Parsimonious Model Selection for Tissue Segmentation and Classification Applications: A Study Using Simulated and Experimental DTI Data

Raisa Z. Freidlin*, Evren Özarslan, Michal E. Komlosh, Lin-Ching Chang, Cheng Guan Koay, Derek K. Jones, and Peter J. Basser

***Abstract*—One aim of this work is to investigate the feasibility of using a hierarchy of models to describe diffusion tensor magnetic resonance (MR) data in fixed tissue. Parsimonious model selection criteria are used to choose among different models of diffusion within tissue. Using this information, we assess whether we can perform simultaneous tissue segmentation and classification. Both numerical phantoms and diffusion weighted imaging (DWI) data obtained from excised pig spinal cord are used to test and validate this model selection framework. Three hierarchical approaches are used for parsimonious model selection: the Schwarz criterion (*SC*), the F-test t-test ($F-t$), proposed by Hext, and the F-test F-test (F–F), adapted from Snedecor. The $F-t$ approach is more robust than the others for selecting between isotropic and general anisotropic (full tensor) models. However, due to its high sensitivity to the variance estimate and bias in sorting eigenvalues, the $F-F$ and *SC* are preferred for segmenting models with transverse isotropy (cylindrical symmetry). Additionally, the *SC* method is easier to implement than the $F-t$ and $F-F$ methods and has better performance. As such, this approach can be efficiently used for evaluating large MRI data sets. In addition, the proposed voxel-by-voxel segmentation framework is not susceptible to artifacts caused by the inhomogeneity of the variance in neighboring voxels with different degrees of anisotropy, which might contaminate segmentation results obtained with the techniques based on voxel averaging.**

***Index Terms*—Diffusion tensor, diffusion tensor imaging (DTI), diffusion tensor magnetic resonance imaging (DT-MRI), hierarchical, magnetic resonance imaging (MRI), model selection, parsimonious, segmentation, tissue classification.**

I. Introduction

DIFFUSION tensor magnetic resonance imaging (DT-MRI) [1] is a noninvasive imaging technique for quantitative analysis of intrinsic features of tissues. DT-MRI has been applied to study the structural organization of skeletal muscles [2], brain [3], spinal cord [4], peripheral nerves [5], intervertebral discs [6], and heart muscle [7], [8]. Based on its extensive use, it is increasingly important to develop new tools for efficient and accurate tissue analysis and segmentation of DT-MRI data, since better characterization of organization of different tissue types may enhance our understanding structure/function relationships in organs. In addition, quantitative tissue segmentation may advance intrasubject comparisons between tissue compartments. Most DT-MRI work to date has focused on characterizing the trace of the diffusion tensor (Tr), the fractional anisotropy (FA), and the fiber orientation of tissue. Comparatively little has been done to identify the underlying microstructure and microstructural models appropriate for each voxel.

Automated tissue segmentation and classification are among the most challenging tasks in DT image analysis. Segmentation separates acquired data into objects, while tissue classification generates meaningful regions of interest. Here, we examine whether parsimonious model selection criteria applied to a hierarchy of diffusion models can simultaneously segment and classify tissues based on their underlying diffusion properties.

A hierarchy of diffusion models and a statistical hypothesis testing framework were used in the context of the first magnetic resonance (MR) measurement of the translational diffusion tensor [1] to determine whether proton diffusion was isotropic in water and anisotropic in a skeletal muscle phantom. Because this study used diffusion spectroscopy sequences with data obtained at high signal-to-noise ratios (SNRs), it was not clear whether such statistical approaches would work at the SNR of clinical or animal images or would behave reliably from voxel to voxel within an image volume. Subsequently, it was used for diffusion models with different degrees of symmetry to characterize different modes of diffusion transport in tissue [9], [10].

Other parsimonious model selection methods have recently been used to analyze DT-MRI data. Alexander *et al.* [12] used a parsimonious modeling framework to test the adequacy of the tensor model applied to human brain tissues. Kroenke *et al.* [12] used a Bayesian model selection approach to analyze different diffusion models in fixed baboon brain. In this work, we test the appropriateness and relative efficiency of four predefined diffusion models: isotropic, general anisotropic, prolate, and oblate. We also compare and contrast three methods for parsimonious model selection, which employ the Schwarz criterion (*SC*) [13], the F-test [14], [15], and the t-test [14].

Manuscript received May 15, 2007; revised August 21, 2007. This work was supported by the Intramural Research Program of the National Institute of Child Health and Development (NICHD) and the Center for Information Technology (CIT), National Institutes of Health, Bethesda, MD. *Asterisk indicates corresponding author.*

*R. Z. Freidlin is with the Telemedicine and Applied Imaging Section, Computational Bioscience and Engineering Laboratory, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892 USA and also with the Department of Electrical and Computer Engineering, George Washington University, Washington DC 20052 USA

E. Özarslan, M. E. Komlosh, L.-C. Chang, C. G. Koay, and P. J. Basser are with the Section on Tissue Biophysics and Biomimetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892 USA.

D. K. Jones is with the Cardiff University Brain and Repair Imaging Centre, School of Psychology, Cardiff University, CF10 3XQ, Cardiff, U.K.

Digital Object Identifier 10.1109/TMI.2007.907294

II. Theory

A. Diffusion Tensor Imaging

The relationship between observed echo attenuation [1], [16], [17], caused by applying diffusion sensitizing gradients along various directions, and the diffusion tensor $\mathbf{D}$ can be characterized by

$$S(\mathbf{G}) = S(0)e^{-\mathbf{b}:\mathbf{D}} \tag{1}$$

where ":" stands for the tensor dot product, $S(\mathbf{G})$ is the observed signal, $S(0)$ is a signal in the absence of the diffusion-weighting gradient, and $\mathbf{b}$ is a matrix whose components are given by

$$b_{ij} = \gamma^2 G_i G_j \delta^2 \left[\Delta - \frac{\delta}{3}\right] \tag{2}$$

where G_i is the magnitude of the diffusion gradient pulse applied in ith direction $(i = x, y, z)$ with duration δ, and Δ is the diffusion time. In (1), $\mathbf{D}$ is a symmetric (3×3) second-order diffusion tensor that has a form

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}. \tag{3}$$

Diagonal elements of the diffusion tensor $\mathbf{D}$ are proportional to the diffusion rate in the collinear directions, while correlations in displacements along orthogonal directions are represented by off-diagonal elements. The six independent elements of $\mathbf{D}$ are sufficient to describe Gaussian molecular diffusivity in three dimensions. Furthermore, given that $\mathbf{D}$ is symmetric and positive definite, it can be characterized bythe three orthonormal eigenvectors, ϵ_1, ϵ_2, and ϵ_3, associated with three positive eigenvalues, λ_1, λ_2, and λ_3. In the matrix form it is represented by

$$\begin{aligned} &\mathbf{DE} = \mathbf{E}\Lambda \text{ or } \mathbf{D} = \mathbf{E}\Lambda\mathbf{E}^{-1} = \mathbf{E}\Lambda\mathbf{E}^{\mathrm{T}} \\ &\mathbf{E} = [\epsilon_1; \epsilon_2; \epsilon_3] \text{ and } \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \end{aligned} \tag{4}$$

where $\mathbf{E}$ is the matrix whose columns are the orthonormal eigenvectors and Λ is the diagonal matrix containing their corresponding eigenvalues. It was first suggested in [18] that in fibrous anisotropic media the eigenvector ϵ_1 associated with the largest eigenvalue λ_1 coincides with the tissue's dominant fiber-tract axis, while the two remaining eigenvectors, ϵ_2 and ϵ_3, define the transverse plane. For the general anisotropic model, a typical observation is that $\lambda_1 > \lambda_2 > \lambda_3$.

It also has long been assumed that some anisotropic tissues, like skeletal muscle [2] and nerve white matter [19], are cylindrically symmetric having a prolate diffusion ellipsoid, i.e., $\lambda_1 > \lambda_2 \approx \lambda_3$. Cylindrical symmetry associated with an oblate diffusion ellipsoid entails $\lambda_1 \approx \lambda_2 > \lambda_3$. The diffusion tensors for both these transversely isotropic models can be written as [20]

$$\mathbf{D}_{o/p}(\alpha, \beta, \theta, \varphi) = \alpha \boldsymbol{\epsilon}_{o/p}(\theta, \varphi)\boldsymbol{\epsilon}_{o/p}(\theta, \varphi)^{\mathrm{T}} + \beta\mathbf{I} \tag{5}$$

where $\boldsymbol{\epsilon}_{o/p}(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^{\mathrm{T}}$ is the unit vector parallel to the axis of symmetry. For the prolate model, $\boldsymbol{\epsilon}_p$ corresponds to $\boldsymbol{\epsilon}_1$, which is parallel to the long axis of the prolate diffusion ellipsoid. This was considered previously in [21] and [22]. However, we show how the same equation describing transverse isotropy can be applied for estimating both the prolabe and oblate tensor models for the oblate model, where $\boldsymbol{\epsilon}_o$ corresponds to $\boldsymbol{\epsilon}_3$, which is parallel to the axis of symmetry of the "pancake" shaped diffusion ellipsoid. Above, $\mathbf{I}$ is the 3×3 identity matrix. For the oblate model, we have $\lambda_1 = \lambda_2 \Rightarrow \alpha = \lambda_3 - \lambda_2, \beta = \lambda_2$. For the prolate model, we have $\lambda_2 = \lambda_3 \Rightarrow \alpha = \lambda_1 - \lambda_2, \beta = \lambda_2$.

In contrast to an anisotropic medium, only one scalar diffusion coefficient D is necessary to describe isotropy. Then (1) reduces to

$$S(\mathbf{G}) = S(0)e^{-bD} \tag{6}$$

where $b = b_{xx} + b_{yy} + b_{zz}$, and D is an apparent diffusion coefficient (ADC). The isotropic diffusion tensor has the form

$$\mathbf{D} = D\mathbf{I}. \tag{7}$$

Equation (7) is also a special case of (5) in which $\alpha = 0$ and $\beta = D$. In this section, we have established a hierarchy of nested models of diffusion given by (3), (5), and (7).

1) Parameters Estimation for the Different Models: A nonlinear least square (NLS) minimization method, proposed by Koay *et al.* [23] is used to estimate each diffusion tensor whose initial guess is obtained from the linear least squares minimization. In the NLS method, (1) takes the form

$$S(\mathbf{G}) = e^{-\mathbf{B}\Psi} \tag{8}$$

where the design matrix $\mathbf{B}$ consists of a list of $\mathbf{b}$-matrix elements for a series of n trials

$$\mathbf{B} = \begin{bmatrix} b_{x_1}^2 & b_{y_1}^2 & b_{z_1}^2 & 2b_{x_1y_1} & 2b_{x_1z_1} & 2b_{y_1z_1} & -1 \\ b_{x_2}^2 & b_{y_2}^2 & b_{z_2}^2 & 2b_{x_2y_2} & 2b_{x_2z_2} & 2b_{y_2z_2} & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{x_n}^2 & b_{y_n}^2 & b_{z_n}^2 & 2b_{x_ny_n} & 2b_{x_nz_n} & 2b_{y_nz_n} & -1 \end{bmatrix} \tag{9}$$

and Ψ consists of six independent parameters of the estimated diffusion tensor $\hat{\mathbf{D}}$ and the estimated log of the signal in the absence of the diffusion-weighting gradient, $\log[S(0)]$. For the

general anisotropic (ga) model Ψ is written as a (7×1) column vector

$$\Psi_{\mathrm{ga}} = [D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}, \log[S(0)]]^{\mathrm{T}}. \quad (10)$$

Since the diffusion tensors for the transversely isotropic models can be estimated from four parameters (5) for cylindrically symmetric oblate (o) and prolate (p) models the number of free parameters we estimate is reduced from 7 [as in (10)] to 5 [24]

$$[\alpha, \beta, \theta, \varphi, \log[S(0)]]^{\mathrm{T}} \quad (11)$$

where the initial guesses of $\alpha, \beta, \theta, \varphi$ are obtained from the previously estimated diffusion tensor, using the following assignments (mathematical support for (13) can be found in [25]).

Oblate $(\lambda_1 \approx \lambda_2)$

$$\begin{aligned} \alpha &= \lambda_3 - \left(\frac{\lambda_1 + \lambda_2}{2}\right) \\ \beta &= \frac{\lambda_1 + \lambda_2}{2} \\ \boldsymbol{\epsilon}(\theta, \varphi) &= \boldsymbol{\epsilon}_3 \end{aligned} \quad (12)$$

Prolate $(\lambda_2 \approx \lambda_3)$

$$\begin{aligned} \alpha &= \lambda_1 - \left(\frac{\lambda_2 + \lambda_3}{2}\right) \\ \beta &= \frac{\lambda_2 + \lambda_3}{2} \\ \boldsymbol{\epsilon}(\theta, \varphi) &= \boldsymbol{\epsilon}_1. \end{aligned} \quad (13)$$

The final $\Psi_{o/p}$ have the same form as Ψ_{ga} (10).

For the isotropic model (I), the number of unknown parameters is 2

$$\Psi_I = \left[\hat{D}, \log[S(0)]\right]^{\mathrm{T}}. \quad (14)$$

We use the following initial guess for the isotropic case: $\hat{D} = (D_{xx} + D_{yy} + D_{zz})/3$ and the design matrix $\mathbf{B}$ is reduced to $n \times 2$, where the first column is a list of averaged diagonal elements of the corresponding $\mathbf{b}$-matrices and the second column consists of -1.

Once the elements of Ψ are estimated for all four models, we can derive the corresponding residual sum of squares (RSS) for each model as

$$\mathrm{RSS}_j = \sum_{i=1}^{n} (S_i(\mathbf{G}) - e^{-\mathbf{B}_i \Psi_j})^2 \quad (15)$$

where $\mathrm{RSS}_j = \{\mathrm{RSS}_{\mathrm{ga}}, \mathrm{RSS}_p, \mathrm{RSS}_o, \mathrm{RSS}_I\}$[1], $\Psi_j = \{\Psi_{\mathrm{ga}}, \Psi_p, \Psi_o, \Psi_I\}$, and $\mathbf{B}_i$ is the ith row of the design matrix, $\mathbf{B}$.

[1]general anisotropic (ga); prolate (p); oblate (o); isotropic models (I)

The unbiased estimate of the residual mean square error for the full tensor model, s^2, on $(n - 7)$ degrees-of-freedom, is defined as

$$s^2 = \frac{\mathrm{RSS}_{\mathrm{ga}}}{n - 7}. \quad (16)$$

The unbiased estimate of the covariance matrix for the full tensor model, Σ_Ψ, is obtained from

$$\Sigma_\Psi = s^2 (\mathbf{J}^{\mathrm{T}} \mathbf{J})^{-1} \quad (17)$$

where $\mathbf{J}$ is a Jacobian matrix of first-order derivatives of $S_i(\mathbf{G}) - e^{-\mathbf{B}_i \Psi}$ with respect to the free parameters evaluated using the optimally estimated free parameters.

With these definitions, we are interested to know whether one could select the model that most faithfully describes the acquired data, which uses the fewest unknown parameters.

B. Hierarchical Approaches to Parsimonious Model Selection

In this work we investigate three schemes for parsimonious model selection: *SC* [13], [10], $F-t$ [14], and $F-F$ [15], [11]. The logical schematic for the *SC*, $F-t$ and $F-F$ methods are shown in Fig. 1(a)–(c), respectively.

The first step for all three approaches is to ensure that the estimated full diffusion tensor passes a χ^2 goodness-of-fit test with the confidence level of 95%. If Ψ is admissible, the next step is to differentiate between the isotropic and anisotropic models. We test the null hypothesis that assumes that the diffusion tensor is isotropic, i.e., that the experimental data can be more economically described by a model with two free parameters (14), rather than the full diffusion tensor model with seven unknowns (10).

The first method we consider is the *SC*, also commonly known as the Bayesian information criterion (BIC). This method works by imposing penalties for models with a larger number of free parameters and a larger mean squared residual error. It is defined as

$$SC_i = \log\left(\frac{\mathrm{RSS}_i}{n}\right) + p_i \frac{\log(n)}{n} \quad (18)$$

where i represents the model type (isotropic, general anisotropic, prolate, or oblate), n is the number of experimental data points, and p is the number of free parameters for the ith model.

The remaining two model selection methods we consider are based upon sequential hypothesis tests. The second method is the $F-t$ hierarchical model selection approach, which is based on the F-test [14]

$$F_{0,\mathrm{iso}}^{F-t} = \frac{(\mathrm{SS}_{\mathrm{ga}} - \mathrm{SS}_I)/(p_{\mathrm{ga}} - p_I)}{(\mathrm{SS}_{\mathrm{Sig}} - \mathrm{SS}_{\mathrm{ga}})/(n - p_{\mathrm{ga}})} \quad (19)$$

where $p_{\mathrm{ga}} = 7$ and $p_I = 2$ are the numbers of the free parameters in the general anisotropic and isotropic models, respectively; n is the number of experimental data points; and $\mathrm{SS}_{\mathrm{Sig}}$, $\mathrm{SS}_{\mathrm{ga}}$, and SS_I are the sums of the squares of the acquired signals, and the fitted signals for the anisotropic and isotropic

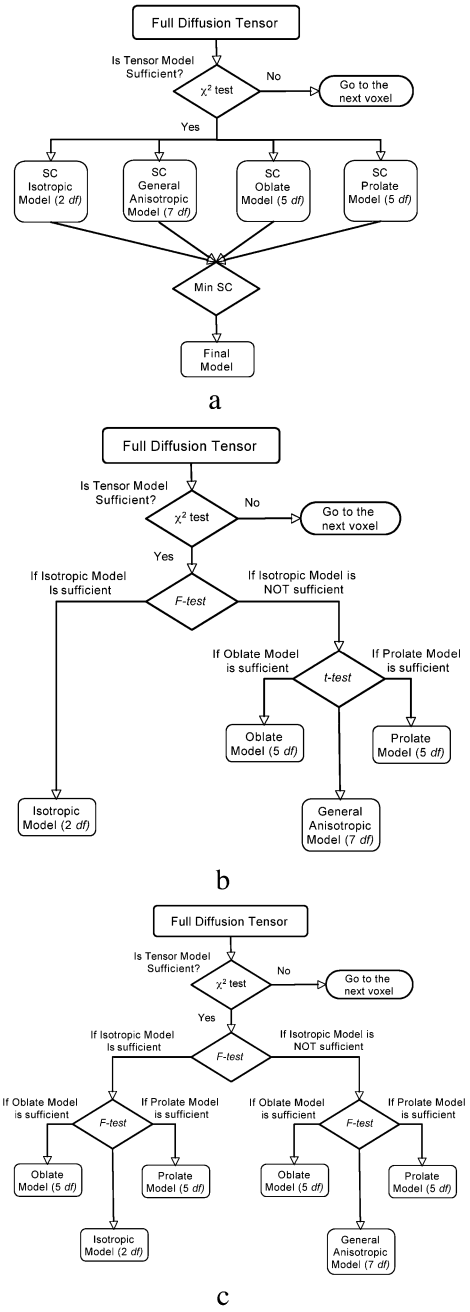

Fig. 1. Schematic hierarchical model selection for (a) *SC*; (b) $F-t$; and (c) $F-F$ approaches.

models, respectively. The $F-F$ approach uses a multivariate F-test [15] performed according to (20)

$$F_{0,\mathrm{iso}}^{F-F} = \frac{(\mathrm{RSS}_I - \mathrm{RSS}_{\mathrm{ga}})/(p_{\mathrm{ga}} - p_I)}{(\mathrm{RSS}_{\mathrm{ga}})/(n - p_{\mathrm{ga}})} \tag{20}$$

where RSS_I and $\mathrm{RSS}_{\mathrm{ga}}$ are the residual sum of squares for the isotropic and general anisotropic models, respectively. Large values for $F_{0,\mathrm{iso}}$, given above, indicate that the isotropic model can be rejected.

If the anisotropic model is accepted, we further investigate whether it is more economical to represent diffusion as being transversely isotropic or cylindrically symmetric, i.e., having either an oblate or prolate-shaped diffusion ellipsoid with one axis of symmetry. The null hypothesis that two eigenvalues are equal, can be evaluated using a number of statistical tests.

To evaluate this hypothesis, in the $F-t$ approach we modify Hext's t-test [14] whose t-statistic is of the form

$$t = \frac{\hat{\lambda}_i - \hat{\lambda}_j}{\sqrt{(\hat{\epsilon}_i - \hat{\epsilon}_j)^{\mathrm{T}} \Sigma_\Psi^{-1} (\hat{\epsilon}_i - \hat{\epsilon}_j)}} \tag{21}$$

where $i = 1$ and $j = 2$ for the oblate model, $i = 2$ and $j = 3$ for the prolate model, Σ_Ψ is the estimate of covariance (17), and $\hat{\epsilon}_i$ is the estimate of the orthonormal eigenvector associated with the corresponding eigenvalue estimate, $\hat{\lambda}_i$ for the full tensor model.

In the third approach, following $F-F$ scheme [1], we use an F-test again to test whether the reduced model for transverse isotropy having five free parameters is more efficient than the full seven parameter tensor model. In this case, the test statistic is given by (22)

$$F_{0,\mathrm{sym}}^{F-F} = \frac{(\mathrm{RSS}_R - \mathrm{RSS}_{\mathrm{ga}})/(p_{\mathrm{ga}} - p_R)}{(\mathrm{RSS}_{\mathrm{ga}})/(n - p_{\mathrm{ga}})} \tag{22}$$

where RSS_R is the residual sum of squares for the reduced model (prolate or oblate for which $p_R = 5$) and $\mathrm{RSS}_{\mathrm{ga}}$ is the residual sum of squares for the general anisotropic model with $p_{\mathrm{ga}} = 7$.

From Fig. 1(c) it can be seen that the $F-F$ approach applies the same scheme, as above, if the isotropic model is accepted.

III. Methods

A. Simulations

To evaluate the various parsimonious model selection approaches, synthetic phantoms [Fig. 2(a)] were generated in MATLAB (The MathWorks, Inc., Natick, MA) by setting the SNR to 15, 25, and 33 ($\mathrm{SNR} = 33$ matches SNR in the excised pig spinal cord DTI data), for a fixed signal intensity, $I_0 = 1000$. The parameters for the general anisotropic and prolate models were chosen to simulate white matter, while the isotropic model parameters simulated gray matter, with values typical for living brain tissue [3]. The oblate model was set to parameters between white and gray matter. The trace of $\mathbf{D}$ Tr for the general anisotropic model was set to 2600×10^{-6} mm^2/s and the FA was set to 0.65($\lambda_1 = 1500 \times 10^{-6}$ mm^2/s, $\lambda_2 = 880 \times 10^{-6}$ mm^2/s, $\lambda_3 = 250 \times 10^{-6}$ mm^2/s), for the prolate model $\mathrm{Tr} = 2100 \times 10^{-6}$ mm^2/s and $\mathrm{FA} = 0.8$, for the oblate model $\mathrm{Tr} = 3000 \times 10^{-6}$ mm^2/s and $\mathrm{FA} = 0.6$, and for the isotropic model $\mathrm{Tr} = 2100 \times 10^{-6}$ mm^2/s and $\mathrm{FA} = 0.3$. Normally distributed random noise was added to the signal intensity in each voxel; the diffusion weighted images were calculated and scaled, as shown in (23). This model assumes that noise is added to the real and imaginary channels independently, and that the MR signal is rectified [26], [27].

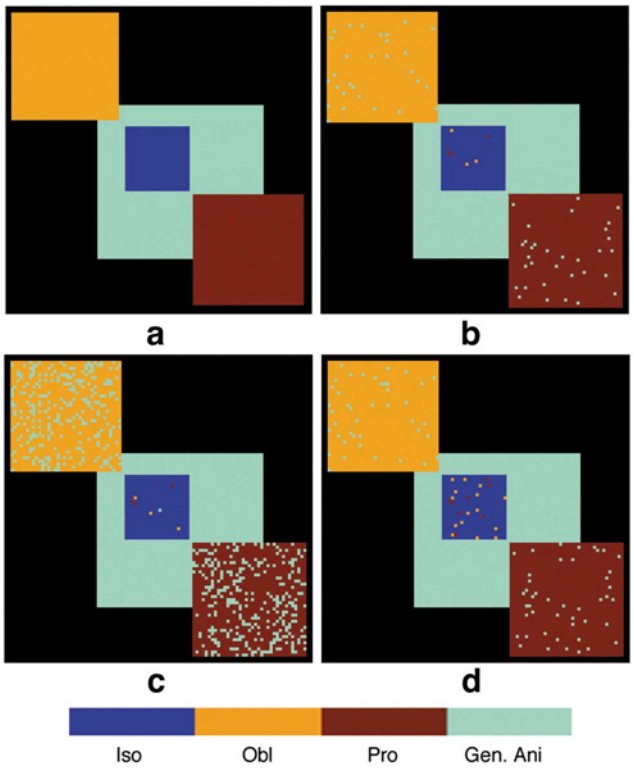

Fig. 2 (a) Noise-free synthetic phantom; Model map results assuming FA $=$ 0.8 in prolate and FA $=$ 0.55 in oblate and general anisotropy regions, respectively, with SNR $=$ 33 for (b) *SC* (c) $F-t$ and (d) $F-F$ methods.

The b-matrix was calculated with the imaging parameters described in Section III-B.

$$\mathrm{DWI} = \sqrt{\mathrm{DWI}_{\mathrm{Re}}^2 + \mathrm{DWI}_{\mathrm{Im}}^2} \tag{23}$$

where

$$\mathrm{DWI}_{\mathrm{Re}} = I_0 e^{-\mathbf{b}:\mathbf{D}} + N_{\mathrm{Re}};$$
$$\mathrm{DWI}_{\mathrm{Im}} = N_{\mathrm{Im}}$$

and N_{Re} and N_{Im} are normally distributed random numbers with mean zero and standard deviation $\sigma = I_0/\mathrm{SNR}$.

The hierarchical methods for parsimonious model selection were applied to the set of 50 reconstructed diffusion-weighted images with four unweighed images.

B. Excised Pig Spinal Cord DTI Experiments

In addition to simulations, we demonstrate our results on experimental MRI data obtained from an excised pig spinal cord fixed with 4% paraformaldehyde solution. Prior to MR data collection, the spinal cord was washed in phosphate-buffered saline (PBS) to avoid signal loss due to fixative-related T_2-shortening [28]. The sample was imaged in a 15-mm NMR tube containing MR-compatible perfluoropolyether oil ("Fomblin"), using a Micro2.5 microscopyprobe (15-mm solenoid coil) with 1450 mT/m 3-axis gradients. A diffusion-weighted spin echo pulse sequence was used with repetition time (TR) $=$ 3500 ms, echo time (TE) $=$ 33 ms, bandwidth=50 kHz, field-of-view

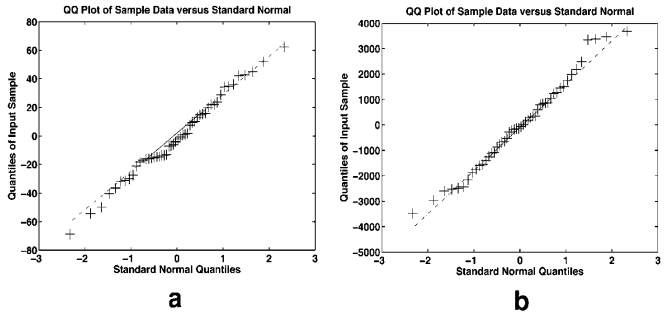

Fig. 3. Q-Q Plot of prolate model residuals in (a) phantom and (b) pig spinal cord versus standard normal.

(FOV) $= 12 \times 12$ mm, matrix $= 128 \times 128$ with seven continuous 1 mm thick slices. Four DWIs per slice were acquired without applying the diffusion sensitizing gradients ($b \approx 0$ s/mm^2), followed by the acquisition of 46 diffusion-weighted images with diffusion gradient strength $(\hat{G}) = 120$ mT/m yielding approximate b-values of 1000 s/mm^2. The number of averages (NEX) was 2. Each of these diffusion-weighted scans were collected with the diffusion gradients applied along a different direction determined from the second-order tessellations of an icosahedron on the surface of a unit hemisphere. The diffusion gradient duration (δ) was 5 ms, and the gradient separation (Δ) was 20 ms. The total imaging time was less than 13 h. At each voxel location in the raw image, the apparent diffusion tensor $\hat{\mathbf{D}}$ was estimated. Tensor-derived parameters, such as the principal directions, ϵ_1, ϵ_2, and ϵ_3, and the corresponding principal diffusivities, λ_1, λ_2, and λ_3, were estimated and passed to the parsimonious model selection algorithm.

IV. Results

Systematic study of the residuals is performed to assess their distribution within different voxels. Fig. 3 shows a representative Q-Q plot of the residuals, indicating that they are normally distributed. Since the residuals from the phantom and the excised pig spinal cord experiments are normally distributed (Fig. 3), and the variance of each measurement is unchanging (homoscedasticity), the use of the hypothesis testing framework given below is well grounded. The confidence interval for all tests was set to 95%.

A. Simulations

Fig. 2(b)–(d) shows graphically the model selection results obtained with the *SC*, F–t and F–F approaches, at SNR $=$ 33. Performance for the model selection framework at SNRs equal to 15, 25, and 33 are shown in Fig. 4(a), (b), and (d), respectively. The true positive and the false negative counts are calculated within the area of the predicted model, while the false positive counts are obtained from the outside regions. The *SC*, $F-t$, and $F-F$ results are similar for the isotropic and general anisotropic models at all SNRs. The isotropic model selection showed consistent results for the true positive and the false negative counts with the averages of 96% and 4% success, respectively. The general anisotropic model performed better with the true positive counts 99.3% success. While the differences between oblate and prolate models for the *SC* and $F-F$

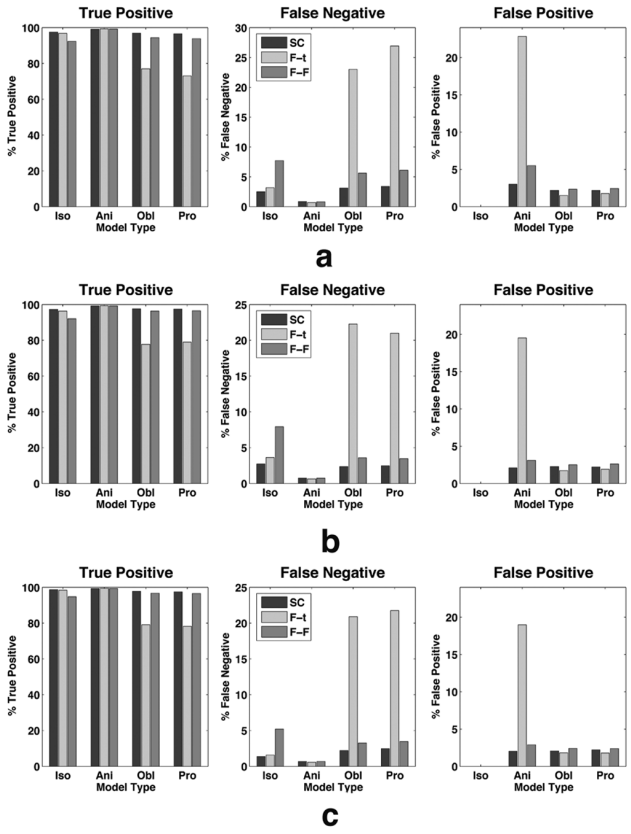

Fig. 4. Performance comparison at (a) $\mathrm{SNR} = 15$, (b) $\mathrm{SNR} = 25$, and (c) $\mathrm{SNR} = 33$, where *SC* is represented by the blue bar, $F - t$ by the green bar, and $F - F$ by the red bar. The true positive and the false negative counts are calculated within the area of the predicted model, while the false positive counts are obtained from the outside regions. In each box the first group of three bars shows the performance of the isotropic model, second group—for the general anisotropic model and the other two groups show the performance of the oblate and prolate models, respectively.

TABLE I
ACCURACY OF A CLASSIFICATION (%) OBTAINED WITH *SC* APPROACH AT $\mathrm{SNR} = 33$

SC

\	Iso	Ani	Obl	Pro
Iso	98.6	0	0.7	0.7
Ani	0	99.3	0.7	0
Obl	0	2.2	97.8	0
Pro	0	2.4	0.1	97.5

TABLE II
ACCURACY OF A CLASSIFICATION (%) OBTAINED WITH $F - t$ APPROACH AT $\mathrm{SNR} = 33$

F-t

\	Iso	Ani	Obl	Pro
Iso	98.4	0.2	0.7	0.7
Ani	0	99.4	0.6	0
Obl	0	20.9	79.1	0
Pro	0	21.8	0	78.2

TABLE III
ACCURACY OF A CLASSIFICATION (%) OBTAINED WITH $F - F$ APPROACH AT $\mathrm{SNR} = 33$

F-F

\	Iso	Ani	Obl	Pro
Iso	94.8	0	3.4	1.8
Ani	0	99.3	0.7	0
Obl	0	3.2	96.8	0
Pro	0	3.5	0	96.5

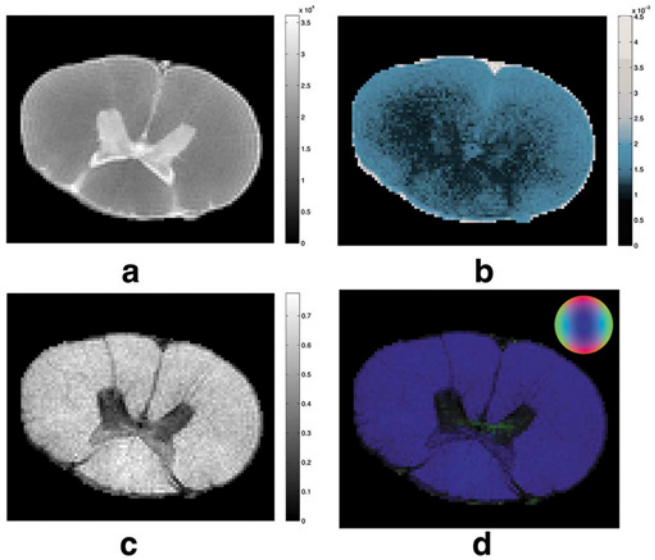

Fig. 5. Excised pig spinal cord images: (a) T_2-weighted amplitude; (b) Trace in mm^2/s; (c) FA; (d) DTI colormap: green—left to right direction, red—up and down direction, blue—through the plane.

approaches were not significant for the true positive and false negative results and their overall errors were around 7% at all SNRs, the $F-t$ approach showed poor performance in identifying the oblate and prolate models with an average error in the range from 21% to 26%. These results for $\mathrm{SNR} = 33$ are shown as confusion matrices in Tables I–III. Here, each column of the confusion matrix represents the true model, and each row represents the results of the parsimonious model classification.

Subsequently, we tested this approach with the fractional anisotropy varying from 0.1 to 0.9. The obtained results were consistent with the described above at $\mathrm{FA} \geq 0.3, \mathrm{FA} \geq 0.4$, and $\mathrm{FA} \geq 0.3$ for the *SC*, $F-t$ and $F-F$ approaches, respectively.

B. Excised Pig Spinal Cord DTI Experiment

Fig. 5 shows (a) the T_2-weighted amplitude image, the orientationally invariant (b) Tr, and (c) FA. The T_2-weighted amplitude image, the Tr and FA maps delineate white and gray matter groups consistent with previous works [29]. Additionally, regions appearing bright in the amplitude image appear dark in the FA maps. Fig. 5(d) is the direction-encoded color map [30]. White matter fiber groups are also easily discernible in this image. The bluish color in white matter groups indicates fibers pointing into the page, consistent with their known anatomy.

To assess the results obtained using the *SC* Fig. 6(a), $F-t$ [Fig. 6(b)], and $F-F$ [Fig. 6(c)] parsimonious model selection methods, we compared our results with the typical DTI maps, i.e., Tr, FA, and color maps [Fig. 5(b)–(d)], respectively. Fig. 7(a) was obtained using the recently developed diffusion orientation transform (DOT) technique [31]. In this image, the orientation profiles computed using the DOT are overlaid on generalized anisotropy (GA) maps displayed using a gray col-

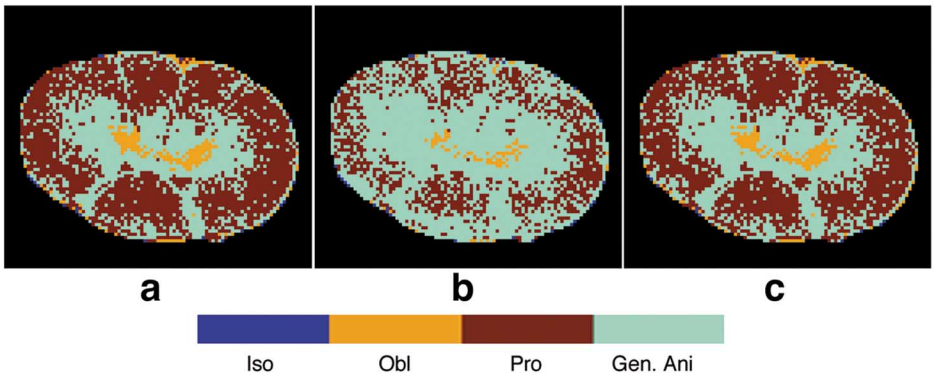

Fig. 6. Parsimonious model maps for excised pig spinal cord using: a) *SC*, b) $F-t$ and c) $F-F$ approaches.

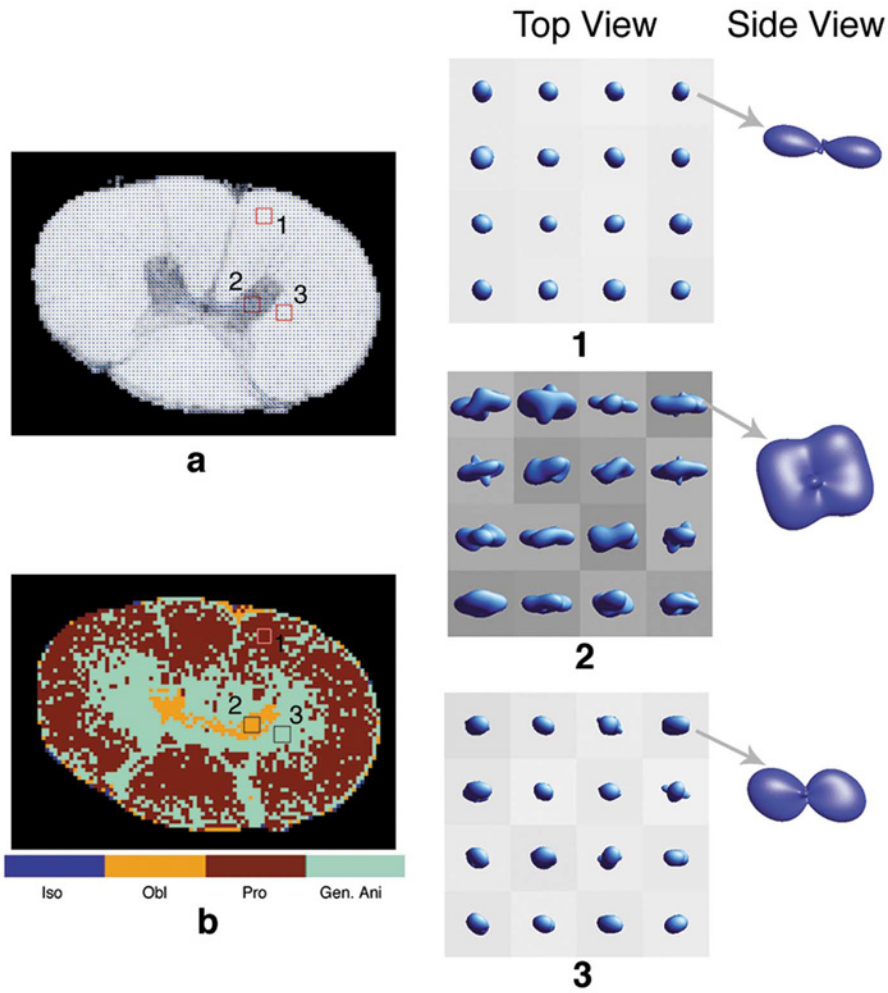

Fig. 7. (a) Diffusion orientation transform (DOT) map: the orientation profiles are overlaid on generalized anisotropy (GA) maps, where blue color represents surfaces of probability profiles and GA is displayed in gray scale. The three insets (4 voxels each) are selected to show detailed 3-D of probability profiles in different regions; (b) Parsimonious model map obtained with the *SC* method. The inset 1 shows the parallel aligned peanut shaped surfaces orientated through the plane (top and side views), which corresponds to the prolate model region in the *SC* map (red color); The inset 2 shows the complicated orientational characteristics, i.e., voxels with more than one fiber direction, which corresponds to the oblate region in the *SC* map (orange color). The inset 3 shows deformed peanuts, which could reflect the presence of more than one fiber orientation. This inset corresponds to the general anisotropic model in the *SC* map (turquoise color).

ormap on the background. The blue color represents surfaces of probability profiles. It is obvious that the coherently oriented and highly anisotropic fibers of white matter appear to be oriented along the direction normal to the image plane (inset 1 in Fig. 7); these regions also have high GA values. In contrast, in most regions of the spinal cord gray matter, there is a great deal of directional heterogeneity in the preference of water diffusion. However, much of the fiber directions appear to be on the image plane suggesting an oblate diffusion profile (inset 2 in Fig. 7), where voxels with more complicated patterns are expected to yield general anisotropy in our model selection framework [inset 3 in Fig. 7(a)]. As can be seen from Fig. 6, the *SC* and $F-F$ approaches consistently selected the prolate model in the areas corresponding to white matter, while the $F-t$ approach frequently failed. Furthermore, although the FA map of the excised pig spinal cord appears to be uniform throughout

the entire white matter region, the Tr map revealed less uniformity in this region. This indicates that some of the regions have more complex structures than the voxels with fibers aligned along one direction. Thus, we may assume that the voxels in the white matter region, which were not selected as prolate, might correspond to the structures with more than one fiber orientation [e.g., *SC* method inset 3 in Fig. 7(b)]. Moreover, more complicated orientational fiber directions in gray matter [e.g., inset 2 in Fig. 7(a)] overlap with the voxels identified as oblate by our model selection frameworks [e.g., *SC* method inset 2 in Fig. 7(b)]. Although the sharpened displacement profiles produced by the DOT [Fig. 7(a)] appear unidirectional in the medial sections of the spinal cord gray matter, the generalized anisotropy (GA) [32] image provided in the background suggests that these orientational features are not as pronounced; this suggests that the variations due to noise were comparable to the overall anisotropy due to diffusional preference. As a result, these voxels are identified as general anisotropic according to our model selection method.

V. Discussion

The aim of this work is to investigate the feasibility of using a parsimonious model selection framework to obtain the most appropriate and economical diffusion model within each voxel of an imaging volume. Analysis of the three proposed model selection methods has shown that the *SC* approach is more robust than the $F-t$ and $F-F$ methods. The t-test applied to selecting models exhibiting transverse symmetry was less successful than the $F-F$ and *SC* approaches, due to its high sensitivity to the variance estimation and the bias owing to sorting the eigenvalues [33], [27]. Since the F–t method requires only one nonlinear fit, rather than the three required for the F–F and *SC* approaches, the former may be computationally less intensive. Currently, we are investigating the effect of variance estimation improvement on the t-test. Monte Carlo simulations have shown that in the F–F method it is unnecessary to perform the F-tests to compare isotropic and oblate/prolate models. Generally, the performance of the *SC* method is better than the other two methods.

Under what circumstances can we justify the use of this hypothesis testing framework for model selection? As long as the residuals are normally distributed, and the variance of each measurement is not changing (homoscedasticity) we can safely use the F-test and t-test formalisms to compare one model to another. In this study, the Q–Q plots and statistical tests supported the use of this sequential hypothesis testing framework for model selection.

While the Bayesian method for model selection [12] is elegant, we have developed this model testing pipeline to accommodate the high throughput of diffusion weighted data of demanding DT-MRI applications. In a typical microimaging study $512 \times 512 \times 64$ voxels might be acquired. The advantage of using this hypothesis testing approach is that most of the calculations required for the hypothesis tests themselves have already been performed during the nonlinear estimation of the free parameters of each model. Thus, the subsequent statistical tests have a small computational overhead. This approach is also easily extended to consider a larger number of nested diffusion models. For instance, one can incorporate models having multiple compartments (such as a CSF compartment), or consider HARDI or CHARMED models within this testing framework. While this initial demonstration uses on diffusion tensor MRI data, the hierarchical testing framework is not limited only to data of this type.

The proposed segmentation methods are based on a voxel-by-voxel analysis. Such approach prevents possible introduction of artifacts due to averaging the neighboring voxels with different variances. The inhomogeneity in the variance is observed in voxels with different degrees of anisotropy [34].

What are the prospects for extending this approach to clinical and *in vivo* biological MRI applications? Provided that the conditions for normally distributed residuals and stable variances for DWI in time are met, this analysis pipeline could be used with *in vivo* data as well. However, one should establish that these conditions are satisfied before using this approach. With the methods introduced in this work, we performed a plot study using a clinical data set, where artifacts, such as physiological noise, small scale motion, and eddy current distortion were ameliorated prior to performing statistical model selection. Preliminary results (not reported here) are promising and will be the subject of the future work.

It has been shown that the segmentation obtained from our parsimonious model selection scheme can be used to inform an unsupervised tissue clustering algorithm, also based on multivariate hypothesis testing [35], [36]. We used the parsimonious model selection framework to identify seed regions for unsupervised tissue clustering algorithms to ensure that the variance of each measurement in the seeding region is uniform (homoscedasticity), and that the distribution of diffusion tensor parameters is similar, making the clustering algorithm more reliable.

Furthermore, model selection maps can be useful for unsupervised tractography, as well. For example, voxels selected as prolate are better candidates for seed voxels from which to launch fibers, while the voxels identified as general anisotropic or oblate will indicate changes in structure, thus requiring more detailed analysis.

In addition, while FA, trace, and skewness [37] maps individually provide useful scalar information about isotropic and anisotropic regions, the parsimonious model selection map distills this information into a single map.

VI. Conclusion

The maps produced by the proposed parsimonious model selection schemes provide useful information about underlying tissue microstructure in each voxel. The simplicity and speed of applying the F- and t-tests and the *SC* make the proposed approaches feasible for large DWI data sets routinely encountered in high resolution microscopic DT-MRI studies or in clinical DT-MRI applications. The results of the phantom simulations increase our confidence in model selection schemes based upon statistical hypothesis tests. When applied to *ex vivo* tissue specimens, where background noise is the primary artifact and other systematic artifacts can be remedied, this approach should work robustly. In clinical applications, however, where other systematic artifacts can corrupt DWI data, this approach may be more

problematic. When using DWI data from living tissue, tests for Gaussianity of the distribution of residuals and a careful assessment of the degree of homoscedasticity should be performed prior to applying our model selection approaches to ensure data integrity. Our expectation is that these model selection procedures may lead to improved methods of automatic region of interest (ROI) delineation and classification of different tissue types in DT-MRI volume data sets.

ACKNOWLEDGMENT

R. Z. Freidlin would like to thank Dr. R. L. Martino and K. Kempner for their support and encouragement on this work. The authors wish to thank W. Jarisch, who originally interested them in parsimonious model selection methods, and particularly, in the use of the Schwarz criterion, Dr. C. Pierpaoli, Dr. U. Nevo and Dr. Y. Assaf for helpful discussions throughout this study, and L. Salak for editing this paper. The authors also would also like to thank the people who were involved in providing this research with the biological specimens. R. R. Clevenger, T. J. Hunt, G. J. Zywicke, A. D. Zetts, K. Keeran, S. M. Kozlov, and K. R. Jeffries, from LAMS, NHLBI, for supplying fixed pig spinal cord tissue.

REFERENCES

[1] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *J. Magn. Reson. B.*, vol. 103, no. 3, pp. 247–254, Mar. 1994.

[2] U. Sinha and L. Yao, "In vivo diffusion tensor imaging of human calf muscle," *J. Magn. Reson. Imag.*, vol. 15, no. 1, pp. 87–95, Jan. 2002.

[3] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. D. Chiro, "Diffusion tensor MR imaging of the human brain," *Radiology*, vol. 201, no. 3, pp. 637–648, Dec. 1996.

[4] C. A. Clark and D. J. Werring, "Diffusion tensor imaging in spinal cord: Methods and applications—A review," *NMR Biomed.*, vol. 15, no. 7–8, pp. 578–586, 2002.

[5] M. Skorpil, M. Karlsson, and A. Nordell, "Peripheral nerve diffusion tensor imaging," *Magn. Reson. Imag.*, vol. 22, no. 5, pp. 743–745, Jun. 2004.

[6] E. W. Hsu and L. A. Setton, "Diffusion tensor microscopy of the intervertebral disc anulus fibrosus," *Magn. Reson. Med.*, vol. 41, no. 5, pp. 992–999, May 1999.

[7] J. Dou, T. G. Reese, W.-Y. I. Tseng, and V. J. Wedeen, "Cardiac diffusion MRI without motion effects," *Magn. Reson. Med.*, vol. 48, no. 1, pp. 105–114, Jul. 2002.

[8] T. G. Reese, R. M. Weisskoff, R. N. Smith, B. R. Rosen, R. E. Dinsmore, and V. J. Wedeen, "Imaging myocardial fiber architecture in vivo with magnetic resonance," *Magn. Reson. Med.*, vol. 34, no. 6, pp. 786–791, Dec. 1995.

[9] P. J. Basser, "Testing for and exploiting microstructural symmetry to characterize tissues via diffusion tensor MRI," in *ISMRM 4th Sci. Meeting*, 1996, vol. 2, p. 1323.

[10] D. Shrager, D. K. Jones, S. Pajevic, P. Munson, and P. J. Basser, "When is a gaussian displacement distribution adequate to describe water diffusion in tissues?," in *ISMRM Workshop Diffusion MRI: Biophys. Issues: What Can We Measure?*, 2002, pp. 21–25.

[11] D. C. Alexander, G. J. Barker, and S. R. Arridge, "Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data," *Magn. Reson. Med.*, vol. 48, no. 2, pp. 331–340, Aug. 2002.

[12] C. D. Kroenke, G. L. Bretthorst, T. E. Inder, and J. J. Neil, "Modeling water diffusion anisotropy within fixed newborn primate brain using Bayesian probability theory," *Magn. Reson. Med.*, vol. 55, no. 1, pp. 187–197, Jan. 2006.

[13] G. Schwarz, "Estimating the dimension of the model," *Ann. Statist.*, vol. 6, pp. 461–468, 1978.

[14] G. R. Hext, "The estimation of second-order tensors, with related tests and designs," *Biometrika*, vol. 50, pp. 353–357, 1963.

[15] G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8th ed. Malden, MA: Blackwell, 1989.

[16] H. Y. Carr and E. M. Purcell, "Effects of diffusion on free precession in nuclear magnetic resonance experiments," *Phys. Rev.*, vol. 94, no. 3, pp. 630–438, May 1954.

[17] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *J. Chem. Phys.*, vol. 42, no. 1, pp. 288–292, Jan. 1966.

[18] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, no. 1, pp. 259–267, Jan. 1994.

[19] J. Coremans, R. Luypaert, F. Verhelle, T. Stadnik, and M. Osteaux, "A method for myelin fiber orientation mapping using diffusion-weighted MR images," *Magn. Reson. Imag.*, vol. 12, no. 3, pp. 443–454, 1994.

[20] S. G. Advani and C. L. Tucker, "The use of tensors to describe and predict fiber orientation in short fiber composites," *J. Rheol.*, vol. 31, no. 8, pp. 751–784, Nov. 1987.

[21] D. Alexander, *An Introduction to Computational Diffusion MRI: The Diffusion Tensor and Beyond, Ser. Visualization and Image Processing of Tensor Fields*. New York: Springer, 2005.

[22] O. Friman and C.-F. Westin, "Uncertainty in white matter fiber tractography," *Med. Image. Comput. Comput. Assist. Interv. Int. Conf. Med. Image. Comput. Comput. Assist. Interv.*, vol. 8, pt. 1, pp. 107–114, 2005.

[23] C. G. Koay, L.-C. Chang, J. D. Carew, C. Pierpaoli, and P. J. Basser, "A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging," *J. Magn. Reson.*, vol. 182, pp. 115–125, Jul. 2006.

[24] E. W. Hsu and S. Mori, "Analytical expressions for the NMR apparent diffusion coefficients in an anisotropic system and a simplified method for determining fiber orientation," *Magn. Reson. Med.*, vol. 34, no. 2, pp. 194–200, Aug. 1995.

[25] O. Friman, G. Farneback, and C.-F. Westin, "A bayesian approach for stochastic white matter tractography," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 965–978, Aug. 2006.

[26] R. M. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Med. Phys.*, vol. 12, no. 2, pp. 232–233, 1985.

[27] C. Pierpaoli and P. J. Basser, "Toward a quantitative assessment of diffusion anisotropy," *Magn. Reson. Med.*, vol. 36, no. 6, pp. 893–906, Dec. 1996.

[28] T. M. Shepherd, P. E. Thelwall, P. E. Stanisz, and S. J. Blackband, "Chemical fixation alters the water microenvironment in rat cortical brain slices—Implications for MRI contrast mechanisms," in *Proc. Int. Soc. Magn. Reson. Med.*, 2005, no. 13, p. 619.

[29] B. A. Inglis, L. Yang, E. D. Wirth, D. Plant, and T. H. Mareci, "Diffusion anisotropy in excised normal rat spinal cord measured by NMR microscopy," *Magn. Reson. Imag.*, vol. 15, no. 4, pp. 441–450, 1997.

[30] S. Pajevic and C. Pierpaoli, "Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: Application to white matter fiber tract mapping in the human brain," *Magn. Reson. Med.*, vol. 42, no. 3, pp. 526–540, Sep. 1999.

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

[32] E. Özarslan, B. C. Vemuri, and T. H. Mareci, "Generalized scalar measures for diffusion MRI using trace, variance and entropy," *Magn. Reson. Med.*, vol. 53, no. 4, pp. 866–876, 2005.

[33] H. Van Der Vaart, Some results on the probability distribution of the latent roots of a symmetric matrix of continuously distributed elements, and some applications to the theory of response surface estimation Inst. Statistics, Univ. North Carolina, Chapel Hill, 1958.

[34] J. D. Carew, C. G. Koay, G. Wahba, A. L. Alexander, M. E. Meyerand, and P. J. Basser, The asymptotic behavior of the nonlinear estimators of the diffusion tensor and tensor-derived quantities with implications for group analysis Dept. Statistics, Univ. Wisconsin, Madison, 2006.

[35] R. Z. Freidlin, Y. Assaf, and P. J. Basser, "Multivariate hypothesis testing of DTI data for tissue clustering," *IEEE Int. Symp. Biomed. Imag.: Macro to Nano (ISBI)*, pp. 776–779, Apr. 2007.

[36] R. Z. Freidlin, Y. Assaf, and P. J. Basser, "Multivariate hypothesis testing for tissue clustering and classification: A DTI study of excised rat spinal cord," in *Joint Annu. Meeting ISMRM-ESMRMB*, May 2007, p. 625.

[37] T. E. Conturo, R. C. McKinstry, E. Akbudak, and B. H. Robinson, "Encoding of anisotropic diffusion with tetrahedral gradients: A general mathematical diffusion formalism and experimental results," *Magn. Reson. Med.*, vol. 35, no. 3, pp. 399–412, Mar. 1996.