Quantification of Measurement Error in DTI: Theoretical Predictions and Validation*

Casey Goodlett[1], P. Thomas Fletcher[2], Weili Lin[3], and Guido Gerig[1,4]

[1] Department of Computer Science, University of North Carolina
[2] School of Computing, University of Utah
[3] Department of Radiology, University of North Carolina
[4] Department of Psychiatry, University of North Carolina

Abstract. The presence of Rician noise in magnetic resonance imaging (MRI) introduces systematic errors in diffusion tensor imaging (DTI) measurements. This paper evaluates gradient direction schemes and tensor estimation routines to determine how to achieve the maximum accuracy and precision of tensor derived measures for a fixed amount of scan time. We present Monte Carlo simulations that quantify the effect of noise on diffusion measurements and validate these simulation results against appropriate in-vivo images. The predicted values of the systematic and random error caused by imaging noise are essential both for interpreting the results of statistical analysis and for selecting optimal imaging protocols given scan time limitations.

1 Introduction

Diffusion tensor MRI (DT-MRI) has become a critical technology for studying white matter in-vivo. In clinical studies, derived tensor measures such as anisotropy measures or mean diffusivity are commonly used for voxel-wise and region-based statistical analysis [1]. In order to understand the significance of statistical analysis, the precision and accuracy of the measurements must be well understood. DT-MRI is particularly sensitive to error introduced by imaging noise for two reasons. First, since multiple diffusion-weighted images are needed, each individual image must be acquired relatively quickly, reducing the signal-to-noise ratio (SNR) for each image. Secondly, unlike structural MRI where intensities are primarily used to establish contrast between tissue types, DT-MRI measures quantitative physical properties requiring a more careful evaluation of noise. This paper both simulates the influence of Rician noise on tensor derived measures and evaluates the simulation against in-vivo experiments.

Several studies have investigated the effects of noise on tensor measurements through theory and Monte Carlo simulation [2,3,4]. Jones and Basser showed how

Rician noise tends to underestimate high values of the apparent diffusion coefficient (ADC) along gradient directions [5]. Basu, Fletcher, and Whitaker used the Rician noise model to demonstrate statistical bias and develop a regularization filter for diffusion-weighted images [6]. Fillard et al. combined a maximum likelihood (ML) tensor estimator with a regularization function to jointly smooth and estimate a tensor field [7]. This paper builds on previous work by combining a comparison of gradient direction schemes with tensor estimation methods to study the error in diffusion tensors given noisy image acquisition in a clinical framework. In this paper we work within the assumption of a single diffusion tensor model per voxel and do not consider high angular resolution diffusion imaging (HARDI). Our simulations show that increasing the number of gradient directions reduces the bias introduced by the orientation of the tensor. Furthermore, we demonstrate the increased variability caused by linear least squares estimation on sequences with many gradient directions. We also present a novel validation of the simulation with experiments using in-vivo data.

2   Methods

The estimated diffusion tensor can be understood as a function of observed diffusion-weighted MR intensities via the Stejskal-Tanner equation

$$S_i = S_0 \exp\left(-b\mathbf{g}_i \mathbf{D} \mathbf{g}_i^T\right). \tag{1}$$

Diffusion weighted images are acquired by computing the magnitude of the Fourier transform of a measured k-space signal. Pure Johnson noise in both the real and and imaginary components of k-space is well-approximated by a Gaussian distribution, and noise in the magnitude signal S_i is well characterized by a Rician distribution. A noisy measurement R of an underlying signal A in the diffusion-weighted image is a random variable given by

$$R = \sqrt{(A+X)^2 + Y^2}, \quad X, Y \sim N(0, \sigma^2), \tag{2}$$

where X and Y are Gaussian random variables. The probability density function for a Rician random variable R with true intensity A and noise variance σ^2 is

$$f(x|A, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + A^2}{2\sigma^2}\right) I_0\left(\frac{xA}{\sigma^2}\right), \tag{3}$$

where I_0 is the zero-order modified Bessel function of the first kind. The Rician distribution is equivalent to the Rayleigh distribution when $A = 0$, and converges to a Gaussian distribution as $\sigma/A \to 0$. For low signals the expected value of the measurement is greater than the true signal, $\mathrm{E}[R] > A$.

The apparent diffusion coefficient (ADC) in a direction $\mathbf{g}_i$ is measured in a voxel by the ratio of a baseline signal S_0 and an attenuated diffusion-weighted signal S_i. Because Rician random variables with low A have a positive bias, we are likely to observe a higher signal than the true intensity. Overestimation of an intensity S_i causes an underestimation of diffusion in the direction g_i because of

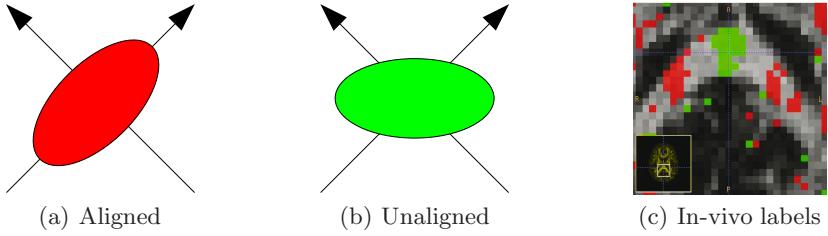(a) Aligned   (b) Unaligned   (c) In-vivo labels

Fig. 1. Example of tensors with principal eigenvector aligned or unaligned with a gradient direction. Figure 1(c) shows the computed classes in in-vivo experiments.

the exponential decay in the Stejskal-Tanner equation (1). Measurements of low diffusion have lower attenuation and correspondingly less bias. The tendency to underestimate high ADC values causes two major challenges for reliable measurements in DTI. First, mean diffusivity is likely to be underestimated for regions with high diffusion. This is a particular problem in the cerebrospinal fluid (CSF), where the ADC is high in all directions. Secondly, anisotropy can be underestimated depending on the alignment of the principal diffusion direction of a highly anisotropic tensor with the gradient directions. As shown in Figure 1, the maximum measured ADC of a highly anisotropic tensor depends on the gradient direction sampling, and the bias in the measured FA value depends on the orientation of the tensor.

A minimum of six gradient images plus a baseline image are required to estimate the parameters of the diffusion tensor. Since diffusion-weighted images typically have low SNR, acquiring more images than the minimum seven is desirable to improve SNR and obtain more robust measures of diffusivity and anisotropy. To improve SNR, repetitions of gradient directions can be acquired and are typically processed by averaging of repetitions on-scanner or offline where corresponding images are registered and averaged. Alternatively, additional diffusion-weighted images can be acquired using more gradient directions, and the additional observations are combined in the tensor estimation. In images with multiple repetitions of the same gradient direction, the signals are typically averaged. However, the mean of the signals is a poor estimator of the true signal A, because of the bias in the Rician distribution. Averaging of the intensities across repetitions consequently tends to overestimate the signal and underestimate the ADC in voxels with low SNR.

For gradient schemes with more than the minimal number of gradient directions, several methods exist in the literature for estimating diffusion tensors from the diffusion-weighted images. The most common approach has been a linear least squares (LLS) estimator for the tensor parameters $\boldsymbol{\beta}$ from the log of the observed signal intensities $\mathbf{S}$ with baseline signal S_0. The matrix $\mathbf{X}$, which is Nx6, encapsulates the b-value and gradient directions.

$$\hat{\boldsymbol{\beta}}_{\text{lls}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\ln\mathbf{S} - \ln S_0). \tag{4}$$

To avoid the re-weighting penalties associated with the log, non-linear least squares (NLS) estimation optimizes the objective function

$$f_{\text{nls}}(\boldsymbol{\beta}) = ||\mathbf{S} - S_0 \exp{(\mathbf{X}\boldsymbol{\beta})}||^2 \tag{5}$$

on the diffusion-weighted signal. Our implementation uses Levenberg-Marquadt optimization. Salvador et al. proposed a weighted least squares (WLS) estimator

$$\hat{\boldsymbol{\beta}}_{\text{wls}} = (\mathbf{X}^T\mathbf{W}^2\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{W}^2(\ln{\mathbf{S}} - \ln{S_0})) \tag{6}$$

$$\mathbf{W} = \text{Diag}(\mathbf{X}\hat{\boldsymbol{\beta}}_{\text{lls}}) \tag{7}$$

based on the log Rician probability distribution [8]. Our implementation of this method uses iterative weight refinement until the tensor values converge.

An ML estimate of the diffusion tensor using the log-likelihood function of the Rician distribution was proposed by Fillard et al. [7]. We use a similar method that does not use a spatial regularization term or the log of the tensor matrix. The ML method explicitly accounts for the noise model and uses an estimate of the noise level computed from the background of the image. The ML estimation of tensor parameters is obtained by numerical optimization of the log-likelihood function

$$\log L(\boldsymbol{\beta}) = \sum_i \log\left(\frac{S_i}{\sigma^2}\right) - \frac{S_i^2 + S_0^2 e^{2X_i\boldsymbol{\beta}}}{2\sigma^2} + \log\left(I_0\left(\frac{S_i S_0^2 e^{2X_i\boldsymbol{\beta}}}{\sigma^2}\right)\right), \tag{8}$$

where X_i is a row of the matrix $\mathbf{X}$ and N is the number of gradient directions. Our implementation uses a gradient descent optimizer to maximize the objective function. We evaluate the different tensor estimation methods across several different gradient sampling schemes using a Monte Carlo framework for simulating the effect of imaging noise on derived properties.

3   Simulation

Our Monte Carlo simulations compute the distribution of estimated tensors from the predicted signal of a given true tensor with added Rician noise. The true tensor has a fixed trace of $2.1x10^{-3}mm^2/s$, which is a typical value for white matter. Several levels of anisotropy and orientation are simulated. The simulations use a b-value of $1000\ s/mm^2$, a noise level of $\sigma = 27$ estimated from the background of the image, and a baseline signal of 250.

The simulations show that positive bias of Rician noise at low signal level can lead to underestimation of FA and trace. Furthermore, the orientation of the tensor within the gradient fields correlates with the bias, and statistical comparison of structures with different fiber orientations is potentially biased. Many common clinical gradient schemes use the minimal six gradient directions and with these schemes the expected bias depends on the orientation of the fiber structure within the magnetic field. Figure 2 shows the bias and variability of the FA and trace of a fixed diffusion tensor as the tensor rotates in space relative
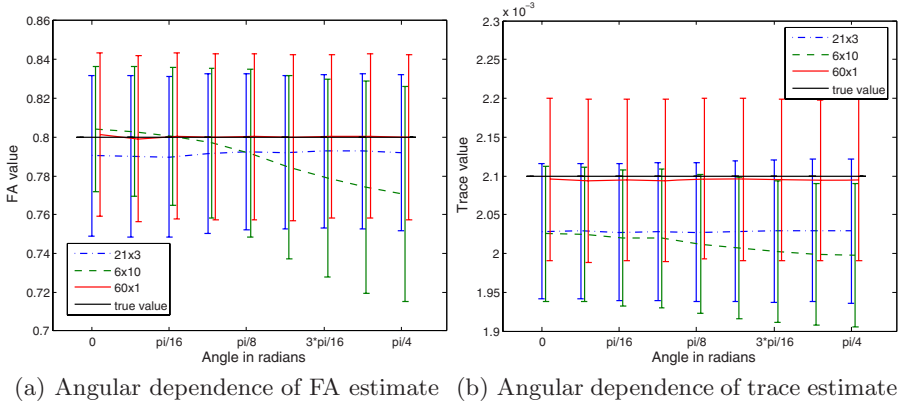
(a) Angular dependence of FA estimate   (b) Angular dependence of trace estimate

Fig. 2. The orientation dependence of estimated FA and trace under different sampling schemes. The horizontal axis ranges from the principal diffusion direction unaligned with any gradient direction for $x = 0$ to being perfectly aligned with one of the gradient directions for a rotation of $x = \pi/4$. The noise level is $\sigma = 27$ as determined from our collected data. Notice in the 6 direction scan the difference in mean estimated FA of .04 (5%) between the same tensor orientated at 0 and $\pi/4$ radians. Weighted least squares estimation was used for this simulation.

to a gradient direction. The simulations show that FA is substantially correlated with orientation. The trace estimate has less bias due to orientation, but the trace is underestimated in gradient schemes with repeated directions. Figure 3 shows the simulated distribution of FA, trace, and the Frobenius norm of the difference between the estimated and true tensor using a 60 direction protocol. Weighted least squares and maximum likelihood perform similarly, while linear least squares has more variability, non-linear least squares tends to have a lower estimate of trace.

The simulations demonstrate the bias due to orientation of tensor derived measures when using protocols with a minimal number of gradient directions. Furthermore, when protocols with many gradient directions are used linear least squares estimation can increase variability, and non-linear least squares estimation can underestimate trace. The simulations predict that the minimum error is introduced by using as many isotropic non-repeated gradient directions as possible with weighted or ML estimation methods.

4   Experiments and Validation

We acquired test sets of data under different imaging protocols to compare with the Monte Carlo simulations. Three sets of images of a single healthy adult volunteer were acquired on a Siemens Allegra 3T head-only scanner. The scanning time for each sequence was approximately 12 minutes. Diffusion weighted images were acquired with an isotropic resolution of $2x2x2mm$ resolution and image size 128x128x39. Three different sequences were used: 6 directions with 10

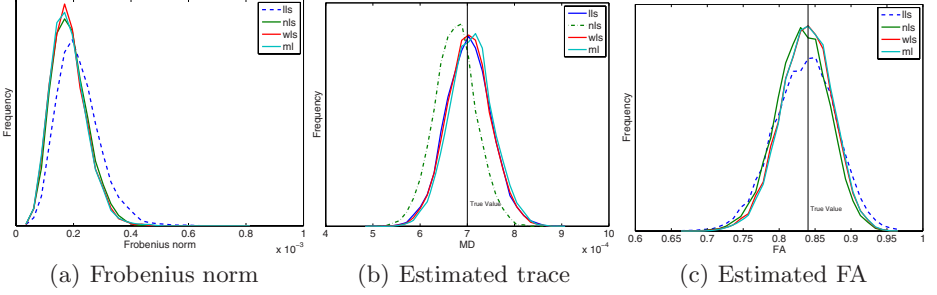(a) Frobenius norm   (b) Estimated trace   (c) Estimated FA

Fig. 3. Distribution of estimated values in Monte Carlo simulations of a tensor with FA=0.8 using 10,000 repetitions and a 60 direction protocol: (a) the Frobenius norm of the difference between the estimated and true tensor, (b) the estimated trace, and (c) the estimated FA. The weighted, non-linear, and ML estimation techniques reduce the variance of estimated FA and decrease the Frobenius norm of the error in the estimated tensor. Non-linear least squares tends to underestimate the trace.

repetitions, 21 directions with 3 repetitions, and 60 directions with 1 repetition. All scans were of approximately equal time to demonstrate the trade-off between image repetition and acquiring more gradient directions. To eliminate bias from differences in the baseline images, the 14 acquired baseline images were registered to a T2 atlas using a rigid transformation and normalized mutual information. The baseline images were averaged to produce a common baseline image. The transformation from each baseline image to the atlas was also applied to the diffusion-weighted images in the corresponding set, and the gradient directions were corrected by the rotation component of the transformation. For sequences with repeated directions, the corresponding gradient directions were averaged. The noise level σ was estimated from the background of the images.

A white matter segmentation was created by co-registration of a T1 structural image to the averaged baseline image and applying a tissue segmentation tool. A label image was created from the six direction image by identifying voxels within the white matter segmentation which are highly aligned or highly un-aligned with the closest gradient direction. The angle for each voxel is given by $\theta = \arccos(\min_i(\hat{e}_1 \cdot \boldsymbol{g}_i))$, where $\hat{e}_1$ is the estimated principal eigenvector. In the six direction scan nearby gradients are separated by $pi/2$, and as a result the maximum angle between the principal eigenvector and the nearest gradient direction is $\pi/4$ so $\theta \in [0, \pi/4]$. The threshold for aligned tensors was $\theta < \pi/16$ and for unaligned tensors $\theta > 3\pi/16$. Figure 1(c) shows the labels overlaid on the FA image.

The difference between the histograms of FA for aligned and unaligned tensors decreases with an increase in the number of gradient directions as shown in Figure 4. Table 1 lists the mean difference between aligned and unaligned voxels using different gradient schemes and tensor estimation methods. The experimental results confirm the simulation prediction of underestimated anisotropy in tensors aligned in the six direction scan, because the difference decreases as
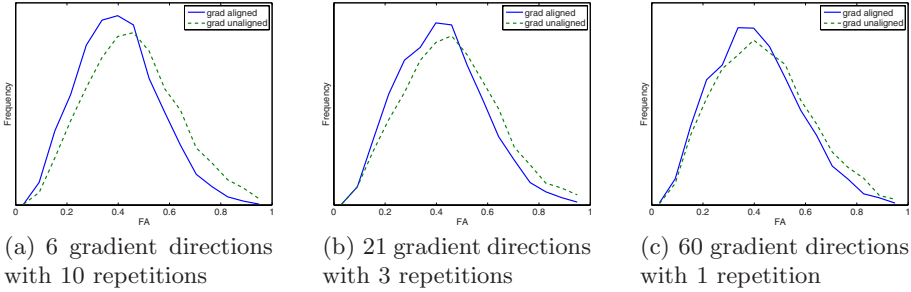
(a) 6 gradient directions with 10 repetitions

(b) 21 gradient directions with 3 repetitions

(c) 60 gradient directions with 1 repetition

Fig. 4. The histograms of estimated FA using weighted least squares estimation for aligned and unaligned tensors with (a) 6 gradient directions with 10 repetitions (b) 21 gradient directions with 3 repetitions and (c) 60 gradient directions with 1 repetition. Notice the significant reduction in difference between the two histograms as the number of gradient directions increases, because of a decrease in noise bias. Quantitative results are given in Table 1.

Table 1. Mean difference of FA values between aligned and unaligned voxels. The difference between the mean FA decreases with an increase in the number of gradient directions, confirming the simulation prediction that some of the difference is due to the choice of gradient sampling. All least squares methods are equivalent in the six direction scan because an exact solution exists.

	6 directions (10 reps)	21 directions (3 reps)	60 directions (1 reps)
LLS:	0.053	0.030	0.019
NLS:	0.053	0.032	0.020
WLS:	0.053	0.032	0.021
ML:	0.045	0.033	0.021

Table 2. Mean and variance of trace values in white matter across gradient sampling schemes and estimation methods. The six direction scan has the highest estimated trace. In the 21 and 60 direction scans the non-linear least squares method has a lower estimate of the trace than the other methods. The maximum-likelihood has the highest estimate of trace of all the methods.

	6 directions (10 reps)	21 directions (3 reps)	60 directions (1 reps)
LLS:	2.373e-03 (7.2071e-08)	2.304e-03 (7.7886e-08)	2.307e-03 (2.2293e-07)
NLS:	2.373e-03 (7.2071e-08)	2.289e-03 (7.7122e-08)	2.235e-03 (2.0839e-07)
WLS:	2.373e-03 (7.2071e-08)	2.307e-03 (7.8263e-08)	2.311e-03 (2.2847e-07)
ML:	2.472e-03 (8.5249e-08)	2.383e-03 (1.0698e-07)	2.326e-03 (2.4044e-07)

the number of gradient direction increases. The experimental estimate of tensor trace is higher for low directions images, which is different than the simulation prediction. The difference could be due to the assumption of a single tensor model in the simulation. We conclude that studies comparing different regions could be substantially biased by the orientation of the tissue within the magnetic field. Therefore studies relying on statistical analysis of anisotropy measures should

use as many gradient directions as possible instead of repeating a minimum number of gradient directions. The results from the analysis of the in-vivo scans confirms the results of the simulation experiments.

5   Conclusions

In this paper we have evaluated the magnitude of error in DTI measurements caused by imaging noise through Monte Carlo simulations and in-vivo experiments. We have shown that low direction schemes introduce a statistical bias with a clinically relevant magnitude. Furthermore, we have shown that standard linear least squares tensor estimation introduces additional variability in tensor estimation. Understanding the magnitude of these two effects is critical for interpreting the results of statistical analysis. For new imaging studies these results indicate that scans with non-repeated isotropic gradient sampling should be preferred over protocols with a small number of repeated gradient directions, and that weighted least squares or ML tensor estimation should be preferred. Future work will extend the analysis to errors which are introduced by EPI distortion, subject motion, and inter-subject registration.

References

1. Kubicki, M., Park, H., Westin, C., Nestor, P., Mulkern, R., Maier, S., Niznikiewicz, M., Connor, E., Levitt, J., Frumin, M., Kikinis, R., Jolesz, F., McCarley, R., Shenton, M.: DTI and MTR abnormalities in schizophrenia: Analysis of white matter integrity. NeuroImage 26(3), 1109–1118 (2005)
2. Bastin, M.E., Armitage, P.A., Marshall, I.: A theoretical study of the effect of experimental noise on the measurement of anisotropy in diffusion imaging. Magn Reson Imaging 16(7), 773–785 (1998)
3. Jones, D.K.: The Effect of Gradient Sampling Schemes on Measures Derived From Diffusion Tensor MRI: A Monte Carlo Study. Magnetic Resonance in Medicine 51(4), 807–815 (2004)
4. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. Magnetic Resonance in Medicine 44(1), 41–50 (2000)
5. Jones, D.K., Basser, P.J.: Squashing peanuts and smashing pumpkins: How noise distorts diffusion-weighted MR data. Magnetic Resonance in Medicine 52(5), 979–993 (2004)
6. Basu, S., Fletcher, P.T., Whitaker, R.: Rician Noise Removal in Diffusion Tensor MRI. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 117–125. Springer, Heidelberg (2006)
7. Fillard, P., Arsigny, V., Pennec, X., Ayache, N.: Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In: ISBI 2006, Crystal Gateway Marriott, Arlington, Virginia, USA. LNCS, pp. 786–789 (2006)
8. Salvador, R., Peña, A., Menon, D.K., Carpenter, T.A., Pickard, J.D., Bullmore, E.T.: Formal characterization and extension of the linearized diffusion tensor model. Human Brain Mapping 24(2), 144–155 (2005)