

Assessment of Bias for MRI Diffusion Tensor Imaging Using SIMEX

Carolyn B. Lauzon^{1,2}, Andrew J. Asman¹, Ciprian Crainiceanu³,
Brian C. Caffo³, and Bennett A. Landman^{1,2}

¹ Department of Electrical Engineering, Vanderbilt University, Nashville, TN, USA

² Institute of Imaging Science, Vanderbilt University, Nashville, TN, USA

³ Department of Biostatistics, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA

Abstract. Diffusion Tensor Imaging (DTI) is a Magnetic Resonance Imaging method for measuring water diffusion *in vivo*. One powerful DTI contrast is fractional anisotropy (FA). FA reflects the strength of water's diffusion directional preference and is a primary metric for neuronal fiber tracking. As with other DTI contrasts, FA measurements are obscured by the well established presence of bias. DTI bias has been challenging to assess because it is a multi-variable problem including SNR, six tensor parameters, and the DTI collection and processing method used. SIMEX is a modern statistical technique that estimates bias by tracking measurement error as a function of added noise. Here, we use SIMEX to assess bias in FA measurements and show the method provides; i) accurate FA bias estimates, ii) representation of FA bias that is data set specific and accessible to non-statisticians, and iii) a first time possibility for incorporation of bias into DTI data analysis.

Keywords: DTI, FA, bias, SIMEX, parameter estimation, bias correction, diffusion, tensor, imaging.

1 Introduction

Radiological practice appears to be on the cusp of reliably and routinely using imaging biomarkers as surrogates for traditional tissue biopsy (e.g., virtual biopsy). One promising magnetic resonance imaging (MRI) technique, diffusion tensor imaging (DTI) provides unique, non-invasive contrasts sensitive to tissue micro-architecture, which can be correlated with cellular organization – for example, as seen through histological analysis [1]. However, quantification of observed data is an *estimation* process and numeric values must be viewed in light of their distributional properties. Extensive research has been invested in characterizing the impacts of noise and study design on DTI contrasts[2, 3], and it is well recognized that both lead to changes in precision (*variance*) as well as accuracy (*bias*) of diffusion derived measures. Much progress has been made on variability analysis in DTI using Monte Carlo and boot-strap techniques [4], but quantitative *post hoc* assessment of the bias in diffusion estimated parameters has remained elusive.

Bias is explicitly defined as the difference between the expected value of an experimentally observed measure, $\mathbf{E}(\Theta_{\text{obs}})$, and the true value of that measure, Θ_{truth} ; $\text{bias} = \mathbf{E}(\Theta_{\text{obs}}) - \Theta_{\text{truth}}$. As seen, bias causes the statistical average of a measured DTI contrast to converge on the wrong value. Also, differing propensities for motion and/or underlying anatomy could lead to systematically different tensor bias between disease and healthy subgroups, thus statistically impacting analyses. Bias in DTI is a multi-variable function including imaging noise, field in homogeneity, DTI experimental parameters, and the underlying biological diffusion parameters. No analytic equation for bias in DTI is known and no published method to account for bias in empirical DTI data has been tried. Herein, we present a statistical approach to estimate the expected level of bias in DTI contrasts on a voxel-wise basis.

In this initial evaluation, we focus on quantifying bias in the commonly studied and important DTI contrast fractional anisotropy (FA) as measured using the traditional log-linear minimum mean-squared error (LLMMSE) tensor estimation framework. Significant bias in FA has been well documented both experimentally and through computational studies[3, 5]. Within a clinical and research SNR range (~15:1 – 40:1), FA bias is generally a monotonic decreasing function, with lower FA values tending to have larger bias than higher FA values.

To estimate FA bias, we test the SIMulation EXtrapolation method (SIMEX) which emerged from the statistical literature as a way to compensate for bias due to additive noise within a regression framework[6, 7]. However, it can also be viewed as a general procedure to extrapolate the degree of bias in an estimate given noisy data by studying the sensitivity of the model fitting procedure when *additional synthetic* noise is added. Two key assumptions of SIMEX are (i) the bias as a function of SNR is monotonic and smooth and (ii) the noise level of the observed data is well understood and estimated. Fortunately, (i) is generally holds and for (ii), noise estimation has been well-studied in the DTI literature and several robust methods exist to estimate the noise field given empirical data [8-10].

This manuscript is organized as follows. First, we briefly review the diffusion tensor formulation and describe how SIMEX bias estimation can be applied to bias estimation for FA. Second, we present a simulation based on empirical data which enables knowledge of the true bias level for test comparisons with the SIMEX estimated bias. Third, we present single and multi-voxel demonstrations of SIMEX estimation of bias in FA. We close with a brief critique and discussion of potential applications for this approach and avenues for continuing research.

2 Theory

2.1 Diffusion Tensor Imaging

In DTI, magnetic gradients (\mathbf{g}) are used to label the relative position of water at the start and end of an experiment. Given the random thermal motions of water, the labels add incoherently and result in an attenuated, diffusion weighted image (dwi). In general, the greater probabilistic movement, the greater the signal attenuation. In the tensor model, the dwi will be a function of the unweighted signal intensity, symbolized as S_0 , diffusion direction and magnitude (encoded in diffusion tensor \mathbf{D}), the experimental timing and gradient strength (encoded in parameter b), and the direction of the gradient applied (encoded in \mathbf{g}). For a single voxel we write

$$dwi_j = S_o \exp(-b \mathbf{g}_j^T \cdot \mathbf{D} \cdot \mathbf{g}_j). \quad (1)$$

Here, j indexes the gradient directions. We define the observed data for a single voxel, \vec{X}_{obs} , as a vector with elements x_i , representing dwi_j for $j=1, 2, \dots, m$, and S_o .

$$\vec{X}_{\text{obs}} = \{dwi |_{j=1}^m, S_o\} = \{x_i |_{i=1}^{m+1}\}. \quad (2)$$

\mathbf{D} is a 3×3 symmetric matrix whose Eigenvalues, λ_1 , λ_2 , and λ_3 , describe the three principal axis of diffusion. Herein, we consider the FA of the tensor which is the normalized standard deviation of the Eigenvalues and reflects the relative degree of orientation (i.e., 0: no orientation preference = isotropic diffusion, 1: infinite directional preference = 1-D diffusion).

$$FA = \sqrt{\frac{3}{2} \frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \quad (3)$$

2.2 SIMEX Applied to DTI

The premise of SIMEX is simple. Bias can be understood by adding noise to data in incremental amounts and measuring the resulting contrast. The trend in the contrast with added noise should enable prediction of the contrast with 'removed' noise. For simplicity, the following description assumes a single DTI experiment is performed (e.g. single subject, one DTI dataset) and describes the single-voxel case, though the analysis can be extended to multiple experiments and multiple voxels.

Let a truth data set be described by zero experimental noise

$$\vec{X}_{\text{truth}} = \{x_i |_{i=1}^{m+1}\} \text{ s. t. } (\sigma_E = 0), \quad (4)$$

where σ_E is the standard deviation of the experimental noise. The function \mathbf{T} mapping \vec{X}_{truth} to FA determines the calculated ground truth FA value, FA_{truth} ,

$$FA_{\text{truth}} = \mathbf{T}(\{\vec{X}_{\text{truth}}\}). \quad (5)$$

The observed data of an actual experiment contains experimental noise and is described by,

$$\vec{X}_{\text{obs}} = \vec{X}_{\text{truth}} + \sigma_E \vec{Z}. \quad (6)$$

Here, \vec{Z} is a vector of $m+1$ random drawings from a standard normal distribution (and noise is independent across voxels when expanded to the multi-voxel case). The resulting observed FA from the experiment is then,

$$FA_{\text{obs}} = \mathbf{T}(\{\vec{X}_{\text{obs}}\}). \quad (7)$$

In SIMEX, synthetic normally distributed noise with standard deviation $\omega^{1/2} \sigma_E$ is added to \vec{X}_{obs} , and the new simulated data is written as,

$$\vec{X}_{\text{M.C.}}(\omega) = \vec{X}_{\text{obs}} + \omega^{1/2} \sigma_E \vec{Z}_{\text{M.C.}}. \quad (8)$$

M.C. stands for ‘Monte Carlo’ (Note, M.C. replaces ‘b’ as defined in the SIMEX literature[7]). The simulated FA is then described by,

$$FA_{M.C.}(\omega) = \mathbf{T}(\{\bar{X}_{M.C.}(\omega)\}) . \tag{9}$$

The variance of the simulated data as a function of ω is

$$\begin{aligned} \text{var}(\bar{X}_{M.C.}(\omega)) &= \\ \text{var}(\sigma_E \bar{Z} + \omega^{1/2} \sigma_E \bar{Z}_{M.C.}) &= \\ \sigma_E^2(1 + \omega) . \end{aligned} \tag{10}$$

A key observation is to note that at the value $\omega = -1$, (which cannot be simulated) the variance goes to 0. For repeated Monte Carlo simulations of $FA_{M.C.}(\omega)$, a mean value can be calculated, $\bar{FA}_{M.C.}$, and the trend approximated by fitting $\bar{FA}_{M.C.}(\omega)$ with a standard non-linear equation, $\bar{FA}(\omega)$,

$$\bar{FA}(\omega) = a + \frac{b}{1 + \omega} \tag{11}$$

The FA at zero noise can then be estimated by solving eq. 11 for $\omega = -1$,

$$FA_{SIMEX} = \bar{FA}(-1) . \tag{12}$$

The bias in the data can then be estimated,

$$\text{Estimated Bias} = FA_{obs} - FA_{SIMEX} . \tag{13}$$

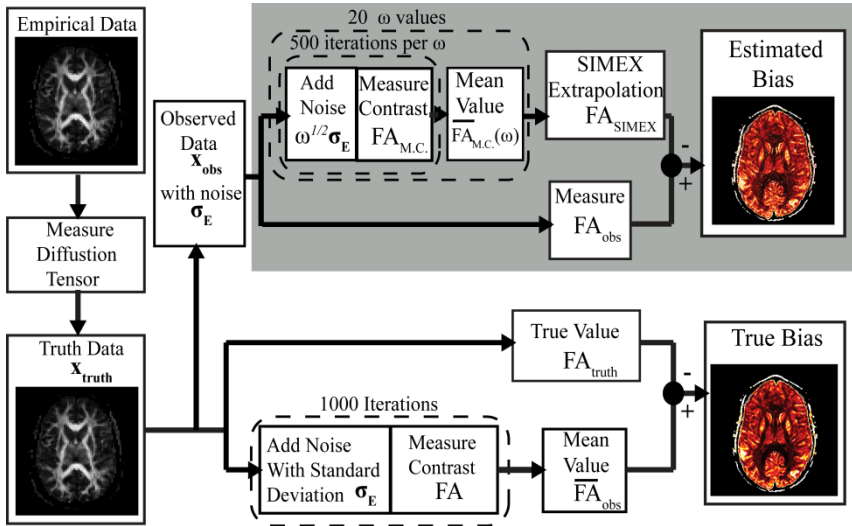


Fig. 1. Flow chart of steps to create SIMEX Estimated Bias and True Bias maps. Variables refer to terms defined explicitly by equations in the Theory section.

The true bias can also be calculated if the truth data is known,

$$\text{True Bias} = \mathbf{E}(\text{FA}_{\text{obs}}) - \text{FA}_{\text{Truth}} . \quad (14)$$

If the true extrapolant function could be used in place of eq. 11, then FA_{SIMEX} would be an asymptotically unbiased estimator of FA_{truth} . Under the condition where the true but unknown extrapolant is smooth w.r.t. ω , the approximating extrapolant, eq. 11, can provide approximately unbiased results.

3 Methods and Results

The steps used here to evaluate SIMEX on FA are summarized in the flow chart of Fig. 1 and exemplified by individual voxels in Fig. 2. Unless otherwise stated, all processing and analysis was performed in Matlab 2010 (Mathworks, Natick, MA). Diffusion tensor estimates were calculated by fitting the diffusion model (eq.1) to the data using a simple LLMMSE. In all cases, noise was added to data using a Rician distribution, which is approximately Gaussian for $\text{SNR} > 5:1$ [11].

Empirical Data. Empirical data (KKI2009-33-DTI) were downloaded from the Multi-Modal MRI Reproducibility study available online at www.nitrc.org [12]. Full collection details are provided in the reference. Briefly, DTI data was collected using a multi-slice, single-shot, echo planar imaging (EPI) sequence. Thirty two gradient directions were used with $b = 700 \text{ s/mm}^2$ ($m = 32$, $b = 700$, eq. 1 & 2). Five S_0 images were collected and averaged into a single S_0 . The resulting images consisted of sixty-five transverse slices with a field of view = 212×212 reconstructed to 256×256 .

Creating Noiseless Ground Truth (eq. 4). The 32 dwi images were registered to the single S_0 image using FLIRT affine registration (FMRIB, Oxford, UK). For spatial consistency, the resulting affine transformation matrix was applied to the gradient table. A single axial slice of interest was selected and diffusion tensor estimates (\mathbf{D}) for that slice were calculated. Using the S_0 image, the estimated diffusion tensor, and the given \mathbf{g} and b as inputs to eq. 1, dwi images were simulated. The S_0 image and the 32 noiseless dwi images make the noiseless ground truth dataset, $\bar{\mathbf{X}}_{\text{truth}}$.

Creating Observed Data (eq. 6). To create the observed data, $\bar{\mathbf{X}}_{\text{obs}}$, noise was added to the ground truth data set such that $\text{SNR} = 35:1$. The signal strength was estimated from the average signal intensity of the masked S_0 image. Noise, $\sigma_E = \text{signal}/35$, was added to the ground truth data to create the 32 dwi images and the single S_0 image.

Calculating the Estimated Bias (eq. 13). To calculate FA_{SIMEX} , first a plot of $\bar{\text{FA}}_{\text{M.C.}}(\omega)$ needed to be constructed. 500 Monte Carlo simulations for 20 different ω values were performed and $\bar{\text{FA}}_{\text{M.C.}}(\omega)$ calculated from the average. Each simulation consisted of a noise adding step to create $\mathbf{X}_{\text{M.C.}}$ (eq. 8), followed by a calculation of $\text{FA}_{\text{M.C.}}$. Example $\bar{\text{FA}}_{\text{M.C.}}(\omega)$ for three voxels are shown in Fig. 2A. $\bar{\text{FA}}(\omega)$ was fit using the empirical non-linear equation (eq. 11). The fit equation was extrapolated to $\omega = -1$, producing FA_{SIMEX} (Fig. 2B). The FA_{obs} from eq. 13 was calculated directly from diffusion tensor fits to the observed data set, $\bar{\mathbf{X}}_{\text{obs}}$. The difference of FA_{SIMEX} and FA_{obs} was taken and the Estimated Bias calculated.

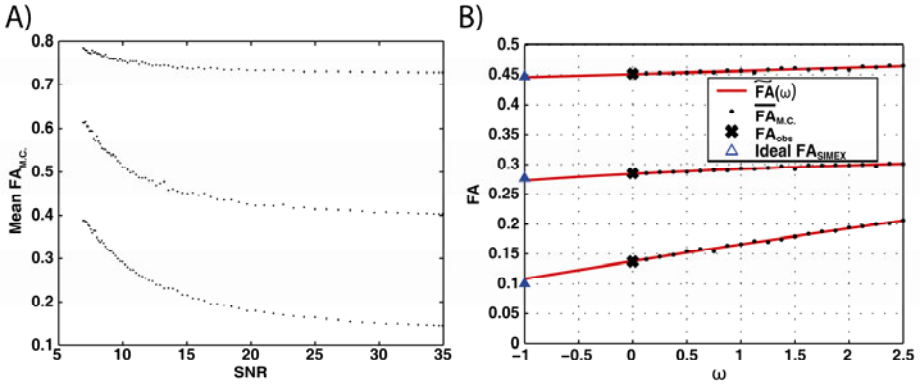


Fig. 2. (A) Exploratory SIMEX plots of simulated $FA_{M.C.}$ as a function of SNR for three voxels from the simulated observed data. 1000 Monte Carlo simulations were averaged for each $\overline{FA}_{M.C.}$. The FA_{obs} is the point at the highest SNR value (SNR = 35:1). An upward bias with decreasing SNR is observed for all three voxels. (B) $FA_{M.C.}(\omega)$ and the extrapolation fit for three voxels in the observed data set. The x-axis for extrapolation is $\omega \sim 1/SNR$ and the fit is extrapolated to $\omega = -1$. The ideal FA_{SIMEX} (blue triangles) represents the extrapolated FA_{SIMEX} value if the true bias was correctly estimated.

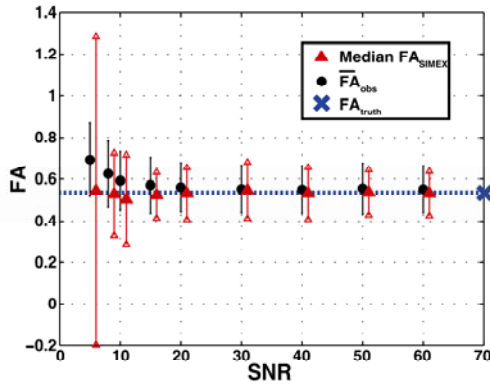


Fig. 3. Simulated outcomes for SIMEX at several SNR values. \overline{FA}_{obs} values for each SNR value (circles) were calculated starting from a voxel in the simulated truth data. The \overline{FA}_{obs} values are shown with the standard deviation (error bars) from 1000 simulations run at each SNR value. 100 observations at each SNR value were randomly chosen for SIMEX. The median FA_{SIMEX} values (triangles) are shown with the standard deviation (error bars). The median value was chosen due to the decreased robustness of only 100 simulations.

Calculating True Bias (eq. 14). The average of 1000 iterations was used to estimate the expected value of FA_{obs} , $\overline{FA}_{obs} \sim E(FA_{obs})$. Each simulation consisted of first adding noise (SNR = 35:1) to \overline{X}_{truth} followed by calculation of FA. FA_{truth} was calculated directly from \overline{X}_{truth} . The True Bias was calculated from the difference between \overline{FA}_{obs} and FA_{truth} . Exploratory M.C. simulations applying SIMEX to different SNR levels of \overline{X}_{obs} for a

single ground truth voxel are shown in Fig. 3. The True Bias and Estimated Bias are compared in Fig. 4. The results shown in Fig. 4 were typical for multiple test trials (data not shown).

4 Discussion

Potential systematic differences in distributional properties of tensor-derived contrasts pose serious hazards for statistical analysis. Understanding “statistical artifacts” in DTI has long been an essential challenge in proper interpretation of DTI contrasts [13]. Routine approaches for interpreting contrasts (e.g., regions of interest, voxel-based morphometry, tract-based spatial statistics) hinge upon the assumption of equivalent (or zero) bias between populations. If this assumption is violated then the inferences (and p-values) may be suspect. In any empirical experiment, there will be variations in bias due to subject characteristics, protocols, scanner stability, hardware/software changes, etc. Without a method to estimate bias in measures derived from individual empirical datasets, evaluation of any change requires an extensive simulation study based on a range of potential anatomical models.

With the proposed SIMEX approach, it is possible to estimate bias in contrasts on a voxel-by-voxel level given the empirical data. SIMEX proved a sensitive method for detecting FA bias as evidenced by its success on the high SNR (and therefore relatively low bias) dataset used here (Fig. 4). Although bias may be negligible for some individual cases at high SNR values (such as exemplified for high SNR values in Fig. 3), bias may become increasingly important at lower SNR values or in hypothesis test settings where even a small bias can significantly inflate the Type-I error rate. To that end, the individual voxel test case (Fig. 3) suggests SIMEX to be a promising technique for empirical data collected at lower SNR levels (stable median FA_{SIMEX} values) as well as a reliable method for repeated use on individual datasets within grouped data (stable error bars). The instability of the SIMEX estimate for $SNR \leq 10:1$ (Fig. 3) is likely to be a result of the non-monotonic behavior of bias at extremely low SNR values [5].

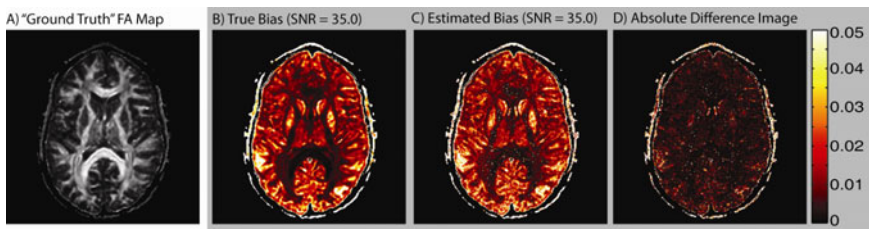


Fig. 4. (A) The ‘ground truth’ FA map (FA_{truth}) of the slice selected for data analysis. (B) The True Bias map for an SNR of 35:1 is compared to (C) the absolute value of the Estimated Bias at SNR = 35:1. The absolute value difference map was calculated from the raw bias data (not the difference in the absolute values). For a significant majority of voxels the True Bias is positive and the Estimated Bias is greater than the True Bias.

We note that at these early stages, we are not specifically recommending using the estimated bias measures to correct for voxel-wise bias. Optimization of user input SIMEX parameters (e.g. the approximation function in eq. 11 or the number of Monte Carlo iterations) and repeat testing on empirical data of various SNR levels is underway. Refinements may help decrease the variability in the SIMEX estimate (Fig. 3 red error bars) and mediate some of the isolated outliers seen in Fig. 4D. Additionally, methods for constructing error estimates of the measured SIMEX bias exist and have yet to be tested in the context of DTI[6]. With continued future refinement and careful incorporation of robust/advanced tensor estimators it may be possible to produce unbiased contrast estimates without the disadvantages of direct maximum likelihood methods[14].

In summary, a method for direct, empirical assessment of bias in tensor derived contrasts opens many possibilities for useful statistical assessments and improved analysis approaches. Further investigation is warranted into bias estimation with other tensor contrasts, using advanced tensor fitting methods, and examining quantitative tissue models beyond the tensor formulation.

Acknowledgments. This work was supported in part by NIH/NINDS 1R01NS056307 and NIH/NINDS 1R21NS064534.

References

1. Le Bihan, D., van Zijl, P.: From the diffusion coefficient to the diffusion tensor. *NMR Biomed* 15, 431–434 (2002)
2. Landman, B.A., Farrell, J.A., Jones, C.K., Smith, S.A., Prince, J.L., Mori, S.: Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T. *Neuroimage* 36, 1123–1138 (2007)
3. Jones, D.K., Basser, P.J.: Squashing peanuts and smashing pumpkins: How noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine* 52, 979–993 (2004)
4. Whitcher, B., Tuch, D.S., Wisco, J.J., Sorensen, A.G., Wang, L.: Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging. *Hum Brain Mapp* 29, 346–362 (2008)
5. Landman, B.A., Farrell, J.A., Huang, H., Prince, J.L., Mori, S.: Diffusion tensor imaging at low SNR: nonmonotonic behaviors of tensor contrasts. *Magn. Reson Imaging* (2008)
6. Carroll, R.J.: *Measurement error in nonlinear models: a modern perspective*. CRC Press, Boca Raton (2006)
7. Cook, J., Stefanski, L.: Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association* 89 (1994)
8. Landman, B.A., Bazin, P.L., Smith, S.A., Prince, J.L.: Robust estimation of spatially variable noise fields. *Magn. Reson Med.* 62, 500–509 (2009)
9. Rajan, J., Poot, D., Juntu, J., Sijbers, J.: Noise measurement from magnitude MRI using local estimates of variance and skewness. *Phys Med Biol* 55, N441–N449 (2010)

10. Aja-Fernandez, S., Tristan-Vega, A., Alberola-Lopez, C.: Noise estimation in single- and multiple-coil magnetic resonance data based on statistical models. *Magn. Reson Imaging* 27, 1397–1409 (2009)
11. Gudbjartsson, H., Patz, S.: The Rician distribution of noisy MRI data. *Magn Reson Med.* 34, 910–914 (1995)
12. Landman, B., Huang, A., Gifford, A., Vikram, D., Lim, I., Farrell, J., Bogovic, J., Hua, J., Chen, M., Jarso, S., Smith, S., Joel, S., Mori, S., Pekar, J., Barker, P., Prince, J., van Zijl, P.: Multi-Parametric Neuroimaging Reproducibility: A 3T Resource Study. *Neuroimage* 4, 2854–2866 (2011)
13. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. *Magn. Reson Med.* 44, 41–50 (2000)
14. Anderson, A.W.: Theoretical analysis of the effects of noise on diffusion tensor imaging. *Magn Reson Med.* 46, 1174–1188 (2001)