

RESEARCH ARTICLE

Reliable Dual Tensor Model Estimation in Single and Crossing Fibers Based on Jeffreys Prior

Jianfei Yang^{1,2}, Dirk H. J. Poot^{1,3}, Matthan W. A. Caan², Tanja Su², Charles B. L. M. Majoie², Lucas J. van Vliet¹, Frans M. Vos^{1,2*}

1 Quantitative Imaging Group, Department of Imaging Physics, Delft University of Technology, 2628CJ Delft, The Netherlands, **2** Department of Radiology, Academic Medical Center Amsterdam, 1105AZ Amsterdam, The Netherlands, **3** BIGR, Department of Medical Informatics and Radiology, Erasmus MC, Rotterdam, The Netherlands

* f.m.vos@tudelft.nl



 OPEN ACCESS

Citation: Yang J, Poot DHJ, Caan MWA, Su T, Majoie CBLM, van Vliet LJ, et al. (2016) Reliable Dual Tensor Model Estimation in Single and Crossing Fibers Based on Jeffreys Prior. PLoS ONE 11(10): e0164336. doi:10.1371/journal.pone.0164336

Editor: Christophe Lenglet, University of Minnesota, UNITED STATES

Received: August 28, 2015

Accepted: September 25, 2016

Published: October 19, 2016

Copyright: © 2016 Yang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The MRI data underlying datasets B-C were obtained from subjects who did not consent to their data being publically available. However, excerpts of the data are available on request from the authors (Frans M. Vos email: f.m.vos@tudelft.nl).

Funding: This work was supported from Fonds Nuts Ohra grant number 1002-028. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Purpose

This paper presents and studies a framework for reliable modeling of diffusion MRI using a data-acquisition adaptive prior.

Methods

Automated relevance determination estimates the mean of the posterior distribution of a rank-2 dual tensor model exploiting Jeffreys prior (JARD). This data-acquisition prior is based on the Fisher information matrix and enables the assessment whether two tensors are mandatory to describe the data. The method is compared to Maximum Likelihood Estimation (MLE) of the dual tensor model and to FSL's ball-and-stick approach.

Results

Monte Carlo experiments demonstrated that JARD's volume fractions correlated well with the ground truth for single and crossing fiber configurations. In single fiber configurations JARD automatically reduced the volume fraction of one compartment to (almost) zero. The variance in fractional anisotropy (FA) of the main tensor component was thereby reduced compared to MLE. JARD and MLE gave a comparable outcome in data simulating crossing fibers. On brain data, JARD yielded a smaller spread in FA along the corpus callosum compared to MLE. Tract-based spatial statistics demonstrated a higher sensitivity in detecting age-related white matter atrophy using JARD compared to both MLE and the ball-and-stick approach.

Conclusions

The proposed framework offers accurate and precise estimation of diffusion properties in single and dual fiber regions.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Diffusion-weighted magnetic resonance imaging (DW-MRI) can provide unique information about the integrity of white matter structures in the brain. Conventionally, the diffusion is described by a single rank-2 diffusion tensor [1], which is estimated from diffusion-weighted images (DWIs). There is an ongoing debate on how to effectively characterize the diffusivities in voxels containing complex anatomical structures such as crossing fibers. In these voxels the diffusion profile is not adequately described by a single rank-2 tensor [2,3].

Several sophisticated models for the diffusion in white matter have shown the potential to estimate more plausible anatomical properties of complex tissue structures, for instance the 'ball-and-stick' model [4], the composite hindered and restricted model of diffusion (CHARMED)[5], and the neurite orientation dispersion and density imaging approach (NODDI) [6]. Although many more techniques were proposed, the aforementioned techniques relate to our approach. A comprehensive overview is presented in [7].

The 'ball-and-stick' approach models the diffusion by one isotropic compartment and a set of linear, 1D diffusion profiles. The method is widely applied and works well for reconstructing the orientations of fiber bundles [8], even though these fiber bundles are not represented by full, 3D diffusion profiles.

CHARMED explicitly models the apparent slow diffusing component arising from *restricted*, intra-axonal diffusion (a non-Gaussian process). It yielded unbiased estimation of the orientations of two or more fiber compartments with low angular uncertainty. However, the application of CHARMED is challenging due to measurements at very high b-values causing both signal-to-noise ratio and scanning time limitations.

A clinically feasible technique for in vivo neurite orientation dispersion and density imaging (NODDI) [6] was proposed. NODDI adopts a tissue model that distinguishes three types of microstructural environment: intra-cellular, extra-cellular, and cerebrospinal fluid (CSF) compartments. The signal of intra-cellular diffusion is described by zero radius cylinders [9] (like in the ball-and-stick approach); the extra-cellular part is modeled by anisotropic, Gaussian diffusion and the CSF compartment is modeled as isotropic Gaussian diffusion. Experiments showed that quantities derived from NODDI such as the neurite orientation dispersion provided more specific markers of brain tissue microstructure than standard indices from classical, single-tensor DTI [6].

Other methods [10] [11] [12] were proposed that also aim to recover specific tissue parameters from the diffusion signal, such as cell size and cell density. However, these models are not directly compartment-specific. Instead, a multiple-tensor model [13] describes the complex, full diffusion shape by a weighted sum of 3D tensors. We mean with full diffusion shape the 3D probability density function representing the position of water molecules after a certain diffusion time. We assume that an anisotropic Gaussian distribution (mathematically a rank-2 tensor model) accurately approximates this distribution in single fiber voxels. The ball-and-stick model represents the diffusion as a line-like profile, thereby ignoring a 'lateral' diffusion component. This is an important difference because pathology leads to an increase of the lateral diffusivity, which can be measured with our representation. The model is an intuitive, physical representation and a natural extension of the classical single tensor. Also, it does not necessarily require extremely high b-values. Previously, we introduced an optimization framework that rendered a constrained dual tensor model (DTM) as well as a set of diffusion weighting parameters, such that both the diffusion *shape* and the diffusion *orientations* of crossing fibers could be accurately estimated[3].

Unfortunately, even a constrained DTM is prone to overfitting in areas containing a single fiber bundle, causing biased volume fractions and imprecise diffusivity estimates. As such, a

new challenge arises: how to automatically adapt the model complexity to warrant an accurate characterization of the underlying diffusion processes?

Many model selection methods were introduced in the DWI field, for instance based on constrained spherical deconvolution (CSD) [14], the Bayesian information criterion (BIC) and the generalization-error [15]. A limitation of the CSD approach is that it requires tuning of a threshold to reject small contributions. Furthermore, BIC is influenced by the estimated likelihood and non-estimated factors such as the number of parameters and the sample size. The generalization-error method is a non-local model selection technique. Importantly, all these approaches involve model selection techniques that make hard decisions to select an appropriate model.

Automatic relevance determination (ARD) aims to eliminate the redundant parameters in a complex model, such that the simplified model yields a better description of the data [16]. Behrens [8] adopted ARD for assessing the appropriate number of fiber orientations in each voxel for fiber tracking. This method ensures that if there is no evidence for a second fiber orientation in the data, the volume fraction attributed to this fiber will automatically be forced to zero. ARD methods assume a prior distribution for the model parameters. A Gaussian distribution is a straightforward choice for a prior [16]. Such a prior may involve hyper-parameters to tune its shape. Previous ARD approaches [8,17,18] involved marginalization (integration) over the hyper-parameters to get a prior for each parameter separately. Such a prior is likely to be sub-optimal for different diffusion geometries since potential correlations between the parameters of complete models are ignored.

We present a new framework for data-acquisition adaptive estimation of the diffusion shape of simple *and* complex white matter structures. We consider the method data-acquisition adaptive as it takes properties of the data-acquisition into account such as the number of gradient directions, the number of b-values used and the noise level. The method is based on ARD for a rank-2 dual tensor model and assesses whether two anisotropic tensors are ‘mandatory’ to model the acquired diffusion-weighted signals. Our ARD estimates the mean of the posteriori distribution, i.e. the model parameters given the data, exploiting Jeffreys prior [19] [20]: ‘JARD’. This data-acquisition adaptive prior is based on the Fisher’s information matrix [21]. Previous work on ARD for diffusion-weighted MRI primarily focused on the accurate reconstruction of fiber *orientations* based on the ball-and-stick model [8,17]. This rank-1 tensor model is not appropriate for estimating the diffusion shape as reflected by a rank-2 tensor model. The proposed JARD method is particularly suited for application in comparative studies in which the goal is to assess subtle differences in diffusion *shape* between patients and matched controls.

Methods

The JARD framework for estimation of the diffusion shape processes every voxel in the same way. It estimates the parameters of a constrained dual tensor model (DTM) by computing the mean of the posterior distribution sampled by a Markov Chain Monte Carlo (MCMC) approach. The algorithm is initialized by applying the constrained DTM to the measured diffusion-weighted signals using maximum likelihood estimation (MLE). The prior on the parameters in the MCMC sampling is based on the non-informative Jeffreys prior. This prior forces parameters, particularly the volume fraction, towards zero when there is little to no information in them. This will be verified experimentally in the Experiments and Results section.

Dual Tensor Diffusion Model

We assume that the diffusion in every fiber bundle is mono-exponential and Gaussian. The diffusion-weighted signal in all voxels is initially modeled by a so-called dual tensor model

(DTM) [3]. This model contains signal contributions of up to two fiber bundles and an isotropic component and is given by

$$S_{\theta,j} = S_0 \left(\sum_{i \in \{1,2,iso\}} f_i \exp(-b_j \mathbf{g}_j^T \mathbf{D}_i \mathbf{g}_j) \right), \tag{1}$$

where $S_{\theta,j}$ is the diffusion-weighted signal given parameter vector θ for diffusion weighting b_j in gradient direction \mathbf{g}_j and S_0 the signal without diffusion weighting. \mathbf{D}_1 and \mathbf{D}_2 are rank-2 tensors to model the anisotropic diffusion in each fiber, \mathbf{D}_{iso} is the amount of isotropic diffusion (i.e., $D_{iso} \cdot \mathbf{I}_{3 \times 3}$, D_{iso} representing the scalar amount of isotropic diffusion), and f_i represents the volume fraction of component \mathbf{D}_i . Note that the DTM in Eq (1) reduces to a single tensor model (STM)—reflecting a single fiber—if $f_1 > 0$ and $f_2 = 0$ or vice versa. The volume fraction parameters play an essential role in our JARD scheme.

Maximum Likelihood Estimation of a constrained DTM

The measured diffusion weighted image (DWI) $\tilde{S}_{j,\sigma}$ with diffusion weighting b_j in direction \mathbf{g}_j is corrupted by Rician noise of standard deviation σ [22]. Therefore, the probability density function (PDF) for $\tilde{S}_{j,\sigma}$ is given by

$$p(\tilde{S}_{j,\sigma} | \theta) = \frac{\tilde{S}_{j,\sigma}}{\sigma^2} \exp\left(-\frac{\tilde{S}_{j,\sigma}^2 + S_{\theta,j}^2}{2\sigma^2}\right) I_0\left(\frac{\tilde{S}_{j,\sigma} S_{\theta,j}}{\sigma^2}\right), \tag{2}$$

with $I_0(\cdot)$ the zero-th order modified Bessel function of the first kind. The DWIs are statistically independent, so that the joint probability density function $p(\tilde{\mathbf{S}}_\sigma | \theta)$ of the signal profile $\tilde{\mathbf{S}}_\sigma$ is given by the product of the marginal distributions for the measured signals $\tilde{S}_{j,\sigma}$ in each of the N_g diffusion weighted directions \mathbf{g}_j :

$$p(\tilde{\mathbf{S}}_\sigma | \theta) = \prod_{j=1}^{N_g} p(\tilde{S}_{j,\sigma} | \theta) \tag{3}$$

Here, $p(\tilde{\mathbf{S}}_\sigma | \theta)$ is the likelihood function of θ given $\tilde{\mathbf{S}}_\sigma$. The underlying parameter values can be inferred by maximizing this likelihood function [23]

$$\hat{\theta}_{MLE} = \arg \max_{\theta} \{p(\tilde{\mathbf{S}}_\sigma | \theta)\}. \tag{4}$$

Maximum likelihood estimation (MLE) has a number of favorable statistical properties in the estimation of diffusion properties in crossing bundles [3]. First, under very general conditions, MLE asymptotically reaches the Cramér-Rao lower bound (CRLB). The CRLB is a theoretical lower bound on the variance of any unbiased estimator. Second, the MLE is consistent, which means that it asymptotically ($N_g \rightarrow \infty$) converges to the true value of the parameter in a statistically well-defined way [24].

The dual tensor model given in (1) should be parameterized such that its parameter values reside in a well-defined range. In previous work [3], we parameterized the tensor \mathbf{D}_i as follows: $\mathbf{D}_i = \mathbf{R}_i^T \mathbf{E}_i \mathbf{R}_i$, where $\mathbf{E}_i = \text{diag}(\lambda_{i,1}, \lambda_{i,2}, \lambda_{i,3})$ is a diagonal matrix with the eigenvalues of the tensor \mathbf{D}_i on its diagonal. The non-negativity constraint that is imposed on the estimated diffusivity values is accomplished by employing an exponential mapping [3]. The matrices $\mathbf{R}_i = \mathbf{R}_{i,1} \mathbf{R}_{i,2} \mathbf{R}_{i,3}$ describe rotations around the x -, y - and z -axes: $\mathbf{R}_x(\alpha_1) = \mathbf{R}_x(\alpha_1) \mathbf{R}_y(\alpha_2) \mathbf{R}_z(\alpha_3 \pm \alpha_4/2)$. The first two rotations $\mathbf{R}_x(\alpha_1) \mathbf{R}_y(\alpha_2)$ determine the orientation of the plane in which the first principal eigenvectors of both tensor reside. $\mathbf{R}_z(\alpha_3 + \alpha_4/2)$ and $\mathbf{R}_z(\alpha_3 - \alpha_4/2)$ denote the in-plane rotations of the first principal eigenvector of the two tensors. As such, the parameter vector to be

estimated for a dual-tensor model MLE becomes

$$\boldsymbol{\theta} = \left\{ \underbrace{f_1, f_2, f_{iso}}_{\text{volume fractions}}, \underbrace{\lambda_{1, //}, \lambda_{1, \perp 1}, \lambda_{1, \perp 2}}_{\text{diag}(\mathbf{E}_1)}, \underbrace{\lambda_{2, //}, \lambda_{2, \perp 1}, \lambda_{2, \perp 2}}_{\text{diag}(\mathbf{E}_2)}, \underbrace{D_{iso}, \alpha_1, \alpha_2, \alpha_3, \alpha_4}_{\text{fiber orientations}} \right\}. \quad (5)$$

However, MLE does not necessarily yield useful estimates. A potential error in the estimated parameters is greatly influenced by the degrees of freedom (DOFs) and the covariance (s) between parameters. We demonstrated that restricting the DOFs by imposing constraints on the DTM greatly reduces the covariance between parameters. The experiments in [3] showed that precise and accurate estimation can be achieved if we apply the following constraints:

$$\lambda_{1, //} = \lambda_{2, //}, \quad (6)$$

$$\begin{aligned} \lambda_{1, \perp 1} &= \lambda_{1, \perp 2} \equiv \lambda_{1, \perp} \\ \lambda_{2, \perp 1} &= \lambda_{2, \perp 2} \equiv \lambda_{2, \perp} \end{aligned} \quad (7)$$

$$D_{iso} = C_{free-water} \quad (8)$$

$$f_1 + f_2 + f_{iso} = 1. \quad (9)$$

Eq (6) imposes that the “unrestricted” diffusivity (i.e., free diffusivity) along the fibers, denoted by the first eigenvalues of \mathbf{D}_1 and \mathbf{D}_2 , are equal. Eq (7) states that the diffusion perpendicular to the fiber orientation is assumed to be axially symmetric, which models the average shape of axons. Eq (8) defines that D_{iso} equals $C_{free-water} = 3 \times 10^{-3} \text{ mm}^2 \text{ s}^{-1}$, the diffusivity of free water at body temperature 37°C, and Eq (9) states that the two anisotropic tensors plus the isotropic compartment fill the entire volume of each voxel.

Constraining the DTM cannot avoid the inherent risk of overfitting. This happens when a complex model is fitted to simple data, e.g. fitting multiple tensors to data of a single fiber bundle. Typically, this yields an increase of the variance and the covariance of the parameters, but also leads to biased diffusivity estimates.

Automatic Relevance Determination

Bayes factors offer an alternative to model selection by the classical likelihood test [25]. It computes the evidence for a model to be used in model selection. However, calculating the evidence for any model requires integration over all model parameters, weighted by the parameter priors. This is computationally unfeasible, especially with high-dimensional parameter spaces for which no analytical solution exists. ARD was introduced exactly to cope with such issues [16] [8] [17] [26] [27]. Compared to the Bayes factors approach, ARD does not fit competing potential models to the data and compares them on the basis of the residual after fitting. Instead, ARD always fits a complex model to the data and forces irrelevant parameters to zero, so that a complex model reduces to a simpler one.

Our JARD estimates the mean of the posterior distribution of the constrained DTM based on Bayes’ theorem [16]. The posterior distribution $p(\boldsymbol{\theta}|\tilde{\mathbf{S}}_\sigma)$ is

$$p(\boldsymbol{\theta}|\tilde{\mathbf{S}}_\sigma) = \frac{p(\tilde{\mathbf{S}}_\sigma|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\tilde{\mathbf{S}}_\sigma)}, \quad (10)$$

where $p(\tilde{\mathbf{S}}_\sigma|\boldsymbol{\theta})$ is the aforementioned likelihood function, $p(\boldsymbol{\theta})$ the prior probability of $\boldsymbol{\theta}$ in the

DTM, and $p(\tilde{\mathbf{S}}_\sigma)$ the evidence for the DTM. As the evidence term in Eq (10) is constant for any measured signal, the posterior probability distribution in JARD becomes

$$p(\boldsymbol{\theta}|\tilde{\mathbf{S}}_\sigma) \propto p(\tilde{\mathbf{S}}_\sigma|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{11}$$

Our framework estimates the posterior distribution given the data, which is influenced by the likelihood function and the prior. We introduce a data-acquisition adaptive prior for DTM parameters based on Jeffreys theorem (see next subsection). It allows simplifying a complex model to a simple model by automatically forcing volume fractions, which are not supported by data to zero.

JARD employs a Markov Chain Monte Carlo (MCMC) technique with Metropolis-Hasting sampling of the posterior distribution $p(\boldsymbol{\theta}|\tilde{\mathbf{S}}_\sigma)$ [28] [29]. The MCMC draws 5000 samples from the posterior distribution in the nine-dimensional parameter space. The algorithm is listed in Algorithm 1. It is initialized by MLE of the constrained DTM. The final JARD estimate $\hat{\boldsymbol{\theta}}_{JARD}$ is the mean of 3000 accepted samples after a burn-in period of 2000.

If the posterior estimates of *both* anisotropic fractions lie in a small interval around their MLE value, then this would indicate that the prior did not significantly change the outcome and that fitting the initial dual tensor model was justified. Reversely, if the posterior estimate for *one* of the two anisotropic fractions does not significantly differ from zero, then its corresponding tensor compartment can be treated an unnecessary parameter. In such a case, the estimation essentially returns a ‘single-tensor’ model.

Jeffreys Prior

Methods to choose the prior for a Bayesian analysis can be divided into two groups: informative and non-informative priors [27] [30]. An informative prior expresses specific, definite information about a variable, whereas an uninformative (or diffuse) prior expresses only general information about a variable. We aim to introduce a new, data-acquisition adaptive prior, which makes JARD non-informative. Specifically, we adopt Jeffreys non-informative prior $p_J(\boldsymbol{\theta})$ which can be written as:

$$p_J(\boldsymbol{\theta}) \propto \det(\mathbf{I}(\boldsymbol{\theta}))^{1/2}, \tag{12}$$

where $\mathbf{I}(\boldsymbol{\theta})$ denotes the Fisher information matrix given by

$$\mathbf{I}(\boldsymbol{\theta}) = -\mathbb{E}_s \left\{ \frac{\partial^2 \ln(p(\mathbf{S}|\boldsymbol{\theta}, \sigma))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right\}. \tag{13}$$

The Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ provides the amount of expected information about the parameter vector $\boldsymbol{\theta}$ in measurements. By definition, it is influenced by properties of the data-acquisition such as the number of data points and the noise. In general, Jeffreys prior is in agreement with one’s intuition that if a parameter is necessary, it must be supported by the data. Poot [31] showed that the Fisher information matrix for Rice distributed measurements given by Eq (13) can be efficiently computed.

We aim to exploit Jeffreys prior in Bayesian estimation to discriminate between configurations that yield a degenerate or a non-degenerate model. The preferred prior must convey support for a dual tensor representation in crossing fibers, e.g. represented by two tensors at 90-degree angles. Therefore, the prior should be flat in such a configuration, making that it only mildly affects the posterior distribution, which would peak near the initial dual tensor parameters obtained by MLE. Reversely, consider an actual single tensor configuration, e.g. a dual tensor model consisting of two tensors with the same shape and orientation and having

$f_{1,2} = 0.45$. In this case the dual-tensor model is degenerate and the prior must be harsh, promoting a near-zero volume fraction in the posterior distribution.

In order to meet these preferred properties, we use a prior based on Jeffreys prior:

$$p(\theta) = \det(\mathbf{I}(\theta))^{-1/2}, \tag{14}$$

Fig 1 illustrates the shape of the prior. Notice in Fig 1(A) that as the volume fraction $f_1 \approx 0.45$ the prior is rather flat. This is because in a true dual tensor configuration, there is no correlation between the model parameters, as such yielding a maximum for $\det(\mathbf{I}(\theta))$. Reversely, as the volume fraction f_1 approaches zero to establish a single tensor configuration, there will be large correlation between several parameters yielding a very small $\det(\mathbf{I}(\theta))$ and in turn a sharp rise in the prior probability. Additionally, Fig 1(B) shows that as orientation divergence α_4 decreases to zero, the prior increases in the direction of $f_1 = 0$. As such, the prior will favor a ‘simpler’ configuration (i.e. push f_2 to zero) as long as the decrease in goodness of fit is more than compensated for by an increase in the prior probability. Note that the bathtub shape of the prior probability in Fig 1(A) indicates that the roles of the two volume fractions can be exchanged. The deviation from a pure symmetric curve is due to the difference in FA of the two crossing fiber bundles.

A nice property of Jeffreys prior is that it automatically adapts to data acquisition conditions. This is exactly why we describe our prior as data-acquisition adaptive. For example, a decreased number of gradient directions or an increased noise level will reduce the Fisher information of a dual tensor model. This corresponds to larger theoretical minimum variance and a reduced support of that model. Therefore, the curve representing the lower SNR in Fig 1(C) (blue) starts to increase at a larger α_4 than that corresponding to the higher SNR (red). This shows that a higher SNR yields a better orientation sensitivity in detecting and estimating crossing fibers. In effect, our prior enforces that the dual-tensor model reduces to the single tensor representation unless there is sufficient support for two non-zero tensors in the data. Furthermore, a higher b-value by itself enhances the support for a two-tensor representation (see [13] [32] [33]). However, such a higher b-value usually comes with a lower signal to noise ratio. Therefore, the effect of varying the b-values is more difficult to predict. In previous work, we found that imaging at two b-values (1000, 3000) is appropriate for precise estimation of diffusion parameters in fiber crossings.

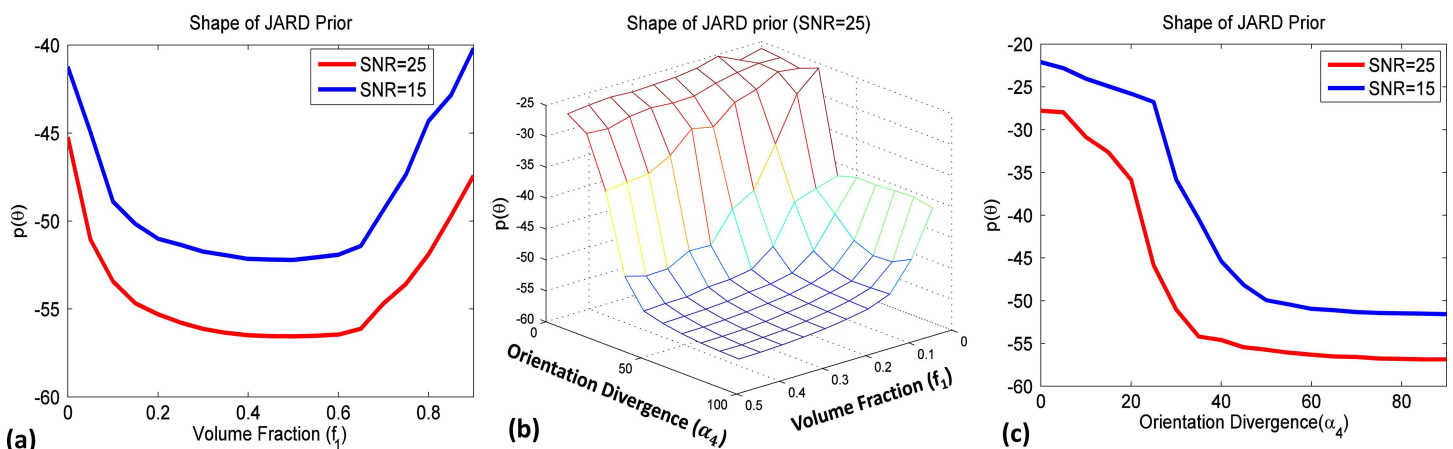


Fig 1. Plots of the prior probability for a model of two crossing fibers (detailed parameters, see Table 1). a) Prior probability for crossing fibers at 90-degree angle as a function of f_1 . b) Prior probability as a function of both f_1 and α_4 . c) Prior probability for crossing fibers of equal volume fraction as a function of α_4 . The blue and red curves in subplots (a) and (c) were generated using respectively SNR = 15 and SNR = 25.

doi:10.1371/journal.pone.0164336.g001

Algorithm 1. Markov Chain Monte Carlo with Metropolis-Hastings sampling
 Algorithm for estimating the mean of a multivariate posterior distribution with our prior using a Markov Chain Monte Carlo method employing Metropolis-Hastings sampling. The 'proposal' distribution $Q(\theta' | \theta_t) = \theta_t + \Delta \cdot \mathbf{N}(0, 1)$ with $\mathbf{N}(0, 1)$ denoting a multivariate Gaussian distribution with zero mean and standard variance 1; and Δ the step size for parameter vector θ . The step sizes for all parameters are: 10^{-5} (with unit $mm^2 s^{-1}$) for the diffusivity parameters, 10^{-2} (with unit rad) for the angles, and 10^{-2} for the volume fractions.

```

For all voxels
   $\theta_0 = \hat{\theta}_{MLE}$  // Initialize vector
  For t = 0 to N-1
     $\theta' = Q(\theta' | \theta_t)$  // Draw candidate from 'proposal' distribution
     $p(\theta') = \det(\mathbf{I}(\theta'))^{-1/2}$  // Calculate our prior
     $p(\theta' | \hat{\mathbf{S}}) = p(\hat{\mathbf{S}} | \theta') p(\theta')$  // Calculate posterior probability
     $\alpha = p(\theta' | \hat{\mathbf{S}}) / p(\theta_t | \hat{\mathbf{S}})$  // Calculate the acceptance ratio
    If  $\alpha > 1$  then
       $\theta_{t+1} = \theta'$  // Accept the candidate vector
    Else
       $r = U(0, 1)$  // Draw random variable r between 0 and 1;
      If  $r \leq \alpha$ 
         $\theta_{t+1} = \theta'$  // Accept the candidate vector
      Else
         $\theta_{t+1} = \theta_t$ ; // Keep the previous vector
      Endif
    Endif
  Endfor
   $\hat{\theta}_{ARD} = \frac{1}{N - N_{burn-in}} \sum_{t=N_{burn-in}}^{N-1} \theta_t$  // Compute mean of accepted samples after burn-in
Endfor
  
```

Results

All experiments below were carried out on a DELL laptop computer with an Intel i7-2620CPU @2.7GHz and 4 GB RAM running a Windows-7 64-bit operating system. The method was implemented in MATLAB_R2014b. The average execution time on our brain image data was 6 voxel/s. Applying our method to one DTI volume (from dataset A, see below) took approximately one day by parallel processing of voxels on a cluster of 12 Intel Pentium cores. Clearly, a further speedup can be achieved by involving more cores because the execution time decreases linearly with the number of cores. Furthermore, the presented number is measured based on interpreted MATLAB code. We expect that a further speedup can be achieved by compiling the MATLAB code.

In the first part of this section we evaluate the performance of estimating the parameters of our constrained dual tensor model by JARD and by MLE on simulated data. We studied the differences between JARD and MLE as a function of the volume fraction for simulated crossing fibers under realistic conditions. Diffusion measurements were simulated by means of the model presented in Eq (1). The parameters of crossing fibers are listed in Table 1 and are in agreement with the work of Pierpaoli [34] who reported diffusivities ranging from 0.25×10^{-3} to $1.5 \times 10^{-3} mm^{-2} s$. At the given diffusivities $FA_1 = 0.91$ and $FA_2 = 0.67$. The SNR (defined by S_0/σ) was 25 [3]. Two measurements at $b = 0 mm^{-2} s$ were simulated. Furthermore, 92 gradient directions were adopted for each of two b -values ($1.0 \cdot 10^3 mm^{-2} s$ and $3.0 \cdot 10^3 mm^{-2} s$), homogeneously distributed over the surface of a sphere. These settings are identical to dataset B (see below) and equal to those reported in [3].

In the second part of this section we demonstrate the potential of the proposed framework for some neuroimaging applications. Therefore, we evaluated the performance on varying

Table 1. Model parameters for generating synthetic data. The units of the diffusion parameters λ are $10^{-3} \text{ mm}^2 \text{ s}^{-1}$.

Parameters	Value (θ)
$\lambda_{1//}$	1.480
$\lambda_{1,\perp 1}$	0.15
$\lambda_{1,\perp 2}$	0.12
$\lambda_{2//}$	1.400
$\lambda_{2,\perp 1}$	0.4
$\lambda_{2,\perp 2}$	0.38
$\alpha_{1,2}$	Random
α_3	0.8π
α_4	Variable
f_1	Variable
f_{iso}	0.1
D_{iso}	3.0
S_0	250
σ	10

doi:10.1371/journal.pone.0164336.t001

types of brain datasets to verify whether a reliable estimation could be achieved. We applied JARD and MLE to the genu of the corpus callosum (CC) representing a single fiber region enclosed at both ends by a crossing with the corticospinal tract (CST). Subsequently, we show a neuroimaging application of our proposed framework, i.e. automatic estimation of diffusion properties.

Three different datasets, acquired with different acquisition protocols, were adopted to explore the two methods. Dataset A concerned diffusion data from one subject of the Human Connectome Project (HCP) [35]. The relevant acquisition parameters of this dataset were: three b -values 1000, 2000 and 3000 s/mm^2 , 90 gradient directions per b -shell and three measurements at $b = 0$ per b -shell, TE/TR 89.5/5520 ms , voxel size $1.25 \times 1.25 \times 1.25 \text{ mm}^3$. Dataset B was acquired from one control subject (see also [3]). The acquisition parameters: two b -values 1000 and 3000 s/mm^2 , 92 gradient directions per b -shell and one measurements at $b = 0$ per b -shell, TE/TR 84/3800 ms , voxel size $1.7 \times 1.7 \times 2.2 \text{ mm}^3$. Dataset C consisted of data from 24 healthy controls from an ongoing DTI study into the effects of HIV on the brain [36]. The acquisition parameters were: two b -values at 1000 and 2000 s/mm^2 , 64 gradient directions b -shell and one measurements at $b = 0$ per b -shell, voxel size $2.0 \times 2.0 \times 2.0 \text{ mm}^3$. The SNR for each of the three datasets was found to be higher or equal than 20. This was determined by fitting a single tensor model to a selected region of the CC, after which we took the ratio between $S_{v,0}$ and the model residual as the estimated SNR.

JARD versus MLE: simulation experiment

The volume fraction of the constituting compartments of a dual tensor model is a crucial parameter for the modeling of simple and complex fiber geometries. Therefore, we evaluated the performance of the methods as the volume fractions of the two compartments were varied. To assess the robustness for variations in volume fraction, for each volume fraction and divergences of 90 degrees and 45 degrees we generated 100 realizations with the parameters given in Table 1. For each realization $\hat{\theta}_{JARD}$ and $\hat{\theta}_{MLE}$ were computed.

Fig 2 shows boxplots depicting the results of dual tensor estimation by JARD (red) and MLE (blue). Since the estimation procedure assigns a random label to the first and second tensor, we need to assign the two estimated tensors to the corresponding ground truth

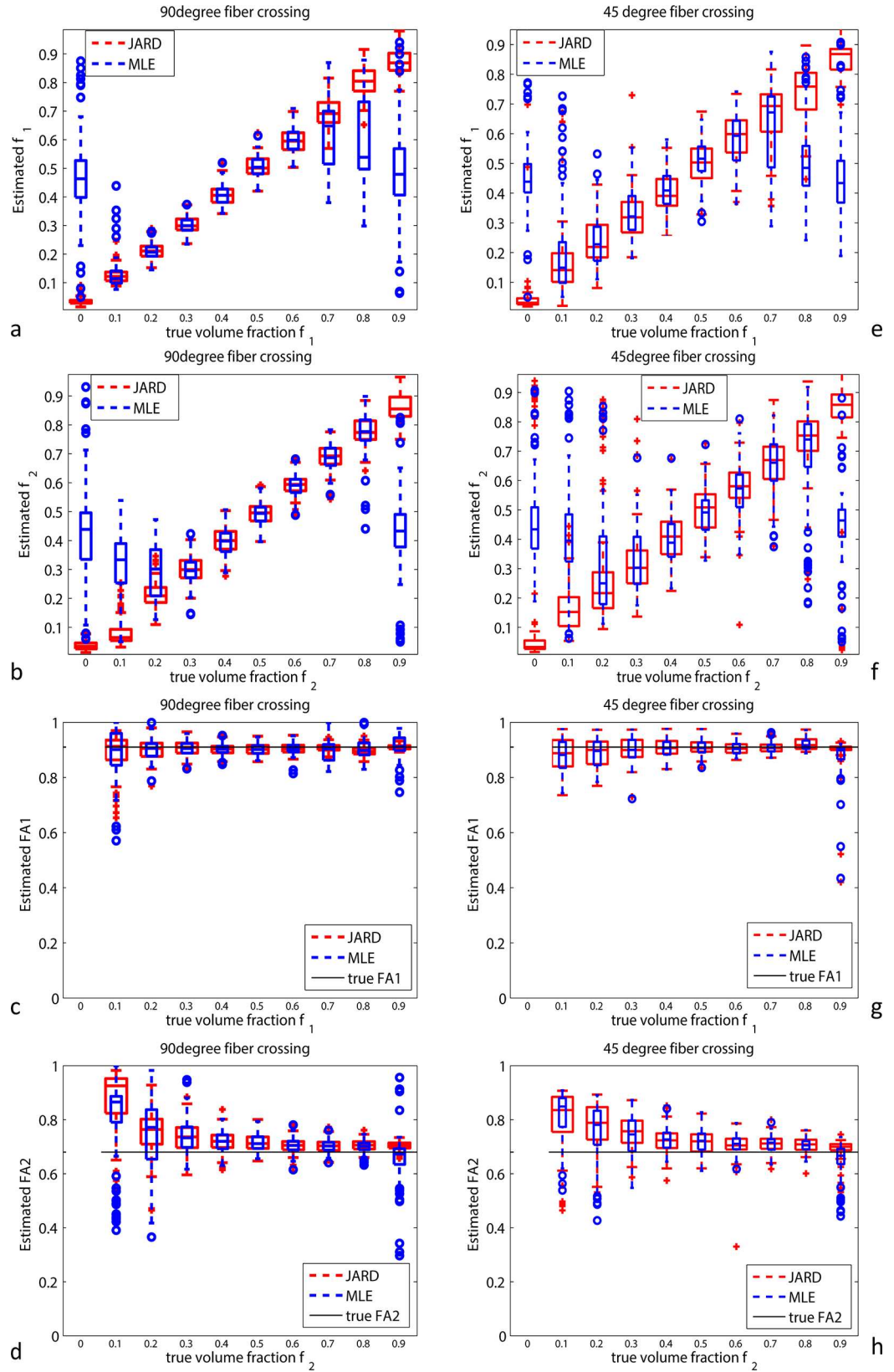


Fig 2. Results of dual tensor model estimation by JARD (red) and MLE (blue) on 100 noisy realizations (SNR 25) of two crossing fibers (divergence: 90 degrees in a-d and 45 degrees in e-h) as a function of volume fraction. The boxplots show: a) and e) volume fraction of the first compartment; b) and f) volume fraction of the second compartment; c) and g) fractional anisotropy (FA) of the first component; d) and h) FA of the second component. The simulated model listed in Table 1 and acquisition parameters accords with Dataset B. The boxes display the median and 25th, respectively 75th percentiles of the data distribution; whiskers extend to 1.5 times the interquartile range; values outside these ranges are indicated as individual points.

doi:10.1371/journal.pone.0164336.g002

compartment. The estimated tensors are sorted by tensor similarity based on the Frobenius norm: the tensor with the smallest Frobenius norm with respect to the ground truth of compartment 1 received the label 1. Fig 2A–2D) show the results for the 90° crossing. The estimated volume fraction by JARD nearly ideally correlates with the ground truth over the entire range of volume fractions, both for single fiber ($f_1 = 0 \vee f_1 = 0.9$) and crossing fiber configurations ($f_1 \in (0, 0.9)$), see Fig 2A and 2B). Clearly, MLE yields a random f_1 as the true volume fraction approaches that of a single fiber configuration.

Fig 2(C) shows that the median FA_1 estimation is almost identical for both estimation methods in crossing fiber configurations, irrespective of the volume fraction. Fig 2(D) shows that the estimated FA_2 converges (almost) to the true value as the true f_2 increases. The FA estimation in single fibers ($f_1 = 0.9 \vee f_2 = 0.9$) appears equally unbiased for both methods, but a *considerably larger spread* is encountered with MLE than with JARD. Note that the estimated FA 's with $f_1 = 0 \vee f_2 = 0$ essentially represent degenerate measurements and are therefore not shown in the graphs. In the absence of a ground truth this can be detected with JARD as the corresponding volume fraction is automatically forced to (near) zero (see Fig 2 (A)). MLE does not offer such a mechanism, which might lead to a fictitious fiber compartment.

Fig 2E–2H) show the results for the 45° crossing. This figure shows similar trends as obtained for the 90° crossing in Fig 2A–2D), albeit with a larger spread. Specifically, for the 45° crossing JARD yielded $f_1 = 0.05 \pm 0.06, f_2 = 0.87 \pm 0.06$ (mean \pm standard deviation) at a true f_1 of 0.0. Instead, MLE gave $f_1 = 0.47 \pm 0.14, f_2 = 0.44 \pm 0.15$ at true $f_1 = 0.0$. JARD yielded $FA_1 = 0.89 \pm 0.07$ at a true $f_1 = 0.1$ while $FA_2 = 0.68 \pm 0.06$ as the true $f_2 = 0.9$. Alternatively, MLE gave $FA_1 = 0.89 \pm 0.09$ at true $f_1 = 0.1$ while $FA_2 = 0.64 \pm 0.08$ as the true $f_2 = 0.9$. Notice that there is not a significant bias in the FA of a tensor if its volume fraction is larger than 0.2. In particular, the FA of a tensor is not biased in case the other component has zero volume fraction, i.e. the single fiber situation. Only in the case where there is a small second component, particularly with true $f_2 = 0.1 \dots 0.2$, the FA of this tensor is slightly biased.

The performance of the two methods in single fiber regions is further corroborated in Fig 3. It shows the dual tensor model estimation by JARD (red) and MLE (blue) on a single fiber with a small isotropic compartment while only the single fiber's FA is varied.

Fig 3(A) shows that the estimated f_1 with JARD improves with increasing FA and approximates the ground truth. MLE essentially yields a random estimate of f_1 irrespective of the actual FA. Fig 3(B) confirms that the estimation of FA_1 remains largely unbiased with both methods. Clearly, the spread in the estimated FA_1 with MLE is much larger than with JARD. Particularly, JARD yielded $f_1 = 0.79 \pm 0.08$ at a true FA = 0.3 and $f_1 = 0.87 \pm 0.04$ at a true FA = 0.95 (true $f_1 = 0.9$). Instead, MLE gave $f_1 = 0.56 \pm 0.17$ at true FA = 0.3 and $f_1 = 0.52 \pm 0.11$ at true FA = 0.95. Furthermore, JARD yielded $FA_1 = 0.35 \pm 0.04$, at a true FA = 0.3 and $FA_1 = 0.95 \pm 0.01$, at a true FA = 0.95. Similarly, MLE gave $FA_1 = 0.38 \pm 0.23$, at a true FA = 0.3 and $FA_1 = 0.95 \pm 0.02$, at a true FA = 0.95.

The effect of varying the fiber orientation was assessed by generating diffusion measurements through Eq (1) using the parameter values given in Table 1, with $\alpha_4 \in \{0, 20^\circ, 40^\circ, 90^\circ\}$. For each such configuration 100 noisy realizations were generated. Subsequently, the model

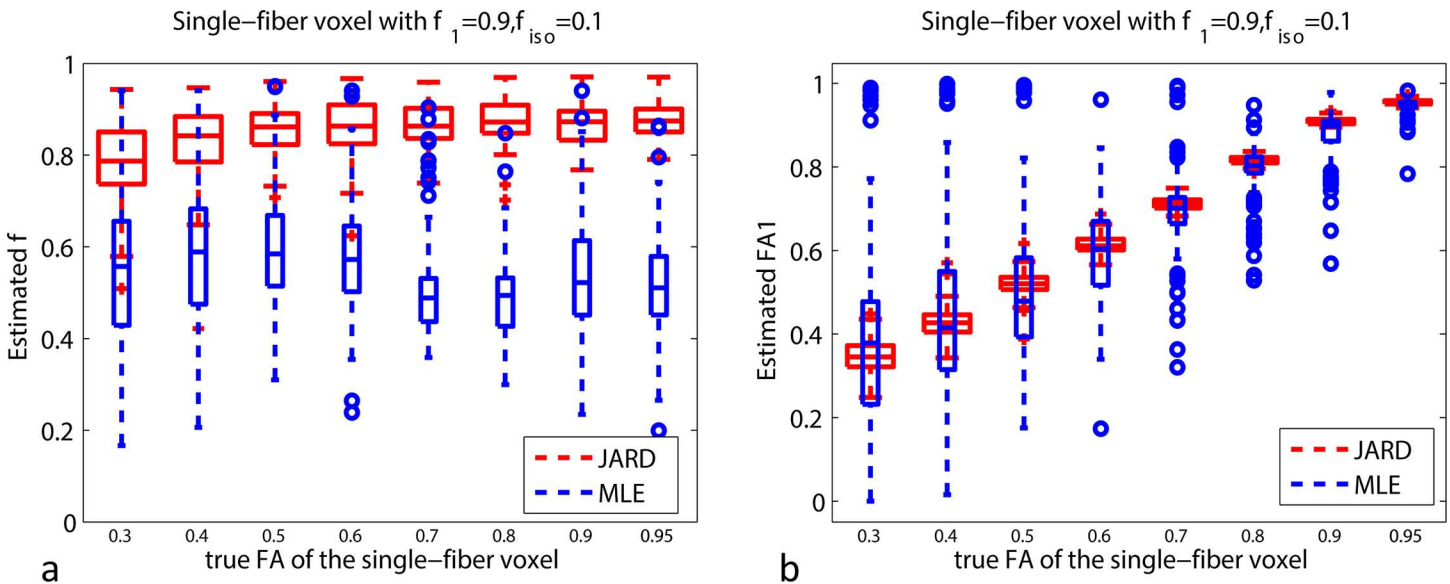


Fig 3. Results of dual tensor model estimation by JARD (red) and MLE (blue) on 100 noisy realizations (SNR 25) of a single fiber and an isotropic compartment as a function of FA. a) Volume fraction of the first compartment; b) Fractional Anisotropy (FA) of the first component. The simulated model are listed in Table 1 and acquisition parameters accords with Dataset B.

doi:10.1371/journal.pone.0164336.g003

parameters were estimated by means of MLE and ARD using both the prior from Behrens’ paper (i.e. $f^{-1}(1-f)^{-1}$) as well as Jeffreys prior, referred to as BARD and JARD respectively.

The outcome of this experiment is presented in Fig 4. It contains scatterplots of estimated orientations of the largest eigenvectors of the tensors.

The results confirm that the three methods yield very similar performance as $\alpha_4 \geq 20^\circ$. However, the proposed JARD clearly yields the most precise orientation estimation if $\alpha_4 = 0$: the tensors with a large volume fraction assert a single orientation, while the scattered orientations all have a very small volume fraction.

Summarizing, the graphs demonstrate two things: 1) JARD facilitates accurate estimation of volume fractions especially with increasingly unbalanced real volume contributions, in which case MLE grossly fails; 2) the estimation of FA by JARD shows a much narrower distribution than by MLE, especially in single fibers

JARD versus MLE: brain imaging

To demonstrate the performance of JARD and MLE on brain data, one subject was randomly selected from each of the three aforementioned datasets (A, B, and C). Fig 5 shows approximately corresponding coronal-views of regions of interest where the corpus callosum (CC) crosses the corticospinal tract (CST). Specifically, it contains the JARD (left) and MLE (middle) and FSLs ARD [8] (ARD-FSL, right) estimates in this slice. From top to bottom are shown data from respectively datasets A, B, and C. A single fiber region, i.e. the central part of CC, is indicated by the yellow ellipse and the region where CC and CST cross by the green circle. For JARD and MLE, the line segments visualize the orientations of the largest eigenvectors of the underlying tensors and the length of each line segment is scaled by the volume fraction. For FSL-ARD the line segments visualize stick orientations, scaled with the estimated volume fractions. Therefore, we applied the command “bedpostX” in FSL with parameter values: number of fibers = 2, weight = 1, Burn in = 1000, and switching to multi-shell model and rician noise (leaving model noise floor off) for a comparable outcome.

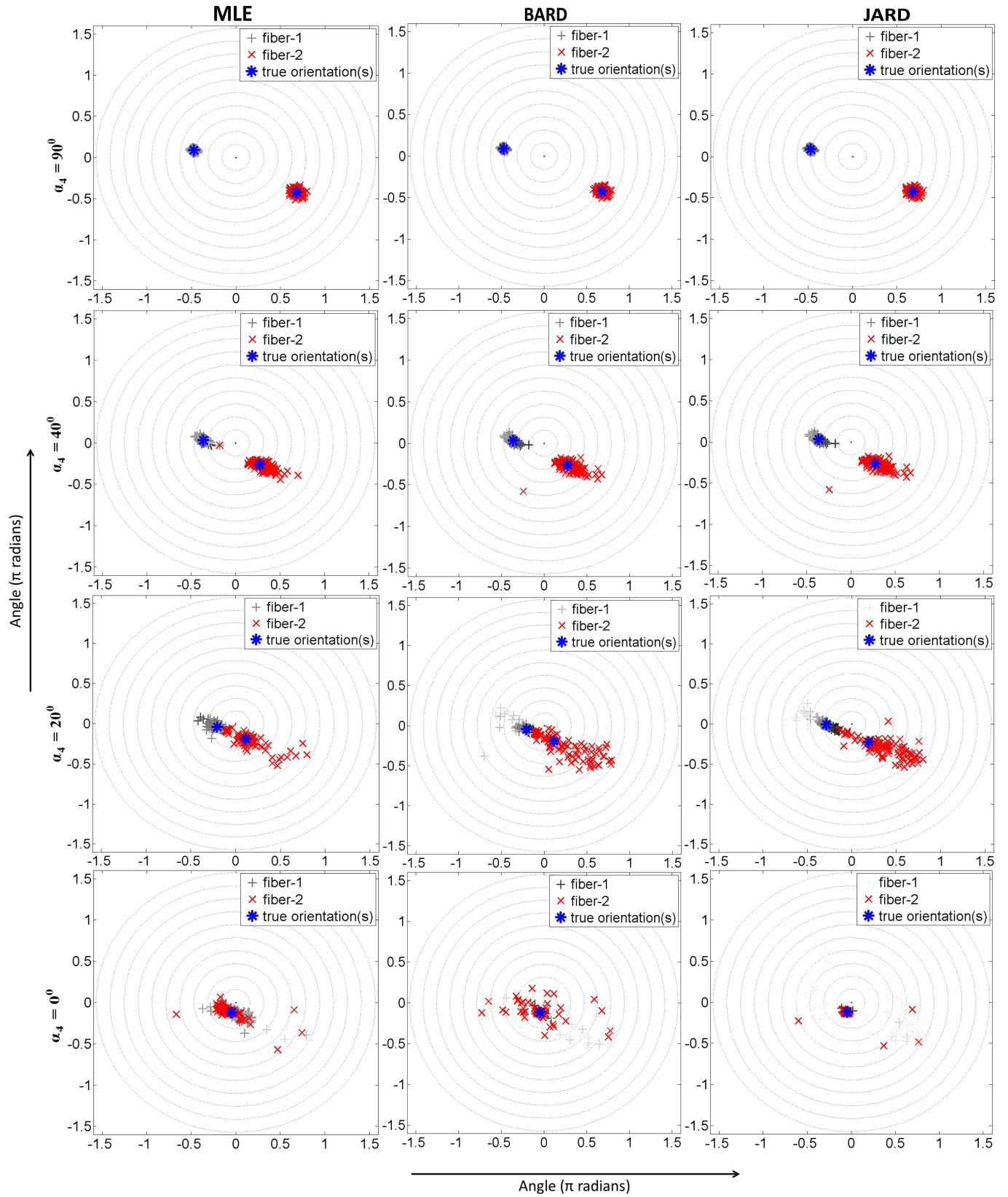


Fig 4. Scatterplots of estimated orientations by MLE and ARD using Behrens prior (BARD) and Jeffreys prior (JARD) for with varying α_4 . The two axes in each subplot represent angles (unit: π rad). The orientation of the tensor with the largest volume fraction is shown in red, the other in gray. The intensity of each symbol is scaled by the volume fraction of the tensor.

doi:10.1371/journal.pone.0164336.g004

Clearly, JARD forces the volume fraction of one tensor compartment nearly to zero in the single fiber region of the CC. Evidently, MLE yields an erratic outcome regarding both volume fraction and fiber orientation in the same region. Furthermore, notice that FSL-ARD often returns two fiber orientations in this region. In crossings like the region where CC crosses with CST, JARD, MLE and FSL-ARD yield a similar outcome regarding fiber orientations. These trends can be observed for each of the three datasets.

The performance of FSL-ARD is influenced by the so-called ARD weight: a higher *weight* reduces the number of secondary fibres per voxel.

Fig 6 shows the influence of the FSL-ARD weight on FSL's performance. In this experiment, we examined the same regions as shown in Fig 5 and used three different FSL-ARD weights.

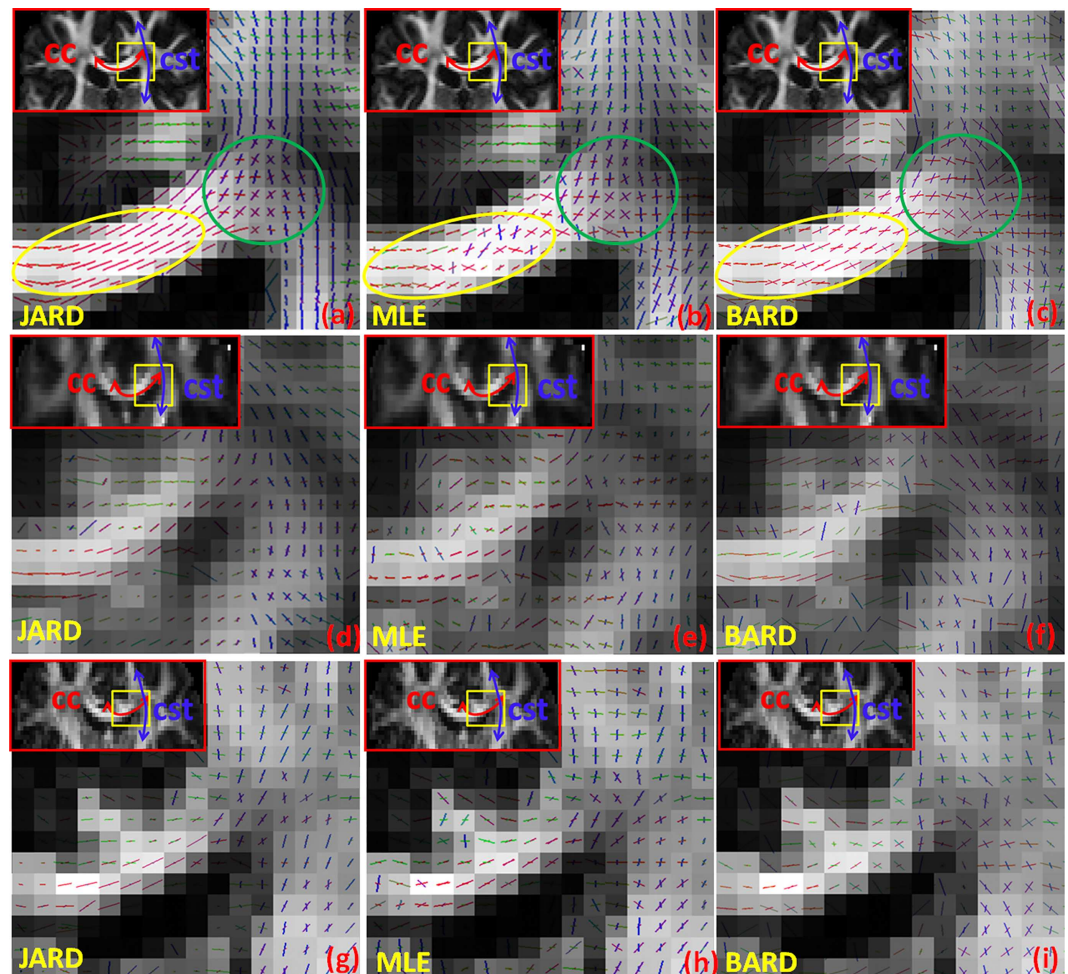


Fig 5. Results of our JARD (a, d, g), MLE (b, e, h) and FSL-ARD using the ball-and-stick model (c, f, i) in a region where the corpus callosum (CC) crosses the corticospinal tract (CST) in three randomly selected subjects from Dataset A: (a,b,c), Dataset B: (d, e, f), Dataset C (g, h, i). The length of a line segment is proportional to the corresponding volume fractions; the colors indicate the orientation of fibers: medio-lateral (red), anterior-posterior (green), and superior-inferior (blue).

doi:10.1371/journal.pone.0164336.g005

It can be noticed that a larger FSL-ARD weight reduces the number of modeled fibers, which is preferred in single fiber regions such as the central part of the corpus callosum. However, a larger weight simultaneously decreases the number of modeled fibers in crossing regions. Furthermore, the proper FSL-ARD weight varies for the different datasets: 10 for dataset A, 1 for dataset B and C. For instance, Fig 6(C) shows that weight = 10 yields good model selection in the single fiber region of Dataset A (see the red circle). However, at this same weight spurious second fibers are not effectively restrained in Dataset B (see the red circle in Fig 6(F)). On the other hand, at this weight no secondary fibers are detected in crossing regions (indicated by yellow circles in (f) and (i)). Observe that our framework (as reported in Fig 5) yields a better performance than FSL-ARD, even at the optimal weight settings, and does not require parameter tuning.

JARD versus MLE: Fractional Anisotropy along the CC

Fig 7 shows the estimated FA along the Genu of the corpus callosum (GCC) in one subject from dataset B. The GCC is indicated by the yellow trajectory superimposed on the red-colored structure of the inset. Notice that the center of the GCC is a single fiber bundle but there is a crossing with the CST at its lateral sides. The FA along the tract was estimated by the proposed

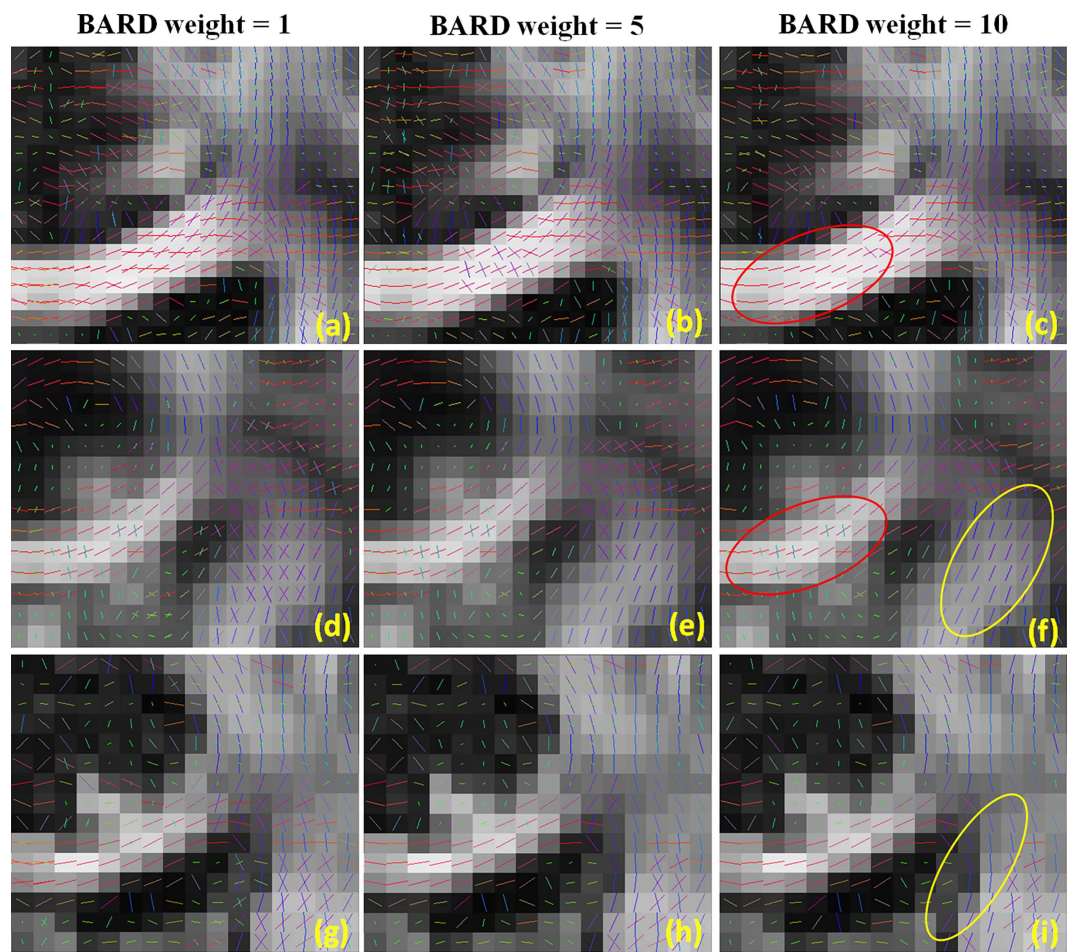


Fig 6. Results of FSL's ball-and-stick parameter estimation for three different FSL-ARD weights. The ROIs and datasets are the same as in Fig 5.

doi:10.1371/journal.pone.0164336.g006

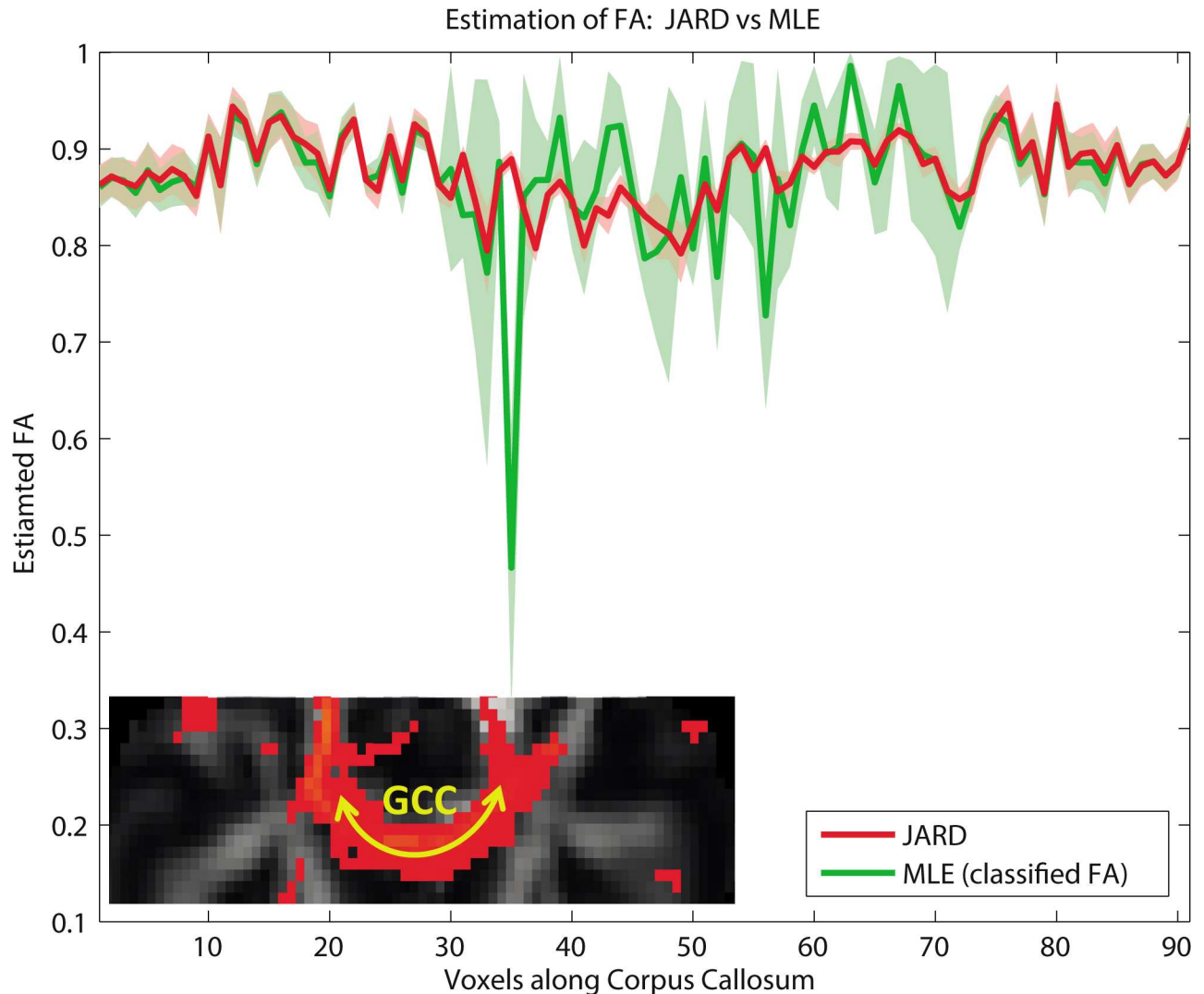


Fig 7. Estimated FA by JARD (red) and MLE (green) as a function of position along the genu of the corpus callosum (GCC trajectory is indicated by the yellow arrow on the inset). The estimated uncertainty (indicated by the colored background) was calculated by \pm the square root of the Cramer-Rao lower bound. The trajectory along the GCC consists of a single fiber region in the middle (30–70) and enclosed by crossing fibers region (with CST) on both ends. The figure displays the FA of the estimated tensor compartments assigned to the GCC by front evolution for both JARD and MLE.

doi:10.1371/journal.pone.0164336.g007

JARD framework (red) and MLE (green). We had to select the tensor compartment that corresponds to the GCC since the labels assigned by MLE and JARD are random. To solve this, the FA belonging to CC was selected based on “front evolution” [37]. In front evolution, the estimated tensor of one compartment is randomly chosen as the reference tensor. Then, the tensor of a neighborhood voxel with the smallest Frobenius norm to the reference tensor receives the same label. After processing all neighbors of the current front as such, these neighbors become the new reference tensors for the next iteration. The green (MLE) and red (JARD) areas indicate the uncertainty in the estimated value as quantified by the square root of the CRLB.

Confirming the above findings, estimation by JARD yields a rather constant FA with small variance. In contrast, MLE yields FA values with a larger variance, particularly in the central part of GCC. We attribute this to overfitting.

JARD versus MLE: Dual-tensor FA and volume-fraction maps

FA as well as volume-fraction maps have been used to detect white matter changes [38] [37]. Here, we will display FA and volume fraction maps generated by JARD and MLE and point out the differences.

Specifically, Fig 8 shows color-coded RGB maps encoding FA in the red channel and the corresponding volume fraction in the green channel. Left images show the outcome by JARD, right images by MLE; top images reflect the first tensor, bottom images denote the second tensor. The ordering of the tensors was performed by front evolution [37].

Regarding the JARD outcomes, one can observe that in Fig 8(A) and Fig 8(C) grey matter is dark, representing both small f_1 , FA_1 and small f_2 , FA_2 . Furthermore, in Fig 8(A) single fiber regions, particularly the central part of CC, are yellowish due to simultaneously large f_1 and FA_1 ; in Fig 8(C) the regions are greenish because of a small f_2 . In Fig 8(A) crossing fiber regions are yellowish again reflecting large f_1 and FA_1 ; in Fig 8(C) these regions are slightly more greenish, because of the smaller f_2 .

In contrast, MLE does not specifically force the volume fraction of one tensor to zero in single fiber and gray matter regions. Therefore, the corpus callosum in Fig 8 (D) contains more yellow spots than Fig 8(C) that reflect large f_2, FA_2 . At the same time gray matter regions in Fig 8(B) and 8(D) contain reddish spots due to the small FA for substantial volume fractions f_1, f_2 .

In general, Fig 8 confirms that MLE estimation is not able to automatically cope with single fiber regions whereas JARD suppresses the volume fraction of one tensor in such areas. In a region in the corpus callosum (indicated by the red circle) we measured $FA_1 = 0.94 \pm 0.05$, $f_1 = 0.96 \pm 0.03$ and $FA_2 = 0.93 \pm 0.06$, $f_2 = 0.04 \pm 0.01$ (mean \pm standard deviation) using JARD. Similarly, we measured $FA_1 = 0.95 \pm 0.04$, $f_1 = 0.65 \pm 0.12$ and $FA_2 = 0.94 \pm 0.04$, $f_2 = 0.47 \pm 0.14$ after MLE. In a region with crossing fibers (indicated with the blue circle) we measured $FA_1 = 0.87 \pm 0.02$, $f_1 = 0.66 \pm 0.01$ and $FA_2 = 0.94 \pm 0.04$, $f_2 = 0.37 \pm 0.03$ using JARD. Similarly, we measured $FA_1 = 0.88 \pm 0.02$, $f_1 = 0.66 \pm 0.02$ and $FA_2 = 0.94 \pm 0.05$, $f_2 = 0.36 \pm 0.03$ using MLE.

TBSS based on dual tensor FA and volume-fraction maps

For a statistical analysis of the FA and volume fractions with age, we included the healthy controls from dataset C [36]. The subjects aged between 45 and 50 (12 subjects, mean-age: 46.2, standard deviation 1.49) and those aged between 65 and 75 (12 subjects, mean-age: 68.2, standard deviation 3.72) were selected from the full control group. All data was registered to the MNI152 standard space using FNIRT [39]. We used the version of FNIRT that is implemented in FSL (version 5.0.7). Subsequently, differences between the two age groups were analyzed by means of the classical TBSS technique, i.e. single tensor analysis, [38] as well as the extended TBSS method for the dual tensor models [39]. Compared to the classical TBSS method, the extended TBSS technique employs 'front evolution' to avoid swapping of the two anisotropic components between the different images. Extended TBSS was used to analyze differences in the dual tensor volume fractions (comparable to [37]) as well as differences in the dual tensor FA maps between the two age groups. Importantly, our approach facilitates such a separate analysis of tensor volume and shape. Notice that the volume fraction used in [37] is a different variable, representing the stick strength in a ball-and-stick model.

Fig 9 shows an axial slice containing multiple areas with crossing fibers and regions with just a single fiber. Fig 9(A) shows the classical, single tensor TBSS analysis and Fig 9(B) shows the extended TBSS analysis of the volume fraction estimated from FSL's ball-and-stick model. Fig 9C and 9E contain the extended TBSS analysis of the dual tensor FA maps, and Fig 9D and 9F) the extended TBSS analysis of the dual tensor volume fraction maps. The dual tensor

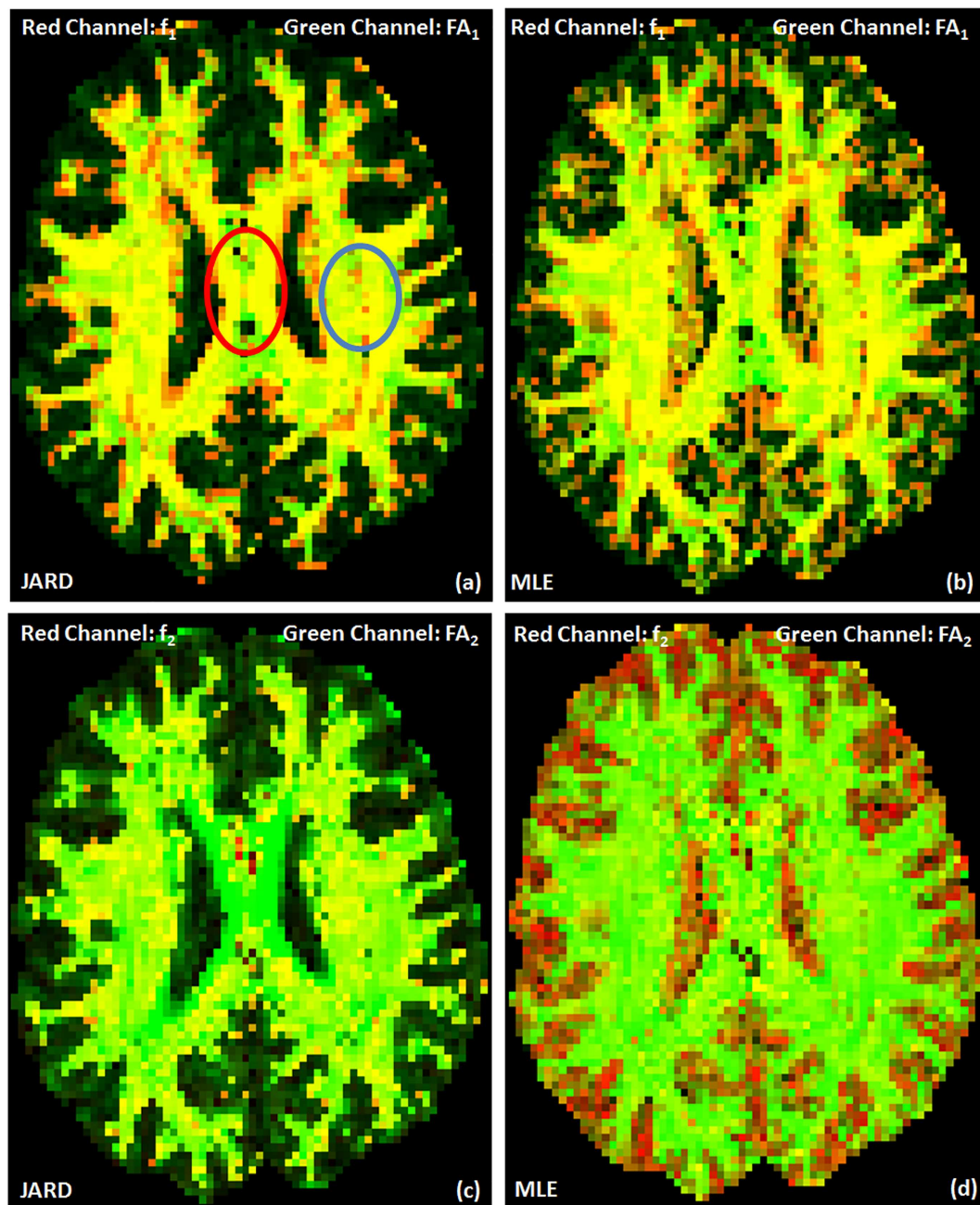


Fig 8. Color-coded output displaying the FA (green channel) and the corresponding volume fraction (red channel) for ARD and MLE. The tensor compartments were classified into first and second by front evolution. FA of first tensor and its corresponding volume fraction by JARD (a) and by MLE (b); FA of the second tensor and its corresponding volume fraction by JARD (c) and by MLE (d). A percentile stretch was performed on the data for more contrast. This stretching put the 5% lowest signal values to zero and the 5% highest values to 1; the remaining values were linearly mapped in between. The red circle indicates a single fiber region, the blue circle a region with crossing fibers.

doi:10.1371/journal.pone.0164336.g008

estimations in Fig 9C and 9D were obtained by JARD, those in Fig 9E and 9F by means of MLE. The red-yellow colored regions in Fig 9A–9F identify regions where the differences are significant.

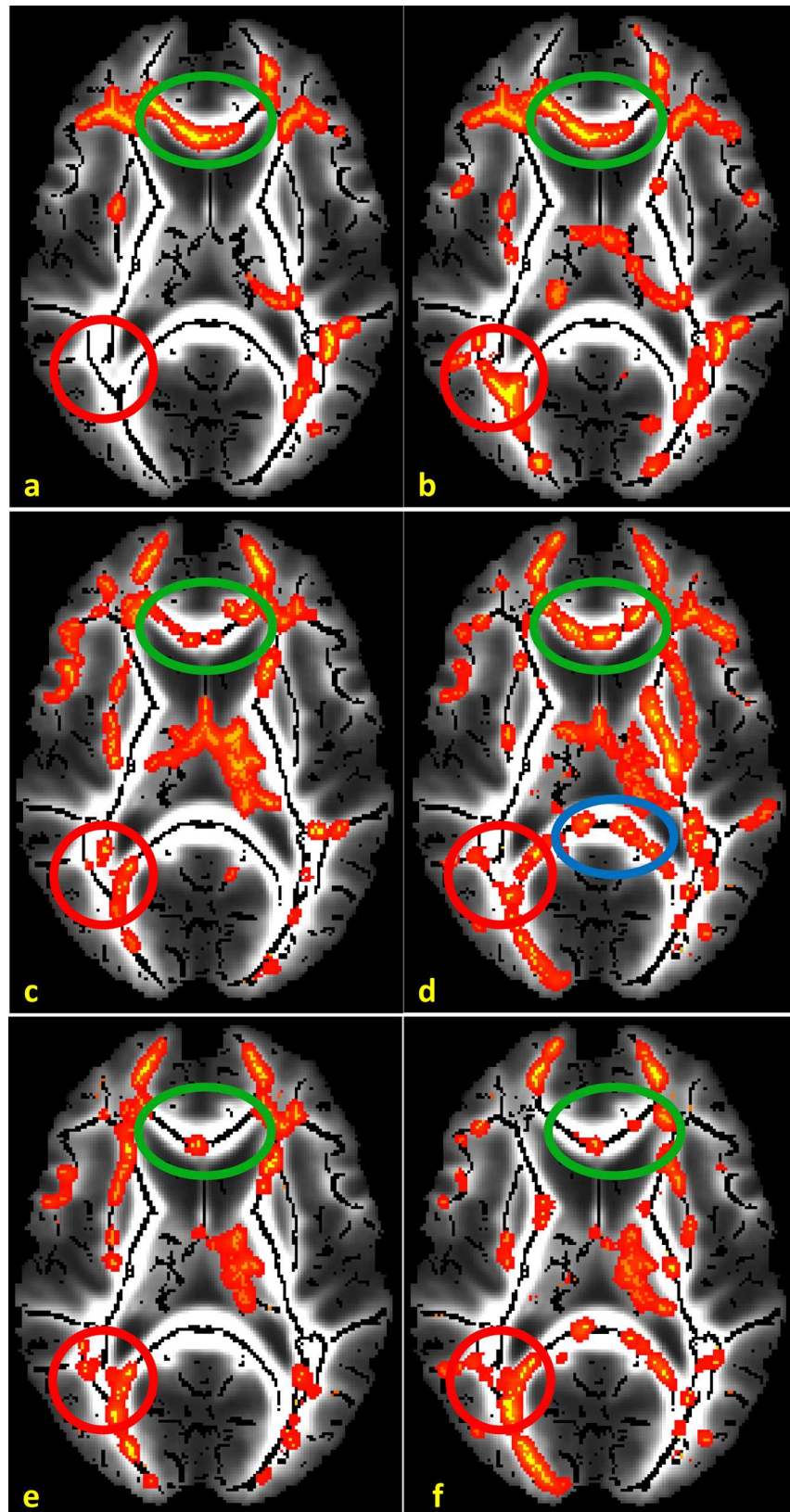


Fig 9. Results of TBSS on the two age groups of dataset C: classical, single tensor TBSS analysis (a); extended TBSS analysis of the volume fractions derived from FSL's ball-and-stick model (b); extended TBSS analysis of the fiber-specific FA and volume fraction maps given by JARD (c, d) and MLE (e, f). Red-Yellow (region/tract) indicates where the FA or the volume fraction of the younger subgroup is significant larger than that of the older subgroup.

doi:10.1371/journal.pone.0164336.g009

It can be observed that the single tensor approach does not detect significant differences in the crossing fiber regions indicated by the red ellipse (Fig 9(A)). By comparison, the analysis of FSL's ball-and-stick method yields remarkably more regions with significant differences, particularly in fiber crossings such as in the red ellipse (Fig 9(B)). The JARD method (Fig 9C and 9D)) finds slightly more significant differences compared to both the classical approach and FSL's ball-and-stick method in single tensor regions, see the right part of the green circle (Fig 9C and 9D)) and also in the a blue circle (Fig 9(D)). This signifies that the superfluous parameters are effectively eliminated by JARD in single fiber regions. In single fiber regions the extended TBSS analysis based on MLE (Fig 9E and 9F)) yields smaller regions with significant differences compared to the classical approach (Fig 9(A)). We attribute this to the large variability in such regions that we already observed with MLE e.g. in Fig 2. Furthermore, we found the expected similarities regarding detected differences in crossing regions (e.g. the red circles) between the MLE and JARD. This indicates that Jeffreys prior does not affect the MLE outcome in such regions. All significant differences are negative, i.e. reduced FA and volume fraction with increasing age. This outcome confirms the finding that significant age-related white matter atrophy was found in the corpus callosum [40].

Discussion

We developed a new framework for estimating the parameters of a constrained dual tensor model in diffusion-weighted MRI. It automatically determines to what extent the diffusion in a voxel should be modeled by one or two rank-2 tensors. In essence, the complexity of the model is implicitly inferred with JARD in a Bayesian probabilistic manner. An initial guess of the diffusion model is obtained by fitting a dual tensor model to the data with MLE. Subsequently, a new automated relevance determination method assesses whether two tensors are 'mandatory' to model the data. If this is not the case, the volume fraction of the superfluous tensor automatically reduces to (nearly) zero.

The proposed framework extends previous work by Behrens et al and Jbabdi et al [38] [37]. A crucial difference is that we employ a rank-2 tensor model, whereas the previous works concerned ball-and-stick models. As such, we aim to recover the full diffusion shape. In FSL 5.0.9, released after the initial submission of this manuscript, bedpostx was extended to include axially symmetric tensors (see <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/WhatsNew#anchor1>). This might reduce the number of voxels where bedpostx finds multiple fibers in the corpus callosum (as in Fig 5). However, our model still differs from this new bedpostx in two important ways: (1) bedpostx forces both tensors to have the same perpendicular diffusivity and hence FA, and (2) our prior is data-acquisition adaptive. Furthermore, an important novelty of our work is that Jeffreys non-informative prior is exploited in JARD, yielding an alternative for previously used informative priors. It facilitates accurate and precise estimation of the volume fractions as well as the diffusion properties with a DTM in single fiber and crossing fiber regions. Jeffreys prior is based on the Fisher's information matrix which accounts for properties of the data acquisition, such as diffusion weighting b -value, the gradient directions and the effective SNR. Therefore we call this method data-acquisition adaptive.

We demonstrated that both in a central part of the corpus callosum (single fiber) as well as in a region where the corpus callosum crosses with the corticospinal tract, the configuration

inferred by our method corresponds to the expected neuro-anatomy [41] [42]. The proposed framework has been compared with direct MLE of the same dual tensor model. Several differences between the proposed framework and MLE were observed. In regions that were considered to contain just a single fiber, MLE typically inferred a large volume fraction for both tensor components (see Fig 5). Here, the proposed framework yielded a single tensor representation by diminishing the volume fraction of the second tensor component. Furthermore, the FA estimated by MLE showed much more variation than the FA estimated by JARD in such a region (Fig 7). In regions that were considered to contain crossing fibers, the results of MLE and JARD were similar.

There are a few limitations of our method. Firstly, we assume a mono-exponential decay along the eigenvectors of the three compartments up to $b = 3000 \text{ s/mm}^2$. Measuring at higher b -values will certainly introduce sensitivity to different compartments such as the myelin sheet [33] with the associated restricted and hindered diffusion [5]. In the latter case, the Gaussian diffusion assumption is no longer valid. However, investigating such non-Gaussian diffusion is beyond the scope of this work. As such, we follow [3] and [43]. Secondly, recent studies reported the presence of a three-way crossing of fiber bundles [14] [44]. In our framework a dual-tensor model is employed to characterize voxels encompassing crossing fibers. The reason for limiting the number of anisotropic components to two is the limited SNR in our HARDI data. Notice that whereas [14] and [44] only recover the fiber orientation, we aim to reconstruct the full diffusion shape, which requires a higher SNR. In [3] we showed that estimating a dual rank-2 tensor model already requires HARDI at two b -values, data of sufficient SNR, and some model restrictions. The latter is needed to ensure stability as the number of model parameters may approach or even surpass the number of degrees of freedom present in the data. Therefore, fitting a triple rank-2 tensor model to voxels with a three-way-crossing will be even more challenging [13]. Developing methods for estimation of the diffusion properties in three-way-crossing fiber bundles will remain an important challenge for future research.

Diffusion imaging may reveal several aspects to white matter integrity: (I) locations of alterations; (II) which fiber tract is affected; (III) the exact change in diffusion. Previously, many solutions were already proposed for the first two aspects. Our work focused on all three aspects. Particularly, our framework may aid a more accurate characterization of the diffusion shape.

Acknowledgments

This work was supported by Fonds Nuts Ohra under grant number 1002–028.

Author Contributions

Conceptualization: JY DP FV LvV.

Data curation: FV MC TS CM.

Formal analysis: JY DP FV LvV.

Funding acquisition: FV LvV.

Investigation: JY DP FV LvV.

Methodology: JY DP FV LvV.

Project administration: FV LvV.

Resources: FV MC TS CM.

Software: JY DP MC.

Supervision: FV LvV.

Validation: JY FV LvV.

Visualization: FV.

Writing – original draft: JY.

Writing – review & editing: FV LvV.

References

1. Basser PJ, Mattiello J, Le Behan D. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*. 1994; 66: p. 259–267. doi: [10.1016/S0006-3495\(94\)80775-1](https://doi.org/10.1016/S0006-3495(94)80775-1) PMID: [8130344](https://pubmed.ncbi.nlm.nih.gov/8130344/)
2. Alexander DC, Barker GJ. Optimal imaging parameters for fiber orientation estimation in diffusion MRI. *Neuroimage*. 2005; 27(2): p. 357–367. doi: [10.1016/j.neuroimage.2005.04.008](https://doi.org/10.1016/j.neuroimage.2005.04.008) PMID: [15921931](https://pubmed.ncbi.nlm.nih.gov/15921931/)
3. Caan MW, Khedoe H, Poot D, Den Dekker A, Olabarriga S, Grimbergen K, et al. Estimation of diffusion properties in crossing fiber bundles. *IEEE Trans Med Imaging*. 2010; 29(8): p. 1504–1515. doi: [10.1109/TMI.2010.2049577](https://doi.org/10.1109/TMI.2010.2049577) PMID: [20562045](https://pubmed.ncbi.nlm.nih.gov/20562045/)
4. Behrens TEJ, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, et al. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic resonance in medicine*. 2003; 50(5): p. 1077–1088. doi: [10.1002/mrm.10609](https://doi.org/10.1002/mrm.10609) PMID: [14587019](https://pubmed.ncbi.nlm.nih.gov/14587019/)
5. Assaf Y, Basser PJ. Composite hindered and restricted model of diffusion (CHARMD) MR imaging of the human brain. *Neuroimage*. 2005; 27: p. 48–58. doi: [10.1016/j.neuroimage.2005.03.042](https://doi.org/10.1016/j.neuroimage.2005.03.042) PMID: [15979342](https://pubmed.ncbi.nlm.nih.gov/15979342/)
6. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*. 2012; 61(4): p. 1000–1016. doi: [10.1016/j.neuroimage.2012.03.072](https://doi.org/10.1016/j.neuroimage.2012.03.072) PMID: [22484410](https://pubmed.ncbi.nlm.nih.gov/22484410/)
7. Panagiotaki E, Schneider T, Siow B, Hall MG, Lythgoe MF, Alexander DC. Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *Neuroimage*. 2012; 59(3): p. 2241–2254. doi: [10.1016/j.neuroimage.2011.09.081](https://doi.org/10.1016/j.neuroimage.2011.09.081) PMID: [22001791](https://pubmed.ncbi.nlm.nih.gov/22001791/)
8. Behrens T, Berg H, Jbabdi S, Rushworth M, Woolrich M. Probability diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage*. 2007; 34: p. 144–155. doi: [10.1016/j.neuroimage.2006.09.018](https://doi.org/10.1016/j.neuroimage.2006.09.018) PMID: [17070705](https://pubmed.ncbi.nlm.nih.gov/17070705/)
9. Zhang H, Hubbard PL, Parker GJ, Alexander DC. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *Neuroimage*. 2011; 56(3): p. 1301–1315. doi: [10.1016/j.neuroimage.2011.01.084](https://doi.org/10.1016/j.neuroimage.2011.01.084) PMID: [21316474](https://pubmed.ncbi.nlm.nih.gov/21316474/)
10. Alexander DC. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine*. 2008; 60(2): p. 439–448. doi: [10.1002/mrm.21646](https://doi.org/10.1002/mrm.21646) PMID: [18666109](https://pubmed.ncbi.nlm.nih.gov/18666109/)
11. Barazany D, Basser PJ, Assaf Y. In vivo measurement of axon diameter distribution in the corpus callosum of rat brain. *Brain*. 2009; 132(5): p. 1210–1220.
12. Zhang H, Dyrby TB, Alexander DC. Axon diameter mapping in crossing fibers with diffusion MRI. In *In Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*; 2011: Springer Berlin Heidelberg. p. 82–89.
13. Kreher B, Schneider JF, Mader I, Martin E, Henning J, Il'yasov KA. Multitensor approach for analysis and tracking of complex fiber configurations. *Magn Reson Med*. 2005; 54: p. 1216–1225. doi: [10.1002/mrm.20670](https://doi.org/10.1002/mrm.20670) PMID: [16200554](https://pubmed.ncbi.nlm.nih.gov/16200554/)
14. Jeurissen B, Leemans A, Jones DK, Tournier JD, Sijbers J. Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. *Human brain mapping*. 2013; 32(3): p. 461–479.
15. Scherrer B, Taquet M, Warfield SK. Reliable selection of the number of fascicles in diffusion images by estimation of the generalization error. In *In Information Processing in Medical Imaging*; 2013: Springer Berlin Heidelberg.
16. Neal RM. Bayesian learning for neural network: PhD thesis; 1995.
17. Jbabdi S, Sotiropoulos SN, Savio AM, Graña M, Behrens TE. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magn Reson Med*. 2012; 68(6): p. 1846–1855. doi: [10.1002/mrm.24204](https://doi.org/10.1002/mrm.24204) PMID: [22334356](https://pubmed.ncbi.nlm.nih.gov/22334356/)

18. Nummenmaa A, Auranen T, Hamalainen MS, Jaakelainen IP, Sams M, Vehtari A, et al. Automatic relevance determination based on hierarchical Bayesian MEG inversion in practice. *Neuroimage*. 2007; 37: p. 876–889. doi: [10.1016/j.neuroimage.2007.04.021](https://doi.org/10.1016/j.neuroimage.2007.04.021) PMID: [17627847](https://pubmed.ncbi.nlm.nih.gov/17627847/)
19. Jeffreys H. *Theory of Probability* London: Oxford University Press; 1961.
20. Jeffreys H. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. London A*. 1946; 186: p. 453–461.
21. Akaike H. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*. New York: Springer; 1998. p. 199–213.
22. Gudbjartsson H, Patz S. The Rician distribution of noise MRI data. *Magn Reson Med*. 1995; 34: p. 910–914. PMID: [8598820](https://pubmed.ncbi.nlm.nih.gov/8598820/)
23. Sijbers J, Den Dekker AJ, Scheunders P, Van Dyck D. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magn Reson Med*. 2004; 51(3): p. 586–594. doi: [10.1002/mrm.10728](https://doi.org/10.1002/mrm.10728) PMID: [15004801](https://pubmed.ncbi.nlm.nih.gov/15004801/)
24. Van den Bos A. *Parameter Estimation for Scientific and Engineers* Hoboken, NJ: Wiley; 2007.
25. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. *Annals of internal medicine*. 1999; 130(12): p. 1005–1013. PMID: [10383350](https://pubmed.ncbi.nlm.nih.gov/10383350/)
26. MacKay D. Probable networks and plausible predictions: a review of practical Bayesian methods for supervised neural networks. *Comput. in Neur. System*. 1995; 6: p. 496–505.
27. Bernardo J, Smith A. *Bayesian Theory* New York: Wiley; 2000.
28. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH. Equation of state calculations by fast computing machines. *The journal of chemical physics*. 1953; 21: p. 1087–1093.
29. Hasting WK. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*. 1970; 57: p. 97–109.
30. Kass R, Wasserman L. Formal rules for selecting prior distributions: a review and annotated bibliography. *Jasa*. 1996; 91: p. 1343–1370.
31. Poot DHJ. *Advances in the reconstruction and statistical processing of magnetic resonance images*; 2010.
32. B S, SK. W. Why multiple b-values are required for multi-tensor models. Evaluation with a constrained log-Euclidean model. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on 2010 Apr 14*. IEEE.; 2010. p. pp. 1389–1392.
33. Peled S, Friman O, Jolesz F, Westin CF. Geometrically constrained two tensor model for crossings tracts in DWI. *Magn Reson Imaging*. 2006; 24: p. 1263–1270. doi: [10.1016/j.mri.2006.07.009](https://doi.org/10.1016/j.mri.2006.07.009) PMID: [17071347](https://pubmed.ncbi.nlm.nih.gov/17071347/)
34. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. *Radiology*. 1996; 201(3): p. 637–648. doi: [10.1148/radiology.201.3.8939209](https://doi.org/10.1148/radiology.201.3.8939209) PMID: [8939209](https://pubmed.ncbi.nlm.nih.gov/8939209/)
35. Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, et al. The Human Connectome Project: a data acquisition perspective. *Neuroimage*. 2012; 62: p. 2222–2231. doi: [10.1016/j.neuroimage.2012.02.018](https://doi.org/10.1016/j.neuroimage.2012.02.018) PMID: [22366334](https://pubmed.ncbi.nlm.nih.gov/22366334/)
36. Su T, Caan M, Wit FJS, Geurtsen G, Schmand B, . . . Majoie C. White matter abnormalities in males with suppressed HIV-infection on Combination Antiretroviral Therapy compared to representative controls. In *In Proc 23th ISMRM*; 2014. p. 185.
37. Jbabdi S, Behrens TE, Smith SM. Crossing fibres in tract-based spatial statistics. *Neuroimage*. 2010; 49(1): p. 249–256. doi: [10.1016/j.neuroimage.2009.08.039](https://doi.org/10.1016/j.neuroimage.2009.08.039) PMID: [19712743](https://pubmed.ncbi.nlm.nih.gov/19712743/)
38. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*. 2006; 31(4): p. 1487–1505. doi: [10.1016/j.neuroimage.2006.02.024](https://doi.org/10.1016/j.neuroimage.2006.02.024) PMID: [16624579](https://pubmed.ncbi.nlm.nih.gov/16624579/)
39. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. *Fsl*. *Neuroimage*. 2012; 62(2): p. 782–790. doi: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015) PMID: [21979382](https://pubmed.ncbi.nlm.nih.gov/21979382/)
40. Vernooij MW, de Groot M, van der Lugt A, Ikram MA, Krestin GP, Hofman A, et al. White matter atrophy and lesion formation explain the loss of structural integrity of white matter in aging. *Neuroimage*. 2008; 43(3): p. 470–477. doi: [10.1016/j.neuroimage.2008.07.052](https://doi.org/10.1016/j.neuroimage.2008.07.052) PMID: [18755279](https://pubmed.ncbi.nlm.nih.gov/18755279/)
41. Zeki S. *A Vision of the Brain*.: Blackwell scientific publications; 1993.
42. S W, PC VZ, LM. NP. *MRI atlas of human white matter*: Amsterdam: Elsevier; 2005 May 11.
43. Taquet M, Scherrer B, Boumal N, Macq B, Warfield SK. Estimation of a multi-fascicle model from single b-value data with a population-informed prior. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI*; 2013: Springer Berlin Heidelberg. p. 695–702.

44. Sotiropoulos SN, Jbabdi S, Xu J, Andersson JL, Moeller S, Auerbach EJ, et al. Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*. 2013; 80: p. 125–143. doi: [10.1016/j.neuroimage.2013.05.057](https://doi.org/10.1016/j.neuroimage.2013.05.057) PMID: [23702418](https://pubmed.ncbi.nlm.nih.gov/23702418/)