Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Modeling the orientation distribution function by mixtures of angular central Gaussian distributions

Karsten Tabelow[1], Henning U. Voss[2], Jörg Polzehl[1] ,

submitted: November 1, 2010

[1] Weierstrass Institute
for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117
Berlin, Germany
E-Mail: karsten.tabelow@wias-berlin.de
E-Mail: joerg.polzehl@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

No. 1559
Berlin 2010

2010 *Mathematics Subject Classification.* 62P10, 92C55, 62H11, 62H35, 62G05 .

Key words and phrases. diffusion weighted imaging, high angular resolution, orientation distribution function, angular central Gaussian distribution, tensor mixture model, order selection.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

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

Abstract

In this paper we develop a tensor mixture model for diffusion weighted imaging data using an automatic model selection criterion for the order of tensor components in a voxel. We show that the weighted orientation distribution function for this model can be expanded into a mixture of angular central Gaussian distributions. We show properties of this model in extensive simulations and in a high angular resolution experimental data set. The results suggest that the model may improve imaging of cerebral fiber tracts. We demonstrate how inference on canonical model parameters may give rise to new clinical applications.

1 Introduction

Since the introduction of its basic principles in the 1980's [Bihan and Breton, 1985, Merboldt et al., 1985, Taylor and Bushell, 1985] diffusion weighted imaging (DWI) has evolved into a versatile tool for in-vivo examination of tissue micro-structure in the human brain [Bihan et al., 2001] and spinal cord [Clark et al., 1999]. The large interest in this technique arises from the fact that nuclear magnetic resonance (NMR) is sensitive to *the diffusion of molecules*, usually water, in complex systems. Hence, DWI probes microscopic structures well beyond typical image resolutions through water molecule displacement. This is formally reflected by the three-dimensional Fourier transform relating the normalized diffusion weighted signal E to the diffusion propagator P, the spin displacement probability function. As the typical voxel extension is much larger than the length scale of the microscopic structure the diffusion weighted signal is always an aggregate quantity over a rich structure in the voxel volume.

Diffusion weighted data is measured using the pulsed gradient spin echo sequence [PGSE, Stejskal and Tanner, 1965] with a more or less large number of diffusion gradients and can thus be very high-dimensional. It was the development of diffusion tensor imaging [DTI, Basser et al., 1994a,b] that made DWI feasible for clinical and neuroscience applications. (See also Mori [2007] and Johansen-Berg and Behrens [2009] for an introduction.) In DTI, the high-dimensional diffusion weighted data is reduced to a Gaussian distribution model for free anisotropic diffusion, which can be considered as an approximation of hindered diffusion as observed in neuronal tissue [Tuch, 2002]. Within this model, diffusion is completely characterized by a rank-2 diffusion tensor $\mathcal{D}$, a symmetric positive definite 3×3 matrix with six independent components.

The diffusion tensor model describes diffusion completely if the microscopic diffusion properties within a voxel are homogeneous and non-restricted. In the presence of partial volume effects, such as two or more fiber bundles with different directions in one voxel, or more complex intra-voxel structure, such as crossing fibers, this model is only a rough approximation.

To be able to image complex fiber structures with approaches beyond the Gaussian diffusion model is of high value both in basic neuroscience and in clinical applications. Crossing fibers now can be resolved on MRI scanners optimized for clinical applications but not necessarily for performance in a research setting [Perrin et al., 2005, Yeh et al., 2010]. For example, using

a 4 Tesla MRI scanner and clinically feasible scan times, crossing fibers could be identified in the cingulum bundle, an important region in the pathology of various forms of dementia [Nezamzadeh et al., 2010]. In general, it can be anticipated that locally defined diffusion anisotropy indices, quantitative markers of fiber integrity that characterize the severity or progression of disease [Fritzsche et al., 2010], can be made more specific with non-Gaussian approaches. For example, in Batten disease, a neurodegenerative disease that causes premature death in children, a refined picture of the anatomic distribution of neurodegeneration [Dyke et al., 2007] may contribute to a better targeting of genetic or cellular therapies to appropriate locations in the brain [Sondhi et al., 2001]. Another example is brain surgery planning; although for most major white matter tracts DTI based tracking seems to be sufficient, surgical planning could benefit from resolving more complex fiber pathways at the interface between white and gray matter [Nimsky et al., 2007], as regions showing functional activity are often conveniently used as seed regions for fiber tracking.

There has been conducted a tremendous amount of research to overcome the limitations of DTI for the last decade [see e.g. Johansen-Berg and Behrens, 2009]. Most of the methods developed for DWI require a larger number of diffusion weighted gradients to be applied and hence increasing scanning time. While it seems impossible to give a complete overview here, we point the attention to a selection of methods, which are more related to this paper. Some of them try to overcome the limitations of DTI by modeling the diffusion weighted signal E with higher order tensors [Liu et al., 2003, Özarslan and Mareci, 2003], with a mixture of tensors [Tuch, 2002, Tournier et al., 2004, Alexander, 2005, Hosey et al., 2005, Özarslan et al., 2006, Alexander, 2006, Behrens et al., 2007, Jian et al., 2007, Leow et al., 2009], or by expanding the data into spherical harmonics [Frank, 2002].

Some methods explore the Fourier relation of the diffusion signal E and the diffusion propagator P directly to infer on the tissue micro-structure. For diffusion spectrum imaging (DSI) one measures diffusion weighted images on a Cartesian grid of diffusion gradients to perform the Fourier transform [Wedeen et al., 2005]. Hybrid diffusion imaging (HYDI) uses the measurements on spheres in q-space [Wu and Alexander, 2007]. Other methods exploit the mono-exponential signal decay [Özarslan et al., 2006] to determine the diffusion propagator, or the solution of the diffusion equation in a spherical harmonics expansion [Descoteaux et al., 2010].

Tuch [2004] showed that the radial projection of the diffusion propagator, the orientation distribution function (ODF), already contains information on the angular distribution of the diffusion directions. This gave rise to q-ball imaging (QBI) for which the measurement on a single q-shell is sufficient to determine the ODF. Again, using the spherical harmonics expansion [Anderson, 2005, Hess et al., 2006, Descoteaux et al., 2007] significantly reduces the computational effort for the determination of the ODF in QBI.

In this paper we analyze a *tensor mixture model* with a variable maximum order of mixture components depending on the number of measured diffusion weighted gradients. We consider a restricted tensor model with a re-parametrization of the general mixture model and use a *model selection criterion* to automatically select the *optimal number* of tensor components. It is usually thought [Johansen-Berg and Behrens, 2009] that tensor mixture models need a prespecified number of tensor components and are becoming unstable if more than two components are present [Tuch et al., 2002, Tournier et al., 2004]. A model selection for the number of components of a tensor mixture has been introduced so far only in the Bayesian context [Hosey et al., 2005, Behrens et al., 2007]. We will present extensive numerical and

experimental results to demonstrate the properties of our tensor mixture model. Furthermore we show that an appropriately defined weighted orientation distribution function (wODF) can be written as a mixture of *angular central Gaussian distribution functions* on the sphere. In contrast to spherical harmonics expansion, the components of this expansion are itself proper probability functions on the sphere and can be easily interpreted. Last but not least we present properties of quantities that naturally derive from the model and which could possibly used as diffusion anisotropy indices for clinical applications.

2 On the orientation distribution function

Let $P(\vec{r}, \vec{r}', \tau)$ denote the diffusion propagator, i.e., the probability for a particle (spin) to diffuse from position $\vec{r}'$ to $\vec{r}$ in time τ. Furthermore, with a slight abuse of notation, let $P(\vec{R}, \tau)$ be the ensemble average (EAP) over a (voxel) volume V

$$P(\vec{R}, \tau) = \int_{\vec{r}' \in V,\ \vec{R} = \vec{r} - \vec{r}'} P(\vec{r}, \vec{r}', \tau) p(\vec{r}') d\vec{r}' \tag{1}$$

of the propagator with the initial particle density $p(\vec{r}')$. Abstracting from the radial information the orientation distribution function (ODF) is defined as the radial projection of the diffusion propagator P [Tuch, 2002],

$$\psi(\vec{u}, \tau) = \int_0^\infty P(r\vec{u}, \tau) dr \ . \tag{2}$$

This function is defined on the unit sphere $|\vec{u}| = 1$. Note that this is basically the integral projection of the diffusion propagator onto the plane $\vec{u}^\perp$ perpendicular to $\vec{u}$ evaluated at the center.

For the special case of free anisotropic Gaussian diffusion with no boundaries covered by the diffusion tensor model the diffusion propagator is given by [Tuch, 2002]:

$$P(\vec{R}, \tau) = P(r\vec{u}, \tau) = \frac{1}{\sqrt{\det \mathcal{D}(4\pi\tau)^3}} \exp\left(-r^2 \frac{\vec{u}^T \mathcal{D}^{-1} \vec{u}}{4\tau}\right) \ ,$$

leading to the special form of the ODF [Tuch, 2004]

$$\psi(\vec{u}, \tau) = \frac{1}{\sqrt{\det \mathcal{D}(4\pi\tau)^3}} \sqrt{\frac{\pi\tau}{\vec{u}^T \mathcal{D}^{-1} \vec{u}}} \ .$$

In the special case of isotropic diffusion, i.e., $\mathcal{D} = D\mathcal{I}$, this further simplifies to

$$\psi(\vec{u}, \tau) = \frac{1}{4\pi} \frac{1}{2\tau D}$$

and the integral over the unit sphere S evaluates to

$$\int_{\mathcal{S}} \psi(\vec{u}, \tau) d\vec{u} = \frac{1}{2\tau D} \ .$$

Here, the denominator is the one-dimensional mean squared displacement as given by the Einstein relation [Einstein, 1905].

2.1 Weighted orientation distribution function

The orientation distribution function (2) is not dimensionless, depends on the diffusion time τ and has to be normalized for visualization [Tuch, 2002]. A close look quickly reveals that the initial definition (2) simply neglects the symmetry of the problem and the fact that P is a probability density. We therefore review a different definition as was first proposed by Wedeen et al. [2005] and further investigated by Aganj et al. [2010], and Barnett [2009].

The ensemble average $P(\vec{R}, \tau)$, see Eq. (1), of the diffusion propagator is a probability density:

$$\int_{\mathbb{R}^3} P(\vec{R}, \tau)\, d\vec{R} = 1 \ .$$

Expressing this integral in spherical coordinates $d\vec{R} = r^2 dr d\Omega_r$ with the radial part r and the solid angle Ω_r,

$$\int_{\mathcal{S}} \int_0^\infty r^2 P(r, \Omega_r, \tau)\, dr\, d\Omega_r = 1 \ ,$$

the radial integral, which is the (spherical) projection of P onto the (unit) sphere $\mathcal{S}$, can be defined as the *weighted (or marginal) orientation distribution function* (wODF) $\psi_w(\Omega_r, \tau) = \psi_w(\vec{u}, \tau)$:

$$\psi_w(\Omega_r, \tau) = \int_0^\infty r^2 P(r, \Omega_r, \tau)\, dr = \int_0^\infty r^2 P(r\vec{u}, \tau)\, dr \ .$$

Note the additional weighting factor r^2 compared to the usual ODF definition (2) resulting from the spherical projection. This definition is intrinsically normalized, i.e.,

$$\int_{\mathcal{S}} d\Omega_R\, \psi_w(\Omega_R, \tau) = 1 \ .$$

In the special case of anisotropic Gaussian diffusion the wODF is given by

$$\psi_w(\vec{u}, \tau) = \frac{1}{4\pi\sqrt{\det \mathcal{D}}} \sqrt{\frac{1}{(\vec{u}^T \mathcal{D}^{-1} \vec{u})^3}} \ .$$

2.2 wODF for elliptical distributions

Definition. A probability density P is said to be in the class of elliptical distributions in 3 dimensions, if

$$P(\vec{R}) = (\det \Psi)^{-1/2} h\left((\vec{R} - \vec{\mu})^\top \Psi^{-1} (\vec{R} - \vec{\mu})\right)$$

for some function h, a 3×3 positive definite matrix Ψ and some centrality parameter $\vec{\mu}$.

Lemma. *For any elliptical distribution P with $\vec{\mu} = 0$ the weighted orientation distribution function is*

$$\psi_w(\vec{u}) = \frac{1}{4\pi} (\det \Psi)^{-1/2} \sqrt{\frac{1}{(\vec{u}^\top \Psi^{-1} \vec{u})^3}} \ .$$

Proof: With substitution $\widetilde{r} = r(\vec{u}\Psi^{-1}\vec{u})^{1/2}$ we have

$$\begin{aligned}
1 &= \int_{\mathcal{S}} \int_0^\infty r^2 P(r\vec{u}) dr d\vec{u} \\
&= \int_{\mathcal{S}} \int_0^\infty \widetilde{r}^2 (\vec{u}\Psi^{-1}\vec{u})^{-1} (\det \Psi)^{-1/2} h(\widetilde{r}^2) (\vec{u}\Psi^{-1}\vec{u})^{-1/2} d\widetilde{r} d\vec{u} \\
&= \int_{\mathcal{S}} \frac{1}{4\pi} (\det \Psi)^{-1/2} (\vec{u}\Psi^{-1}\vec{u})^{-3/2} d\vec{u} \int_0^\infty 4\pi\, \widetilde{r}^2 h(\widetilde{r}^2) d\widetilde{r}.
\end{aligned}$$

the first integrand is the density of the *angular central Gaussian distribution* [Mardia and Jupp, 2000], a probability distribution on the sphere $\mathcal{S}$. Hence both integrals equal 1 and the second integral does not depend on h.

Since the Weibull distribution is an elliptical distribution, the result can e.g. be directly applied to the model in Jian et al. [2007].

Remark: Note that for the unweighted ODF, the result is

$$\psi(\vec{u}) = (\det \Psi)^{-1/2} \sqrt{\frac{1}{\vec{u}^\top \Psi^{-1} \vec{u}}} \int_0^\infty h\left(r^2\right)\, dr \ ,$$

but the integral over r cannot be calculated in general.

3 Estimating the mixture of diffusion tensors

The mixture model of p diffusion tensors $\mathcal{D}_k$ in its general form [see e.g. Tuch, 2002, Alexander, 2005, 2006, Özarslan et al., 2006],

$$E(\vec{q}) = \sum_{k=1}^{p} w_k \exp(-b\vec{q}^\top \mathcal{D}_k \vec{q}), \quad w_k \geq 0, \quad \sum_{k=1}^{p} w_k = 1 \ ,$$

is too flexible, leading to severe identifiability and numerical problems [Tuch, 2002, Johansen-Berg and Behrens, 2009]. Several authors therefore suggest to restrict the number of components p to 2 [Alexander, 2005, 2006, Özarslan et al., 2006] or to approximate a solution by estimating the ODF spherical harmonics expansion and subsequent deconvolution [Tournier et al., 2004]. Others consider special continuous tensor mixture distributions [Jian et al., 2007], or restrictions for the tensor eigenvalues [Leow et al., 2009].

3.1 Tensor mixture model

We reduce the complexity of the model by considering all diffusion tensors $\mathcal{D}_k$ in a voxel (but not in different voxel) to have identical eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ and to be prolate [$\lambda_1 > \lambda_2 = \lambda_3$, Basser et al., 1994a] as in Tournier et al. [2004]. The rationale behind this model is first the assumption of anisotropic Gaussian diffusion in each fiber component of a voxel with common voxel-specific diffusivity but distinct directions and second the rotational symmetry of the diffusion in the plane perpendicular to the main diffusion direction ($\lambda_2 = \lambda_3$). The reduced model has the form

$$E(\vec{q}) = \sum_{k=1}^{p} w_k \exp\left(-b\vec{q}^\top \left((\lambda_1 - \lambda_2) d_k d_k^\top + \lambda_2 \mathcal{I}\right) \vec{q}\right), \quad w_k \geq 0, \quad \sum_{k=1}^{p} w_k = 1 \ ,$$

where $d_k = (\sin(\phi_k)\cos(\eta_k), \sin(\phi_k)\sin(\eta_k), \cos(\phi_k))$ are the main eigen vectors of the tensors characterizing the main diffusion direction in the corresponding fiber compartments and $\lambda_1 > \lambda_2 > 0$ are the common eigenvalues characterizing the shape of the prolate tensors.

The model can be re-parametrized in the form

$$E(\vec{q}) = \sum_{k=1}^{p} \widetilde{w}_k \exp(-b\vartheta(\vec{q}^\top d_k)^2), \quad \widetilde{w}_k \geq 0, \quad \sum_{k=1}^{p} \widetilde{w}_k < 1 \quad \vartheta > 0,$$

with $\lambda_2 = -\frac{1}{b} log \sum_{k=1}^{p} \widetilde{w}_k$, $\lambda_1 = \vartheta + \lambda_2$ and $w_k = \widetilde{w}_k / \sum_{l=1}^{p} \widetilde{w}_l$. In this reformulation the model has $2p+1$ nonlinear parameters and p linear parameters with constraints.

Parameter estimates for $\theta = (\vartheta, \widetilde{w}, \phi, \eta)$ are obtained by solving the optimization problem

$$\widehat{\theta} = \operatorname{argmin}_{\substack{\vartheta > 0 \\ 0 \leq \phi < \pi \\ 0 \leq \eta < 2\pi}} \operatorname{argmin}_{\widetilde{w}_k \geq 0} \sum_{j=1}^{N} \left(E(\vec{q}_j) - \sum_{k=1}^{p} \widetilde{w}_k \exp(-b\vartheta(\vec{q}_j^{\,\top} d_k)^2) \right)^2 \tag{3}$$

with N diffusion gradient directions $\vec{q}_j$. This is a separable nonlinear optimization problem [see e.g. Mullen, 2008]. The inner optimization is a linear problem with constraints and can be solved efficiently using e.g. the non-negative least squares algorithm of Lawson and Hanson [1995].

The choice of initial estimates for the optimization problem (3) may be crucial for the convergence of the utilized optimization algorithm. For a specified maximal order p_{max} we choose the initial estimates using information obtained within the tensor model. Let $\lambda_{il} (l = 1, \ldots, 3)$ be the eigenvalues of the estimated diffusion tensor in voxel i. The parameter ϑ_i is initialized as the maximum of $\lambda_{i1} - 0.5(\lambda_{i2} + \lambda_{i3})$ and the 0.1-quantile of these values for all voxel with tensor fractional anisotropy (FA) larger than 0.3. Initial angles ϕ_{ik} and η_{ik} in voxel i are obtained by minimizing

$$\operatorname{argmin}_{\widetilde{w}_k \geq 0} \sum_{j=1}^{N} \left(E(\vec{q}_{ij}) - \sum_{k=1}^{p_{max}} \widetilde{w}_k \exp(-b\vartheta_i(\vec{q}_j^{\,\top} d_k)^2) \right)^2$$

over a set of randomly chosen directions $d_k (k = 1, \ldots, p_{max})$ from a uniform distribution on the sphere. In case of tensor FA larger than 0.3, the first principal direction obtained within the tensor model is explicitly included as candidate direction. The sampling of directions is done under the constraint $|d_k^\top d_l| < 0.866$ for $k \neq l$ corresponding to a minimum angle of directions of 30 degrees.

The optimization problem (3) is then solved starting from these initial estimates using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method with analytic gradients.

3.2 Model selection

For a sequence of mixture orders p decreasing from a maximum value p_{max} to 1 we get a set of nested tensor mixture models. In order to select a model providing a *best* description of the data we employ information based criteria that penalize for overfitting. Suitable criteria in this context are of the form

$$C(p, N) = -2N \log L(\widehat{\theta}) + \Phi(p, N) \ ,$$

where $L(\widehat{\theta})$ is the likelihood function evaluated at the parameter values estimated by maximum likelihood, i.e., here for a Gaussian error model. The selected model will minimize $C(p, N)$. The term $\Phi(p, N)$ penalizes for the higher uncertainty of parameter estimates and overfitting associated with the complexity of the respective model.

The choice of the penalty $\Phi(p, N)$ depends on the objective of the model selection, i.e., the definition of a *best* model. The *Akaike information criterion (AIC)* [Akaike, 1974] employs

$$\Phi(p, N) = 2(3p + 1) \ ,$$

with $3p + 1$ being the number of parameters in the model, and focuses on a best description of the data by the chosen model in terms of the expected Kullback-Leibler discrepancy, i.e., $-2\mathbf{E} \log L(\widehat{\theta})$. It provides a model that is suitable for the prediction of values of $E(\vec{q})$ in additional gradient directions. A modification of AIC using

$$\Phi(p, N) = N \frac{1 + (3p + 1)/N}{1 - (3p + 3)/N}$$

(AIC$_c$) [Hurvich and Tsai, 1989] is based on a second order approximation of the expected Kullback-Leibler discrepancy and corrects for a small sample (small N) bias of AIC.

An alternative maximizing the probability of selecting the correct model if one of the candidate models is adequate is the *Bayes information criterion (BIC)* [Schwarz, 1978] with

$$\Phi(p, N) = \log(N)(3p + 1) \ .$$

BIC should be used for model selection if the interest is in selecting the order of mixture components. For a more comprehensive discussion see Linhart and Zucchini [1986], Burnham and Anderson [2002], Claeskens and Hjort [2008].

We select an appropriate model order p in each voxel separately by evaluating the model selection criteria, i.e., BIC, AIC or AIC$_c$, within the following sequential approach:

For each voxel i

1. Initialize $p_{best} = 0$ and compute $C(p_{best}, N)$,
2. Set model order to $p = p_{max}$ and initialize estimates of $\vartheta_i^{(p)}$, $\phi_{ik}^{(p)}$ and $\eta_{ik}^{(p)}$ as described above.
3. Estimate parameters for the model of order p by solving (3) using BFGS optimization with analytic gradients.
4. If $C(p, N) < C(p_{best}, N)$ then set $p_{best} = p$.
5. Re-order $\widehat{\phi}_{ik}^{(p)}$ and $\widehat{\eta}_{ik}^{(p)}$ $(k = 1, \ldots, p)$ in descending order of weights $\widehat{\widetilde{w}}_{ik}^{(p)}$ and set p to the minimum of $p - 1$ and the number of non-zero weights $\widehat{w}_{ik}^{(p)}$. Specify the estimate of ϑ and the first, after the re-ordering, p estimated angles ϕ and η as initial values and continue in step 3 if $p > 0$.

The algorithm provides us with an estimated mixture order p_{best} and parameter estimates $\widehat{\theta}_i^{p_{best}} = (\widehat{\vartheta}_i^{p_{best}}, \widehat{\widetilde{w}}_{ik}^{p_{best}}, \widehat{\phi}_{ik}^{p_{best}}, \widehat{\eta}_{ik}^{p_{best}})$ for $k = 1, \ldots, p_{best}$ in voxel i.

3.3 wODF estimate

Given the estimates for the mixture order p_{best} and the parameters $\widehat{\theta}_i = \widehat{\theta}_i^{p_{best}}$ the weighted orientation distribution function (wODF) can be written as

$$\psi_w(\vec{u}) = C \sum_{k=1}^{p_{best}} \widehat{w}_k \sqrt{\frac{1}{(\vec{u}^\top \mathcal{D}_k^{-1} \vec{u})^3}}$$

with $C \equiv \frac{1}{4\pi}(\det \mathcal{D}_k)^{-1/2} = \frac{1}{4\pi}\widehat{\lambda}_1^{-1/2}\widehat{\lambda}_2^{-1}$ and $\mathcal{D}_k = \widehat{\vartheta}\widehat{d}_k\widehat{d}_k^\top - \frac{1}{b}log\sum_{k=1}^{p}\widehat{\widetilde{w}}_k\mathcal{I} = (\widehat{\lambda}_1 - \widehat{\lambda}_2)\widehat{d}_k\widehat{d}_k^\top + \widehat{\lambda}_2\mathcal{I}$. The wODF in the considered model is therefore a mixture of *angular central Gaussian distributions* [Mardia and Jupp, 2000], where the directional parameters $\widehat{d}_k$ can be directly interpreted as main fiber direction within the voxel.

4 Inference on the tensor mixture

The tensor mixture model as described in this paper provides the possibility to define several contrasts. While it is beyond the scope of this paper to evaluate their clinical relevance we here define the *effective order of the model* (EO), the *fractional anisotropy* (FA), and the *eigenvalues* (EV).

4.1 Effective order of the model (EO)

Consider e.g. the situation that in some voxel a model of order $p_{best} = 3$ has been selected by the model selection criterion described in the preceeding section. The weights w_k of the three components can be different. Intuitively we would probably agree the model to be of order 3 if all weights are approximately $1/3$, but would consider the order very close to 2 if two weights are say 0.48 and the third is only 0.04.

Let $w_1 \geq w_2 \geq \cdots \geq w_{p_{best}}$ be the ordered estimated weights in a tensor mixture model with maximum order p_{best}. We define the effective order (EO) of an estimated tensor mixture model as

$$EO = \sum_{k=1}^{p_{best}-1} k^2(w_k - w_{k+1}) = \sum_{k=1}^{p_{best}} (2k-1)w_k \ .$$

It is easy to show that $\sum_{k=1}^{p_{best}} w_k = 1$ yields $1 \leq EO \leq p_{best}$ with $EO = p_{best}$ if and only if $w_k \equiv 1/p_{best}$.

4.2 Fractional anisotropy (FA)

The fractional anisotropy for the single diffusion tensor model [Mori, 2007] canonically simplifies in the tensor mixture model as $\lambda_2 = \lambda_3$ and all mixture components are defined by the same tensor:

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \langle\lambda\rangle)^2 + 2\cdot(\lambda_2 - \langle\lambda\rangle)^2}{\lambda_1^2 + 2\cdot\lambda_2^2}} = \frac{(\lambda_1 - \lambda_2)}{\sqrt{\lambda_1^2 + 2\cdot\lambda_2^2}} \ .$$

Here, $\lambda_1 \geq \lambda_2$ are the (ordered) eigenvalues, with the mean diffusivity $\langle\lambda\rangle = \frac{1}{3}(\lambda_1 + 2\lambda_2)$.

4.3 Eigenvalues (EV)

The eigenvalues of a single fiber diffusion tensor are directly related to the diffusion coefficients in the main diffusion direction and in the plane perpendicular to this direction. The tensor mixture model considered here tries to model different fiber components in a voxel. The relation between the estimated eigenvalues λ_1 and λ_2 and the associated diffusion coefficients can therefore be expected to be more realistic than in the single diffusion tensor model.

5 Examples

We first investigate properties of the tensor mixture model by simulation experiments before applying it to experimental DWI data.

Computation are performed with our R-package dti [Tabelow and Polzehl, 2010] (version 0.9-3) on an Intel(R) Xeon(R) CPU 5160 3GHz and 24 GByte of memory. The operating system was OpenSuse 11.0 with R version 2.10.1. The package is freely available on CRAN (`http://cran.r-project.org`) and NITRC (`http://www.nitrc.org`).

Our current implementation uses R with FORTRAN. Computation time for 1.4 Million voxel in the experimental data has been approximately $6h$ for $p_{max} = 2$, $15h$ for $p_{max} = 3$, $26h$ for $p_{max} = 4$, and $40h$ for $p_{max} = 5$. The implementation has the potential of speeding up by a factor $3 - 4$ as first trials showed. This re-implementation will be part of one of the next version of the R-package dti making it more feasible for practical applications. Further significant speed up can achieved by parallelization of the algorithm.

5.1 Simulation

In a first setting we generate mixtures of two prolate tensors with fractional anisotropy $FA = 0.8$, product of eigenvalues $\lambda_1\lambda_2^2 = 3.2$, $b = 1$ (Note, that the units can be chosen arbitrarily, as long as $b\lambda$ is set to realistic values.) and mixture coefficients $w_k \equiv 0.5$. Diffusion weighted data are generated for different angles δ between the main tensor orientations (91 values from 0 to $\pi/2$) using 35, 60 and 140 gradient directions and SNR = 30, with SNR defined as ratio between non-diffusion weighted image S_0-value and error standard deviation in K-space. Gradient directions are chosen to minimize symmetrized Coulomb forces on a sphere following a proposal by [Jones et al., 1999]. For each angle we performed 1000 simulations. The model order was selected by BIC. The results for 35, 60 and 140 gradient directions are summarized in the left, central and the right column of Figure 1. The upper row illustrates the accuracy of the estimated angle between mixture components while the lower row shows the estimated effective order. Density values for varying δ are coded by greyvalues. Additionally $0.025, 0.25, 0.5, 0.75$ and 0.975 quantiles are provided as functions of δ. The results suggest that for the specified SNR and FA the tensor mixture model provides an angle resolution of about 40 degrees. Additional simulations show that angular resolution is mainly limited by FA and number of measured gradients.

Figure 2 provides results for varying number of gradients, ranging from 12 to 150 in a similar form for mixture order 2 and 3 models, FA and eigenvalues as before and SNR = 30. Mixture

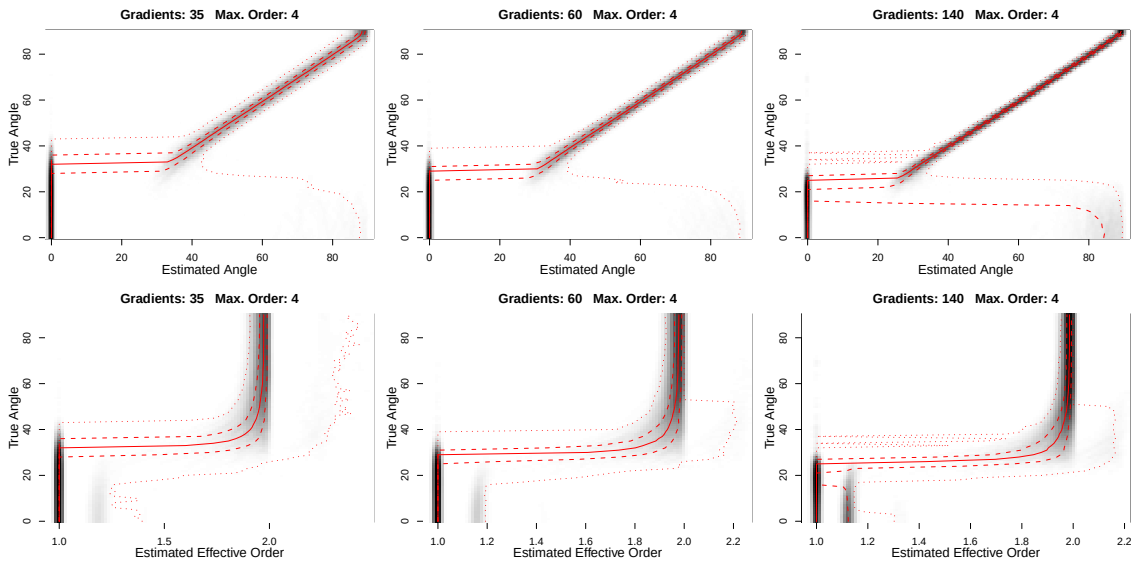

Figure 1 – *Density of estimated angle between mixture components (top) and estimated order of mixtures (bottom) for true angles $\delta = 60$ ranging from 0 to 90 degrees. SNR = 30, mixture coefficients $w_k \equiv 0.5$, and true model order $p = 2$. From left to right: results for 35, 60 and 140 gradient directions. Densities are estimated from 1000 simulations with maximum order 4 models. Dotted curves correspond to 0.025 and 0.975 quantiles, dashed curves to quartiles and solid curves to the median of the estimated quantities.*

coefficients are $w_k \equiv 0.5$, $w_k \equiv 1/3$, and angles between directions are $\delta = 60$ and $\delta_{1,2}, \delta_{1,3} = 60, \delta_{2,3} \approx 75$, for models of order 2 and 3, respectively. A maximum number of 4 or 5 mixture components was specified and the model order again selected by BIC.

In a third simulation we investigate the quality of reconstruction in dependence of SNR, true number of mixtures and specified maximum order. The number of gradients was fixed to 60, $FA = 0.8$ and $\lambda_1\lambda_2^2 = 3.2$. For mixtures of order 2 the true angle is 60 degrees while for mixtures of order 3 the three angles between tensor orientations are 60, 60 and approximately 75 degrees respectively. Results in terms of estimated angle between the first two tensor orientations and estimated effective order are provided in Figure 3. The upper two rows correspond to mixtures of order 2 while the lower two rows show corresponding results for mixtures of order 3. Note that in case of two mixture components the effective order and angle are correctly identified for SNR > 15 and that results do not vary with the specified maximum model order. For mixtures with three components the correct model is found for SNR > 25 and a specified maximum number of components ≥ 3. Results are slightly improved in case of a larger specified maximum number of components.

5.2 Experimental data

5.2.1 Methods

MRI images were obtained from a healthy male volunteer in the age group 40 - 45 within an Institutional Review Board approved research protocol at Weill Cornell Medical College. Images were acquired on a 3.0 Tesla General Electric Excite MRI scanner using an 8-channel

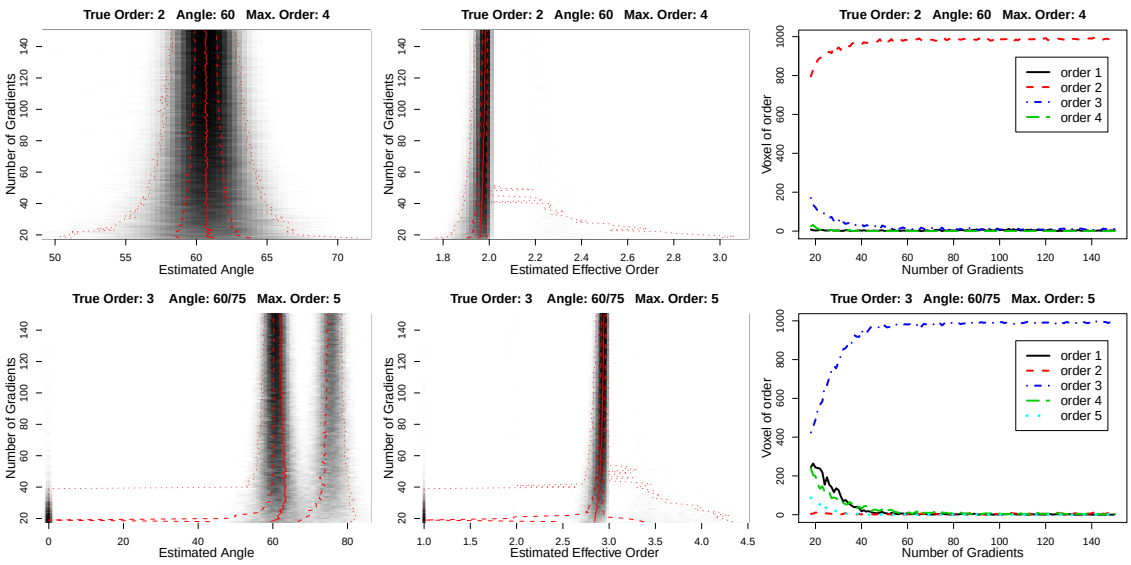

Figure 2 – *Density of estimated angle between mixture components (left), estimated order of mixtures (center) and number of voxel with estimated order for varying number of gradients (18-150) (right). Specifications for first row are: angle $\delta = 60$ degrees, SNR = 30, mixture coefficients $w_k \equiv 0.5$, true model order $p = 2$ and maximal model order 4. Bottom row: angles $\delta_{1,2}, \delta_{1,3} = 60, \delta_{2,3} \approx 75$ degrees, SNR = 30, mixture coefficients $w_k \equiv 1/3$, true model order $p = 3$ and maximal model order 5. Densities are estimated from 1000 simulations. Dotted curves in the left and center plots correspond to 0.025 and 0.975 quantiles, dashed curves to quartiles and solid curves to the median of the estimated quantities.*

receive-only head coil. First, a localizer scan was obtained to prescribe the position of the subsequent DWI scan. For the DWI scan, a single-shot spin-echo EPI sequence with 10 images without diffusion weighting and 140 diffusion gradient directions, which were approximately isotropically distributed over the sphere, was used, with an echo and repetition time of TE = 73.2 ms and TR = 14000 ms, respectively. 66 axial slices were scanned with no skip and an acquisition matrix size of 128 $\times$ 128. Images were zero-filled to an image matrix size of 256 $\times$ 256, yielding an effective resolution of 0.898 $\times$ 0.898 $\times$ 1.800 mm^3. The b-value in the diffusion weighted images was 1000 s/mm^2, the parallel imaging acceleration factor was 2, and the total scan time for this scan was 36 min.

We now show several examples of the results of the tensor mixture model of maximum order 4 in brain regions that have been shown to contain interesting crossing structures previously. Computation has been performed in voxel within a mask for the brain determined by thresholding the S_0-image.

5.2.2 Results

Using the tensor mixture model described in this paper we show in the Figures 4-7 results for several brain regions. Figure 8 illustrates results within a spatially extended region where a voxelwise analysis identifies tensor mixtures of order 3, i.e. crossings of three fibers.

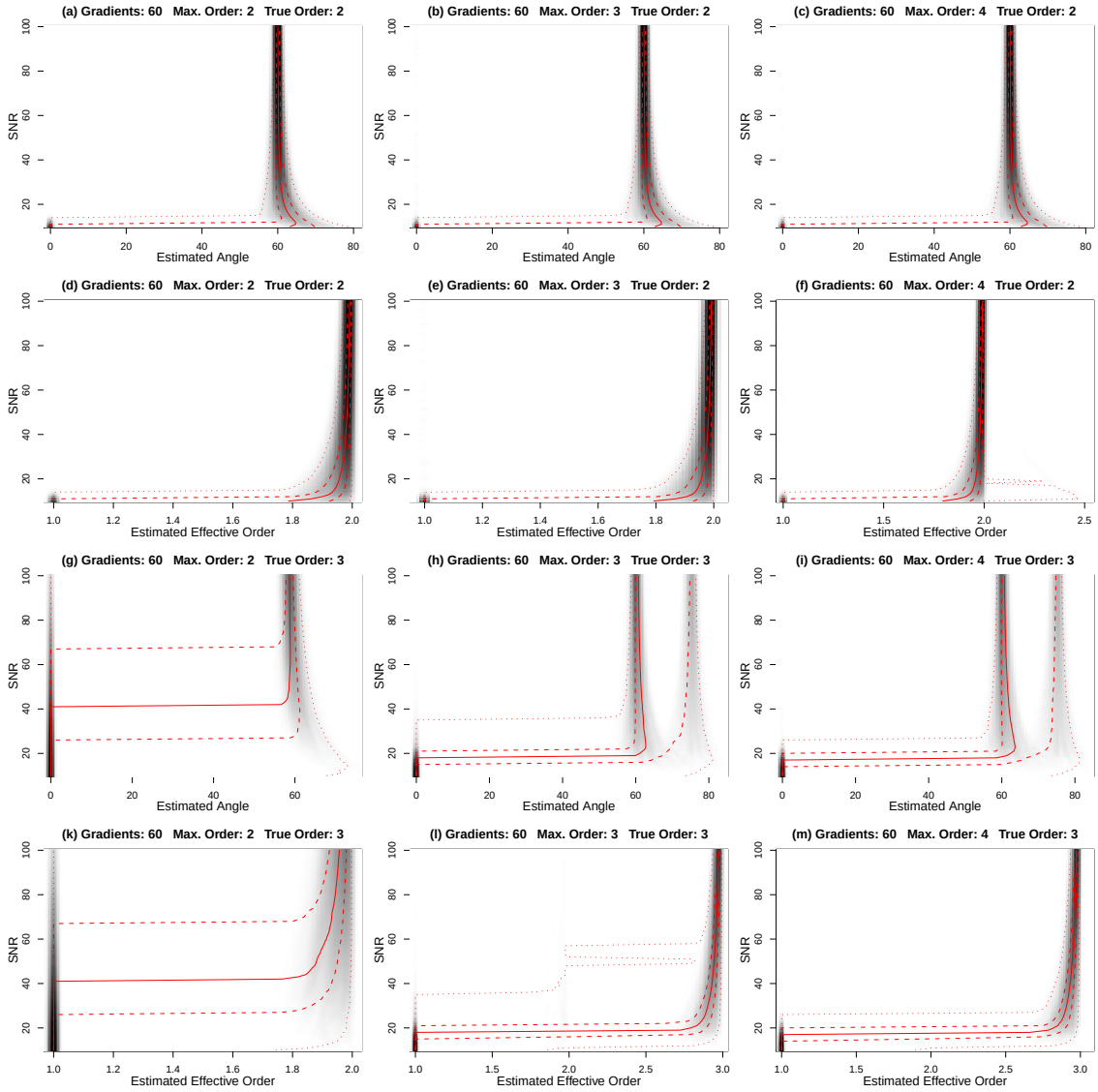

Figure 3 – *Density of estimated angle between mixture components (a-c,g-i) and estimated order (d-f, k-m) for varying SNR* $\in (10, 100)$*, true model order 2 (a-f) and 3 (g-m) and specified maximum number of mixture components 2 (a,d,g,k), 3 (b,e,h,l) and 4 (c,f,i,m). Results are based on 1000 simulations using 60 gradients. For order* $p = 2$ *models the angle between main directions is* $\delta = 60$ *degrees while for order* $p = 3$ *models the angles between directions are* $\delta_{1,2}, \delta_{1,3} = 60$ *and* $\delta_{2,3} \approx 75$ *degrees. Mixture coefficients were specified as* $w_i \equiv 1/p$*. Dotted curves correspond to* $0.0.25$ *and* 0.975 *quantiles, dashed curves to quartiles and solid curves to the median of the estimated quantities.*

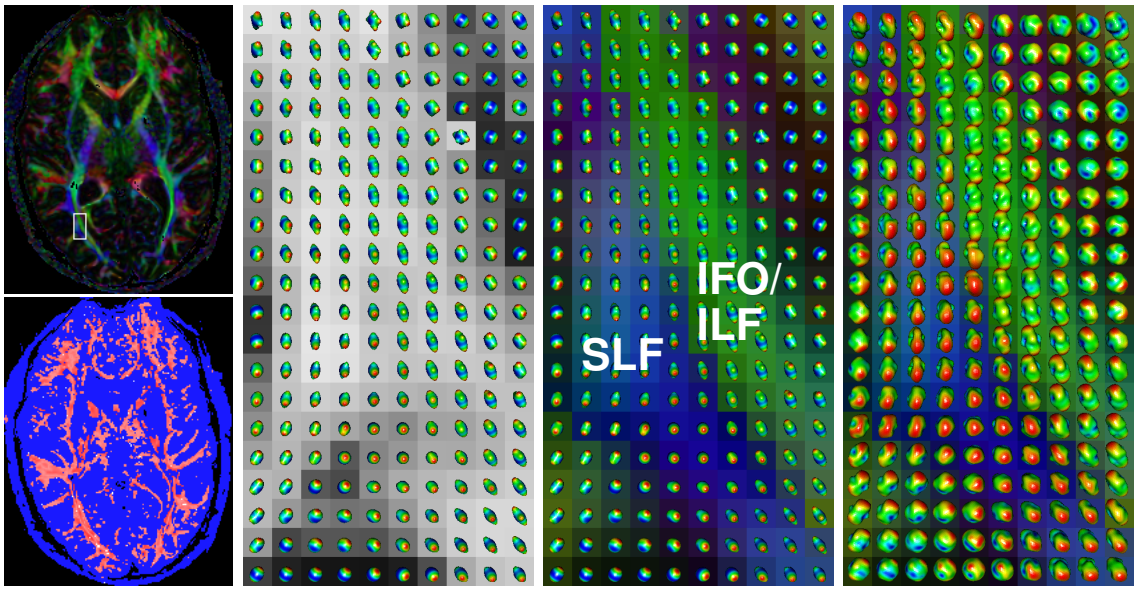

Figure 4 – *Example of two adjacent fiber tracts oriented in an angle to each other. Labeled are the inferior fronto-occipital fasciculus (IFO, mostly green), the inferior longitudinal fasciculus (ILF, mostly green), and the superior longitudinal fasciculus (SLF, mostly blue). From left to right: Color coded directional FA map from single diffusion tensor model (upper). Effective order (lower, for used color bar cf. Figure 9). Estimated wODF obtained using the tensor mixture model described in this paper over estimated FA from the same model. Same but over color coded directional FA map from single diffusion tensor model. For comparison the estimated wODF using spherical harmonics expansion [Aganj et al., 2010] is shown in the rightmost figure. For color, see online version of this paper.*

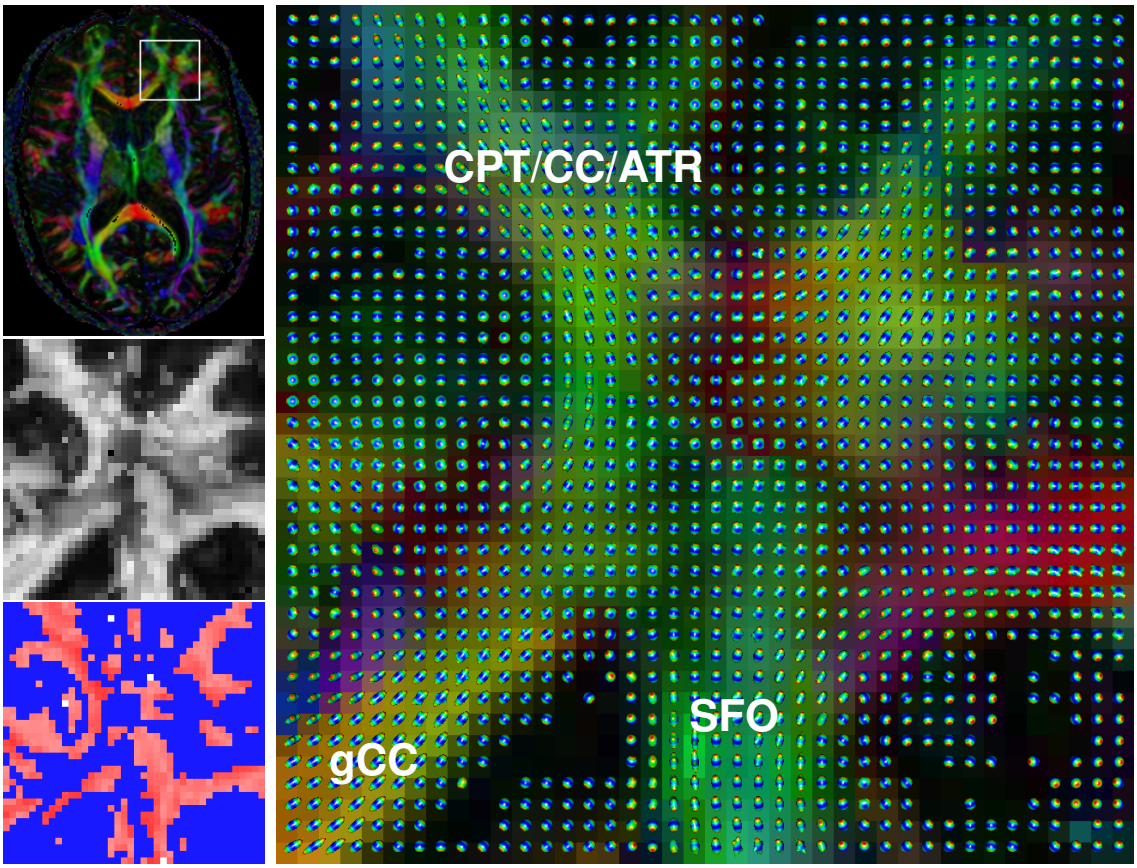

Figure 5 – *Example for the intersection of fiber tracts in the frontal part of the brain. Labeled are the genu of the corpus callosum (gCC), the corpus callosum (CC), the corticopontine tract (CPT), the anterior thalamic radiation (ATR), and the superior fronto-occipital fasciculus (SFO). This example has been inspired by [Descoteaux et al., 2010]. Left column from top to bottom: Color coded directional FA map from single diffusion tensor model, estimated FA from tensor mixture model and estimated effective order for the specified subregion. Note, that the blue area corresponds to tensor order 1. Right subfigure: estimated wODF over color coded directional FA map from single diffusion tensor model in selected subregion. wODF with estimated FA smaller than 0.05 are not shown. For color, see online version of this paper.*

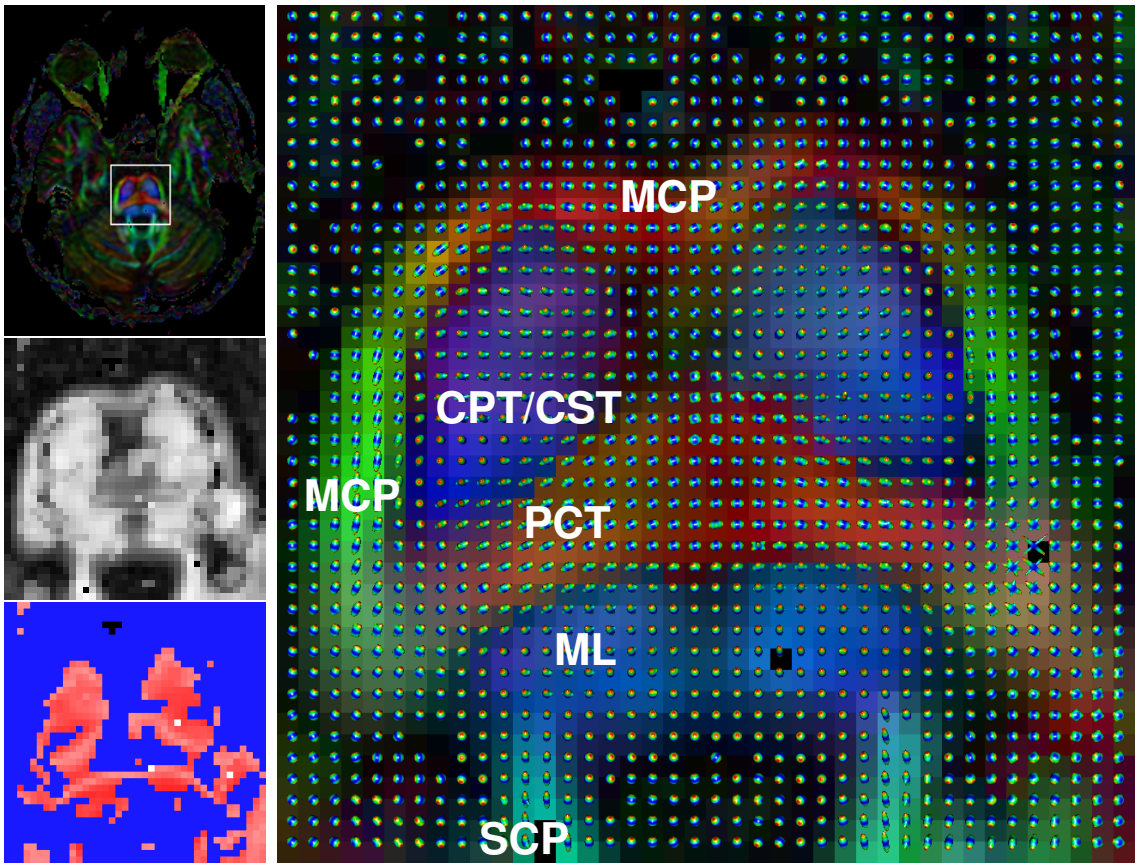

Figure 6 – *Fiber tracts in the midbrain. In this area there should be many voxel with partial voluming effects due to the dense packing of fiber tracts, and the estimation of the number of tensor mixture components is expected to be particularly useful. Labeled are: Corticopontine tract (CPT), corticospinal tract (CST), medial lemniscus (ML), middle cerebellar peduncle (MCP), pontine crossing tract (CPT), superior cerebellar peduncle (SCP). Figure organization as in Figure 5. Note, that the blue area corresponds to tensor order 1. wODF with estimated FA smaller than 0.05 are not shown. For color, see online version of this paper.*

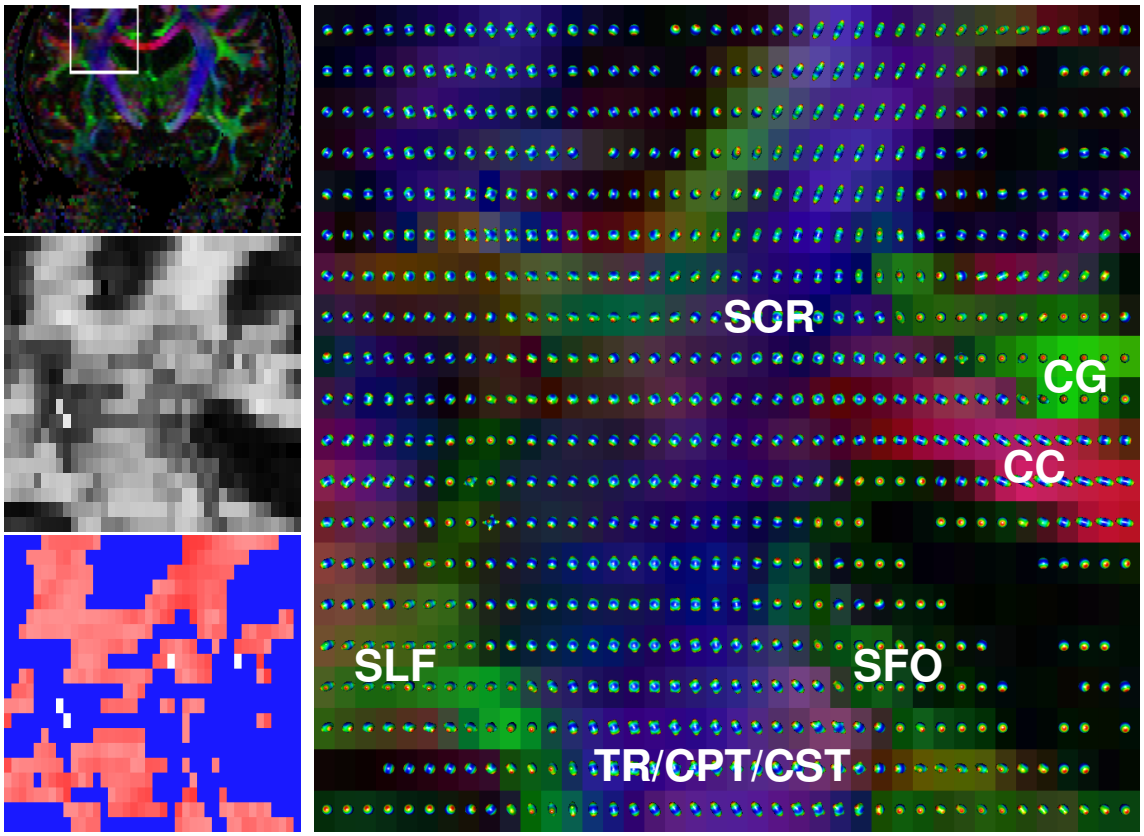

Figure 7 – *Some of the major fiber bundles in the brain. Labeled are: Cingulum (CG), corpus callosum (CC), corticopontine tract (CPT), corticospinal tract (CST), superior corona radiata (SCR), superior fronto-occipital fasciculus (SFO), superior longitudinal fasciculus (SLF), thalamic radiation (TR). This example has been inspired by [Kaden et al., 2008]. Figure organization as in Figure 5. Note, that the blue area corresponds to tensor order 1. wODF with estimated FA smaller than 0.05 are not shown. For color, see online version of this paper.*

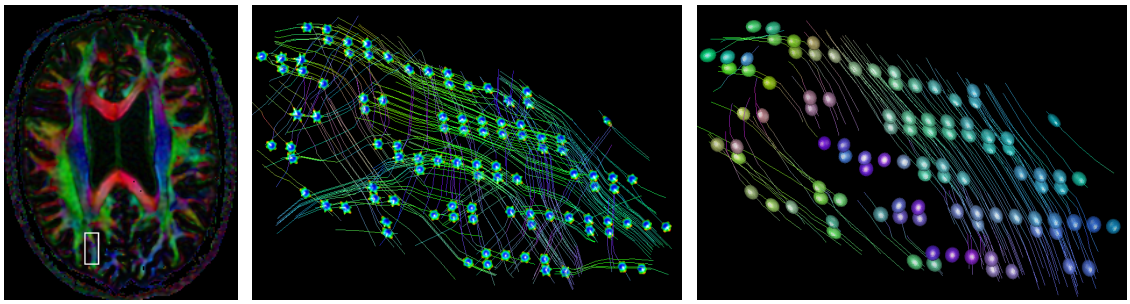

Figure 8 – *Fiber tracking at the interface between gray and white matter near the forceps major, where, within a contiguous set of 99 voxel, the number of mixture components was selected as 3. Left: Color coded FA map and intersection with a region of interest. The region exceed over 5 slices (2 below and 2 above the slice shown). Central: estimated ODF's for voxel with estimated order 3 and local fiber tracks starting from these voxel. Right: estimated tensors for the same set of voxel and local fiber tracking results using the tensor model. For color, see online version of this paper.*

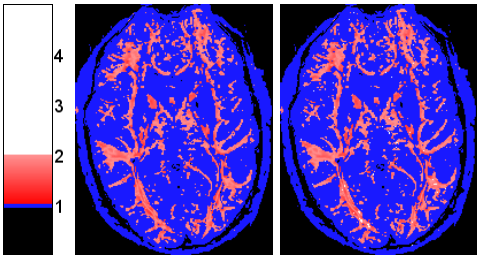

Figure 9 – *Effective order of the estimated tensor mixture model in a slice in the experimental data as described in the text. The maximum order for the model is 2 (left) and 5 (right). The color scheme is normalized to the same range in all cases (see legend on left; blue corresponds to order 1 models, different intensities of red to effective order between 1 and 2, white for orders above 2). Thus, independent from the chosen maximum order, the estimated order is a "stable" quantity. For color, see online version of this paper.*

maxcomp	0	1	2	3	4	5
1	159	1454401	0	0	0	0
2	595	1193761	260204	0	0	0
3	167	1196071	255677	2645	0	0
4	150	1193221	258180	2937	72	0
5	150	1192515	258760	3086	27	22

Table 1 – *Number of voxel in the experimental data set with estimated order of tensor mixture model depending on the maximum number of tensor components (maxcomp). One can see that increasing the maximum order only slightly affects the results for lower orders. A model of order 0 (isotropic tensor, $\lambda_1 = \lambda_2 = \lambda_3$) has been assigned, if it provides the best fit.*

5.2.3 Inference on experimental data

At first glance one could think that increasing the maximum order for the model estimation would include more and more model components, due to the noise, but in fact it does not. The model selection rule described in this paper leads to stable estimates of the tensor mixture model order as can be seen in Figure 9 and Table 1. The effective order is a smooth quantity, which also includes the effect of partial volume effects due to the limited voxel resolution in the DWI compared to the mean size of the anisotropic structure of the brain.

In Table 2 we show the number of voxel with an estimated order of the tensor mixture depending on the number of gradients by selecting a subset of diffusion weighted images from the complete data set. These are chosen such that the corresponding gradient directions are approximately uniformly distributed on the sphere. Requiring five measured values per model parameter as a rule of thumb the maximum number of components for the algorithm can be restricted to $\frac{N-5}{15}$ to ensure sufficient information for the parameter estimation and model selection. For comparison we show results for computations with maximum order 5 for all cases.

In Figure 10 we show a comparison between the fractional anisotropy in the diffusion tensor model and in the tensor mixture model described in this paper. Due to largely reduced partial volume effects and a more appropriate (while still not fully correct) modeling, the fractional anisotropy in the tensor mixture model is larger than in the single diffusion tensor model and much more homogeneous in brain white matter, probably reflecting more precisely the

number of gradients	maxcomp	0	1	2	3	4	5
140	5	1	4984	2985	30	0	0
105	5	2	5420	2562	16	0	0
70	4	2	6117	1872	9	0	0
70	5	2	6115	1874	9	0	0
56	3	1	6412	1576	11	0	0
56	5	0	6391	1595	13	1	0
42	2	2	6588	1410	0	0	0
42	5	0	6570	1395	35	0	0
35	2	4	6872	1124	0	0	0
35	5	4	6843	1102	48	3	0

Table 2 – *Selected orders for a region of 8000 voxel (taken as the region from Figure 5 plus two more slices in each direction) for different numbers of maximum components in the tensor mixture model and different subsamples of gradient directions. It shows the stability of estimates also for lower numbers of gradients and its independence from the chosen maximum number of tensor mixture components.*

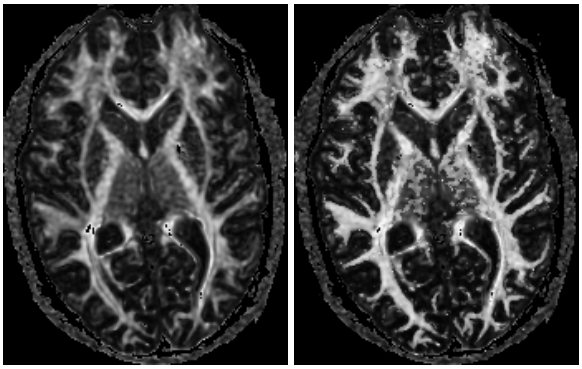

Figure 10 – *Left: FA for the diffusion tensor model. Right: FA estimated in a tensor mixture model with maximum order 5.*

anisotropy properties of the distinct fiber bundles than the single diffusion tensor model.

In Figure 11 we show the estimated eigenvalues for a slice in the experimental data compared with the largest and smallest eigenvalue in the single diffusion tensor model. The more homogeneous and sharper white matter indicates a more adequate modeling of partial volume effects by the tensor mixture model. The effect is especially visible at structural borders.

6 Conclusion

In this paper we considered and analyzed a *tensor mixture model* for modeling diffusion weighted data for brain and spinal chord imaging. In contrast to common view the order of the model (or number of mixing components) need not be predetermined but can be selected from a series of nested models of decreasing order by using a *model selection rule* based on the *Bayes information criterion (BIC)*.

We showed that in this model the weighted orientation distribution function is a mixture of

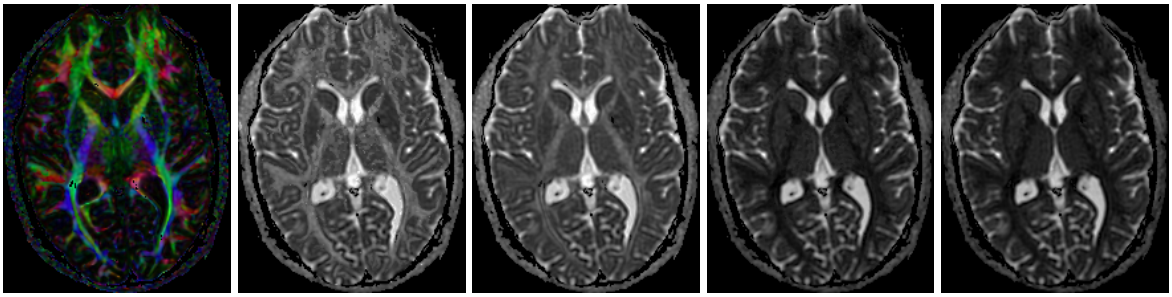

Figure 11 – *From left to right: color-coded FA of the diffusion tensor model, first eigenvalue estimated by maximum order 5 tensor mixture model, first eigenvalue estimated using the diffusion tensor model, second eigenvalue estimated by maximum order 5 tensor mixture model, smallest eigenvalue estimated within the diffusion tensor model.*

angular central Gaussian distributions, which are probability functions on the sphere. They can, in contrast to other functional expansions, be easily interpreted and show no negativity artefact's as occur using e.g. spherical harmonics expansion.

Based on the mixture model, we defined a number of natural contrasts, namely the *effective order of the model* (EO), the *fractional anisotropy* (FA), and the *eigenvalues* (EV). These contrasts might have the potential to become more sensitive indices in clinical contexts or basic research than the conventional anisotropy indices.

We showed in a series of examples based on extensive simulated and experimental data that the model leads to stable estimates. The fact that the mixture models are determined *independently in each voxel* but nevertheless show spatially extended regions with similar model parameters (tensor directions and eigenvalues) is a strong hint that the mixture components indeed correspond to fiber bundles. Our results suggest that specifying a larger maximum number of mixture components provides improved results, mainly due to the increased number of initial directions, while model selection using BIC effectively avoids overfitting. The estimates can surely be further improved by incorporating knowledge from adjacent voxel as in [Tabelow et al., 2008] for the single diffusion tensor model but this is beyond the scope of this paper and will be studied in future work.

The stability of the estimates for different numbers of gradients, SNR, and maximum order of the model gives rise to the speculation that the model may be feasible to detect complex fiber architecture also in clinical DWI settings with lower number of diffusion gradients then in typical HARDI measurements.

The tensor mixture model naturally extends to other sampling schemes in q-space, i.e. does not depend on measurements on a single q-shell. It can be directly applied to multiple q-shell data, radial or Cartesian sampling schemes.

There is of course a price to be paid, as the formulation as a non-linear optimization problem leads to high computational costs. However, using an optimized implementation with efficient memory usage and the fact that the problem can be canonically parallelized we expect the calculation to become feasible for common desktop hardware soon.

Acknowledgments

This work is supported by the DFG Research Center MATHEON. HUV is supported by NSF award 0956306 and Nancy M and Samuel C Fleming Research Scholar Award in Intercampus Collaborations, Cornell University.

References

I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel. Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. *Magn Reson Med*, 64:554–556, 2010.

H. Akaike. A new look at the statistical identification model. *IEEE Trans Automat Contr*, 19: 716–723, 1974.

D.C. Alexander. Multiple-fiber reconstruction algorithms for diffusion MRI. *Ann N Y Acad Sci*, 1064:113–133, 2005.

D.C. Alexander. An introduction to computational diffusion MRI: the diffusion tensor and beyond. In J. Weickert and H. Hagen, editors, *Visualization and Image Processing of Tensor Fields*. Springer, 2006.

A. Anderson. Measurements of fiber orientation distributions using high angular resolution diffusion imaging. *Magn Reson Med*, 54:1194–1206, 2005.

A. Barnett. Theory of Q-ball imaging redux: Implications for fibre tracking. *Magn Reson Med*, 62:910–923, 2009.

P.J. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophys J*, 66:259–267, 1994a.

P.J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion *tensor* from the NMR spin echo. *J Magn Reson*, 103:247–254, 1994b.

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

D. Le Bihan and E. Breton. Imagerie de diffusion in vivo par résonance magnétique nucléaire (invivo magnetic resonance imaging of diffusion). *C R Acad Sci*, 301:1109–1112, 1985.

D. Le Bihan, J.F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: Concepts and applications. *J Magn Reson*, 13:534–546, 2001.

K.P. Burnham and D. Anderson. *Model Selection and Multi-Model Inference*. Springer, 2002.

G. Claeskens and N.L. Hjort. *Model Selection and Model Averaging*. Cambridge University Press, 2008.

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.

M. Descoteaux, R. Deriche, D. Le Bihan, J.-F. Mangin, and C. Poupon. Multiple q-shell diffusion propagator imaging. *Med Image Anal*, 2010. doi: 10.1016/j.media.2010.07.001.

J.P. Dyke, H.U. Voss, D. Sondhi, N.R. Hackett, S. Worgall, L.A. Heier, B.E. Kosofsky, A.M. Ulŭg, D.C. Shungu, X. Mao, R.G. Crystal, and D. Ballon. Assessing disease severity in late infantile neuronal ceroid lipofuscinosis using quantitative MR diffusion-weighted imaging. *Am J Neuroradiol*, 28(7):1232–1236, 2007.

A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann Phys*, 4:549–560, 1905.

L.R. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magn Reson Med*, 47:1083–1099, 2002.

K.H. Fritzsche, F.B. Laun, H.-P. Meinzer, and B. Stieltjes. Opportunities and pitfalls in the quantification of fiber integrity: What can we gain from Q-ball imaging? *Neuroimage*, 51(1): 242–251, 2010.

C. Hess, P. Mukherjee, E. Han, D. Xu, and D. Vigneron. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magn Reson Med*, 56:104–117, 2006.

T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn Reson Med*, 54:1480–1489, 2005.

C.M. Hurvich and C.-L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76:297–307, 1989.

B. Jian, B.C. Vemuri, E.Özarslan, P.R. Carney, and T.H. Mareci. A novel tensor distribution model for the diffusion-weighted mr signal. *Neuroimage*, 37:164–176, 2007.

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

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.

E. Kaden, A. Anwander, and T.R. Knoesche. Variational inference of the fiber orientation density using diffusion MR imaging. *Neuroimage*, 42(4):1366–1380, 2008.

C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*. Classics in Applied Mathematics. SIAM, Philadelphia, 1995.

A. Leow, S. Zhu, K. McMahon, G. de Zubicaray, M. Meredith, M. Wright, and P. Thompson. The tensor distribution function. *Magn Reson Med*, 61:205–214, 2009. University of California, Los Angeles, CA, USA.

H. Linhart and W. Zucchini. *Model selection*. Wiley, 1986.

C. Liu, R. Bammer, and M.E. Moseley. Generalized diffusion tensor imaging (GDTI): a method for characterizing and imaging diffusion anisotropy caused by non-gaussian diffusion. *Isr J Chem*, 43:145–154, 2003.

K.V. Mardia and P.E. Jupp. *Directional Statistics*. Wiley, 2000.

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

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

K.M. Mullen. *Separable nonlinear models: theory, implementation and applications in physics and chemistry*. PhD thesis, Vrije Universiteit Amsterdam, 2008.

M. Nezamzadeh, V.J. Wedeen, R. Wang, Y. Zhang, W. Zhan, K. Young, D.J. Meyerhoff, M.W. Weiner, and N. Schuff. In-vivo investigation of the human cingulum bundle using the optimization of MR diffusion spectrum imaging. *Eur. J. Radiol.*, 75(1):E29–E36, 2010.

C. Nimsky, O. Ganslandt, and R. Fahlbusch. Implementation of fiber tract navigation. *Neurosurgery*, 61(1, Suppl. S):306–317, 2007.

E. Özarslan and T.H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magn Reson Med*, 50:955–965, 2003.

E. Özarslan, T.M. Shepherd, B.C. Vemuri, S.J. Blackband, and T.H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *Neuroimage*, 31:1086–1103, 2006.

M. Perrin, C. Poupon, B. Rieul, P. Leroux, A. Constantinesco, J.-F. Mangin, and D. Le Bihan. Validation of q-ball imaging with a diffusion fibre-crossing phantom on a clinical scanner. *Philos Trans R Soc Lond B Biol Sci*, 360(1457):881–891, 2005.

G. Schwarz. Estimating the dimension of a model. *Ann Statist*, 6:461–64, 1978.

D. Sondhi, N.R. Hackett, R.L. Apblett, S.M. Kaminsky, R.G. Pergolizzi, and R.G. Crystal. Feasibility of gene therapy for late neuronal ceroid lipofuscinosis. *Arch Neurol*, 58(11):1793–1798, 2001.

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*, 2010. URL `http://CRAN.R-project.org/package=dti`. R package version 0.9-3.

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

D.G. Taylor and M.C. Bushell. The spatial mapping of translational diffusion coefficients by the NMR imaging technique. *Phys Med Biol*, 30:345–349, 1985.

J.-D. Tournier, F. Calamante, D.G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *Neuroimage*, 23:1176–1185, 2004.

D.S. Tuch. *Diffusion MRI of Complex Tissue Structure*. PhD thesis, Massachusetts Institute of Technology, 2002.

D.S. Tuch. Q-ball imaging. *Magn Reson Med*, 52:1358–1372, 2004.

D.S. Tuch, T.G. Reese, M.R. Wiegell, N. Makris, J.W. Belliveau, and V.J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn Reson Med*, 48:577–582, 2002.

V.J. Wedeen, P. Hagmann, W.Y. Tseng, T.G. Reese, and R.M. Weisskopf. Mapping complex tissue architecture with diffusion spectrum imaging. *Magn Reson Med*, 54:1377–1386, 2005.

Y.-C. Wu and A.L. Alexander. Hybrid diffusion imaging. *Neuroimage*, 36:617–629, 2007.

C.-H. Yeh, J.-D. Tournier, K.-H. Cho, C.-P. Lin, F. Calamante, and A. Connelly. The effect of finite diffusion gradient pulse duration on fibre orientation estimation in diffusion MRI. *Neuroimage*, 51(2):743–751, 2010.