structure, also at the crossings. For diffusion measurements, the phantom is placed inside a water-filled cylindrical container made of Plexiglas. A detailed description of design, construction and data acquisition of the diffusion phantom is given in Poupon et al. (2008) and Poupon et al. (2010).

Here, we consider a data acquisition of the phantom that was obtained on a $64 \times 64 \times 3$ grid with $3 \times 3 \times 3$ mm^3 isotropic voxel-size using 64 gradients at $b = 1500$ s/mm^2. Figure 3.5 shows the fiber pathways (Figure 3.5a) and image slices of the diffusion-weighted acquisition for $b = 0$ s/mm^2 (Figure 3.5b) and $b = 1500$ s/mm^2 for one gradient (Figure 3.5c). From the DWI acquisition we computed a set of fiber ODFs using the standard CSD algorithm from MRtrix as well as a set of spatially regularized fiber ODFs using the algorithm by Hohage and Rügge (2015). Both ODF data sets are sampled at 642 points on the sphere as described in Section 2.5.2. Standard and spatially regularized ODFs are visualized in the previous chapter in Figures 2.6b and 2.6c for the two rectangular regions indicated in Figure 3.5a. Tractography experiments and results are discussed in Sections 4.3.2 and 5.7.2.

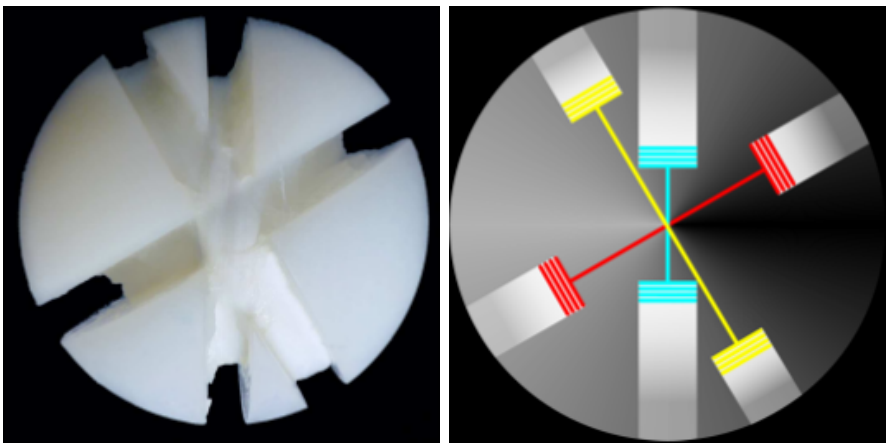
3.2.2. Spherical Diffusion Phantom

Manufacturing and diffusion-weighted MR measurements of a spherical diffusion phantom simulating crossing fiber bundles were performed by members of

[1] http://www.tractometer.org/original_fibercup/data/

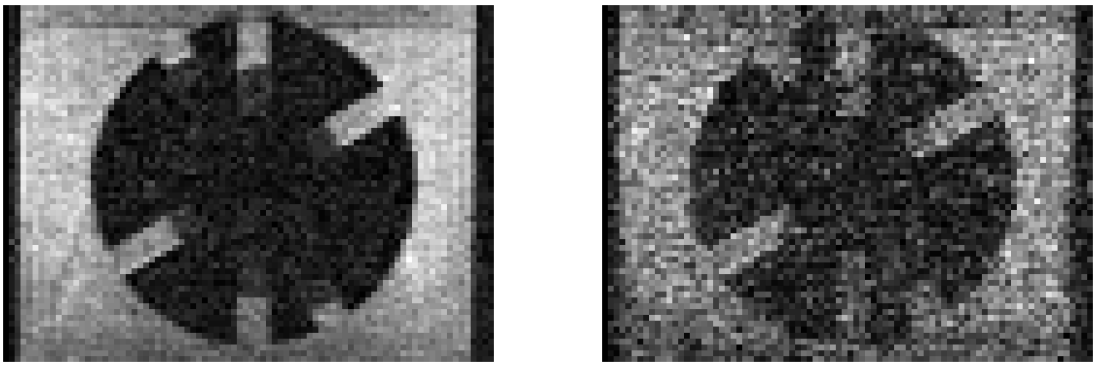
(a) fiber pathways **(b)** $b = 0$ s/mm^2 **(c)** $b = 1500$ s/mm^2 (for a single gradient direction)

Figure 3.5.: Fiber pathways (*left*) and Diffusion MRI acquisition (*middle and right*) for one slice in the $x - y$-plane.



(a) phantom **(b)** crossing structure

Figure 3.6.: Spherical diffusion phantom construction.



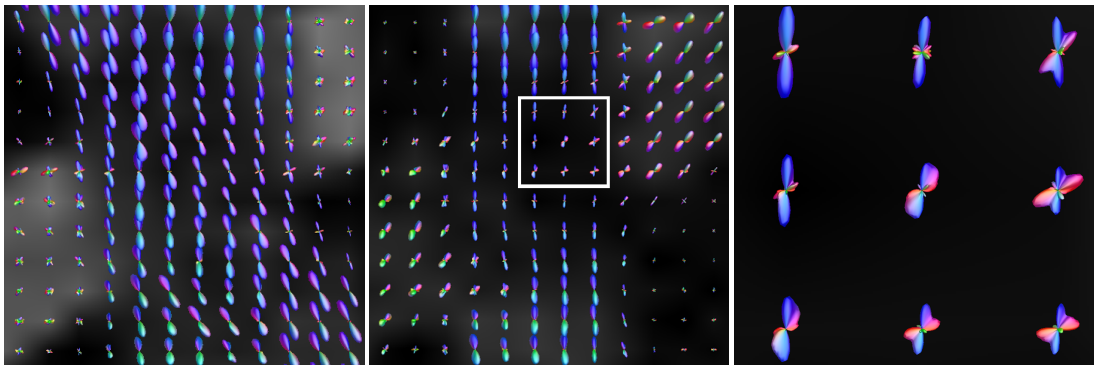
(a) $b = 0$ s/mm^2

(b) $b = 1000$ s/mm^2 (for a single gradient direction)

Figure 3.7.: Diffusion MRI acquisition in the mid-coronal image plane.

the Biomedizinische NMR Forschungs GmbH in accordance with the procedure described in Moussavi-Biugui et al. (2011). The phantom, illustrated in Figure 3.6a, consists of a spherical polyamide spindel where thin polyfil fibers are winded around forming three bundles. In each of two opposite regions on the spindle surface, a 60° crossing is formed by interleaving polyfil fibers of the bundles marked in blue and red in Figure 3.6b. Similarly, the bundles colored in blue and yellow in Figure 3.6b form a 30° crossing located directly below the 60° crossing in the direction towards the center of the sphere, respectively. In total, there are four crossings, two with a 30° and two with a 60° crossing angle. There is no area where all three bundles cross.

We use a diffusion-weighted data set of the phantom containing 92 diffusion-weighted images at $b = 1000$ s/mm^2 and one unweighted image. The spatial resolution is $84 \times 100 \times 61$, and voxel-size is $1.8 \times 1.8 \times 1.8$ mm^3. Two-dimensional image slices of the diffusion data are shown in Figure 3.7. We computed an ODF reconstruction with spatial regularization (see. Section 2.5.1). Though benefiting from the regularization when comparing it to standard CSD ODFs (not shown), the reconstruction suffers from artifacts due to the relatively high noise level of the data set. Figure 3.8 shows the ODF reconstruction in two coronal slices containing 30° and 60° crossing regions, respectively. As can be observed in Figure 3.8b, the 60° crossing could be resolved. That is, the ODFs in the crossing region have local maxima pointing in the directions of the fiber bundles. On the other hand, Figure 3.8a shows that the 30° crossing could not be resolved properly. That means that, instead of having maxima in the directions of the two crossing fiber bundles, most ODFs in the 30° crossing region have a single maximum pointing bidirectionally in a direction resembling the average between both fiber bundle directions. This may be caused by the fact that a b-value of 1000 is relatively low for ODF reconstruction. For instance, Tournier



(a) 30° crossing **(b)** 60° crossing **(c)** 60° crossing (detail)

Figure 3.8.: ODF reconstruction.

et al. (2004) examine the effects of different b-values on ODF reconstructions obtained with CSD. They conclude that a b-value chosen too low can result in an angular dependency that is too small to resolve the different fiber orientations. Tractography results for these data are discussed in Section 4.3.6.

3.3. Human Connectome Project

The *Human Connectome Project* (HCP) are projects conducted by neuroscientists from 9 different universities and other research institutions and funded by the *National Institutes of Health* (NIH) with the objective of investigating the human connectome (Van Essen et al., 2013). Data resulting from these research projects are made publicly available for download at `humanconnectome.org`.

We use a data set containing preprocessed diffusion and structural (T1w) MRI scans (Andersson et al., 2003; Sotiropoulos et al., 2013; Milchenko and Marcus, 2013; Glasser et al., 2013). The considered acquisition has spatial dimensions $145 \times 174 \times 145$, a voxel-size of $1.25 \times 1.25 \times 1.25$ mm^3 and consists of 18 scans at $b = 0$ s/mm^2 and diffusion-weighted scans at $b = 1000$ s/mm^2, $b = 2000$ s/mm^2 and $b = 3000$ s/mm^2 using 90 gradients, respectively. A brain mask is provided with the data set. We obtained a white matter mask using FAST in FSL (Zhang et al., 2001) from the T1 weighted image with default options. In a postprocessing step we manually dilated the mask because it appeared too restrictive in some regions. Moreover, FSL/FAST was used to generate tissue partial volume maps for anatomically informed tractography. In Figures 2.2 and 2.8, diffusion-weighted MR images and tissue partial volume maps are illustrated, respectively. The spatially regularized ODFs reconstructed from the data were sampled at 642

directions. Figure 4.11 shows a coronal view of the ODF data in a part of the *centrum semiovale*.

4. A Bayesian Approach for Neighborhood-Informed Tractography

In Section 2.7.1, we describe a class of widely-used local streamline tractography algorithms. Though being fast and easy to implement, these algorithms have the disadvantage of accumulating DT or ODF data errors due to the exclusively local treatment of the data in each iteration step. In this chapter, we focus on the improvement of ODF-based streamline techniques by including information obtained from the local neighborhood to refine the tracking direction. The *guiding by extrapolation* strategy, introduced in Section 4.1, integrates information about the curvature of the recently generated fiber fragment into the step-wise tracking process by extrapolating a polynomial of degree two, curve-fitted to a certain number of previously tracked path points. In Section 4.2, we describe the *forward search* strategy which explores the region ahead by characterizing candidate tracts of a certain length originating from the current path point. By means of Bayesian statistics, an a-posteriori probability, depending on prior and likelihood, is assigned to each candidate tract. We model the likelihood in such a way that it represents the plausibility of a candidate tract with respect to the underlying field of ODFs. Including the *guiding by extrapolation* strategy described above, the prior distribution assigns to the candidate tracts a probability considering its curvature. Maximum a-posteriori estimation is used to derive an optimal (deterministic) tracking direction in each iteration step. Moreover, we obtain a probabilistic tracking method by drawing from the marginalized posterior distribution. In Section 4.3, tests and comparisons on simulated data, diffusion phantom data, and *in vivo* data show the advantages of our approach.

This chapter has in part been published in the journal *Medical Image Analysis* (Schomburg and Hohage, 2017). Basic concepts and early results have appeared in abstract form (Schomburg et al., 2014, 2015).

4.1. Guiding by Extrapolation

Using the same notation as in Section 2.7.1, we seek to improve the tracking direction $d^{(k)}$ of the streamline iteration scheme (2.16). Let the ODF-based main diffusion direction $d_\psi^{(k)} \in \mathcal{U}$ in iteration k be as described in Algorithm 2.1. As a first attempt in obtaining a more stable tracking direction in the presence of noise, we search for the direction that optimizes the regularized problem

$$d^* = \underset{d\in\mathbb{R}^3}{\operatorname{argmin}} \left\| d - d_\psi^{(k+1)} \right\|_2^2 + \alpha \left\| d - d^{(k)} \right\|_2^2 \tag{4.1}$$

where the weighting factor $\alpha > 0$ balances closeness to the main diffusion direction $d_\psi^{(k+1)} \in \mathcal{U}$ at point $x^{(k+1)}$ and the previous direction $d^{(k)}$, thus regulating the curvature of the reconstructed fiber track. Solving (4.1) yields as tracking direction in iteration k the weighted sum

$$d^* = \frac{1}{1+\alpha}\left(d_\psi^{(k+1)} + \alpha d^{(k)}\right), \quad d^{(k+1)} = \frac{d^*}{\|d^*\|_2} \tag{4.2}$$

of main diffusion direction and former direction. This formulation is related to the ideas of Weinstein et al. (1999) and Westin et al. (2002) who added a stabilizing term to the principle eigenvector for DT tracking due to tensor ambiguities.

Instead of only using information about the previous tracking direction, it can be beneficial to look back a few more steps and consider a certain (small) number of former directions. Hence, we deduce an extrapolation strategy to include information about the curvature of the previously obtained fiber fragment. For this purpose, a smooth curve

$$\mathbb{R} \to \mathbb{R}^3, \quad t \mapsto \gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t)),$$

is fitted to the N previously tracked path points $x^{(k-i)} \in \Omega$, $i = 0, 1, \ldots, N-1$ using the curve-fitting strategy explained below, which is in this situation more robust than for instance a polynomial or spline interpolation. The functions γ_1, γ_2 and γ_3 are chosen as polynomials of degree 2 to obtain curvature and keep oscillating to a minimum. Thus, we want to find the coefficients $c^{(j)} \in \mathbb{R}^3$ of the three polynomials $\gamma_j\left(t, c^{(j)}\right) = c_0^{(j)} + c_1^{(j)} t + c_2^{(j)} t^2$, $j = 1, 2, 3$ such that

$$c^{(j)} = \underset{c\in\mathbb{R}^3}{\operatorname{argmin}} \sum_{i=0}^{N-1} w_i \left(\gamma_j\left(t_i, c\right) - x_j^{(k-i)}\right)^2$$

with monotonically decreasing weights $w_i > 0$, $w_i \geq w_{i+1}$ for $i = 0, 1, \ldots, N-1$

and parameter values $t_i = -\lambda i$, where $\lambda \geq 0$ is the step length as introduced in Equation (2.16). For our tests we have chosen linearly decreasing weights $w_i = (N - i)/N$. The three weighted least-squares problems are formulated as matrix equations $W^{\frac{1}{2}} V c^{(j)} = W^{\frac{1}{2}} x$ with

$$W = \begin{pmatrix} w_0 & & & \\ & w_1 & & \\ & & \ddots & \\ & & & w_{N-1} \end{pmatrix} \in \mathbb{R}^{N \times N}, \quad x = \begin{pmatrix} x_j^{(k)} \\ x_j^{(k-1)} \\ \vdots \\ x_j^{(k-(N-1))} \end{pmatrix} \in \mathbb{R}^{N \times 1},$$

and Vandermonde matrix

$$V = \begin{pmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{N-1} & t_{N-1}^2 \end{pmatrix} \in \mathbb{R}^{N \times 3}$$

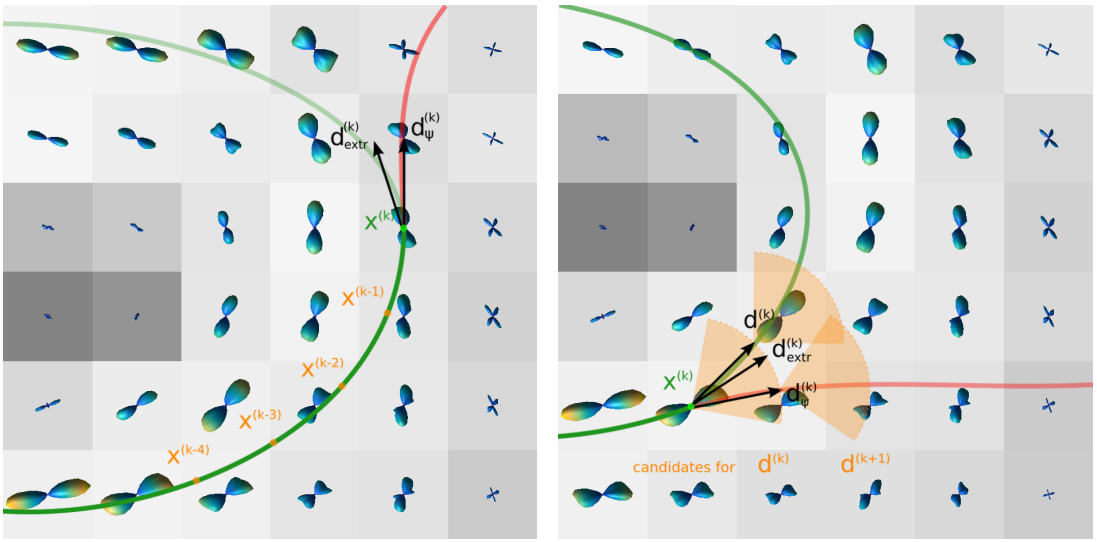
and are solved using QR-factorization, respectively. The obtained curve $\gamma(t)$ is extrapolated at $t_{-1} = \lambda$ to determine a *guiding direction*

$$d_{\text{extr}}^{(k)} := \frac{\gamma(t_{-1}) - x^{(k)}}{\|\gamma(t_{-1}) - x^{(k)}\|_2}$$

at point $x^{(k)}$. The *guiding by extrapolation* method is illustrated in Figure 4.1a using the five previously tracked path points $x^{(k)}, \ldots, x^{(k-4)}$. A curve, illustrated in dark green, is fitted to the five points. Via extrapolation of that curve, we obtain the guiding direction $d_{\text{extr}}^{(k)}$. The new direction $d^{(k)}$ is computed using $d_{\text{extr}}^{(k)}$ instead of $d^{(k-1)}$ in (4.2). The new path point $x^{(k)}$ computed using the guiding direction will be located close to the course of the true fiber depicted in green. On the other hand, following the local direction $d_{\psi}^{(k)}$ that maximizes the ODF at point $x^{(k)}$ causes a deviation from the true track indicated by the red curve.

4.2. Forward Search Method

For increased robustness against noise and artifacts we expand the algorithm described above by a method incorporating diffusion information of the region that lies ahead. Basically, a set of candidate fiber fragments of a certain length continuing from the current path point $x^{(k)}$ is determined based on anatomical plausibility, and a probability is assigned to each of these paths according to ODF data. The fiber tracking algorithm continues in the direction of the most



(a) Guiding by extrapolation **(b)** Forward search method

Figure 4.1.: Exemplary illustration of the methods described in Sections 4.1 and 4.2.

likely candidate path.

The idea of our approach is illustrated in Figure 4.1b for two search steps. From a path point $x^{(k)}$, the green line depicts the course of the true fiber, whereas the red line indicates a wrong track that a simple tracking algorithm without forward search would follow given the underlying field of ODFs. $d_{\psi}^{(k)}$ represents the locally optimal tracking direction at $x^{(k)}$ and $d_{\text{extr}}^{(k)}$ is the guiding direction described in Section 4.1. The orange area fanning out from $x^{(k)}$ indicates the region around $d_{\text{extr}}^{(k)}$ from which candidates for the direction $d^{(k)}$ are selected. The region angle is preselected to restrict curvature of the candidate tracts to an anatomically reasonable degree. The procedure is repeated for all $x^{(k+1)} = x^{(k)} + \mu d^{(k)}$ with valid candidate directions of $d^{(k)}$ and step length $\mu > 0$. For the sake of clarity, the figure shows areas of valid candidates for the direction $d^{(k+1)}$ originating only from two of the candidate points of $x^{(k+1)}$. Taking into account this whole set of fiber fragments of length two and the related ODF values, the most promising candidate direction of $d^{(k)}$ is selected for calculation of the new tracking direction. Note that for forward search we introduce a second step length μ that can be chosen independently from λ. Hence, the curve parameterization used for curve fitting in Section 4.1 is not entirely equidistant, but the distance $|t_i - t_{i-1}|$ of successive points is equal to λ for the track up to $x^{(k)}$ and equal to μ for the candidate tracks generated by forward search.

In the following, we formalize this idea using Bayesian statistics and deduce both a deterministic and a probabilistic tractography algorithm.

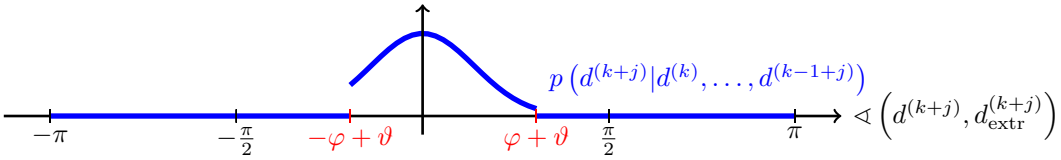


Figure 4.2.: Illustration of the transition probabilities. The angle between former and guiding direction is defined by $\vartheta := \sphericalangle\left(d^{(k+j-1)}, d^{(k+j)}_{\text{extr}}\right)$. In this example, we chose $\sigma = \frac{\pi}{4}$, $a \approx 0.8869$, and $\varphi = \frac{\pi}{4}$ and $\vartheta = \frac{\pi}{18}$.

4.2.1. Bayesian Framework

Let $\mathcal{U}$ denote a set of sampled directions on the unit sphere, as described in Section 2.5.2. In iteration $k-1$, a fiber track proceeding from a starting point $x^{(0)}$ in either positive or negative direction is uniquely determined by step length λ and the directions $d^{(0)}, d^{(1)}, \ldots, d^{(k-1)}$. We define the (unknown) fiber fragment of length n continuing this already tracked path by the array

$$\mathcal{F}^n := \left(d^{(k)}, d^{(k+1)}, \ldots, d^{(k+n-1)}\right) \in \mathcal{U}^n, \quad \mathcal{U}^n = \mathcal{U} \times \ldots \times \mathcal{U}$$

of successive directions. Moreover, $\mathcal{D}$ represents the *observations*, that is the ODF data providing information about directional diffusion of water at each voxel of a volumetric data set given on a domain $\Omega = [1, N_x] \times [1, N_y] \times [1, N_z]$. We are interested in the (a-posteriori) probability $P(\mathcal{F}^n|\mathcal{D})$ that the fiber fragment adopts a certain configuration given the diffusion data. By Bayes' theorem the a-posteriori probability is given by

$$P(\mathcal{F}^n|\mathcal{D}) = \frac{P(\mathcal{F}^n)P(\mathcal{D}|\mathcal{F}^n)}{P(\mathcal{D})}$$

with the prior-distribution $P(\mathcal{F}^n)$, the likelihood $P(\mathcal{D}|\mathcal{F}^n)$, and the normalizing constant $P(\mathcal{D})$. We model the prior-distribution to represent the probability of a certain fiber configuration $\mathcal{F}^n$ by considering its anatomical plausibility with respect to tract curvature. Therefore, we introduce the prior distribution as a product

$$\begin{aligned} P(\mathcal{F}^n) &:= p\left(d^{(k)}\right) \cdot p\left(d^{(k+1)}|d^{(k)}\right) \cdot \ldots \cdot p\left(d^{(k+n-1)}|d^{(k)},\ldots,d^{(k+n-2)}\right) \\ &= \prod_{j=0}^{n-1} p\left(d^{(k+j)}|d^{(k)},\ldots,d^{(k-1+j)}\right) \end{aligned} \tag{4.3}$$

of transition probabilities

$$p\left(d^{(k+j)}|d^{(k)},\ldots,d^{(k-1+j)}\right) \tag{4.4}$$

$$= \begin{cases} a \cdot \exp\left(-\frac{\sphericalangle\left(d^{(k+j)}, d_{\text{extr}}^{(k+j)}\right)^2}{\sigma^2} \right) & \text{if } \left| \sphericalangle\left(d^{(k+j)}, d^{(k+j-1)}\right) \right| \leq \varphi \\ 0 & \text{else} \end{cases}$$

given by the probability density of a truncated normal distribution with $\sigma^2 > 0$ and zero mean, depending on the angle between $d^{(k+j)}$ and the guiding direction $d_{\text{extr}}^{(k+j)}$ defined in Section 4.1 determined at step $k+j$. Hence, the less the direction $d^{(k+j)}$ deviates from the previously tracked path represented here by the guiding direction, the larger the transition probability. The angle $\varphi \in (0, \frac{\pi}{2}]$ indicates the maximally permitted deviation of $d^{(k+j)}$ from the direction $d^{(k+j-1)}$. As illustrated in Figure 4.2, the normal distribution is truncated according to the choice of φ. The normalizing constant a is selected such that $\sum_{d^{(k+j)} \in \mathcal{U}} p\left(d^{(k+j)} | d^{(k)}, \ldots, d^{(k-1+j)}\right) = 1$ for any $j = 1, \ldots, n-1$. This implies that $\sum_{\mathcal{F}^n \in \mathcal{U}^n} P(\mathcal{F}^n) = 1$.

We model the likelihood by

$$P(\mathcal{D}|\mathcal{F}^n) := \prod_{j=0}^{n-1} \psi\left(x^{(k+j)} + \frac{\mu}{2} d^{(k+j)}, d^{(k+j)}\right)$$

in terms of orientation distribution function $\psi : \Omega \times \mathcal{S}^2 \to \mathbb{R}$. The likelihood is normalized such that $\sum_{\mathcal{F}^n \in \mathcal{U}^n} P(\mathcal{D}|\mathcal{F}^n) = 1$. As illustrated in Figure 4.3 for $j = 0$, the ODF is evaluated at the midpoint of the points $x^{(k+j)}$ and $x^{(k+j+1)} = x^{(k+j)} + \mu d^{(k+j)}$ which, compared to simply taking $\psi\left(x^{(k+j)}, d^{(k+j)}\right)$, increases stability against local errors of the ODF due to noise. Our setting uses linear interpolation between the ODFs at the grid points. Hence, the ODF value obtained at $x^{(k+j)} + \frac{\mu}{2} d^{(k+j)}$ in direction $d^{(k+j)}$ is equal to $0.5 \cdot \left(\psi\left(x^{(k+j)}, d^{(k+j)}\right) + \psi\left(x^{(k+j+1)}, -d^{(k+j)}\right)\right)$. The likelihood $P(\mathcal{D}|\mathcal{F}^n)$ describes the probability distribution of data $\mathcal{D}$ conditioned that the true fiber continues as $\left(d^{(k)}, d^{(k+1)}, \ldots, d^{(k+n-1)}\right)$. It can be understood as the probability that a given fiber configuration is plausible regarding the fiber orientations given by the field of ODFs. Note that the purpose of local tractography as opposed to a (truly) global tractography is to recover only partial information about the fiber structure of the brain, which are not sufficient to fully explain the measured ODFs up to measurement errors. Therefore, we cannot simply compute the likelihood function from a fully specified set of probability distributions, but the definition of the likelihood has to involve some modeling.

The a-posteriori probability is computed as the product of prior and likelihood

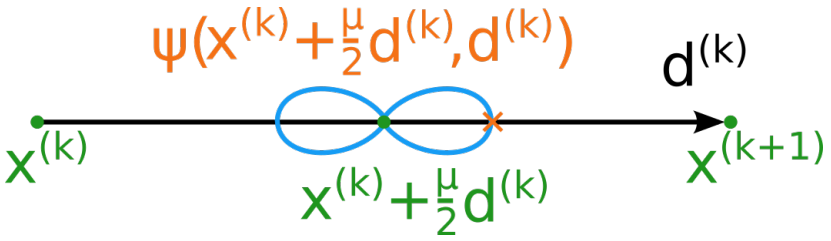


Figure 4.3.: ODF evaluation

and normalized such that $\sum_{\mathcal{F}^n \in \mathcal{U}^n} P(\mathcal{F}^n|\mathcal{D}) = 1$. We determine the maximum a-posteriori estimate

$$\hat{\mathcal{F}}_n^{\text{MAP}} = \left(\hat{d}_{\text{MAP}}^{(k)}, \hat{d}_{\text{MAP}}^{(k+1)}, \ldots, \hat{d}_{\text{MAP}}^{(k+n-1)}\right) \in \underset{\mathcal{F}^n \in \mathcal{U}^n}{\text{argmax}}\, P(\mathcal{F}^n|\mathcal{D}) \tag{4.5}$$

and select $\hat{d}_{\text{MAP}}^{(k)}$ as the new tracking direction $d^{(k)}$ of our deterministic tracking algorithm.

In the following, we show that this expression can be rewritten as an optimization problem similar to (4.1) where the objective function consists of a regularizing term that promotes closeness to the guiding direction and a data term to reduce deviations from the main diffusion direction. First, we reformulate the maximum a-posteriori (4.5) estimate as

$$\begin{aligned}
\hat{\mathcal{F}}_n^{\text{MAP}} &\in \underset{\mathcal{F}^n \in \mathcal{U}^n}{\text{argmax}}\, P(\mathcal{F}^n|\mathcal{D}) \\
&= \underset{\mathcal{F}^n \in \mathcal{U}^n}{\text{argmin}} [-\log(P(\mathcal{F}^n|\mathcal{D}))] \\
&= \underset{\mathcal{F}^n \in \mathcal{U}^n}{\text{argmin}} [\underbrace{-\log\left(P(\mathcal{F}^n)P(\mathcal{D}|\mathcal{F}^n)\right)}_{=:\mathcal{G}(\mathcal{F}^n)}].
\end{aligned} \tag{4.6}$$

Inserting the expressions for prior and likelihood we obtain for the objective function

$$\begin{aligned}
&\mathcal{G}\left(\mathcal{F}^n = \left(d^{(k)}, d^{(k+1)}, \ldots, d^{(k+n-1)}\right)\right) \\
&= -\log\left(\prod_{j=0}^{n-1} p\left(d^{(k+j)}|d^{(k+1)}, \ldots, d^{(k+j-1)}\right) \cdot \psi\left(x^{(k+j)} + \frac{\mu}{2} d^{(k+j)}, d^{(k+j)}\right)\right) \\
&= \begin{cases} \sum_{j=0}^{n-1} \frac{\sphericalangle\left(d^{(k+j)}, d_{\text{extr}}^{(k+j)}\right)^2}{\sigma^2} - \log(a) - \log\left(\psi\left(x^{(k+j)} + \frac{\mu}{2} d^{(k+j)}, d^{(k+j)}\right)\right) \\ \qquad \text{if } \left|\sphericalangle\left(d^{(k+j)}, d^{(k+j-1)}\right)\right| \leq \varphi \\ \infty \qquad \text{else,} \end{cases}
\end{aligned}$$

for $j = 0, \ldots, n-1$. Finally, substituting in equation (4.6) yields

$$\hat{\mathcal{F}}_n^{\mathrm{MAP}} \in \underset{\substack{\left(d^{(k)},\ldots,d^{(k+n-1)}\right)\in\mathcal{U}^n \\ \sphericalangle\left(d^{(k+j)},d^{(k+j-1)}\right)\leq\varphi}}{\operatorname{argmin}} \sum_{j=0}^{n-1} \frac{\sphericalangle\left(d^{(k+j)}, d_{\mathrm{extr}}^{(k+j)}\right)^2}{\sigma^2} - \log\left(\psi\left(x^{(k+j)} + \frac{\mu}{2} d^{(k+j)}, d^{(k+j)}\right)\right)$$

for the maximum a-posteriori estimate an expression comprising a sum of data term and regularization term similar to (4.1).

Alternatively, we can regard whole regions combining possible tract configuration for a certain starting orientation, instead of evaluating the possibility of single tract configurations, to increase stability against data errors. We define the marginal distribution at path point $x^{(k)}$ as

$$\begin{aligned}\tilde{\psi}\left(x^{(k)}, d^{(k)}\right) &= P\left(\bigcup_{d^{(k+1)}\in\mathcal{U}} \cdots \bigcup_{d^{(k+n-1)}\in\mathcal{U}} \mathcal{F}^n = \left(d^{(k)},\ldots,d^{(k+n-1)}\right) \middle| \mathcal{D}\right) \\ &= \sum_{d^{(k+1)}\in\mathcal{U}} \cdots \sum_{d^{(k+n-1)}\in\mathcal{U}} P\left(\mathcal{F}^n = \left(d^{(k)},\ldots,d^{(k+n-1)}\right) \middle| \mathcal{D}\right) \quad (4.7)\end{aligned}$$

with $\mathcal{F}^n$ disjoint. It describes the probability that the true fiber matches one of the fibers passing through $d^{(k)}$ by taking into account the averaged probabilities of subsequent configurations $\left(d^{(k+1)},\ldots,d^{(k+n-1)}\right)$. We derive a probabilistic interpretation of the tractography algorithm proposed above by drawing the new tracking direction $d^{(k)}$ from the marginalized probability distribution stated in (4.7).

Example for $n = 2$

For a better understanding, we illustrate two steps of the method in Figure 4.4 by means of a tree diagram where the root is given by the path point $x^{(k)}$. The nodes on the first level represent the candidate points of $x^{(k+1)}$. The edges linking these points with the root are each associated with a candidate direction $d^{(k)}$ and carry as weights the product of transition probability defined in (4.4) and the value of the local orientation distribution function describing the likeliness of the connection. Repeating this procedure, new candidates for the directions $d^{(k+1)}$ and points $x^{(k+2)} = x^{(k+1)} + \mu d^{(k+1)}$ are calculated from each candidate point of $x^{(k+1)}$ determined in the first step. The weights indicate the likeliness of a link between nodes $x^{(k+1)}$ and $x^{(k+2)}$. Overall, we can compute the probabilities of

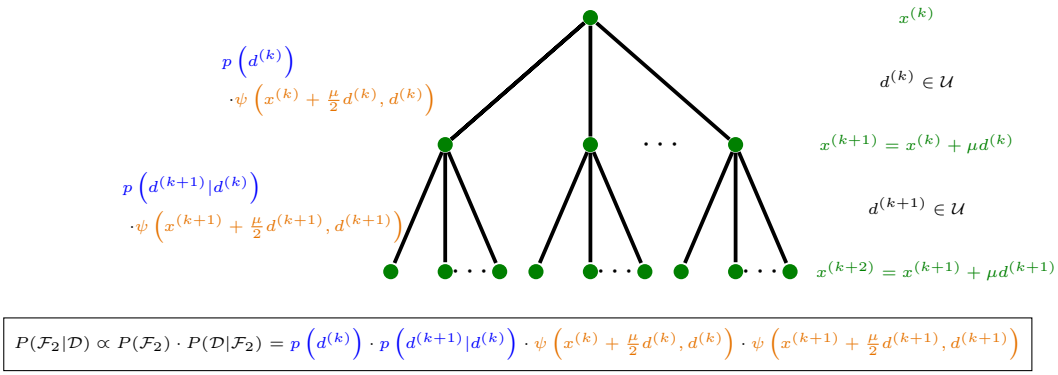


Figure 4.4.: Tree diagram of the forward search method for two steps.

connections between root and candidate nodes $x^{(k+2)}$ by either the product of edge weights of the path $x^{(k)} \to x^{(k+1)} \to x^{(k+2)}$ given by

$$\begin{aligned} &P(\mathcal{F}_2)\cdot P(\mathcal{D}|\mathcal{F}_2) \\ &= p\left(d^{(k)}\right)\cdot p\left(d^{(k+1)}|d^{(k)}\right)\cdot\psi\left(x^{(k)}+\frac{\mu}{2}d^{(k)},d^{(k)}\right)\cdot\psi\left(x^{(k+1)}+\frac{\mu}{2}d^{(k+1)},d^{(k+1)}\right), \end{aligned}$$

or by the marginal distribution

$$\begin{aligned} &\sum_{d^{(k+1)}\in\mathcal{U}} P\left(\mathcal{F}_2=\left(d^{(k)},d^{(k+1)}\right)\right)\cdot P\left(\mathcal{D}|\mathcal{F}_2=\left(d^{(k)},d^{(k+1)}\right)\right) \\ &= p\left(d^{(k)}\right)\cdot\psi\left(x^{(k)}+\frac{\mu}{2}d^{(k)},d^{(k)}\right) \\ &\quad\cdot\sum_{d^{(k+1)}\in\mathcal{U}} p\left(d^{(k+1)}|d^{(k)}\right)\cdot\psi\left(x^{(k+1)}+\frac{\mu}{2}d^{(k+1)},d^{(k+1)}\right), \end{aligned}$$

combining weights of subtrees from nodes $x^{(k+1)}$, respectively.

4.2.2. Computation of the Overall Tracking Direction

We sample the ODF $\psi : \Omega \times \mathcal{S}^2 \to \mathbb{R}$ at the vertices $u \in \mathcal{U}$ of a triangulation of the sphere, as described in Section 2.5.2. Thus, the maximum a-posteriori estimate (4.5) provides the optimal tracking direction, $\hat{d}^{(k)}_{\mathrm{MAP}}$, from only a finite set of sampling points. To obtain an optimal tracking direction on $\mathcal{S}^2$ instead of $\mathcal{U}$, we linearly interpolate on the tessellated surface of the sphere using barycentric coordinates. For this purpose, consider a triangle defined by its vertices v_1, v_2

and v_3. Each point v in the triangle plane or on its edges can be described as a barycentric combination

$$v = \sum_{i=1}^{3} b_i v_i$$

of the vertices, where $0 \leq b_i \leq 1$ and $\sum_{i=1}^{3} b_i = 1$. For each of the, usually 5 to 6, triangles containing $\hat{d}_{\mathrm{MAP}}^{(k)}$ as one of its vertices we assume linearity of the marginal distribution $\tilde{\psi}$ defined in (4.7) and determine the barycentric coordinates b_i, $i = 1, 2, 3$, of point v solving

$$\underset{\substack{b_1,b_2,b_3\in[0,1]\\ \sum_{i=1}^{3} b_i=1}}{\operatorname{argmin}} \left\{ -\sum_{i=1}^{3} b_i \tilde{\psi}(x^{(k)}, v_i) + \beta \left\| \sum_{i=1}^{3} b_i v_i - d_{\mathrm{extr}}^{(k)} \right\|_2^2 \right\}$$

with $\beta > 0$. Choosing the overall minimum comparing results for the neighboring triangles that share vertex $\hat{d}_{\mathrm{MAP}}^{(k)}$, we obtain the final tracking direction $d^{(k+1)} \in \mathcal{S}^2$ as a refinement of $\hat{d}_{\mathrm{MAP}}^{(k)}$. To solve these constrained quadratic minimization problems, we implemented an active set strategy. Typically, it finds the global minimum in only 2–3 iteration steps.

4.2.3. Stopping Criteria

The proposed algorithm stops when reaching a region outside a previously defined valid domain $W \subseteq \Omega$. In general, this domain can be defined as a $N_x \times N_y \times N_z$ binary mask M. For instance, M can be a binary brain mask with elements

$$M_{\mathrm{B}}(i,j,k) = \begin{cases} 1 \text{ if voxel } (i,j,k) \text{ is inside the brain} \\ 0 \text{ else} \end{cases}$$

for $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$, $k = 1, \ldots, N_z$, or a subset of M_{B} derived from binary tissue segmented maps, or tissue partial volume maps. For this purpose, let $I_{\mathrm{WM}}, I_{\mathrm{GM}}$, and I_{CSF} denote tissue partial volume images as described in Section 2.6.2. Then we define,

$$M_{\mathrm{T}}(i,j,k) = \begin{cases} 0 \text{ if } I_{\mathrm{GM}}(i,j,k) \geq \theta_{\mathrm{GM}} \text{ and } I_{\mathrm{CSF}}(i,j,k) \geq \theta_{\mathrm{CSF}} \\ 1 \text{ else} \end{cases}$$

for $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$, $k = 1, \ldots, N_z$, where θ_{GM} and θ_{CSF} are threshold values e.g. $\theta_{\mathrm{GM}} = \theta_{\mathrm{CSF}} = 0.5$, similar to Smith et al. (2012) or Girard et al. (2014). The algorithm breaks if $M_{\mathrm{T}}(i,j,k) = 0$. Note that in the case of binary

GM and CSF maps, $\theta_{\mathrm{GM}} = \theta_{\mathrm{CSF}} = 1$. In Section 4.3, we describe the masks used for the respective data sets in more details. Besides using a predefined mask, we also experimented with thresholds on the a-posteriori probability. Results generated in this way are discussed in Section 4.3.

4.2.4. Including Anatomical Priors in the Bayesian Framework

It is also possible to include additional anatomical prior information in the posterior probability to inform the algorithm about tissue types present in its neighborhood. As a first approach, we have achieved this by weighting the transition probabilities of the prior distribution with a combined white and gray matter binary (see results for the *Fiber Cup* phantom) or partial volume map (see results for *in vivo* data), that is, Equation (4.3) is extended to

$$P(\mathcal{F}^n) = \prod_{j=0}^{n-1} \underbrace{A\left(x^{(k+j+1)}\right) p\left(d^{(k+j)} | d^{(k+1)}, \ldots, d^{(k-1+j)}\right)}_{=:\tilde{p}\left(d^{(k+j)} | d^{(k+1)}, \ldots, d^{(k-1+j)}\right)}$$

where map $A = I_{\mathrm{GM}} + I_{\mathrm{WM}}$ is interpolated at $x^{(k+j)}$. The weighted transition probabilities $\tilde{p}$ are normalized such that $\sum_{\mathcal{F}^n \in \mathcal{U}^n} P(\mathcal{F}^n) = 1$. The voxel weights in A are chosen as the sum of WM and GM proportions, since we do not want the algorithm to avoid entering GM. Termination of tracks in GM regions is ensured by the stopping criteria described in 4.2.3.

4.2.5. Algorithm

We summarize the tractography algorithm (Algorithm 1) combining both methods described above, *forward search* and *guiding by extrapolation*. The following steps are executed for a maximum number of iterations or while the current fiber path point $x^{(k)}$ lies inside the physiologically reasonable region $W \subseteq \Omega$ determined according to Section 4.2.3.

Algorithm 4.1. (Tractography with 2.1 and 2.2)
Input:
$x^{(0)} \in \Omega$ - starting point
$d^{(0)} \in \mathcal{U}$ - starting direction
n - # forward search steps
N - # extrapolation points
$\lambda, \mu > 0$ - step lengths

$\varphi \in [0, \pi/2)$ - angle
type - deterministic or probabilistic

$x^{(0)} = x^{(1)} + \lambda d^{(0)}$ 1
$k = 1$ 2
while $x^{(k)} \in W$ 3
  for each $\mathcal{F}^n$ 4
    $[P(\mathcal{F}^n), P(\mathcal{D}|\mathcal{F}^n)] = \text{forwardSearch}\Big(x^{(k)}, \mathcal{F}^n, \varphi, n, N, \mu\Big)$ 5
  end 6
  normalize $P(\mathcal{F}^n)$ and $P(\mathcal{D}|\mathcal{F}^n)$ 7
  $P(\mathcal{F}^n|\mathcal{D}) = P(\mathcal{F}^n)P(\mathcal{D}|\mathcal{F}^n)$ 8
  normalize $P(\mathcal{F}^n|\mathcal{D})$ 9
  if type == deterministic % det. algorithm 10
    $\hat{\mathcal{F}}^n_{\text{MAP}} \in \underset{\mathcal{F}^n \in \mathcal{U}^n}{\text{argmax}} P(\mathcal{F}^n|\mathcal{D})$ where $\hat{\mathcal{F}}^n_{\text{MAP}} = \Big(\hat{d}^{(k)}_{\text{MAP}}, \ldots, \hat{d}^{(k+n-1)}_{\text{MAP}}\Big)$ 11
    $d^{(k)} = \text{refine}\Big(\hat{d}^{(k)}_{\text{MAP}}\Big)$ % see 4.2.2 12
  else % prob. algorithm 13
    draw $d^{(k)} \in \mathcal{U}$ from probability distribution 14
      $\sum_{d^{(k+1)} \in \mathcal{U}} \cdots \sum_{d^{(k+n-1)} \in \mathcal{U}} P(\mathcal{F}^n|\mathcal{D})$ where $\mathcal{F}^n = \Big(d^{(k)}, \ldots, d^{(k+n-1)}\Big)$ 15
  end 16
  $x^{(k+1)} = x^{(k)} + \lambda d^{(k)}$ 17
  $k = k + 1$ 18
end 19

Output: $\Big(x^{(0)}, x^{(1)}, \ldots\Big)$ - fiber path

First, a forward search is performed, exploring the neighborhood in front of the current point $x^{(k+1)}$. For this purpose, the posterior probability $P(\mathcal{F}^n|\mathcal{D})$ is determined for each fiber fragment $\mathcal{F}^n$ where n denotes the number of forward search steps as shown in Section 4.2. The forward search algorithm (Algorithm 4.2) is explained in detail below. The deterministic version of Algorithm 4.1 is executed, if type = deterministic is specified, in lines 11 and 12. Here, a new direction $d^{(k+1)}$ is selected such that it equals the candidate direction of the first forward search step corresponding to the fiber fragment $\mathcal{F}^n$ that maximizes the a-posteriori probability. Subsequently, $d^{(k+1)}$ is refined according to the method explained in Section 4.2.2. The probabilistic version of Algorithm 4.1 where a new direction $d^{(k+1)}$ is drawn from the marginal distribution is specified in lines 14 and 15. Then, a new path point $x^{(k+1)}$ is computed in line 17.

Algorithm 4.2 summarizes the forward search method that estimates the posterior probability $P(\mathcal{F}^n|\mathcal{D})$ for each of the candidate fiber fragments $\mathcal{F}^n$. In line 1, values for prior and likelihood are initialized with one. Then, for each forward

search step $j = 0, \ldots, n-1$ a guiding direction at $x^{(k+j)}$ is determined from the N previously tracked points using the method described in Section 4.1. In lines 4 to 8, the transition probability of $d^{(k+j)}$ is specified according to its deviation from the guiding direction. In line 9, the value for the prior is updated by multiplying the transition probability. Similarly, the likelihood is updated in line 10 where $\psi : \mathcal{S}^2 \times \mathbb{R}^2 \to \mathbb{R}$ denotes the orientation distribution function.

Algorithm 4.2. forwardSearch
Input:
$x^{(k)}$ - path point at iteration k
$\mathcal{F}^n$ - fiber fragments
φ - angle
n - # forward search steps
N - # extrapolation steps
μ - step length

initialize $P(\mathcal{F}^0) = 1, P(\mathcal{D}|\mathcal{F}^0) = 1$ 1
for $j = 0, \ldots, n-1$ 2
  $d_{\text{extr}}^{(k+j)} = \text{extrGuiding}\left(x^{(k+j)}, \ldots, x^{(k+j-N+1)}\right)$ 3
  if $\sphericalangle\left(d^{(k+j)}, d^{(k+j-1)}\right) \leq \varphi$ 4
    $p\left(d^{(k+j)}|d^{(k)}, \ldots, d^{(k-1+j)}\right) = \exp\left(-\frac{\sphericalangle\left(d^{(k+j)}, d_{\text{extr}}^{(k+j)}\right)^2}{\sigma^2}\right)$ 5
  else 6
    $p\left(d^{(k+j)}|d^{(k)}, \ldots, d^{(k-1+j)}\right) = 0$ 7
  end 8
  $P(\mathcal{F}^n) = P(\mathcal{F}^{n-1}) \cdot p\left(d^{(k+j)}|d^{(k)}, \ldots, d^{(k-1+j)}\right)$ 9
  $P(\mathcal{D}|\mathcal{F}^n) = P(\mathcal{D}|\mathcal{F}^{n-1}) \cdot \psi\left(x^{(k+j)} + \frac{\mu}{2} d^{(k+j)}, d^{(k+j)}\right)$ 10
end 11

Output: $P(\mathcal{F}^n)$, $P(\mathcal{D}|\mathcal{F}^n)$

In Algorithm 4.3, we summarize the steps necessary to compute a guiding direction, as described in Section 4.1.

Algorithm 4.3. extrGuiding
Input: $x = \left(x^{(k)}, \ldots, x^{(k-N+1)}\right)$

for j = 1,2,3
  Solve $W^{\frac{1}{2}} V c^{(j)} = W^{\frac{1}{2}} x$
  $\gamma_j\left(t, c^{(j)}\right) = c_0^{(j)} + c_1^{(j)} t + c_2^{(j)} t^2$
end

$\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$
$d_{\text{extr}}^{(k)} := \frac{\gamma(t_{-1}) - x^{(k)}}{\left\|\gamma(t_{-1}) - x^{(k)}\right\|_2}$

Output: $d_{\text{extr}}^{(k)}$

4.3. Experiments and Results

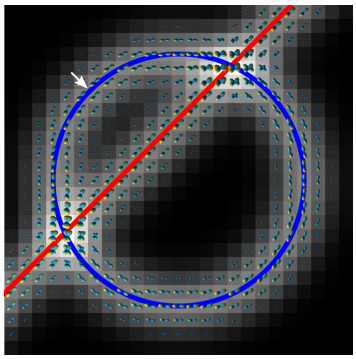
At first, we show results for applying our method to an academic example of an ODF field consisting of a circular bundle crossed by a straight bundle. This simple ODF field is simulated by an own implementation. The data set and the procedure used for data simulation are described in Section 3.1.1. Furthermore, we tested our algorithm on two simulated data sets, the ISBI 2013 challenge training data set and the ISMRM 2015 tractography challenge data, outlined in Sections 3.1.2 and 3.1.3, respectively, on acquisitions of two diffusion phantoms, the *Fiber Cup* diffusion phantom (see Section 3.2.1), and a spherical diffusion phantom (see Section 3.2.2), as well as on *in vivo* data provided by the HCP (see Section 3.3).

Our algorithm has originally been developed based upon ODF reconstructions obtained from high-angular-resolution diffusion imaging (HARDI) data using constrained spherical deconvolution (CSD) with spatial regularization as described by Hohage and Rügge (2015). To demonstrate that our algorithm is not limited to data obtained from that method but can be applied independently from the fiber ODF acquisition technique, we also include some results using ODF reconstructions generated by the CSD algorithm included in the software package MRtrix[1].

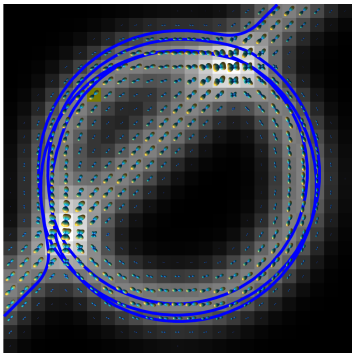
4.3.1. Numerical ODF Data

Creation and properties of the simulated data set considered here are detailed in Section 3.1.1. In Figure 4.5, results are illustrated from applying to the data (a) the basic tractography strategy stated in Algorithm 2.1 in Section 2.7.1, (b) an expansion of this basic algorithm by our extrapolation guided approach described in Section 4.1, and (c) the deterministic version of our tractography algorithm summarized in Section 4.2.5, respectively. As seed point, we choose a

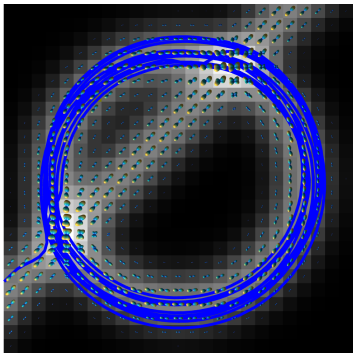
[1] http://www.mrtrix.org/



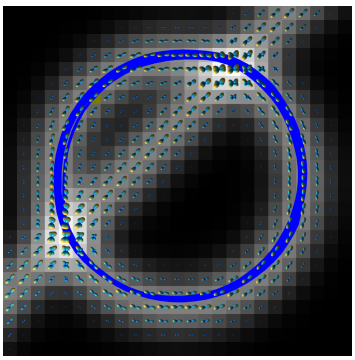
(a) Ground truth.



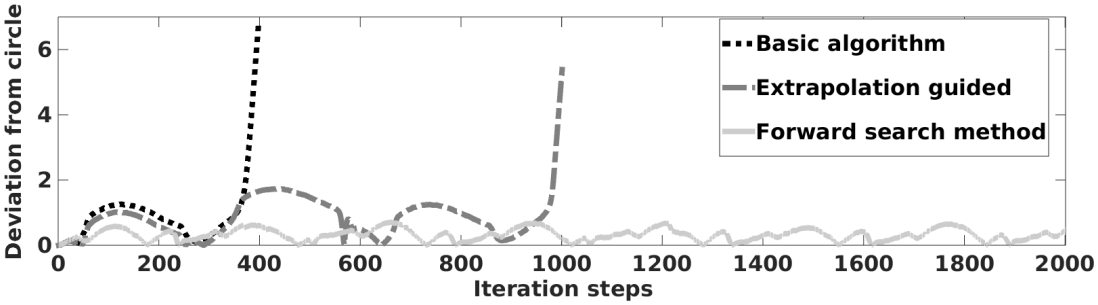
(b) Basic algorithm.



(c) Extrapolation guided.



(d) Forward search.



(e) Minimum distance $\min_t \left\| x^{\text{circle}}(t) - x^{(k)} \right\|_2$ of the counter-clockwise track from the ground truth curve $x^{\text{circle}}(t)$.

Figure 4.5.: Comparison of results for fiber tracking on the simulated data for three different methods with step length $\lambda = 0.2$. Each method is used to generate one track (plotted in blue) in clockwise and counter-clockwise direction, respectively, starting from a seed point on the top left highlighted in yellow in Subfigures 4.5b, 4.5c, 4.5d and by the white arrow in Figure 4.5a. The maximum number of iterations is set to 2000.

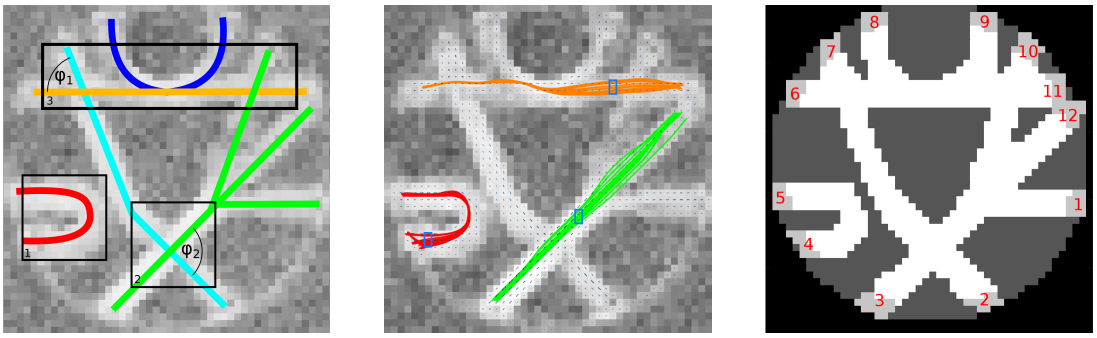
point on the top right of the circle highlighted by the white arrow in Figure 4.5a. The step lengths λ and μ are chosen as $0.2 \times$ voxel-size and the maximum number of iterations is set to 2000. The basic tractography algorithm whose result is shown in Figure 4.5b successively steps in the direction maximizing the local ODF. The optimal direction is searched for in each iteration step in a region of $\varphi = \pi/36$ around the former tracking direction, thus allowing for a maximum of 5° difference between consecutive tracking directions. This curvature restriction is chosen to avoid problems at the crossings while it still allows enough flexibility in tracking the circle with radius $\approx 9 \times$ voxel-size (discretizing the ground truth circle with step-length $0.2 \times$ voxel-size results in circle segments of angle $\approx 1.27°$). The algorithm used to generate the track in Figure 4.5c merely differs in the fact that in each iteration step k the maximum is located in a region of $\varphi = \pi/36$ around the *guiding direction* $d_{\text{extr}}^{(k)}$ determined by extrapolation of a quadratic polynomial fitted through the $N = 6$ previously tracked points. Figure 4.5d shows results for the *forward search* algorithm without further refinement, i.e. simply taking the direction obtained from the maximum a-posteriori estimate. We choose the further parameters as follows: The number of forward search steps is set to $n = 2$, we compute the *guiding direction* from $N = 6$ previously tracked path points as explained above, and the standard deviation of the Gaussian transition probabilities (see equation (4.4)) is set to $\sigma = 2\pi$. Moreover, the angle indicating the size of the candidate region around $d_{\text{extr}}^{(k)}$ is set to $\varphi = \pi/36$. For 2562 sampling points on the sphere, this results in approximately 5-6 candidate directions per cone-shaped region. Besides, this choice allows for approximately the same maximum deviation between successive tracking directions as for both other methods regarded in this section. Regarding the tractography results of the different algorithms we observe the following: Applying the basic algorithm, the evolving fiber track is spiraling outward and deflects after several rounds at the margin of one of the crossing structures, as can be seen in Figure 4.5b. For instance, the track that is running in counter-clockwise direction breaks after 397 iterations ($\approx$ 1.5 circles). As shown in Figure 4.5c, including a *guiding direction* by extrapolation the tracking algorithm performs much better. However, also severe deviations from the ground truth can be discovered that finally result in deflection from the ring and subsequent termination in counter-clockwise direction, though only after 1000 iteration ($\approx$ 3.5 circles). In comparison, in Figure 4.5d it can be observed that when applying two steps of *forward search* the tracked path shows hardly any deviations from the ground truth even after the maximum of 2000 iterations ($\approx$ 7 circles). In Figure 4.5e, the deviation of the current path point $x^{(k)}$ from the ground truth circle $x^{\text{circle}}(t)$ calculated as $\min_t \left\| x^{\text{circle}}(t) - x^{(k)} \right\|_2$ is plotted against the number of iteration steps k for the three different methods.

4.3.2. *Fiber Cup* Phantom

In this section, we analyze the performance of our method on the Fiber Cup diffusion phantom data described in Section 3.2.1. We test both deterministic and probabilistic version of Algorithm 4.1 and compare them to the basic methods stated in Algorithms 2.1 and 2.2, and to state-of-the-art algorithms.

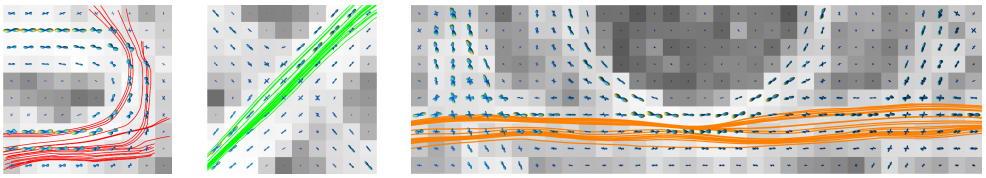
Qualitative Analysis

In Figure 4.6a, the *Fiber Cup* phantom is illustrated showing ground truth and three regions where we analyze the tractography results and compare them to the deterministic SD_Stream with 4-th order Runge-Kutta integration and the probabilistic IFOD-2 algorithm, both included in the MRtrix software. The first region consists of a U-shaped bundle, the second region includes an approximately 90°-fiber-crossing and the third region a crossing of approximately 70°. In Figure 4.6b, results of the deterministic version of our proposed algorithm are shown using ODF data obtained using CSD with spatial regularization. The parameter settings of the tractography algorithm are as follows: We chose step length $\lambda = 0.5 \times$ voxel-size for tract evolution, $n = 2$ *forward search* steps, *forward search* step length $\mu = 1 \times$ voxel-size, standard deviation $\sigma = \pi$ for the transition probabilities, regularization parameter $\beta = 0.5$ for the computation of the overall tracking direction, and the $N = 6$ previously tracked points for extrapolation of a *guiding direction*. The angle indicating the maximum allowed deviation of candidate directions from the guiding direction is set to $\varphi = \pi/9$, resulting in approximately 20 candidate directions per cone-shaped region. The algorithm terminates if it enters the regions resembling gray matter and CSF defined in Figure 4.6c, or if it leaves the circular phantoms boundaries. Moreover, the data has been zero-padded in the third dimension, which is originally only three slices thin, to avoid early termination due to exiting the spatial dimensions of the phantom. Figure 4.7 compares results of a basic tractography algorithm and the forward search algorithm with respect to the three rectangular regions in Figure 4.6a. The seed points are the same as those used in Figure 4.6b. In the first row of Figure 4.7, the outcome of the basic tracking method labeled Algorithm 2.1 is shown using spatially regularized ODF data. The basic algorithm successively steps with step length $\lambda = 0.5 \times$ voxel-size in the direction that maximizes the local ODF. In each iteration step, the maximizing direction is located in a region of $\varphi = \pi/9$ around the previous tracking direction, which allows for a maximum of 20° difference between successive tracking directions. Regarding the sharp bending in Region 1, it can be observed that all tracks terminate before completing the U-turn. On the other hand, the 90°-crossing

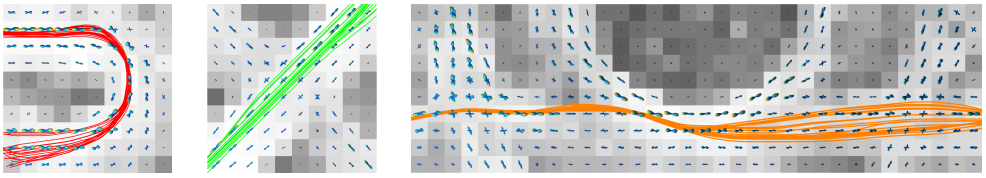


(a) Ground truth. **(b)** Results. **(c)** Segmentation.

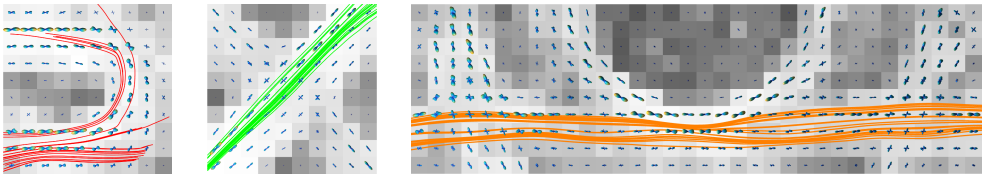
Figure 4.6.: GFA map of the *Fiber Cup* phantom overlaid with *left:* ground truth fibers, three regions of interest and angles $\varphi_1 \approx 70°$, $\varphi_2 \approx 90°$, and *middle:* fiber tracking results using our proposed deterministic *forward search* method with 20 seed points randomly selected in each of the areas framed in blue. *right:* Mask defining white matter (*white*), gray matter (*light gray*) and fluid filled regions (*dark gray*). The 12 seed regions used for quantitative analysis are highlighted in red.



(a) Basic algorithm



(b) Forward search

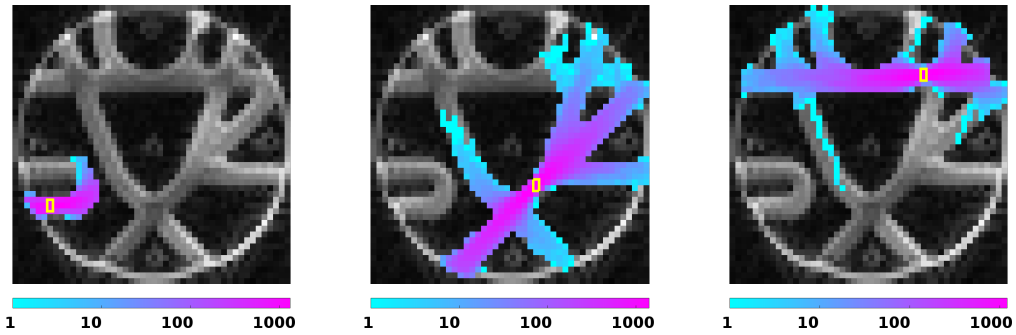


(c) SD_Stream -rk4

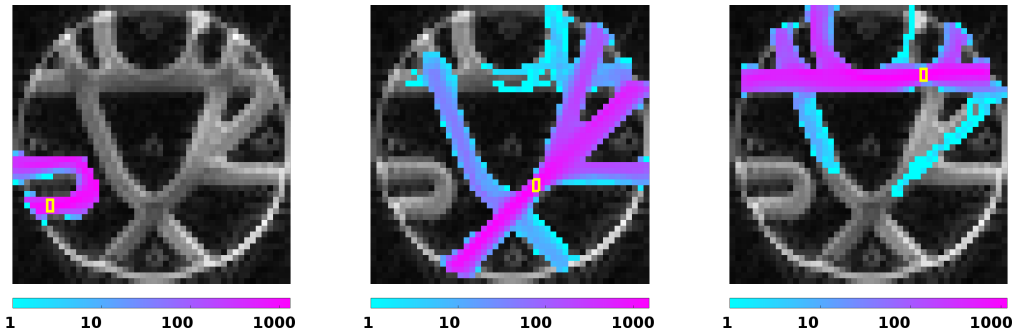
Figure 4.7.: Deterministic tractography results for the three regions specified in Figure 4.6a using fiber ODF data obtained with spatially regularized CSD. *Top row*: Basic algorithm with step length $\lambda = 0.5 \times$ voxel-size. *Middle*: Forward search method with $\lambda = 0.5 \times$ voxel-size. *Bottom row:* MRtrix SD_Stream algorithm with 4-th order Runge-Kutta integration, step size $0.5 \times$ voxel-size and default maximum angle between successive steps (45°).

and the more challenging 70°-crossing in the third column are correctly passed by all tracks. Shown in the second row are enlarged image details of the results of our *forward search* algorithm presented in Figure 4.6b. Here, not only do all tracks perfectly pass the 90°-crossing and 70°-crossing, but also the U-turn is tracked by all fibers without deviation. In the bottom row, results for applying the SD_Stream algorithm with 4-th order Runge-Kutta integration included in MRtrix are shown. The step size is chosen, equal to the other two methods, as $0.5 \times$ voxel-size . According to the step size, the default angle between successive steps is used ($90° \times$ step-size/voxel-size $= 45°$). The algorithm performs very good for the straight bundles. Regarding the sharp bending, SD_Stream performs better than the basic algorithm, but still most tracks break without completing the turn.

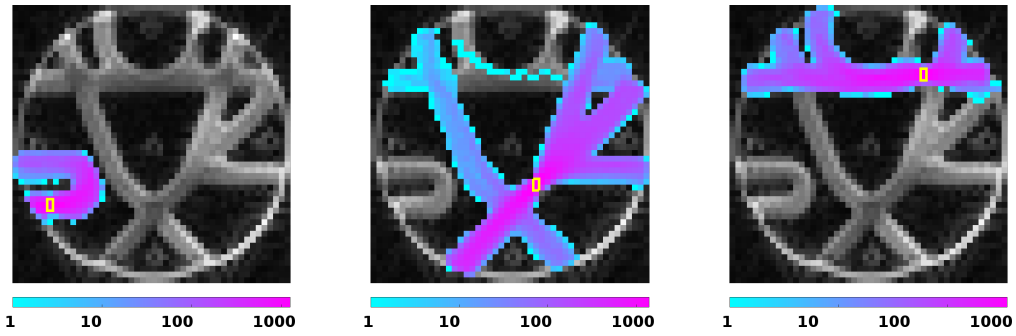
In Figure 4.8, we compare the probabilistic version of our *forward search* method to the simple probabilistic tractography algorithm referred to as Algorithm 2.2, and to the MRtrix IFOD-2 method. Tracking starts from 2000 seed points randomly selected in each of the three regions that where also used for the deterministic tracking in Figure 4.6b. The basic probabilistic algorithm regards the ODF at the current spatial position as a probability distribution restricted to a region of maximum deviation of 20° from the former tracking direction. In each iteration, the track evolves by stepping with step length $\lambda = 0.5 \times$ voxel-size in a direction drawn from this distribution. The results are shown in Figure 4.8a in the form of general fractional anisotropy (GFA) maps with voxels colored according to the number of tracks passing through. The number of voxel visits is summed up along the three slices of the data set in the third dimension. Similarly, the results for the probabilistic forward search algorithm setting $\lambda = 0.5 \times$ voxel-size, $\mu = 1 \times$ voxel-size, $n = 2$, $N = 6$, $\sigma = \pi$, and $\varphi = \pi/9$ are presented in Figure 4.8b. Regarding the first region of interest, none of the 2000 tracks generated by the basic algorithm manages the U-turn without deflection, while it is passed by 1050 of 2000 tracks produced by the forward search method. Starting in the second seed region, 807 of 2000 tracks resulting from the basic algorithm and 1708 of 2000 tracks resulting from the forward search algorithm correctly pass the 90°-crossing. Furthermore, after the 70°-crossing in Region 3, we count 87 of 2000 tracks from the basic algorithm and 1275 of 2000 tracks generated by our algorithm that had started in the third region. In Figure 4.8c, the results for IFOD-2 using the default parameter settings ($\lambda = 0.5 \times$ voxel-size, 45° maximum angle between successive steps) are shown. Here, 320 of 2000 tracks manage to track the U-turn, 1163 of 2000 correctly pass the 90°-crossing, and 154 of 2000 the 70°-crossing.



(a) Basic probabilistic algorithm (number of track visits)



(b) Probabilistic forward search algorithm (number of track visits)



(c) MRtrix IFOD-2 (number of track visits)

Figure 4.8.: Probabilistic tractography results for the three regions specified in Figure 4.6a, comparing a basic probabilistic algorithm (Figure 4.8a), our proposed method (Figure 4.8b), and the MRtrix IFOD-2 algorithm with default parameters (Figure 4.8c). Tracking starts from 2000 seed points randomly selected from inside the region defined in yellow. The voxels are colored according to the number of track visits on a logarithmic scale.

Quantitative Analysis

For the *Fiber Cup* phantom, we model white matter, gray matter, and CSF as shown in Figure 4.6c. The fiber ODF data was zero-padded in third dimension as described above and set equal to zero in CSF regions. This causes a weighting of candidate directions according to this information, as theoretically described in Section 4.2.4. For quantitative analysis, 100 seed points per region were randomly selected on the interface between white and gray matter. Tracks were generated from these (identical) seed points using the deterministic and probabilistic basic in-house algorithms described above, the deterministic and probabilistic versions of the proposed algorithm, and both the deterministic SD_Stream with 4-th order Runge-Kutta integration and the probabilistic IFOD-2 included in MRtrix. Moreover, we ran our proposed method on (spatially regularized) ODF data where the CSF regions where not set to zeros for comparison. The results are stated in the column with heading *not anatomically-informed*. The algorithms stop if entering CSF regions, reach gray matter, or the outer bounds of the circular phantom.

We used Algorithm 2.3 to categorize the tracks and computed the Tractometer values defined in Section 2.7.3 to evaluate the results. *Valid connections* (VC), *invalid connections* (IC), *no connections* (NC), *valid connection to connection ratio* (VCCR) and *connection to seed ratio* (CSR) are listed in percent in Tables 4.1 and 4.2 for the performance on the overall dataset of the deterministic and probabilistic algorithms, respectively. In Table 4.1, we observe that the proposed

	spatially regularized fODF				standard fODF		
	basic	proposed		SD_Stream -rk4	basic	proposed	SD_Stream -rk4
			na				
VC	40.25	**90.58**	75.33	54.65	31.17	**87.42**	32.43
IC	2.33	6.25	2.08	0.00	11.08	7.00	2.54
NC	57.42	**3.17**	22.58	43.67	57.75	**5.58**	62.66
VCCR	94.52	93.55	97.31	100.00	73.77	92.59	92.74
CSR	42.58	**96.83**	77.42	55.58	42.25	**94.42**	35.82

Table 4.1.: Tractometer values for the three deterministic algorithms compared on the *Fiber Cup* phantom (na = not anatomically-informed). All data are given in percent.

method has the highest percentage of VC and the lowest percentage of NC using both standard and spatially regularized fODFs. The VCCR is above 90% for all methods, whereas the proposed method achieves the highest CSR. MRtrix could not generate tracks from approximately 2% of the seed points for SD_Stream which does not have a significant influence on the CSR values. Note, that we did *not* include these as *no connections*. Regarding results for the probabilistic

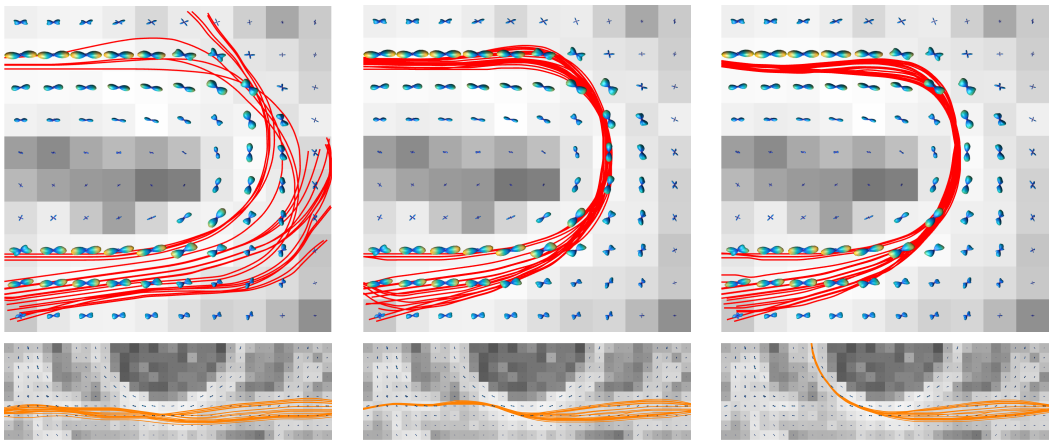
	spatially regularized fODF				standard fODF		
	basic	proposed		IFOD-2	basic	proposed	IFOD-2
			na				
VC	0.58	**56.83**	26.92	12.83	0.83	**51.25**	9.92
IC	0.67	25.67	10.08	11.17	1.25	25.25	7.42
NC	98.75	**17.50**	63.00	74.92	97.92	**23.50**	81.75
VCCR	46.67	68.89	72.75	53.47	40.00	66.99	57.21
CSR	1.25	**82.50**	37.00	24.26	2.08	**76.50**	17.49

Table 4.2.: Tractometer values for the three probabilistic algorithms compared on the *Fiber Cup* phantom (na = not anatomically-informed). All data are given in percent.

algorithms in Table 4.2, the proposed method has again the highest percentage of VC and CSR and the lowest percentage of NC. The rather high amount of invalid connections is due to the fact that overall very few tracks terminate in white matter. Essentially, invalid connections can be reduced in the proposed method (though at the cost of VC) by terminating if the posterior probability falls below a certain threshold. The VCCR is comparatively high with around 70%. Comparing the two settings of the proposed method, we observe that especially the probabilistic algorithm benefits from using information about CSF regions in the tracking process.

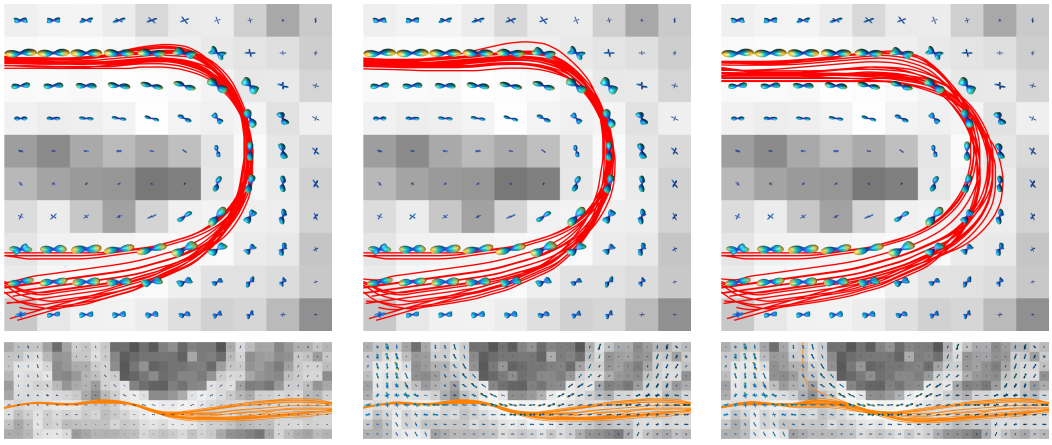
Effects of Parameter Changes

Figure 4.9 illustrates the effects of parameter changes with respect to two of the regions considered above. All other parameter choices are equal to those described above. Different choices for the forward search step length μ, are shown in Figure 4.9a. It is not advisable to chose μ much smaller than the actual step length λ. The result could be improved by increasing the number of forward steps n, but this would increase computation time. On the other hand, selecting μ to large causes inaccuracies because of skipping the immediate neighborhood. Different settings for parameter σ are shown in Figure 4.9b. A decrease in clustering of tracks can be observed with decreasing σ, since the impact of the guiding direction becomes stronger while the information obtained by the ODF field is down weighted. These qualitative results are supported by a quantitative evaluation according to the definitions of Section 2.7.3, which is summarized in Table 4.3. One can observe that choosing the forward search step length μ smaller than the actual step length λ (here: $\mu = 0.2$, $\lambda = 0.5$) can cause problems in those bundles with high curvature (4-5 and 8-9), since then forward searching is too local and has no effect. On the other hand, if the immediate neighborhood is skipped by selecting μ much larger than λ (here: $\mu = 1.5$, $\lambda = 0.5$) difficulties in tracking bundle 6-11, which is straight but includes crossing



$\mu = 0.2$ $\mu = 1$ $\mu = 1.5$

(a) Effect of changing the forward search step length parameter μ.



$\sigma = \pi$ $\sigma = \pi/2$ $\sigma = \pi/4$

(b) Effects of changing parameter σ.

Figure 4.9.: Effects of parameter change regarding the proposed method. The other parameters are chosen as described in Section 4.3.2

and kissing structures, can occur. The effects of changing parameter σ, which were visually distinguishable in Figure 4.9 as the degree of clustering of tracks, are insignificant with respect to the quantitative results.

4.3.3. ISBI 2013 HARDI Reconstruction Challenge Data

We also tested and compared our algorithm quantitatively on simulated data used for a reconstruction challenge at ISBI 2013. A more detailed description of the data set can be found in Section 3.1.2. The tractography algorithms are initialized from 100 seed points per seed region using the same parameter settings as in Section 4.3.2 for the *Fiber Cup* phantom. We define the seed regions as gray matter regions and obtain a binary white matter mask from thresholding ADC, i.e. setting all voxel with ADC ≥ 0.6 to one. The algorithm stops when the track leaves white matter. In Table 4.4, the Tractometer values described in Section 2.7.3 are presented. Among the deterministic algorithms, our proposed deterministic method performs only slightly better for the SNR = 30 data with respect to valid connections. On the other hand, regarding the results for the data set with SNR = 10 the proposed algorithm stands out with more than 20% more VC than the basic deterministic Algorithm 2.1 and the SD_Stream algorithm. The results for the proposed probabilistic algorithm show significantly more VC than for the basic probabilistic Algorithm 2.2 and IFOD-2 for both data sets. Furthermore, NC are increasing for SD_Stream and IFOD-2 with decreasing SNR while remaining almost constant for the proposed methods. The relatively high percentage of IC connections for the proposed probabilistic method can be reduced by for instance terminating the algorithm if the posterior probability is below a threshold. In the column headed *cutoff*, metrics for the probabilistic method where tracks are terminated if the posterior is below 0.005 are given. Though this results in a decrease of valid connections and an increase of no connections, the proportion of invalid connections is reduced by about 40%.

4.3.4. ISMRM 2015 Tractography Challenge Data

Furthermore, we tested our algorithm on the ISMRM 2015 tractography challenge data described in Section 3.1.3 and compared the results quantitatively to the basic deterministic and probabilistic algorithms described in Section 2.7.1, and to SD_Stream and IFOD-2 from MRtrix. Tractography starts from the endpoints of 200 randomly selected GT tracks per bundle. The parameter settings of the algorithms are the same as in Sections 4.3.2 and 4.3.3. Tractography terminates if a track leaves the WM mask defined by all voxels containing ground

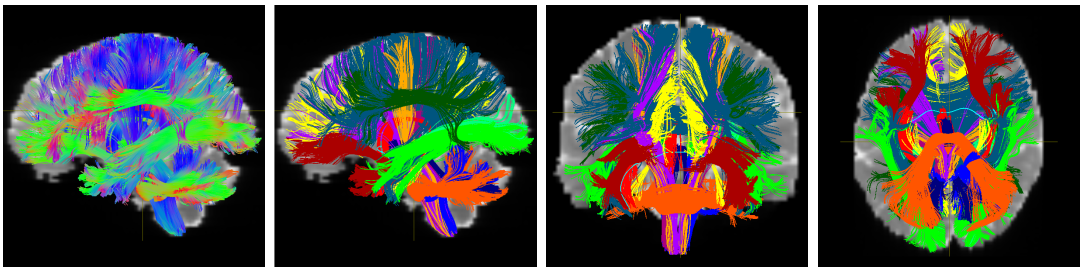


Figure 4.10.: Valid connection (VC) tracks generated for the ISMRM 2015 tractography challenge data by the proposed algorithm.

truth fiber tracks, or if the ODF evaluated at the current tracking direction is below a threshold. The threshold is set to 0.08 for basic and proposed algorithms, and to the MRtrix default for SD_Stream and IFOD-2. The threshold was not selected equally for all algorithm, because the MRtrix algorithms yielded better results for their default cutoff value than for the value more suitable for the other algorithms. Moreover, track length is restricted to a minimum of 5 and a maximum of $100 \times$ voxel-size. For the quantitative evaluation we define end regions as voxels with endpoints of the respective ground truth bundles. Tracks are categorized using Algorithm 2.3. In Table 4.5, the results of the Tractometer evaluation are listed. In addition to VC, IC, NC, VCCR, and CSR, we also counted the number of valid bundles (VB) and invalid bundles (IB) that were detected by the algorithms. As some end regions are overlapping, an invalid track can be assigned to more than one invalid bundle. IB contains all possible combinations and is thus rather high for all algorithms. Regarding the deterministic tracking algorithms, our proposed method generates about 30% more VC and more than 70% less NC than the other two methods. All three algorithms perform similar with respect to IC, whereas VCCR is slightly higher and CSR considerable higher for the proposed algorithm. Of the 25 ground truth bundles, all but the *posterior commissure* are detected. Comparing the probabilistic algorithms, the proposed method and IFOD-2 perform rather similar. On the other hand, the simple probabilistic algorithm produces a lot of NC, verifying the advantage of tracking algorithms of higher order. Figure 4.10 shows the tracks generated by the proposed deterministic algorithm that were assigned to the *valid connection* category.

4.3.5. *In Vivo* Human Brain

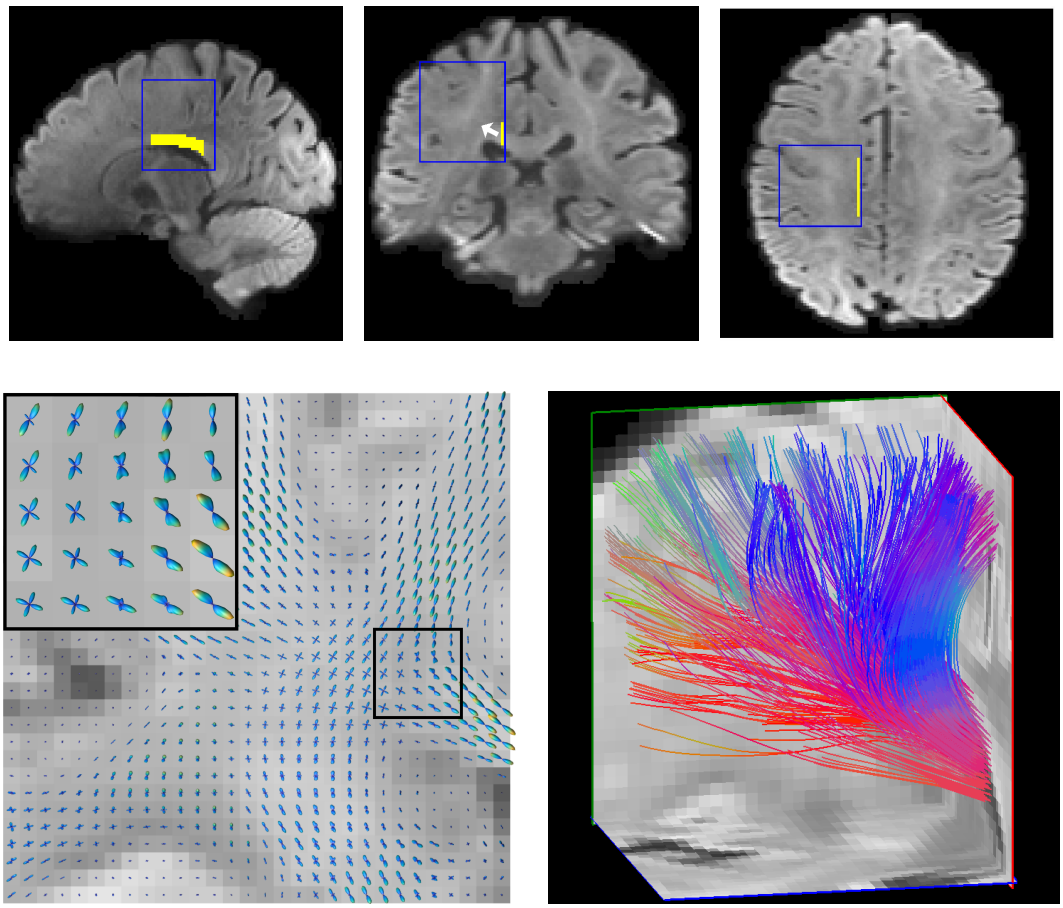
In this section we test our method on an *in vivo* data set provided by the Human Connectome Project. A detailed description of data properties and preprocessing steps can be found in Section 3.3. ODF reconstruction in a part of the

centrum semiovale is depicted in Figure 4.11, along with a corresponding deterministic tracking experiment using the proposed algorithm originating from a seed region in the corpus callosum. Parameter settings for both deterministic and probabilistic version of our algorithm (summarized in Section 4.2.5) are as follows: We chose a tract propagation step size of $\lambda = 0.5 \times$ voxel-size, forward search with $n = 2$ steps and step length $\mu = 1 \times$ voxel-size, $N = 5$ path points for the extrapolation strategy, and regularization parameter $\beta = 0.03$. Tractography experiments were performed seeding from three different regions: the corpus callosum in the mid-sagittal slice, the pyramidal tracts, and the cingulum bundles. The exact locations are illustrated in the right column of Figure 4.12 as projections onto a sagittal, coronal and transversal slice, respectively. The prior distribution was adapted to the different brain regions with respect to the standard deviation using $\sigma = \pi/14$ for the tracks starting in the corpus callosum, $\sigma = \pi/18$ for the pyramidal tracts, and $\sigma = \pi/2$ for the cingulum bundles. The angle determining the difference from the former tracking direction was set consistently to $\varphi = \pi/18$, corresponding to approximately 7 candidate directions per cone-shaped region. Tracking starts in left and right direction from the seed region in the corpus callosum of the mid-sagittal slice, in (+/-)-direction of principle diffusion from the seed region of the pyramidal tracts, and in anterior and posterior direction from the seed region in the cingulum bundles. The tracking algorithm stops when reaching a voxel outside the valid domain $W \subset \Omega$ obtained by combining the white matter mask with GFA and ADC thresholding. In Figure 4.12, results for the deterministic basic tractography Algorithm 2.1 and for our proposed deterministic algorithm are compared. The basic algorithm generates streamlines by following with step size $\lambda = 0.5 \times$ voxel-size the ODF maximizing direction that lies in a region of $\varphi = \pi/18$ around the former stepping direction. Thus, both algorithms allow for a maximum of 10° difference between successive steps. For each voxel in the seed region, 9 tracks were generated, respectively. The fibers generated from seeding in the corpus callosum appear similar. Compared to the tractogram on the right obtained from our proposed algorithm, the lateral bundles in the tractogram generated by the basic algorithm appear more sparse. Vice versa, the vertical projections are a bit more pronounced for the basic algorithm. Increasing parameter σ, our method approaches the outcome of the basic algorithm. The results for the pyramidal tracts demonstrate the potential of our method to increase lateral projection fibers. The result obtained from the basic method captures merely a vertical course of fibers, whereas using the proposed algorithm the fiber tracks show a broader dispersion towards the cortex. Regarding the outcome for the cingulum bundle, the result from the proposed forward search method appears smoother with less spurious tracks. Furthermore, the challenging sharp bendings are managed much better as compared to the basic algorithm. In Figure 4.13, results for different choices of parameter σ are shown. It can be observed, that a small value for σ promotes smooth tracks, but results in earlier termination of the curvy regions of the cin-

gulum bundle. On the other hand, a larger value for σ reduces the reconstruction of lateral connections in the corpus callosum or pyramidal tracts. Results for a first approach to include anatomical information in our tractography algorithm are shown in Figure 4.14. Therefore, the prior probability was in each forward search step multiplied by the sum of WM and GM partial volume map at the respective path point. Tracking terminates when GM or CSF partial volume maps are ≥ 0.5 or if the brain region is left. The results show less spurious tracks, but also an earlier termination of tracks. Probabilistic tractography experiments using our proposed algorithm are illustrated in Figure 4.15. Here, 90 tracks were generated from each voxel in the seed region, respectively. Next to the plotted tracks, a GFA map with voxels colored according to the number of track visits is shown for the three regions of interest, respectively. For these probability maps, the number of tracks passing through is summed up along 20 slices whose exact location is highlighted in the S_0 images on the right side.

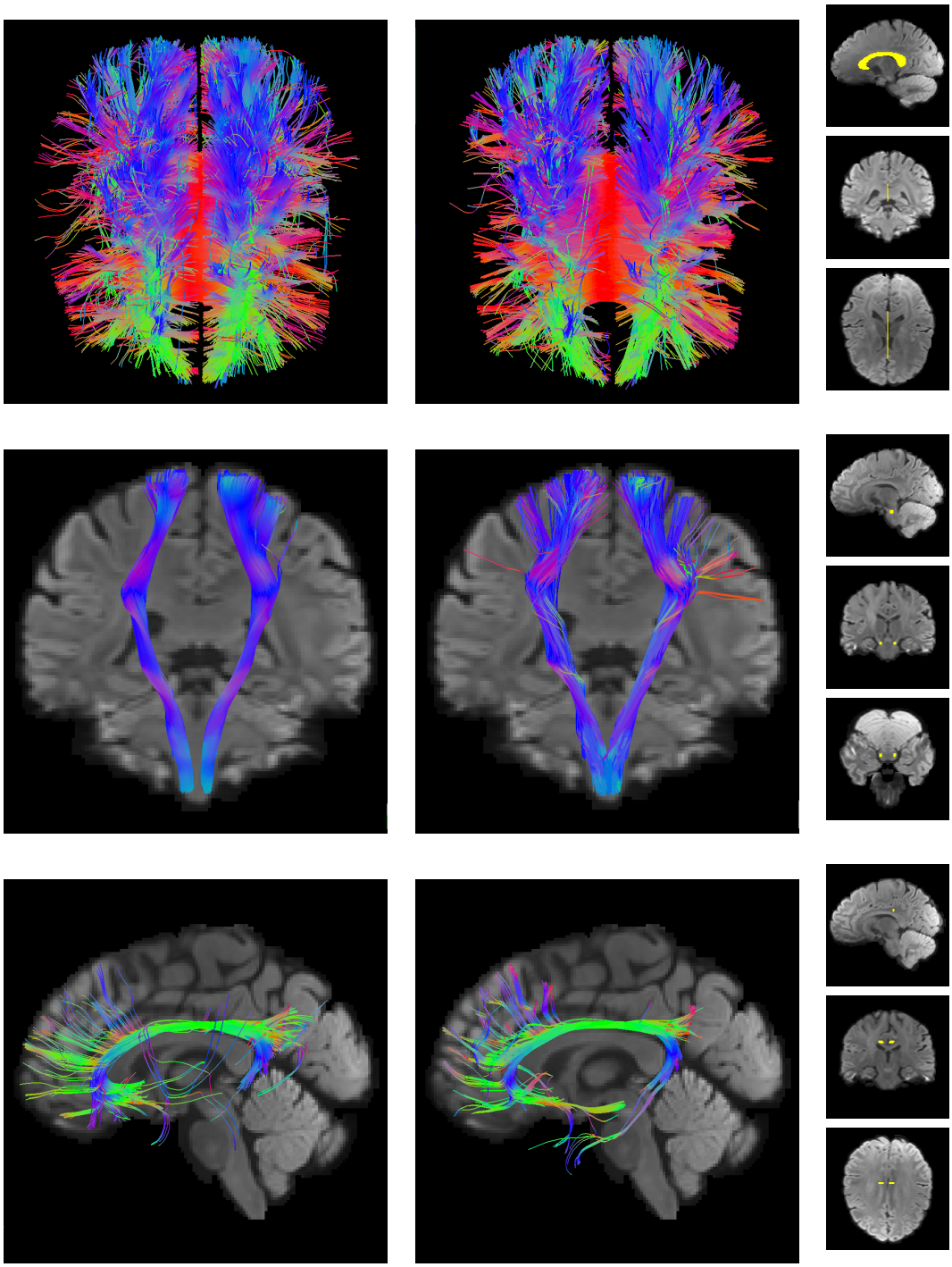
4.3.6. Spherical Diffusion Phantom

Furthermore, we tested the proposed method on data obtained from the spherical diffusion phantom described in Section 3.2.2. Computation and visualization were performed using the diffusion MRI postprocessing kit *pk*. The tractography algorithm parameters were chosen as in Section 4.3.2 for the Fiber Cup phantom, that is $\lambda = 0.5 \times$ voxel-size, $n = 2$, $\mu = 1 \times$ voxel-size, $\sigma = \pi$, $\beta = 0.5$, $N = 6$, and $\varphi = \pi/9$. The algorithm terminates if $\mathrm{FA}_\psi < 0.2$. For comparison, we also computed results using the DT-based FACT algorithm included in *pk*. Here, the algorithm stops if the angle between successive directions is larger than 20°, and $\mathrm{FA}_D < 0.08$. We experienced that increasing the FA threshold resulted in early termination of tracks in the crossing regions. Figures 4.16 and 4.17 illustrate the results for seeding in the bundle that crosses both other bundles at 60° and 30° angle, respectively. The exact seeding ROI is located in the mid-coronal slice and highlighted in red in Figure 4.16a. Tractography starts bi-directionally from the points in the center and at the corners of each voxel in the ROI. In Figures 4.16b and 4.16c, the results are shown after the tracks passed through the crossing for the first time. The bundle resulting from the FACT algorithm fans out at the crossing, i.e. some tracks pass the crossing correctly while others make a wrong turn. This is due to the fact that, as described in Section 2.4, the diffusion tensors at the crossings are isotropic. On the other hand, the crossings are passed correctly without deviations by the proposed ODF-based algorithm. Figure 4.17 shows results where the maximum number of iteration steps of the tracking algorithm is set to 500, but otherwise the parameter setting is the same as above. 500 iteration steps equal approximately 2-3 circles around the sphere. Now, also the proposed algorithm produces a few tracks that make a wrong turn



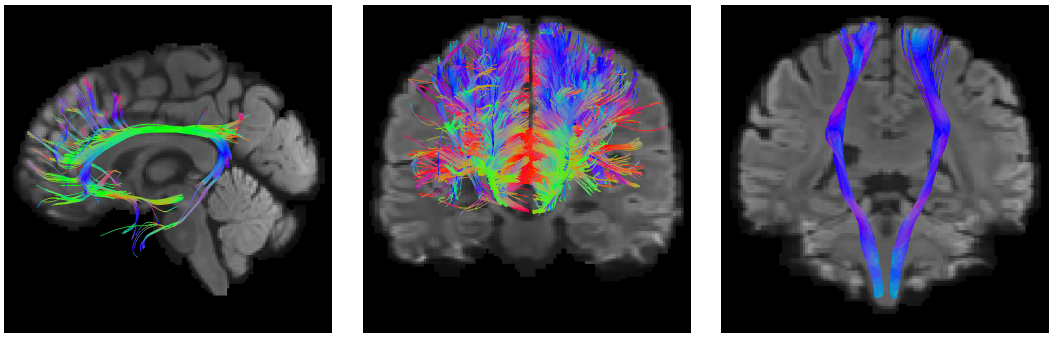
ODF-field in a coronal slice Tracking result

Figure 4.11.: ODFs (*bottom left*) and deterministic tractography using the proposed algorithm (*bottom right*) in a part of the centrum semiovale. The location and seed region for tractography are highlighted above.

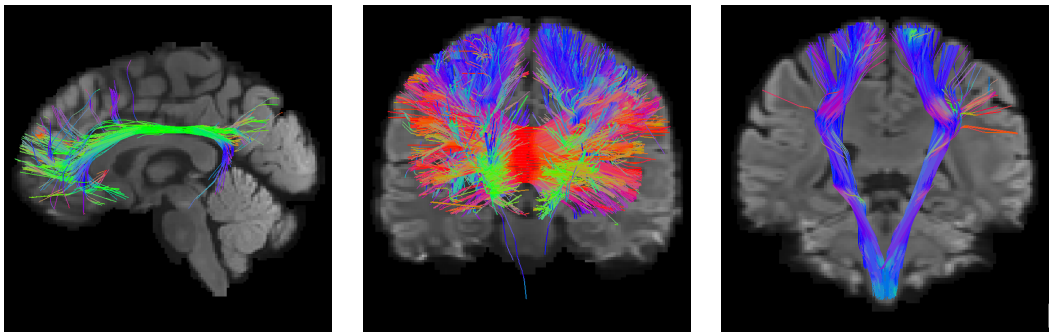


Basic algorithm Forward search

Figure 4.12.: Deterministic tractography results for basic and the proposed forward search method with extrapolation guiding. *First row*: seeding in the center sagittal slice of the corpus callosum, *second row*: seeding in the pyramidal tracts, *third row*: seeding in the cingulum bundles. The seed regions are highlighted in the last column, respectively.

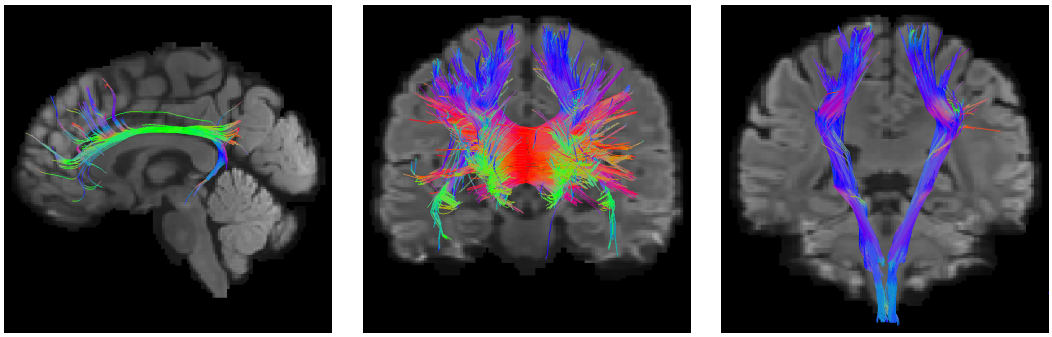


$\sigma = \pi$



$\sigma = \pi/18$

Figure 4.13.: Comparison of different choices for parameter σ. Seed regions and the other parameters are equal to those used for *forward search* in Figure 4.12.



cingulum bundle corpus callosum pyramidal tracts

Figure 4.14.: Results for the anatomically-informed deterministic forward search method using the same parameter settings as for the results in Figure 4.12.

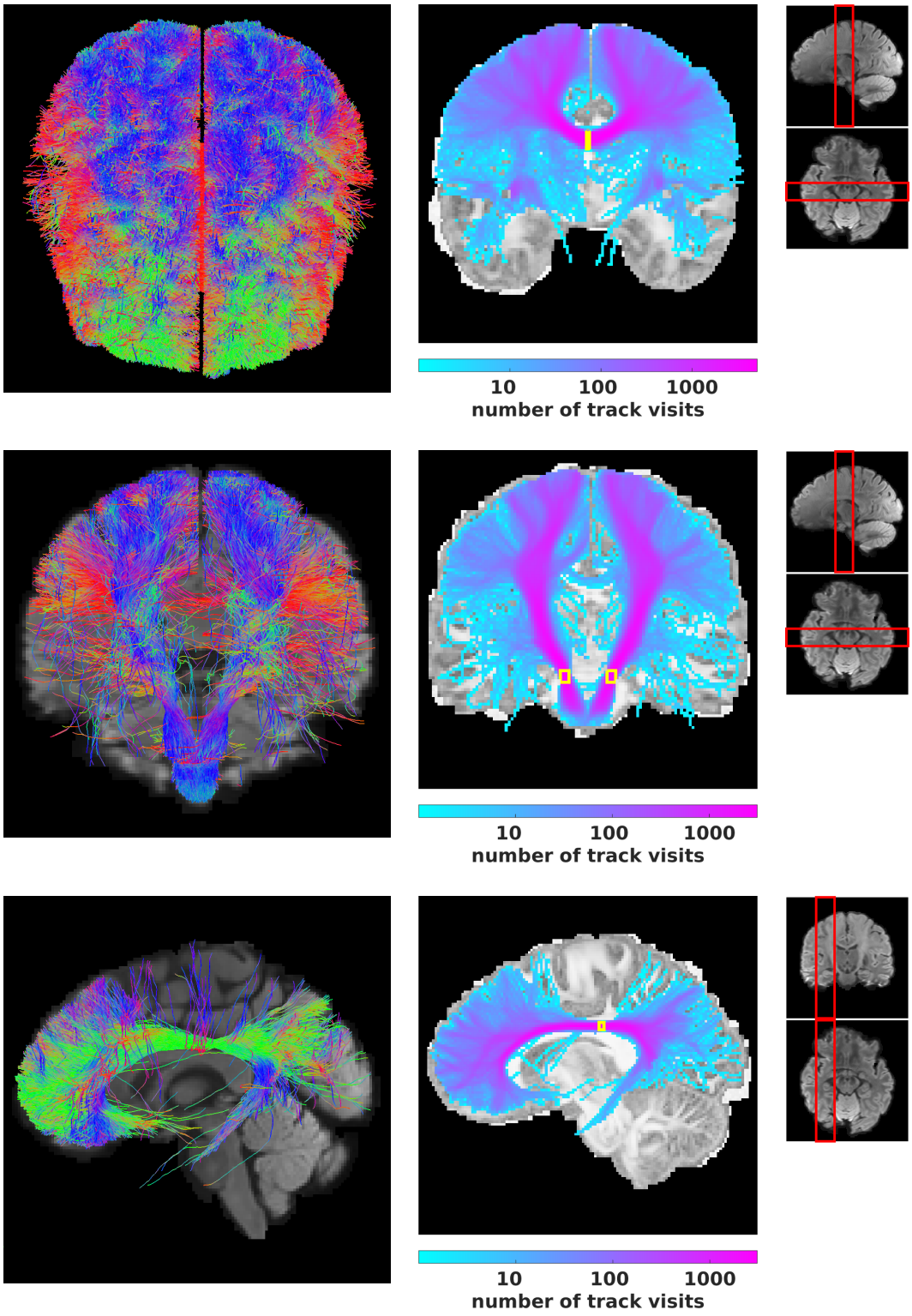
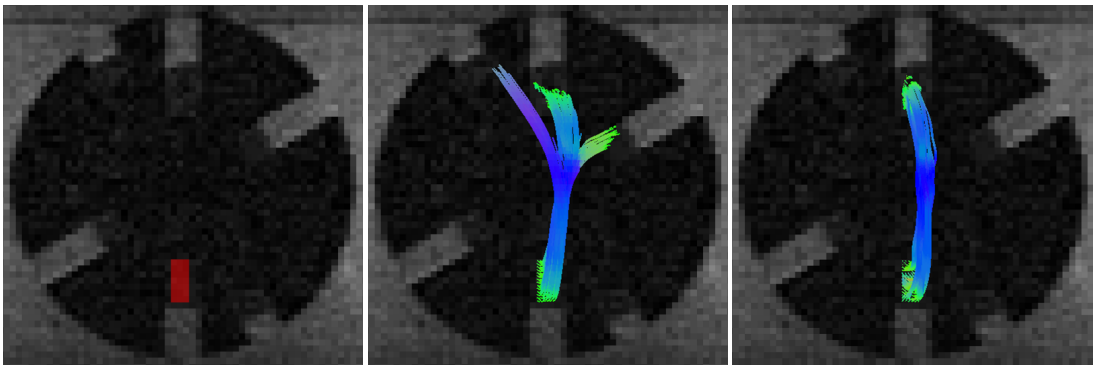
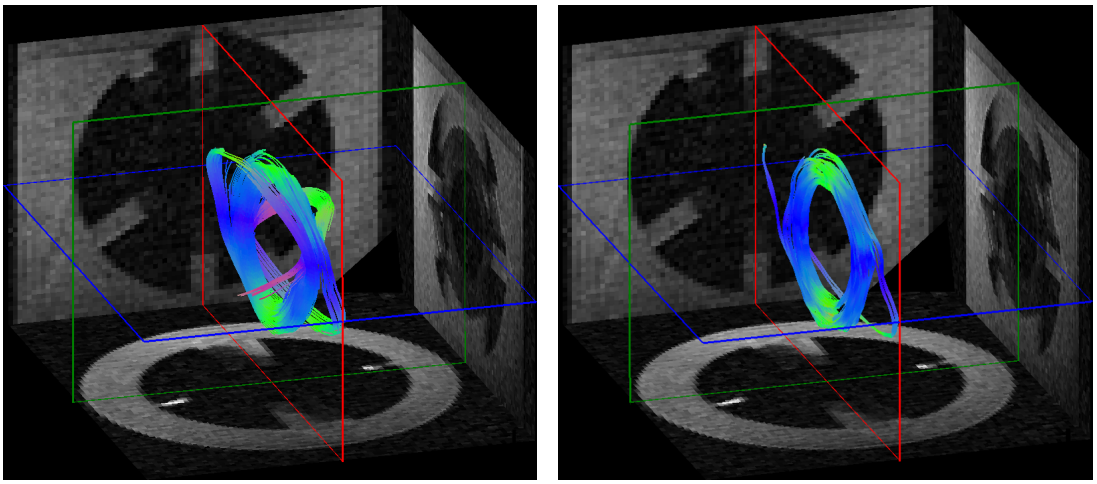


Figure 4.15.: Probabilistic tractography results using the proposed algorithm. *First row*: seeding in the corpus callosum, *second row*: seeding in the pyramidal tracts, *third row*: seeding in the cingulum bundles. The seed regions are the same as in Figure 4.12. The colored GFA maps in the second column show on a logarithmic scale the number of track visits per voxel summed up along the 20 slices highlighted in red on the right side, respectively. A projection of the seed region from these 20 slices is highlighted in yellow.



(a) seedregion **(b)** DT-based **(c)** proposed

Figure 4.16.: Tractography results of the crossing region after the first pass through.



(a) DT-based **(b)** proposed

Figure 4.17.: 3D view of tractography results for maxIter = 500.

at the 30° crossing, which is mainly due to the improperly resolved 30° crossing of the ODF reconstruction discussed in Section 3.2.2.

4.4. Discussion

Qualitative, Quantitative Analysis and Comparison

We compared our algorithm to a purely local streamline method, and to SD_-Stream with 4-th order Runge-Kutta integration and IFOD-2 from MRtrix.

Qualitatively, the proposed deterministic and probabilistic methods show a clear advantage with respect to the number of valid tracks. For quantitative evaluation, we regarded the Tractometer values introduced by Côté et al. (2013). Here, our method has the highest percentage of *valid* and the lowest percentage of *no connections* for experiments performed on *Fiber Cup*, ISBI phantom, and ISMRM challenge data. Since almost no tracks terminate before reaching the gray matter regions defined for the phantoms, the method shows also a certain increase of invalid tracks. That number can be reduced by introducing a cut-off threshold, as described below in the discussion part about stopping criteria. Moreover, we show qualitative results for *in vivo* data obtained from the HCP for three different bundles. Compared to the purely local streamline method, results benefit from the use of the proposed algorithm. For instance, the fiber tracks tracing the pyramidal tracts fan out more broadly towards the cortex and also in the tracking results of the corpus callosum, laterality is more pronounced. Moreover, tracking of the cingulum bundles involving difficult crossings as well as narrow curving structures yields reasonable results. A limitation is, that *in vivo* the parameter σ (standard deviation) that regulates the narrowness of the truncated Gaussian transition probabilities forming the prior distribution may have to be adapted. However, promising quantitative results for the ISMRM 2015 tractography challenge data show that the proposed algorithm also performs well on data simulating realistic WM fiber bundles with constant parameter σ.

Parameter Choice

Table 4.6 summarizes the parameters of the proposed algorithm and suggests ranges for n, μ, σ, φ, N, and β for a fixed step length of $\lambda = 0.5 \times$ voxel-size. Parameter φ defines the maximum deviation angle of the candidate directions from the former direction. Decreasing φ results in smooth tracks, where increasing it promotes structures of high curvature. We suggest to choose φ between 18° and 30°. The forward search step length μ has been introduced to the algorithm to decouple forward search from actual step length λ for more flexibility. The smaller the value for μ is chosen, the more forward search steps n are necessary for a good result. Since computation time increases with n, the number of forward search steps, we recommend to use only a subset of most promising candidate directions, or a smaller angle φ, when selecting $n > 3$. On the other hand, setting μ larger than one voxel-size leads to inaccuracies due to skipping the immediate neighborhood. Therefore, we suggest to choose $\mu \in [\lambda, \text{voxel-size}]$. Regarding extrapolation, we suggest to choose the number of previous points included in the curve fitting process between 3 and 10 for a step length of $0.5 \times$ voxel-size. Parameter σ controls the shape of the Gaussian prior distribution. Increasing σ increases the influence of the guiding direction but reduces the influence of the

fiber ODF information on the overall tracking direction. It has to be adapted according to data set and ODF reconstruction to find a trade-off between promoting straight and curved structures.

Stopping Criteria

Regarding the majority of results shown in the previous sections, the proposed algorithm is terminated when it encounters a region which is not supported by a predefined binary mask. However, it is also possible to define a threshold on the a-posteriori probability to terminate or even reject a track that is not supported by the local neighborhood. For the simulated data set, we showed that stopping the algorithm if the a-posteriori probability at the current path point fell below a certain value, could decrease the amount of invalid connections.

Including Anatomical Priors

We showed that anatomical information of binary or tissue partial volume maps can be included in the Bayesian framework to inform tractography, similar to the particle filtering algorithm presented by Girard et al. (2014). Especially for the phantom data, we observed that our proposed algorithm performed better with respect to the number of valid connections when information about the location of modeled CSF and white matter was incorporated into the prior distribution. For the *in vivo* data set, we showed first results of a straightforward approach to inform our algorithm by partial volume maps.

4.5. Related Work

Previously, different approaches for neighborhood informed tractography have been introduced, e.g. in Savadjiev et al. (2008); Tournier et al. (2010); Dhollander et al. (2014); Smith et al. (2012); Zhang et al. (2009); Pontabry and Rousseau (2011); Rowe et al. (2013); Girard et al. (2014). Zhang et al. (2009); Pontabry and Rousseau (2011); Rowe et al. (2013); Girard et al. (2014) combine probabilistic streamline tracking with a particle filter (sequential Monte Carlo methods). There, a finite number of particles is propagated from a seed point by drawing samples from an importance function and assigning weights to the evolving path according to a recursively defined posterior distribution. The methods

differ in the diffusion model that contributes to the posterior distribution as observation density. For instance, instead of using a single tensor diffusion model as in Zhang et al. (2009), Pontabry and Rousseau (2011) extend the method incorporating the Q-ball diffusion model. Subsequently, Zhang et al. (2009); Pontabry and Rousseau (2011); Girard et al. (2014) perform a resampling step where particles with low weights are eliminated and particles with high weights are emphasized. In contrast, the method described by Rowe et al. (2013) uses particle filtering only over a short distance in each streamline iteration step to explore the neighborhood and select the new propagation direction accordingly. Particle filtering methods are similar to our approach in the sense of a forward exploration scheme derived from a stochastic framework, but differ in the underlying Bayesian model. Yet another similar method with respect to forward exploration, though not based upon a Bayesian approach, is the second order integration probabilistic streamline method proposed in Tournier et al. (2010). Here, candidate paths originating from the current point are chosen as circular arcs tangent to the current tracking direction and evaluated based on a probability density function. Another difference compared to our approach is that these two methods only generate probabilistic tracks while we derive both a deterministic and a probabilistic algorithm. Track orientation distribution (TOD) based tractography recently introduced by Dhollander et al. (2014) can be seen as another related approach. The TOD, a function on the 5d spatio-angular domain, indicates the existence of fiber tracks along certain orientations. Since TODs are obtained from continuity constrained short-tracks tractograms, TOD-based tractography is informed by the track-like structure of the local neighborhood leading to more coherent tractography results. A different concept of neighborhood informed tractography is described by Savadjiev et al. (2008). There, an ODF-based labeling scheme is presented assigning to each voxel one of the four associated tract configurations *cross*, *fan*, *single*, or *unknown*. To distinguish between fanning and single fiber tract, a differential geometry framework is used that approximates tracts in each voxel by helical curves supported by a local neighborhood. A deterministic tracking algorithm is then applied based upon this advance information. These latter two methods have in common that, compared to Rowe et al. (2013); Tournier et al. (2010) and the algorithm proposed in this thesis, neighborhood exploration is carried out in a preprocessing step before the actual tractography algorithm is applied. Beyond using only the information provided by the diffusion signal, Smith et al. (2012) and Girard et al. (2014) propose methods for anatomically-informed tractography. For instance, Smith et al. (2012) introduce a back-tracking strategy which cuts and re-tracks streamlines that are not sufficiently supported by anatomical image data. Girard et al. (2014) extend this idea including tissue partial volume information derived from a T1-weighted image in the weighting process of the particle filtering algorithm.

bundle		$\mu = 0.2$	$\mu = 1$	$\mu = 1.5$	$\sigma = \frac{\pi}{4}$	$\sigma = \frac{\pi}{2}$	$\sigma = \pi$
3-(1,10,12)	VC	94.25	97.50	98.00	97.25	97.75	97.50
	IC	0.00	0.50	1.50	0.50	0.50	0.50
	NC	5.75	2.00	0.50	2.25	1.75	2.00
	VCCR	100.00	99.49	98.49	99.49	99.49	99.49
	CSR	94.25	98.00	99.50	97.75	98.25	98.00
2-7	VC	90.50	100.00	100.00	100.00	100.00	100.00
	IC	6.00	0.00	0.00	0.00	0.00	0.00
	NC	3.50	0.00	0.00	0.00	0.00	0.00
	VCCR	93.78	100.00	100.00	100.00	100.00	100.00
	CSR	96.50	100.00	100.00	100.00	100.00	100.00
4-5	VC	14.00	100.00	100.00	99.50	100.00	100.00
	IC	0.00	0.00	0.00	0.00	0.00	0.00
	NC	86.00	0.00	0.00	0.50	0.00	0.00
	VCCR	100.00	100.00	100.00	100.00	100.00	100.00
	CSR	14.00	100.00	100.00	100.00	99.50	100.00
6-11	VC	86.50	61.00	0.00	72.00	70.50	61.00
	IC	5.50	32.50	95.50	22.00	25.50	32.50
	NC	8.00	6.50	4.50	6.00	4.00	6.50
	VCCR	94.02	65.24	0.00	76.60	73.44	65.24
	CSR	92.00	93.50	95.50	94.00	96.00	93.50
8-9	VC	12.50	87.50	61.00	86.00	85.50	87.50
	IC	9.00	4.00	0.00	5.00	4.00	4.00
	NC	78.50	8.50	39.00	9.00	10.50	8.50
	VCCR	58.14	95.63	100.00	94.51	95.53	95.63
	CSR	21.50	91.50	61.00	91.00	89.50	91.50
all	VC	65.33	90.58	76.17	92.00	91.92	90.58
	IC	3.42	6.25	16.42	4.67	5.08	6.25
	NC	31.25	3.17	7.42	3.33	3.00	3.17
	VCCR	95.03	93.55	82.27	95.17	94.76	93.55
	CSR	68.75	96.83	92.58	96.67	97.00	96.83

Table 4.3.: Tractometer values for different choices of μ and σ. All data are given in percent. The bundles are defined by connected end regions numbered as in Figure 4.6c.

	deterministic			probabilistic			
	basic	proposed	SD_Stream -rk4	basic	proposed	cutoff	IFOD-2
SNR 30							
VC	71.65	**78.85**	76.05	49.30	**68.70**	42.85	51.10
IC	14.20	16.05	12.47	19.75	22.75	**13.75**	21.85
NC	14.15	**5.10**	11.47	30.95	**8.50**	43.40	27.05
VCCR	83.46	83.09	86.89	71.40	75.12	**75.71**	70.05
CSR	85.85	**94.90**	87.53	69.05	**91.50**	56.60	72.95
SNR 10							
VC	55.50	**66.95**	52.08	36.05	**51.70**	35.23	29.70
IC	29.80	27.70	19.57	28.70	39.70	**21.79**	23.00
NC	14.70	**4.95**	28.35	35.25	**8.20**	42.88	47.20
VCCR	65.06	70.73	72.69	55.68	56.56	**61.79**	56.36
CSR	85.30	**95.03**	71.65	64.75	**91.77**	57.05	52.75

Table 4.4.: Tractometer values comparing deterministic and probabilistic algorithms on the ISBI phantom. Fiber ODFs were derived with standard CSD using MRtrix. All data are given in percent.

	deterministic			probabilistic		
	basic	proposed	SD_Stream -rk4	basic	proposed	IFOD-2
VC	56.83	**80.03**	61.83	1.19	**21.86**	20.92
IC	13.17	13.19	13.16	7.53	**16.75**	17.44
NC	30.00	**6.78**	25.01	91.28	61.39	61.65
VCCR	81.18	**85.85**	82.45	13.66	**56.61**	54.54
CSR	59.31	**79.60**	67.62	4.76	28.30	34.93
VB	24	24	24	11	23	23
IB	423	348	442	253	501	532

Table 4.5.: Tractometer values comparing deterministic and probabilistic algorithms on the ISMRM 2015 tractography challenge data. Fiber ODFs were derived with standard CSD using MRtrix. The numbers for VC, IC, NC, VCCR, and CSR are given in percent. VB and IB denote the number of detected valid and invalid bundles, respectively.

Param.	Function	Range
n	number of forward search steps	$0 < n \leq 3$ (larger where appropriate)
μ	forward search step size	$\mu \in [\lambda, \text{voxel-size}]$
σ	standard deviation of Gaussian function of prior distribution	$\sigma \in]0, 2\pi]$
φ	maximum deviation of candidate directions from former direction	$\varphi \in [\frac{\pi}{10}, \frac{\pi}{6}]$
N	number of points used to determine guiding direction	$3 \leq N \leq 10$
β	smoothing parameter	$0 < \beta < 1$

Table 4.6.: Suggested parameter ranges for $\lambda = 0.5 \times$ voxel-size.

5. Microstructure-Informed Global Tractography

Previously, we have described and developed tractography strategies that propagate streamlines based on diffusion information of the local neighborhood. In this chapter, we focus on the global support of a tractogram by the underlying DW-MRI data. The methods described in Daducci et al. (2013, 2015) and Pestilli et al. (2014) seek to model the diffusion MR signal from a large set of candidate fibers obtained from an arbitrary tractography algorithm. In an optimization procedure, weights to the candidate fibers are adjusted, such that the modeled signal best fits the measured one. By this means, these methods can serve to quantify the quality of a tractogram, as well as filter out tracks that do not well support the diffusion data. In Sections 5.1 and 5.2, we summarize the approach described in Daducci et al. (2015), where the DW-MR signal is modeled by three compartments to account for tissue microstructure, and the model fitting problem is expressed as a convex ℓ_1-norm minimization problem. By the three compartment model, the signal is modeled by an intra-axonal, an extra-axonal, and an isotropic part. Only the intra-axonal part is described by the tractogram, but the other two parts are related to the voxels in the image domain where spatial regularization would be more appropriate. In Section 5.3, we expand on the approach stated in Daducci et al. (2015) by formulating the problem as a Tikhonov-type regularization problem where fiber weights are regularized by ℓ_1, and voxel weights by a H^s norm to promote spatial smoothness. Moreover, we outline the theory of the alternating direction method of multipliers (ADMM) in Section 5.4, and use it as a tool to efficiently solve our large scale optimization problem in Section 5.5. In Section 5.6, we propose strategies to further reduce computation time of the algorithm presented in Section 5.5. Experiments and results are discussed in Section 5.7.

5.1. Modeling the DW-MR Signal

Based upon the *StickZeppelinBall* compartment model described in Section 2.8, and using the notions of Tournier et al. (2004), Daducci et al. (2015) represent the diffusion signal in white matter as a combination of restricted, hindered, and isotropic contributions. Before summarizing the overall signal model, we first describe the three individual signal parts in detail. Let $R : \mathcal{S}^2 \to \mathbb{R}$ denote the response function describing the signal attenuation profile for a fiber or fiber bundle, as stated in Section 2.5.1. Furthermore, let the set of candidate fibers be defined by $\mathcal{F} := \left\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{n_f}\right\}$ with $\mathcal{F}_i \in \mathbb{R}^{n_i \times 3}$ as introduced in (2.14), and the set of fiber populations in a voxel, for instance obtained as the principle diffusion directions from the fiber ODF, by $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{n_b}\} \subseteq \mathcal{S}^2$. For reasons of clarity, the modeled signal is in this section considered only as a function of $q \in \mathcal{S}^2$ and for a fixed voxel. In the next section, we state an expression for the signal depending on both direction $q \in \mathcal{S}^2$ and voxel $v \in \mathbb{R}^3$.

The predicted intra-axonal diffusion signal S_{IC} in each voxel consists of the signal contributions of the candidate fibers. In gradient direction $q \in \mathcal{S}^2$, it is defined by

$$S_{\mathrm{IC}}(q) := \sum_{\mathcal{F}_i \in \mathcal{F}} \ell_i f_i^{\mathrm{IC}} R_i^{\mathrm{IC}}(q), \tag{5.1}$$

where $\ell_i \geq 0$ is the length of the fiber segment and $f_i^{\mathrm{IC}} \geq 0$ denotes the weight or volume fraction, both corresponding to fiber $\mathcal{F}_i$ in the respective voxel. The response function R_i^{IC} for fiber $\mathcal{F}_i$ is defined as in Equation (2.20) with principle diffusion direction v_1 aligned with the local fiber orientation of $\mathcal{F}_i$. For fibers $\mathcal{F}_i$ which do not cross the voxel, the signal contribution is zero. Similar to equation (5.1), extra-axonal diffusion in a voxel is modeled by

$$S_{\mathrm{EC}}(q) := \sum_{\mathcal{B}_j \in \mathcal{B}} f_j^{\mathrm{EC}} R_j^{\mathrm{EC}}(q), \tag{5.2}$$

with weights $f_j^{\mathrm{EC}} \geq 0$ and where the rotated response function R_j^{EC} is defined as in Equation (2.22) with v_1 matching the direction of the respective principle diffusion directions $\mathcal{B}_j$ in that voxel. Finally, the predicted isotropic signal

$$S_{\mathrm{ISO}}(q) := f^{\mathrm{ISO}} R^{\mathrm{ISO}}(q) \tag{5.3}$$

is modeled by the isotropic response function R^{ISO} defined in Equation (2.23) and scaled by $f^{\mathrm{ISO}} \geq 0$. Overall, the voxel-wise diffusion signal attenuation in gradient direction $q \in \mathcal{S}^2$ is given by the sum of the signals defined in Equations

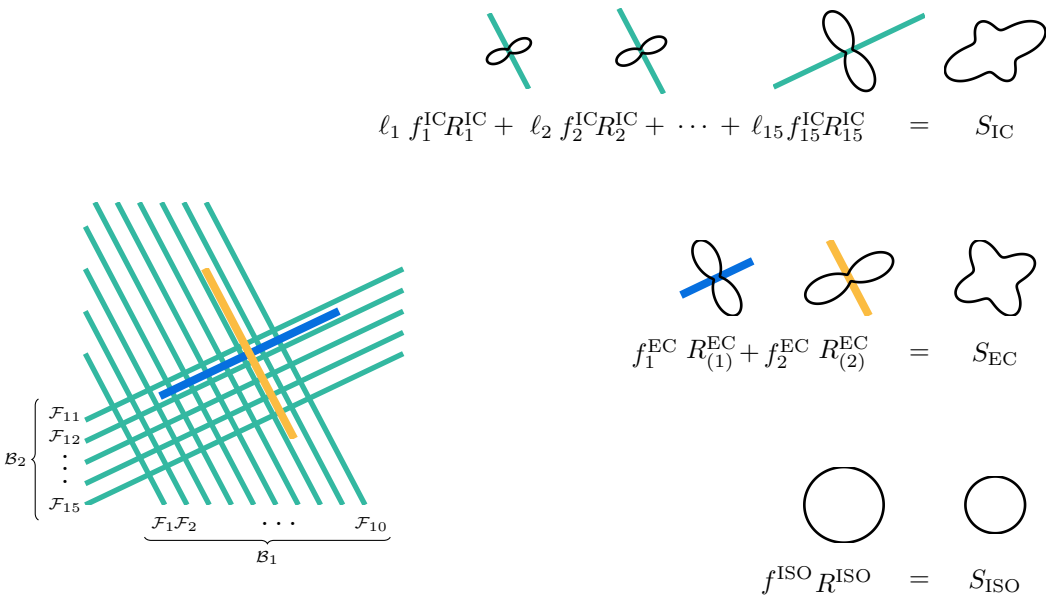


Figure 5.1.: Two-dimensional exemplary illustration of the modeled diffusion signal. *Left:* A pixel containing two crossing fiber bundles. The corresponding principle diffusion directions are delineated in orange and blue, respectively. *Right:* Intra-axonal, extra-axonal, and isotropic signal components as linear combinations of response functions.

(5.1),(5.2), and (5.3), that is

$$\frac{\tilde{S}(q)}{\tilde{S}_0} = S_{\mathrm{IC}}(q) + S_{\mathrm{EC}}(q) + S_{\mathrm{ISO}}(q) \tag{5.4}$$

with diffusion-weighted signal $\tilde{S} : \mathcal{S}^2 \to \mathbb{R}$ and unweighted signal $\tilde{S}_0$. Figure 5.1 illustrates the composition of the modeled diffusion signal by means of a two-dimensional example of two crossing fiber bundles in a pixel.

5.2. Global Optimization Problem

Let the measured DW-MR signal be given by $S : \mathcal{S}^2 \times \Omega \to \mathbb{R}$, as defined in Chapter 2, Equation (2.3). Extending the expression for the predicted voxel-wise signal in (5.4), the modeled signal attenuation for all voxels $v \in \Omega$ and gradient directions $q \in \mathcal{S}^2$ can be written as

$$\frac{\tilde{S}(q,v)}{\tilde{S}_0(v)} = \sum_{\mathcal{F}_i \in \mathcal{F}} \ell_i(v) f_i^{\mathrm{IC}} R_i^{\mathrm{IC}}(q,v) + \sum_{\mathcal{B}_j \in \mathcal{B}(v)} f_j^{\mathrm{EC}}(v) R_j^{\mathrm{EC}}(q,v) + f^{\mathrm{ISO}}(v) R^{\mathrm{ISO}}(q,v) \tag{5.5}$$

with response functions $R : \mathcal{S}^2 \times \Omega \to \mathbb{R}$, and where $\ell_i(v)$ is the length of fiber $\mathcal{F}_i$ in voxel v, $f_j^{\mathrm{EC}}(v)$ is the weight corresponding to fiber population $\mathcal{B}_j$ in voxel v, and $\mathcal{B}(v)$ denotes the set of principle diffusion directions in voxel v. Given ℓ_i, R_i^{IC}, R_j^{EC}, and R^{ISO}, the aim is to find weights f_i^{IC}, f_j^{EC}, and, f^{ISO}, such that the predicted signal best fits the measured signal S. That is, we want to solve

$$\underset{f_i^{\mathrm{IC}}, f_j^{\mathrm{EC}}, f^{\mathrm{ISO}} \geq 0}{\operatorname{argmin}} \mathcal{D}\left(\tilde{S}\left(q, v; f_i^{\mathrm{IC}}, f_j^{\mathrm{EC}}, f^{\mathrm{ISO}}\right) - S(q, v)\right) + \alpha \mathcal{R}\left(f_i^{\mathrm{IC}}, f_j^{\mathrm{EC}}, f^{\mathrm{ISO}}\right)$$

with data fidelity function $\mathcal{D}$, and possibly a regularization term $\mathcal{R}$ where $\alpha \geq 0$. Note that, while the f_i^{IC} weight the individual candidate fibers, each crossing several voxels, the weights f_j^{EC} and f^{ISO} correspond to the individual voxels, respectively.

5.2.1. Discretization and Matrix-Vector Notation

Let n_d denote the number of gradient directions, n_v the number of voxels, n_f the number of candidate fibers in $\mathcal{F}$, and n_b the predefined maximum number of considered bundles in $\mathcal{B}(v)$. The representation of the estimated signal (5.5) can be formulated as matrix-vector product

$$[A_{\mathrm{IC}}|A_{\mathrm{EC}}|A_{\mathrm{ISO}}]x = Ax = y_{\mathrm{est}}$$

with the block matrix A consisting of $A_{\mathrm{IC}} \in \mathbb{R}^{n_d \cdot n_v \times n_f}$, $A_{\mathrm{EC}} \in \mathbb{R}^{n_d \cdot n_v \times n_b \cdot n_v}$, and $A_{\mathrm{ISO}} \in \mathbb{R}^{n_d \cdot n_v \times n_v}$, containing for each voxel and each gradient direction the values of $\ell_i R_i^{\mathrm{IC}}$, R_j^{EC}, and R^{ISO}, respectively. The weights f_i^{IC}, f_j^{EC}, and, f^{ISO} are stored in vector

$$x = [x^{\mathrm{IC}}|x^{\mathrm{EC}}|x^{\mathrm{ISO}}] \in \mathbb{R}^{n_f + n_b \cdot n_v + n_v}$$

where

$$\begin{aligned}
x^{\mathrm{IC}} &= \left(f_{(1)}^{\mathrm{IC}}, f_{(2)}^{\mathrm{IC}}, \ldots, f_{(n_f)}^{\mathrm{IC}}\right)^\top \\
x^{\mathrm{EC}} &= \left(x_1^{\mathrm{EC}}(1), \ldots, x_{n_b}^{\mathrm{EC}}(1), \ldots, x_1^{\mathrm{EC}}(n_v), \ldots, x_{n_b}^{\mathrm{EC}}(n_v)\right)^\top \\
&\quad \text{with } x_j^{\mathrm{EC}} = \left(f_{(j)}^{\mathrm{EC}}(v_1), f_{(j)}^{\mathrm{EC}}(v_2), \ldots, f_{(j)}^{\mathrm{EC}}(v_{n_v})\right)^\top, \quad j = 1, \ldots, n_b \\
x^{\mathrm{ISO}} &= \left(f^{\mathrm{ISO}}(v_1), f^{\mathrm{ISO}}(v_2), \ldots, f^{\mathrm{ISO}}(v_{n_v})\right)^\top.
\end{aligned} \tag{5.6}$$

Figure 5.2 illustrates the structure of matrix A and corresponding vectors x and y_{est}. In accordance with modeled signal y_{est}, let y denote the measured DW-MRI

signal in vector notation.

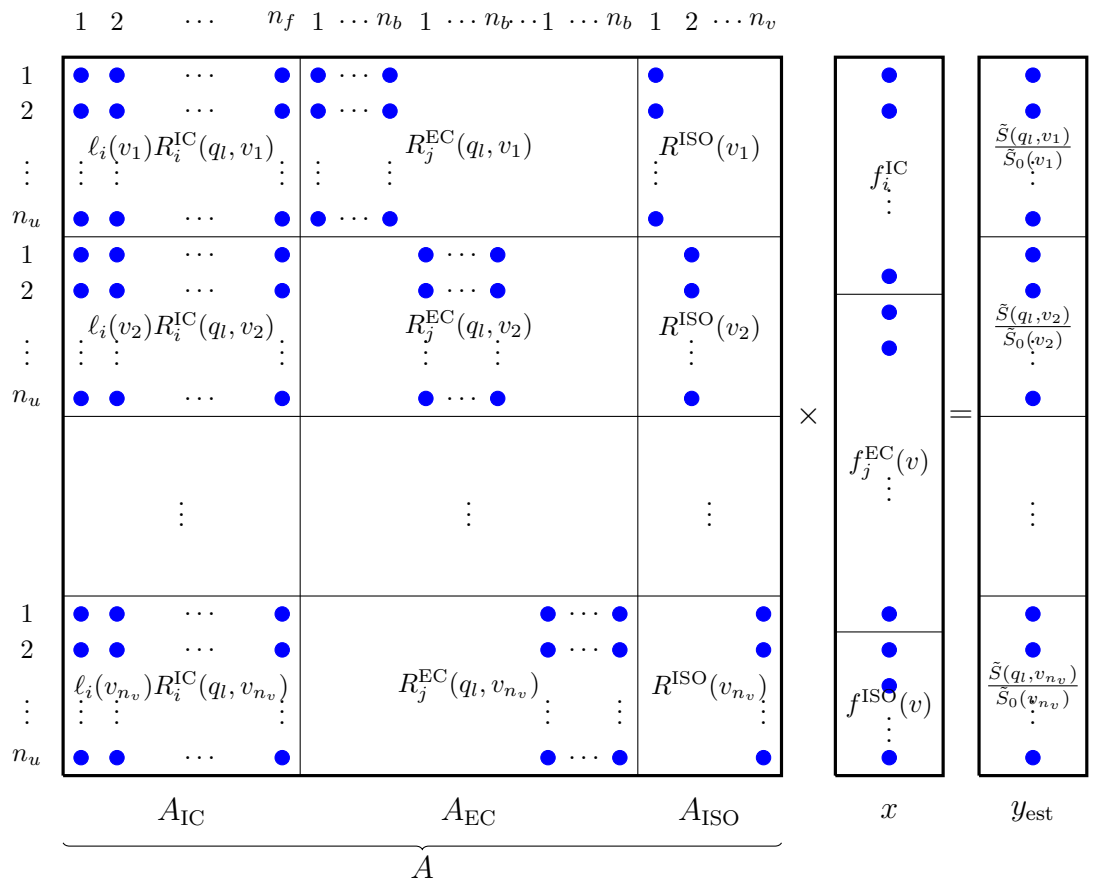


Figure 5.2.: Structure of matrix A, where $i = 1, 2, \ldots, n_f$ is the fiber index, $j = 1, 2, \ldots, n_b$ is the bundle index, and $l = 1, 2, \ldots, n_u$ is the gradient direction index.

5.2.2. Convex Optimization Problem Formulation

In Daducci et al. (2015), two formulations of the model fitting problem, *non-negative least-squares* (NNLS)

$$\underset{x \geq 0}{\operatorname{argmin}} \|Ax - y\|_2^2 \tag{5.7}$$

and *basis pursuit de-noise* (BPDN)

$$\underset{x \geq 0}{\operatorname{argmin}} \left\{ \|x\|_1 \text{ subject to } \|Ax - y\|_2 \leq \epsilon \right\}, \quad \epsilon \geq 0, \tag{5.8}$$

where $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$, are analyzed. The constrained minimization problem (5.8) is closely related to the penalized minimization problem

$$\underset{x \geq 0}{\operatorname{argmin}} \{\|Ax - y\|_2 + \alpha \|x\|_1\} \tag{5.9}$$

with $\alpha > 0$. In fact, the unconstrained problem formulation (5.9) can be regarded as the problem of minimizing the Lagrange function

$$\mathcal{L}(x, \mu) = \|x\|_1 + \lambda \left(\|Ax - y\|_2 - \epsilon\right), \quad \lambda > 0,$$

of constrained problem (5.8) over $x \in \mathbb{R}^n_{\geq 0}$ and shifting the (modified) Lagrange multiplier in front of the ℓ_1-norm term ($\alpha = 1/\lambda$). Under certain assumptions, a minimizer $x^* \in \mathbb{R}^n_{\geq 0}$ of (5.8) also solves (5.9) and vice versa (see for example Hiriart-Urruty and Lemarechal (1993, VII), Ciak et al. (2013)).

These problem formulations either do not include regularization (NNLS), or impose ℓ_1 regularization upon all weights (BPDN). While ℓ_1 regularization is useful with respect to fiber weighting coefficients x^{IC}, for instance to filter out anatomically implausible tracts of a tractogram, a regularization term that enforces spatial smoothness would be more suitable regarding the voxel weighting coefficients x_j^{EC} and x^{ISO}. In the next section, we take this into account and formulate a new approach by adding a Sobolev norm term to the optimization problem to promote smoothness of x_j^{EC} and x^{ISO}.

5.3. Sobolev Norm Regularization

In this section, we define the Sobolev space H^s and consider different representations of the associated norm $\|\cdot\|_{H^s}$.

Definition 5.1. For $s \geq 0$, the Sobolev space $H^s(\mathbb{R}^n)$ is defined by

$$H^s(\mathbb{R}^n) := \left\{ f \in L^2(\mathbb{R}^n) : \left(1 + |\xi|^2\right)^{s/2} (\mathscr{F} f)(\xi) \in L^2(\mathbb{R}^n) \right\}, \quad \xi \in \mathbb{R}^n$$

with the norm

$$\|f\|_{H^s(\mathbb{R}^n)} := \left\| \mathscr{F}^{-1} \left[\left(1 + |\xi|^2\right)^{s/2} \mathscr{F} f \right] \right\|_{L^2(\mathbb{R}^n)}$$

where $\mathscr{F}$ denotes the n-dimensional Fourier transform.

The H^s-norm can also be expressed in terms of the Bessel-potential operator

$(I - \Delta)^s$, which is given by

$$(I - \Delta)^s f = \mathscr{F}^{-1}\left[\left(1 + |\xi|^2\right)^{s/2} \mathscr{F} f\right],$$

where Δ describes the Laplace operator. Regarding our application, we consider a function $f : \Omega \to \mathbb{R}$ assigning an intensity to each point in a cuboid domain $\Omega = [1, n_x] \times [1, n_y] \times [1, n_z] \subset \mathbb{R}^3$. In three dimensions, the Laplace operator of f in Cartesian coordinates is given by

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2},$$

that is, the sum of (unmixed) second-order partial derivatives. Since the intensities are only given at discrete grid points (voxels), we deduce a discretized expression for the H^s norm below.

5.3.1. Discretization of the Laplacian

In the following, we summarize the discretization of the Laplacian by finite differences, which is well documented, for instance in the book by Hanke-Bourgeois (2009, XV,83). First, we consider the one-dimensional case where f is given on an equidistant grid

$$\{x_i = ih : i = 1, \ldots, n_x\}$$

with $h > 0$. We approximate the second derivative of f at point x_i by the central difference quotient

$$D_h^2 [f] (x_i) = \frac{1}{h^2} (f(x_i - h) - 2f(x_i) + f(x_i + h)) \approx f''(x_i), \tag{5.10}$$

for $i = 2, \ldots, n_x - 1$. Using Neumann boundary conditions and approximating the first derivatives of f at x_1 and x_{n_x} by forward and backward differences, respectively, we obtain

$$\begin{aligned} f''(x_1) &\approx \frac{1}{h^2} (f(x_2) - f(x_1)) \\ f''(x_{n_x}) &\approx \frac{1}{h^2} (f(x_{n_x - 1}) - f(x_{n_x})) \end{aligned} \tag{5.11}$$

for the second derivative at the boundaries. Rewriting (5.10) and (5.11) as matrix-vector multiplication, we obtain the well-known formulation of the one-

dimensional discrete Laplace operator equation

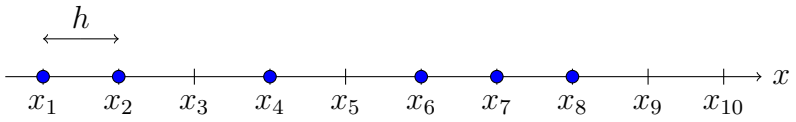
$$\mathbf{f}'' \approx \Delta^{1\mathrm{D}}\mathbf{f} \quad \text{with} \quad \Delta^{1\mathrm{D}} = \frac{1}{h^2}\begin{pmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{pmatrix} \in \mathbb{R}^{n\times n} \tag{5.12}$$

with

$$\mathbf{f} := (f(x_1), \ldots, f(x_{n_x}))^\top \quad \text{and} \quad \mathbf{f}'' := (f''(x_1), \ldots, f''(x_{n_x}))^\top. \tag{5.13}$$

The image data we want to apply the Laplacian to is, as a whole, given on a three dimensional equidistant grid in the domain Ω. However, since the problem size of our optimization problem can be dramatically reduced by only including voxels where fiber tracks are passing through, we need a modified version of (5.12) where the Laplacian is applied to not necessarily connected subdomains of Ω. For this purpose, the number of neighboring points is determined for each grid point in the set of included points, and the rows of the linear Laplace operator are modified accordingly. In one dimension, there are only three different cases: (1) The grid point can have two neighbors, i.e. it is an inner points, (2) one neighbor, i.e. it is a point at the boundary, or (3) no neighbors, then it is a single-point domain and the approximative second derivative is zero. The following examples illustrates the structure of matrix $\Delta^{1\mathrm{D}}$.

Example 5.2. Consider f given at the following grid points, where only the points highlighted by blue dots are of interest.



The corresponding discrete Laplacian applied to $\mathbf{f} = (\, f(x_1),\, f(x_2),\, f(x_4),\, f(x_6),\, f(x_7),\, f(x_8)\,)^\top$ is the 6×6 matrix

$$\Delta^{1\mathrm{D}} = \frac{1}{h^2}\begin{pmatrix} -1 & 1 & & & & \\ 1 & -1 & & & & \\ & & 0 & & & \\ & & & -1 & 1 & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -1 \end{pmatrix}$$

(a) grad $f^\top n = 0$ on $\partial\Omega$ **(b)** $\frac{\partial f}{\partial y}(x, y + \frac{h}{2}) \approx \frac{f(x,y)-f(x,y+h)}{h} = 0$

Figure 5.3.: Exemplary illustration of Neumann zero boundary conditions. *Left:* Smooth boundary. *Right:* Approximation used here.

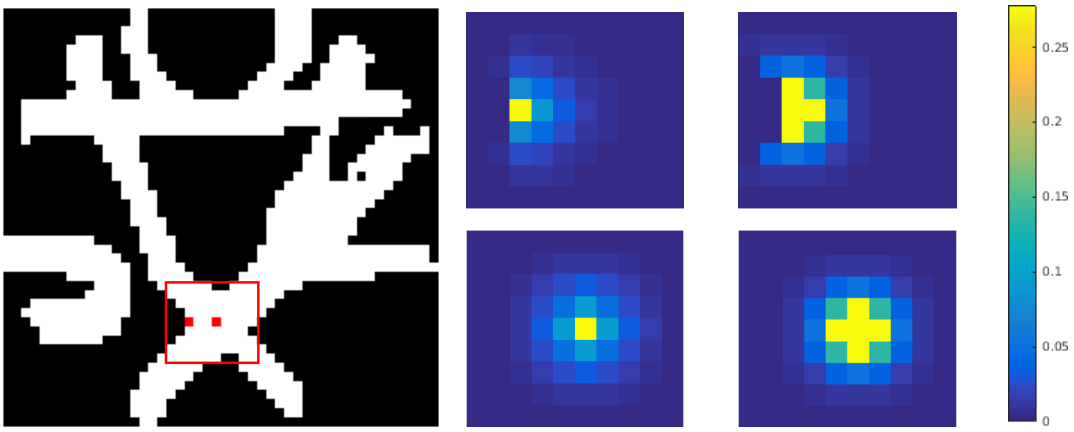
This approach can easily be extended to 2 or 3 dimensions by consecutively assigning linear indices to the grid points of the subregions of Ω that are of interest. Let N denote the overall number of included grid points. Then $\mathbf{f} \in \mathbb{R}^N$ and $\mathbf{f}'' \in \mathbb{R}^N$ contain the respective function values arranged as in (5.13). Assuming that grid width h is the same in all dimensions, the matrices $\Delta^{\mathrm{2D}}, \Delta^{\mathrm{3D}} \in \mathbb{R}^{N\times N}$ describing the 2D and 3D Laplace operator, respectively, are structured as follows: Let $\mathcal{N}_k \subset \{1, 2, \ldots, N\}$ define the set of indices corresponding to neighboring grid points of point k. Then, row k of $\Delta^{\{\mathrm{2D,3D}\}}$, which multiplied to $\mathbf{f}$ yields element $\mathbf{f}''(k)$, is given by

$$\Delta^{\{\mathrm{2D,3D}\}}(k, j) = \frac{1}{h^2} \cdot \begin{cases} -|\mathcal{N}_k| & \text{if } j = k \\ 1 & \text{if } j \in \mathcal{N}_k, \\ 0 & \text{else,} \end{cases}$$

that is, each row of Δ^{2D} contains up to 5 entries different from zero, each row of Δ^{3D} up to 7, and the row entries sum up to zero.

Implementation Details

In general, imposing Neumann zero boundary conditions implies that the derivative of a function f given on a domain Ω vanishes on the boundary $\partial\Omega$ in direction of the exterior normal n, as illustrated in Figure 5.3a. Note that for our application the domain Ω is discretized on a uniform grid. Hence, the boundary $\partial\Omega$ is not smooth, as in Figure 5.3a, but piecewise constant, as exemplarily shown in Figure 5.3b for the two-dimensional case, and we approximate the derivatives along the equidistant grid points. Our implementation of the discrete Laplacian is similar to MATLAB's `delsq` with an extension to three dimensions. Despite the relatively coarse approximation of the boundary itself and the first derivative



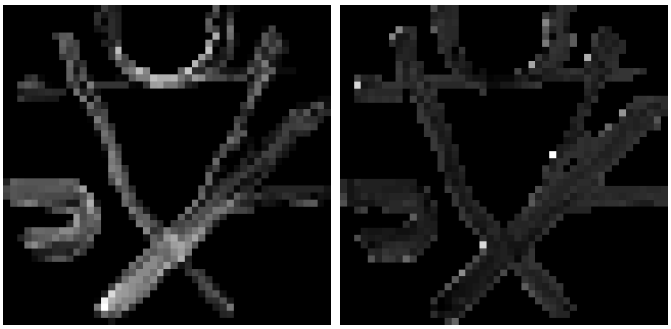
(a) Ω (white regions) **(b)** G derived from discretization. **(c)** G computed directly.

Figure 5.4.: Green's function G corresponding to differential operator $(\lambda^2 I - \Delta^{2D})$ for fixed $\lambda > 0$ evaluated at the two sample pixels highlighted in red.

on the boundary, we obtain good results with this approach. Green's functions from a partial differential equations perspective are briefly reviewed in Appendix B. In our context, from a Bayesian perspective, Green's function G is the kernel of the prior covariance operator, i.e. $G(x, y)$ indicates how much the values at points x and y are correlated in the chosen prior distribution. In Figure 5.4b, we show the Green's function (see Appendix B) computed from the operator $(\lambda^2 I - \Delta^{2D})$ discretized as described here for two sample points in the domain illustrated in Figure 5.4a. As a comparison, the free Green's function corresponding to $(\lambda^2 I - \Delta^{2D}) \in \mathbb{R}^2$ given by Bessel functions is shown in Figure 5.4c. In view of the above interpretation, the width of Green's function seems plausible. Moreover, as opposed to some other boundary conditions and implementations, the width of Green's function is not significantly influenced by the boundary, which is reasonable in our context.

Experimental Motivation for Using H^s-Norm Regularization

Consider the Fiber Cup phantom data where the fiber ground truth is known, as described in Section 3.2.1. Furthermore, let a tractogram with valid and invalid tracks be given for this data set. Correspondingly, $A = [A^{IC}|A^{ISO}]$ denotes the dictionary matrix where the extra-cellular part is omitted for simplicity, and whose structure is explained by Figure 5.2. To specify which of the tracks belongs to which category, we define index sets I_{VC} and I_{IC} containing the indices of valid and invalid connections, respectively. To acquire a data set with known optimal solution x_{true} for testing purposes, we use the following simulation strategy:



(a) $\sum_{\mathcal{F}_i \in \mathcal{F}} \ell_i(v) f_i^{\mathrm{IC}}$ **(b)** $f^{\mathrm{ISO}}(v)$

Figure 5.5.: Intracellular and isotropic contributions for the data set simulated from Fiber Cup data, shown for one slice in the $x - y$ plane.

1. Compute the weights for the valid connection tracks solving

$$\overline{x}^{\mathrm{IC}}(I_{\mathrm{VC}}) = \underset{x \geq 0}{\operatorname{argmin}} \, \|A_{\mathrm{IC}}(:, I_{\mathrm{VC}})x - y\|_2^2.$$

2. Set the weights for the invalid tracks to zero: $\overline{x}^{\mathrm{IC}}(I_{\mathrm{IC}}) = 0$.

3. Note that we want that $Ax = A_{\mathrm{IC}}x^{\mathrm{IC}} + A_{\mathrm{ISO}}x^{\mathrm{ISO}} = y$, and compute the weights of the isotropic contributions by solving

$$\overline{x}^{\mathrm{ISO}} = \underset{x \geq 0}{\operatorname{argmin}} \, \|A_{\mathrm{ISO}}x - b\|_2^2,$$

   where $b = y - A_{\mathrm{IC}}x^{\mathrm{IC}}$.

4. Obtain the simulated signal $\tilde{y} = Ax_{\mathrm{true}} + \epsilon$ where $x_{\mathrm{true}} = [\overline{x}^{\mathrm{IC}}, \overline{x}^{\mathrm{ISO}}]$ and ϵ denotes Gaussian white noise.

As can be seen in Figure 5.5, the isotropic weights obtained by this procedure appear smooth apart from a few outliers at the borders of the image domain. Thus, regularization promoting spatial smoothness of the voxel weights appears plausible.

5.3.2. Overall Problem Formulation

Equipped with a discrete formulation of the Sobolev norm, we return to the optimization problem and associated notation introduced in Section 5.2. We split vector x into intracellular, extracellular and isotropic components as in (5.6) for

regularization, and introduce Sobolev norm terms for spatial smoothing, thus estimating x by solving

$$\underset{x\geq 0}{\operatorname{argmin}} \frac{1}{2}\left\|Ax-y\right\|_2^2 + \alpha\left\|x^{\mathrm{IC}}\right\|_1 + \sum_{j=1}^{n_b}\frac{\beta_j}{2}\left\|x_j^{\mathrm{EC}}\right\|_{H^s}^2 + \frac{\gamma}{2}\left\|x^{\mathrm{ISO}}\right\|_{H^s}^2 \tag{5.14}$$

where $\alpha, \beta_j, \gamma \geq 0$, and $\|\cdot\|_{H^s}$ denotes the Sobolev norm. Note that matrix A is ill-conditioned.

In the next section, we introduce the alternating direction method of multipliers (Boyd et al., 2011) as a strategy to solve problem (5.14).

5.4. Alternating Direction Method of Multipliers

In this section, we describe the theory of the *alternating direction method of multipliers* (ADMM), based upon the article by Boyd et al. (2011), in order to solve our minimization problem (5.14). Concepts from convex optimization needed in this section are summarized in Appendix C. The ADMM solves the equality-constrained optimization problem

$$\begin{aligned} &\text{minimize } f(x)+g(z) \\ &\text{subject to } M_1x+M_2z=c \end{aligned} \tag{5.15}$$

with $x\in\mathbb{R}^n$, $z\in\mathbb{R}^m$, $M_1\in\mathbb{R}^{p\times n}$, $M_2\in\mathbb{R}^{p\times m}$, $c\in\mathbb{R}^p$, and convex functions $f:\mathbb{R}^n\to\mathbb{R}$, $g:\mathbb{R}^m\to\mathbb{R}$.

Applying basic Lagrange duality theory, the associated Lagrangian function is given by

$$\mathcal{L}_0(x,z,\lambda) = f(x)+g(z)+\lambda^\top(M_1x+M_2z-c) \tag{5.16}$$

where $\lambda\in\mathbb{R}^p$ denotes the Lagrange multiplier vector, also called *dual variable*. The dual problem corresponding to (5.15) can be formulated as

$$\text{maximize } h(\lambda) \tag{5.17}$$

with Lagrange dual function $h:\mathbb{R}^p\to\mathbb{R}$ defined by

$$h(\lambda) = \inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m}\mathcal{L}_0(x,z,\lambda).$$

A point (x,z) is called *primal feasible* if it satisfies the constraints of the primal

problem (5.15). If a point $\lambda \in \mathrm{dom}(h)$, it is called *dual feasible*. Furthermore, a *primal optimal* point denotes a primal feasible point which minimizes problem (5.15), while a *dual optimal* point is dual feasible and maximizes (5.17). The dual optimization problem can be solved using dual gradient/subgradient ascent strategies of the form

$$\lambda^{k+1} = \lambda^k + \alpha^k \operatorname{grad}_\lambda h(\lambda^k)$$

with step length $\alpha^k > 0$, $k = 0, 1, 2, \ldots$. Since the Lagrangian is separable with respect to x and z, the gradient can be formulated as

$$\begin{aligned}
\operatorname{grad}_\lambda h(\lambda) &= \operatorname{grad}_\lambda \inf_{x,z} \mathcal{L}_0(x, z, \lambda) \\
&= \operatorname{grad}_\lambda \left(\inf_x \left\{ f(x) + \lambda^T M_1 x \right\} + \inf_z \left\{ g(z) + \lambda^\top M_2 z \right\} - \lambda^\top c \right) \\
&= \operatorname{grad}_\lambda \left(f(\overline{x}) + \lambda^\top M_1 \overline{x} + g(\overline{z}) + \lambda^\top M_2 \overline{z} - \lambda^\top c \right) \\
&= M_1 \overline{x} + M_2 \overline{z} - c
\end{aligned}$$

where

$$\begin{aligned}
\overline{x} &= \operatorname*{argmin}_x \left\{ f(x) + \lambda^\top M_1 x \right\} \\
\overline{z} &= \operatorname*{argmin}_z \left\{ g(z) + \lambda^\top M_2 z \right\}
\end{aligned}$$

resulting in the iteration steps

$$\begin{aligned}
x^{k+1} &= \operatorname*{argmin}_x \mathcal{L}_0 \left(x, z^k, \lambda^k \right) \\
z^{k+1} &= \operatorname*{argmin}_z \mathcal{L}_0 \left(x^{k+1}, z, \lambda^k \right) \\
\lambda^{k+1} &= \lambda^k + \alpha^k \left(M_1 x^{k+1} + M_2 z^{k+1} - c \right).
\end{aligned}$$

However, this algorithm converges only under very strong assumptions. For more robustness and better convergence properties, the *augmented* Lagrangian is used instead of (5.16), resulting in the ADMM algorithm.

The augmented Lagrangian function related to problem (5.15) is given by

$$\mathcal{L}_\rho(x, z, \lambda) := f(x) + g(z) + \lambda^\top (M_1 x + M_2 z - c) + \frac{\rho}{2} \|M_1 x + M_2 z - c\|_2^2$$

with a linear Lagrange term involving the dual variable $\lambda \in \mathbb{R}^p$, and a quadratic penalty term weighted by the penalty parameter $\rho > 0$. For $\rho = 0$, the augmented Lagrangian is equal to the standard Lagrangian stated in (5.16). Applying the dual ascent strategy from above to the modified problem with augmented La-

grangian yields the ADMM algorithm where variables x (by x-minimization of $\mathcal{L}_\rho$), z (by z-minimization of $\mathcal{L}_\rho$), and λ (by a dual update), are alternatingly updated as follows:

Algorithm 5.3. (ADMM)
Given $z^0 \in \mathbb{R}^m$, $\lambda^0 \in \mathbb{R}^p$, $\rho > 0$:
for $k = 0, 1, 2, \ldots$

$$x^{k+1} := \underset{x}{\operatorname{argmin}}\, \mathcal{L}_\rho \left(x, z^k, \lambda^k\right)$$
$$z^{k+1} := \underset{z}{\operatorname{argmin}}\, \mathcal{L}_\rho \left(x^{k+1}, z, \lambda^k\right)$$
$$\lambda^{k+1} := \lambda^k + \rho \left(M_1 x^{k+1} + M_2 z^{k+1} - c\right)$$

end

Due to the distributed structure of the algorithm, it is especially well suited for applications to large scale problems. In the following, we review some essential convergence properties, and state necessary and sufficient optimality conditions for Algorithm 5.3.

Theorem 5.4. *(Convergence of ADMM, Boyd et al. (2011))*
If both assumptions

(1) $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *and* $g : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ *are proper, convex, and lower semi-continuous, and*

(2) $\mathcal{L}_0$ *has a saddle point*

hold, then as $k \to \infty$*, the iterates of Algorithm 5.3 satisfy*

(i) $r^k \to 0$ *(residual convergence),*

(ii) $f(x^k) + g(z^k) \to p^*$ *(objective convergence),*

(iii) $\lambda^k \to \lambda^*$ *(dual variable convergence)*

where $r^k = M_1 x^k + M_2 z^k - c$ *denotes the residual,* $p^* := \inf \{f(x) + g(z) \mid M_1 x + M_2 z = c\}$ *is an optimal value of* (5.15)*, and* λ^* *is an optimal selection for the dual variable.*

Theorem 5.5. *(Necessary and sufficient optimality conditions for ADMM, Boyd et al. (2011))*

If x^* and z^* are primal optimal points, and λ^* is dual optimal point, then primal feasibility is given by

$$M_1 x^* + M_2 z^* - c = 0$$

and dual feasibility is given by

$$0 \in \partial f(x^*) + M_1^\top \lambda^*$$
$$0 \in \partial g(z^*) + M_2^\top \lambda^*$$

where ∂ denotes the subdifferential.

Besides, if x can be split into subvectors $x = (x_1, \ldots, x_N)$, and f and $\|M_1 x\|_2^2$ are separable accordingly, then the x-minimization iteration step in Algorithm 5.3 can be partitioned into the updates of the N subvectors. Hence, computation time can be reduced due to a decrease in problem size and by executing the N subproblems in parallel.

5.5. Solving the Model-Fitting Problem

We efficiently solve inverse problem (5.14) by ADMM. First, consider (5.14) without non-negativity constraint. To obtain the ADMM formulation of the problem, we substitute

$$f(x) := \frac{1}{2}\|Ax - y\|_2^2,$$
$$g(z) := \underbrace{\alpha \left\| z^{\mathrm{IC}} \right\|_1}_{g_1(z^{\mathrm{IC}})} + \underbrace{\sum_{j=1}^{n_b} \frac{\beta_j}{2} \left\| z_j^{\mathrm{EC}} \right\|_{H^s}^2}_{g_2\left(z_1^{\mathrm{EC}}\right)+\ldots+g_{n_b+1}\left(z_{n_b}^{\mathrm{EC}}\right)} + \underbrace{\frac{\gamma}{2} \left\| z^{\mathrm{ISO}} \right\|_{H^s}^2}_{g_{n_b+2}(z^{\mathrm{ISO}})},$$

in the objective, and $M_1 := I$, $M_2 := -I$, $c := 0$ in the equality constrained formulation of (5.15). Let the indicator function $\mathcal{I}_+ : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be defined by

$$\mathcal{I}_+(z) := \begin{cases} 0, & \text{if } z \geq 0 \\ \infty & \text{else.} \end{cases}$$

To account for non-negativity of z, and (indirectly) x, we replace g by an extended version $\tilde{g}(z) = g(z) + \mathcal{I}_+(z)$. In summary, problem (5.14) in ADMM form is given

by

$$\begin{aligned} &\underset{x,z}{\operatorname{argmin}}\, f(x) + \tilde{g}(z) \\ &\text{subject to } x = z. \end{aligned} \tag{5.18}$$

Accordingly, using the identity

$$\lambda^\top a + \frac{\rho}{2}\|a\|^2 = \frac{\rho}{2}\left(\|a + \lambda/\rho\|^2 - \|\lambda/\rho\|^2\right),$$

the augmented Lagrangian with scaled dual variable $u := \lambda/\rho$ is given by

$$\begin{aligned} \mathcal{L}_\rho(x,z,u) &= f(x) + \tilde{g}(z) + \frac{\rho^{\mathrm{IC}}}{2}\left(\left\|x^{\mathrm{IC}} - z^{\mathrm{IC}} + u^{\mathrm{IC}}\right\|_2^2 - \left\|u^{\mathrm{IC}}\right\|_2^2\right) \\ &\quad + \frac{\rho^{\mathrm{EC}}}{2}\left(\sum_{j=1}^{n_b}\left\|x_j^{\mathrm{EC}} - z_j^{\mathrm{EC}} + u_j^{\mathrm{EC}}\right\|_2^2 - \left\|u_j^{\mathrm{EC}}\right\|_2^2\right) \\ &\quad + \frac{\rho^{\mathrm{ISO}}}{2}\left(\left\|x^{\mathrm{ISO}} - z^{\mathrm{ISO}} + u^{\mathrm{ISO}}\right\|_2^2 - \left\|u^{\mathrm{ISO}}\right\|_2^2\right) \\ &= f(x) + \tilde{g}(z) + \frac{1}{2}\left(\left\|D^{\frac{1}{2}}(x - z + u)\right\|_2^2 - \left\|D^{\frac{1}{2}}u\right\|_2^2\right) \end{aligned}$$

with x, z and u partitioned as defined in (5.6), penalty parameters ρ^{IC}, ρ^{EC}, $\rho^{\mathrm{ISO}} > 0$, and diagonal matrix

$$D^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\rho^{\mathrm{IC}}} I_{n_f} & & \\ & \sqrt{\rho^{\mathrm{EC}}} I_{n_b n_v} & \\ & & \sqrt{\rho^{\mathrm{ISO}}} I_{n_v} \end{pmatrix}$$

where I_n denotes the $n \times n$ identity matrix. Minimizing $\mathcal{L}_\rho(x,z,u)$ alternately for x and z, and updating u, yields the ADMM iterations. Below, we regard the x- and z-minimizations in more detail, and summarize our method in Algorithm 5.6.

x-Minimization The x-minimization problem is given by

$$\overline{x} = \underset{x}{\operatorname{argmin}}\left\{f(x) + \frac{1}{2}\left\|D^{\frac{1}{2}}(x - z + u)\right\|_2^2\right\}.$$

Since the gradient of the objective can be computed as

$$\nabla_x\left\{f(x) + \frac{1}{2}\left\|D^{\frac{1}{2}}(x - z + u)\right\|_2^2\right\} = A^\top Ax - A^\top y + D^{\frac{1}{2}}D^{\frac{1}{2}}(x - z + u),$$

we obtain

$$\overline{x} = \left(A^\top A + D^{\frac{1}{2}} D^{\frac{1}{2}}\right)^{-1} \left(A^\top y + D^{\frac{1}{2}} D^{\frac{1}{2}} (z - u)\right).$$

Thus, in ADMM iteration k the new iterate x^{k+1} is given by

$$x^{k+1} = \left(A^\top A + D^{\frac{1}{2}} D^{\frac{1}{2}}\right)^{-1} \left(A^\top y + D^{\frac{1}{2}} D^{\frac{1}{2}} \left(z^k - u^k\right)\right).$$

z-Minimization The z-minimization problem

$$\overline{z} = \underset{z}{\operatorname{argmin}} \left\{ \tilde{g}(z) + \frac{1}{2} \left\| D^{\frac{1}{2}} (x - z + u) \right\|_2^2 \right\}$$

is split up into $n_b + 2$ subproblems

$$\overline{z}^{\mathrm{IC}} = \underset{z^{\mathrm{IC}}}{\operatorname{argmin}} \left\{ \alpha \left\| z^{\mathrm{IC}} \right\|_1 + \mathcal{I}_+(z^{\mathrm{IC}}) + \frac{\rho^{\mathrm{IC}}}{2} \left\| x^{\mathrm{IC}} - z^{\mathrm{IC}} + u^{\mathrm{IC}} \right\|_2^2 \right\}, \tag{5.19}$$

$$\overline{z}_j^{\mathrm{EC}} = \underset{z_j^{\mathrm{EC}}}{\operatorname{argmin}} \left\{ \frac{\beta_j}{2} \left\| z_j^{\mathrm{EC}} \right\|_{H^s}^2 + \mathcal{I}_+(z_j^{\mathrm{EC}}) + \frac{\rho^{\mathrm{EC}}}{2} \left\| x_j^{\mathrm{EC}} - z_j^{\mathrm{EC}} + u_j^{\mathrm{EC}} \right\|_2^2 \right\}, \tag{5.20}$$

$$j = 1, \ldots, n_b$$

$$\overline{z}^{\mathrm{ISO}} = \underset{z^{\mathrm{ISO}}}{\operatorname{argmin}} \left\{ \frac{\gamma}{2} \left\| z^{\mathrm{ISO}} \right\|_{H^s}^2 + \mathcal{I}_+(z^{\mathrm{ISO}}) + \frac{\rho^{\mathrm{ISO}}}{2} \left\| x^{\mathrm{ISO}} - z^{\mathrm{ISO}} + u^{\mathrm{ISO}} \right\|_2^2 \right\} \tag{5.21}$$

which can be solved independently. A solution to (5.19) is given by

$$\overline{z}^{\mathrm{IC}} = \mathcal{S}^+_{\alpha/\rho^{\mathrm{IC}}} \left(x^{\mathrm{IC}} + u^{\mathrm{IC}}\right)$$

where

$$\mathcal{S}_a^+(x_i) := \begin{cases} x_i - a, & \text{if } x_i > a \\ 0, & \text{else} \end{cases}$$

denotes a variation of the soft-threshold operator ensuring non-negativity. Neglecting regularization parameters β_j and γ, problems (5.20) and (5.21) can equivalently be represented by constrained quadratic optimization problems of the form

$$\begin{aligned} &\underset{v}{\operatorname{argmin}} \left\{ \frac{1}{2} v^\top Q^{\{\mathrm{EC,ISO}\}} v - \rho v^\top w \right\} \\ &\text{subject to } v \geq 0 \end{aligned}$$

with $Q_j^{\mathrm{EC}} = \beta_j (I - \Delta)^s + \rho_{\mathrm{EC}} I$ and $Q^{\mathrm{ISO}} = \gamma (I - \Delta)^s + \rho_{\mathrm{ISO}} I$. This problem can be solved using, for instance, active-set, gradient-projection or interior-point methods. In summary, the ADMM iterations are given by:

Algorithm 5.6. (ADMM applied to problem (5.14))
Given $z^0, u^0 \in \mathbb{R}^{n_f+n_b n_v+n_v}$, $\rho_{\mathrm{IC}}, \rho_{\mathrm{EC}}, \rho_{\mathrm{ISO}} > 0$:
for $k = 0, 1, 2, \ldots$

x-minimization:

$$x^{k+1} = \left(A^\top A + D^{\frac{1}{2}} D^{\frac{1}{2}}\right)^{-1} \left(A^\top y + D^{\frac{1}{2}} D^{\frac{1}{2}} \left(z^k - u^k\right)\right) \tag{5.22}$$

z-minimization:

$$\begin{aligned}
\left(z^{\mathrm{IC}}\right)^{k+1} &= \mathcal{S}^+_{\alpha/\rho^{\mathrm{IC}}} \left(\left(x^{\mathrm{IC}}\right)^{k+1} + \left(u^{\mathrm{IC}}\right)^k\right) \\
\left(z_j^{\mathrm{EC}}\right)^{k+1} &= \underset{z_j^{\mathrm{EC}} \geq 0}{\operatorname{argmin}} \left\{ \frac{\beta_j}{2} \left(z_j^{\mathrm{EC}}\right)^\top Q_j^{\mathrm{EC}} z_j^{\mathrm{EC}} - \rho^{\mathrm{EC}} \left(z_j^{\mathrm{EC}}\right)^\top \left(\left(x_j^{\mathrm{EC}}\right)^{k+1} + \left(u_j^{\mathrm{EC}}\right)^k\right) \right\}, \\
& \qquad j = 1, \ldots, n_b \\
\left(z^{\mathrm{EC}}\right)^{k+1} &= \left(\left(z_1^{\mathrm{EC}}(1)\right)^{k+1}, \ldots, \left(z_{n_b}^{\mathrm{EC}}(1)\right)^{k+1}, \ldots, \left(z_1^{\mathrm{EC}}(n_v)\right)^{k+1}, \ldots, \left(z_{n_b}^{\mathrm{EC}}(n_v)\right)^{k+1}\right)^\top \\
\left(z^{\mathrm{ISO}}\right)^{k+1} &= \underset{z^{\mathrm{ISO}} \geq 0}{\operatorname{argmin}} \left\{ \frac{\gamma}{2} \left(z^{\mathrm{ISO}}\right)^\top Q^{\mathrm{ISO}} z^{\mathrm{ISO}} - \rho^{\mathrm{ISO}} \left(z^{\mathrm{ISO}}\right)^\top \left(\left(x^{\mathrm{ISO}}\right)^{k+1} + \left(u^{\mathrm{ISO}}\right)^k\right) \right\} \\
z^{k+1} &= \left[\left(z^{\mathrm{IC}}\right)^{k+1} \mid \left(z^{\mathrm{EC}}\right)^{k+1} \mid \left(z^{\mathrm{ISO}}\right)^{k+1}\right]
\end{aligned}$$

dual update:

$$u^{k+1} = u^k + x^{k+1} - z^{k+1}$$

end

Convergence

The following Theorem ensures residual convergence, objective convergence, and dual variable convergence of Algorithm 5.6.

Theorem 5.7. *Assumptions (1) and (2) of Theorem 5.4 hold for optimization problem* (5.18).

Proof. The proof is stated in Appendix D. □

In this section, we have derived the ADMM algorithm for the solution of our optimization problem (5.14). However, the straight-forward solution of the x-minimization step, involving the solution of a large-scale problem, is computationally very expensive. Hence, in the next section, we focus on more efficient computation strategies regarding this algorithm step.

5.6. Strategies for Efficient x-Minimization

In practice, x-minimization step (5.22) requires the solution of a large system of linear equations. In this section, we discuss strategies for efficient computation of this large-scale problem, for instance by exploiting the special structure of dictionary matrix A.

5.6.1. Using a Precomputed Cholesky Decomposition

Since $\left(A^\top A + D^{\frac{1}{2}} D^{\frac{1}{2}}\right)$ is symmetric positive definite for $\rho_{\mathrm{IC}}, \rho_{\mathrm{EC}}, \rho_{\mathrm{ISO}} > 0$, the *Cholesky decomposition* $\left(A^\top A + D^{\frac{1}{2}} D^{\frac{1}{2}}\right) = LL^\top$ with lower triangular matrix $L \in \mathbb{R}^{N \times N}$, $N := n_f + n_b n_v + n_v$, exists. Using a precomputed Cholesky decomposition, (5.22) can be performed more efficiently by forward substitution solving $Ly = A^\top y + D^{\frac{1}{2}} D^{\frac{1}{2}} \left(z^k - u^k\right)$ for y, followed by a back substitution step solving $L^\top x^{k+1} = y$ for x^{k+1}. Hence, in each ADMM iteration the x-minimization step has the computational complexity $\mathcal{O}(N^2)$. Algorithms calculating the Cholesky decomposition usually have order of N^3 complexity, but this computation is only required once prior to the ADMM iterations.

However, since L is usually a dense matrix, an *incomplete* Cholesky decomposition can be more suitable depending on the degree of sparsity of $\left(A^\top A + D^{\frac{1}{2}} D^{\frac{1}{2}}\right)$. An incomplete Cholesky decomposition can simply be obtained by calculating only the entries in L that coincide with non-zero elements in A and setting the remaining elements in L equal to zero. It is most commonly used as preconditioner for the preconditioned conjugate gradient (pCG) algorithm (Golub and van Loan, 1983, 10.3).

5.6.2. Dimension Reduction Using Truncated SVDs

Dictionary matrix $A \in \mathbb{R}^{n_d n_v \times N}$, where $N := n_f + n_b n_v + n_v$ for simplicity, can be considered as a block matrix

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_{n_v} \end{bmatrix} \text{ with } A_i \in \mathbb{R}^{n_d \times N}, i = 1, \ldots, n_v. \tag{5.23}$$

Hence, each block matrix contains the diffusion information with respect to the n_d gradient directions for a certain voxel.

Computing the singular value decomposition (SVD) for each block matrix A_i, we observe that the singular values rapidly decrease towards zero. This property motivates replacing A_i by a reduced rank approximation to reduce computation time. Recall that the *truncated* SVD is defined as follows:

Definition 5.8. (truncated SVD) Let $M = U\Sigma V^\top$ denote the SVD of matrix $M \in \mathbb{R}^{m\times n}$ of rank $p \leq \min\{m, n\}$. A truncated SVD of M is given by

$$M_k = U_k \Sigma_k V_k^\top, \quad k \ll p, \qquad \text{with } \Sigma_k = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_k \end{pmatrix}$$

where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$ are the k largest singular values of M, and $U_k = [u_1, \ldots, u_k] \in \mathbb{R}^{m\times k}$, $V_k = [v_1, \ldots, v_k] \in \mathbb{R}^{n\times k}$ consist of the corresponding singular vectors of M.

Note that by the Eckard-Young theorem, matrix M_k is the closest rank-k approximation to M.

Let the truncated SVD of matrix A_i be given by $A_{k_i} = U_{k_i}\Sigma_{k_i}V_{k_i}^\top$, where $U_{k_i} \in \mathbb{R}^{n_d\times k_i}$, $V_{k_i} \in \mathbb{R}^{N\times k_i}$, and $\Sigma_{k_i} \in \mathbb{R}^{k_i\times k_i}$. In (5.22) we require the matrix product $A^\top A$, which can be represented as follows using the block diagonal structure of A stated in (5.23):

$$A^\top A = \sum_{i=1}^{n_v} A_i^\top A_i \approx \sum_{i=1}^{n_v} V_{k_i}\Sigma_{k_i}U_{k_i}^\top U_{k_i}\Sigma_{k_i}V_{k_i}^\top = \sum_{i=1}^{n_v} V_{k_i}\Sigma_{k_i}^2 V_{k_i}^\top = VBV^\top \quad (5.24)$$

where $U_{k_i}^\top U_{k_i} = I_{k_i}$ due to orthogonality of U_{k_i}, and defining block matrix $V := [V_1, V_2, \ldots, V_{n_v}] \in \mathbb{R}^{N\times k}$, $k = \sum_i^{n_v} k_i$, and block diagonal matrix $B \in \mathbb{R}^{k\times k}$ with diagonal block, thus $\text{diag}(B) = \left[\Sigma_1^2, \Sigma_2^2, \ldots, \Sigma_{n_v}^2\right]$.

Reformulating the inverse in x-minimization step (5.22) of Algorithm 5.6, we obtain

$$\begin{aligned}\left(A^\top A + D^{\frac{1}{2}}D^{\frac{1}{2}}\right)^{-1} &= \left(D^{\frac{1}{2}}\left(D^{-\frac{1}{2}}A^\top A D^{-\frac{1}{2}} + I_N\right)D^{\frac{1}{2}}\right)^{-1} \\ &= D^{-\frac{1}{2}}\Big(\underbrace{D^{-\frac{1}{2}}A^\top A D^{-\frac{1}{2}}}_{\tilde{A}^\top\tilde{A}} + I_N\Big)^{-1} D^{-\frac{1}{2}}\end{aligned}$$

with diagonal matrix $D^{\frac{1}{2}}$, and defining $\tilde{A} := AD^{-\frac{1}{2}}$. Similar to A_i, the singular values of many of the matrices $\tilde{A}_i$ decrease rapidly towards zero. Computing the truncated SVDs of $\tilde{A}_i$ we get $\tilde{A}^\top \tilde{A} \approx \tilde{V}\tilde{B}\tilde{V}^\top$ as in (5.24), and the following theorem holds:

Theorem 5.9. *Let $\tilde{V} = QR$ denote the QR-factorization of $\tilde{V}$ where $R \in \mathbb{R}^{k\times k}$ and $Q \in \mathbb{R}^{N\times k}$, and define $Z := R\tilde{B}R^\top$. Then, $\tilde{V}\tilde{B}\tilde{V}^\top = QZQ^\top$ and the inverse of matrix $\left(QZQ^\top + I_N\right)$ is given by*

$$\left(QZQ^\top + I_N\right)^{-1} = Q\left[(Z + I_k)^{-1} - I_k\right]Q^\top + I_N. \tag{5.25}$$

Proof. The first assertion follows immediately from substituting $\tilde{V} = QR$ and $Z = R\tilde{B}R^\top$:

$$\tilde{V}\tilde{B}\tilde{V}^\top = QR\tilde{B}(QR)^\top = QR\tilde{B}R^\top Q^\top = QZQ^\top$$

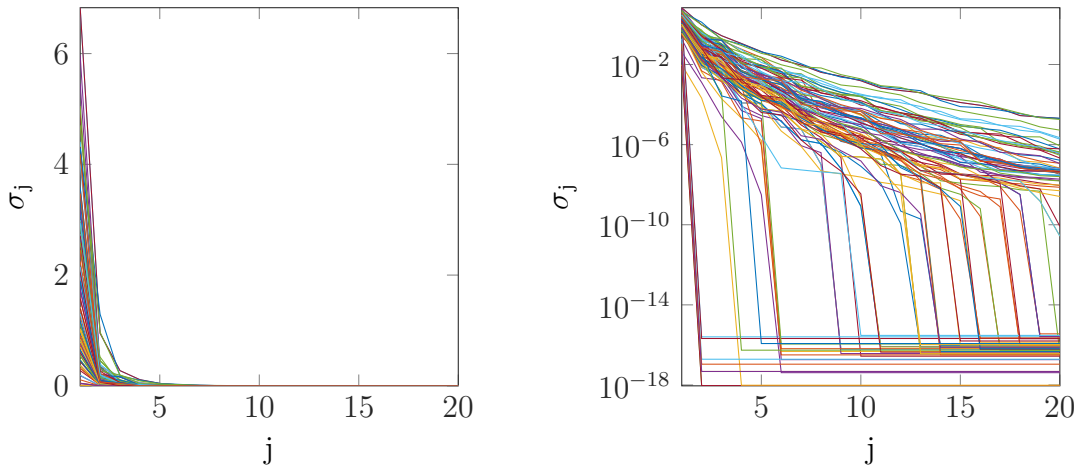
To prove that equation (5.25) holds true, we show that multiplying the inverse by $QZQ^\top + I_N$ yields the identity:

$$\begin{aligned}
&\left(QZQ^\top + I_N\right)\left(Q\left[(Z + I_k)^{-1} - I_k\right]Q^\top + I_N\right)\\
&= Q\Big(Z\underbrace{Q^\top Q}_{I_k}\left[(Z + I_k)^{-1} - I_k\right] + Z + (Z + I_k)^{-1} - I_k\Big)Q^\top + I_N\\
&= Q\Big(Z\,(Z + I_k)^{-1}\underbrace{-Z + Z}_{=0} + (Z + I_k)^{-1} - I_k\Big)Q^\top + I_N\\
&= Q\underbrace{(Z + I_k - (Z + I_k))}_{=0}(Z + I_k)^{-1}Q^\top + I_N\\
&= I_N
\end{aligned}$$

□

Hence, by applying (5.25), the inversion of $N \times N$ matrix $\left(\tilde{A}^\top\tilde{A} + I_N\right)$ is reduced to matrix-vector multiplications and the inversion of $k \times k$ matrix $(Z + I_k)$ with $k \ll N$. Again, a Cholesky decomposition of $(Z + I_k)$ can be precomputed to speed up ADMM iterations. Then, the computational complexity of applying the approximated inverse is $\mathcal{O}(N+Nk+k^2)$ in each ADMM iteration step. Since $k \ll N$, both the calculation of the Cholesky decomposition and the application of the inverse matrix are much more efficient compared to performing the computations with matrix $\left(\tilde{A}^\top\tilde{A} + I_N\right)$ as described in Section 5.6.1.

Instead of using direct solvers, (5.25) can also function as a preconditioner for pCG. However, this is only favorable if k is very small, i.e. only few singular values are kept.



(a) linear scale **(b)** logarithmic scale

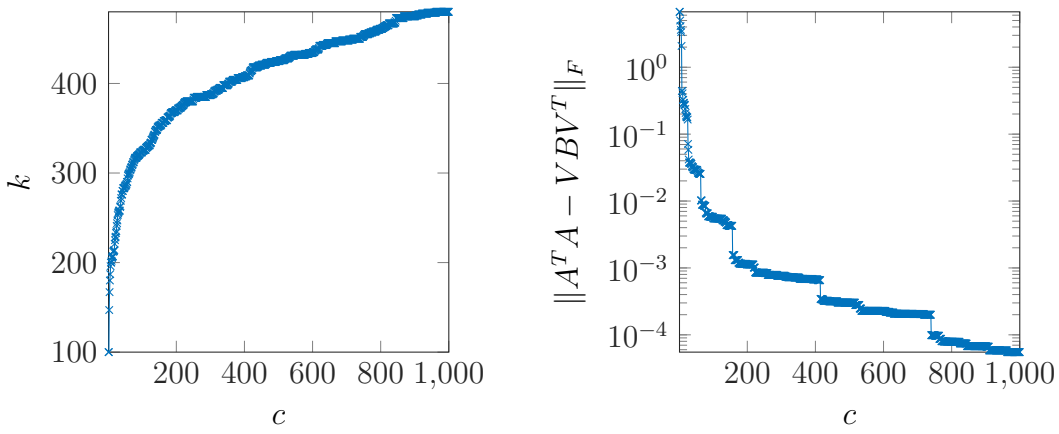
Figure 5.6.: The 20 largest singular values σ_j, $j = 1, \ldots, 20$, of each matrix A_i, $i = 1, \ldots, 100$. Each plot corresponds to a different A_i.

Numerical Experiments

We regard the *Fiber Cup* data set with 64 gradients and b0-image, and a corresponding tractogram consisting of 7500 evenly distributed tracks. 100 white-matter voxels are randomly selected, and a test matrix $A = [A_{\mathrm{IC}}|A_{\mathrm{ISO}}]$ of size $(n_v \cdot n_d) \times (n_f + n_v)$ with $n_v = 100$, $n_d = 65$, $n_f = 7500$, modeling restricted and isotropic diffusion, is generated. In Figure 5.6, the 20 largest singular values are plotted on linear and logarithmic scale for each block matrix $A_i \in \mathbb{R}^{n_d \times n_f + n_v}$, $i = 1, \ldots, 100$. Given a threshold $c > 1$ on the condition of matrix A_i, let $\sigma_{k_i}(A_i)$ be the smallest singular value of A_i such that $\sigma_1(A_i)/\sigma_{k_i}(A_i) < c$. For each block A_i, all singular values smaller than σ_{k_i} are truncated. The number of all singular values included in the truncated SVDs of A_i, $i = 1, \ldots, 100$, defined by k, is plotted against different values of c in Figure 5.7a. Correspondingly, Figure 5.7b shows the difference between $A^\top A$ and its approximation by $VBV^\top$. In Figure 5.8a, the relative error $e = \|x - \tilde{x}\|_2^2 / \|x\|_2^2$ between true x solving $(A^\top A + \rho I_N)x = A^\top y$, and solution $\tilde{x}$ obtained using (5.25) is plotted for $\rho = 1$ and different values of c. Figure 5.8b shows e for $\rho = \sigma_1^2(A)/(c-1)$.

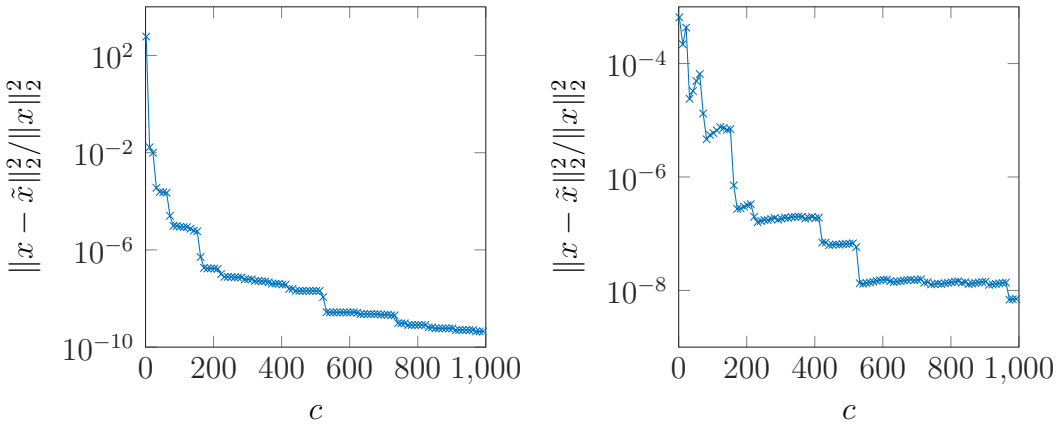
5.7. Numerical Experiments and Results

We tested the methods derived in Sections 5.5 and 5.6 on simulated and real diffusion data. Experiments and results are presented in this section.



(a) number of included singular values **(b)** approximation error

Figure 5.7.: Number of singular values $k = k_1 + \ldots + k_{n_v}$ (*left*), and approximation error in Frobenius norm (*right*), for different condition thresholds c on the truncated SVD of A_i.



(a) $\rho = 1$ **(b)** $\rho = \sigma_1^2(A)/(c-1)$

Figure 5.8.: Relative error of the solution, for different condition thresholds c on the truncated SVD of A_i.

5.7.1. Simulated Test Data

For first tests of the ADMM algorithm's properties on data with known solution, we generate a simple data set which is not based on a physical model. The resulting system of linear equations considered in this section is much smaller than it is for real data sets used in Section 5.7.2.

Signal Modeling

First, we test our algorithm on a simple simulated data set that we obtain as follows. Let n fibers, distributed on the surface of the unit ball, be given by the curves $\gamma_k : [0,1] \to \mathcal{S}^2$, $k = 1, \dots, n$. We define the signal $Y_k : \mathcal{S}^2 \to \mathbb{R}$ originating from fiber k by

$$Y_k(u) = \int_0^1 \exp\left\{-\frac{1}{\sigma^2}|\gamma_k(t) - u|^2\right\} |\gamma_k'(t)| dt.$$

Regarding the signal at m points on the sphere, the dictionary matrix can be written as

$$A = \begin{pmatrix} Y_1(u_1) & \dots & Y_n(u_1) \\ \vdots & & \vdots \\ Y_1(u_m) & \dots & Y_n(u_m) \end{pmatrix} \tag{5.26}$$

where $u_1, \dots, u_m \in \mathcal{S}^2$. A subset of the n fibers is specified as valid, the remaining fibers as invalid. Accordingly we define a weight vector

$$(x_{\text{true}})_k = \begin{cases} 1 & \text{if fiber } k \text{ is valid} \\ 0 & \text{else} \end{cases}$$

specifying the contribution of each fiber to the overall signal. Furthermore, we define the overall (*observed*) signal by

$$Y = Ax_{\text{true}} + \epsilon \tag{5.27}$$

with Gaussian white noise ϵ. After discretization, we obtain data as depicted in Figure 5.9.

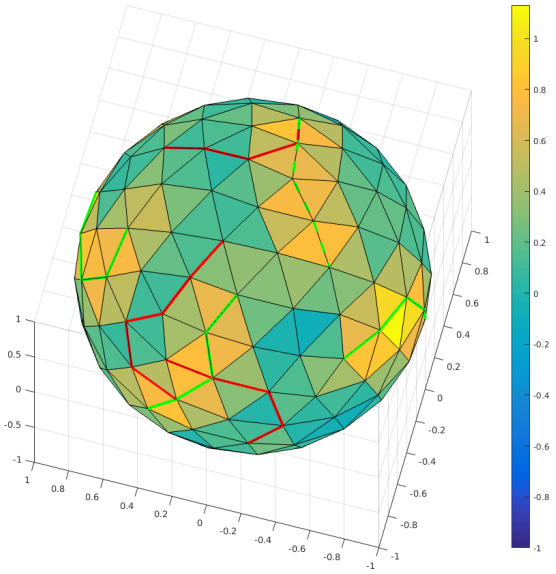


Figure 5.9.: Simulated data on the sphere with Gaussian noise added such that PSNR $\approx$ 20 dB. The triangles are colored according to signal Y. True fibers are plotted in green, false fibers in red. Dashed green fibers represent true fibers with overlapping to a certain extent with a false fiber.

Data and Results

In this section, we present results for a data set obtained as described above with $n_d = 162$ sampling points on the sphere and $n_f = 10$ fibers. Matrix $A_{\text{IC}} \in \mathbb{R}^{162\times 10}$ is obtained as in (5.26), $A_{\text{ISO}} \in \mathbb{R}^{162\times 1}$ is set to a vector with constant entries. Extra-cellular contributions were not simulated to keep this first example simple. As can be observed in Figure 5.10, dictionary matrix $[A_{\text{IC}}, A_{\text{ISO}}]$ is sparse except for the last column. Half of the simulated fibers are categorized as valid, the other half as invalid. This classification is encoded as ones and zeros in vector x^{IC}. The isotropic part x^{ISO} consists of a single component that is set to one. The signal Y is simulated as stated in (5.27) with $x_{\text{true}} = [x^{\text{IC}}, x^{\text{ISO}}]$ and additive noise ϵ chosen such that PSNR = 20dB.

We use the proposed ADMM algorithm (Algorithm 5.6) to solve

$$\underset{x\geq 0}{\operatorname{argmin}} \frac{1}{2}\|Ax - Y\|_2^2 + \alpha\|x^{\text{IC}}\|_1 + \frac{\beta}{2}\left(x^{\text{ISO}}\right)^2,$$

setting $\rho_{\text{IC}} = 0.9$ and $\rho_{\text{ISO}} = 0.1$. Figure 5.11 shows the relative squared distance between x_{true} and the solution x of the ADMM minimization for different choices of $\alpha, \beta \in [0, 1]$ after 20 ADMM iterations. The minimum distance is attained for $\alpha = 0.13$ and $\beta = 0.22$. In Figure 5.12, relative error, norm of the

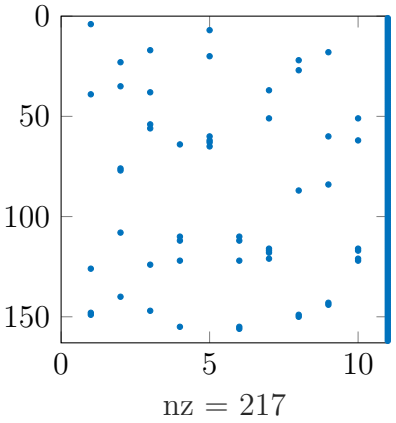


Figure 5.10.: Nonzero entries of matrix A considered in Section 5.7.1.

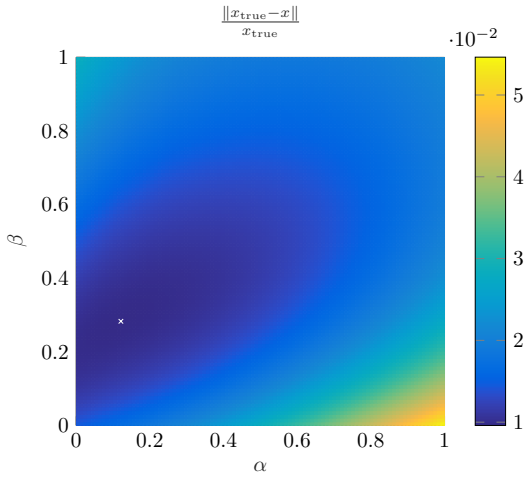


Figure 5.11.: Relative error for different choices of parameters α and β for simulated data on the sphere with PSNR $\approx$ 20dB. The white cross highlights the minimum value.

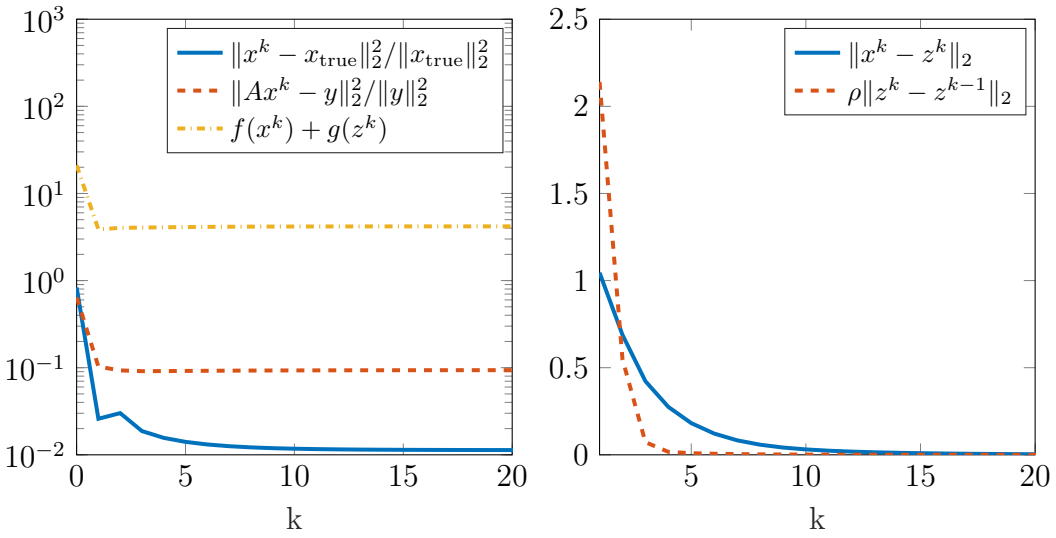


Figure 5.12.: Norm of the relative error, norm of relative residual, and objective (*left*), primal and dual residual norm (right) plotted against ADMM iteration k.

relative residual, objective, primal and dual residual of the ADMM optimization problem are plotted against the number of iterations for $\alpha = 0.13$ and $\beta = 0.22$. We observe that, as stated in Theorem 5.4, the objective converges, and primal and dual residual norm converge to zero. In Table 5.1, solutions for different choices of α and number of ADMM iterations are compared to the true solution. Moreover, the solution of $Ax = Y$ obtained using MATLAB's `mldivide` is stated for comparison. As expected, increasing α and/or the number of iterations increases sparsity of the solution. A larger parameter α also results in a stronger enforcement of the minimization of the ℓ_1 norm relative to the minimization of the data misfit term. Hence, the components of x^{IC} become small while norm of error and residual increase. To extract the true fibers from a solution vector it is advisable to set a threshold close to zero and discard those tracks corresponding to weights below that threshold. For instance, with respect to the solution for $\alpha = 0.13$ and $k = 20$, discarding all tracks with absolute weights smaller or equal to 0.0028 yields all valid fibers. Furthermore, compared to the solution without regularization, the solution of the ADMM algorithm shows a great improvement.

Overall, we notice that for this first test problem the ADMM algorithm converges after less than 20 iteration steps to a good approximation to the true solution from which all tracks can correctly be classified as valid or invalid.

	$\alpha = 0.13$		$\alpha = 0.3$		no reg.
	$k = 20$	$k = 100$	$k = 20$	$k = 100$	(direct)
$x_{\text{true}} =$	$x^k =$	$x^k =$	$x^k =$	$x^k =$	$x =$
0	−0.0010	−0.0000	−0.0011	0.0000	−0.0426
1	1.0585	1.0586	1.0295	1.0296	1.0582
0	0.0017	0.0018	0.0008	0.0000	0.0010
1	1.0149	1.0150	0.9964	0.9965	1.0129
1	1.1292	1.1290	1.1044	1.1042	1.1495
1	0.8493	0.8494	0.8251	0.8252	0.8496
0	−0.0005	0.0000	−0.0004	−0.0000	−0.0548
0	−0.0007	−0.0000	−0.0009	0.0000	0.0066
0	−0.0028	−0.0000	−0.0028	0.0000	−0.2042
1	0.8807	0.8804	0.8630	0.8628	0.9157
1	1.0470	1.0459	1.0830	1.0821	1.2575
$\frac{\|x^k - x_{\text{true}}\|_2^2}{\|x_{\text{true}}\|_2^2}$	0.00992	0.009899	0.011338	0.011309	0.028082
$\frac{\|Ax^k - y\|_2^2}{\|y\|_2^2}$	0.093129	0.09321	0.09371	0.093792	0.089613

Table 5.1.: Comparison between true solution, solution of ADMM for two different choices of α and different numbers of iterations, and the solution obtained from applying MATLAB's `mldivide` to the unregularized problem.

5.7.2. Diffusion-Weighted MR Data

In the following, we apply the proposed ADMM algorithm to two different diffusion-weighted MR data sets. First, we consider an *in-vivo* data set for which we present qualitative results in comparison with results obtained using the COMMIT algorithm by Daducci et al. (2015). Then, we present experiments and quantitative results for a diffusion phantom data set with known ground truth.

in-vivo Data

We tested our algorithm on *in-vivo* human brain data that was provided with the COMMIT software package which is available at https://github.com/daducci/COMMIT. The diffusion MR data contains 30 diffusion-weighted measurements at $b = 700$ s/mm^2, 60 at $b = 2000$ s/mm^2, and 10 unweighted measurements, and has a spatial resolution of $106 \times 106 \times 60$. Moreover, a corresponding tractogram consisting of $n_f = 283522$ tracks is given. All b_0-measurements are averaged resulting in one merged b_0-image, and the diffusion-weighted images are normalized to the b_0-image. The domain containing fiber tracks includes $n_v = 53021$ voxels. Hence, the dictionary matrix, obtained using the default set-

tings of the COMMIT software for this data set, is given by $A = [A_{\mathrm{IC}}|A_{\mathrm{EC}}|A_{\mathrm{ISO}}]$ with block matrix $A_{\mathrm{IC}} \in \mathbb{R}^{n_v n_u \times n_f}$ encoding the intra-cellular components, block matrix $A_{\mathrm{EC}} \in \mathbb{R}^{n_v n_u \times 145471}$ for the extra-cellular components, and isotropic part $A_{\mathrm{ISO}} \in \mathbb{R}^{n_v n_u \times 106042}$. Matrix A_{EC} encodes for each voxel between 1 and 3 fiber populations, matrix A_{ISO} consists of two isotropic contributions per voxel. The structure of the dictionary matrix is illustrated in Figure 5.2 at the beginning of this chapter.

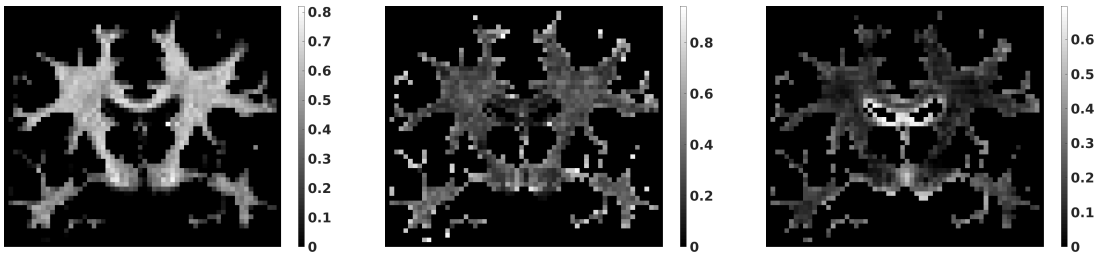
All regularization parameters are set to the same value, that is $\alpha = \beta_i = \gamma = 0.1$, $i = 1, 2, 3$, and for the Sobolev norm we select $s = 1$. The penalty parameters of the augmented Lagrangian are chosen as $\rho^{\mathrm{IC}} = \rho^{\mathrm{EC}} = \rho^{\mathrm{ISO}} = 1$. Apart from the penalty parameters, ADMM expects initial solutions z^0 and u^0 as input. For our tests, we chose z^0 as a vector with identical positive components, and set $u^0 = 0$. As the NNLS algorithm used by the COMMIT framework breaks per default if

$$\frac{|\|Ax^k - y\|_2 - \|Ax^{k-1} - y\|_2|}{\|Ax^k - y\|_2} < 10^{-4}$$

we included the same criteria with our algorithm for comparison. For a coronal image slice, Figure 5.13 compares voxel-wise intra-cellular, extra-cellular and isotropic contributions between the proposed method solving minimization problem 5.14 with ADMM, and for the COMMIT algorithm. The contributions are derived from solution vector x in the following way: For voxel $v \in \Omega$, the isotropic contributions are defined by $\sum_{i=1}^{n_f} \ell_i(v) f_i^{\mathrm{IC}}$, the extracellular contributions are $(f_1^{\mathrm{EC}} + f_2^{\mathrm{EC}} + f_3^{\mathrm{EC}})(v)$ and the isotropic contribution is given by $f^{\mathrm{ISO}}(v)$. For both methods, the values of the IC contributions are highest in regions with dense fiber bundles. The extra-cellular contributions tend to decrease slightly with an increase in IC values, and the isotropic weights are small in regions with fibers and high in regions close to boundaries where partial volume effects are expected. It can be observed that due to H^s-norm regularization the images of EC and ISO contributions resulting from the proposed method are much smoother than the images obtained from COMMIT showing some unnaturally high peaks. RMSE and NRMSE for both methods are shown in Figure 5.14 for the same coronal slice as above.

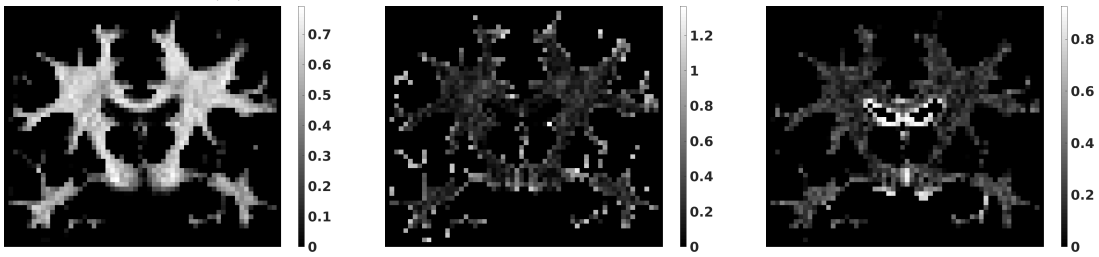
Fiber Cup Phantom Data

In this section we present experiments and results using the Fiber Cup phantom data set described in Section 3.2.1. A deterministic streamline tractography algorithm was used to generate a large number of tracks. As the fiber ground



IC contributions EC contributions ISO contributions

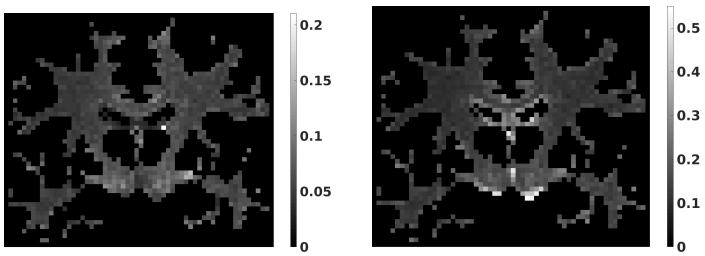
(a) $\ell_1 - H^s$-regularized problem solved with ADMM.



IC contributions EC contributions ISO contributions

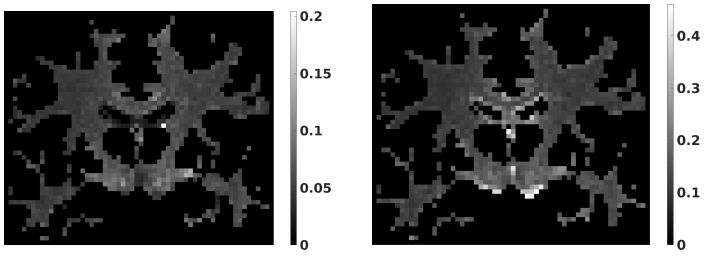
(b) COMMIT

Figure 5.13.: Comparison of intracellular (IC), extracellular (EC), and isotropic (ISO) contributions for a coronal slice.



RMSE NRMSE

(a) $\ell_1 - H^s$-regularized problem solved with ADMM.



RMSE NRMSE

(b) COMMIT

Figure 5.14.: Comparison of RMSE and NRMSE for a coronal slice.

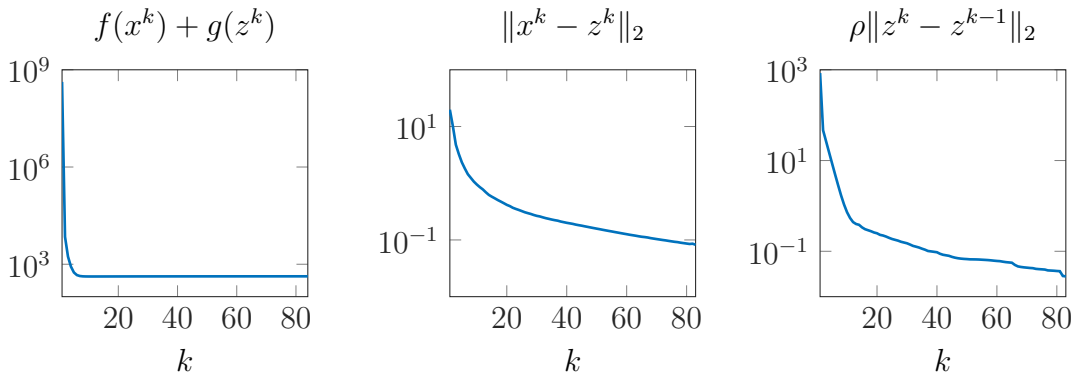


Figure 5.15.: Objective, primal and dual residual norm plotted against ADMM iteration k for the ADMM algorithm with $\rho^{IC} = \rho^{ISO} = 1$ applied to the $\ell_1 - H^s$-regularized problem with $\alpha = \gamma = 0.1$.

truth of the data set is known, the tracks can be assigned to one of the three categories, *valid connection* (VC), *invalid connection* (IC) and *no connection* (NC) explained in Section 2.7.3. From this tractogram, we selected $n_f = 19480$ tracks, 10480 valid connections and 9000 invalid connections, evenly distributed in the white matter regions. The number of voxels containing tracks is $n_v = 2005$. We consider only intracellular and isotropic contributions for simplicity and obtain dictionary matrix $A = [A_{\text{IC}}|A_{\text{ISO}}] \in \mathbb{R}^{130325\times 21485}$ using the COMMIT software by Daducci et al. (2015) which is available at https://github.com/daducci/COMMIT. Correspondingly, vector y containing the measured DW-MRI data is reduced to the $n_v = 2005$ voxels containing tracks, thus $y \in \mathbb{R}^{130325}$. Here, we discuss results obtained from applying ADMM algorithm 5.6 to solve problem

$$\underset{x\geq 0}{\operatorname{argmin}} \frac{1}{2}\|Ax - y\|_2^2 + \alpha\|x^{\text{IC}}\|_1 + \frac{\gamma}{2}\|x^{\text{ISO}}\|_{H^s}^2$$

with A and y as described above, and where $s = 1$ for the Sobolev norm. As for the *in-vivo* data, z^0 is chosen as vector with constant entries, and $u^0 = 0$.

In Figure 5.15, objective, primal and dual residual norm are plotted against ADMM iteration k for $\rho^{IC} = \rho^{ISO} = 1$ and $\alpha = \gamma = 0.1$. The objective changes only slightly for $k > 10$, while primal and dual residual norms tend more slowly towards zero.

With respect to the choice of the regularization parameter, our tests showed that good results can be obtained by setting $\alpha = \gamma \in [2 \cdot 10^{-2}, 2 \cdot 10^{-1}]$. Table 5.2 summarizes results for fixed penalty parameter $\rho^{\text{IC}} = \rho^{\text{ISO}} = 1$ and a collection of choices of regularization parameters α and γ. Furthermore, results obtained

		VC			
		%	#	$\overline{\text{RMSE}}$	$\overline{\text{NRMSE}}$
ℓ_1, H^1 (ADMM)	$\alpha = \gamma$				
	0.04	81.41	3670	0.036 +/- 0.055	0.197 +/- 0.110
	0.06	85.57	3838	0.036 +/- 0.055	0.203 +/- 0.109
	0.08	87.68	4243	0.036 +/- 0.055	0.210 +/- 0.108
	0.10	**89.65**	4582	0.039 +/- 0.055	0.218 +/- 0.107
	0.12	88.12	4998	0.040 +/- 0.055	0.227 +/- 0.106
ℓ_1 (ADMM)	α				
	0.04	82.72	4563	0.035 +/- 0.055	0.197 +/- 0.110
	0.06	86.27	4720	0.036 +/- 0.055	0.203 +/- 0.110
	0.08	88.40	5435	0.037 +/- 0.055	0.211 +/- 0.109
	0.10	73.17	6421	0.039 +/- 0.054	0.221 +/- 0.108
	0.12	65.82	6523	0.040 +/- 0.054	0.232 +/- 0.106
COMMIT	(NNLS)	74.86	542	0.037 +/- 0.055	0.206 +/- 0.110

Table 5.2.: Results for the Fiber Cup data set with an input tractogram consisting of 10480 valid connections and 9000 invalid connections.

using the COMMIT software package, where a NNLS problem is solved as stated in (5.7), and from the solution of a solely ℓ_1-regularized problem

$$\underset{x \geq 0}{\operatorname{argmin}} \frac{1}{2}\|Ax - y\|_2^2 + \alpha\|x\|_1 \tag{5.28}$$

are included as comparison. As the COMMIT software package only provides a solver for NNLS, we used an own implementation for the solution of problem (5.28) which uses ADMM. All algorithms considered in Table 5.2 break if stopping rule 5.7.2 is satisfied. For the $\ell_1 - H^s$ regularized problems that means that ADMM stops after around 80 iterations. As described above, this criterion was chosen for reasons of comparability. In fact, more sophisticated stopping conditions involving primal and dual residual norms could be used for ADMM. The initial tractogram has 53.90% of valid and 46.10% of invalid connections. After the optimization procedure of the respective algorithm where the weights x^{IC} are adjusted, all tracks with $x^{\text{IC}} \leq 0$ are discarded from the tractogram. For the remaining tracks, valid connections in percent are stated in Table 5.2. Using our proposed algorithm, the highest percentage of valid connections, 89.65%, can be obtained. Hence, the percentage of invalid connections can be reduced by about 10% compared to the best solution of the solely ℓ_1-regularized problem,

and by almost 60% compared to NNLS. Furthermore, the proposed algorithm is relatively stable with respect to parameter changes. For $\alpha = \gamma$ between 0.04 and 0.12, the valid connections are above 80%. Using only ℓ_1-regularization, the range of regularization parameters yielding a good result is much narrower. As can be observed in the column showing the absolute number of valid tracks after optimization, many tracks - also valid tracks - are discarded, as they possibly do not yield new information for the description of the diffusion data. The *root mean squared error* (RMSE) for a voxel $v \in \mathbb{R}^3$ is defined by

$$\mathrm{RMSE}(v) = \sqrt{\frac{1}{n_u} \sum_{i=1}^{n_u} \left(S(v, q_i) - \tilde{S}(v, q_i) \right)}$$

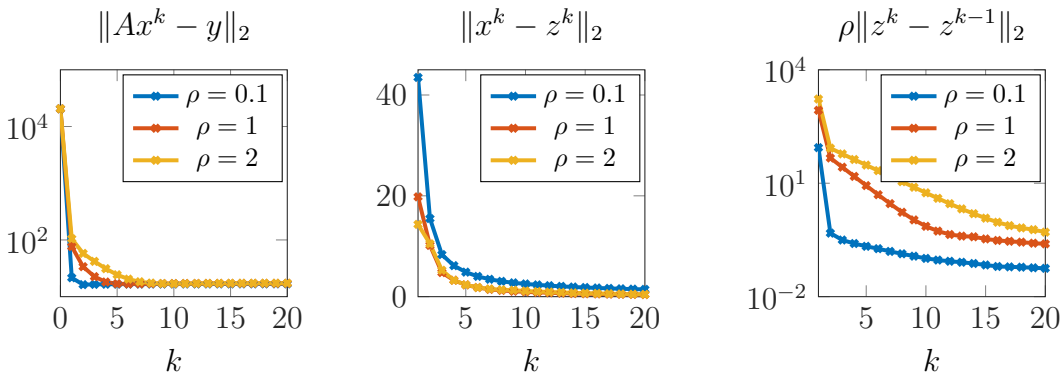
where n_u is the number of gradient directions. Accordingly, the *normed root mean squared error* (NRMSE) is given by

$$\mathrm{NRMSE}(v) = \frac{\mathrm{RMSE}(v)}{\sqrt{\frac{1}{n_u} \sum_{i=1}^{n_u} S(v, q_i)}}.$$

In the last two columns of Table 5.2, mean RMSE and NRMSE, and corresponding standard deviations are shown. As would be expected, the errors increase with increasing regularization parameters.

Different choices of ρ Decreasing $\rho = \rho^{\mathrm{IC}} = \rho^{\mathrm{ISO}}$ increases the ill-conditioning of $(A^\top A + \rho I)$, which results in slower convergence of the CG method in the x-minimization step. Hence, the computation time that is needed to execute a single ADMM iteration increases. On the other hand, the data error decreases faster as ADMM proceeds, as can be observed in Figure 5.16a. Moreover, if ρ is small primal infeasibility is less penalized. It can be seen in Figure 5.16b that the smaller ρ is chosen, the slower the primal residual norm decreases in the first few ADMM iteration steps. Dual feasibility by definition becomes small for smaller ρ, as illustrated by Figure 5.16c. From this property, as for instance described in Boyd et al. (2011), strategies can be derived which adapt ρ in each iteration to keep primal and dual residual norm in a certain range relative to each other to improve convergence of the ADMM algorithm.

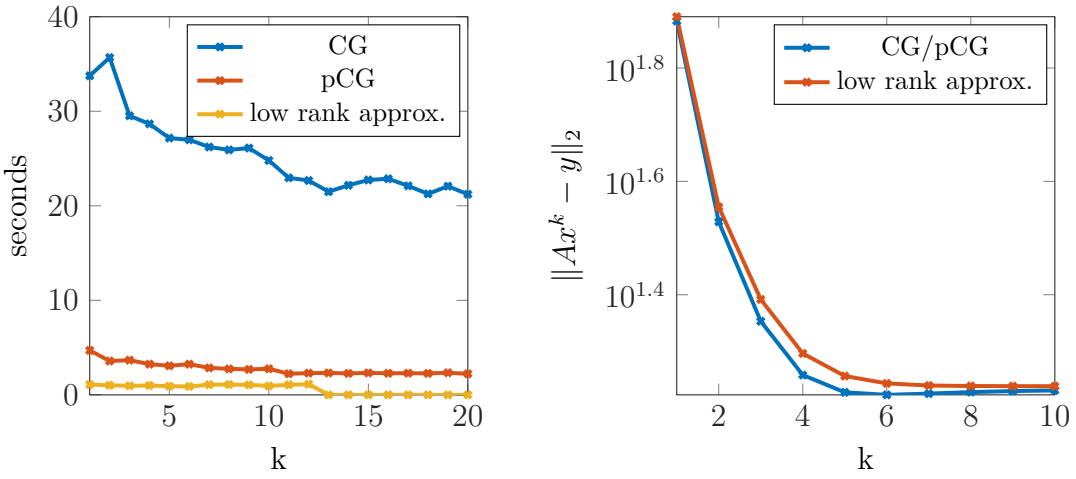
Different x-minimization strategies In Section 5.6, we considered strategies to perform the x-minimization step of the proposed algorithm more efficiently by using truncated SVDs. Note that, although time consuming, the initial computation of SVDs of the block matrices of $AD^{\frac{1}{2}}$ can be seen as a preprocessing step if a good choice for ρ^{IC} and ρ^{ISO} is known, and the resulting matrix can be



(a) data error **(b)** primal residual norm **(c)** dual residual norm

Figure 5.16.: Effects of changing parameter $\rho = \rho^{\mathrm{IC}} = \rho^{\mathrm{ISO}}$.

saved, and loaded before starting ADMM. To analyze performance, 20 ADMM iterations with $\rho^{\mathrm{IC}} = \rho^{\mathrm{ISO}} = 1$ and $\alpha = \gamma = 0.1$ were executed with the following x-minimization methods, respectively: CG without preconditioning, CG using the inverse stated in Equation (5.25) with $k_i = 1$ for all $i = 1, \ldots, n_v$ as preconditioner, and the low rank approximation approach using the inverse stated in Equation (5.25) where the SVDs of matrices A_i are truncated according to a threshold $c = 10$ on the condition (compare the numerical experiments section at the end of Section 5.6.2), respectively. The CG algorithm stops if the relative residual is smaller than 10^{-6}. In Figure 5.17, we compare computation time and data error for the different approaches. Preconditioning yields a great benefit with respect to computation time compared to the unconditioned CG method. In each ADMM iteration, the standard version needs on average about 15 times more CG iterations than the preconditioned one. The low rank approximation method is fastest but less accurate regarding the data error. Note, that for the low rank approximation method between one and four singular values were included in the truncated SVDs of the matrices A_i defined in (5.23). Increasing the number of singular values used to compute the inverse representation in (5.25) yields more accurate results at the cost of a higher computation time.



(a) computation time **(b)** data error

Figure 5.17.: Comparison of different computation methods for the x-minimization step of the ADMM algorithm.

6. Summary

The aim of this thesis was to develop ODF-based tractography algorithms which are more robust in the presence of noise and artifacts. Furthermore, we investigated possibilities for further reducing the number of incorrect tracks using a global framework.

Semi-local tractography methods can be regarded as a compromise between reliable but computationally expensive global methods, and fast but noise susceptible local methods. In Chapter 4, we have proposed semi-local strategies based on ODF data where streamline fiber tracking is informed by the local neighborhood to improve performance in the presence of noise and partial voluming. Instead of only incorporating information about the previous tracking direction in the streamlining process, the *guiding by extrapolation* approach seeks to include curvature information of the subsequently tracked fiber fragment. The *forward search* strategy embeds this guidance and, furthermore, allows for the exploration of the region that lies ahead. Our algorithms have been derived from a Bayesian paradigm which has always turned out to be superior to ad hoc solutions. Qualitative results on simulated and diffusion phantom data demonstrate that the proposed method can keep up with state-of-the-art higher order methods, and shows clear advantages compared to a purely local streamline method. Quantitative experiments show, that by using the proposed algorithm, the number of tracks terminating before reaching the designated regions simulating cortical gray matter can be drastically reduced. Accordingly, the number of valid connections increases. Benefits can also be observed in experiments on *in vivo* data, when compared to a simple local streamline method.

To filter out invalid tracks resulting from local tractography methods, we use a global preprocessing approach. In Chapter 5, we extend the method introduced in Daducci et al. (2015) by additional Sobolev-norm regularization to promote spatial smoothness of weights in the solution vector that correspond to image voxels. We developed and implemented algorithms for the efficient solution of the resulting non-trivial, large-scale minimization problem for both the solely ℓ_1- as well as the ℓ_1- and H^s-norm regularized formulation using the alternating direction method of multipliers. For the proposed algorithm, residual, objective, and dual variable convergence hold. A low rank approximation strategy which

takes advantage of the special structure of the dictionary matrix A by computing truncated SVDs of block matrices of A speeds up computation of the expensive *x-minimization* step of the ADMM algorithm. The low rank approximation can either be used in a direct optimization procedure or, if the problem size is very large, it can be used as a preconditioner for the conjugate gradient method which leads to a significant increase in computational efficiency. Qualitative results on real data show the applicability of the algorithm to large *in vivo* data sets. Furthermore, quantitative results for diffusion phantom data with known ground truth clearly show the benefits of the proposed method when compared to the results of the solely ℓ_1-regularized problem and the method provided with the COMMIT software package by Daducci et al. (2015).

An interesting starting point for future work would be a more extensive testing of the proposed algorithms on *in vivo* data sets obtained with clinically realistic acquisition parameters. These tests could serve to derive stable default parameter values for the algorithms to improve user-friendliness for clinical applications. To further improve usability for practical applications, both semi-local and global algorithms can in part be designed to be executed in parallel.

A. Anatomical Terms of Location

Throughout this thesis, the following locational terminology is used to describe, for instance, the position of image slices or the course of fiber bundles in the brain.

Important directional terms are:

- *superior* (= above) and *inferior* (= below)
- *anterior* (= in front) and *posterior* (= to the back)

The three image planes, illustrated in Figure A.1, are defined as follows:

- *sagittal* plane: the $y-z$ plane, separating left from right
- *coronal* plane: the $x-y$ plane, separating anterior from posterior
- *transverse* or *axial* plane: the $x-z$ plane, separating superior from inferior

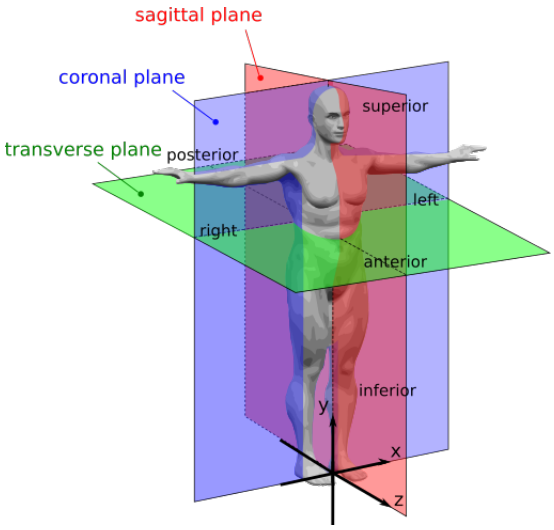


Figure A.1.: Anatomical planes in the human body. (Image adapted from commons.wikimedia.org.)

B. Green's Functions

The *Green's function* can be regarded as a tool for the solution of linear differential equations. In Section 5.3.1 we examine the difference between the Green's function of the discretized modified Helmholtz operator $(\lambda^2 I - \Delta)$ on a domain with Neumann boundary conditions, and the *fundamental solution*, also called *free* Green's function, which does not accept boundary conditions. Below, we define the Green's function and briefly motivate its use. Green's functions and their applications are for example covered in the books by Roach (1970) and Arfken (1968).

Let L be a linear elliptic differential operator on $\mathbb{R}^n$. A function G defined on

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x \neq y\}$$

is called fundamental solution or free Green's function if

$$LG(\cdot, y) = \delta_y$$

where δ_y is the Dirac delta distribution at y. This characterization can be used to solve differential equations of the form

$$Lu(x) = f(x), \tag{B.1}$$

for $x \in \mathbb{R}^n$. In fact, u can be expressed by

$$u(x) = \int_{\mathbb{R}^n} G(x, y) f(y) dy. \tag{B.2}$$

Hence, Equation (B.1) can be solved using formulation (B.2) if a Green's function is known. For instance, consider the homogeneous modified Helmholtz equation

$$(\lambda^2 I - \Delta)u = 0$$

with $\lambda > 0$. In two dimensions

$$G(x, y) = K_0(\lambda|x - y|),$$

and in three dimensions

$$G(x, y) = K_{1/2}(\lambda|x - y|),$$

where K_α is the modified bessel function of the second kind of order α (see Arfken (1968, Ch. 11, Ch. 15)).

Let us now consider a boundary value problem of the form

$$\begin{aligned} Lu(x) &= f(x) \quad && \text{for } x \in \Omega \\ Du(x) &= 0 && \text{for } x \in \partial\Omega \end{aligned} \tag{B.3}$$

with a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ and a boundary condition operator D. A function G on

$$\{(x, y) \in \Omega^2 : x \neq y\}$$

is called Green's function for the boundary value problem (B.3) if $G(\cdot, y)$ solves (B.3) with $f = \delta_y$ for all $y \in \Omega$. Similar to above, the solution of (B.3) can be formulated as

$$u(x) = \int_\Omega G(x, y) f(y) dy.$$

C. Concepts from Convex Optimization

In this section, we summarize fundamentals of convex optimization arising in the theory of the alternating direction method of multipliers in Section 5.4, and in its application to our optimization problem regarded in Section 5.5. The contents of this section are based on the books by Rockafellar (1970), Boyd and Vandenberghe (2004), and Hiriart-Urruty and Lemarechal (1993).

Convex sets and convex functions

Definition C.1. (convex set)
A set $C \subset \mathbb{R}^n$ is *convex* if

$$\lambda x + (1 - \lambda)y \in C$$

for any $x, y \in C$ and $0 < \lambda < 1$.

Example C.2. $\emptyset$ and $\mathbb{R}^n$ are convex sets.

Definition C.3. (proper function)
A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is called *proper* if it is not identical to ∞.

Definition C.4. (convex function)
A proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is *convex* if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

for all $x, y \in \mathbb{R}^n$ and all $0 \leq \lambda \leq 1$.

Epigraph

Definition C.5. (epigraph)
For a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, the set

$$\text{epi}f := \{(x, \mu) | \mu \in \mathbb{R}, x \in \text{dom}(f), f(x) \leq \mu\} \subset \mathbb{R}^{n+1}$$

defines the *epigraph* of f.

Theorem C.6. *A proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex if and only if* epi$f$ *is convex in $\mathbb{R}^n \times \mathbb{R}$.*

Proof. See e.g. Hiriart-Urruty and Lemarechal (1993, Ch. IV, Sec. 1, Proposition 1.1.6). □

Lower semi-continuity

Definition C.7. (lower semi-continuous function)
A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is *lower semi-continuous* if

$$\liminf_{y \to x} f(y) \geq f(x)$$

for all $x \in \mathbb{R}^n$.

Theorem C.8. *A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is lower semi-continuous if and only if* epi$f$ *is closed in $\mathbb{R}^n \times \mathbb{R}$.*

Proof. See e.g. Hiriart-Urruty and Lemarechal (1993, Ch. IV, Sec. 1, Proposition 1.2.1). □

Lemma C.9. *Let $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex and lower semi-continuous functions, and $\alpha > 0$. Then*

- *αf is convex and lower semi-continuous, and*
- *$f + g$ is convex and lower semi-continuous.*

Proof. See e.g. Hiriart-Urruty and Lemarechal (1993, Ch. IV, Sec. 2, Proposition 2.1.1). □

Example C.10.

- Norms on $\mathbb{R}^n$ are convex and lower semi-continuous.
- Quadratic functions $f : \mathbb{R}^n \to \mathbb{R}$ of the form $f(x) = \frac{1}{2}x^\top A^\top x + d^\top x + c$ with symmetric positive semi-definite linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$ are convex and lower semi-continuous.
- The indicator function of a (nonempty) closed and convex set C,

$$\tilde{\mathcal{I}}_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{else,} \end{cases}$$

  is convex and lower semi-continuous.

See e.g. Hiriart-Urruty and Lemarechal (1993, Ch. IV, Sec. 1.3) for more details.

Subdifferentiability

The concept of subdifferentiability generalizes the gradient to convex functions which are not differentiable.

Definition C.11. (subdifferential)
$g \in \mathbb{R}^n$ is called a *subgradient* of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ at x_0 if

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

for all $x \in \mathbb{R}^n$. The *subdifferential* $\partial f(x_0)$ denotes the set of all subgradients of f at x_0.

Example C.12. The subdifferential of $f(x) = |x|$ is given by

$$\partial f(x_0) = \begin{cases} \{-1\} & \text{at } x_0 < 0 \\ [-1, 1] & \text{at } x_0 = 0 \\ \{1\} & \text{at } x_0 > 0. \end{cases}$$

Existence of a solution

Consider a convex, lower semi-continuous and proper function $f : \mathcal{C} \to \mathbb{R}$ where $\mathcal{C} \subseteq \mathbb{R}^n$ is nonempty, convex and closed. We are interested in the existence of a

solution of the general minimization problem

$$\operatorname*{Inf}_{x\in\mathcal{C}} f(x). \tag{C.1}$$

Therefore, we consider the definition of a coercive function below:

Definition C.13. A function $f : \mathcal{C} \to \mathbb{R}$ is called *coercive* if

$$\lim_{\substack{\|x\|\to\infty \\ x\in\mathcal{C}}} f(x) = \infty.$$

Theorem C.14. *If $f : \mathcal{C} \to \mathbb{R}$ is coercive, then the minimization problem* (C.1) *has at least one solution.*

Proof. See e.g. Ekeland and Temam (1976, Ch. II, Sec. 1, Proposition 1.2). □

Lagrange-Duality

Definition C.15. (saddle point)
A point $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ is called a saddle point of a function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ if

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*)$$

for all $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$.

Definition C.16. (primal and dual problem)
As above, let $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and define the functions

$$\begin{aligned} F(x) &:= \sup_{\lambda\in\mathbb{R}^m} \mathcal{L}(x, \lambda), \quad x \in \mathbb{R}^n \\ G(\lambda) &:= \inf_{x\in\mathbb{R}^m} \mathcal{L}(x, \lambda), \quad \lambda \in \mathbb{R}^m. \end{aligned}$$

The optimization problem

$$\min_{x\in\mathbb{R}^n} F(x) \tag{C.2}$$

is called *primal problem*, while

$$\max_{\lambda\in\mathbb{R}^m} G(\lambda) \tag{C.3}$$

is the corresponding *dual problem*. Furthermore, let

$$\alpha := \inf_{x \in \mathbb{R}^n} F(x)$$
$$\beta := \sup_{\lambda \in \mathbb{R}^m} G(\lambda)$$

denote the optimal values of (C.2) and (C.3), respectively.

Theorem C.17. *(weak duality)*
The optimal value of the dual problem is a lower bound on the optimal value of the primal problem, that is $\beta \leq \alpha$.

Proof. See e.g. Geiger and Kanzow (2002, Ch. 6, Theorem 6.10). □

Theorem C.18. *(strong duality)*
(x^*, λ^*) *is a saddle point of* $\mathcal{L}$ *if and only if* x^* *is a primal optimal point (i.e. a minimizer of the primal problem),* λ^* *is a dual optimal point (i.e. a maximizer of the dual problem), and* $\alpha = \beta$.

Proof. See e.g. Hiriart-Urruty and Lemarechal (1993, Ch. VII, Sec. 4, Theorem 4.2.5). □

Let $f : \mathcal{C} \to \mathbb{R}$ be a function and consider the general minimization problem with linear inequality and equality constraints:

$$\begin{aligned} \min_{x \in \mathcal{C}} \; & f(x) \\ \text{subject to} \quad & A_{\mathcal{I}} x - b_{\mathcal{I}} \leq 0 \\ & A_{\mathcal{E}} x - b_{\mathcal{E}} = 0 \end{aligned} \tag{C.4}$$

with $A_{\mathcal{I}} \in \mathbb{R}^{m_{\mathcal{I}} \times n}$, $A_{\mathcal{E}} \in \mathbb{R}^{m_{\mathcal{E}} \times n}$, $b_{\mathcal{I}} \in \mathbb{R}^{m_{\mathcal{I}}}$, $b_{\mathcal{E}} \in \mathbb{R}^{m_{\mathcal{E}}}$. The *Lagrange function* corresponding to problem (C.4) is defined by

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \mu^\top (A_{\mathcal{I}} x - b_{\mathcal{I}}) + \lambda^\top (A_{\mathcal{E}} x - b_{\mathcal{E}}) \tag{C.5}$$

where $\mu \in \mathbb{R}^{m_{\mathcal{I}}}$ and $\lambda \in \mathbb{R}^{m_{\mathcal{E}}}$ are called Lagrange multipliers.

Theorem C.19. *Let function* f *be convex. Then* (x^*, μ^*, λ^*) *is a saddle point of* $\mathcal{L}$ *in* (C.5) *if and only if* x^* *is an optimal point of minimization problem* (C.4)*, and* μ^* *and* λ^* *are Lagrange multipliers.*

Proof. See e.g. Hiriart-Urruty and Lemarechal (1993, Ch. VII, Sec. 4, Theorem 4.4.3). □

D. Proof of Theorem 5.7

Recall that minimization problem (5.14) is given by

$$\underset{x\geq 0}{\operatorname{argmin}} \frac{1}{2}\left\|Ax-y\right\|_2^2+\alpha\left\|x^{\mathrm{IC}}\right\|_1+\sum_{j=1}^{n_b}\frac{\beta_j}{2}\left\|x_i^{\mathrm{EC}}\right\|_{H^s}^2+\frac{\gamma}{2}\left\|x^{\mathrm{ISO}}\right\|_{H^s}^2 \tag{D.1}$$

where $\alpha,\beta_j,\gamma>0$, and that in ADMM form the problem is written as

$$\begin{aligned}\underset{x,z}{\operatorname{argmin}} \frac{1}{2}\|Ax-y\|_2^2+\underbrace{\alpha\left\|z^{\mathrm{IC}}\right\|_1}_{g_1(z^{\mathrm{IC}})}+\underbrace{\sum_{j=1}^{n_b}\frac{\beta_j}{2}\left\|z_j^{\mathrm{EC}}\right\|_{H^s}^2}_{g_2\left(z_1^{\mathrm{EC}}\right)+\ldots+g_{n_b+1}\left(z_{n_b}^{\mathrm{EC}}\right)}+\underbrace{\frac{\gamma}{2}\left\|z^{\mathrm{ISO}}\right\|_{H^s}^2}_{g_{n_b+2}(z^{\mathrm{ISO}})}+\mathcal{I}_+(z)\\ \text{subject to } x=z.\end{aligned} \tag{D.2}$$

To show that this problem satisfies the assumptions for the ADMM convergence results in Theorem 5.4, we have to verify

(1) $f(x)=\frac{1}{2}\|Ax-y\|_2^2$ and $g(z)=\alpha\left\|z^{\mathrm{IC}}\right\|_1+\sum_{j=1}^{n_b}\frac{\beta_j}{2}\left\|z_j^{\mathrm{EC}}\right\|_{H^s}^2+\frac{\gamma}{2}\left\|z^{\mathrm{ISO}}\right\|_{H^s}^2+\mathcal{I}_+(z)$ are lower semi-continuous, proper, and convex functions.

(2) $\mathcal{L}_0(x,z,u)=f(x)+g(z)+\lambda^\top(x-z)$ has a saddle point.

Proof of (1) Function

$$f(x)=\frac{1}{2}\|Ax-y\|_2^2=\frac{1}{2}x^\top A^\top Ax-y^\top Ax+\frac{1}{2}y^\top y$$

and the squared Sobolev norm terms of function g, having the form

$$\|z\|_{H^s}^2=\|(\Delta+I)^{s/2}\|_2^2=z^\top(\Delta+I)^s z,$$

are convex and lower semi-continuous according to Example C.10. Being a norm on $\mathbb{R}^{n_f}$, the first term of function g, $g_1\left(z^{\mathrm{IC}}\right)=\|x\|_1$, is also convex and lower semi-continuous (see e.g. Example C.10). Furthermore, the indicator function is $\mathcal{I}_+(z)$ is convex and lower semi-continuous (see e.g. Example C.10). Hence,

function g is convex and lower semi-continuous by Theorem C.9. Clearly, functions f and g are proper, as both attain values different from ∞ for any $x \in \mathbb{R}$ and $z \in \mathbb{R}_{\geq 0}$.

Proof of (2) To show that $\mathcal{L}_0(x, z, \lambda)$ has a saddle point, we first show the existence of a solution of the corresponding minimization problem. We consider the minimization problem formulation (D.1) and define by

$$F(x) := \frac{1}{2} \|Ax - y\|_2^2 + \alpha \left\|x^{\mathrm{IC}}\right\|_1 + \sum_{j=1}^{n_b} \frac{\beta_j}{2} \left\|x_j^{\mathrm{EC}}\right\|_{H^s}^2 + \frac{\gamma}{2} \left\|x^{\mathrm{ISO}}\right\|_{H^s}^2$$

the objective function of (D.1). From above we know that F is proper, convex, and lower semi-continuous. Moreover, the set of feasible points of minimization problem (D.1), $\mathcal{C} := \mathbb{R}^n_{\geq 0}$, is convex, closed and nonempty. We notice that

$$\begin{aligned} F(x) &\geq \alpha \left\|x^{\mathrm{IC}}\right\|_1 + \sum_{j=1}^{n_b} \frac{\beta_j}{2} \left\|x_j^{\mathrm{EC}}\right\|_{H^s}^2 + \frac{\gamma}{2} \left\|x^{\mathrm{ISO}}\right\|_{H^s}^2 \\ &\geq c \left(\left\|x^{\mathrm{IC}}\right\|_1 + \left\| \left(\left(x^{\mathrm{EC}}\right)^\top, \left(x^{\mathrm{ISO}}\right)^\top \right) \right\|_{H^s}^2 \right) \end{aligned} \tag{D.3}$$

with $c = \min\left\{\alpha, \frac{\beta_j}{2}, \frac{\gamma}{2}\right\}$. To show that $F : \mathcal{C} \to \mathbb{R}$ is coercive, we consider a sequence (x^ℓ) with $\|x^\ell\| \to \infty$ and show that $F(x^\ell) \to \infty$ as $\ell \to \infty$. For the sake of clearness, we define $x_1 := x^{\mathrm{IC}}$ and $x_2 := \left(\left(x^{\mathrm{EC}}\right)^\top, \left(x^{\mathrm{ISO}}\right)^\top \right)$. Using triangle inequality and equivalence of all norms on $\mathbb{R}^n$ it can be shown that

$$\|x^\ell\| = \left\| \begin{pmatrix} x_1^\ell \\ x_2^\ell \end{pmatrix} \right\| \leq C \left(\|x_1^\ell\| + \left\|x_2^\ell\right\| \right)$$

with a real constant $C > 0$. Hence, at least one of the two norms $\left\|x_1^\ell\right\|$ or $\left\|x_2^\ell\right\|$ diverges to ∞.

First, we consider the case that $\left\|x_1^\ell\right\| \to \infty$. From (D.3) we deduce that

$$F(x^\ell) \geq c \left\|x_1^\ell\right\|_1 .$$

As all norms are equivalent on $\mathbb{R}^n$, it follows that

$$F(x^\ell) \geq c \left\|x_1^\ell\right\|_1 \to \infty \text{ as } \ell \to \infty.$$

Now, consider the second case, i.e. $\left\|x_2^\ell\right\| \to \infty$. As $\left\|x_2^\ell\right\| \to \infty$, $\left\|x_2^\ell\right\| \geq 1$ for all elements of the sequence above a certain element. That is, we can write

$$F(x^\ell) \geq c \left\|x_2^\ell\right\|_{H^s}^2 \geq c \left\|x_2^\ell\right\|_{H^s}.$$

Similar to above, we obtain

$$F(x^\ell) \geq c \left\|x_2^\ell\right\|_{H^s} \to \infty \text{ as } \ell \to \infty.$$

Hence, $F(x^\ell) \to \infty$ as $\ell \to \infty$ which means that $F : \mathcal{C} \to \mathbb{R}$ is coercive. With Theorem C.14 it follows that minimization problem (D.1) has at least one solution.

Now, let x^* be a solution of (D.1). Then (x^*, z^*) where $x^* = z^*$ is a solution of (D.2). Hence, the objective of (D.2) is finite at (x^*, z^*). Recall that (D.2) is a convex minimization problem with linear equality constraints. Furthermore, let λ^* denote a Lagrange multiplier. Then, Theorem C.19 implies that (x^*, z^*, λ^*) is a saddle point of the unaugmented Lagrangian $\mathcal{L}_0$.

List of Abbreviations

ADC	Apparent Diffusion Coefficient
ADMM	Alternating Direction Method of Multipliers
BPDN	Basis Pursuit Denoise
CG/pCG	(preconditioned) Conjugate Gradients
CSD	Constrained Spherical Deconvolution
CSF	Cerebrospinal Fluid
CSR	Connection to Seed Ratio
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Imaging
EC	Extra Cellular (as superscript)
FA	Fractional Anisotropy
GFA	General Fractional Anisotropy
GM	Gray Matter
HARDI	High Angular Resolution Diffusion Imaging
HCP	Human Connectome Project
IB	Invalid Bundles
IC	Invalid Connections

IC	Intra Cellular (as superscript)
ISO	Isotropic (as superscript)
MRI	Magnetic Resonance Imaging
NC	No Connections
NNLS	Nonnegative Least Squares
NRMSE	Normalized Root Mean Squared Error
ODF/fODF	(fiber) Orientation Distribution Function
PSNR	Peak Signal to Noise Ratio
RF	Radio Frequency
RMSE	Root Mean Squared Error
ROI	Region Of Interest
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
VB	Valid Bundles
VC	Valid Connections
VCCR	Valid Connection to Connection Ratio
WM	White Matter

Bibliography

Alexander, D.C., 2008. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. Magn. Reson. Med. 60, 439–448.

Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. NeuroImage 20, 870–888.

Arfken, G., 1968. Mathematical Methods for Physicists. Academic Press.

Basser, P.J., Mattiello, J., LeBihan, D., 1994. MR Diffusion Tensor Spectroscopy and Imaging. Biophys. J. 66, 259–267.

Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. 44, 625–632.

Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system–a technical review. NMR Biomed. 15, 435–455.

Behrens, T.E.J., Johansen Berg, H., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage 34, 144–155.

Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088.

Berman, J.I., Chung, S.W., Mukherjee, P., Hess, C.P., Han, E.T., Henry, R.G., 2008. Probabilistic streamline q-ball tractography using the residual bootstrap. NeuroImage 39, 215–222.

Björnemo, M., Brun, A., Kikinis, R., Westin, C.F., 2002. Regularized stochastic

white matter tractography using diffusion tensor MRI, in: Proc. Med. Image. Comput. Comput. Assist. Interv., Springer. pp. 435–442.

Bloch, F., 1946. Nuclear induction. Phys. Rev. 70, 460.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 3, 1–122.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.

Campbell, J.S.W., Siddiqi, K., Rymar, V.V., Sadikot, A.F., Pike, G.B., 2005. Flow-based fiber tracking with diffusion tensor and q-ball data: Validation and comparison to principal diffusion direction techniques. NeuroImage 27, 725–736.

Chao, Y.P., Chen, J.H., Cho, K.H., Yeh, C.H., Chou, K.H., Lin, C.P., 2008. A multiple streamline approach to high angular resolution diffusion tractography. Med. Eng. Phys. 30, 989–996.

Ciak, R., Shafei, B., Steidl, G., 2013. Homogeneous penalizers and constraints in convex image restoration. J. Math. Imaging Vis. 47, 210–230.

Close, T.G., Tournier, J.D., Calamante, F., Johnston, L.A., Mareels, I., Connelly, A., 2009. A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. NeuroImage 47, 1288–1300.

Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E., 1999. Tracking neuronal fiber pathways in the living human brain. Proc. Natl. Acad. Sci. USA 96, 10422–10427.

Côté, M.A., Girard, G., Boré, A., Garyfallidis, E., Houde, J.C., Descoteaux, M., 2013. Tractometer: Towards validation of tractography pipelines. Med. Image Anal. 17, 844–857.

Daducci, A., Dal Palu, A., Lemkaddem, A., Thiran, J.P., 2013. A convex optimization framework for global tractography, in: Proc. 10th Intl. Symposium on Biomed. Imaging, pp. 524–527.

Daducci, A., Dal Palù, A., Lemkaddem, A., Thiran, J.P., 2015. COMMIT: Con-

vex optimization modeling for microstructure informed tractography. IEEE T. Med. Imaging 34, 246–257.

Descoteaux, M., Deriche, R., Knösche, T.R., Anwander, A., 2009. Deterministic and probabilistic tractography based on complex fibre orientation distributions. IEEE T. Med. Imaging 28, 269–286.

Dhollander, T., Emsell, L., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P., 2014. Track orientation density imaging (TODI) and track orientation distribution (TOD) based tractography. NeuroImage 94, 312–336.

Ekeland, I., Temam, R., 1976. Convex Analysis and Variational Problems. North-Holland Publishing Company.

Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., Reisert, M., Sakaie, K., Tensaouti, F., Yo, T., Mangin, J.F., Poupon, C., 2011. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. NeuroImage 56, 220–234.

Fillard, P., Poupon, C., Mangin, J.F., 2009. A novel global tractography algorithm based on an adaptive spin glass model, in: Proc. Med. Image. Comput. Comput. Assist. Interv., Springer. pp. 927–934.

Friman, O., Farnebäck, G., Westin, C.F., 2006. A Bayesian Approach for Stochastic White Matter Tractography. IEEE T. Med. Imaging 25, 965–978.

Geiger, C., Kanzow, C., 2002. Theorie und Numerik restringierter Optimierungsaufgaben. Springer.

Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M., 2014. Towards quantitative connectivity analysis: Reducing tractography biases. NeuroImage 98, 266–278.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al., 2013. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 80, 105–124.

Golub, G.H., van Loan, C.F., 1983. Matrix Computations. The Johns Hopkins University Press.

Haacke, E.M., Brown, R.W., Thompson, M.R., Venkatesan, R., 1999. Magnetic

Resonance Imaging: Physical Principles and Sequence Design. John Wiley and Sons.

Hagmann, P., Thiran, J.P., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., Meuli, R., 2003. DTI mapping of human brain connectivity: Statistical fibre tracking and virtual dissection. NeuroImage 19, 545–554.

Hanke-Bourgeois, M., 2009. Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens. Vieweg+Teubner.

Hiriart-Urruty, J.B., Lemarechal, C., 1993. Convex Analysis and Minimization Algorithms I: Fundamentals (Grundlehren Der Mathematischen Wissenschaften). Springer.

Hohage, T., Rügge, C., 2015. A coherence enhancing penalty for diffusion MRI: Regularizing property and discrete approximation. SIAM J. Imaging Sci. 8, 1874–1893.

Jansons, K.M., Alexander, D.C., 2003. Persistent angular structure: New insights from diffusion magnetic resonance imaging data. Inverse Probl. 19, 1031.

Jbabdi, S., Bellec, P., Toro, R., Daunizeau, J., Pélégrini-Issac, M., Benali, H., 2008. Accurate anisotropic fast marching for diffusion-based geodesic tractography. J. Biomed. Imaging 2008, 2.

Jeurissen, B., Leemans, A., Jones, D.K., Tournier, J.D., Sijbers, J., 2011. Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. Hum. Brain Mapp. 32, 461–479.

Jones, D.K., 2008. Tractography gone wild: Probabilistic fibre tracking using the wild bootstrap with diffusion tensor MRI. IEEE T. Med. Imaging 27, 1268–1274.

Kaden, E., Knösche, T.R., Anwander, A., 2007. Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion MR imaging. NeuroImage 37, 474–488.

Koch, M.A., Norris, D.G., Hund-Georgiadis, M., 2002. An investigation of functional and anatomical connectivity using magnetic resonance imaging. NeuroImage 16, 241–250.

Kreher, B.W., Mader, I., Kiselev, V.G., 2008. Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. Magn. Reson. Med. 60, 953–963.

Lazar, M., Alexander, A.L., 2002. White matter tractography using randomvector (RAVE) perturbation, in: Proc. Intl. Soc. Mag. Reson. Med. 10, p. 539.

Lazar, M., Alexander, A.L., 2005. Bootstrap white matter tractography (BOOTTRAC). NeuroImage 24, 524–532.

Le Bihan, D., Breton, E., 1985. Imagerie de diffusion in-vivo par résonance magnétique nucléaire. C. R. Acad. Sci. 93, 27–34.

McGraw, T., Vemuri, B.C., Chen, Y., Rao, M., Mareci, T., 2004. DT-MRI denoising and neuronal fiber tracking. Med. Image Anal. 8, 95–111.

Merboldt, K.D., Hanicke, W., Frahm, J., 1985. Self-diffusion NMR imaging using stimulated echoes. J. Magn. Reson. (1969) 64, 479–486.

Milchenko, M., Marcus, D., 2013. Obscuring surface anatomy in volumetric imaging data. Neuroinformatics 11, 65–75.

Mori, S., 2007. Introduction to Diffusion Tensor Imaging. Elsevier.

Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C.M., 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. 45, 265–269.

Moussavi-Biugui, A., Stieltjes, B., Fritzsche, K., Semmler, W., Laun, F.B., 2011. Novel spherical phantoms for q-ball imaging under in vivo conditions. Magn. Reson. Med. 65, 190–194.

Neher, P.F., Laun, F.B., Stieltjes, B., Maier-Hein, K.H., 2014. Fiberfox: Facilitating the creation of realistic white matter software phantoms. Magn. Reson. Med. 72, 1460–1470.

Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C., 2012. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. NeuroImage 59, 2241–2254.

Parker, G.J.M., Alexander, D.C., 2005. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. Philos. T. R. Soc. B 360, 893–902.

Parker, G.J.M., Haroon, H.A., Wheeler-Kingshott, C.A.M., 2003. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. J. Magn. Reson. Im. 18, 242–254.

Parker, G.J.M., Wheeler-Kingshott, C.A.M., Barker, G.J., 2002. Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. IEEE T. Med. Imaging 21, 505–513.

Perrin, M., Poupon, C., Cointepas, Y., Rieul, B., Golestani, N., Pallier, C., Rivière, D., Constantinesco, A., Le Bihan, D., Mangin, J.F., 2005. Fiber tracking in q-ball fields using regularized particle trajectories, in: Proc. Inf. Process. Med. Imaging., Springer. pp. 52–63.

Pestilli, F., Yeatman, J.D., Rokem, A., Kay, K.N., Wandell, B.A., 2014. Evaluation and statistical inference for human connectomes. Nat. Methods 11, 1058–1063.

Pontabry, J., Rousseau, F., 2011. Probabilistic tractography using q-ball modeling and particle filtering, in: Proc. Med. Image. Comput. Comput. Assist. Interv., Springer. pp. 209–216.

Poupon, C., Laribiere, L., Tournier, G., Bernard, J., Fournier, D., Fillard, P., Descoteaux, M., Mangin, J.F., 2010. A diffusion hardware phantom looking like a coronal brain slice, in: Proc. Intl. Soc. Mag. Reson. Med. 18, p. 581.

Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.F., 2008. New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. Magn. Reson. Med. 60, 1276–1283.

Ramirez-Manzanares, A., Rivera, M., 2006. Basis Tensor Decomposition for Restoring Intra-Voxel Structure and Stochastic Walks for Inferring Brain Connectivity in DT-MRI. Int. J. Comput. Vision 69, 77–92.

Reisert, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., Kiselev, V., 2011. Global fiber reconstruction becomes practical. NeuroImage 54, 955–962.

Roach, G.F., 1970. Green's Function: Introductory Theory with Applications. Van Nostrand.

Rockafellar, R.T., 1970. Convex Analysis. Princeton Univ. Press.

Rowe, M., Zhang, H.G., Oxtoby, N., Alexander, D.C., 2013. Beyond crossing fibers: Tractography exploiting sub-voxel fibre dispersion and neighbourhood structure, in: Proc. Inf. Process. Med. Imaging., Springer. pp. 402–413.

Rügge, C., 2015. Spatial Coherence Enhancing Reconstructions for High Angular Resolution Diffusion MRI. Ph.D. thesis. Georg-August-Universität Göttingen.

Savadjiev, P., Campbell, J.S.W., Descoteaux, M., Deriche, R., Pike, G.B., Siddiqi, K., 2008. Labeling of ambiguous subvoxel fibre bundle configurations in high angular resolution diffusion MRI. NeuroImage 41, 58–68.

Schomburg, H., Hohage, T., 2017. Semi-local tractography strategies using neighborhood information. Med. Image Anal. 38, 165–183.

Schomburg, H., Rügge, C., Hofer, S., Frahm, J., Hohage, T., 2015. Probabilistic fiber tractography using neighborhood information, in: Proc. Intl. Soc. Mag. Reson. Med. 23, p. 2838.

Schomburg, H., Rügge, C., Hohage, T., 2014. A semi-local tractography approach using neighborhood information, in: Proc. Intl. Soc. Mag. Reson. Med. 22, p. 2668.

Slichter, C.P., 1990. Principles of Magnetic Resonance. Springer.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2012. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. NeuroImage 62, 1924–1938.

Sotiropoulos, S.N., Moeller, S., Jbabdi, S., Xu, J., Andersson, J.L., Auerbach, E.J., Yacoub, E., Feinberg, D., Setsompop, K., Wald, L.L., et al., 2013. Effects of image reconstruction on fiber orientation mapping from multichannel diffusion MRI: Reducing the noise floor using SENSE. Magn. Reson. Med. 70, 1682–1689.

Taylor, D.G., Bushell, M.C., 1985. The spatial mapping of translational diffusion coefficients by the NMR imaging technique. Phys. Med. Biol. 30, 345.

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

Tournier, J.D., Calamante, F., Connelly, A., 2010. Improved probabilistic

streamlines tractography by 2nd order integration over fibre orientation distributions, in: Proc. Intl. Soc. Mag. Reson. Med. 18, p. 1670.

Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A., 2003. Diffusion-weighted magnetic resonance imaging fibre tracking using a front evolution algorithm. NeuroImage 20, 276–288.

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

Tuch, D.S., 2004. Q-ball imaging. Magn. Reson. Med. 52, 1358–1372.

Twieg, D.B., 1983. The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. Med. Phys. 10, 610–621.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al., 2013. The WU-Minn Human Connectome Project: An overview. NeuroImage 80, 62–79.

Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54, 1377–1386.

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

Weinstein, D., Kindlmann, G., Lundberg, E., 1999. Tensorlines: Advection-diffusion based propagation through diffusion tensor fields, in: Proc. 10th IEEE Vis. Conf., IEEE Computer Society Press. pp. 249–253.

Westin, C.F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R., 2002. Processing and visualization for diffusion tensor MRI. Med. Image Anal. 6, 93–108.

Zhang, F., Hancock, E.R., Goodlett, C., Gerig, G., 2009. Probabilistic white matter fiber tracking using particle filtering and von Mises-Fisher sampling. Med. Image Anal. 13, 5–18.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through

a hidden Markov random field model and the expectation-maximization algorithm. IEEE T. Med. Imaging 20, 45–57.

Danksagung

An dieser Stelle möchte ich mich noch bei allen bedanken, die mich während meiner Promotionszeit in Göttingen unterstützt haben. Mein Dank gilt insbesondere meinem Betreuer Professor Thorsten Hohage für geduldige Erklärungen, Ratschläge, Diskussionen und Motivation. Außerdem möchte ich mich bei meinem Zweitbetreuer Professor Jens Frahm und den Mitgliedern seiner Biomedizinische NMR Forschungs GmbH, insbesondere bei Sabine Hofer und Andreas Merrem, für Zusammenarbeit, Erklärungen zu Biologie und Physik, und für die Bereitstellung von Datensätzen bedanken.

Mein Dank gilt auch den aktuellen und ehemaligen Mitgliedern der Arbeitsgruppe Inverse Probleme für die gute Arbeitsatmosphäre und viele interessante Gespräche. Besonders hervorgehoben sei hier Christoph, der stets ein offenes Ohr für meine Fragen hatte und mir geduldig sein Promotionsthema zur ODF Rekonstruktion näher brachte. Außerdem möchte ich mich bei Carolin, mit der ich mir fast drei Jahre ein Büro geteilt habe und mit deren Hilfe ich mich sehr schnell am Institut für Numerische und Angewandte Mathematik eingelebt habe, für die vielen Gespräche und natürlich für unser gemeinsames Schwimmtraining zum Ausklingen des Arbeitstages bedanken.

Mein besonderer Dank gilt meinen Eltern und Thomas, auf die ich mich stets verlassen kann.

Zuletzt möchte ich mich beim GRK 1023 für die finanzielle Unterstützung bedanken.