A Bayesian method with reparameterization for diffusion tensor imaging - art. no. 69142J

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: ZHOU, D. ... et al., 2008. A Bayesian method with reparameterization for diffusion tensor imaging [69142J]. Medical Imaging 2008: Image Processing, 69142J (March 11, 2008); doi:10.1117/12.771697

Additional Information:


Copyright 2008 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Metadata Record: https://dspace.lboro.ac.uk/2134/17102

Version: Published

Publisher: © Society of Photo-optical Instrumentation Engineers (SPIE)

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
A Bayesian method with reparameterization for diffusion tensor imaging

Diwei Zhoua, Ian L. Drydena, Alexey Koloydenko a, Bai Li b
aDept. of Mathematical Sciences, Univ. of Nottingham, University Park, Nottingham, UK, NG7 2RD, bDept. of Computer Science & IT, Univ. of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, UK, NG8 1BB

ABSTRACT

A multi-tensor model with identifiable parameters is developed for diffusion weighted MR images. A new parameterization method guarantees the symmetric positive-definiteness of the diffusion tensor. We set up a Bayesian method for parameter estimation. To investigate properties of the method, Monte Carlo simulated data from three distinct DTI direction schemes have been analyzed. The multi-tensor model with automatic model selection has also been applied to a healthy human brain dataset. Standard tensor-derived maps are obtained when the single-tensor model is fitted to a region of interest with a single dominant fiber direction. High anisotropy diffusion flows and main diffusion directions can be shown clearly in the FA map and diffusion ellipsoid map. For another region containing crossing fiber bundles, we estimate and display the ellipsoid map under the single tensor and double-tensor regimes of the multi-tensor model, suitably thresholding the Bayes factor for model selection.

Keywords: Diffusion tensor imaging, multi-tensor model, Bayesian method, Monte Carlo simulation

1. INTRODUCTION

Diffusion magnetic resonance imaging (MRI) is a non-invasive method in medical and biological imaging analysis for quantifying the motion of water molecules in biological systems. The tissue structure determines molecular motion. Therefore, diffusion MRI provides information about the microstructure and organization of tissues[1].

In diffusion MRI, the Brownian displacements of water molecules over a fixed time are normally distributed with zero mean, but with a standard deviation that is proportional to the apparent diffusion coefficient (ADC) and diffusion time, i.e. \( N(0,2Dt) \). In a free biological environment, the measured ADC is a single (scalar) constant which is independent of the orientation of tissue. In general, molecular diffusion need not be isotropic (i.e. the same in all directions). Anisotropy may result from the barriers of biological tissue[1] where water molecules move within some preferred directions. Developing a model for capturing diffusion anisotropy has been an important topic in diffusion MRI.

1.1 Diffusion tensor imaging

In the presence of anisotropy, a single ADC cannot adequately characterize the water diffusion. An advanced model that reflects the directional dependence of diffusion properties is the diffusion tensor[1]:

\[
D = \begin{pmatrix}
D_{xx} & D_{xy} & D_{xz} \\
D_{yx} & D_{yy} & D_{yz} \\
D_{zx} & D_{zy} & D_{zz}
\end{pmatrix}
\] (1)

This diffusion tensor is represented by a symmetric positive-definite matrix. Three diagonal elements, \( D_{xx}, D_{yy} \) and \( D_{zz} \), represent molecular diffusivity along \( x, y \) and \( z \) directions, respectively, in the reference frame. Diffusion tensor MRI assumes that molecular displacement is a zero-mean trivariate Gaussian distribution[3], and its covariance matrix is proportional to \( D \).
The eigen-structure of $D$ gives the picture of molecular diffusion along different directions. Since $D$ is a 3x3 symmetric positive-definite matrix, its three eigenvalues ($\lambda_1, \lambda_2, \text{and} \lambda_3$) are positive, and the corresponding eigenvectors ($v_1, v_2$ and $v_3$) are orthogonal (we also normalize $v_1, v_2$ and $v_3$ to have unit length). More importantly, the eigenvectors and eigenvalues of $D$ coincide with the main diffusion directions and associated diffusivities, respectively, in the tissue\cite{3}. In particular, the principal eigenvector (i.e. the one corresponding to the largest eigenvalue), say, $v_1$ represents the main fiber direction, and $\lambda_1$ is the diffusivity along the main fiber direction.

### 1.2 Diffusion ellipsoid and diffusion anisotropy indices

A diffusion tensor can be visualized graphically by a diffusion ellipsoid (see Fig. 1). In a particular voxel, the eigenvectors of $D$ define three semi-axes of the ellipsoid, and the lengths of the semi-axes are proportional to the square roots of the eigenvalues\cite{3}. To quantitatively and easily measure and monitor diffusion properties, several scalar invariants, or indices, have been derived from $D$\cite{4}\cite{5}\cite{6}. Fractional anisotropy (FA) is one of the most popular diffusion anisotropy indices used by the MRI community, and is given in (2) below.

$$\text{FA} = \sqrt{\frac{3}{2} \sum_{k=1}^{3} (\lambda_k - \hat{\lambda})^2}$$  

(2)

First, FA gives a high white-gray matter contrast, i.e., FA=1 describes the fully anisotropic diffusion, and FA=0 is for complete isotropy. FA maps are therefore intuitive to interpret when the white matter is rendered white and gray matter - dark (see Fig. 2).

![Fig. 1 Diffusion ellipsoid](image1.png)  
![Fig. 2 FA map](image2.png)

### 1.3 Diffusion tensor estimations

Several models and methods have been proposed to estimate the diffusion tensor(s). For the fiber bundles with one dominant orientation, the linear least squares method (LLS)\cite{7}\cite{8} and the nonlinear least squares method (NLS)\cite{8} have been used to estimate a single $D$ from the single diffusion tensor model. The constrained LLS, the constrained NLS and Cholesky parameterization\cite{9} have been explored for guaranteeing positive eigenvalues. For brain regions where some fiber bundles may be crossing, kissing and diverging with more than one orientation\cite{5}, the multi-tensor model has been developed\cite{10}\cite{11} to analyze the diffusion properties of these complex fiber configurations.

### 2. MODELS

#### 2.1 Single diffusion tensor model

Under the 3D Gaussian assumption of molecule displacement, the mean $\mu_i$ of the resulting diffusion weighted signal $S_i$ corresponding to the $i$th diffusion gradient direction $g_i$ (unit vector) can be modeled as\cite{12}

$$S_i = \mu_i + \epsilon_i = S_0 \exp(-bg_i^T Dg_i) + \epsilon_i, \quad i = 1, \ldots, N,$$

(3)
where \( S_0 \) is the signal without diffusion gradient applied (i.e. \( b=0 \)). Roughly, \( b \) (b-value) characterizes the gradient pulses used in MRI sequence, \( N \) is the number of gradient directions, and \( \mu_i \) is the mean of \( S_i \). For each voxel, the noise variables \( \epsilon_i \)'s are commonly assumed to be independent and identically distributed (i.i.d.) Gaussians, \( \epsilon_i \sim N(0, \sigma^2) \). Thus, \( S_i \)'s are also independent Gaussian variables, i.e., \( S_i \sim N(\mu_i, \sigma^2) \). Here, the coefficients of \( D \) and \( \sigma^2 \) are the only unknown parameters.

### 2.2 Multiple compartments model

The single-tensor model is commonly carried out under the assumption that the principal eigenvector is aligned along the dominant fiber direction. However, there are regions in the brain with more than one distinct fiber orientation captured in a single voxel. Thus, a voxel in general can contain \( m \geq 1 \) distinct compartments. Modeling diffusion within \( j \)th compartment by a Gaussian distribution with covariance matrix \( D_j \) and assuming no molecular exchange between compartments, we obtain a mixture of the \( m \) Gaussians for the overall diffusion process. Then the mean of \( i \)th diffusion-weighted signal can be modeled as \( ^2 \):

\[
\mu_i = \text{Mean}(S_i) = \sum_{j=1}^{m} a_j S_0 \exp(-b g_i^T D_j g_i),
\]

where the weights \( a_j \in [0,1] \), \( \sum_{j=1}^{m} a_j = 1 \) of the individual compartments are also known as “volume fractions”.

However, the parameter set of the multi-compartment model is non-identifiable. The example below shows that a typical setting of the parameters (\( a_1, a_2, \ldots, a_m, D_1, D_2, \ldots, D_m \)) (with the proper constraints) is not identifiable. Fix \( b = S_0 = 1 \), and \( m = 2 \). Let \( a_1 = 0.2 \), then \( a_2 = 0.8 \). \( D_1 \) and \( D_2 \) are set as follows:

\[
\mu_i = 0.2 \times \exp\left(-g_i^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} g_i\right) + 0.8 \times \exp\left(-g_i^T \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix} g_i\right),
\]

(5)

Since \( g_i^T \cdot g_i = 1 \), \( \mu_i \) can also be written as follows:

\[
\mu_i = 0.1 \times \exp\left(-g_i^T \begin{pmatrix} 1 - \log 2 & 0 & 0 \\ 0 & 2 - \log 2 & 0 \\ 0 & 0 & 3 - \log 2 \end{pmatrix} g_i\right) + 0.9 \times \exp\left(-g_i^T \begin{pmatrix} 4 - \log \frac{8}{9} & 0 & 0 \\ 0 & 5 - \log \frac{8}{9} & 0 \\ 0 & 0 & 1 - \log \frac{8}{9} \end{pmatrix} g_i\right),
\]

(6)

Thus, two distinct settings of the parameters result in the identical model. This is just one example, but the lack of identifiability is inherent in all multi-tensor models, which we now explore.

### 2.3 Multi-tensor model

In order to restore the identifiability of the parameters in a multi-tensor model, we develop a reparameterization of the model. Namely, if \( b > 0 \), let \( q_j = -\log(a_j / b) \), \( q_j \geq 0 \). Define

\[
D_j' = D_j + q_j I_{3x3}
\]

(7)

It can be seen that \( D_j' \) is symmetric and positive-definite. Hence, the multi-tensor model can be defined as
\[ \mu_j = \begin{cases} \sum_{i=1}^{m} S_0 \exp(-b g_i^T D_j^* g_i) & \text{if } b > 0 \\ \frac{S_0}{\sqrt{b}} & \text{if } b = 0 \end{cases} \]  

(8)

To guarantee the symmetric positive-definiteness of \( D_j^* \), an overparameterized embedding is used, i.e., \( D_j^* = Q_j \cdot Q_j^T \), where \( Q_j \) is a general 3x3 matrix. Then,

\[ \mu_j = \begin{cases} \sum_{i=1}^{m} S_0 \exp(-b g_i^T D_j^* Q_j^T g_i) & \text{if } b > 0 \\ \frac{S_0}{\sqrt{b}} & \text{if } b = 0 \end{cases} \]  

(9)

In particular, when \( m=1 \), \( \mu_j = S_0 \exp(-b g_i^T D_j^* Q_j^T g_i) \). When \( m=2 \), \( \mu_j = S_0 \exp(-b g_i^T D_j^* Q_j^T g_i) + S_0 \exp(-b g_i^T D_j^* Q_j^T Q_j^T g_i) \).

Note that \( Q_j \) and \( Q_j R_j \), where \( R_j \in O(3) \) result in the same model. If we ‘rotate’ \( Q_j \) so that \( Q_j R_j \) is lower triangular with positive diagonals, then we have the Cholesky decomposition[9] of \( D_j^* \), which is one way to ensure identifiability with the correct dimension of the parameter space. However, we shall actually keep the high dimensional embedding and the matrix \( R_j \) is then a nuisance parameter matrix, which will be controlled through specification of the prior in a Bayesian model.

3. BAYESIAN METHODS

3.1 Single tensor model

A Bayesian method is proposed to estimate the parameters of the single tensor model. A major advantage of this method is that it incorporates the constraints into the prior distributions of the parameters. The posterior distribution of the parameters expresses all revised knowledge in light of the data.

Since \( S_i \)’s are independent, the likelihood function of \((Q, \sigma^2)\) given \( N \) acquisitions \( S = (S_1, S_2, ..., S_N) \) is given by

\[ L(Q, \sigma^2 | S) = \prod_{i=1}^{N} f(S_i | Q, \sigma^2) \]  

(10)

In particular, we assume \( Q \) and \( \sigma^2 \) to be independent a priori and their priors are chosen below according to our initial beliefs.

\[ \text{vec}(Q) \sim N_{9} (\text{vec}(I_{3x3}), \xi^2 I_{9x9}) \), and  
\[ \sigma^2 \sim \text{Inv-Gamma}(\alpha, \beta), \]  

(11)

(12)

where \( \text{vec}(Q) \) vectorizes \( Q \) by stacking the columns of \( Q \). \( I_{3x3} \) and \( I_{9x9} \) are 3x3 and 9x9 identity matrices, respectively. We will assume large \( \xi \), and so the prior uncertainty about \( Q \) is high.

According to Bayes’ theorem,

\[ \text{posterior} \propto \text{likelihood} \times \text{prior}, \]  

(13)

we can obtain the posterior distribution \( P(Q, \sigma^2 | S) \). The estimates of \( Q \) (and, subsequently estimates of \( D \)) and \( \sigma^2 \) can be obtained by maximizing the posterior distribution. Alternatively we can obtain samples from the posterior distribution using Markov chain Monte Carlo simulation.
3.2 Double tensor model

The double tensor model is

\[ S_i = S_0 \exp(-bg_i^T D_i^* g_i) + S_0 \exp(-bg_i^T D_i^* g_i) + \epsilon_i. \] (14)

By parameterizing \( D_1^* = Q_1 \cdot Q_1^T \) and \( D_2^* = Q_2 \cdot Q_2^T \), we obtain

\[ S_i = S_0 \exp(-bg_i^T Q_i Q_i^T g_i) + S_0 \exp(-bg_i^T Q_i Q_i^T g_i) + \epsilon_i, \] (15)

which is also invariant under the rotations and reflections of \( Q_i \). Again we control this rotation/reflection by introducing a joint prior for \( Q_1 \) and \( Q_2 \). We assume

\[
\begin{align*}
\text{vec}(Q_1) &\sim N_6(\text{vec}(I_{3 \times 3}), \xi_1^2 1_{9 \times 9}), \\
\text{vec}(Q_2) &\sim N_6(\text{vec}(I_{3 \times 3}), \xi_2^2 1_{9 \times 9}), \\
\text{vec}(Q_1 - Q_2) &\sim N_6(\text{vec}(0_{3 \times 3}), \xi_2^2 1_{9 \times 9}), \quad \text{and} \\
\sigma^2 &\sim \text{Inv-Gamma}(\alpha, \beta) .
\end{align*}
\] (16)-(19)

Let \( \rho \) be the coefficient of correlation between \( Q_{1ij} \) and \( Q_{2ij} \), \( 1 \leq i, j \leq 3 \), then \( \rho = 1 - \frac{\xi_2^2}{2 \xi_1^2} \).

4. RESULTS

This section includes two parts, a simulation study for comparison of DTI direction schemes (§4.1), and an application to real data (§4.2).

4.1 Simulations

The purpose of the study is to compare three DTI direction schemes using Bayesian estimation for the signal tensor model and double-tensor model. A direction scheme is a collection of \( g_i = (g_x, g_y, g_z) \in \mathbb{R}^3, \quad i = 1, \ldots, N \). We consider spherical schemes only, that is \( g_x^2 + g_y^2 + g_z^2 = 1 \) (and \( g = -g \)). The accuracy of DTI measurements depends on DTI direction scheme applied. Uniformly distributed diffusion directions schemes are used currently to provide equal noise levels for arbitrarily oriented fibers. Three schemes of interest are Philips 15 (Fig. 3 (a)), Philips 32 (Fig. 3 (b)), and a Uniform 32 directions scheme. The Uniform 32 directions scheme is more uniformly distributed over the sphere than the two Philips schemes.

Fig. 3 Three DTI Direction Schemes: (a) Philips 15 direction scheme (b) Philips 32 direction scheme (c) Uniform 32 direction scheme.
For the single tensor model, we choose the following diffusion tensor:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$  

The eigenvalues are 1, 2 and 3, and the corresponding eigenvectors are along $x$, $y$ and $z$ axis respectively (Fig. 4).

![An ellipsoid representation of $D$](image)

We simulate our data $n=100$ times according to the multivariate Gaussian distribution, $N(\mu, \sigma^2 I_{NXN})$ with $\mu = (\mu_1, \mu_2, \ldots, \mu_N)$, for each of the three direction schemes and for a range of $\sigma^2$.

We allowed $M=1, 5, 10, 15, 30$ replicates for each of the $N$ directions, and $n=100$ Monte Carlo simulations were performed for each value of $M$. In practice, however, $M$ is usually one. Fig. 5(a)(b) shows the root mean squared errors (RMSE) of LLS and Bayesian estimators of $D$ for the single tensor model ($\sigma^2 = 5$). The definition of RMSE is

$$RMSE(\hat{D}) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} || \hat{D}(k) - D ||^2}.$$  

(21)

where $\hat{D}(k)$ is the estimator from the $k$th simulation. Both RMSE plots support that the Uniform 32 scheme gives lower RMSE than the Philips schemes. Importantly, the Bayesian estimator outperforms the LLS one in all the three schemes.

![Plots of RMSE of $D$ for three DTI direction schemes: (a) RMSE for LLS estimator (b) RMSE for Bayesian estimator.](image)

For the double-tensor model, we define $D_1$ and $D_2$ as follows:

$$D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
and the corresponding ellipsoids are shown in Fig. 6.

Bayesian estimation is now used and 100 Monte Carlo simulations ($\sigma^2 = 5$) were also carried out for the double-tensor model describing the diffusion behavior in crossing fiber bundles. The tensors $D_1$ and $D_2$ were estimated from $M=1, 5, 10, 15$ and $30$ replicates and the respective RMSEs as well as the means of the individual non-zero tensor coefficients are shown in Fig. 8 and Fig. 7, respectively. For both $D_1$ and $D_2$, the Uniform 32 scheme provides nearly perfect estimators. Indeed, all of the six non-zero coefficients are estimated with no bias under the Uniform 32 scheme. On the other hand, Philips 15 and 32 schemes noticeably underestimate the two larger coefficients of the second tensor while overestimating the remaining four coefficients. The Uniform 32 scheme also results in lower RMSE when compared with the Philips schemes (see Fig. 8). It is also noticeable that RMSEs with the Uniform 32 and Philips 15 schemes decrease as sample size $M$ increases from 1 to 30. However, Philips 32 scheme shows growing error at sample size 15. There are clear difficulties in estimating $D_1$ for the Philips 32 scheme, which has directions far from parallel to the principal eigenvector of $D_1$ (along the $z$-axis).

Fig. 6 Ellipsoids for two defined diffusion tensors: $D_1$ is oblate and $D_2$ is prolate.

Fig. 7 Mean of 100 estimated $D_1$ and $D_2$: (a) Mean $D_1$ coefficient (b) Mean $D_2$ coefficient

Fig. 8 Root Mean Squared Error: (a) RMSE ($D_1$) (b) RMSE ($D_2$).
4.2 Real data

In this section, the DTI model and multi-tensor model with the Bayesian method are applied for two regions of interest (ROI). The Academic Radiology Department of Queens Medical Center, University of Nottingham provides a set of the MR images with the Uniform 32 DTI direction scheme from a healthy human brain. The MR images were acquired using a spin echo EPI (echo planar imaging) sequence with diffusion weighting gradients applied with a weighting factor of \( b = 1000 \text{ s/mm}^2 \). 52 interleaved contiguous transaxial slices were acquired throughout the subject’s head in a matrix of 112x112 (interpolated to 224x224) with an acquisition voxel size of 1x1x2 mm\(^3\). For each slice, the acquisition was repeated to acquire diffusion weighted images in the 32 non-collinear directions (the Uniform scheme 32, see Fig. 3 (c)), as well as one volume with no diffusion weighting (\( b = 0 \)).

Fig. 9 (a) (b) shows the coronal view of the brain with FSL\(^{[14]}\) and a zoomed ROI. This ROI contains part of corpus callosum and corona radiata. FSL’s outputs are obtained with LLS estimation: red line segments represent the principal eigenvector (main fiber orientation), and FA value for each voxel is grey colored. Fig. 10 and Fig. 11 show FA map and ellipsoid map obtained with MATLAB by applying Bayesian estimation to the single DTI model. Higher anisotropy corresponding to higher FA value has brighter color in FA map. Both Fig. 9 (b) and Fig. 10 reveal high anisotropic diffusion in corpus callosum and part of corona radiata. In Fig. 11, two diffusion flows in corpus callosum and corona radiata can be observed clearly with larger volumes of diffusion ellipsoids and consistent principal diffusion directions.

![Fig. 9 FSL’s coronal view 1 (front to back): red lines represent principal eigenvectors, and background is FA map. (a) Whole brain (b) region of interest: part of corpus callosum and corona radiata.](image)

![Fig. 10 FA map: Bayesian estimation for single tensor.](image)
The other ROI is shown in Fig. 12 and is the crossing part of corpus callosum and corona radiata. Fig. 13(a) shows the ellipsoid map with single-tensor fitting. Higher anisotropy is represented by larger ellipsoid. In this ROI, more than one fiber orientation is supposed. Therefore, multi-tensor model (m = 2) with Bayesian estimation can be applied. Fig. 13(b) shows the Bayesian estimates of the principal direction map of selected single and double-tensor models. The threshold of the Bayes factor for models selection is 100. The double tensor model does reflect the main directions of the crossing fibers, although the estimates are somewhat noisy. It is clear that the method would benefit from spatial regularization, which can be added in a straightforward way by specifying a Gaussian Markov random field prior for the parameter matrix $Q$. This work is currently underway.
5. CONCLUSION

A multi-tensor model was derived from the multi-compartment model with improved identifiability of the parameters. Parameterizing the diffusion tensor by the product of a 3x3 matrix and its transpose guarantees the symmetric positive-definite property. A least squares method and a Bayesian method with suitable prior distributions were used to estimate the parameters (the diffusion tensor coefficients and the variance of noise) for the single DTI model and the multi-tensor model.

For applications, three DTI direction schemes were compared using our models and methods. The Uniform 32 scheme with 32 uniformly distributed directions works better than the other two schemes. A real MR dataset then was studied by fitting single and double tensors to the fiber bundle with one dominant orientation and crossing fiber bundles, respectively. The diffusion directions (eigenvectors) and the associated diffusivities (eigenvalues) from the Bayesian estimators then were visualized by means of the FA map and diffusion ellipsoid map of two regions of interest.

The model can be easily extended to include spatial regularity through a Gaussian Markov random field prior for $Q$. Also, simulation from the posterior is relatively straightforward using Markov chain Monte Carlo method, giving the additional information about the uncertainty in the estimates of the diffusion tensors. Such uncertainty can be built in to subsequent fiber tracking algorithms, and this work will be reported elsewhere.

ACKNOWLEDGMENT

We would like to thank Dorothee Auer, Paul Morgan, Christopher Tench and Stamatios Sotiropoulos of the School of Medical and Surgical Sciences, the University of Nottingham for providing the data and helpful discussions. This study is funded by the European Commission FP6 Human Resources and Mobility program to provide training and mobility activities.
REFERENCES