Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 161 (2003) 1–14
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Parametric and non-parametric statistical analysis of DT-MRI data☆

Sinisa Pajevic[a] and Peter J. Basser[b,*]

[a] *Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892-5772, USA*
[b] *Section on Tissue Biophysics and Biomimetics, Laboratory of Integrative and Medical Biophysics, National Institute of Child Health and Human Development, National Institutes of Health, NIH/13 South Dr./Building 13, Room 3W16, Bethesda, MD 20892-5772, USA*

Received 23 May 2000; revised 29 November 2002

Abstract

In this work parametric and non-parametric statistical methods are proposed to analyze Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) data. A Multivariate Normal Distribution is proposed as a parametric statistical model of diffusion tensor data when magnitude MR images contain no artifacts other than Johnson noise. We test this model using Monte Carlo (MC) simulations of DT-MRI experiments. The non-parametric approach proposed here is an implementation of bootstrap methodology that we call the DT-MRI bootstrap. It is used to estimate an empirical probability distribution of experimental DT-MRI data, and to perform hypothesis tests on them. The DT-MRI bootstrap is also used to obtain various statistics of DT-MRI parameters within a single voxel, and within a region of interest (ROI); we also use the bootstrap to study the intrinsic variability of these parameters in the ROI, independent of background noise. We evaluate the DT-MRI bootstrap using MC simulations and apply it to DT-MRI data acquired on human brain in vivo, and on a phantom with uniform diffusion properties.
Published by Elsevier Science (USA).

Keywords: Diffusion; Tensor; MR; MRI; Bootstrap; Multivariate; Normal; Distribution

1. Introduction

Applications of Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) [3] have grown significantly in recent years. This is because from the measured effective diffusion tensor, $\mathbf{D}^{\text{eff}}$, one can calculate new and useful parameters such as the trace of the diffusion tensor [4], trace($\mathbf{D}^{\text{eff}}$), the relative and fractional anisotropies, the three principal diffusivities, and other rotationally invariant quantities calculated from them [5]. Additionally, from the principal directions of the diffusion tensor one can determine fiber direction [3] and even follow fiber tract trajectories [6–14]. Nevertheless, only recently have attempts been made to quantify statistical uncertainties of tensor-derived quantities or to characterize statistical distributions of diffusion tensor data and of quantities derived from them [15–21]. Knowing the uncertainties, higher moments, and probability distributions of various DT-MRI parameters could improve our ability to glean more information from DT-MRI data and to design DT-MRI experiments more efficiently, particularly longitudinal or multi-site studies.

The main goals of this work are (i) to propose the appropriate parametric statistical model of artifact-free diffusion tensor data containing only thermal Johnson noise (i.e., with no motion and other artifacts), and to test the parametric model using Monte Carlo (MC) methods, and (ii) to develop, apply, and assess the efficacy of non-parametric methodologies to analyze experimental DT-MRI data. In this paper the assessment is performed only on the diffusion tensor elements and some simple, local, voxel-based parameters derived from them. This non-parametric methodology described here can also be used on other voxel-based quantities, as well as non-local, region-based quantities (e.g., lattice index [21] or fiber tracts [14,22]).

☆ Some material presented previously [1,2].

* Corresponding author. Fax: 1-301-435-5035.
E-mail addresses: pajevic@nih.gov (S. Pajevic), pjbasser@helix.nih.gov (P.J. Basser).

1090-7807/03/$ - see front matter. Published by Elsevier Science (USA).
doi:10.1016/S1090-7807(02)00178-7

2. Theory

2.1. Parametric model for noise in DT-MRI

Background noise in magnitude reconstructed MRI images is known to conform to a Rician distribution [23,24], which can be written as

$$p(M,A,\sigma) = \frac{M}{\sigma^2} e^{-(M^2+A^2)/2\sigma^2} I_0\left(\frac{MA}{\sigma^2}\right), \quad (1)$$

where $I_0(x)$ is the modified Bessel function of the first kind, M is the magnitude, A is the parameter that controls the mean of the distribution, and σ is the standard deviation of the experimental noise. In DT-MRI, components of the diffusion tensor are estimated from noisy diffusion weighted (DW) magnitude images using regression analysis, generally multivariate linear regression of the log-linearized DW magnitude signals [25]. In this work we propose the appropriate parametric distribution of noise in diffusion tensor data within a voxel, assuming that the magnitude images are Rician distributed.

For signal-to-noise (S/N) ratios greater than 3, the Rician distribution is well approximated by a Normal distribution with mean $\sqrt{A^2+\sigma^2}$, and variance σ^2, which is the variance of the signal in each of the quadrature channels [23,24]. However, the relationship between the magnitude of the NMR signal and the diffusion tensor components in the absence of noise is nonlinear. Thus, arriving at an analytical expression for the noise distribution of the individual diffusion tensor elements is problematic. However, since regression analysis is usually performed using a large number of independent diffusion weighted signals or images (i.e., >7), we expect that the distribution of the diffusion tensor elements will be Gaussian owing to the Central Limit Theorem, even for small S/N ratios. This result applies for both linear and nonlinear regression (due to a property of least square minimization). In other words, we expect that the probability distribution of the six independent components of $\mathbf{D}^{\text{eff}}$, which can be written as a 6-dimensional vector, are distributed according to a Multivariate Normal distribution.

This distribution is described by two parameters: a 6×6 covariance matrix, $\boldsymbol{\Sigma}$, and a 6-dimensional mean vector, $\boldsymbol{\mu} = \{\overline{D}_{xx}, \overline{D}_{yy}, \overline{D}_{zz}, \overline{D}_{xy}, \overline{D}_{xz}, \overline{D}_{yz}\}$. For a random vector $\mathbf{x} = \{D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}\}$ the Multivariate Normal distribution is written as

$$p(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{|\Sigma|^{-1/2}}{(2\pi)^3} e^{-1/2(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}. \quad (2)$$

For both log-linear and nonlinear regression, which are commonly used in DT-MRI, an approximate relationship between $\boldsymbol{\Sigma}$ in Eq. (2) and σ of DW images in Eq. (1) can be derived [25]. In multivariate linear regression, the 6×6 covariance matrix can be expressed as

$$\boldsymbol{\Sigma} = (\boldsymbol{B}^{\mathrm{T}} \Sigma_{\mathrm{M}}^{-1} \boldsymbol{B})^{-1}, \quad (3)$$

where $\boldsymbol{B}$ is the experimental B-matrix ($6 \times N_{\mathrm{M}}$) described in Section 3.1.3, and Σ_{M} is the $N_{\mathrm{M}} \times N_{\mathrm{M}}$ (diagonal) covariance matrix containing the experimental variances of each of the N_{M} independent measurements. The diagonal terms of Σ_{M} are usually weighted by the square of the magnitude of the corresponding MR signal-to-correct for distortions introduced in log-linearized data [26]. Approximate expressions can also be derived for nonlinear regression which uses a large number of observations [27,28].

If a homogeneous and sufficiently large region of interest (ROI) containing identically distributed diffusion tensors can be found, estimating $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ is straightforward. However, this is generally not possible in practice, since material properties and composition of a tissue are usually not known a priori, and uniformity cannot be assumed. Instead, such estimates can be obtained using non-parametric empirical statistical methods, one of which is described in Section 3.

2.2. Non-parametric analysis using bootstrap

Even when we know the parametric distribution of noise in diffusion tensor data, for most useful DT-MRI parameters, estimates of errors and probability distributions cannot be calculated since their parametric distributions are not known a priori. Generally, this is because these quantities are complicated nonlinear functions of the diffusion tensor elements. Therefore, using an empirical approach like the bootstrap is required.

As a general non-parametric approach to analyze DT-MRI data, we propose a particular implementation of the bootstrap method [29,30]. Bootstrap analysis is an empirical, non-parametric technique commonly used to obtain various uncertainty measures of a given statistic when the underlying statistical model is not known. Fig. 1 illustrates graphically the general bootstrap procedure, in which many bootstrap estimates (θ_i^*) of a given statistic θ are obtained by randomly drawing with replacement from the original sample. These estimates are used to obtain the standard errors, bias, confidence intervals, probability distributions, and other measures of uncertainty for a given statistic. In the same way we use ordinary samples to infer statistical properties of a population, we use bootstrap samples to infer the properties of our sample. The bootstrap estimate of the standard error (SE) is the non-parametric maximum likelihood estimate of the true SE. The bootstrap can also be used to determine the bias in the estimate of a given statistic.

In this work we propose the acquisition scheme in a DT-MRI experiment that enables us to analyze DT data using the bootstrap method. The implementation of this

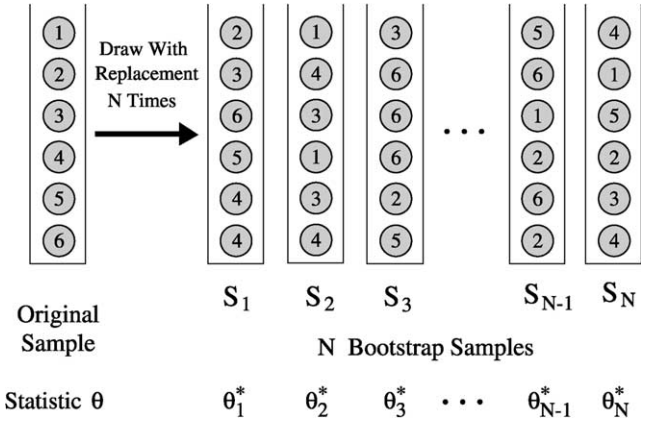

Fig. 1. A graphical illustration of the bootstrap samples obtained by drawing with replacement from the original sample. By repeating the process N times, one can obtain N bootstrap estimates (θ_i^*) of a given statistic θ. These estimates can then be used to obtain the standard errors, confidence intervals, probability distributions, and other measures of uncertainty of a given statistic.

bootstrap scheme, that we call the DT-MRI bootstrap is described below.

3. Methods

3.1. DT-MRI methods

3.1.1. Generation of synthetic DWI data

A set of DWIs generated using Monte Carlo methods is also obtained. Using assumed "ideal" diffusion tensor data within an imaging volume, known Gaussian noise is added in quadrature to the NMR echo, as described elsewhere [21]. A wide range of imaging parameters and b-matrices was simulated, including the imaging parameters of the human data acquisitions described below, and, for each, a set of noisy DWIs was generated.

3.1.2. Diffusion weighted image (DWI) acquisition

Healthy volunteers were scanned using a 1.5 T GE Signa Horizon EchoSpeed scanner equipped with a 2.2 G/cm gradient set, according to an approved NIH clinical protocol. A set of DWIs were acquired in six isotropically distributed directions, using an interleaved, spin-echo, echo-planar sequence, employing navigator echo correction, as described in [31]. For each direction the measurement was repeated $n = 4$ times. DWI parameters were as follows: FOV = 22 cm, TE = 78 ms, TR > 5 s with cardiac gating, voxel size = 3.5 mm × 1.75 mm × 1.75 mm, data matrix = 128 × 128. The strength of diffusion weighting as measured by trace(**b**) (where **b** is the b-matrix [25,32–34] calculated for each DWI) was varied from approximately 0 to 1000 s/mm^2.

To validate the bootstrap approach we performed DT-MRI scans on two different uniform phantoms, one performed on the 1.5 T scanner described above, and the other on a 3 T GE scanner with a 4 G/cm gradient set. The uniform phantom was a cylindrical glass container filled with a highly viscous polyvinyl alcohol (PVA) solution (1% concentration) with uniform diffusion properties. We used different numbers of gradient directions, each having a large number of repetitions. The phantom data shown in this paper were acquired using 12 gradient directions plus a non-diffusion weighted set, each repeated 14 times.

3.1.3. Estimation of the diffusion tensor from DWIs

The effective diffusion tensor, $\mathbf{D}^{\text{eff}}$, is estimated in each voxel from DWIs in which diffusion gradients have been applied in at least six oblique, non-collinear directions, as specified by the symmetric 3×3 b-matrix, **b** [25]. This quantity should not be confused with the experimental B-matrix, which is $6 \times N_{\text{M}}$ matrix, where N_{M} is the number of measurements containing the six independent elements of **b** in each row. From the B-matrix and the N_{M} measured diffusion-weighted signals, $\mathbf{D}^{\text{eff}}$ is estimated using weighted multivariate linear regression (*regress* routine in IDL, Research Systems), or Levenberg–Marquardt nonlinear regression [25]. Theoretically, the estimated parameters are expected to conform to a Multivariate Normal distribution for both regressions, and approximate expressions for the errors in the estimates can be derived, which can be found in many textbooks on regression analysis and are not repeated here [35,27,28].

There are many ways to design a DT-MRI experiment since one has control over the diffusion gradient strength, direction, and gradient pulse parameters, which are all embodied in **b**. In order to use the bootstrap the DWI acquisition must be repeated several times for each choice of **b**, as described below.

3.2. DT-MRI bootstrap analysis

Fig. 2 illustrates our implementation of the bootstrap in DT-MRI indicating how to perform the bootstrap analysis for a particular DT-MRI experimental design which consists of N_{M} DWIs defined by the corresponding b-matrices, **b**, and in which there are M non-collinear diffusion gradient directions. Together with the b-matrix with zero diffusion weighting, this yields $N_{\text{S}} = M + 1$ distinct **b**. For the ith distinct b-matrix, $\mathbf{b}_i$, the measurement is repeated n_i times, designated as subset S_i. All such subsets, $S_i, i = 0, \ldots, M$ constitute the *original design* of the experiment and thus $N_{\text{M}} = \sum_{i=0}^{M} n_i$. In some designs all of the b-matrices can be distinct ($n_i = 1$, for all i); however, to perform bootstrap analysis at least some of the b-matrix subsets must have $n_i > 1$. In our bootstrap implementation we require all b-matrix subsets to have more than one image acquired, and thus, in many situations, the actual *bootstrap measurements* will not conform to the original measurement design.

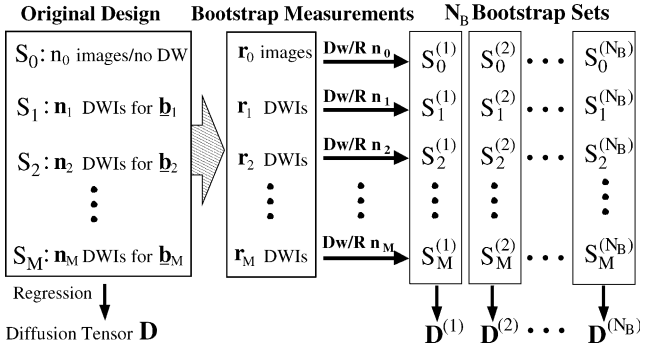

Fig. 2. A graphical illustration of the proposed DTI bootstrap. In this scheme, the *original design* to be studied needs to be extended in some situations to accommodate bootstrap analysis. The *bootstrap measurements* that need to be acquired preserve the number of distinct b-matrix groups, or subsets S_i (total number is $M+1$). The only difference between the original design and the bootstrap measurements design is that the number of repetitions for each group is usually increased and is now r_i, instead of n_i. From this pool of r_i bootstrap measurements, n_i are drawn with replacement (Dw/R) for each subset S_i to form a single *bootstrap set*. This Dw/R is repeated N_B times.

The difference is that the number of repeated acquisitions for each subset will now be r_i, such that $r_i \geqslant n_i$ and we require, $r_i > 1$ for all i. Note, if the original design uses sufficiently large n_i values, no additional images need to be acquired just for the sake of the bootstrap analysis.

Naturally, the reliability of the bootstrap estimate increases monotonically with r_i. From the bootstrap measurements the N_B individual *bootstrap sets* will be drawn randomly with replacement (Dw/R), each matching the structure of the original set, i.e., the subset S_i in the bootstrap sets will contain exactly n_i measurements ($r_i = n_i$ for all i). This is important since the statistical properties of DT-MRI data are dependent on the values of n_i and the total number of images acquired.

Since acquiring additional repeated images for performing bootstrap analysis is expensive, we are interested in knowing how to choose optimal values for r_i. We postulate that with an increasing number of different gradient directions fewer repeated images are needed and that the total number of distinct combinations that a given setup in Fig. 2 can produce is a measure of the richness of the bootstrap samples. We tested this hypothesis with extensive MC simulations of the DT-MRI bootstrap methodology in which the number of different gradient directions, M, as well as the number of repetitions, $n_i = n$ for all i, were varied. The bootstrap estimates were made repeatedly, and with each new MC evaluation, the original diffusion tensor was rotated randomly.

Although, the bootstrapping procedure described here is usually performed on any single voxel within an imaging volume, it can be extended to include several voxels in a ROI, which we generally call *ROI bootstrap*. We distinguish two different implementations of the ROI bootstrap: (i) The *compound ROI bootstrap* (C-ROI bootstrap) and (ii) the *pooled ROI bootstrap* (P-ROI bootstrap).

The C-ROI bootstrap is used when the ROI is assumed to have homogeneous diffusion and statistical properties. Thus the r_i images acquired for each subset S_i will actually be a collection of all corresponding DW values for the voxels in that ROI. The estimation power of the standard error and other statistics, using the C-ROI bootstrap is much greater than that of the simple ROI analysis. If the ROI is not homogeneous, the C-ROI cannot determine the underlying probability density distribution of the voxels in the ROI due to the Central Limit Theorem, i.e., the bootstrap estimate of such distribution is biased to be Gaussian. This property of the C-ROI bootstrap can be used to study the power of the single voxel bootstrap in detecting deviations from Normality.

To determine the underlying distribution in the ROI qualitatively, one should use the P-ROI bootstrap, where the bootstrap estimates from each voxel in the ROI are pooled together. In the limit of large number of bootstrap replicates, N_B, this is equivalent to choosing a voxel randomly first and then Dw/R only the DW values for that voxel. We use the P-ROI bootstrap only qualitatively, since pooling the bootstrap samples from several different original samples will overestimate variability. In the case of a simple bootstrap for an ROI that has identically distributed voxels the estimate of standard deviation is larger by a factor of $\sqrt{2}$ (41%).

Although, both the single voxel bootstrap methodology, and the C-ROI bootstrap, tend to give Normally distributed estimates of the underlying probability density function (PDF) of $\mathbf{D}^{\text{eff}}$, they can be used to estimate the probability distribution of the quantities derived from the diffusion tensor, which are often not Normally distributed. We use the bootstrap methodology, as well as MC simulations, to investigate the probability distributions of the eigenvalues of the diffusion tensor, $\lambda_1, \lambda_2, \lambda_3$ (ordered by magnitude), of trace($\mathbf{D}^{\text{eff}}$), and of a measure of diffusion anisotropy, the relative anisotropy (RA) [36]

$$\text{RA} = \sqrt{\text{Var}(\lambda)}/\langle\lambda\rangle, \tag{4}$$

where $\langle\lambda\rangle = \text{trace}(\mathbf{D}^{\text{eff}})/3$, is the average of the three eigenvalues, and $\text{Var}(\lambda)$ is the eigenvalues variance [37]

$$\text{Var}(\lambda) = \frac{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}{3}. \tag{5}$$

All these quantities derived from the diffusion tensor are local, voxel-based. We consider the non-local, region-based quantities in Section 5.

3.3. Testing Normality

To determine whether a Multivariate Gaussian distribution is the appropriate parametric statistical model

of ideal diffusion tensor data, we generated noisy diffusion tensor data using the MC simulations of the DT-MRI experiments in which the noise in the DWIs conforms to a Rician distribution. Specifically, we test whether $\mathbf{D}^{\text{eff}}$ conforms to the parametric model given in Eq. (2). We simulated different types of diffusion processes. Both isotropic and anisotropic diffusion tensors with values within the range found in living human brain were used [38]. A wide range of S/N values (defined as the S/N value for the unweighted signal, trace($\mathbf{b}$) = 0), was used (1–100); other imaging parameters were explored too (e.g., the number of images acquired for regression N_{M} was varied from 7 to 100).

If the eigenvalues of the tensor conform to a Normal distribution then we expect $\text{Var}(\lambda)$ in Eq. (5) and RA^2 to conform to the scaled central chi-square distribution, i.e., $\chi^2(sx, v)$. Here, v is the number of degrees of freedom, and the scale parameter s accounts for the fact that $\text{Var}(\lambda)$ is not normalized, and that RA^2 is normalized by $\langle\lambda\rangle^2$ instead of having each individual term normalized by its σ^2 in the standard definition of χ^2. Throughout the rest of the paper we will call this distribution simply the χ^2_s distribution. Note that, even if the distribution of eigenvalues slightly deviate from Normality, $\text{Var}(\lambda)$ and RA^2 could still conform to the χ^2_s distribution. We test this hypothesis using Monte Carlo simulations.

Finally, we compared the sample covariance matrix elements to the theoretically predicted values using standard linear regression error analysis [25].

3.4. Estimating true variability in the tissue

Because diffusion properties can vary from voxel to voxel, the spatial heterogeneity can be a source of statistical variability in an ROI. The single voxel bootstrap methodology provides intravoxel estimates of experimental noise, which can be used with the ROI estimates of noise to obtain the true variability of various DT-MRI derived parameters. Since the spatial variability and experimental variability of DWIs are assumed to be statistically independent we can express the true standard deviation of a property of the tissue as

$$\sigma_{\text{T}} = \sqrt{\sigma^2_{\text{ROI}} - \sigma^2_{\text{E}}}, \tag{6}$$

where σ_{ROI} is the ROI estimate of the standard deviation, and σ_{E} is the effective experimental intravoxel standard deviation for the ROI. If σ_i is the standard deviation for the ith voxel in the ROI then σ_{E} can be expressed as

$$\sigma_{\text{E}} = \sqrt{\frac{1}{N_{\text{ROI}}}\sum_{i=1}^{N_{\text{ROI}}}\sigma_i^2}, \tag{7}$$

where N_{ROI} is the number of voxels in the ROI. In practice, the σ_i are the bootstrap estimates of the standard deviation within each voxel. This equation is generally valid for situations with independent sources of noise. However, when the actual σ_{T} is negligible, negative arguments of the square root in Eq. (6) can appear, due to the randomness of the estimates of σ_{ROI} and σ_{E}, which are now close to each other. Zeroing the argument of the square root for such cases produces a biased estimate.

If σ_{E} is estimated using the DT-MRI bootstrap, static systematic errors that do not change as measurements are repeated (e.g., Nyquist ghosts, eddy-currents, etc.) will not be captured and will cause the true tissue variability to be overestimated in Eq. (6). If the spatial variability of those static systematic errors, σ_{S}, could be measured independently then the σ_{E} used in Eq. (6), $\sigma_{\text{E}} = \sqrt{\sigma^2_{\text{B}} + \sigma^2_{\text{S}}}$, where σ_{B} is the bootstrap estimate of experimental error. In this work we perform measurements on phantoms with uniform diffusion properties, which can give us an estimate of σ_{S}. We do not have an estimate of σ_{S} for human data and, thus, the values of σ_{T} we report in Section 4 are overestimated.

4. Results

4.1. Monte Carlo simulation results

Monte Carlo simulations of DT-MRI experiments demonstrate that the six independent components of $\mathbf{D}^{\text{eff}}$ are distributed according to a Multivariate Gaussian. In Fig. 3 we juxtapose empirical probability distributions of the diffusion tensor components (open circles) and Multivariate Gaussian fits (solid lines) for $\boldsymbol{\mu} = \{1104, 661, 329, 180, 135, 77\}\,\mu\text{m}^2/\text{s}$, S/N = 20, and 28 DW images. Monte Carlo simulations gave similarly good fits to the Multivariate Gaussian distribution for all paradigmatic tissue types in human brain, and all simulated experimental setups, even in cases when S/N is slightly less than 2 and only 7 images are used in the regression. All the cases in which the deviation from Normality is observed are not of practical interest (S/N $\approx$ 1, and $N_{\text{M}} = 7$).

We obtain estimates of the sample covariance matrix empirically, using bootstrap analysis or Monte Carlo methods, and analytically, using a linear regression model prediction given by Eq. (3). The linear model estimates of log-linearized data are biased at low S/N, particularly the variances of the diagonal elements of the tensor. We found that the standard linear regression predictions underestimate the diagonal elements of the covariance matrix by approximately 20% for S/N = 20. The prediction for the standard deviation of the trace is approximately one half of that determined from the samples.

Our ability to predict $\boldsymbol{\Sigma}$ degrades as diffusion attenuation (i.e., trace($\mathbf{b}$)) increases to the point where the NMR signal approaches the background noise level

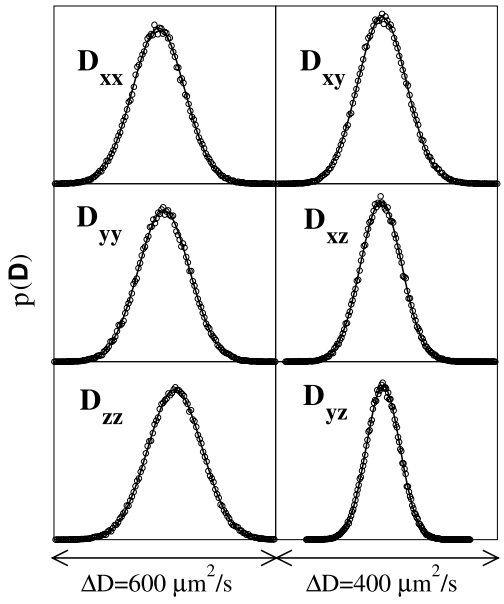

Fig. 3. Probability distributions of the components of $\mathbf{D}^{\text{eff}}$ obtained from Monte Carlo simulations. The mean vector is $\boldsymbol{\mu} = \{1104, 661, 329, 180, 135, 77\}\,\mu\text{m}^2/\text{s}$, S/N = 20, and $N_{\text{M}} = 28$. The solid line indicates a fit to a Gaussian distribution. The distributions represent a histogram obtained using 50,000 Monte Carlo replicates of a DT-MRI experiment. ΔD is the range of diffusion values displayed in the graph.

(very low S/N). This is because in the low S/N regime, the log-linearized data deviate significantly from the linear model with additive noise. The use of weighting factors for correcting this [26] is only approximate, and does not completely restore the predictive power of the linear regression formula. Thus, it is important to have a non-parametric estimate of errors, such as bootstrap estimates.

Fig. 4a shows the Monte Carlo simulations of the slightly anisotropic diffusion tensor ($\lambda_1 = 750\,\mu\text{m}^2/\text{s}$, $\lambda_2 = 700\,\mu\text{m}^2/\text{s}$, $\lambda_3 = 650\,\mu\text{m}^2/\text{s}$, and the corresponding RA = 0.058) for three different experimental situations in terms of the number of gradient directions, and the number of repetitions, each having signal-to-noise ratio, S/N = 25. The empirically obtained probability density function (PDF) for the square of the relative anisotropy, RA^2, is fitted to the χ^2_s distribution with the degree of freedom parameter, v, and the scale parameter, s. Similarly, Fig. 4b shows the MC simulations of the same experimental situations but for a more anisotropic tensor ($\lambda_1 = 1050\,\mu\text{m}^2/\text{s}$, $\lambda_2 = \lambda_3 = 525\,\mu\text{m}^2/\text{s}$, and the corresponding RA = 0.35).

Fig. 5 shows the true probability density function (a) of trace($\mathbf{D}^{\text{eff}}$) and (b) of a tensor component D_{xx} (thick dashed line), where 8 repetitions were used for each of the 6 + 1 directions. Juxtaposed are two bootstrap estimates of the corresponding PDFs of trace($\mathbf{D}^{\text{eff}}$) and D_{xx} obtained from two different samples of DWIs used for bootstrapping. This figure indicates that the bootstrap gives reliable estimates of the variance and the shape of

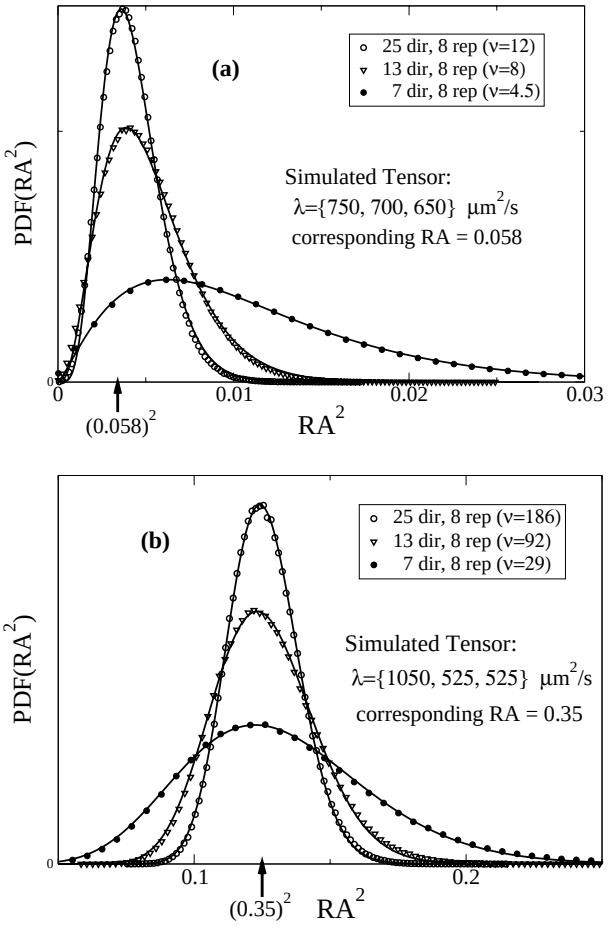

Fig. 4. Probability distributions of the square of the relative anisotropy (RA). The figures were obtained by fitting the χ^2_s distribution to the empirical density distributions obtained from Monte Carlo simulations of (a) a nearly isotropic tensor with eigenvalues $\{750, 700, 650\}\,\mu\text{m}^2/\text{s}$, (RA = 0.058, $\text{RA}^2 = 0.0034$) and (b) an anisotropic tensor with eigenvalues $\{1050, 525, 525\}\,\mu\text{m}^2/\text{s}$, (RA = 0.35, $\text{RA}^2 = 0.125$).

the PDF, but not the mean. This result is expected since the mean of a given statistic depends on the mean of the original bootstrap sample, and on the intrinsic bias of a given statistic. The diffusion tensor components and trace are not biased and the peak of the PDF in Fig. 5 coincides with the estimate of the original sample. On the other hand, the peak in the PDF for RA does not correspond to the RA of the original sample, since RA is biased statistic. Also, the spread of the peaks of the individual PDFs is an illustration of why P-ROI produces biased (larger) estimate of the true variability, because it includes the same variability twice.

We also used MC simulations to study the relationship between the quality of the bootstrap estimates and the number of gradient directions and measurement repetitions. We simulated experimental setups with the number of directions chosen from {7,10,13,25,61,113} and the number of repetitions from {2,3,4,5,7,10,14,20}. The results we show are for the simulated data with eigenvalues $\{800, 700, 600\}\,\mu\text{m}^2/\text{s}$ and with S/N = 25, but

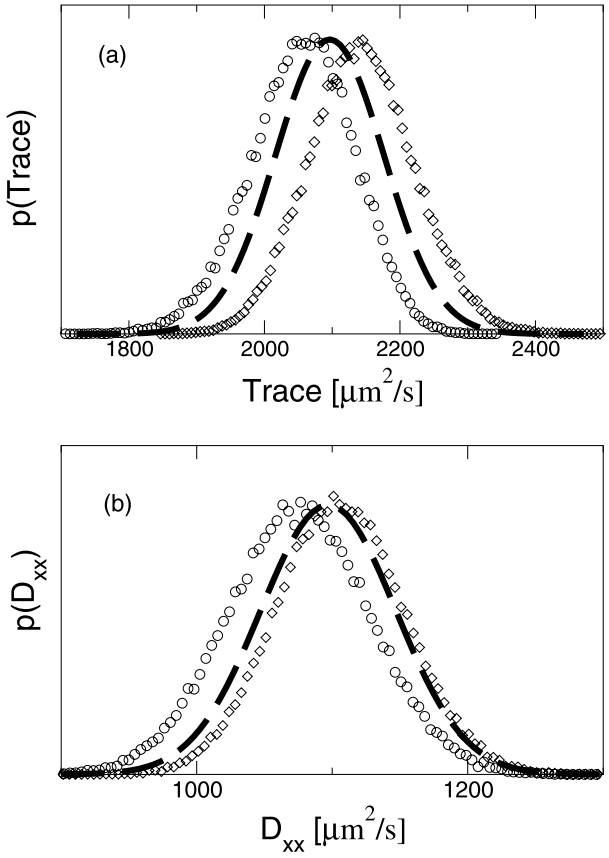

Fig. 5. Comparison of the true distribution (thick dashed line) of (a) trace($\mathbf{D}^{\text{eff}}$) and (b) the diffusion tensor component, D_{xx}, to two independent bootstrap estimates of the same distribution (circles and diamonds). The means of the distribution differ but the bootstrap estimates of the standard error and the shape of the distribution are in good agreement. The original DWI sample used for bootstrap analysis was simulated using a typical experimental B-matrix and a diffusion tensor with $\boldsymbol{\mu} = \{1100, 0, 0, 500, 0, 500\}\,\mu\text{m}^2/\text{s}$, i.e., a tensor with $\text{trace}(\mathbf{D}^{\text{eff}}) = 2100\,\mu\text{m}^2/\text{s}$.

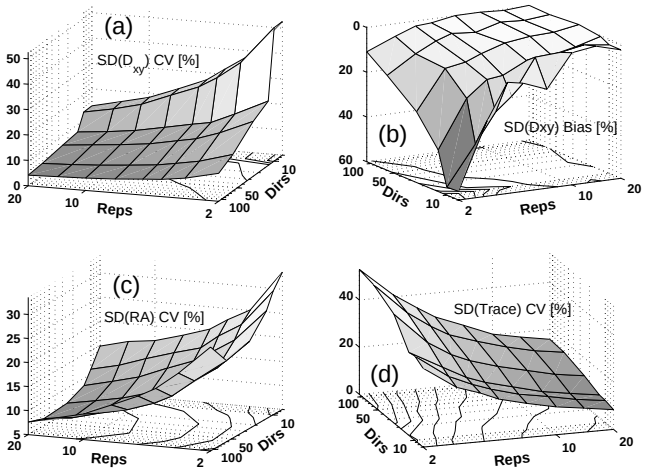

Fig. 6. How the quality of the bootstrap estimates depends on the number of different gradient directions and the number of repetitions for each of the directions. Figures (a), (c), and (d) show the dependence of the precision of a given bootstrap estimate on a given statistic, θ_{B}, expressed through $\text{CV} = \text{SD}(\theta_{\text{B}})/\theta_{\text{B}}$. The statistics used are (a) standard deviation of an off-diagonal element, $\text{SD}(D_{xy})$; (c), standard deviation of relative anisotropy, SD(RA); and (d) standard deviation of the trace, SD(trace). Figure (b) shows the bias of the estimates, $\text{bias} = (\theta_{\text{B}} - \theta_{\text{B}}^{\text{true}})/\theta_{\text{B}}^{\text{true}}$ expressed as a percent.

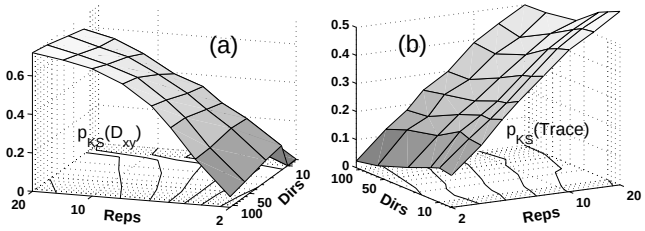

Fig. 7. Another measure of the quality of the bootstrap estimate. The measure is the p-value obtained from the Kolmogorov–Smirnov test when comparing (a) PDF(D_{xy}) and (b) PDF(trace). Note that the larger the p-value, the more accurate the bootstrap estimate.

similar results are observed for a wide range of simulated tensors and experimental parameters. As a measure of the quality of the bootstrap estimate of a given statistic, θ_{B} we chose the coefficient of variation defined as $\text{CV}(\theta_{\text{B}}) = \text{SD}(\theta_{\text{B}})/\text{E}(\theta_{\text{B}})$. The standard deviation and the expectation values were determined empirically using many Monte Carlo repetitions of the experiment. Our hypothesis that with an increasing number of different gradient directions, fewer r_i images are needed does not hold for all of the statistics, e.g., for the SD(trace) in Fig. 6d. As additional verification of this finding, we use another measure of the quality of the bootstrap estimates, which compares the bootstrap estimate of the PDF of a given statistic to the true one (empirically determined using MC). The measure we plot is the it p-value obtained from the Kolmogorov–Smirnov (KS) test when comparing the two distributions, shown in Fig. 7 for $\text{SD}(D_{xy})$ and SD(trace). The new measure is consistent with the previous one.

4.2. Bootstrap estimation results

4.2.1. Uniform phantom

Fig. 8 shows the images of the trace for two different PVA phantoms. Fig. 8c and d demonstrate that the measured trace($\mathbf{D}^{\text{eff}}$) of the uniform phantoms was in fact not very uniform. In all attempts we were unable to obtain phantom data with sufficient uniformity, suitable for verification of the bootstrap results. These non-uniformities are present to a certain extent in every DT-MRI measurement and are due to susceptibility differences, Nyquist ghosts, partial volume effects, eddy-currents, etc. Thus, it is not surprising that the estimates of SD(trace) obtained from a large ROI in the uniform phantom were much larger than the bootstrap estimates. For the phantom acquired on the 1.5 T scanner the ROI estimates were 50% larger than the single voxel bootstrap estimates, while for the 3 T scanner data the estimates were 150% larger.

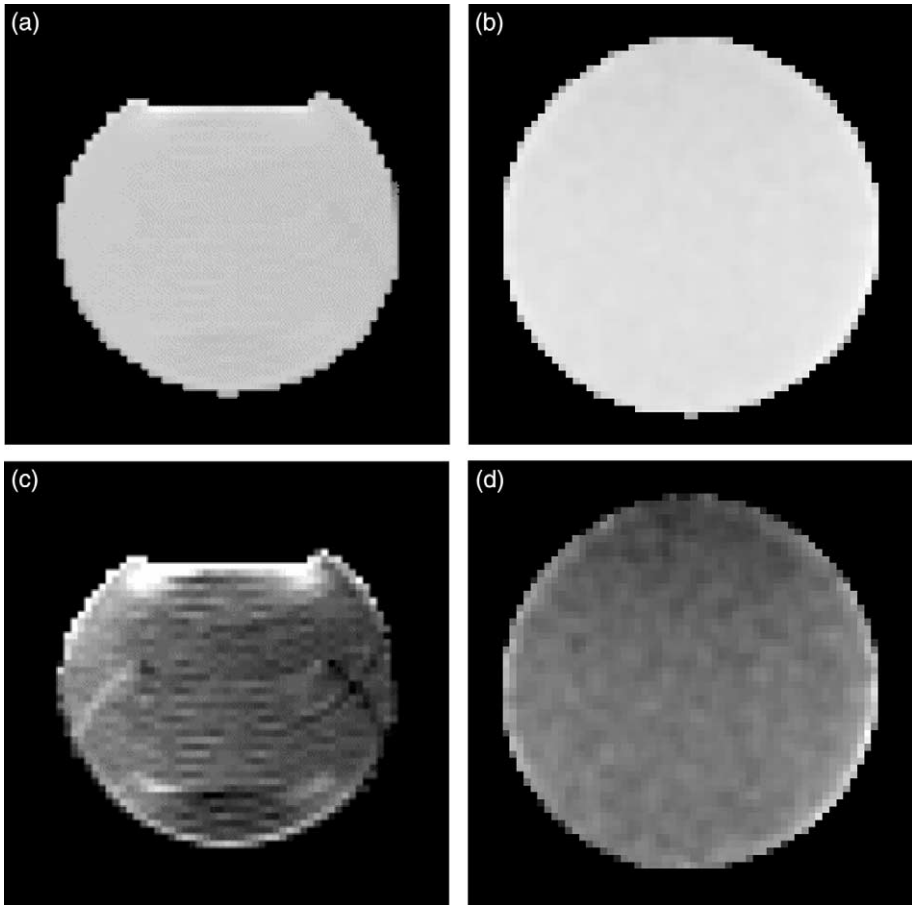

Fig. 8. The trace of the diffusion tensor shown for two different uniform phantoms: (a) acquired with a 3 T scanner and (b) with a 1.5 T scanner. Black corresponds to trace = 0, and white to Max(trace). Figures (c) and (d) show the same data except that black corresponds to the min(trace) value within the phantom. The black-white range in the scaled images (c) and (d) is approximately 10 times narrower than the full range of values shown in (a) and (b). The images illustrate that the "uniform" phantoms were not truly uniform.

To make an indirect validation of the bootstrap approach we studied the dependence of the ROI estimate on the size of the ROI (in terms of the number of voxels within ROI, n_{vox}). Since, for the smaller ROIs, the estimate of SD is imprecise (for Normal data error $= 100\%/\sqrt{2(n_{\text{vox}}-1)}$) we placed many approximately circular ROIs with n_{vox} voxels randomly across the uniform phantom. An average from 300 such estimates is plotted in Fig. 9 (open circles) vs. n_{vox}. We see that extrapolating this curve to $n_{\text{vox}} = 1$ yields an estimate very close to the value of the single voxel bootstrap estimates, whose range is designated by the dashed lines.

4.2.2. Human data: single voxel bootstrap

We also tested the bootstrap method using experimental DT-MRI patient data (repetitions $n = 4$), both with the single voxel bootstrap, and with the ROI bootstrap. We found that the DT-MRI bootstrap estimate of PDF($\mathbf{D}^{\text{eff}}$) in the majority of voxels is a Multivariate Normal distribution. Fig. 10 shows typical results for the white matter voxels, and similar results were observed in gray matter. The cerebrospinal fluid (CSF) voxels, although often conforming to the Normal distribution, were much more likely to deviate significantly from it.

Off-diagonal elements were more likely to deviate from a Normal distribution than the diagonal elements. Off-diagonal elements may be more susceptible to noise because their relationship to the NMR signal is not as direct as that of the diagonal elements. While diagonal elements can be observed with only one diffusion weighting gradient, the observation of the off-diagonal elements requires at least two different gradient directions. We attribute deviations from Normality in the diffusion tensor elements to motion and other systematic artifacts, which are exogenous to the Rician noise model, and difficult to describe parametrically.

Fig. 11 shows estimates of the PDF for trace($\mathbf{D}^{\text{eff}}$) in gray matter, white matter, and CSF. The gray and white matter distributions are very well described by a univariate Gaussian [39]. The CSF distributions in many situations qualitatively fit the Normal distribution, but are rejected by rigorous tests of Normality, such as the KS test, including the distribution shown in this figure.

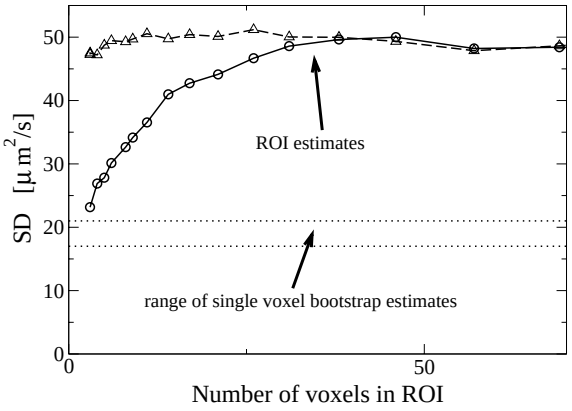

Fig. 9. Plot of the estimates of SD of trace for the uniform phantom acquired on a 3 T scanner. The ROI estimate of SD for a given ROI size is actually an average estimate collected from 300 different ROIs of the same size and placed randomly within the uniform phantom. The straight, parallel, dotted lines in the figure represent a range of the bootstrap estimates of SD(trace) collected over 20 randomly selected voxels. The dashed line with open triangles represents test results obtained by applying the same procedure to a simulated, truly homogeneous ROI with simulated $\sigma = 50\,\mu m^2/s$ which indicates that the significant drop in the value of estimated SD is not an artifact of the procedure.

Similarly, Fig. 12 shows estimates of the PDF for trace($\mathbf{D}^{eff}$) for the three tissue types, and for the cases where clear deviation from Normality is observed. These cases illustrate how the bootstrap methodology could be used for detecting systematic artifacts in DT-MRI data. In white matter voxels that we studied this deviation

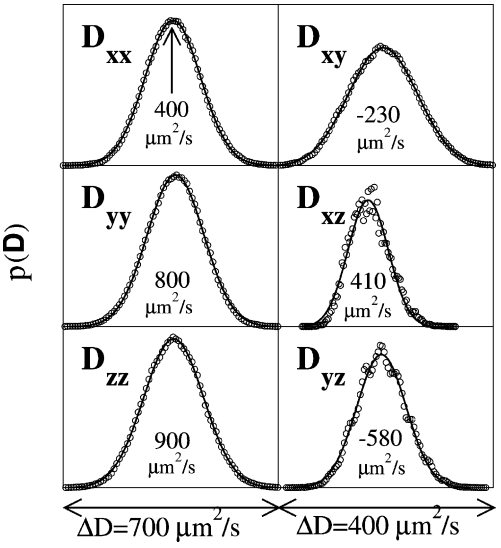

Fig. 10. The probability density function of the diffusion tensor for a voxel within white matter, obtained by applying the single voxel bootstrap to the human data. This is a typical result, for both, white and gray matter voxels, for which deviation from Normality is rarely observed. The solid line indicates fit to a Gaussian distribution. The numbers indicate the mean value of a given diffusion tensor component in $\mu m^2/s$. The ΔD is the range of diffusion values displayed in the graph.

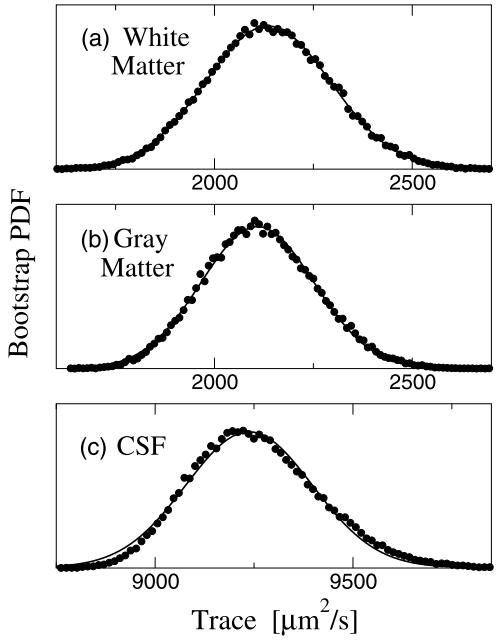

Fig. 11. A single voxel bootstrap estimate of the probability density function of the trace of the diffusion tensor. The graphs indicate that the trace($\mathbf{D}^{eff}$) is Normally distributed. This is true for most, but not all voxels. The CSF bootstrap analysis was much more likely to deviate from Normality than the gray or white matter voxels.

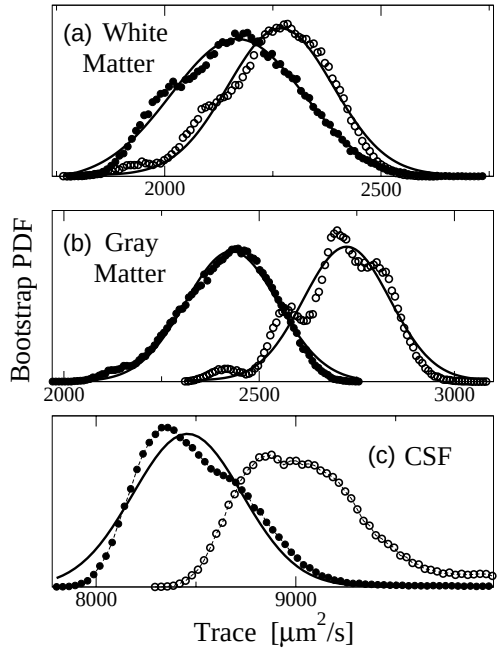

Fig. 12. A single voxel bootstrap estimate of the probability density function of the trace of the diffusion tensor for three different tissue types and for the voxels that clearly deviate from Normality. For each tissue type two examples are shown. The solid lines indicate fits to a Gaussian distribution to a given PDF. The fit to CSF data represented by open circles is not shown because it failed due to a long tail of this distribution, extending up to the trace value of $12{,}000\,\mu m^2/s$.

occured in approximately 2% of voxels studied, in gray matter for 5% of voxels, and in CSF for 25% of voxels.

4.2.3. Human data: ROI bootstrap

Fig. 13 shows some of the ROIs used for the ROI bootstrap. In Fig. 14a the probability distribution of trace($\mathbf{D}^{\text{eff}}$) obtained within a single voxel, located approximately in the center of the ROI CC, is fit very well by a Gaussian with $\sigma = 130\,\mu m^2/s$. This ROI CC is drawn in corpus callosum, and in close proximity to cerebrospinal fluid (CSF) to include the partial volume effects. Fig. 14b shows a probability distribution of trace($\mathbf{D}^{\text{eff}}$) obtained with the C-ROI bootstrap and the P-ROI bootstrap. The ROI standard deviation of the trace is $\sigma_{CC} = 902\,\mu m^2/s$, which is much larger than any of the single voxel bootstrap estimates of $SD(\text{trace}) = 116 \pm 24\,\mu m^2/s$, for the voxels in the ROI CC. Both the C-ROI and the P-ROI bootstrap estimates of the PDF(trace) reveal the existence of different statistical modes, which explains the high value of the ROI estimate of σ.

Fig. 15 shows the results obtained using single voxel, the C-ROI and the P-ROI bootstrap analysis of D_{xy}. The ROIs are drawn carefully within purportedly uniform white matter (internal capsule, see the ROIs IC1 and IC2 in Fig. 13). Fig. 15c also demonstrates that the C-ROI bootstrap does not reveal the bimodal distribution, when both IC1 and IC2 ROIs are analyzed simultaneously; instead, it produces an effective Normal distribution (thick solid line). The large spread of each mode, in comparison to a single voxel estimate, indicates that, in general, it is very difficult to find an ROI which has homogeneous diffusion properties, hence, it is rarely advisable to perform ROI analysis of diffusion tensor elements. The situation improves somewhat when the

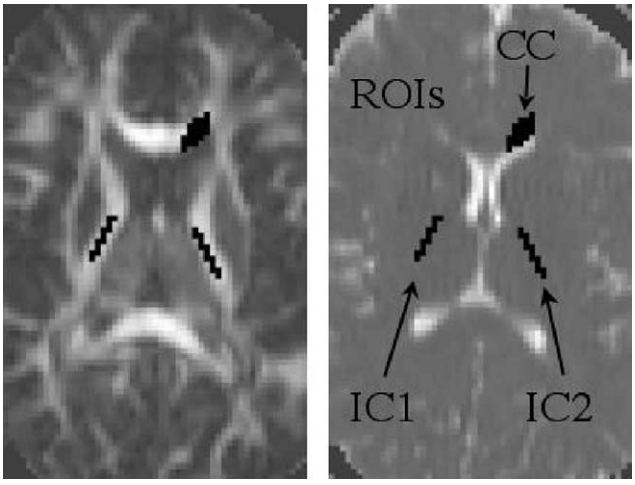

Fig. 13. ROIs used for the demonstration of the ROI bootstrap displayed as black areas on images of lattice anisotropy index (left) and trace (right). The ROI CC (labeled on the trace image only, but displayed on both images) is selected within corpus callosum and close to CSF and is intentionally drawn sloppily to increase the partial volume effects. It contains 20 voxels. The ROIs IC1 and IC2 are drawn carefully within the internal capsule on both hemispheres. The IC1 ROI contains 12 voxels and IC2 contains 13 voxels.

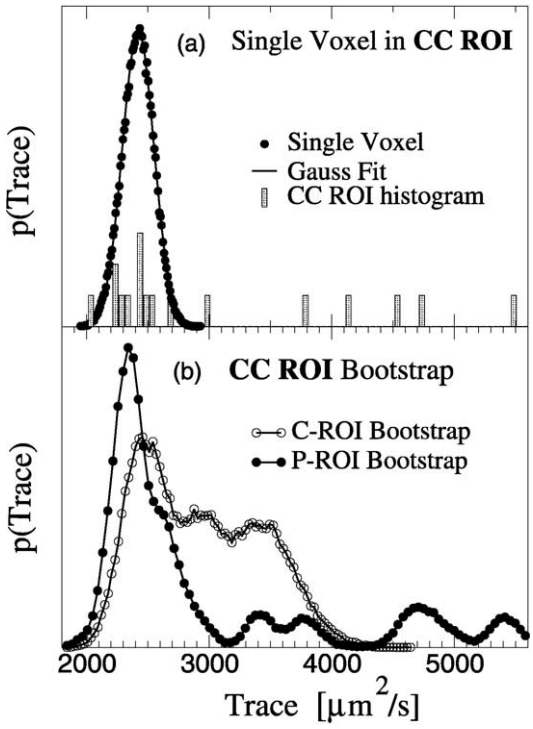

Fig. 14. Comparison of the probability density function of trace($\mathbf{D}^{\text{eff}}$) for (a) single voxel bootstrap and (b) C-ROI and P-ROI bootstrap for voxels within in the corpus callosum, but close to CSF. The bars in (a) represent the histogram of the trace values in CC ROI (three of the voxels fall outside the displayed range). The ROI bootstrap analysis reveals the existence of different statistical modes, which suggests that any ROI statistic, such as ROI standard deviation, is very inaccurate.

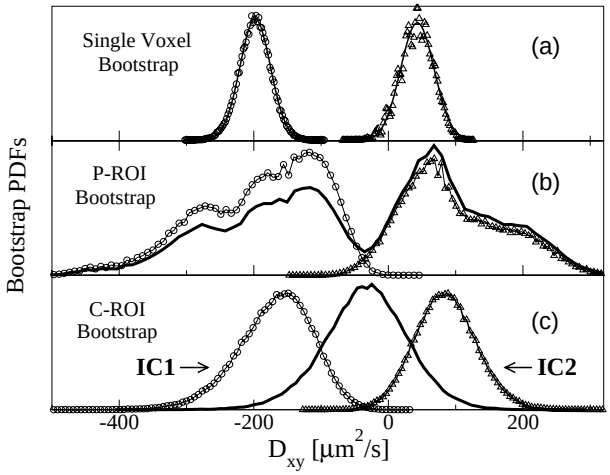

Fig. 15. A bootstrap estimate of the probability density function of diffusion tensors in an ROI drawn within the internal capsule. The open circles designate analysis on IC1; open triangles designate analysis on IC2. The solid thick line in (b) and (c) indicates analysis of the compound ROI (ICTOT) consisting of all voxels in IC1 and IC2. Note that the C-ROI analysis on ICTOT produces effective Normal distribution, despite the underlying bimodal distribution in ICTOT.

rotationally invariant quantities are used, but they also show a significant spread when the ROI bootstrap is used. Fig. 16 shows results of applying bootstrap analysis to investigate the distribution of the eigenvalues of the diffusion tensor for voxels and ROIs containing white matter, gray matter, and CSF. The C-ROI estimates of SE of the rotationally invariant quantities (eigenvalues) also overestimates the single voxel estimates, indicating the presence of non-homogeneities in the used ROIs. In isotropic media, such as gray matter and CSF, there is a significant difference among the means of the three eigenvalues. Moreover, there is negative skewness of the distribution of the largest eigenvalue, and the positive skewness of the smallest eigenvalue, although this is not clearly visible on the graph.

Fig. 17 shows the bootstrap estimates of the probability distribution of the square of relative anisotropy, RA^2 (see Eq. (4)), a voxel-based measure of anisotropy. Some other measures of anisotropy, such as the lattice index[21] which are not voxel-based, are more difficult to analyze using the bootstrap methodology. The distribution of RA^2 is well-described by the χ^2_s distribution, as confirmed using the KS test for this distribution. This experimentally confirms the results of MC simulations of DT-MRI measurements, showing that the $\mathrm{Var}(\lambda)$ in Eq. (5), and the RA^2, conform to the χ^2_s distribution (see Fig. 4).

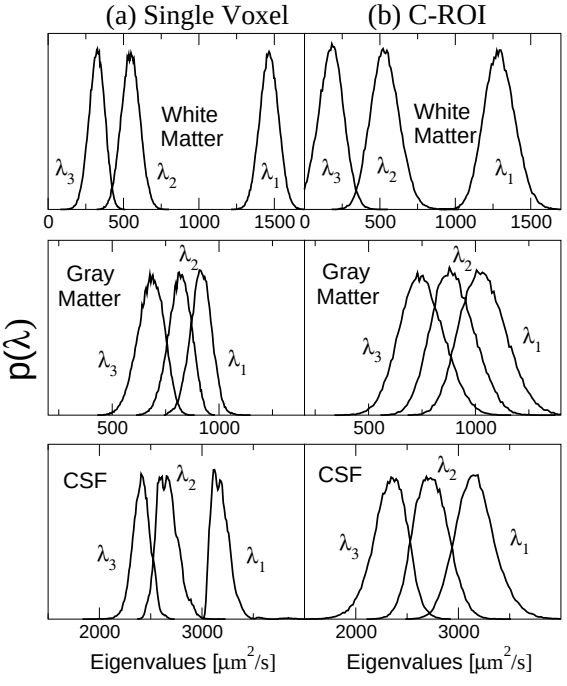

Fig. 16. Bootstrap estimate of the probability density function of diffusion tensor eigenvalues within (a) a single voxel and (b) corresponding ROIs (the white matter ROI was IC1, shown in Fig. 13). It is evident that the ROI estimates of variability and those of a single voxel disagree. This disparity points to a significant intrinsic variability of the rotationally invariant quantities.

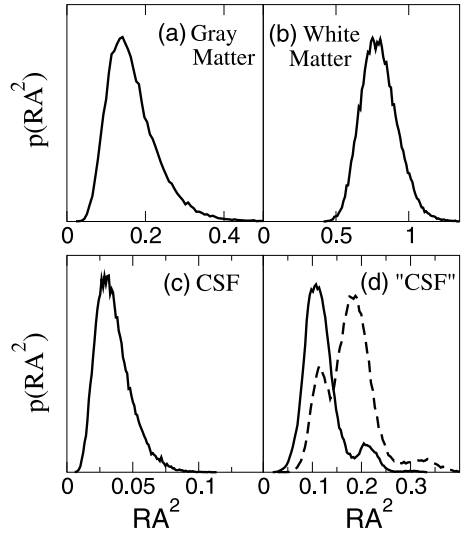

Fig. 17. Single voxel bootstrap estimates of the probability density function of RA^2 for the main tissue types: (a) gray matter; (b) white matter; and (c) CSF. (d) Two examples where the voxel suffers from motion or other artifacts and does not conform to the Multivariate Normal Distribution.

The distribution in Fig. 17, obtained within the CSF compartment, deviates significantly from the expected χ^2_s distribution; this is attributed to artifacts in the original DWI data, which could be due to CSF flow and/or the decreased sensitivity of the DTI experiment to measure accurately the diffusion of CSF.

4.2.4. Estimates of tissue variability

Using the bootstrap analysis we can assess the variability of parameters derived from the diffusion tensor, by using Eq. (6), with the caveat that those estimates could contain the systematic artifacts described above. We selected many ROIs in different regions of the brain and measured the σ_T and σ_E for several rotationally invariant quantities. If the ROI is drawn carefully, the experimental error is comparable to the variability in the tissue. For example, the standard deviations we obtained from the trace bootstrap analysis on these ROIs are: in corpus callosum $\sigma_T = 248 \pm 22\,\mu m^2/s$ and $\sigma_E = 122 \pm 5\,\mu m^2/s$, in gray matter regions $\sigma_T = 280 \pm 25\,\mu m^2/s$ and $\sigma_E = 110 \pm 6\,\mu m^2/s$, in CSF $\sigma_T = 410 \pm 95\,\mu m^2/s$ and $\sigma_E = 225 \pm 11\,\mu m^2/s$. In a region that is drawn across both white and gray matter (brain parenchyma) we obtain a statistically insignificant change of true variability i.e., $\sigma_T = 270 \pm 24\,\mu m^2/s$ and $\sigma_E = 120 \pm 5\,\mu m^2/s$. This is a statistical confirmation of the previous finding that the trace is uniform in human brain parenchyma [38]. Although the σ_T in CSF appears larger than in the parenchyma, note that the percent variability of the trace in CSF is smaller owing to a five times larger trace ($10{,}000\,\mu m^2/s$ vs. $2100\ \mu m^2/s$). These results demonstrate the use of Eq. (6). A more detailed

study on the intrinsic tissue variability of the variety of the diffusion tensor based quantities is under way.

5. Discussion

Knowing the parametric distribution of the tensor enables us to design and apply a variety of statistical tests to assess the significance of their differences between voxels or their spatial variability within different tissue regions. For example, the Multivariate Normal distribution of $\mathbf{D}^{\text{eff}}$ implies that trace($\mathbf{D}^{\text{eff}}$) is distributed according to univariate Normal distribution [39]. This allows us to formulate hypotheses about the moments of trace($\mathbf{D}^{\text{eff}}$) within particular voxels or within ROIs, and to test them quantitatively for the first time. Since the first demonstration of trace images [3], it has been widely reported that trace($\mathbf{D}^{\text{eff}}$) is uniform in normal brain parenchyma, drops substantially in ischemic regions during acute stroke and increases significantly in chronic stroke; however, whether these variations in a given region were due to alterations in a tissue's physiologic state, its physical properties, or background noise, is hard to determine. With a parametric distribution for trace($\mathbf{D}^{\text{eff}}$) in hand, we can test clinically or biologically relevant hypotheses, such as whether trace($\mathbf{D}^{\text{eff}}$) is uniform in normal brain parenchyma, whether it drops significantly in ischemic areas, or whether it is significantly elevated in diseased areas of brain parenchyma. This framework has already been described in [39] and illustrates the power of hypothesis testing in characterizing the variability of trace($\mathbf{D}^{\text{eff}}$).

We also confirmed that RA^2 is distributed according to the χ_s^2 distribution, or equivalently the relative anisotropy index, RA, is distributed according to the scaled central χ-distribution $\chi_s(x,v) \equiv \chi_s^2(x^2,v)$. Very good fits to χ_s^2 were obtained for different levels of experimental precision and for different types of diffusion (isotropic and anisotropic). However, the values for the degrees of freedom obtained from the fits are very often not integers, complicating the interpretation of v.

Although our experiments with the uniform phantoms did not produce the desired experimental verification (except indirectly, see Fig. 9), we still consider the bootstrap a reliable methodology for estimating errors and other statistics of DT-MRI quantities. We have verified using extensive Monte Carlo studies that the DT-MRI bootstrap can produce reliable estimates, some of which are summarized in Figs. 5–7. The results on the uniform phantom clearly show that when using Eq. (6) for estimating the true variability of the tissue, the static systematic artifacts that appear consistently in every repeated measurement, such as Nyquist ghosts, will not be interpreted as measurement noise, but will falsely increase the intrinsic variability of the tissue.

The requirement of the DT-MRI bootstrap that the same DWI acquisition be repeated one or more times for each choice of **b**, is a potential drawback since it has been suggested that applying gradients in more than 6 directions is beneficial for estimating the diffusion tensor [18,40]. However, we do not propose using the bootstrap scheme with every measurement, but rather as a special procedure for the evaluation of the noise properties of a given experimental DT-MRI design. Additionally, we showed in Figs. 6 and 7 that when studying protocols which contain many distinct non-collinear directions, for some of the statistics, fewer repetitions will be required for each direction to achieve the same reliability of the bootstrap. But, as can be seen in those figures, a simple relationship between the number of directions, the number of repetitions, and the quality of the bootstrap estimates is not apparent and depends on other details of experimental design (B-matrix, etc.). Finding how a particular experimental design influences the precision of the bootstrap goes beyond the scope of this paper and will be the subject of future work. For example, sudden changes in the quality of the bootstrap estimates when going from 13 to 25 directions may be caused by utilizing different methods for generating a non-collinear set of B-matrices. Additionally, one can question whether the measures of bootstrap quality that we use are appropriate. One conclusion that can be drawn from Figs. 6 and 7 is that increasing the number of repetitions consistently improves the bootstrap estimates.

The bootstrap method is particularly helpful in examining potential artifacts that arise in the acquisition or the analysis of DT-MRI data. For example, the skewness of the eigenvalue distributions, and the bias in their means seen in Fig. 16, are consistent with an artifact described by Pierpaoli et al. [21,38] that arises when eigenvalues are sorted within each voxel by size. Methods attempting to remedy it have been discussed in [41,19].

As Figs. 12, 17, and indirectly Fig. 14, indicate, the Central Limit Theorem does not "hide" some of the more severe deviations from the Rician model. We attribute those deviations to motion and other systematic artifacts, which suggests that testing for a deviation from Normality is a way to detect systematic artifacts in DT-MRI data. But, note that due to the Central Limit Theorem, smaller motion and other artifacts can still be "hidden" in a new effective Normal distribution, as was demonstrated using a bimodal data set and C-ROI bootstrap (Fig. 15).

The bootstrap method is also very suitable for DT-MRI because reliable estimates of measurement errors can be obtained within a single voxel. Even regions thought to be homogeneous may possess heterogeneous diffusion properties. As seen above in Figs. 15 and 16, aggregating voxels within a ROI just increases the SE of

a measured quantity. This is due to the fact that the probability distribution of the measured random variable has two components. One comes from noise in the measurement (*experimental noise*) and the other comes from the inherent spatial variability of the measured quantity within the imaging volume. In DT-MRI these two components have different noise distributions. As reported in Section 4 the true spatial variability (σ_T) for trace($\mathbf{D}^{eff}$) in gray matter, white matter or CSF is comparable to the experimental variability. Hence, when performing the hypothesis tests on ROIs both sources of noise have to be taken into account. In many situations the spatial variability of a given statistic within an ROI is not Normally distributed and thus parametric hypothesis tests cannot be used. However, the bootstrap confidence intervals and hypothesis tests can be constructed in such cases. We do not describe the details of implementation of such tests here since the focus is on proposing and describing the implementation of the bootstrap methodology. We also demonstrated that estimates of the uncertainty and of the probability distributions of measured and calculated variables in DT-MRI can be obtained using bootstrap analysis, provided a sufficient number of images is used in the bootstrap resampling scheme. In fact, without using the bootstrap method within a single voxel, these estimates cannot otherwise be obtained. A standard estimate of the SD of a certain quantity in a DT-MRI study from the n repetitions of an acquisition having the same b-matrix is very imprecise, because typically, $n \ll 10$.

The bootstrap methodology described here can easily be extended to study non-local, region-based quantities. One such quantity is the lattice index [21], another scalar measure of anisotropy that includes the neighboring voxels in its evaluation. Thus, to obtain a bootstrap estimate for a single voxel, all of the neighboring voxels have to be bootstrapped too. In another example, determining the uncertainty of an estimated fiber tract trajectory requires the inclusion of not only the voxels through which a particular fiber tract passes, but also all of the voxels through which the fiber tract could potentially pass due to inherent variability [22]. The safe approach for bootstrapping the non-local measures is to create a bootstrap replicate of the whole imaging volume (all of the voxels). This approach, however, has enormous computational demands and, in many situations, is not viable. The performance and reliability of the bootstrap estimates of the non-local quantities are not studied here.

Generally, the main drawback in using the bootstrap is the computation time required to run it, which is proportional to N_B. For example, for $N_B = 1000$, the bootstrap analysis takes about a half second per voxel on a SPARC ULTRA-60 workstation when implemented in IDL. On parallel computers, however, the bootstrap method can be implemented very efficiently. In the long term, we do not see computation time as a limitation in bootstrap analysis of DT-MRI data.

6. Conclusion

We find parametric and non-parametric statistical methods, particularly used in conjunction, powerful tools in the analysis of DT-MRI data and in the design of DT-MRI acquisitions. We have shown, using Monte Carlo simulations, that the appropriate parametric model for diffusion tensor data is the Multivariate Normal distribution. We also propose an implementation of bootstrap analysis, the DT-MRI bootstrap. We apply it to data from human subjects, confirming that, in the majority of voxels, tensor components are distributed according to a Multivariate Normal distribution. We show how estimates of the probability distributions, outliers, confidence levels, and bias for statistics that arise in DT-MRI can be obtained, and hypotheses tests can be performed using the bootstrap. Using Monte Carlo simulations we confirm that the DT-MRI bootstrap provides reliable estimates of such quantities. The reliability of the bootstrap estimates depends on the number of different gradient directions and the number of repeated measurements for each direction, but the relationship is not simple. In many situations, however, the total number of distinct combinations that bootstrapping allows is a good measure of its reliability, and, in all cases, increasing the number of repetitions improves the performance of the DT-MRI bootstrap. An important finding of our study is that the total intra-voxel variability observed for in vivo diffusion data is not explained solely by the background Johnson noise. Understanding the sources of this additional variability more deeply, whether it be physiological and/or due to systematic artifacts, is an interesting and important problem demanding further study.

Acknowledgments

The authors would like to thank Alan Barnett and Carlo Pierpaoli for providing us with experimental DT-MRI data used for bootstrap analysis, Ferenc Horkay for providing us with the PVA solution phantom, and Richard Shrager for providing the Matlab code used for the *B*-matrix design in some of our simulations.

References

[1] S. Pajevic, P.J. Basser, Non-parametric statistical analysis of diffusion tensor MRI data using the bootstrap method, in: ISMRM Proceedings, 1999, p. 1790.
[2] S. Pajevic, P.J. Basser, Parametric description of noise in diffusion tensor MRI, in: ISMRM Proceedings, 1999, p. 1787.

[3] P.J. Basser, J. Mattiello, D. LeBihan, MR diffusion tensor spectroscopy and imaging, Biophys. J. 66 (1) (1994) 259–267.
[4] P. vanGelderen, H.M. deVleeschouwer, D. DesPres, J. Pekar, P.C. van Zijl, C.T.W. Moonen, Water diffusion and acute stroke, Magn. Reson. Med. 31 (1994) 154–163.
[5] P.J. Basser, Inferring microstructural features and the physiological state of tissues from diffusion-weighted images, NMR Biomed. 8 (7–8) (1995) 333–344.
[6] V.J. Wedeen, T.L. Davis, R.M. Weisskoff, R. Tootell, B.R. Rosen, J.W. Belliveau, White matter connectivity explored by MRI, in: Proceedings of the First International Conference for Functional Mapping of the Human Brain, Paris, 1995, p. P1.36.
[7] P.J. Basser, Fiber-tractography via diffusion tensor MRI (DT-MRI), in: ISMRM Proceedings, Sydney, AU, 1998, p. 1226.
[8] S. Mori, B.J. Crain, P.C.M. van Zijl, 3d brain fiber reconstruction from diffusion MRI, in: Proceedings of International Conference on Functional Mapping of the Human Brain, Montreal, 1998.
[9] D.K. Jones, A. Simmons, S.C.R. Williams, M.A. Horsfield, Non-invasive assessment of structural connectivity in white matter by diffusion tensor mri, in: ISMRM Proceedings, Sydney, AU, 1998, p. 531.
[10] S. Mori, B.J. Crain, V.P. Chacko, P.C.M. van Zijl, Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging, Ann. Neurol. 45 (1999) 265–269.
[11] T.E. Conturo, N.F. Lori, T.S. Cull, E. Akbudak, A.Z. Snyder, J.S. Shimony, R.C. McKinstry, H. Burton, M.E. Raichle, Tracking neuronal fiber pathways in the living human brain, Proc. Natl. Acad. Sci. USA 96 (1999) 10422–10427.
[12] D.K. Jones, A. Simmons, S.C.R. Williams, M.A. Horsfield, Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI, Magn. Reson. Med. 42 (1999) 37–41.
[13] R. Xue, P.C.M. van Zijl, B.J. Crain, M. Solaiyappan, S. Mori, In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging, Magn. Reson. Med. 42 (1999) 1123–1127.
[14] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, A. Aldroubi, In vivo fiber tractography using DT-MRI data, Magn. Reson. Med. 44 (2000) 625–632.
[15] M.E. Bastin, P.A. Armitage, I. Marshall, A theoretical study of the effect of experimental noise on the measurement of anisotropy in diffusion imaging, Magn. Reson. Imaging 16 (1998) 773–785.
[16] J.S. Shimony, R.C. McKinstry, E. Akbudak, J.A. Aronovitz, A.Z. Snyder, N.F. Lori, T.S. Cull, T.E. Conturo, Quantitative diffusion-tensor anisotropy brain MR imaging: normative human data and anatomic analysis, Radiology 212 (1999) 770–784.
[17] K.M. Martin, N.G. Papadakis, C.L. Huang, L.D. Hall, T.A. Carpenter, The reduction of the sorting bias in the eigenvalues of the diffusion tensor, Magn. Reson. Imaging 17 (1999) 893–901.
[18] N.G. Papadakis, D. Xing, C.L. Huang, L.D. Hall, T.A. Carpenter, A comparative study of acquisition schemes for diffusion tensor imaging using MRI, J. Magn. Reson. 137 (1) (1999) 67–82.
[19] P.J. Basser, S. Pajevic, Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise, Magn. Reson. Med. 44 (2000) 41–50.
[20] S. Skare, T. Li, B. Nordell, M. Ingvar, Noise considerations in the determination of diffusion tensor anisotropy, Magn. Reson. Imaging 18 (2000) 659–669.
[21] C. Pierpaoli, P.J. Basser, Toward a quantitative assessment of diffusion anisotropy [published erratum appears in Magn Reson Med 1997 jun;37(6):972], Magn. Reson. Med. 36 (6) (1996) 893–906.
[22] M. Lazar, K.M. Hasan, A. Alexander, Bootstrap analysis of DT-MRI tractography techniques: streamlines and tensorlines, in: ISMRM Proceedings, Glasgow, Scotland, UK, 2001, p. 1527.
[23] R.M. Henkelman, Measurement of signal intensities in the presence of noise in MR images, Med. Phys. 12 (2) (1985) 232–233.
[24] H. Gudbjartsson, S. Patz, The Rician distribution of noisy MRI data [published erratum appears in Magn Reson Med 1996 aug;36(2):332], Magn. Reson. Med. 34 (6) (1995) 910–914.
[25] P.J. Basser, J. Mattiello, D. LeBihan, Estimation of the effective self-diffusion tensor from the NMR spin echo, J. Magn. Reson. B 103 (3) (1994) 247–254.
[26] P.R. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York, 1969.
[27] C.R. Rao, Linear Statistical Inference and its Applications, Wiley and Sons, New York, 1965.
[28] G.A.F. Seber, C.J. Wild, Nonlinear Regression, Wiley and Sons, New York, 1989.
[29] B. Efron, Bootstrap methods: another look at the jackknife, Ann. Statist. 7 (1979) 1–26.
[30] B. Efron, An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.
[31] A. Virta, A. Barnett, C. Pierpaoli, Visualizing and characterizing white matter fiber structure and architecture in the human pyramidal tract using diffusion tensor MRI, Magn. Reson. Imaging 17 (8) (1999) 1121–1133.
[32] J. Mattiello, P.J. Basser, D. LeBihan, Analytical expression for the *b*-matrix in NMR diffusion imaging and spectroscopy, J. Magn. Reson. A 108 (1994) 131–141.
[33] J. Mattiello, P.J. Basser, D. LeBihan, The *b*-matrix in diffusion tensor echo-planar imaging, Magn. Reson. Med. 37 (1997) 292–300.
[34] J. Mattiello, P.J. Basser, D. LeBihan, Analytical calculation of the *b*-matrix in diffusion imaging, in: Diffusion and Perfusion Magnetic Resonance Imaging, Raven Press, New York.
[35] T.W. Anderson, An Introduction to Multivariate Statistics, Wiley and Sons, New York, 1984.
[36] P.J. Basser, C. Pierpaoli, Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI, J. Magn. Res. B 111 (1996) 209–219.
[37] P.J. Basser, New histological and physiological stains derived from diffusion-tensor MR images, Ann. N.Y. Acad. Sci. 820 (1997) 123–138.
[38] C. Pierpaoli, P. Jezzard, P.J. Basser, A. Barnett, G.D. Chiro, Diffusion tensor MR imaging of the human brain, Radiology 201 (3) (1996) 637–648.
[39] P.J. Basser, S. Pajevic, Quantitative statistical tests for assessing changes in the trace of the diffusion tensor: Clinical and biological implications, in: ISMRM Proceedings, Philadelphia, 1999, p. 1789.
[40] D.K. Jones, M.A. Horsfield, A. Simmons, Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging, Magn. Reson. Med. 42 (3) (1999) 515–525.
[41] K.M. Martin, N.G. Papadakis, C.L. Huang, L.D. Hall, T.A. Carpenter, The reduction of the sorting bias in the eigenvalues of the diffusion tensor, Magn. Reson. Imaging 17 (6) (1999) 893–901.