DEPARTMENT OF STATISTICS
University of Wisconsin
1300 University Ave.
Madison, WI 53706

TECHNICAL REPORT NO. 1132

November 2, 2006

The Asymptotic Behavior of the Nonlinear Estimators of the Diffusion Tensor and Tensor-Derived Quantities with Implications for Group Analysis

John D. Carew[1], C. Guan Koay[2], Grace Wahba[3],
Andrew L. Alexander[4], M. Elizabeth Meyerand[4],
Peter J. Basser[2]

[1]Department of Statistics, Department of Biostatistics and Medical Informatics, University of Wisconsin–Madison.

[2]Section on Tissue Biophysics and Biomimetics, Laboratory of Integrative and Medical Biophysics, National Institutes of Health.

[3]Department of Statistics, Department of Biostatistics and Medical Informatics, Department of Computer Sciences, University of Wisconsin–Madison

[4]Department of Medical Physics, University of Wisconsin–Madison.

The Asymptotic Behavior of the Nonlinear Estimators of the Diffusion Tensor and Tensor-Derived Quantities with Implications for Group Analysis

John D. Carew[*], C. Guan Koay[†], Grace Wahba[‡]
Andrew L. Alexander[§], M. Elizabeth Meyerand[§],
Peter J. Basser[†]

November 2, 2006

Key words and phrases: Diffusion tensor imaging (DTI), nonlinear least squares, maximum likelihood estimation, delta method, trace, fractional anisotropy (FA), voxel-based morphometry (VBM), region of interest (ROI), statistical parametric mapping (SPM).

Correspondence:
John D. Carew
Department of Statistics
Medical Sciences Center
1300 University Ave.
Madison, WI 53705
Tel: (608) 265-5742
Fax: (608) 265-9840
jcarew@stat.wisc.edu

---

[*]Department of Statistics, Department of Biostatistics and Medical Informatics, University of Wisconsin–Madison.

[†]Section on Tissue Biophysics and Biomimetics, Laboratory of Integrative and Medical Biophysics, National Institutes of Health.

[‡]Department of Statistics, Department of Biostatistics and Medical Informatics, Department of Computer Sciences, University of Wisconsin–Madison

[§]Department of Medical Physics, University of Wisconsin–Madison.

Abstract

Diffusion tensor imaging (DTI) is a quantitative magnetic resonance imaging (MRI) method that is used to study the microstructural properties of white matter in the brain. Tensor-derived quantities, such as the trace and fractional anisotropy (FA), are important for characterizing the normal, diseased, and developing brain. Consequently, determining the statistical properties of the diffusion tensor estimator is important to assess whether the method of estimation is appropriate and to describe how noise in diffusion-weighted images affects the variability in estimates of the tensor and functions of the tensor estimate. Here we derive asymptotic properties of the nonlinear least squares estimator (NLSE) of the diffusion tensor. We show that the NLSE is consistent and asymptotically normal. To illustrate and validate this framework we derive the asymptotic distributions of trace and FA. We show, with simulations, experimental designs that have asymptotic distributions that are very close to the empirical distributions. The methods described in this paper are applied to estimate the variances of trace and FA in a healthy human volunteer. The variance of the trace and FA are found to vary significantly throughout the brain. This renders many popular tests used in group analysis invalid for DTI data. Unequal variances for statistical tests with tensor-derived quantities is discussed.

1 Introduction

Diffusion tensor imaging (DTI) (Basser *et al.*, 1994a; Basser *et al.*, 1994b) is a quantitative magnetic resonance imaging (MRI) method that is widely used to study the microstructural properties of white matter in the brain. The diffusion tensor, which describes the diffusion properties of the imaged tissue, is known to be proportional to the covariance matrix for Brownian motion of water molecules. Since DTI can provide microstructural information, it can reveal disease-correlated tissue changes that are not evident on conventional MRI. This sensitivity to microstructure and tissue organization makes DTI an important tool for studying neurologic diseases. Several functions of the diffusion tensor (e.g., trace and fractional anisotropy (FA))

2

are important for characterizing neurologic diseases (Le Bihan *et al.*, 2001).

Determining the statistical properties of the diffusion tensor estimator is necessary to determine if the method of estimation is appropriate, and to describe how noise in diffusion-weighted images leads to variability in estimates of the tensor and functions of the tensor estimate. Estimators of the variance of the tensor estimator and functions of the tensor estimator are needed for statistical inference, comparing experimental designs, and tractography. Studying the statistical properties of the nonlinear least squares estimator (NLSE) of the diffusion tensor is particularly timely since it was recently shown that the NLSE outperforms, in terms of mean square error, linear tensor estimators (Koay *et al.*, 2006a).

In this paper we derive asymptotic properties of the NLSE of the diffusion tensor. The asymptotic distributions of the tensor estimator and functions of the tensor estimator are used to obtain estimators for their variances. We show that the NLSE of the diffusion tensor is a maximum likelihood estimator (MLE). This connection allows us to directly apply the theory of maximum likelihood estimation to obtain asymptotic properties. For diffusion tensor estimation, the asymptotic properties are achieved in the limit as the signal to noise ratio (SNR) and the number of samples of a set of diffusion directions go to infinity. Asymptotic properties are commonly used in many statistical procedures and can be very accurate well before the appropriate quantities approach infinity. We assess the utility of the asymptotic approximations with a series of simulations.

We show that the NLSE is consistent and asymptotically normal. Consistency is a desirable property of estimators, which implies that the estimates approach their true, unknown values in a probabilistic sense. Furthermore, certain functions of the tensor estimator are also consistent and asymptotically normal. To illustrate and validate the theory we derive the asymptotic distributions of trace and FA and show, with simulations, experimental designs that have asymptotic distributions that are very close to the empirical distributions. The methods introduced in this paper are applied to estimate the variance of trace and FA from a DTI data set acquired from a healthy human volunteer. Finally, we discuss implications of unequal variances for statistical

tests with tensor-derived quantities.

2 Diffusion Tensor Model

The diffusion tensor model for a signal $S := (S_1, \ldots, S_n)'$ from a single voxel with n diffusion measurements is

$$S = S_0 \exp\left(-\mathbf{X}\boldsymbol{\beta}\right) + \boldsymbol{\epsilon}. \qquad [\mathbf{1}]$$

The symbol $'$ denotes matrix transposition and $\exp()$ denotes the exponential function applied to each element of an array. The signal with no diffusion weighting is denoted by S_0. The diffusion encoding matrix $\mathbf{X}$ is given by

$$\mathbf{X} := \begin{bmatrix} b_1 g_{x1}^2 & b_1 g_{y1}^2 & b_1 g_{z1}^2 & 2b_1 g_{x1} g_{y1} & 2b_1 g_{y1} g_{z1} & 2b_1 g_{x1} g_{z1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_n g_{xn}^2 & b_n g_{yn}^2 & b_n g_{zn}^2 & 2b_n g_{xn} g_{yn} & 2b_n g_{yn} g_{zn} & 2b_n g_{xn} g_{zn} \end{bmatrix},$$

where the b_j, $j = 1, \ldots, n$, are the diffusion weightings and the g_{ij}, $i \in \{x, y, z\}$, are the components of the gradient encoding unit vectors, which specify the direction of the diffusion weighting. The parameter $\boldsymbol{\beta}$ contains (at most) 6 unique elements of the symmetric, positive-definite diffusion tensor, namely

$$\boldsymbol{\beta} := (D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{yz}, D_{xz})',$$

where the diffusion tensor $\mathbf{D}$ is

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

We assume that the errors ϵ in [1] are independent and normal with constant variance, i.e., $\epsilon \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, where $\mathbf{I}_n$ is an $n \times n$ identity matrix. Noise in magnitude MR images formed with quadrature detection follows a Rician distribution (Henkelman, 1985; Koay and Basser, 2006). As the SNR goes to infinity, the noise in magnitude images follows a normal distribution with variance equal to the variance of the normally-distributed noise in the quadrature channels (Gudbjartsson and Patz, 1995). Yet, surprisingly, when the SNR is greater than three, the noise in the SE-EPI MR images is well-approximated by a normal distribution (Gudbjartsson and Patz, 1995). With routine diffusion imaging parameters (e.g., $b \approx 1000$ s/mm^2 or less and the SNR is at least 20). Consequently, the normality and constant variance assumptions are quite sensible, and achieved at low SNR and number of images.

3 Nonlinear Least Squares Tensor Estimator

In this section, we show that the nonlinear least squares estimator (NLSE) of the diffusion tensor is the same as the maximum likelihood estimator (MLE). We start with the maximum likelihood estimator. The diffusion measurements S are viewed as a single sample drawn from a multivariate normal distribution. Under the distributional assumption $\epsilon \sim \mathcal{N}_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, the signal is distributed as $S \sim \mathcal{N}_n(S_0 \exp(-\mathbf{X}\boldsymbol{\beta}), \sigma^2 \mathbf{I}_n)$. Define the parameter to be $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_8)' := (\boldsymbol{\beta}, S_0, \sigma^2)'$. The likelihood, $L(\boldsymbol{\theta}|S)$, is given by

$$L(\boldsymbol{\theta}|S) = (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} \left[S - S_0 \exp(-\mathbf{X}\boldsymbol{\beta})\right]' \left[S - S_0 \exp(-\mathbf{X}\boldsymbol{\beta})\right] \right\}.$$

Taking the logarithm gives the loglikelihood, $l(\boldsymbol{\theta}|S)$,

$$l(\boldsymbol{\theta}|S) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left[S - S_0 \exp(-\mathbf{X}\boldsymbol{\beta})\right]' \left[S - S_0 \exp(-\mathbf{X}\boldsymbol{\beta})\right]. \qquad [2]$$

Note that the NLSE for $\boldsymbol{\beta}$ and S_0 do not depend on σ^2. In these cases, maximizing [2] with respect to $\boldsymbol{\beta}$ and S_0 is the same as minimizing the nonlinear least squares objective function,

$h(\boldsymbol{\beta}, S_0)$, given by

$$h(\boldsymbol{\beta}, S_0) := \frac{1}{2} \left[S - S_0 \exp(-\mathbf{X}\boldsymbol{\beta}) \right]' \left[S - S_0 \exp(-\mathbf{X}\boldsymbol{\beta}) \right]. \qquad [\mathbf{3}]$$

Therefore, the minimizers of [**3**], call them $\hat{\boldsymbol{\beta}}$ and $\hat{S}_0$, are the maximum likelihood estimators of the components $\boldsymbol{\beta}$ of the tensor and S_0. To get the MLE for σ^2, we substitute $\hat{\boldsymbol{\beta}}$ and $\hat{S}_0$ for $\boldsymbol{\beta}$ and S_0, respectively, in equation [**2**] and find the σ^2 that maximizes $l(\boldsymbol{\theta}|S)$. It is easy to show that the MLE for σ^2 is

$$\hat{\sigma}^2_{\mathrm{MLE}} = \frac{1}{n} \left[S - \hat{S}_0 \exp(-\mathbf{X}\hat{\boldsymbol{\beta}}) \right]' \left[S - \hat{S}_0 \exp(-\mathbf{X}\hat{\boldsymbol{\beta}}) \right] = \frac{\mathrm{RSS}}{n}. \qquad [\mathbf{4}]$$

This estimator is the residual sum of squares (RSS) divided by the number of diffusion measurements. The MLE for the error variance σ^2 is not the best estimator, particularly for small n. In the next section we will provide a better estimator that relies on the asymptotic properties of $\hat{\boldsymbol{\beta}}$ and $\hat{S}_0$.

4   Asymptotic Properties of the Tensor Estimator

The NLS estimators of $\boldsymbol{\beta}$, the components of the tensor, and S_0 are the same as the maximum likelihood estimator under the assumption of independent normal errors with constant variance. This fact allows us to use the asymptotic properties of the MLE to obtain the asymptotic properties of the NLS estimators. The two main properties of the MLE are consistency and asymptotic normality. Consistency, also called convergence in probability, is denoted by $\hat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}$. This means that for all $a > 0$, $\lim_{N,\mathrm{SNR}\to\infty} P(|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}| > a) = 0$, where N is the number of samples of the diffusion directions. In words, this means that for large enough N and SNR, the probability that the tensor estimator is more than an arbitrarily small distance away from the true value goes to zero. By theorem 5.1 on page 463 of Lehmann and Casella (1998), we have the following asymptotic properties:

1. $\hat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}$, and

2. $\hat{\boldsymbol{\theta}} \xrightarrow{d} \mathcal{N}_8(\boldsymbol{\theta}, I^{-1}(\boldsymbol{\theta}))$.

The first property states that as the number of samples of the diffusion directions and the SNR increase, the NLSE of the tensor components converges in probability to the true tensor components. The second property implies the asymptotic normality of the NLSE, i.e., $\hat{\boldsymbol{\theta}}$ converges in distribution to a multivariate normal. Convergence in distribution, denoted by $\xrightarrow{d}$, means that as N and SNR go to infinity, the distribution of the tensor estimator (at all continuity points of its distribution) is normal with mean equal to $\boldsymbol{\theta}$. Both of the asymptotic properties of the MLE hold under certain regularity conditions. Under our normality assumption, these regularity conditions are satisfied.

Note that $I^{-1}(\boldsymbol{\theta})$, the covariance matrix of the asymptotic distribution, is the inverse of the Fisher information matrix, which is proportional to the Hessian of the expected loglikelihood. Specifically, from equation 6.11 on page 125 of Lehmann and Casella (1998),

$$I_{ij}(\boldsymbol{\theta}) := -\mathrm{E}\left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} l(\boldsymbol{\theta}|S)\right], \qquad [\mathbf{5}]$$

where $\mathrm{E}[\cdot]$ denotes the expectation operator, i.e., for random variable X, $\mathrm{E}[X] := \int X \, dP$, where P is a probability measure. In the case of equation [$\mathbf{5}$], the expectation is taken with respect to the multivariate normal distribution that is parameterized by $\boldsymbol{\theta}$.

Since we are primarily interested in the properties of the components of the tensor, we can apply a simple transformation $\mathbf{C}$ to $\hat{\boldsymbol{\theta}}$ to get the asymptotic distribution of $\boldsymbol{\beta}$. Let

$$
\mathbf{C} := \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}.
$$

Then, $\mathbf{C}\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\beta}}$ and

$$
\hat{\boldsymbol{\beta}} \xrightarrow{d} \mathcal{N}_6(\boldsymbol{\beta}, \mathbf{C}I^{-1}(\boldsymbol{\theta})\mathbf{C}'). \tag{6}
$$

From equation [6], we see that $\hat{\boldsymbol{\beta}}$ is asymptotically normal with mean $\boldsymbol{\beta}$ and covariance matrix $\mathbf{C}I^{-1}(\boldsymbol{\theta})\mathbf{C}'$.

We will derive the Fisher information matrix for each component, first considering the derivatives with respect to the components of $\boldsymbol{\beta}$. It follows from the Hessian of [3], which was derived by Koay *et al.*(2006b), that

$$
\nabla_{\boldsymbol{\beta}}^2 l(\boldsymbol{\theta}|S) = -\frac{1}{\sigma^2}\nabla_{\boldsymbol{\beta}}^2 h(\boldsymbol{\beta}, S_0)
$$
$$
= -\frac{1}{\sigma^2}\mathbf{X}'(2\tilde{\mathbf{S}}^2 - \mathbf{S}\tilde{\mathbf{S}})\mathbf{X},
$$

where

$$
\mathbf{S} := \begin{bmatrix}
S_1 & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & S_n
\end{bmatrix}, \text{ and } \quad \tilde{\mathbf{S}} := S_0 \begin{bmatrix}
\exp(-\mathbf{X}_1\boldsymbol{\beta}) & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & \exp(-\mathbf{X}_n\boldsymbol{\beta})
\end{bmatrix},
$$

with $\mathbf{X}_i$ denoting the ith row of the diffusion gradient matrix. The notation $\nabla^2_{\boldsymbol{\beta}} h$ denotes a 6×6 matrix with ijth element equal to $\partial^2 h / \partial \beta_i \partial \beta_j$. Thus, the upper left 6×6 block of $I(\boldsymbol{\theta})$ is

$$\frac{1}{\sigma^2} \mathrm{E}[\mathbf{X}'(2\tilde{\mathbf{S}}^2 - \mathbf{S}\tilde{\mathbf{S}})\mathbf{X}] = \frac{1}{\sigma^2} \mathbf{X}'\tilde{\mathbf{S}}^2 \mathbf{X}.$$

In the left hand side of the previous equation, the only random quantity is $\mathbf{S}$. Since $S \sim \mathcal{N}_n(S_0 \exp(-\mathbf{X}\boldsymbol{\beta}), \sigma^2 \mathbf{I}_n)$ and $\mathrm{E}[\mathbf{S}] = \tilde{\mathbf{S}}$, the right hand side follows. For the terms involving S_0 and σ^2, the following can be shown:

$$-\mathrm{E}\left[\frac{\partial^2}{\partial S_0^2} l(\boldsymbol{\theta}|S)\right] = \frac{1}{\sigma^2} \exp(-\mathbf{X}\boldsymbol{\beta})' \exp(-\mathbf{X}\boldsymbol{\beta})$$

$$-\mathrm{E}\left[\frac{\partial^2}{\partial(\sigma^2)^2} l(\boldsymbol{\theta}|S)\right] = \frac{n}{2(\sigma^2)^2}$$

$$-\mathrm{E}\left[\frac{\partial^2}{\partial \sigma^2 \partial S_0} l(\boldsymbol{\theta}|S)\right] = 0$$

$$-\mathrm{E}\left[\frac{\partial^2}{\partial \sigma^2 \partial \beta_i} l(\boldsymbol{\theta}|S)\right] = 0$$

$$-\mathrm{E}\left[\frac{\partial^2}{\partial \beta_i \partial S_0} l(\boldsymbol{\theta}|S)\right] = \frac{S_0}{\sigma^2} \sum_{j=1}^{n} \mathbf{X}_{ji}[\exp(-\mathbf{X}_j \boldsymbol{\beta})]^2.$$

From the last equation, we see that estimates of S_0 are correlated with estimates of $\boldsymbol{\beta}$. From a practical standpoint, this means that experimental designs that yield poor estimates of S_0 will adversely affect estimates of $\boldsymbol{\beta}$. For this reason, experimental designs that include only one image with no diffusion weighting should be avoided.

The MLE for the error variance in equation [4] has a large bias for small n. This causes under estimation of the variance. The consistency and asymptotic normality of the NLSE for S_0 and $\boldsymbol{\beta}$ allow us to get a better estimator. In particular,

$$\frac{\mathrm{RSS}}{\sigma^2} \xrightarrow{d} \chi^2_{n-7},$$

where χ^2_{n-7} denotes a χ^2 distribution with $n - 7$ degrees of freedom. This suggests that the appropriate estimator of the error variance is

$$\hat{\sigma^2} := \frac{\text{RSS}}{n - 7} \qquad [7]$$

rather than the MLE given in [4].

5   Distribution of a Linear Function of the Tensor

The distribution of a linear function of the tensor (e.g., trace) is asymptotically normal. This follows from the well-know property that linear functions of normal random variables are normally-distributed. To obtain the asymptotic distribution of the trace of the tensor estimator, let $\mathbf{c} := [1\ 1\ 1\ 0\ 0\ 0]$. Then, $\widehat{\text{tr}\mathbf{D}} = \mathbf{c}\hat{\boldsymbol{\beta}}$, where tr denotes the trace of a matrix. From equation [6] it follows that

$$\widehat{\text{tr}\mathbf{D}} \xrightarrow{d} \mathcal{N}_1(\text{tr}\mathbf{D}, \mathbf{c}\mathbf{C}I^{-1}(\boldsymbol{\theta})\mathbf{C}'\mathbf{c}'). \qquad [8]$$

The general form of the normal distribution for the trace of the diffusion tensor was previously given in Basser and Pajevic (2003). The difference in our paper is that we prove asymptotic normality and show how the variance of the asymptotic distribution comes from the Fisher information matrix.

To estimate $\text{Var}(\widehat{\text{tr}\mathbf{D}})$, we use the variance in [8] evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, with σ^2 estimated according to [7], i.e.,

$$\widehat{\text{Var}}(\widehat{\text{tr}\mathbf{D}}) = \mathbf{c}\mathbf{C}I^{-1}(\hat{\boldsymbol{\theta}})\mathbf{C}'\mathbf{c}'. \qquad [9]$$

The consistency of the estimator for the variance of trace follows directly from the consistency of the estimator for $\boldsymbol{\theta}$. The variance of the mean apparent diffusion coefficient (ADC), which is defined as $1/3\ \text{tr}\mathbf{D}$, is the variance of the trace of the tensor scaled by $1/9$.

6 Distribution of a Nonlinear Function of the Tensor

In this section the multivariate delta method is used to show that a nonlinear function of the tensor estimate is asymptotically normal. In particular, we are interested in scalar functions of the tensor such as the fractional anisotropy. It is also possible that a vector-valued function of the tensor is of interest, e.g., the primary eigenvector of the tensor. To include this possibility, let

$$f(\hat{\boldsymbol{\beta}}) := (f_1(\hat{\boldsymbol{\beta}}), \ldots, f_r(\hat{\boldsymbol{\beta}}))$$

denote an r-dimensional vector-valued function of the tensor estimate. The multivariate delta method (Theorem 8.22, page 61 (Lehmann and Casella, 1998)) yields

$$\left[(f_1(\hat{\boldsymbol{\beta}}) - f_1(\boldsymbol{\beta})), \ldots, (f_r(\hat{\boldsymbol{\beta}}) - f_r(\boldsymbol{\beta})) \right] \xrightarrow{d} \mathcal{N}_r(\mathbf{0}, \mathbf{B}\mathbf{C}I^{-1}(\boldsymbol{\theta})\mathbf{C}'\mathbf{B}'), \qquad [\mathbf{10}]$$

where $\mathbf{B}_{ij} = \partial f_i / \partial \beta_j$. The matrix $\mathbf{B}$ of partial derivatives must be nonsingular in a neighborhood ω of $\boldsymbol{\beta}$. For scalar-valued f, the derivatives cannot be zero at $\boldsymbol{\beta}$. We also require that $f_1, \ldots, f_r$ are continuously differentiable in ω.

To compute the asymptotic distribution of FA, which is a normalized standard deviation of the eigenvalues of the diffusion tensor, we need to compute the partial derivatives of FA with respect to each of the six components of $\boldsymbol{\beta}$. FA takes values between 0 and 1, where 0 indicates isotropic diffusion and 1 indicates completely anisotropic diffusion. FA is written in terms of the eigenvalues λ_1, λ_2, and λ_3 of the diffusion tensor $\mathbf{D}$, as follows

$$\mathrm{FA}(\lambda_1, \lambda_2, \lambda_3) := \left[\frac{3[(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2]}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)} \right]^{\frac{1}{2}},$$

where $\bar{\lambda} := (1/3)(\lambda_1 + \lambda_2 + \lambda_3)$. We can express FA directly in terms of $\boldsymbol{\beta}$, which is more convenient for computing partial derivatives with respect to the components of $\boldsymbol{\beta}$. In this form

FA is given by

$$\mathrm{FA}(\boldsymbol{\beta}) = \left[\frac{3}{2} \left(1 - \frac{(\mathrm{tr}\mathbf{D})^2}{3\,\mathrm{tr}(\mathbf{D}^2)} \right) \right]^{\frac{1}{2}}. \tag{11}$$

Differentiating [11] gives the following:

$$\frac{\partial \mathrm{FA}}{\partial D_i} = -\frac{1}{2\mathrm{FA}} \left[\frac{\mathrm{tr}\mathbf{D}\,\mathrm{tr}(\mathbf{D}^2) - (\mathrm{tr}\mathbf{D})^2 D_i}{(\mathrm{tr}(\mathbf{D}^2))^2} \right], \quad i \in \{xx, yy, zz\}$$

$$\frac{\partial \mathrm{FA}}{\partial D_j} = \frac{(\mathrm{tr}\mathbf{D})^2 D_j}{\mathrm{FA}(\mathrm{tr}(\mathbf{D}^2))^2}, \quad j \in \{xy, yz, xz\}.$$

Let

$$\mathbf{B} := \left(\frac{\partial \mathrm{FA}}{\partial D_{xx}}, \frac{\partial \mathrm{FA}}{\partial D_{yy}}, \frac{\partial \mathrm{FA}}{\partial D_{zz}}, \frac{\partial \mathrm{FA}}{\partial D_{xy}}, \frac{\partial \mathrm{FA}}{\partial D_{yz}}, \frac{\partial \mathrm{FA}}{\partial D_{xz}} \right)$$

and apply [10] to get the asymptotic distribution of FA. To use this result to estimate the variance of an FA estimate, evaluate the Fisher information at the value of $\hat{\boldsymbol{\theta}}$, with σ^2 estimated according to [7], to get

$$\widehat{\mathrm{Var}}(\mathrm{FA}(\hat{\boldsymbol{\beta}})) = \mathbf{B}\mathbf{C}I^{-1}(\hat{\boldsymbol{\theta}})\mathbf{C}'\mathbf{B}'. \tag{12}$$

It is important to note that when FA is zero (completely isotropic diffusion), the asymptotic variance estimator in equation [12] is singular. This is due to the dependence of the derivatives of FA on the value of FA. When FA goes to zero, the factor 1/FA in the expressions for the derivatives goes to infinity.

7 Validation and Application

7.1 Simulation Study

To validate the range of utility of the asymptotic distributions of trace and FA for finite SNR and N, we simulated diffusion measurements for three cylindrically symmetric tensors with FA= 0.3578, 0.7840, and 0.9623. The trace of each simulated tensor was held constant to 2.189×10^{-3} mm^2/s. The trace and FA values are representative of brain white matter (Pierpaoli

et al., 1996). The SNR, defined as S_0/σ was set to 20 and $S_0 = 1000$. Three acquisition designs with different numbers of directions and diffusion weightings were selected to examine the performance of the asymptotic approximations for typical experimental designs. The designs are as follows.

1. Six directions at each $b = 0, 300, 650, 1000$ s/mm^2 (24 total measurements).

2. 16 directions at each $b = 0, 300, 650, 1000$ s/mm^2 (64 total measurements).

3. 46 directions at each $b = 0, 300, 650, 1000$ s/mm^2 (184 total measurements).

The six and 46 directions sets are icosahedral directions and the 16 direction set is constructed with the six icosahedral directions plus the dodecahedral directions (Hasan *et al.*, 2001). The three designs were applied to each of the three tensors for a total of 9 simulation combinations. These 9 simulations were then repeated with a trace of 1.0945×10^{-3} mm^2/s to mimic conditions of acute ischemia, where trace can be reduced by as much as 50% relative to healthy white matter (Moseley *et al.*, 1990). For each of the 18 simulations, there were 50000 simulated data sets. Data were simulated with Rician noise according to the method described by Pierpaoli and Basser (1996). The NLSE estimates of the simulated tensors were computed with Newton's method (Koay *et al.*, 2006b). Trace and FA estimates were computed for each data set from the tensor estimate. The empirical distributions of the trace and FA estimates were graphically compared to a normal distribution. Sample means of the trace and FA estimates were computed and compared to their true values to check consistency. The sample variances of the estimated trace and FA were compared to the variances of their asymptotic distributions.

To investigate the dependence of the variance of trace and FA on the values of trace and FA, we simulated data sets according to Design 2. For each point on a 15×15 grid, 50000 simulated data sets were generated for equally-spaced trace and FA values. The trace values ranged from 1.0945×10^{-3} mm^2/s to 3.2835×10^{-3} mm^2/s and the FA values ranged from 0.30 to 0.95. All other simulation parameters are as previously specified. For each trace/FA combination, the sample variances of trace and FA were computed. These variances were then plotted as a function of trace and FA.

7.2   Application to Human DTI Data

The methods for estimating the variance of FA were applied to a healthy human male volunteer, age 27, who provided informed consent in accordance with the guidelines of our Institutional Review Board for human subject studies. The images were acquired on a 3 Tesla scanner with a spin echo diffusion weighted imaging (DWI) pulse sequence. The following scan parameters were used: 10 axial slices, 3 mm thick, 5 mm gap, FOV: 240×240 mm, matrix size: 120×120 zero-padded to 256×256 pixels, 16 diffusion directions, TE: 72.3 ms, diffusion weightings $b = 0, 300, 650, 1000$ s/mm^2. The diffusion weightings and gradient direction set are identical to Design 2 of the simulation studies. The scans were cardiac gated using a pulse oximeter attached to the right index finger. The effective TR was 5 heart beats (approximately 4600 ms). No image averaging was performed, i.e., one image per gradient direction per b-value was obtained. The total scan time was 10 min 40 s. The images were first masked to eliminate pixels outside of the brain. The NLSEs of the tensors were computed at each pixel using Newton's method (Koay *et al.*, 2006b). Images of FA and the trace of the tensor estimates were generated. For FA, a second mask based on the trace of the estimated tensors was applied to eliminate pixels that contain isotropic cerebrospinal fluid (CSF). After masking, the variances of the trace and FA estimates were estimated according to [**9**] and [**12**], respectively. Images of $\log_{10}\{\widehat{\mathrm{Var}}(\widehat{\mathrm{tr}\mathbf{D}})\}$ and $\log_{10}\{\widehat{\mathrm{Var}}(\mathrm{FA}(\hat{\boldsymbol{\beta}}))\}$ were generated.

8   Results

The results from the simulation studies show that approximate normality is achieved for as few as six directions with four b-values for each gradient direction. Normal probability plots of trace and FA are given in Figures 1 and 2, respectively. These probability plots assess how well the empirical distributions of the trace and FA estimates follow normal distributions. If the empirical distribution perfectly followed a normal distribution, the estimates would fall on the red diagonal line. The plots of trace (Figure 1) show that the empirical distribution of the

trace estimates is very well-approximated by a normal distribution for Design 1 and each of the FA values. Departures from normality are minor and occur only in the distant tails of the distribution (probability < 0.001 and > 0.999). Trace estimates for the other designs, which use more images, are at least as well-approximated by a normal distribution and are, consequently, omitted. Figure 2 shows normal probability plots for each of the three FA values for Designs 1, 2, and 3. The plots show that the empirical distribution of the FA estimates is well-approximated by a normal distribution. Convergence to normality is slower for very high FA (0.9623). The largest departures from normality are seen in Design 1 with FA=0.9623. These departures occur in the tails of the distribution (probability < 0.05 and > 0.95). In this case, we see slightly slower convergence to normality than in other designs and lower values of FA. Moving to Design 2, the discrepancies for FA=0.9623 are in the distant tails. For FA at 0.3578 and 0.7840, the departures from normality are in the distant tails for all designs. The reduced range of the FA estimates when moving from Design 1 to other designs indicates the reduced variability that comes from acquiring more directions.

The asymptotic variances of trace and FA are very accurate. Table 1 summarizes the simulation results when the trace mimics that of healthy white matter. The sample means of the trace FA estimates are very close to the true value in all of the 18 simulations. The sample means of the trace estimates are all very close to the true value (2.189×10^{-3} mm^2/s) with discrepancies only in the third significant digit. The largest difference between the true FA and the sample means is found in Design 1 with low FA (6.7% over-estimation). Other differences are significantly lower. The closeness of the sample means to the true values of trace and FA are expected from the consistency of the NLSE. The sample variance of the trace and FA estimates, which are good estimates of the true variances, shows that the asymptotic variances are excellent approximations. In all 18 designs, the sample variance of trace differs from the asymptotic variance by 1.61% or less. For low FA (0.3578) the asymptotic variance tends to over-estimate the true FA variance. This bias is highest for Design 1 with FA $= 0.3578$ where the asymptotic variance is 23.8% higher. Moving to Design 2, the over-estimation is reduced to 5.66%. The

over-estimation is not considerable and will not lead to increased type I error rates. For very high FA (0.9623) there is a slight tendency for the asymptotic variance to under-estimate the true variance. This effect is minor with the largest under-estimation at less than 2.36%.

Also seen in Table 1 is that the variances of trace and FA depend of the value of FA. Variance in estimates of trace from tensors with low FA have lower variance than estimates from tensors with higher FA. We found that trace estimates from higher FA tensors are 30% or less variable than trace estimates from lower FA tensors. In the variance of estimates of FA, the opposite effect is seen and is much more dramatic. Estimates of FA from tensors with relatively high FA have lower variance than estimates from tensors with lower FA. In our simulation studies, the variances of FA estimates from tensors with a true FA of 0.3578 are approximately one order of magnitude higher than tensors with a true FA of 0.9623.

The simulation studies for the case of acute ischemia reveal similar behavior as in the case for healthy white matter (Table 2). For the variance of trace, the asymptotic variance is, on average, slightly more accurate than where trace is higher. All of the asymptotic variances have an error of less than 1.36%. For FA, the asymptotic variances are less accurate. Within the same design, tensors with the lowest FA (0.3578) have the highest errors, with the largest error at 39.7% for Design 1. This overestimation is reduced to 13.2% and 4.32% for Designs 2 and 3, respectively.

Plots of the variance of trace and FA as a function of trace and FA (Figure 3) show graphically how the variance of these tensor-derived quantities depends on their values. These plots reiterate the findings from the simulation studies with three values of FA and two values of trace.

The results from the human data illustrate the anatomical dependence of the variances of trace and FA estimates. The SNR in the diffusion weighted images is approximately 22. The trace and FA maps for four slices are in Figures 4 and 5, respectively. The corresponding variance maps are displayed using a $\log_{10}$ scale since they vary over more than one order of magnitude. The variance of trace estimates depends on the tissue type. Comparing the FA map in Figure 5 to the trace variance image in Figure 4, we see that a region of high FA is where the variance
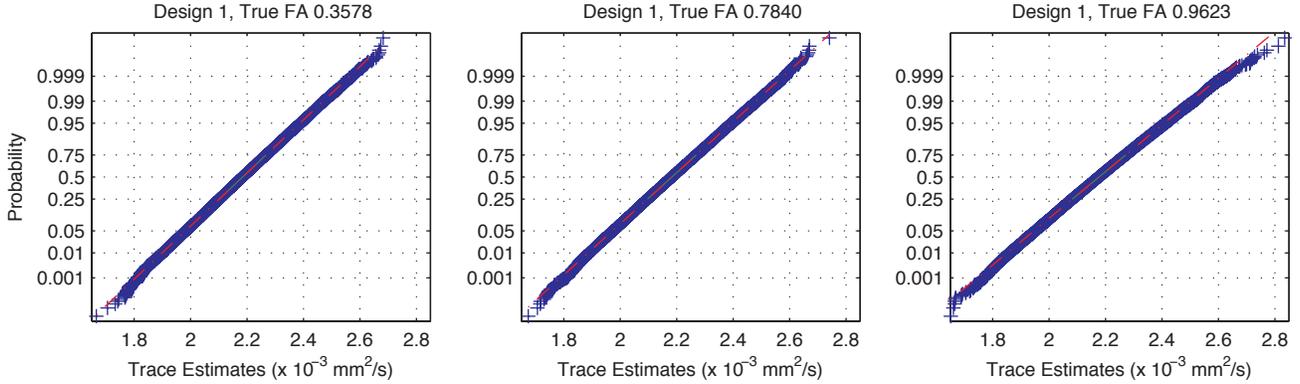
Figure 1: Each plot shows the trace estimates for 50,000 simulated data sets with Design 1 (six directions at each of four *b*-values). The trace estimates are very well-approximated by a normal distribution. Departures from normality are in the distant tails of the distribution ($P < 0.001$, $0.999 < P$).

of trace is also relatively high. The variability of FA estimates is not uniform over the whole brain and, like trace, depends on the tissue type. Lowest variances are seen in regions of high anisotropy (e.g., corpus callosum, internal capsule). The opposite effect is seen with trace and is consistent with the simulation studies.

9   Discussion

One important result of this study is that the NLSE of the diffusion tensor and tensor-derived quantities are consistent and asymptotically normal. The asymptotic distributions are reached as the SNR and number of DWIs increases. We described consistent estimators for the variances of trace and FA. The simulation results show that the asymptotic approximations for trace and FA are accurate. This suggests the utility of the asymptotic variance as the measure of variability in tensor estimates. The asymptotic approximations for trace, which is a linear function of the tensor, perform slightly better than the approximations for FA. This is expected since the use of the delta method for FA adds one additional layer of approximation. The reduction in the variance of trace and FA as the number of unique gradient encoding directions increases has been studied by Jones (2004) with Monte Carlo simulations. Our results on consistency of the
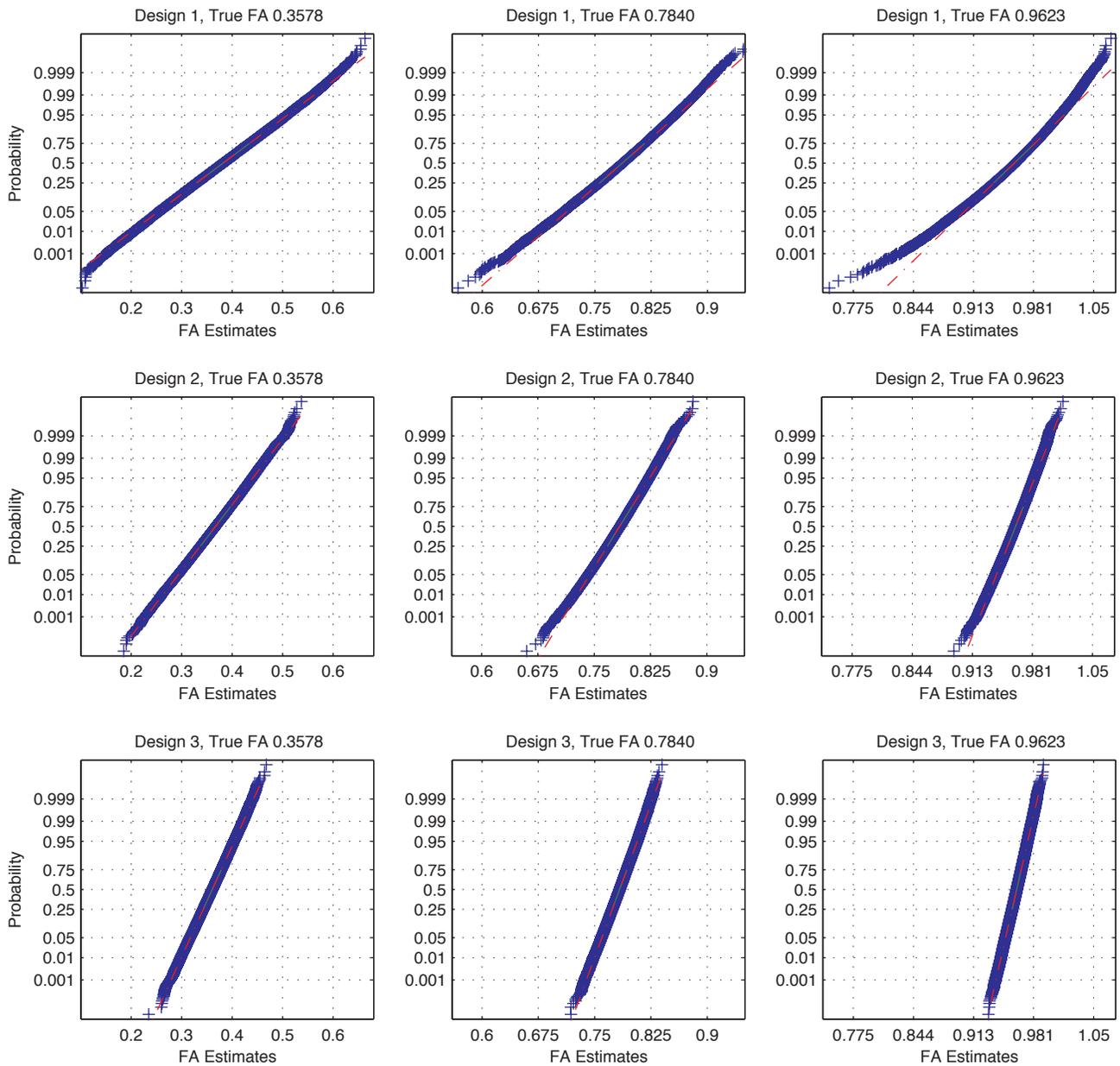
17

Figure 2: Each plot shows the FA estimates for 50,000 simulated data sets. Departures from normality are minor, with the largest occurring in for Design 1 (six directions at each of four b-values) with high FA (0.9623). The discrepancies here occur in the tails of the distribution ($P < 0.05$, $0.95 < P$). Discrepancies for other designs and FA values occur in the far tails of the distribution. Also evident from the decreased range of the FA estimates when moving from Design 1 to Designs 3 and 5 is that the variance of the estimates decreases.

Table 1: Simulation Results When Trace $= 2.189 \times 10^{-3}$ mm^2/s (Parenthetical Quantities are Percent Error)

Design	True FA	FA Sample Mean	FA Sample Variance	FA Asymptotic Variance	Trace Sample Mean	Trace Sample Variance	Trace Asymptotic Variance
1	0.3578	0.3819	5.819×10^{-3}	7.201×10^{-3} (23.8)	2.185×10^{-3}	1.510×10^{-8}	1.513×10^{-8} (0.199)
1	0.7840	0.7854	2.026×10^{-3}	2.109×10^{-3} (4.10)	2.183×10^{-3}	1.626×10^{-8}	1.651×10^{-8} (1.54)
1	0.9623	0.9578	1.231×10^{-3}	1.202×10^{-3} (-2.36)	2.178×10^{-3}	1.953×10^{-8}	1.984×10^{-8} (1.59)
2	0.3578	0.3667	1.980×10^{-3}	2.092×10^{-3} (5.66)	2.180×10^{-3}	5.652×10^{-9}	5.667×10^{-9} (0.265)
2	0.7840	0.7837	6.008×10^{-4}	5.993×10^{-4} (-0.250)	2.178×10^{-3}	6.060×10^{-9}	6.115×10^{-9} (0.908)
2	0.9623	0.9617	1.778×10^{-4}	1.741×10^{-4} (-2.08)	2.174×10^{-3}	6.669×10^{-9}	6.767×10^{-9} (1.47)
3	0.3578	0.3603	7.007×10^{-4}	7.195×10^{-4} (2.68)	2.179×10^{-3}	1.948×10^{-9}	1.971×10^{-9} (1.18)
3	0.7840	0.7830	2.075×10^{-4}	2.057×10^{-4} (-0.867)	2.177×10^{-3}	2.119×10^{-9}	2.127×10^{-9} (0.378)
3	0.9623	0.9618	5.887×10^{-5}	5.810×10^{-5} (-1.31)	2.173×10^{-3}	2.300×10^{-9}	2.337×10^{-9} (1.61)

Table 2: Simulation Results When Trace = 1.0945×10^{-3} mm^2/s (Parenthetical Quantities are Percent Error)

| Design	True FA	FA Sample Mean	FA Sample Variance	FA Asymptotic Variance	Trace Sample Mean	Trace Sample Variance	Trace Asymptotic Variance
1	0.3578	0.4149	9.9154×10^{-3}	1.3856×10^{-2} (39.7)	1.0946×10^{-3}	1.0035×10^{-8}	9.8990×10^{-9} (-1.36)
1	0.7840	0.7915	4.2829×10^{-3}	4.5564×10^{-3} (6.39)	1.0944×10^{-3}	1.0006×10^{-8}	1.0103×10^{-8} (0.969)
1	0.9623	0.9608	2.1175×10^{-3}	2.1176×10^{-3} (0.00047)	1.0949×10^{-3}	1.0436×10^{-8}	1.0522×10^{-8} (0.824)
2	0.3578	0.3803	4.0397×10^{-3}	4.5717×10^{-3} (13.2)	1.0920×10^{-3}	3.6979×10^{-9}	3.7114×10^{-9} (0.365)
2	0.7840	0.7866	1.5569×10^{-3}	1.5775×10^{-3} (1.31)	1.0920×10^{-3}	3.7720×10^{-9}	3.7845×10^{-9} (0.331)
2	0.9623	0.9626	5.6950×10^{-4}	5.6113×10^{-4} (-1.47)	1.0909×10^{-3}	3.9010×10^{-9}	3.8971×10^{-9} (-0.01)
3	0.3578	0.3657	1.5156×10^{-3}	1.5810×10^{-3} (4.32)	1.0913×10^{-3}	1.2939×10^{-9}	1.2910×10^{-9} (-0.22)
3	0.7840	0.7847	5.4062×10^{-4}	5.4435×10^{-4} (0.67)	1.0980×10^{-3}	1.3048×10^{-9}	1.3164×10^{-9} (0.89)
3	0.9623	0.9621	1.9397×10^{-4}	1.9254×10^{-4} (-0.74)	1.0905×10^{-3}	1.3402×10^{-9}	1.3547×10^{-9} (1.08)
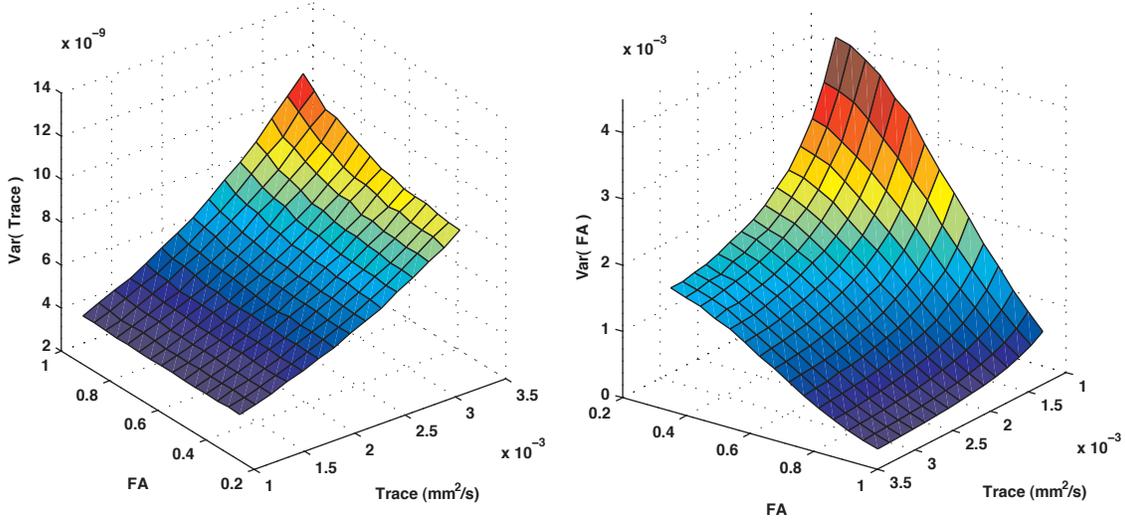
Figure 3: The variances of trace and FA depend on the values of trace and FA. These plots show sample variances of trace and FA for tensors various trace/FA combinations. Each point on the plots is the sample variance of the respective quantity, computed from 50000 data sets simulated with Design 2.

estimators provide a theoretical explanation for his findings.

Unlike bootstrap approaches (Pajevic and Basser, 2003) that are used to obtain estimates of the variance of FA estimates, our method does not require additional diffusion weighted images or intense computation. Since the asymptotic approximations to the variances of NLSEs of trace and FA are accurate, use of the bootstrap may be unnecessary for data collected under routine scan parameters, provided that the systematic artifacts in the DWIs can be shown to be small. However, it is best to first simulate data under a particular design to evaluate the usefulness of the asymptotic approximations.

It is important to note that while the noise in DWIs may be relatively constant throughout the brain, the noise in estimates of the tensor and tensor-derived quantities is not. Our simulation studies show that regions with high FA have less variable FA estimates than regions with low FA. The opposite effect is seen with the trace. When FA is high, the estimates of trace have relatively higher variance. These results are consistent with the findings of Pierpaoli and Basser (1996). The dependence of the tensor-derived quantities on their respective quantities is best
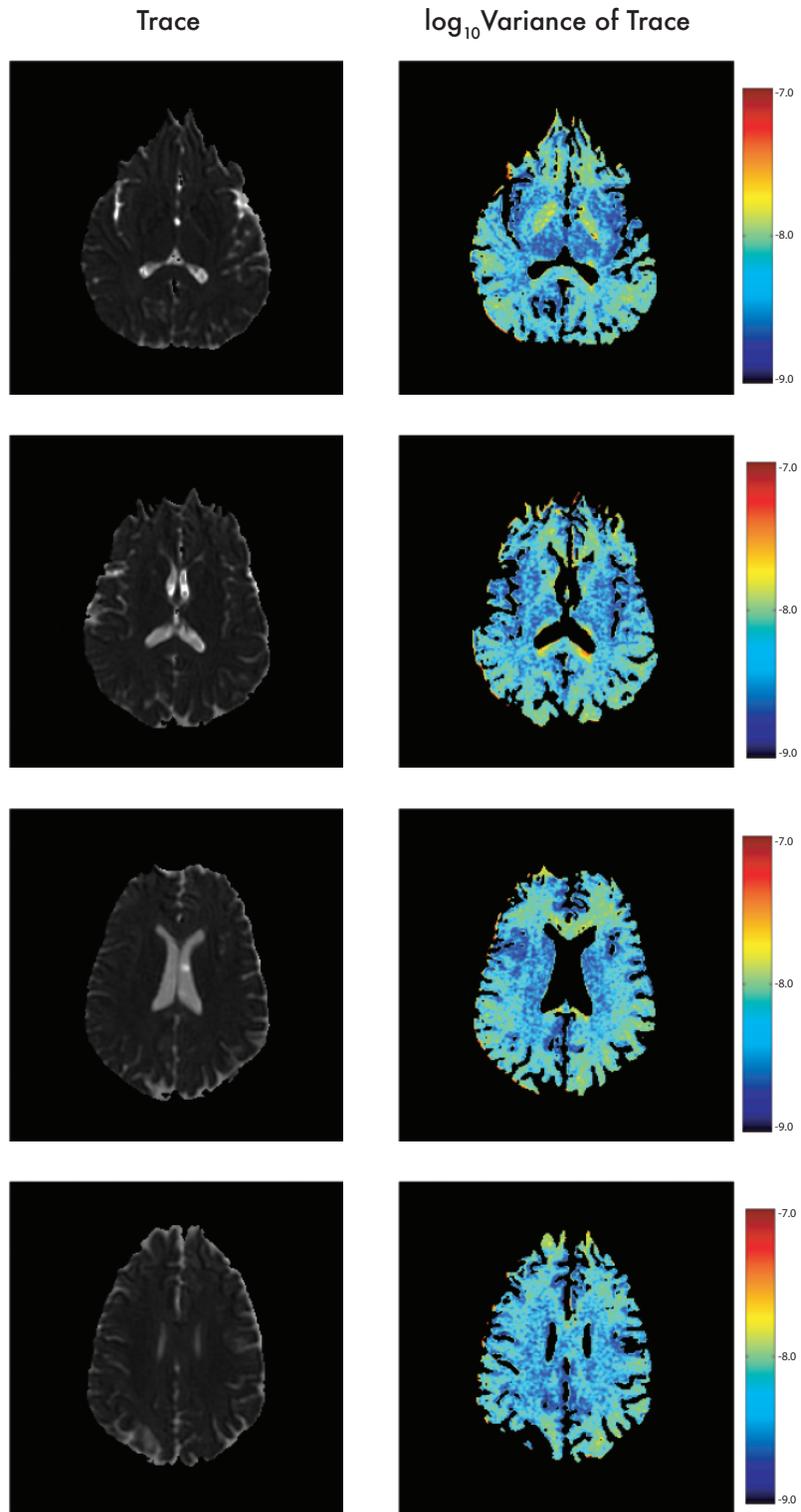
Figure 4: The variances of the trace estimates vary throughout the brain. In some of the highly anisotropic regions (e.g., corpus callosum), trace estimates have higher variance than in more isotropic regions. This is consistent with the simulation results.
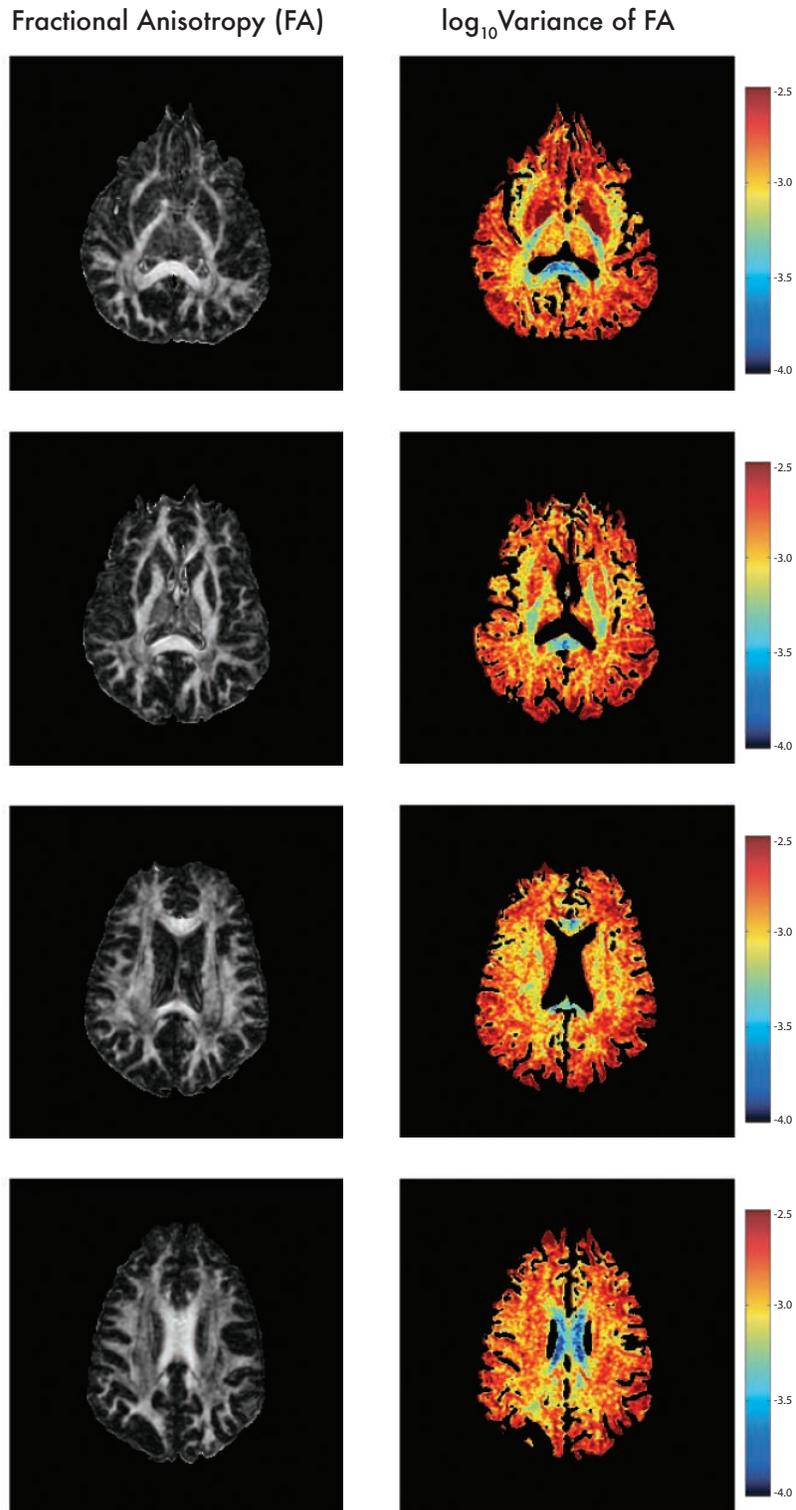
Figure 5: The variance of the FA estimates also vary throughout the brain. Regions where FA is relatively high (e.g., the corpus callosum) have relatively low variance. Within white matter, the FA variances vary by approximately one order of magnitude.

23

illustrated in Figure 3.

The dependence of the variances of trace and FA on the values of trace and FA has important implications for testing for differences in the means of these tensor-derived quantities. The assumption of equal variance that is common to many standard statistical tests used in group analyses could be potentially violated. Suppose that one uses a t-test to test the null hypothesis that the mean of FA in one group is the same as the mean FA in another group. This test is commonly used for both voxel-based morphometry (VBM) and ROI-based group analyses. Jones *et al.* (2005) cites several VBM studies of tensor-derived quantities in the context of examining the effect of the filter size. When the null hypothesis is true, the variances in each group are equal. When the alternative hypothesis, that the means are unequal, is true, then the variances of the FA measurements will differ in the two groups. This aspect of trace and FA makes estimating the null distribution, which is necessary for comparing the test statistic, impossible. The effect of the variance heterogeneity under the alternative hypothesis may be substantial. Variances of trace and FA differ considerably over white matter. For example, the variance of FA varies by roughly one order of magnitude (Figure 5). More work is needed to evaluate the impact of unequal variances on the inferences that are made with simple testing procedures.

Combining estimates of the variances of trace or FA measurements with a weighted linear model may provide one solution to the problem of unequal variances. Take, for example, testing the mean of FA between two groups. We have a set of n_1 estimates from group 1, $\text{FA}_{11}, \ldots, \text{FA}_{1n_1}$, and a set of n_2 estimates from group 2, $\text{FA}_{21}, \ldots, \text{FA}_{2n_2}$. Corresponding to these FA measurements are estimates of their variances, denoted $\xi_{11}, \ldots, \xi_{1n_1}, \xi_{21}, \ldots, \xi_{2n_2}$. These variance estimates can be computed with our estimator in equation [**12**] or other methods such as the bootstrap. Then choose the weights $w_1, \ldots, w_{n_1+n_2}$ as $1/\xi_{11}, \ldots, 1/\xi_{1n_1}, 1/\xi_{21}, \ldots, 1/\xi_{2n_2}$, respectively. Define $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)'$ as the model parameter, where α_0 is the mean FA in group 1

and α_1 is the difference in the means between groups 1 and 2. The parameter $\boldsymbol{\alpha}$ is estimated as

$$\hat{\boldsymbol{\alpha}} := \arg \min_{\boldsymbol{\alpha}} \sum_{i=1}^{n_1+n_2} w_i (\text{FA}_i - \alpha_0 - y_i \alpha_1)^2,$$

where y_i has a value of 1 when the ith observation is from group 2 and otherwise has a value of 0. Then, testing the null hypothesis that the mean FA of group 1 is equal to the mean FA of group 2 is equivalent to testing that $\alpha_1 = 0$. Provided that the variance estimates are good, the residuals from the linear model fit should have zero mean and approximately constant variance. This allows valid testing of the null hypothesis, that $\alpha_1 = 0$, with a t-test. This type of weighted least squares estimation and testing can be performed in many standard statistical packages. More general linear models of FA or other tensor-derived quantities that are fit with weighted least squares are possible. These models may include other factors or covariates.

One limitation of our method is that it assumes the errors in the diffusion weighted images are independent, normal, and have constant variance. The independence assumption is satisfied by the independent acquisition of the DWIs themselves. Ideally, for high SNR, the normality and constant variance assumptions are satisfied. At high values of b ($\gg 1000$ s/mm^2), which cause lower SNR, these assumptions can break down. For most routine DTI scans, the SNR is at least 20 and, consequently, normality is a good assumption. Head motion, eddy current distortion, susceptibility effects, and other systematic artifacts, as well as subtle physiologic noise may also cause violations in these assumptions.

Rician noise has the unique property that the variance of noise in magnitude MR images depends on the mean of the signal even when the noise in the quadrature channels is constant. This is a source of non-constant variance in diffusion-weighted images since directions along the major axis of axons will have lower mean signal than from directions perpendicular to the major axis. The effect of non-constant variance is only substantial for very low SNR (Koay and Basser, 2006) and is not a significant issue for the imaging parameters considered in our study. Nevertheless, our methods can be easily extended to situations where non-constant variance is

an issue. The noise covariance matrix, $\sigma^2 \mathbf{I}_n$, needs to be replaced with a matrix of the form $\sigma^2 \mathbf{w} \mathbf{I}_n$, where $\mathbf{w}$ is a vector of weights, which depend on the SNR of each measurement. Koay and Basser (2006) derived an analytical expression for the weights. The tensor is then estimated iteratively by weighted nonlinear least squares.

Our method can be extended to other functions of the diffusion tensor and to study design of DTI experiments. Other functions of the tensor that are commonly used include the direction of the primary eigenvector of the tensor and relative anisotropy. Methods for testing for group differences in the direction of the primary eigenvector have been described (Schwartzman *et al.*, 2005). The issue of the affect of experimental design on the variance of tensor-derived quantities has been studied with Monte Carlo methods (Papadakis *et al.*, 1999; Skare *et al.*, 2000; Jones, 2004). The asymptotic variances can also be used as a benchmark for comparing designs of DTI experiments.

Finally, Salvador *et al.* (2005) proposed an estimator for the variance of mean ADC based on the linearized tensor model. They did not consider an estimator for the variance of FA. However, by applying the delta method, an estimator for the variance of FA can also be obtained from the linearized tensor model. By substituting the inverse of the Fisher information matrix in equation [**12**] with the estimator for the variance of the tensor from Salvador *et al.* (2005), one obtains an estimator for the variance of FA from the linearized tensor model. It is unknown how the performance of such an estimator compares to our estimator in equation [**12**]. Since the NLSE outperforms the linearized tensor estimator (Koay *et al.*, 2006), we suspect that the FA variance estimator from the NLSE will outperform the FA variance estimator for the linearized model. This requires closer examination.

10 Acknowledgments

11   References

1. Basser PJ, Mattiello J, Lebihan D (1994a) MR diffusion tensor spectroscopy and imaging. *Biophys. J.* **66**, 259-267.

2. Basser PJ, Mattiello J, Lebihan D (1994b) Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. B **103**, 247-254.

3. Basser PJ, Pajevic S (2003) Dealing with uncertainty in diffusion tensor MR data. *Israel J. Chem.* **43(1-2)**, 129-144.

4. Gudbjartsson H, Patz S (1995) The Rician distribution of noisy MRI data. *Magn. Reson. Med.* **16**, 87-90.

5. Henkelman RM (1985) Measurement of signal intensities in the presence of noise in MR images. *Med. Phys.* **12**, 232-233. Erratum in (1986) **13**, 544.

6. Jones DK (2004) The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a monte carlo study. *Magn. Reson. Med.* **51**, 807-815.

7. Jones DK, Symms MR, Cercignani, Howard RJ (2005) The effect of filter size on VBM analyses of DT-MRI data. *NeuroImage* **26**, 546-554.

8. Koay CG, Basser PJ (2006) Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. *J. Magn. Reson.* **179**, 477-482

9. Koay CG, Carew JD, Alexander AL, Basser PJ, Meyerand ME (2006a) Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging. *Magn. Reson. Med.* **55**, 930-936.

10. Koay CG, Chang L-C, Carew JD, Pierpaoli C, Basser PJ (2006b) A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging. *J. Magn. Reson.* **182**, 115-125.

11. Le Bihan D, Mangin J-F, Poupon C, Clark CA, Pappata S, Molko N, Chabriat H (2001) Diffusion tensor imaging: concepts and applications. *J. Magn. Reson. Imag.* **13**, 534-546.

12. Lehmann EL, Casella G (1998) *Theory of Point Estimation, 2nd ed.* (Springer, New York).

13. Moseley ME, Kucharczyk J, Mintorovitch J, et al. (1990) Diffusion-weighted MR imaging of acute stroke: correlation with T2-weighted and magnetic susceptibility-enhanced MR imaging in cats. *AJNR* **11**, 423-429.

14. Papadakis NG, Xing D, Houston GC, *et al.*(1999) The study of rotationally invariant and symmetric indices of diffusion anisotropy. *Magn. Reson. Imag.* **17**, 881-892.

15. Pajevic S, Basser PJ (2003) Parametric and non-parametric statistical analysis of DT-MRI data. *J. Magn. Reson.* **161**, 1-14.

16. Pierpaoli C, Basser PJ (1996) Toward a quantitative assessment of diffusion anisotropy. *Magn. Reson. Med.* **36**, 893-906.

17. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Chiro G (1996) Diffusion tensor MR imaging of the human brain. *Radiology* **201**, 637-648.

18. Salvador R, Pena A, Menon DK, Carpenter TA, Pickard JD, Bullmore ET (2005) Formal characterization and extension of the linearized diffusion tensor model. *Hum. Brain Mapp.* **24**, 144-155.

19. Schwartzman A, Dougherty RF, Taylor JE (2005) Cross-subject comparison of principal diffusion direction maps. *Magn. Reson. Med.* **53**, 1423-1431.

20. Skare S, Hedehus M, Moseley ME, Li TQ (2000) Condition number as a measure of noise performance or diffusion tensor data acquisition schemes with MRI. *J. Magn. Reson.* **147**, 340-352.