Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Position-orientation adaptive smoothing of diffusion weighted magnetic resonance data (POAS)

Saskia Becker[1], Karsten Tabelow[1] , Henning U. Voss[2], Alfred Anwander[3],

Robin M. Heidemann[3] , Jörg Polzehl[1]

submitted: December 8, 2011

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: saskia.becker@wias-berlin.de
E-Mail: joerg.polzehl@wias-berlin.de
E-Mail: karsten.tabelow@wias-berlin.de

[2] Citigroup Biomedical Imaging Center
Weill Cornell Medical College
516 East 72nd Street
New York, NY 10021
USA
E-Mail: hev2006@med.cornell.edu

[3] Max Planck Institute for Human Cognitive and Brain Sciences
Stephanstrasse 1a
04103 Leipzig
Germany
E-Mail: anwander@cbs.mpg.de
E-Mail: heidemann@cbs.mpg.de

No. 1668
Berlin 2011

2010 *Mathematics Subject Classification.* 62P10, 62G05, 22E99 .

Key words and phrases. Diffusion weighted magnetic resonance imaging, POAS, Structural adaptive smoothing, Special Euclidean motion group, Lie groups .

This work was partially supported by the DFG Research Center MATHEON “Mathematics for key technologies” in Berlin and the Stiftung der Deutschen Wirtschaft (SDW).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

ABSTRACT. We introduce an algorithm for diffusion weighted magnetic resonance imaging data enhancement based on structural adaptive smoothing in both space and diffusion direction. The method, called POAS, does not refer to a specific model for the data, like the diffusion tensor or higher order models. It works by embedding the measurement space into a space with defined metric and group operations, in this case the Lie group of three-dimensional Euclidean motion SE(3). Subsequently, pairwise comparisons of the values of the diffusion weighted signal are used for adaptation. The position-orientation adaptive smoothing preserves the edges of the observed fine and anisotropic structures. The POAS-algorithm is designed to reduce noise directly in the diffusion weighted images and consequently also to reduce bias and variability of quantities derived from the data for specific models. We evaluate the algorithm on simulated and experimental data and demonstrate that it can be used to reduce the number of applied diffusion gradients and hence acquisition time while achieving similar quality of data, or to improve the quality of data acquired in a clinically feasible scan time setting.

1. INTRODUCTION

Diffusion weighted magnetic resonance imaging (dMRI) has evolved into a versatile tool for in-vivo examination of micro-structure in the human brain [Le Bihan, 2003], spinal cord [Clark et al., 1999], or muscle tissue [Sinha et al., 2006]. The data is measured using the pulsed gradient spin echo sequence [PGSE, Stejskal and Tanner, 1965] through application of diffusion magnetic field gradients in different directions. The technique is based on the fact that nuclear magnetic resonance is sensitive to the diffusion of molecules, usually water. Each applied magnetic field gradient yields a diffusion weighted image which, together with the non-diffusion weighted image, reveals information about the average displacement of hydrogen protons in the corresponding direction at each voxel.

In the past two decades the development of diffusion tensor imaging (DTI) made dMRI more feasible for a wider range of clinical and neuroscience applications (see Mori [2007] and Johansen-Berg and Behrens [2009] for an introduction). In order to characterize the data by a diffusion tensor at least six gradient directions have to be measured. However, the diffusion tensor model provides only an approximation for the more realistic case of inhomogeneous and restricted diffusion, spurring recent interest in models beyond DTI [Assemlal et al., 2011]. Most of the methods developed to overcome the limitations of DTI require a larger number of diffusion weighted gradients to be applied during measurement and hence a longer acquisition time.

Like all other imaging modalities dMRI suffers from significant noise which may render subsequent analysis or medical decisions more difficult. For DTI it has been shown that noise induces a systematically biased assessment of features, like the fractional anisotropy (FA), which are based on the ordered eigenvalues of the estimated diffusion tensor [Pierpaoli and Basser, 1996, Basser and Pajevic, 2000]. Jones and Basser [2004] showed that noise creates artifacts for dMRI data in general. Last but not least, increasing the spatial resolution inherently decreases the signal-to-noise ratio (SNR). Attempts to achieve higher image resolution collide with the deterioration of the acquired images. Noise reduction is therefore essential.

Besides simple isotropic smoothing, which potentially blurs fine structures, more sophisticated methods for noise reduction have been developed, such as the Perona-Malik filter [Perona and Malik, 1990, Parker et al., 2000], non-linear diffusion approaches [Weickert, 1998, Ding et al., 2005, Duits and Franken, 2011], or the Propagation-Separation approach [Polzehl and Spokoiny, 2006, Tabelow et al., 2008]. Other attempts combine wavelet filtering with subsequent

non-linear diffusion [Lohmann et al., 2010], or perform smoothing in tensor space [Fletcher and Joshi, 2007, Fillard et al., 2007]. It is generally preferable to smooth the diffusion weighted images directly rather than model dependent derived quantities. First, such an approach avoids the bias for model dependent estimates induced by the noise. For DTI it also reduces the probability of estimating a degenerated tensor. Second, if the method does not refer to a specific model, data can be subsequently analyzed with any model.

In this paper we will develop a position-orientation adaptive smoothing algorithm (POAS) for dMRI, that *does not rely* on a specific model like the diffusion tensor or higher order models. This is a significant development from our recent structural adaptive smoothing method for DTI [Tabelow et al., 2008], which was based on modeling of the dMRI data by the diffusion tensor model. Furthermore, we will outline, how the specific geometry of the space in which dMRI is typically measured can be used in order to benefit from the whole information of the data, in position and orientation. Here, we will use results from a recent contribution by Duits and Franken [2011], who developed a smoothing algorithm based on the non-linear diffusion equation. Our method is based on the Propagation-Separation approach [Polzehl and Spokoiny, 2006] and hence, unlike other smoothing methods, includes an intrinsic stopping criterion. The new method POAS is able to reduce noise without blurring the edges of the structures observed with dMRI. We will show through extensive simulation and experimental data, how the position-orientation adaptive smoothing algorithm significantly improves the quality of the diffusion weighted data prior to estimation of diffusion tensors or higher order models.

2. Theory

In the first part of this section we want to discuss the geometry of the space in which dMRI is typically measured. We introduce a discrepancy function on this space following the approach in Duits and Franken [2011]. Then, we present our new position-orientation adaptive smoothing method (POAS) including further comments on the choice of parameter values.

2.1. **The underlying geometry.**

In diffusion weighted imaging data is typically acquired on a regular grid of voxels for varying directions of the diffusion magnetic field gradient at a fixed b-value[1]. The latter can be described by a unit vector in $\mathbb{S}^2$, where $\mathbb{S}^2 := \{\vec{u} \in \mathbb{R}^3 : \|\vec{u}\| = 1\}$ denotes the 2-sphere. Hence, dMRI data form a function $S : \mathbb{R}^3 \times \mathbb{S}^2 \to \mathbb{R}$.

In this section, we want to introduce a discrepancy function on $\mathbb{R}^3 \times \mathbb{S}^2$. On that space there exist neither a metric nor a group operation. So, following the approach in Duits and Franken [2011] we embed the space $\mathbb{R}^3 \times \mathbb{S}^2$ into the Lie group $\mathrm{SE}(3) := \mathbb{R}^3 \rtimes \mathrm{SO}(3)$ of 3 dimensional Euclidean motion, where $\mathrm{SO}(3)$ denotes the 3 dimensional rotation group and $\rtimes$ the semidirect product.

First, we remind of the concept of left coset spaces. The set

$$[g]_H := g \cdot_G H := \{g \cdot_G h : h \in H\}, \quad g \in G,$$

is called left coset of the group G w.r.t. the subgroup H. The space of all left cosets G/H forms a partition of G. Considering the parametrization of $\mathbb{S}^2$ and $\mathrm{SO}(3)$ as given in Appendix A.1, we note that each element $\vec{u}_{(\beta,\gamma)}$ of $\mathbb{S}^2$ depends on two angles β and γ, while the elements

[1] Acquisition schemes that require multiple q-shells or varying b-value are not considered in this paper. The methods proposed here may be applied separately for each b-value.

$\mathbf{R}_{(\alpha,\beta,\gamma)}$ of $\mathrm{SO}(3)$ additionally depend on a third angle α which describes the rotation around the z-axis. More precisely, for any $\vec{u} \in \mathbb{S}^2$ there is a matrix $\mathbf{R} \in \mathrm{SO}(3)$ such that $\mathbf{R}_{\vec{u}}\vec{e}_z = \vec{u}$. This observation allows us to uniquely identify $\mathbb{S}^2$ with the left coset space $\mathrm{SO}(3)/\mathrm{stab}(\vec{e}_z)$ of $\mathrm{SO}(3)$, which is defined as follows:

$$\mathrm{SO}(3)/\mathrm{stab}(\vec{e}_z) := \left\{ \left[\mathbf{R}_{(\alpha,\beta,\gamma)}\right] : \mathbf{R}_{(\alpha,\beta,\gamma)} \in \mathrm{SO}(3) \right\},$$

where $\mathrm{stab}(\vec{e}_z) \subseteq \mathrm{SO}(3)$ denotes the group of rotations around the z-axis and

$$\left[\mathbf{R}_{(\alpha,\beta,\gamma)}\right] := \left[\mathbf{R}_{(\alpha,\beta,\gamma)}\right]_{\mathrm{stab}(\vec{e}_z)} := \{\mathbf{R}_{(\alpha',\beta,\gamma)} : \alpha' \in [0, 2\pi)\}.$$

In other words, we construct equivalence classes $[.]$ by identifying all elements of $\mathrm{SO}(3)$ where the angles β and γ coincide while the angle α is ignored. Then, for any $\vec{u}_{(\beta,\gamma)} \in \mathbb{S}^2$ it holds

$$\vec{u}_{(\beta,\gamma)} \equiv \left[\mathbf{R}_{(\alpha,\beta,\gamma)}\right] \in \mathrm{SO}(3)/\mathrm{stab}(\vec{e}_z), \quad \forall\, \alpha \in [0, 2\pi).$$

Finally, it follows $\mathbb{R}^3 \times \mathbb{S}^2 \cong \mathrm{SE}(3)/(\{\vec{0}\} \rtimes \mathrm{stab}(\vec{e}_z))$ since $\mathbb{R}^3/\{\vec{0}\} = \mathbb{R}^3$. In the following we will write $\mathbb{R}^3 \rtimes \mathbb{S}^2$ instead of $\mathbb{R}^3 \times \mathbb{S}^2$ referring to the group structure which has been introduced on $\mathbb{R}^3 \times \mathbb{S}^2$ by the embedding into $\mathrm{SE}(3)$.

The Riemannian 1-norm relates to the length of the geodesic, i.e. the shortest way within $\mathrm{SE}(3)$, from the identity element e to g. Analogous to the well-known L^p-spaces, where L^2 is often preferred because of its technical advantages, we consider the Riemannian 2-norm instead of the Riemannian 1-norm. It can be interpreted as L^2-analog. In Duits [2005] the Riemannian 2-norm has been suggested as modulus function on $\mathrm{SE}(2)$. On $\mathrm{SE}(3)$ the Riemannian 2-norm of $g := (\vec{v}, \mathbf{R}_{(\alpha,\beta,\gamma)}) \in \mathrm{SE}(3)$ with $\vec{v} := (x, y, z)^T$ can be approximated locally by

$$\|g\|_R \approx \left(\sum_{i=1}^{6} k_i^2 \right)^{1/2}, \tag{1}$$

where the coordinates $\{k_i\}_{i=1}^6$ follow from

$$\prod_{i=1}^{6} \exp(k_i \mathbf{A}_i) = \begin{pmatrix} & & & x \\ & \mathbf{R}_{(\alpha,\beta,\gamma)} & & y \\ & & & z \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv g \in \mathrm{SE}(3)$$

using the matrix representation of $\mathrm{SE}(3)$. The left-invariant basis matrices $\{\mathbf{A}_i\}_{i=1}^6$ and details on the implementation are specified in Appendices A.2 and A.4.

The Riemannian 2-norm induces a discrepancy on $\mathrm{SE}(3)$ by

$$\Delta_{\mathrm{SE}(3)}(g_1, g_2) := \|g_2^{-1} \cdot_{\mathrm{SE}(3)} g_1\|_R, \quad g_1, g_2 \in \mathrm{SE}(3).$$

It can be applied to elements of $\mathbb{R}^3 \rtimes \mathbb{S}^2$ via the introduced embedding as we will discuss in Appendix A.3. However, a slight modification is recommendable since distances in voxel space and in orientation space are not intrinsically related by the measurement process: This is highlighted by different physical measurement units and the fact that there is no natural relation between the spatial distance and the applied magnetic field gradients. Therefore, we introduce

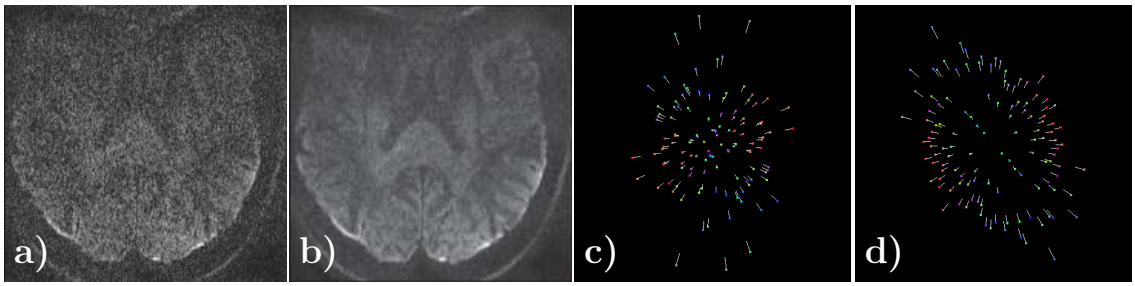

FIGURE 1. Diffusion weighted data S_g. a) Slice of diffusion weighted image for *a single* (arbitrarily selected) diffusion weighting gradient direction. b) Same slice shown after averaging over four subsequent measurements. (A larger version of these images can be found in Figure 7.) c) Diffusion weighted data for all diffusion gradients in a *single voxel* in corpus callosum. The image is created from a 3D plot of the data, where the diffusion weighted values are shown in their corresponding gradient direction. d) Same after averaging over four subsequent measurements.

an additional parameter κ which allows for a balance between distances on the sphere and in space. Finally, we get the following formula

$$\Delta_\kappa(g_1, g_2) := \| \left(\mathbf{R}_{\vec{u}_2}^{-1}(\vec{v}_1 - \vec{v}_2), \mathbf{R}_{\vec{u}_2}^{-1}\mathbf{R}_{\vec{u}_1}\right) \|_{R,\kappa} \approx \left(\sum_{i=1}^{3} k_i^2 + \sum_{i=4}^{6} k_i^2/\kappa^2 \right)^{1/2},$$

where $g_i = (\vec{v}_i, \vec{u}_i) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, $i = 1, 2$, and $\mathbf{R}_{\vec{u}_i} \in \mathrm{SO}(3)$ is any rotation with $\mathbf{R}_{\vec{u}_i}\mathbf{e}_z = \vec{u}_i$ as introduced above.

2.2. **Position-orientation adaptive smoothing (POAS) the diffusion weighted images.** Denote by S_g the measured diffusion weighted data at $g = (\vec{v}, \vec{b}) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, i.e., for some voxel at position $\vec{v}$ and for an applied diffusion weighting gradient in direction $\vec{b}$. The noise distribution of S_g follows a Rician distribution, which can be approximated by a Gaussian distribution for large values of the data [Johansen-Berg and Behrens, 2009].

In order to construct a structural adaptive smoothing algorithm for diffusion weighted data we first formulate a structural assumption on the data. This assumption represents the observation that the parameters of the distribution for the data show similarities in the vicinity of any voxel separated by sharp discontinuities, e.g., at tissue borders, see Figure 1. We therefore assume the existence of a non-trivial partition $\mathcal{U}$ of the space $\mathbb{R}^3 \rtimes \mathbb{S}^2$ into maximal homogeneity compartments. As a consequence for each $g_1 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ there is a vicinity $\mathcal{U}(g_1)$ such that

$$\boldsymbol{E}S_{g_2} = \boldsymbol{E}S_{g_1} \quad \forall\, g_2 \in \mathcal{U}(g_1),$$

where $\boldsymbol{E}$ denotes the expectation with respect to the Rician distribution. This assumption can be extended to more complicated local models, e.g., local linear or local quadratic parameter models. Here, we only consider local constant models for the sake of computational simplicity! See Section 5 for the effects of a misspecified structural assumption.

Our method POAS to infer on expected values $\boldsymbol{E}S_g$, $g \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, and their homogeneity regions $\mathcal{U}(g)$ is based on the Propagation-Separation approach [Polzehl and Spokoiny, 2006] and yields an iterative procedure with a prespecified number of steps k^*. At each iteration step

the aggregated information on the homogeneity regions is used to obtain improved estimates for the image values $\boldsymbol{E}S_g$, using kernel estimation. These estimates are then again used to infer on the homogeneity regions.

The estimate for the image value $\boldsymbol{E}S_{g_1}$ at $g_1 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ and iteration step k is

$$\hat{S}^{(k)}_{g_1} := \sum_{g_2 \in \mathbb{R}^3 \rtimes \mathbb{S}^2} w^{(k)}_{g_1 g_2} S_{g_2} / N^{(k)}_{g_1} \tag{2}$$

with $N^{(k)}_{g_1} := \sum_{g_2} w^{(k)}_{g_1 g_2}$ and local weighting schemes $w^{(k)}_{g_1 g_2}$ for each pair of measurements at $g_1, g_2 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$:

$$w^{(k)}_{g_1 g_2} := K_{loc}\left(\Delta_{\kappa(\vec{b}_1,k)}\left(g_1, g_2\right)/h(\vec{b}_1, k)\right) K_{st}\left(s^{(k)}_{g_1 g_2}/\lambda\right). \tag{3}$$

We call the weights $w^{(k)}_{g_1 g_2}$ adaptive. We use the term non-adaptive if the factor $K_{st}(s^{(k)}_{g_1 g_2}/\lambda)$ is omitted in Eq. (3). In Figure 2 we show the non-adaptive weighting schemes for several bandwidths and values of κ in order to demonstrate effects of the discrepancy function Δ_κ.

Now, we shortly describe the terms used in Eq. (3):

- K_{loc} and K_{st} are kernel functions.
- $\Delta_\kappa\left(g_1, g_2\right)$ is the discrepancy introduced above.
- $\{h(\vec{b}, k)\}_{k=0,\cdots,k^*}$ is an increasing sequence of bandwidths depending on the gradient direction $\vec{b} \in \mathbb{S}^2$.
- The sequence $\{\kappa(\vec{b}, k)\}_{k=0,\cdots,k^*}$ describes for each gradient direction $\vec{b} \in \mathbb{S}^2$ the relation between distances on the sphere and in voxel space.
- The statistical penalty $s^{(k)}_{g_1 g_2}$, defined by

$$s^{(k)}_{g_1 g_2} := N^{(k-1)}_{g_1} \cdot \mathcal{K}\left(\frac{\hat{S}^{(k-1)}_{g_1}}{\hat{\sigma}}, \frac{\hat{S}^{(k-1)}_{g_2}}{\hat{\sigma}}\right),$$

  depends on the Kullback-Leibler distance between the two standardized Rician distributions in g_1 and g_2 with parameters $\hat{S}^{(k-1)}_{g_1}/\hat{\sigma}$ and $\hat{S}^{(k-1)}_{g_1}/\hat{\sigma}$, where $\hat{\sigma}$ denotes an estimate for the scale parameter of the Rician distribution. Evaluating the distance between the estimated image values the statistical penalty tests for our structural assumption. The Kullback-Leibler distance is approximated numerically. $N^{(k)}_g$ is an approximation for the variance reduction at step k.
- λ is the adaptation parameter of the algorithm.

Note that both, the sequence of bandwidths $\{h(\vec{b}, k)\}_{k=0,\cdots,k^*}$ and the sequence of balancing parameters $\{\kappa(\vec{b}, k)\}_{k=0,\cdots,k^*}$ possibly depend on the diffusion gradient direction $\vec{b} \in \mathbb{S}^2$, as most gradient schemes are not fully homogeneous on the sphere.

Starting at very local initial estimates that coincide with non-adaptive-smoothing on the sphere, i.e. setting $s^{(0)}_{g_1 g_2} := 1$ and choosing $h(\vec{b}_l, 0) = 1$, and using the discrepancy derived in the previous section we iterate computation of weights (3) and estimation of $\boldsymbol{E}S_g$ by $\hat{S}^{(k)}_g$ (2), increasing k with each iteration. This defines a position-orientation adaptive smoothing algorithm for diffusion weighted data in the space $\mathbb{R}^3 \rtimes \mathbb{S}^2$.

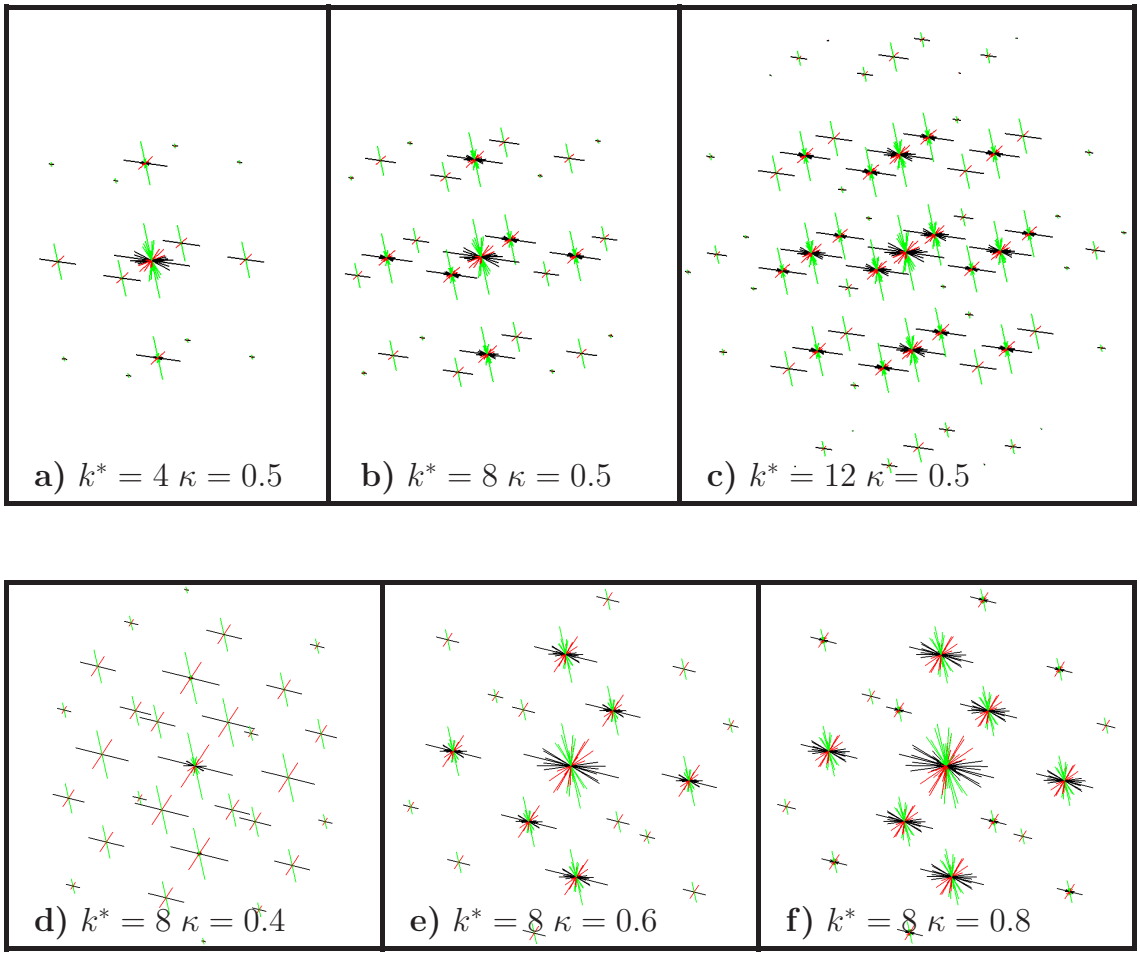

FIGURE 2. Non-adaptive weighting schemes defined by the discrepancy $\Delta_\kappa(g_1, g_2)$ for one voxel and three directions close to the coordinate axes (black - x, green - y and red -z) corresponding to three points $g_1 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$. Weights are shown for the 60 gradient directions used in the experimental data, see Section 3.2. The length of lines corresponds to the weight, the direction to the respective gradient direction in g_2 and the location to $\vec{v}(g_2) - \vec{v}(g_1)$. Upper row: $\kappa_0 = .5$ and number of iterations $k^* = 4$ (a), $k^* = 8$ (b) and $k^* = 12$ (c). Lower row: $k^* = 8$ and varying κ_0, $\kappa_0 = .4$ (d) , $\kappa_0 = .6$ (e) and $\kappa_0 = .8$ (f) Due to the statistical kernel $K_{st}\left(s^{(k)}_{g_1 g_2}\right)$ the adaptive weights are smaller than the non-adaptive weights.

2.3. **Smoothing the non-diffusion weighted images S_0.** Besides the diffusion weighted images S_g, $g \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, the data contain $n_{S_0} \geq 1$ S_0-images derived without applying an additional diffusion gradient. Let $\bar{S}_0 = (\bar{S}_{(\vec{v},0)})_{\vec{v} \in R^3}$ denote the mean over all S_0-images.

The analysis of the diffusion weighted data is usually based on the quotient $S_g / \bar{S}_{(\vec{v}(g),0)}$ where $\vec{v}(g)$ are the spatial (voxel) coordinates in g. Therefore, we have to adaptively smooth the mean image $\bar{S}_0$ in a way that is compatible with the adaptive smoothing of S_g. This can be achieved by defining adaptive weights for $\bar{S}_0$, which are coupled to the adaptive weights of the diffusion

weighted images:

$$\tilde{w}^{(k)}_{\vec{v}_1\vec{v}_2} := K_{loc}\left(||\vec{v}_1 - \vec{v}_2||/h^{(k)}_{\vec{b}_l}\right) K_{st}\left(z^{(k)}_{\vec{v}_1\vec{v}_2}/\lambda\right),$$

with the Euclidean norm $\|.\|$ in $\mathbb{R}^3$ and a statistical penalty

$$z^{(k)}_{\vec{v}_1\vec{v}_2} := \frac{1}{n_{grad}+n_{S_0}}\left[n_{S_0}\cdot\tilde{N}^{(k-1)}_{\vec{v}_1}\cdot\mathcal{K}\left(\frac{\hat{S}^{(k-1)}_{(\vec{v}_1,0)}}{\hat{\sigma}},\frac{\hat{S}^{(k-1)}_{(\vec{v}_2,0)}}{\hat{\sigma}}\right)+\sum_{l=1}^{n_{grad}} s^{(k)}_{g_1^{(l)}g_2^{(l)}}\right]$$

where $g_1^{(l)}, l = 1,\cdots n_{grad}$, are the n_{grad} elements in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ with $\vec{v}(g_1^{(l)}) = \vec{v}_1$. Furthermore we define $s^{(k)}_{g_1g_2}$ as in Section 2.2 and

$$\tilde{N}^{(k)}_{\vec{v}_1} := \sum_{\vec{v}_2}\tilde{w}^{(k)}_{\vec{v}_1\vec{v}_2}.$$

Adaptive estimates for the mean $\bar{S}_0$ are obtained as

$$\hat{S}^{(k)}_{(\vec{v}_1,0)} = \sum_{\vec{v}_2}\tilde{w}^{(k)}_{\vec{v}_1\vec{v}_2}\bar{S}_{(\vec{v}_2,0)}/\tilde{N}^{(k)}_{\vec{v}_1}.$$

2.4. **Choice of parameter values.** The algorithm has a number of parameters, which mostly have only minor influence on the resulting estimates $\hat{S}^{(k^*)}_g$ and $\hat{S}^{(k^*)}_{(\vec{v},0)}$ [Polzehl and Spokoiny, 2006, Tabelow et al., 2008].

The main parameter of the procedure is the adaptation parameter λ which controls the amount of adaptation. If λ is chosen very large, the influence of the value of the statistical penalty on the weights is negligible. If it is chosen too small, the procedure easily adapts to noise, which is equivalent to a random clustering of observed values. Fortunately, λ can be chosen independent of the data by applying the propagation condition [Polzehl and Spokoiny, 2006] to a simulated unstructured situation, i.e., with only one homogeneous region for Rician distributed data. This condition ensures that the quality of the estimates in this situation may deteriorate only by a factor $1+\alpha$ (e.g. $\alpha = 0.1$) in comparison to its, in this case optimal, non-adaptive counterpart. Then, this property also holds for situations with more than one homogeneity region [Polzehl and Spokoiny, 2006], where the structural assumption is fulfilled.

The kernel functions $K : \mathbb{R}^+ \to [0,1]$ should have compact support and be monotone decreasing. The kernel K_{st} should for theoretical reasons exhibit a constant plateau. The exact form of the kernels is not important [see e.g. Section 6.2.3 in Scott, 1992]. Here, we choose them as

$$K_{loc}(x) = \begin{cases} 1-x^2 & x<1 \\ 0 & x\geq 1 \end{cases} \quad \text{and} \quad K_{st}(x) = \begin{cases} 1 & x<0.5 \\ 2-2x & 0.5\leq x<1 \\ 0 & x\geq 1 \end{cases}.$$

For a given gradient direction $\vec{b}$ the sequence of bandwidths $\{h(\vec{b},k)\}_k$, starting with $h(\vec{b},0) = 1$, is chosen such that in case of non-adaptive weights $\bar{w}^{(k)}_{g_1g_2} = K_{loc}\left(\Delta_{\kappa(\vec{b},k)}(g_1,g_2)/h(\vec{b},k)\right)$, it provides a variance reduction $\sum_{g_2}{\bar{w}^{(k)}_{g_1g_2}}^2/\left(\sum_{g_2}\bar{w}^{(k)}_{g_1g_2}\right)^2$ from step $k-1$ to step k by a factor 1.25.

The parameter sequence $\kappa(\vec{b},k)$ in the discrepancy $\Delta_{\kappa(\vec{b},k)}(g_1,g_2)$ in Eq. (3) determines the ratio of the amount of smoothing on the sphere compared to the voxel space, see Section 2.1. $\kappa(\vec{b},k)$ is chosen as $\kappa(\vec{b},k)=\kappa_0/h(\vec{b},k)$ so that the amount of smoothing on the sphere only depends on the distance $||\vec{v}_1-\vec{v}_2||$ in voxel space. For dMRI the nature of the measurement does not relate the radius of the q-shell (given by the b-value) with the voxel extension (given in millimeter). The parameter κ_0 can therefore be freely chosen. Note, that smoothing on the sphere introduces a bias, as the diffusion weighted signal in a fixed voxel is surely a smooth function on $\mathbb{S}^2$. The proposed choice ensures a stabilization of the estimates for the first steps of the method by using a larger portion of the weights on the sphere rather than in $\mathbb{R}^3$. At later iteration steps this portion is reduced leading to a decreased bias and hence more accurate estimates. Based on numerical experience we recommend a value of κ_0 indirect proportional to $\sqrt{n_{grad}}$ for common situations.

Finally, the number of iteration steps k^* defines the maximal bandwidth for the examination in $\mathbb{R}^3 \rtimes \mathbb{S}^2$ and hence determines the computational complexity of the algorithm. In fact, the potential vicinities of voxels which have to be considered, grow with the third power of the bandwidth since the kernel functions are chosen to have compact support. In contrast to smoothing methods based on the diffusion equations, position-orientation adaptive smoothing has an intrinsic stopping criterion, which means that the quality of the best intermediate results holds (up to a constant) for the final estimate, see Polzehl and Spokoiny [2006] for a theoretical study of this property. Within homogeneous parameter regions, POAS behaves almost like non-adaptive smoothing due to the application of the propagation condition for the choice of the adaptation parameter λ. Thus, the number of iteration steps (and correspondingly the maximum bandwidth) should be chosen to ensure a desired smoothness within homogeneous regions and a reasonable computational workload.

In Figure 3 we illustrate the dependence of the smoothing result on the number of iteration steps to shed light into the optimal choice of $k^\star$. We use a region-of-interest in the slice chosen for Figure 6, see Section 4. We show the results for iteration steps $k^\star = 0, 4, 8, 12, 16, 20, 25, 30$. $k^\star = 0$ corresponds to a non-adaptive smoothing on the sphere. Notice that the results with a *smooth* appearance show up for intermediate numbers of iteration steps, i.e., $k^\star = 12$. At later steps, the chosen structural assumption (local constant homogeneity regions) enforces the reconstruction result to a step function with fine steps. Although this might not look as convincing as the smoother result at intermediate steps, the stability of estimates and the bias-variance relation in the method ensures a reasonably small error on the estimates due to the misspecification of the structural assumption, see also Section 5.

2.5. **Variance estimation.** In the statistical penalty the estimation of the variance $\hat{\sigma}^2$ of S_g is crucial. While the parameter λ can be chosen independent of the data at hand, see Section 2.4, we need to determine the variance estimate from the data. In Polzehl and Tabelow [2009] two approaches for voxelwise estimation of the standard deviations have been presented depending on the availability of replicated non-diffusion images, i.e. distinguishing the cases $n_{S_0} > 1$ and $n_{S_0} = 1$. In both cases, σ is estimated as the mode of the distribution of estimates of voxelwise standard deviations. In our experience both approaches lead to similar estimates of σ.

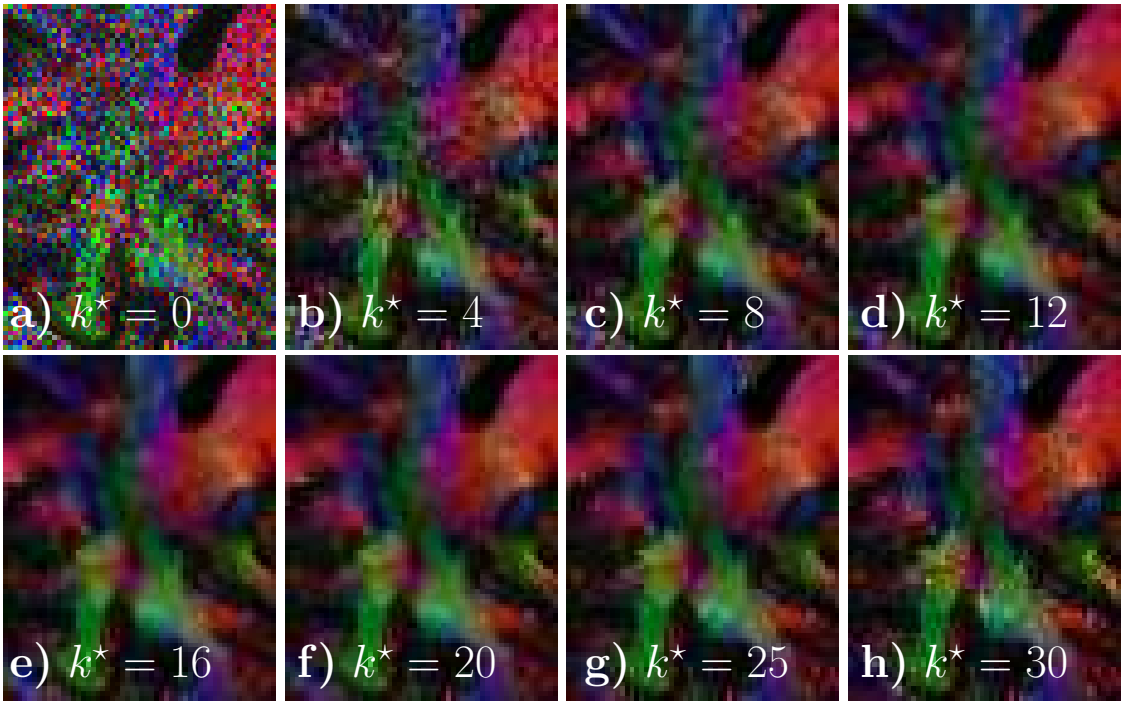

FIGURE 3. Dependence of the smoothing result on the number of iteration steps $k^\star$ (or correspondingly the maximum bandwidth $h^\star$) for the selected region of interest from the slice in Figure 6.

3. DATA

We first investigate properties of the smoothing algorithm by simulation experiments before applying it to experimental diffusion weighted data. For example code, see Appendix B.

3.1. **Simulation.**

We analyze two different artificial data sets to highlight basic properties of the proposed smoothing algorithm.

Example 1. We created an artificial diffusion weighted data set with $32 \times 32 \times 32$ voxels and 42 gradient directions. The latter are chosen to minimize symmetrized Coulomb forces on a sphere following a proposal by Jones et al. [1999]. The data follow a tensor mixture model [Tabelow et al., 2012] that corresponds to fiber bundles of four voxels width completely crossing the cube along the coordinate axes and intermediate areas of again four voxels width in between. The structure contains 22528 voxel with two fiber bundles, 8320 voxel with one fiber bundle and 1920 voxel without fibers. Single fiber bundles are characterized by a prolate diffusion tensor with eigenvalues $(1.4, 0.35, 0.35) \cdot 10^{-3}\,\mathrm{mm}^2/\mathrm{s}$ referred to as typical for white matter [Alexander et al., 2001]. The SNR, defined by the chosen S_0 value and the standard deviation of the noise in k-space, is 10.

Example 2. We created a second artificial diffusion weighted data set with $32 \times 32 \times 11$ voxel and 42 gradient directions. The data contain two homogeneous regions separated by a hyperplane. In each region diffusion weighted data are generated using a tensor mixture model of order two [Tabelow et al., 2012]. Here, tensor mixtures use prolate tensors with eigenvalues

$(1.4, 0.35, 0.35) \cdot 10^{-3}\,\mathrm{mm}^2/\mathrm{s}$. Mixture coefficients in the first region are 0.6 and 0.4 respectively with tensor orientations along the x- and y-axis. The second region has mixture coefficients equal 0.5 and tensor orientations such that a tensor fitted to noiseless data coincides in both regions, see Figure 5 (a) for an illustration. This means that we have no contrast depending on estimated diffusion tensors between both regions which is the worst scenario for the smoothing algorithm described previously in Tabelow et al. [2008]. A color coded FA map obtained from noiseless data would be homogeneous with $\mathrm{FA} \approx 0.38$. The SNR used is 32.

3.2. **Experimental data.** The MR experiment was performed on a 7T whole body MR scanner (MAGNETOM 7T, Siemens Healthcare, Erlangen, Germany) equipped with gradients allowing a peak gradient amplitude of $70\,\mathrm{mT/m}$ with a maximum slew rate of $200\,\mathrm{T/m/s}$ (SC72, Siemens Healthcare, Erlangen, Germany). For signal reception a single channel transmit, 24-channel receive phased array head coil (Nova Medical, Wilmington, MA, USA) was used.

Scans were performed on five healthy young volunteers (age 25 ± 3 years). Written informed consent was obtained from all participants in accordance with the ethical approval from the University of Leipzig. For all acquisitions, an optimized monopolar Stejskal-Tanner sequence [Morelli et al., 2010] was used in conjunction with the ZOOPPA approach [Heidemann et al., 2008] providing an isotropic resolution of $800\,\mu\mathrm{m}$ using the following imaging protocol parameters: 91 slices with 10% overlap and $800\,\mu\mathrm{m}$ isotropic resolution, FOV $143 \times 147\,\mathrm{mm}^2$, TR $14.1\,\mathrm{s}$, TE $65\,\mathrm{ms}$, BW $1132\,\mathrm{Hz/pixel}$, ZOOPPA acceleration factor of 4.6. Diffusion weighted scans were performed with 60 directions with a b-value of $1000\,\mathrm{s/mm}^2$ and 7 interspersed S_0-images. For averaging the scans were repeated 4 times resulting in a total acquisition time of $65\,\mathrm{min}$.

4. Results

Example 1. Figure 4 illustrates results obtained for the first artificial example analyzed within the diffusion tensor model after position-orientation adaptive smoothing. Color coded FA maps are shown for a central slice. For this data tensor estimates have been obtained without noise (a), from noisy data (b), and after smoothing the data using the proposed algorithm with $\lambda = \infty$ (non-adaptive) (c) and $\lambda = 5$ (adaptive) (d). Results for the tensor model based algorithm from Tabelow et al. [2008] (not reported here) are very similar due to a good tensor based contrast between different homogeneous regions in this example. Figure 4 (c) clearly indicated a loss of information due to blurring, if no adaptation is used.

Example 2. We use the second artificial data set to illustrate the difference between the proposed algorithm and our earlier proposal in Tabelow et al. [2008] that was based on information aggregated within the diffusion tensor model. As the method described in this paper does not rely on a specific model for the diffusion weighted data, we show the resulting orientation distribution functions (ODF) estimated in a tensor mixture model [Tabelow et al., 2012], other models like the expansion of the orientation distribution function to spherical harmonics for q-ball imaging [Descoteaux et al., 2007] could have been chosen for visualization with similar results. Figure 5(c) shows that the proposed algorithm is able to remove the distortions caused by noise in Figure 5(b) without blurring the boundary between the two compartments, while the previous algorithm from Tabelow et al. [2008] lacks sensitivity at the discontinuity due to its explicit use of the diffusion tensor model for adaptation, see Figure 5(d).

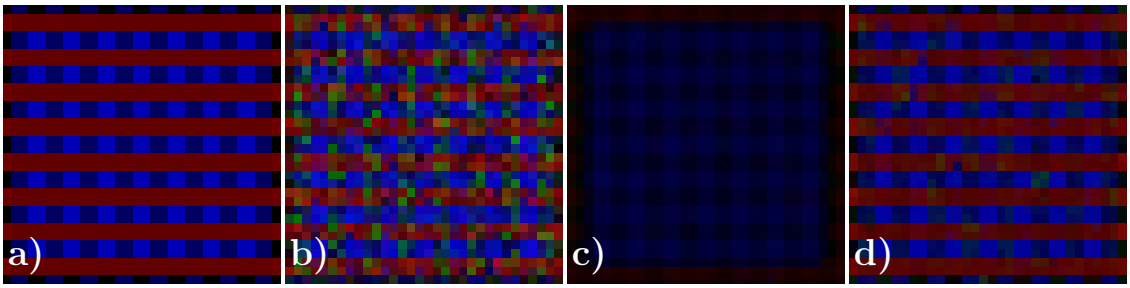

FIGURE 4. Color coded FA maps for example 1 a) for noise-free data, b) for noisy data, c) after smoothing the noisy data with the proposed method but using non-adaptive weights only, d) after position-orientation adaptive smoothing the noisy data as proposed here.

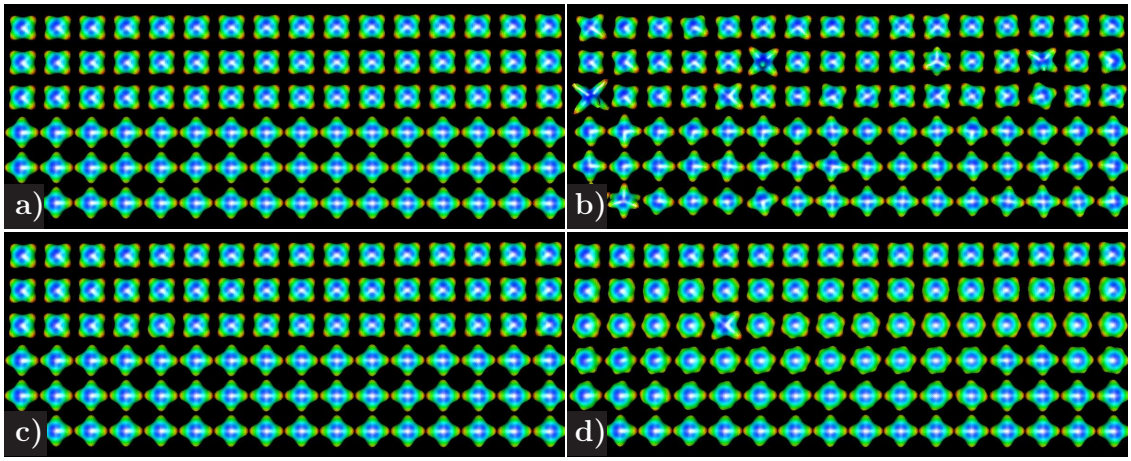

FIGURE 5. Estimated tensor mixture model [Tabelow et al., 2012] a) for noise-free data, b) for noisy data, c) after position-orientation adaptive smoothing as proposed in this paper, d) after smoothing with structural adaptation as proposed previously [Tabelow et al., 2008].

Experimental data. Due to the very high resolution of the experimental diffusion weighted data the SNR for this data set is very low, see Figure 6(a) for a color coded FA map of some axial slice for one subject. However, the position-orientation adaptive smoothing method proposed in this paper using the described choices of parameters (see example code in Appendix B for explicit values) and $k^\star = 12$ iteration steps results in an apparent improvement of the tensor estimates without blurring the fine structures observed in diffusion weighted imaging, see Figure 6(b). The previous version of our method explicitly based on the diffusion tensor model [Tabelow et al., 2008] is not able to remove much of the noise in this data set, see Figure 6(c). The reason is that the former algorithm relies on the diffusion tensor model known to be wrong for large parts of the brain. It obviously fails for very low SNR. For comparison of the results of the smoothing algorithm with a kind of ground truth the measured data set contains four repeated measurements which can lead to a largely improved SNR, see Figure 6(d). This data set with repeated measurements can also be smoothed using the proposed method (Figure 6(e)) and the DTI based algorithm (Figure 6(f)) with apparent improvements. By visual inspection, the

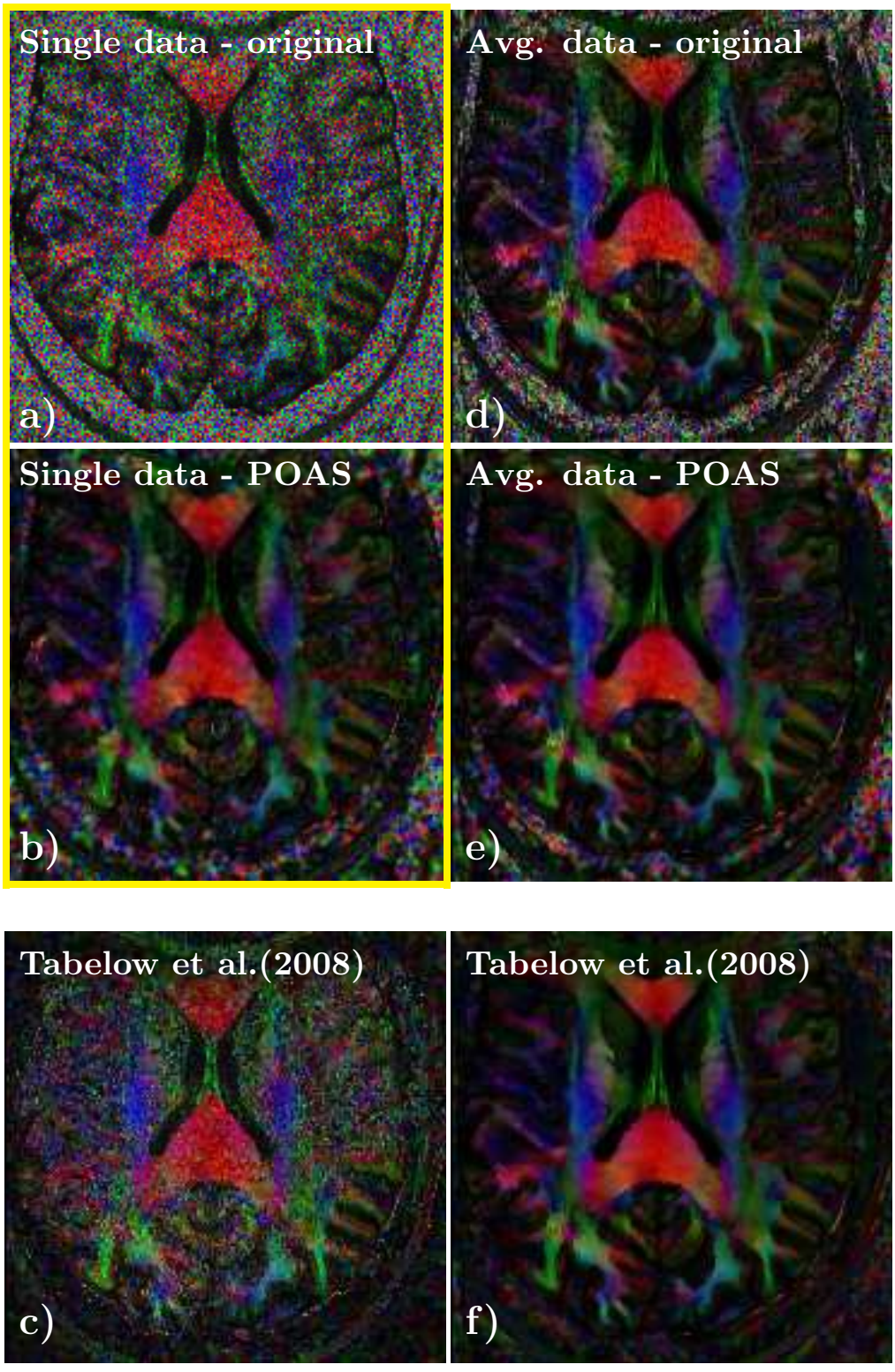

FIGURE 6. Slice showing color coded FA maps generated by estimating the diffusion tensor from the experimental diffusion weighted data for one subject. a) Here, dMRI data from only one measurement has been used. b) This data has been smoothed using the POAS as proposed in this paper. d) In this Figure, the FA is estimated after averaging the dMRI data over four subsequent measurements. Figure e) shows the result after structural adaptive smoothing this averaged data. c) and f) show the estimated FA after smoothing with the method based on DTI [Tabelow et al., 2008] for the single data set and the averaged data set, respectively.

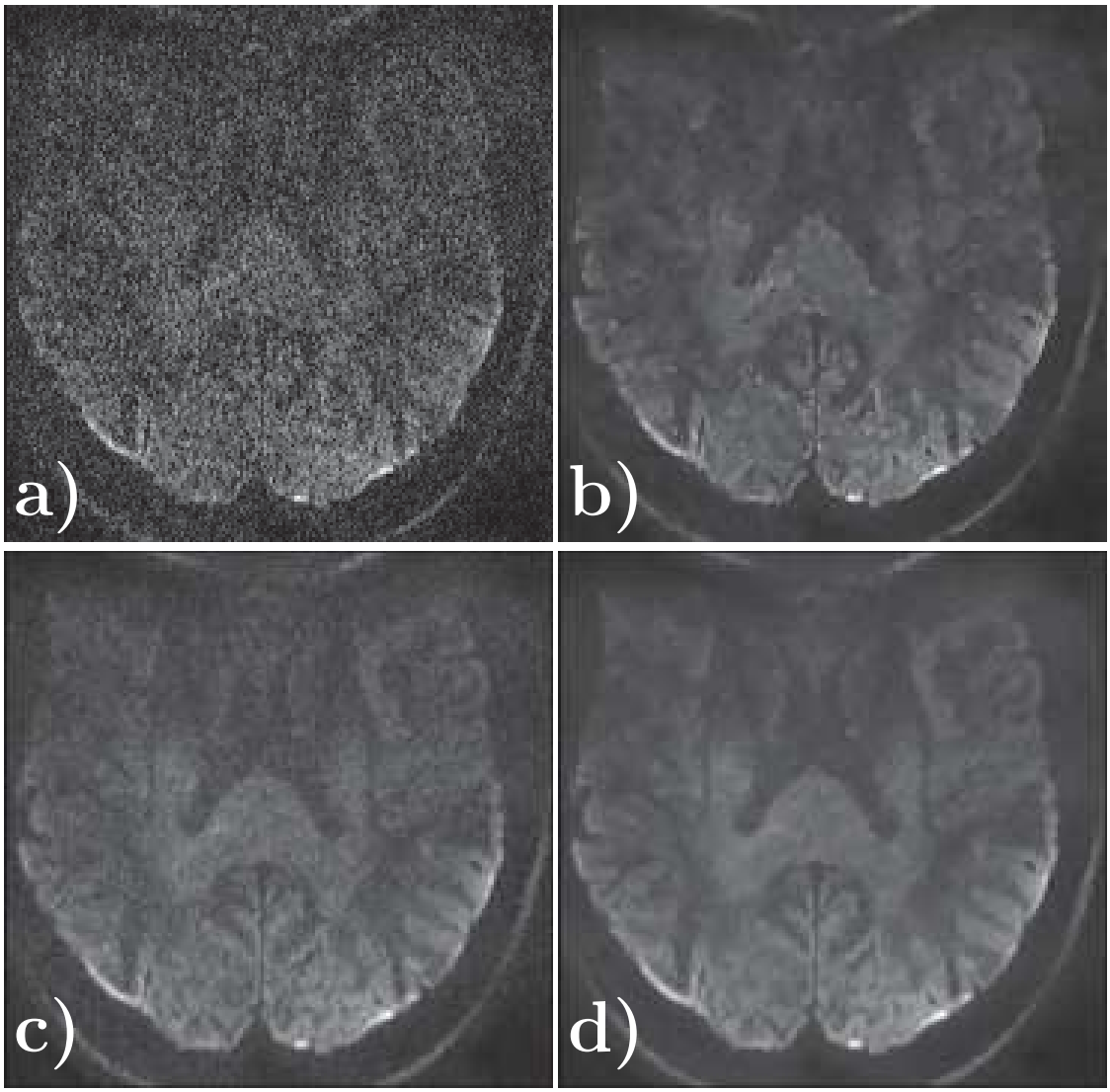

FIGURE 7. a) Slice of diffusion weighted data for one (arbitrarily selected) gradient direction. b) Same diffusion weighted data after POAS as proposed in this paper. c) Same slice shown after averaging over four subsequent measurements (same as Fig. 1). d) Smoothed averaged data set.

quality of the smoothed result based on a single measurement compares well with the repeated data, which required a four times longer acquisition time.

In Figure 7 slices of an (arbitrarily selected) diffusion weighted image S_g are shown (same as Figure 1), together with the results after smoothing. Furthermore, we show the comparison of the diffusion weighted image for the averaged data set with and without smoothing.

Finally, in Figure 8, we provide fiber tracking results using a simple streamline algorithm [Mori et al., 1999, Xue et al., 1999] with $\mathrm{FA} < 0.266$ as stopping rule and a maximum angle of 45 degrees between the line segments. Figure 8(a-d) show only fibers through a region-of-interest within the corpus callosum, while Figure 8(e-h) provides the results for the whole brain. Noise

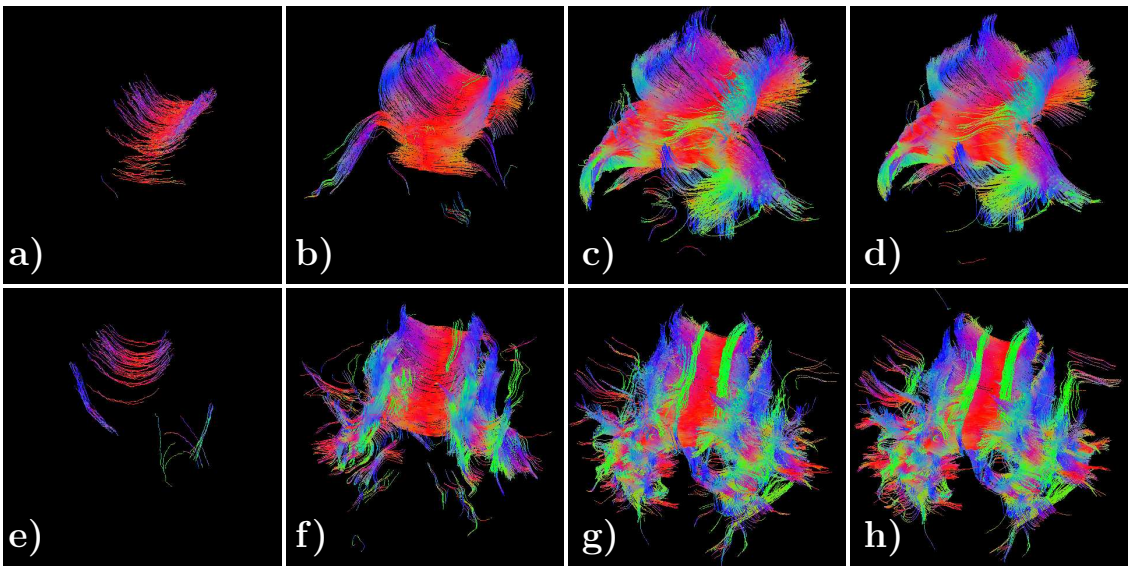

FIGURE 8. Fiber tracking results using the estimated diffusion tensors and a simple streamline algorithm. a)-d) Only fibers through a region-of-interest defined by the center of the corpus callosum are shown. e)-h) All Fibers with a minimum length of 25 fiber segments in the whole data set are shown. a) and e) are the results of the original data, b) and f) after POAS. c) and g) use the averaged diffusion weighted data. d) and h) show the tracks after smoothing this averaged data set.

obviously renders fiber tracking difficult for the single measurement Figure 8(a+e) in contrast to the repeated measurement, see Figure 8(c+g). Position-orientation adaptive smoothing leads to improved tracking results for a single measurement (Figure 8(b+f)) and comparable results for the repeated measurement (Figure 8(d+h)). Of note is that for fiber tracking with the streamline method the repeated measurement still has some advantages over the smoothed single measurement. Other tracking methods could have been used, but this is beyond the scope of this paper. The reduced FA and the low fiber density outside the occipital lobe is due the reduced sensitivity of the receiver coil at 7T.

5. DISCUSSION

Generally, position-orientation adaptive smoothing yields similar results to methods based on anisotropic non-linear diffusion. However, there is an important difference between both classes of algorithms. Diffusion methods for infinite diffusion times end up with a completely homogeneous image and hence require an appropriate stopping criterion. In contrast to this, structural adaptive smoothing based on the Propagation-Separation (PS) approach is known to provide an intrinsic stopping criterion, where the quality of estimates is preserved up to a constant in all later steps of the iteration [Polzehl and Spokoiny, 2006, Tabelow et al., 2008]. As a consequence, the number of iteration steps $k^\star$ or correspondingly the maximum bandwidth in K_{loc} only defines the numerical complexity of the algorithm and the smoothness within homogeneous regions. In cases of a misspecification of the structural assumption, i.e., if the data violates the assumption of local constant image values, the Propagation-Separation approach forces the

resulting reconstruction into a step function rather than a smooth function, which might *look* a worse result, see Figure 3 in Section 2.4. However, as the bias is bounded by the standard deviation of estimates obtained in intermediate steps of the algorithm and due to the stability of the estimates for large bandwidths, the error by the misspecified local model remains small [Polzehl and Spokoiny, 2006]. Alternatively the structural assumption may be replaced by a local smooth assumption, leading to a local polynomial model, on the sphere. However, this would lead to a much higher computational complexity.

Most of the other parameters of the method besides $k^\star$ and κ are of only minor influence on the smoothing results or can be chosen in advance by simulation independent of the data. Further research may concentrate on the choice of the sequence of proportionality parameters $\{\kappa(\vec{b}, k)\}_k$ balancing between distances in the voxel space and on the sphere. Here, we proposed a choice yielding for all iteration steps a constant amount of smoothing on the sphere while the considered region extends only in voxel space, see Section 2.4. In this case the choice of the initial parameter κ_0 is crucial. However, $\{\kappa(\vec{b}, k)\}_k$ could be chosen on its own, independent of the bandwidth $\{h(\vec{b}, k)\}_k$, by multivariate bandwidth selection taking into account the distinct spatial and angular smoothness properties of the signal.

In Tabelow et al. [2008] we first proposed structural adaptive smoothing in the context of the diffusion tensor model for dMRI. There are two important differences and extensions in the method proposed in this paper. First, the algorithm in Tabelow et al. [2008] defines the homogeneity regions and the statistical penalty using the estimated diffusion tensors. Therefore, the application of the algorithm was restricted to DTI, see also Figure 5 and 6(c) for its limitations. The extension proposed in the current paper does not refer to any specific model and can thus be applied to dMRI data prior to any further analysis. Second, the definition of the location weights is based on a discrepancy on $\mathrm{SE}(3)$ (in which the measurement space $\mathbb{R}^3 \rtimes \mathbb{S}^2$ is embedded) rather than on $\mathbb{R}^3$. This includes smoothing on the sphere. The use of the specific geometry for smoothing has been first introduced in Franken [2008]. There, smoothing of the diffusion weighted images is based on the diffusion equation [Weickert, 1998]. The method proposed in this paper relies on a different methodology as it defines local weighting schemes for adaptation to the underlying structure of the data with an intrinsic stopping criterion.

In the presence of partial volume effects the diffusion weighted signal is a superposition of the signals from the corresponding compartments. As a result the areas of local constant diffusion weighted signal may be relatively small: If, e.g., the neighboring voxel contains approximately the same mixture of fibers, the diffusion weighted signals in neighboring voxels are comparable and the statistical penalty leads to a non-adaptive weight as desired. However, if one of the fibers is absent in the neighboring voxel, the diffusion weighted signals will differ significantly and lead to a vanishing weight for this voxel. The result is a non-optimal efficiency for the position-orientation adaptive smoothing. On the other hand the adaptation avoids blurring at these borders with the only effect of a reduced amount of smoothing compared to the (unrealistic) situation where we had complete knowledge on the partial volume effects.

By naturally adapting to the structures of interest at different scales, the algorithm avoids loss of information on size and shape of structures, which is typically observed when using non-adaptive filters. More specific, by its iterative nature, the method accumulates information on the spatial structure at small scales and uses this information to improve estimates at coarser scales. The potential of this method has been shown in Section 4 for simulated data where the

ground truth is known. Note again that the position-orientation adaptive smoothing algorithm does not refer to a specific model for the dMRI data and hence data generation based on such models does not work in favor of the smoothing method. For the experimental data the improvements by POAS especially for very low SNR are apparent.

Another diffusion imaging method that might possibly benefit from adaptive smoothing is the double wave vector dMRI [Koch and Finsterbusch, 2008]. The modification of POAS for this image acquisition will be an interesting subject for future research.

6. Conclusion

In this paper we developed a position-orientation adaptive smoothing algorithm (POAS) based on the Propagation-Separation approach for dMRI data. It has several important properties and advantages:

- The proposed algorithm does not rely on a specific model for the dMRI data. Therefore, after using the method for smoothing the diffusion weighted images any model can be applied to the data, e.g., the diffusion tensor model or higher order models.
- The method has an intrinsic stopping criterion, which means that most of the parameters of the method have only minor influence on the results, while the bandwidth parameter $k^\star$ is limited only by the available computational power and the desired smoothness in homogeneous regions.
- POAS uses the special geometry of the measurement space $\mathbb{R}^3 \rtimes \mathbb{S}^2$.
- POAS is designed to be adaptive to the fine anisotropic structures observed in dMRI by using a statistical penalty. This ensures propagation within homogeneous compartments and separation between distinct compartments avoiding blurring at structural borders.
- The improved quality of the data after smoothing can be used for further analysis or in clinical context for a reduction of number of diffusion weighting gradients and hence acquisition time.

Acknowledgments

This work is supported by the DFG Research Center Matheon and the Stiftung der Deutschen Wirtschaft (SDW). We thank Peter Mathé (WIAS Berlin) and Remco Duits (Eindhoven University of Technology) for helpful discussions.

Appendix A. Mathematical details

In the first two parts of this section we want to summarize some results of Franken [2008], Duits and Franken [2011] on which Section 2.1 is based. Then, we discuss why POAS is a well-defined operation on the left coset space $\mathrm{SE}(3)/(\{0\} \rtimes \mathrm{stab}(\vec{e}_z))$ and how the Riemannian 2-norm can be computed.

A.1. Parametrization of $\mathbb{S}^2$ and $\mathrm{SO}(3)$.

Let

$$\mathbf{R}_\theta^{\vec{e}_x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \mathbf{R}_\theta^{\vec{e}_y} = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \mathbf{R}_\theta^{\vec{e}_z} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

denote the rotations around the x-, y- and z-axis. Then, we parametrize $\mathrm{SO}(3)$ by:

$$(\alpha,\beta,\gamma) \mapsto (\hat{\mathbf{R}}_{(\alpha,\beta,\gamma)}) := \mathbf{R}_\gamma^{\vec{e}_z}\mathbf{R}_\beta^{\vec{e}_y}\mathbf{R}_\alpha^{\vec{e}_z} \in \mathrm{SO}(3), \qquad \beta \neq 0,\pi, \tag{4}$$

$$(\alpha,\beta,\gamma) \mapsto (\mathbf{R}_{(\alpha,\beta,\gamma)}) := \mathbf{R}_\gamma^{\vec{e}_x}\mathbf{R}_\beta^{\vec{e}_y}\mathbf{R}_\alpha^{\vec{e}_z} \in \mathrm{SO}(3), \qquad \beta \neq \tfrac{\pi}{2},\tfrac{3\pi}{2}. \tag{5}$$

For the following parametrization of $\mathbb{S}^2$ it holds

$$(\beta,\gamma) \mapsto \hat{\vec{u}}(\beta,\gamma) := \begin{pmatrix} \cos\gamma\sin\beta \\ \sin\gamma\sin\beta \\ \cos\beta \end{pmatrix} = \hat{\mathbf{R}}_{(\alpha,\beta,\gamma)}\vec{e}_z \in \mathbb{S}^2, \quad \beta \neq 0,\pi,$$

$$(\beta,\gamma) \mapsto \vec{u}(\beta,\gamma) := \begin{pmatrix} \sin\beta \\ -\cos\beta\sin\gamma \\ \cos\beta\cos\gamma \end{pmatrix} = \mathbf{R}_{(\alpha,\beta,\gamma)}\vec{e}_z \in \mathbb{S}^2, \quad \beta \neq \tfrac{\pi}{2},\tfrac{3\pi}{2}.$$

In practice the use of (5) is sufficient since the solutions $k_1,\dots,k_6$ of (1) at the singularity can be set to their unique limits. The parametrization in (4) has the disadvantage not to be well-defined in the identity element $\mathbf{E}_3 \in \mathrm{SO}(3)$.

A.2. **Properties of** $\mathrm{SE}(3)$**.** $\mathrm{SE}(3)$ forms a Lie group with semidirect group product

$$g_1 \cdot_{\mathrm{SE}(3)} g_2 = (\vec{v}_1,\mathbf{R}_1)\cdot_{\mathrm{SE}(3)}(\vec{v}_2,\mathbf{R}_2) = (\vec{v}_1+\mathbf{R}_1\vec{v}_2,\mathbf{R}_1\mathbf{R}_2), \qquad g_1,g_2 \in \mathrm{SE}(3),$$

and the following table of Lie brackets:

$$[X_i,X_j]_{i,j} = \begin{pmatrix} 0 & 0 & 0 & 0 & X_3 & -X_2 \\ 0 & 0 & 0 & -X_3 & 0 & X_1 \\ 0 & 0 & 0 & X_2 & -X_1 & 0 \\ 0 & X_3 & -X_2 & 0 & X_6 & -X_5 \\ -X_3 & 0 & X_1 & -X_6 & 0 & X_4 \\ X_2 & -X_1 & 0 & X_5 & -X_4 & 0 \end{pmatrix},$$

where $\{X_i \in T_e(\mathrm{SE}(3)) : i = 1,...,6\}$ denotes a basis of the corresponding Lie algebra $(T_e(\mathrm{SE}(3)),[\cdot,\cdot])$. Using the matrix representation of $\mathrm{SE}(3)$, which we denote by bold letters, it holds for $(l_1,l_2,l_3,l_4,l_5,l_6) := (x,y,z,\gamma,\beta,\alpha)$ and $g = ((x,y,z)^T,\mathbf{R}_{(\alpha,\beta,\gamma)})$

$$\prod_{i=1}^{6}\exp(l_i\mathbf{X}_i) = \begin{pmatrix} & & & x \\ & \mathbf{R}_{(\alpha,\beta,\gamma)} & & y \\ & & & z \\ 0 & 0 & 0 & 1 \end{pmatrix} =: \mathbf{M}_g \equiv g \in \mathrm{SE}(3), \tag{6}$$

where

$$\exp(\mathbf{M}) := \mathbb{I}_{n\times n} + \sum_{k=1}^{\infty}\frac{\mathbf{M}^k}{k!} \in \mathbb{R}^{n\times n} \text{ for all } \mathbf{M} \in \mathbb{R}^{n\times n}.$$

The corresponding basis matrices of the Lie algebra are

$$\mathbf{X}_1 = \begin{pmatrix} 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \quad \mathbf{X}_2 = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \quad \mathbf{X}_3 = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{pmatrix}$$

$$\mathbf{X}_4 = \begin{pmatrix} 0&0&0&0 \\ 0&0&-1&0 \\ 0&1&0&0 \\ 0&0&0&0 \end{pmatrix} \quad \mathbf{X}_5 = \begin{pmatrix} 0&0&1&0 \\ 0&0&0&0 \\ -1&0&0&0 \\ 0&0&0&0 \end{pmatrix} \quad \mathbf{X}_6 = \begin{pmatrix} 0&-1&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}.$$

Then, the left-invariant vector fields $\{\mathcal{A}_i\}_{i=1}^6$ on $\mathrm{SE}(3)$ are defined for $U \in L^2(\mathrm{SE}(3),\mathbb{R})$ by

$$(\mathcal{A}_i|_g U)(g) = \lim_{h\to 0} \frac{U(g\cdot \exp(hX_i)) - U(g)}{h}, \quad i = 1,...,6,$$

where $\exp : T_e(\mathrm{SE}(3)) \to \mathrm{SE}(3)$ is the exponential map in $\mathrm{SE}(3)$. This yields with respect to the parametrization given in Eq. (5):

$$\begin{aligned}
\mathcal{A}_1|_g &= \cos\alpha\cos\beta\,\partial_x + (\cos\gamma\sin\alpha + \cos\alpha\sin\beta\sin\gamma)\,\partial_y + (\sin\alpha\sin\gamma - \cos\alpha\cos\gamma\sin\beta)\,\partial_z \\
\mathcal{A}_2|_g &= -\sin\alpha\cos\beta\,\partial_x + (\cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma)\,\partial_y + (\sin\alpha\sin\beta\cos\gamma + \cos\alpha\sin\gamma)\,\partial_z \\
\mathcal{A}_3|_g &= \sin\beta\,\partial_x - \cos\beta\sin\gamma\,\partial_y + \cos\beta\cos\gamma\,\partial_z \\
\mathcal{A}_4|_g &= -\cos\alpha\tan\beta\,\partial_\alpha + \sin\alpha\,\partial_\beta + \frac{\cos\alpha}{\cos\beta}\,\partial_\gamma \\
\mathcal{A}_5|_g &= \sin\alpha\tan\beta\,\partial_\alpha + \cos\alpha\,\partial_\beta - \frac{\sin\alpha}{\cos\beta}\,\partial_\gamma \\
\mathcal{A}_6|_g &= \partial_\alpha
\end{aligned} \tag{7}$$

The left-invariant basis matrices $\{\mathbf{A}_i|_g\}_{i=1}^6$ of $\mathrm{SE}(3)$ can be obtained by replacing $\partial_x, \partial_y, \partial_z, \partial_\gamma, \partial_\beta, \partial_\alpha$ by $\mathbf{X}_1, ..., \mathbf{X}_6$ in Eq. (7).

The associated left-invariant dual-frame $\{d\mathcal{A}^i\}_{i=1}^6$ is uniquely determined by

$$\langle d\mathcal{A}^i, \mathcal{A}_j\rangle := d\mathcal{A}^i(\mathcal{A}_j) = \delta^i_j, \quad i,j = 1,...,6,$$

where $\delta^i_j = 1$ if $i = j$ and zero else.

A.3. **Why POAS is well-defined.** We have to show that POAS, introduced in Section 2.2, is well-defined w.r.t. the embedding of $\mathbb{R}^3 \rtimes \mathbb{S}^2$ into $\mathrm{SE}(3)$, i.e. independent of the equivalence classes. Following Duits and Franken [2011], we show further that the method is left-invariant.

Under the notations of Section 2.1 and Appendices A.1 and A.2 the Riemannian 2-norm of $g \in \mathrm{SE}(3)$ is given by

$$\|g\|_R := \inf\{\delta > 0 : \exists\, \varphi \in C^1([0,1], \mathrm{SE}(3)) \text{ with } \varphi(0) = e, \varphi(1) = g,$$
$$\dot\varphi(s) = \sum_{i=1}^6 \varphi_i(s)\mathbf{A}_i|_{\varphi(s)} \text{ and } \left(\int_0^1 \sum_{i=1}^6 |\varphi_i(s)|^2 ds\right)^{1/2} < \delta\}$$

This norm can be described by a metric tensor as we will discuss now.

Let $\mathbf{G} : \mathrm{SE}(3) \times T(\mathrm{SE}(3)) \times T(\mathrm{SE}(3)) \to \mathbb{C}$ denote a metric tensor, where

$$\mathbf{G}_{(\vec v,\vec u)} := \sum_{i,j=1}^6 g_{ij}(\vec v, \mathbf{R}_{\vec u})\, d\mathcal{A}^i \otimes d\mathcal{A}^j \tag{8}$$

for a rotation $\mathbf{R}_{\vec u} \in \mathrm{SO}(3)$ with $\mathbf{R}_{\vec u}\vec e_z = \vec u$. Such metric tensor is left-invariant and well-defined iff $g_{ij}(\vec v, \mathbf{R}_{\vec u}) \equiv g_{ij}$ are constant for all $i,j = 1,...,6$ and

$$\{g_{ij}\}_{ij} = \mathrm{diag}\{g_{11}, g_{11}, g_{33}, g_{44}, g_{44}, g_{66}\} \tag{9}$$

as shown in Duits et al. [2011, Appendix E].[2]

[2]Note, that in Duits et al. [2011, Appendix E] $g_{66} := 0$ since the considered diffusion operator should not be enforced in direction $\mathcal{A}_6$ associated with rotation angle α. However, due to the non-zero commutators of the Lie-group $\mathrm{SE}(3)$ the left-invariant vector fields $\mathcal{A}_4$ and $\mathcal{A}_5$ depend on $\mathcal{A}_6$, see (7), such that this direction can not be avoided completely. Therefore we prefer to allow values $g_{66} \neq 0$ yielding a metric over the whole space $\mathrm{SE}(3)$, or rather $\mathbb{R}^3 \rtimes \mathbb{S}^2$. It can be directly verified, that this does not affect the results in Duits et al. [2011, Appendix E].

For $\mathbf{G} := \mathrm{diag}\{1,1,1,1,1,1\}$ it holds

$$\sum_{i=1}^{6} |\varphi_i(s)|^2 = \mathbf{G}_{\varphi(s)}(\dot{\varphi}(s), \dot{\varphi}(s)),$$

why

$$\|g\|_R = \inf\{ \left(\int_0^1 \mathbf{G}_{\varphi(s)}(\dot{\varphi}(s), \dot{\varphi}(s)) ds \right)^{1/2} \text{ with } \varphi \in C^1\left([0,1], \mathrm{SE}(3)\right), \text{ where}$$

$$\varphi(0) = e, \varphi(1) = g \text{ and } \dot{\varphi}(s) = \sum_{i=1}^{6} \varphi_i(s) \mathbf{A}_i|_{\varphi(s)} \},$$

Hence, the Riemannian 2-norm is left-invariant and well-defined on the space $\mathbb{R}^3 \rtimes \mathbb{S}^2$. The same holds true for $\|.\|_{R,\kappa}$ with $\mathbf{G} := \mathrm{diag}\{1,1,1,\kappa^2,\kappa^2,\kappa^2\}$ yielding the identity

$$\sum_{i=1}^{3} |\varphi_i(s)|^2 + \sum_{i=4}^{6} \kappa^2 |\varphi_i(s)|^2 = \mathbf{G}_{\varphi(s)}(\dot{\varphi}(s), \dot{\varphi}(s)).$$

Next, let us consider the non-adaptive estimator of POAS, i.e. Eq. (2) with $K_{st}(s^{(k)}_{g_1 g_2}/\lambda) \equiv 1$. This corresponds to a (discrete) convolution on $\mathbb{R}^3 \rtimes \mathbb{S}^2$, see Duits and Franken [2011, Section 3] for a definition, with convolution kernel $\Psi(g) := K_{loc}(\|(\vec{v}, \mathbf{R}_{\vec{u}})\|_{R,\kappa(\vec{u},k)}/h(\vec{u},k))/N^{(k)}_g$, $g = (\vec{v}, \vec{u}) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, and $k = 0, ..., k^*$ fixed. Note, that the bandwidths $\{h(\vec{u},k)\}_k$ do not use the embedding of $\mathbb{R}^3 \rtimes \mathbb{S}^2$ into $\mathrm{SE}(3)$ and that the proportionality parameters $\{\kappa(\vec{u},k)\}_k$ depend on the gradient $\vec{u}$ only via the bandwidths. Therefore, we only show left-invariance for the bandwidths. The bandwidths are determined by solving the following equation w.r.t. $h(\vec{u}_1,k)$ for each iteration step k and $g_1 \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, see Section 2.4.

$$\sum_{g_2 \in \mathbb{R}^3 \rtimes \mathbb{S}^2} (\bar{w}^{(k)}_{g_1 g_2})^2 \cdot (\sum_{g_2 \in \mathbb{R}^3 \rtimes \mathbb{S}^2} \bar{w}^{(k)}_{g_1 g_2})^{-2} \stackrel{!}{=} 1.25^k \cdot \sum_{g_2 \in \mathbb{R}^3 \rtimes \mathbb{S}^2} (\bar{w}^{(0)}_{g_1 g_2})^2 \cdot (\sum_{g_2 \in \mathbb{R}^3 \rtimes \mathbb{S}^2} \bar{w}^{(0)}_{g_1 g_2})^{-2},$$

where $\bar{w}^{(k)}_{g_1 g_2}$ denote the non-adaptive weights depending on $h(\vec{u}_1,k)$. We observe that the bandwidths $h(\vec{u}_1,k)$ depend only on the norm $\|.\|_{R,\kappa(\vec{u}_1,k)}$ with $k = 0, ..., k^*$ and on the fixed values $h(\vec{u}_1,0) = 1$ and κ_0. Thus, we get the left-invariance of $\{h(\vec{u},k)\}_k$ and $\{\kappa(\vec{u},k)\}_k$. A convolution on $\mathbb{R}^3 \rtimes \mathbb{S}^2$ is left-invariant and well-defined iff the convolution kernel is left-invariant and well-defined. Here, the convolution kernel Ψ depends directly on the left-invariant and well-defined term $\|.\|_{R,\kappa(\vec{u},k)}/h(\vec{u},k)$ yielding the desired properties of the non-adaptive estimator.

Finally, we look at the statistical penalty $s^{(k)}_{g_1 g_2}$. It is based on the estimates $\hat{S}^{(k-1)}_{g_1}$ and $\hat{S}^{(k-1)}_{g_2}$ from the previous iteration step starting for $k = 0$ with the non-adaptive and hence left-invariant and well-defined estimate $\hat{S}^{(0)}_{g_1} = \sum_{g_2} w^{(0)}_{g_1 g_2} S_{g_2}/N^{(0)}_{g_1}$. Hence, it follows by induction that (2) remains left-invariant and well-defined when using the adaptive weights instead of the non-adaptive ones.

A.4. **Computing the discrepancy $\Delta_\kappa(g_1, g_2)$.** Let

$$\hat{g} := \left(\mathbf{R}_{\vec{u}_2}^{-1}(\vec{v}_1 - \vec{v}_2), \mathbf{R}_{\vec{u}_2}^{-1}\mathbf{R}_{\vec{u}_1}\right)$$

where $g_i = (\vec{v}_i, \vec{u}_i) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$, $i = 1, 2$, and $\mathbf{R}_{\vec{u}_i} \in \mathrm{SO}(3)$ is any rotation with $\mathbf{R}_{\vec{u}_i}\mathbf{e}_z = \vec{u}_i$. We want to calculate the coordinates $\{k_i\}_{i=1,\dots,6}$ satisfying the equation

$$\prod_{i=1}^{6} \exp(k_i \mathbf{A}_i|_{\hat{g}}) = \begin{pmatrix} & & & \hat{x} \\ & \mathbf{R}_{(0,\hat{\beta},\hat{\gamma})} & & \hat{y} \\ & & & \hat{z} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{M}_{\hat{g}} \equiv \hat{g}$$

as introduced in Section 2.1. In other words, we calculate the coordinates of the element $\hat{g}$ w.r.t. the left-invariant basis matrices $\{\mathbf{A}_i|_{\hat{g}}\}_{i=1,\dots,6}$ which have been defined in Eq. (7), see for comparison Eq. (6). Note that the coordinates are independent of rotation angle α since the Riemannian 2-norm is well-defined w.r.t. the embedding of $\mathbb{R}^3 \rtimes \mathbb{S}^2$ into $\mathrm{SE}(3)$. Further, the rotation matrices $\mathbf{R}_{\vec{u}_i} \in \mathrm{SO}(3)$ are defined by $\mathbf{R}_{\vec{u}_i}\mathbf{e}_z = \vec{u}_i$ such that $\alpha_i \in [0, 2\pi)$ can be freely chosen. So, we set $\alpha_1 = \alpha_2 = \hat{\alpha} = 0$ yielding an easier form than (7) for the basis matrices $\{\mathbf{A}_i|_g\}_{i=1,\dots,6}$. Then, it holds

$$\prod_{i=1}^{3} \exp(k_i \mathbf{A}_i|_{\hat{g}}) = \begin{pmatrix} 1 & 0 & 0 & k_1 \cos\hat{\beta} + k_3 \sin\hat{\beta} \\ 0 & 1 & 0 & k_1 \sin\hat{\beta}\sin\hat{\gamma} + k_2 \cos\hat{\gamma} - k_3 \cos\hat{\beta}\sin\hat{\gamma} \\ 0 & 0 & 1 & -k_1 \sin\hat{\beta}\cos\hat{\gamma} + k_2 \sin\hat{\gamma} + k_3 \cos\hat{\beta}\cos\hat{\gamma} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$\prod_{i=4}^{6} \exp(k_i \mathbf{A}_i|_{\hat{g}}) = \begin{pmatrix} & & & 0 \\ & \mathbf{N}_{(k_4,k_5,k_6)} & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\mathbf{N}_{(k_4,k_5,k_6)}$ denotes an appropriate matrix depending only on

$$\mathbf{A}_4|_{\hat{g}} = (\cos\hat{\beta})^{-1}\mathbf{X}_4 - \tan\hat{\beta}\,\mathbf{X}_6, \qquad \mathbf{A}_5|_{\hat{g}} = \mathbf{X}_5, \qquad \mathbf{A}_6|_{\hat{g}} = \mathbf{X}_6.$$

Further, it holds

$$\begin{aligned} \hat{v} &= \mathbf{R}_{(0,\beta_2,\gamma_2)}^{-1}(\vec{v}_1 - \vec{v}_2), \\ \hat{\beta} &= \arcsin[\sin\beta_1 \cos\beta_2 - \cos\beta_1 \sin\beta_2 \cos(\gamma_1 - \gamma_2)], \\ \hat{\gamma} &= \arcsin[\cos\beta_1 \sin(\gamma_1 - \gamma_2)(\cos\hat{\beta})^{-1}] \end{aligned}$$

and

$$\mathbf{R}_{(0,\beta,\gamma)}^{-1} = \begin{pmatrix} \cos\beta & \sin\beta\sin\gamma & -\sin\beta\cos\gamma \\ 0 & \cos\gamma & \sin\gamma \\ \sin\beta & -\cos\beta\sin\gamma & \cos\beta\cos\gamma \end{pmatrix}.$$

So, we can deduce directly that

$$(k_1, k_2, k_3)^T = \mathbf{R}_{(0,\hat{\beta},\hat{\gamma})}^{-1}\hat{v}.$$

The equation $\mathbf{N}_{(k_4,k_5,k_6)} = \mathbf{R}_{(0,\hat{\beta},\hat{\gamma})}$ is solved numerically yielding the coordinates k_4, k_5 and k_6.

APPENDIX B. CODE SNIPPETS

Computations are performed with our R-package **dti** [Tabelow and Polzehl, 2011] (version 1.0) on an Intel(R) Xeon CPU, Six-Core 2933MHz, 24 GB RAM. The operating system was OpenSuse 11.3 with R version 2.12.1. The R-package is freely available on CRAN (`http://cran.r-project.org`) and NITRC (`http://www.nitrc.org`) and has been described in detail in Polzehl and Tabelow [2011]. Our current implementation uses R [R Development Core Team, 2010] with FORTRAN. Computation time for 3 Million voxels (approx. 180 million points in $\mathbb{R}^3 \rtimes \mathbb{S}^2$) in the experimental data using $k^* = 12$ has been approximately $2\,\mathrm{h}$.

In the following example the data, here in DICOM format, are supposed to reside in a directory `s0002`. The number of slices in the scan is 91, the file `b-directions.txt` contains the gradient directions including the non-diffusion weighted gradients in the order in which they where scanned. The following code reads the data, performs the position-orientation adaptive smoothing described in this paper, estimates the diffusion tensors and fractional anisotropy and displays the color-coded FA for slice number 57.

```
> library(dti)
> gradient <- read.table("b-directions.txt")
> dwiobj <- readDWIdata(gradient, dirlist = "s0002",
                        format = "DICOM", nslices = 91,
                        level = 0)
> dwiobj.s <- dwi.smooth(dwiobj, kstar = 12, kappa = .4,
                         sigma = 60, lambda = 10)
> dtiobj.s <- dtiTensor(dwiobj.s)
> dtiind.s <- dtiIndices(dtiobj.s)
> plot(dtiind.s, slice = 57)
```

For more information on the usage of the package see the help files of the package **dti** and Polzehl and Tabelow [2011].

REFERENCES

A. L. Alexander, K. Hasan, M. Lazar, J. S. Tsuruda, and D. L. Parker. Analysis of partial volume effects in diffusion-tensor MRI:. *Magn Reson Med*, 45:770–780, 2001.

H.-E. Assemlal, D. Tschumperlé, L. Brun, and K. Siddiqi. Recent advances in diffusion MRI modeling: Angular and radial reconstruction. *Med Image Anal*, 15(4):369–396, 2011.

P. J. Basser and S. Pajevic. Statistical artefacts in diffusion tensor MRI (DT-MRI) caused by background noise. *Magn Reson Med*, 44(1):41–50, 2000.

C.A. Clark, G.J. Barker, and P.S. Tofts. Magnetic resonance diffusion imaging of the human cervical spinal cord in vivo. *Magn Reson Med*, 41:1269–1273, 1999.

M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast and robust analytical q-ball imaging. *Magn Reson Med*, 58:497–510, 2007.

Z. Ding, J. C. Gore, and A. W. Anderson. Reduction of noise in diffusion tensor images using anisotropic smoothing. *Magn Reson Med*, 53(2):485–490, 2005.

R. Duits. *Perceptual Organization in Image Analysis.* PhD thesis, Eindhoven University of Technology, Eindhoven, The Netherlands, 2005.

R. Duits and E. Franken. Left-invariant diffusions on the space of positions and orientations and their application to crossing-preserving smoothing of HARDI images. *International Journal of Computer Vision*, 92(3):231–264, 2011.

R. Duits, E. Creusenand A. Ghosh, and T. Dela Haije. Diffusion, convection and erosion on $\mathrm{SE}(3)/(0 \times \mathrm{SO}(2))$ and their applications to the enhancement of crossing fibers. *ArXiv, http://arxiv.org/pdf/1103.0656v5*, 2011.

P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with Log-Euclidean metrics. *IEEE Trans Med Imaging*, 26(11):1472–1482, 2007.

P. T. Fletcher and S. Joshi. Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing*, 87:250–262, 2007.

E. M. Franken. *Enhancement of crossing elongated structures in images*. PhD thesis, Eindhoven University of Technology, 2008.

R.M. Heidemann, F. Fasano, M. Vogler, C. Leuze, J. Pfeuffer, and R. Turner. Improving image quality by combining outer volume suppression and parallel imaging: zoomed EPI with GRAPPA at 7T. In *Proceedings of the 16th Scientific Meeting ISMRM, Toronto.*, page 1284, 2008.

H. Johansen-Berg and T. E. J. Behrens. *Diffusion MRI: From Quantitative Measurement to In-Vivo Neuroanatomy*. Academic Press, 2009.

D. K. Jones and P. J. Basser. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. *Magn Reson Med*, 52:979–993, 2004.

D. K. Jones, M. A. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn Reson Med*, 42:515–525, 1999.

M. A. Koch and J. Finsterbusch. Compartment size estimation with double wave vector diffusion-weighted imaging. *Magn Reson Med*, 60(1):90–101, 2008.

D. Le Bihan. Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci*, 4(6):469–480, Jun 2003.

G. Lohmann, S. Bohn, K. Müller, R. Trampel, and R. Turner. Image restoration and spatial resolution in 7-tesla magnetic resonance imaging. *Magn Reson Med*, 64(1):15–22, 2010.

J. N. Morelli, V. M. Runge, T. Feiweier, J. E. Kirsch, K. W Williams, and U. I. Attenberger. Evaluation of a modified stejskal-tanner diffusion encoding scheme, permitting a marked reduction in te, in diffusion-weighted imaging of stroke patients at 3 t. *Invest Radiol*, 45(1):29–35, 2010.

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

S. Mori, B. J. Crain, V. P. Chacko, and P. C. M. van Zijl. Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. 45(2):265–269, 1999.

G. J. Parker, J. A. Schnabel, M. R. Symms, D. J. Werring, and G. J. Barker. Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging. *J. Magn. Res. Imag.*, 11:702–710, 2000.

P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Machine Intell*, 12(7):629–639, 1990.

C. Pierpaoli and P. J. Basser. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med*, 36(6):893–906, Dec 1996.

J. Polzehl and V. Spokoiny. Propagation-separation approach for local likelihood estimation. *Probab Theory Relat Fields*, 135:335–362, 2006.

J. Polzehl and K. Tabelow. Structural adaptive smoothing in diffusion tensor imaging: The **R** package *dti*. *J Stat Software*, 31(9):1–23, 2009.

J. Polzehl and K. Tabelow. Beyond the gaussian model of diffusion-weighted imaging: The package dti. *J Stat Software*, 44(12):1–26, 2011.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, 1992.

S. Sinha, U. Sinha, and V. R. Edgerton. In vivo diffusion tensor imaging of the human calf muscle. *J Magn Reson Imaging*, 24(1):182–190, 2006.

E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J Chem Phys*, 42:288–292, 1965.

K. Tabelow and J. Polzehl. *dti: DTI/DWI Analysis*, 2011. URL `http://CRAN.R-project.org/package=dti`. R package version 1.0.

K. Tabelow, J. Polzehl, V. Spokoiny, and H. U. Voss. Diffusion tensor imaging: Structural adaptive smoothing. *Neuroimage*, 39:1763–1773, 2008.

K. Tabelow, H. U. Voss, and J. Polzehl. Modeling the orientation distribution function by mixtures of angular central gaussian distributions. *J Neurosci Meth*, 203(1):200–211, 2012.

J. Weickert. *Anisotropic Diffusion in Image Processing*. ECMI. Teubner-Verlag, Stuttgart, 1998.

R. Xue, P. C. M. van Zijl, B. J. Crain, M. Solaiyappan, and S. Mori. In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging. *Magn Reson Med*, 42(6):1123–1127, 1999.