Introduction to Diffusion Tensor Imaging Mathematics: Part III. Tensor Calculation, Noise, Simulations, and Optimization

PETER B. KINGSLEY

Department of Radiology, North Shore University Hospital, 300 Community Drive, Manhasset, New York 11030, and Department of Radiology, New York University School of Medicine, New York, New York

ABSTRACT: The mathematical aspects of diffusion tensor magnetic resonance imaging (DTMRI, or DTI), the measurement of the diffusion tensor by magnetic resonance imaging (MRI), are discussed in this three-part series. Part III begins with a comparison of different ways to calculate the tensor from diffusion-weighted imaging data. Next, the effects of noise on signal intensities and diffusion tensor measurements are discussed. In MRI signal intensities as well as DTI parameters, noise can introduce a bias (systematic deviation) as well as scatter (random deviation) in the data. Propagation-of-error formulas are explained with examples. Step-by-step procedures for simulating diffusion tensor measurements are presented. Finally, methods for selecting the optimal b factor and number of $b = 0$ images for measuring several properties of the diffusion tensor, including the trace (or mean diffusivity) and anisotropy, are presented. Concepts Magn Reson Part A 28A: 155–179, 2006

KEY WORDS: diffusion; diffusion tensor; DTI; least-squares fit; Monte Carlo simulations; noise; optimization; tensor calculation

INTRODUCTION

This three-part series discusses the mathematical aspects of diffusion tensor magnetic resonance imaging (DTMRI, or DTI). Part I explained a few concepts about diffusion and its measurement with magnetic resonance imaging (MRI). Vectors and tensors were introduced in both two dimensions (2D) and three dimensions (3D), and their rotations were discussed. Rotationally invariant properties of the tensor were described. The calculation of eigenvectors and eigenvalues from the tensor was explained. A list of errors in DTI-related publications was included as an appendix. Part II began with a discussion of different ways to evaluate the degree of anisotropy of a tensor, which is commonly expressed as a diffusion anisotropy index (DAI). Calculation of the diffusion-weighting b

Received 12 September 2005; revised 12 December 2005; accepted 12 December 2005

Correspondence to: Peter Kingsley: E-mail: pkingsle@nshs.edu

Concepts in Magnetic Resonance Part A, Vol. 28A(2) 155–179 (2006)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/cmr.a.20050

factor was explained next. The last section explored different gradient sampling schemes for diffusion-weighted imaging (DWI) and DTI.

Part III of this series explains how the tensor is calculated from DWI data in six or more directions, including weighted and unweighted linear and nonlinear least-squares fits. Noise in MRI images, and how noise in the raw data propagates into the tensor and parameters derived from the tensor, are discussed next. An explanation of Monte Carlo computer simulations of DTI measurements is provided. These simulations are important for confirming propagation-of-error predictions and for understanding the bias in DTI parameters that is not predicted by standard propagation-of-error formulas. Finally, an explanation of how to optimize DTI data acquisition is based on the information from noise propagation and computer simulations.

CALCULATING THE DIFFUSION TENSOR FROM DWI DATA

The purpose of this section is to explain and compare published methods of calculating the diffusion tensor from DWI data.

Vector Representations of the Tensor, the Gradient Encoding Directions, and the b Matrix

The diffusion tensor can be calculated from DWI data collected with diffusion-sensitizing gradients in six or more directions. For notational convenience, the following formulas assume that there are N total measurements, M with $b > 0$ and $N - M$ with $b = 0$.

One approach (*1*) is to represent the six distinct tensor elements and the logarithm of the $b = 0$ signal intensity as a seven-element column vector $\boldsymbol{\alpha}$:

$$\alpha = [D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}, \ln(S_0)]^{\mathrm{T}} \quad [1]$$

where S_0 is the signal intensity with $b = 0$. Each individual $\mathbf{b}$ matrix is represented by a six-element row vector $\mathbf{b}_i$

$$\mathbf{b}_i = (b_{xxi}, b_{yyi}, b_{zzi}, 2b_{xyi}, 2b_{xzi}, 2b_{yzi}) \quad [2]$$

which is part of a seven-element row vector $\mathbf{B}_i$:

$$\mathbf{B}_i = (-b_{xxi}, -b_{yyi}, -b_{zzi}, -2b_{xyi}, -2b_{xzi}, -2b_{yzi}, 1) \quad [3]$$

These row vectors are combined into one large $N \times 7$ $\mathbf{B}$ matrix:

$$\mathbf{B} = \begin{pmatrix} -b_{xx1} & -b_{yy1} & -b_{zz1} & -2b_{xy1} & -2b_{xz1} & -2b_{yz1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -b_{xxN} & -b_{yyN} & -b_{zzN} & -2b_{xyN} & -2b_{xzN} & -2b_{yzN} & 1 \end{pmatrix} \quad [4]$$

In the absence of noise, the logarithms of the predicted signal intensities are given by an $N \times 1$ column vector ξ:

$$\xi_i = \mathbf{B}_i\alpha = \ln(S_i) = \ln(S_0) - \mathbf{b}_i{:}\mathbf{D} \quad [5]$$

$$\xi_i = -b_{xxi}D_{xx} - b_{yyi}D_{yy} - b_{zzi}D_{zz} - 2b_{xyi}D_{xy} - 2b_{xzi}D_{xz} - 2b_{yzi}D_{yz} + \ln(S_0) \quad [6]$$

$$\xi = \mathbf{B}\boldsymbol{\alpha} = \begin{pmatrix} -b_{xx1}D_{xx} - b_{yy1}D_{yy} - b_{zz1}D_{zz} - 2b_{xy1}D_{xy} - 2b_{xz1}D_{xz} - 2b_{yz1}D_{yz} + \ln(S_0) \\ \vdots \\ -b_{xxN}D_{xx} - b_{yyN}D_{yy} - b_{zzN}D_{zz} - 2b_{xyN}D_{xy} - 2b_{xzN}D_{xz} - 2b_{yzN}D_{yz} + \ln(S_0) \end{pmatrix} \quad [7]$$

The $\mathbf{B}$ matrix has a row for each measurement, both $b = 0$ and $b > 0$, and the $\boldsymbol{\xi}$ vector contains an element for each measurement. The noisy observed data are represented as an $N \times 1$ column vector $\mathbf{x}$:

$$\mathbf{x} = \begin{pmatrix} \ln(S_1) \\ \ln(S_2) \\ \vdots \\ \ln(S_N) \end{pmatrix} \quad [8]$$

The noisy data for each acquisition can then be expressed as

$$\mathbf{x} = \mathbf{B}\alpha + \eta = \xi + \eta \quad [9]$$

where $\boldsymbol{\eta}$ is a noise vector.

Another approach (*2–4*) is to represent the tensor as a six-element column vector, $\mathbf{d}$:

$$\mathbf{d} = [D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}]^{\mathrm{T}} \quad [10]$$

Each individual encoding matrix is represented as a six-element row matrix, $\mathbf{H}_i$, which is derived from the normalized gradient components g_{xi}, g_{yi}, and g_{zi}.

$$\mathbf{H}_i = [g_{xi}^2, g_{yi}^2, g_{zi}^2, 2g_{xi}g_{yi}, 2g_{xi}g_{zi}, 2g_{yi}g_{zi}] \quad [11]$$

The $\mathbf{H}_i$ vectors can then be combined into one large $M \times 6$ matrix, $\mathbf{H}$.

$$\mathbf{H} = [\mathbf{H}_1^{\mathrm{T}}\mathbf{H}_2^{\mathrm{T}} \cdot \cdot \cdot \mathbf{H}_M^{\mathrm{T}}]^{\mathrm{T}} \quad [12]$$

$$\mathbf{H} = \begin{pmatrix} g_{x1}^2 & g_{y1}^2 & g_{z1}^2 & 2g_{x1}g_{y1} & 2g_{x1}g_{z1} & 2g_{y1}g_{z1} \\ g_{x2}^2 & g_{y2}^2 & g_{z2}^2 & 2g_{x2}g_{y2} & 2g_{x2}g_{z2} & 2g_{y2}g_{z2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{xM}^2 & g_{yM}^2 & g_{zM}^2 & 2g_{xM}g_{yM} & 2g_{xM}g_{zM} & 2g_{yM}g_{zM} \end{pmatrix} \quad [13]$$

This time the observed data are expressed as the individual measured apparent diffusion coefficients (*ADC*s) instead of the logarithms of the individual signal intensities:

$$\mathbf{Y}_i = \ln(S_0/S_i)/b \quad [14]$$

$$\mathbf{Y} = [\ln(S_0/S_1)/b, \ln(S_0/S_2)/b, \ldots, \ln(S_0/S_M)/b]^{\mathrm{T}} \quad [15]$$

The noisy data for each acquisition can then be expressed as

$$\mathbf{Y} = \mathbf{Hd} + \eta \quad [16]$$

Although these two approaches have some similarities (cf. Eq. [1] with Eq. [10], Eq. [3] with Eq. [11], Eq. [4] with Eq. [13], and Eq. [9] with Eq. [16]), there are noteworthy differences. In the second approach (Eqs. [10–16]), all the $b = 0$ images must be averaged to produce a single value of S_0. In the first method (Eqs. [1–9]), the separate $b = 0$ acquisitions can be combined, or they can be treated individually. However, errors in the tensor elements and derived parameters will be less if the $b = 0$ images (and any other repeated data points) are combined before calculating the tensor.

Four types of solutions to Eqs. [9] and [16] will be considered here. First, an exact solution exists when there are only six directions. Second, an unweighted linear least-squares fit can be calculated (*4*). Third, a weighted linear least-squares fit can be calculated (*1*). Fourth, a nonlinear least-squares fit can be calculated (*5*).

Exactly Six Directions with the H Matrix

With exactly six directions, there is an exact analytic solution to Eq. [16]. The six measured *ADC*s can be written as

$$D_i = b(g_{xi}^2 D_{xx} + g_{yi}^2 D_{yy} + g_{zi}^2 D_{zz} + 2g_{xi}g_{yi}D_{xy} + 2g_{xi}g_{zi}D_{xz} + 2g_{yi}g_{zi}D_{yz}) \quad [17]$$

where g_{xi}, g_{yi}, and g_{zi} are the normalized gradient components. These six equations in six unknowns can be solved by standard methods, such as Cramer's rule (*6*). Because it is impossible to distinguish $\boldsymbol{\eta}$ from $\mathbf{d}$, $\boldsymbol{\eta} = 0$ in Eq. [16], and $\mathbf{d}$ can be determined by multiplying both sides of Eq. [16] by the inverse of $\mathbf{H}$, $\mathbf{H}^{-1}$:

$$(\mathbf{H}^{-1}\mathbf{H})\mathbf{d} = \mathbf{d} = \mathbf{H}^{-1}\mathbf{Y} \quad [18]$$

For the gradient scheme in Table II-6, the measured *ADC*s are given by

$$D_1 = (D_{xx} + 2uD_{xy} + u^2D_{yy})/(1 + u^2) \quad [19]$$

$$D_2 = (D_{xx} - 2uD_{xy} + u^2D_{yy})/(1 + u^2) \quad [20]$$

$$D_3 = (D_{yy} + 2uD_{yz} + u^2D_{zz})/(1 + u^2) \quad [21]$$

$$D_4 = (D_{yy} - 2uD_{yz} + u^2D_{zz})/(1 + u^2) \quad [22]$$

$$D_5 = (D_{zz} + 2uD_{xz} + u^2D_{xx})/(1 + u^2) \quad [23]$$

$$D_6 = (D_{zz} - 2uD_{xz} + u^2D_{xx})/(1 + u^2) \quad [24]$$

$$D_{\mathrm{av}} = (D_1 + D_2 + D_3 + D_4 + D_5 + D_6)/6 = (D_{xx} + D_{yy} + D_{zz})/3 \quad [25]$$

Equation [25] is an explicit demonstration of the ability to calculate D_{av} directly from encoding schemes 6p and 6v without calculating the tensor elements (see part II). The tensor elements can be computed as

$$h(u) = (1 + u^2)/[2(1 + u^6)] \quad [26]$$

$$h(1) = \tfrac{1}{2} \quad [27]$$

$$D_{xx} = h[D_1 + D_2 - u^2(D_3 + D_4) + u^4(D_5 + D_6)] \quad [28]$$

$$D_{yy} = h[D_3 + D_4 - u^2(D_5 + D_6) + u^4(D_1 + D_2)] \quad [29]$$

$$D_{zz} = h[D_5 + D_6 - u^2(D_1 + D_2) + u^4(D_3 + D_4)] \quad [30]$$

$$D_{xy} = (D_1 - D_2)(1 + u^2)/4u \quad [31]$$

$$D_{xz} = (D_3 - D_4)(1 + u^2)/4u \quad [32]$$

$$D_{yz} = (D_5 - D_6)(1 + u^2)/4u \quad [33]$$

After the tensor elements have been calculated, the eigenvalues, DAIs, and eigenvectors can be calculated as described in parts I and II.

When data are collected in each gradient direction and in each negative direction, cross-terms can be eliminated by taking the geometric mean of the signal intensities of each +/− pair, or the arithmetic mean of their calculated *ADC*s, before the tensor is calculated (Eq. [II-141]). If the arithmetic mean is used instead of the geometric mean, or the individual data sets are all treated equally in calculating the tensor elements, the cross-terms theoretically are not completely eliminated, but the residual error is likely to be much less than the original errors, and much less than the noise.

Unweighted Linear Least-Squares Fit with the H Matrix

With more than six gradient directions, the **H** matrix is not a square matrix, so there is no true inverse $\mathbf{H}^{-1}$. However, Eq. [16] can still be solved by calculating a pseudoinverse $\mathbf{H}^{\psi}$ such that

$$\mathbf{H}^{\psi}\mathbf{H} = \mathbf{I}_{6\times 6} \quad [34]$$

even though $\mathbf{H}\,\mathbf{H}^{\psi}$ does not necessarily equal $\mathbf{I}_{M\times M}$. One way to do this is to form the matrix product $\mathbf{H}^{\mathrm{T}}\mathbf{H}$, which is a square 6×6 matrix and therefore has a true inverse (*4*). From Eq. [9],

$$\mathbf{H}^{\mathrm{T}}\mathbf{Y} = \mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{d} \quad [35]$$

$$(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{d} = \mathbf{d} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{Y} \quad [36]$$

$$\mathbf{H}^{\psi} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}} \quad [37]$$

$$\mathbf{H}^{\mathrm{T}}\mathbf{H} = \begin{pmatrix}
\sum_{i=1}^{M} g_{xi}^4 & \sum_{i=1}^{M} g_{xi}^2 g_{yi}^2 & \sum_{i=1}^{M} g_{xi}^2 g_{zi}^2 & 2\sum_{i=1}^{M} g_{xi}^3 g_{yi} & 2\sum_{i=1}^{M} g_{xi}^3 g_{zi} & 2\sum_{i=1}^{M} g_{xi}^2 g_{yi} g_{zi} \\
\sum_{i=1}^{M} g_{yi}^2 g_{xi}^2 & \sum_{i=1}^{M} g_{yi}^4 & \sum_{i=1}^{M} g_{yi}^2 g_{zi}^2 & 2\sum_{i=1}^{M} g_{xi} g_{yi}^3 & 2\sum_{i=1}^{M} g_{xi} g_{yi}^2 g_{zi} & 2\sum_{i=1}^{M} g_{yi}^3 g_{zi} \\
\sum_{i=1}^{M} g_{xi}^2 g_{zi}^2 & \sum_{i=1}^{M} g_{yi}^2 g_{zi}^2 & \sum_{i=1}^{M} g_{zi}^4 & 2\sum_{i=1}^{M} g_{xi} g_{yi} g_{zi}^2 & 2\sum_{i=1}^{M} g_{xi} g_{zi}^3 & 2\sum_{i=1}^{M} g_{yi} g_{zi}^3 \\
2\sum_{i=1}^{M} g_{xi}^3 g_{yi} & 2\sum_{i=1}^{M} g_{xi} g_{yi}^3 & 2\sum_{i=1}^{M} g_{xi} g_{yi} g_{zi}^2 & 4\sum_{i=1}^{M} g_{xi}^2 g_{yi}^2 & 4\sum_{i=1}^{M} g_{xi}^2 g_{yi} g_{zi} & 4\sum_{i=1}^{M} g_{xi} g_{yi}^2 g_{zi} \\
2\sum_{i=1}^{M} g_{xi}^3 g_{zi} & 2\sum_{i=1}^{M} g_{xi} g_{yi}^2 g_{zi} & 2\sum_{i=1}^{M} g_{xi} g_{zi}^3 & 4\sum_{i=1}^{M} g_{xi}^2 g_{yi} g_{zi} & 4\sum_{i=1}^{M} g_{xi}^2 g_{zi}^2 & 4\sum_{i=1}^{M} g_{xi} g_{yi} g_{zi}^2 \\
2\sum_{i=1}^{M} g_{xi}^2 g_{yi} g_{zi} & 2\sum_{i=1}^{M} g_{yi}^3 g_{zi} & 2\sum_{i=1}^{M} g_{yi} g_{zi}^3 & 4\sum_{i=1}^{M} g_{xi} g_{yi}^2 g_{zi} & 4\sum_{i=1}^{M} g_{xi} g_{yi} g_{zi}^2 & 4\sum_{i=1}^{M} g_{yi}^2 g_{zi}^2
\end{pmatrix} \quad [38]$$

This corresponds to an unweighted linear least-squares fit of the logarithms of the signal intensities. The pseudoinverse can also be calculated by singular value decomposition (SVD), which decomposes a matrix **H** into the product of three matrices (*7*).

$$\mathbf{H} = \mathbf{U}\mathbf{W}\mathbf{V}^{\mathrm{T}} \quad [39]$$

$$\mathbf{U} = M \times 6 \text{ column-orthogonal} \quad [40]$$

$$\mathbf{V} = 6 \times 6 \text{ (row- and column-) orthogonal} \quad [41]$$

$$\mathbf{W} = 6 \times 6 \text{ diagonal} \quad [42]$$

Since $\mathbf{V}$ is orthogonal and $\mathbf{U}$ is column orthogonal,

$$\mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{I}_{6\times 6} \quad [43]$$

$$\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_{6\times 6} \quad [44]$$

Because $\mathbf{W}$ is diagonal, the diagonal elements of $\mathbf{W}^{-1}$ are

$$(\mathbf{W}^{-1})_{ij} = 1/\mathbf{W}_{ij} \quad [45]$$

Thus, $\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^{\mathrm{T}}$ is the pseudoinverse of $\mathbf{H}$, $\mathbf{H}^{\psi}$, because

$$\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^{\mathrm{T}}\mathbf{H} = \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{W}\mathbf{V}^{\mathrm{T}} = \mathbf{I}_{6\times 6} \quad [46]$$

Clearly, the pseudoinverse calculated by these two methods must be equal, so

$$(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}} = \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^{\mathrm{T}} = \mathbf{H}^{\psi} \quad [47]$$

Exactly Six Directions with the B Matrix

With exactly six directions, there is an exact solution to Eq. [9], which can be written as

$$(\mathbf{B}^{-1}\mathbf{B})\alpha = \alpha = \mathbf{B}^{-1}\mathbf{x} \quad [48]$$

Because it is impossible to distinguish noise $\boldsymbol{\eta}$ from the true tensor $\boldsymbol{\alpha}$, $\boldsymbol{\eta}$ in Eq. [9] is set equal to 0. This solution must be the same as the solution calculated with the $\mathbf{H}$ matrix, so Eqs. [19–33] apply.

Unweighted Linear Least-Squares Fit with the B Matrix

The pseudoinverse of $\mathbf{B}$, $\mathbf{B}^{\psi}$, can be calculated as it was for the $\mathbf{H}$ matrix, yielding

$$(\mathbf{B}^{\mathrm{T}}\mathbf{B})^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{B}\alpha = \alpha = (\mathbf{B}^{\mathrm{T}}\mathbf{B})^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{x} \quad [49]$$

$$(\mathbf{B}^{\mathrm{T}}\mathbf{B})^{-1}\mathbf{B}^{\mathrm{T}} = \mathbf{B}^{\psi} \quad [50]$$

The $\mathbf{B}^{\psi}$ matrix can also be calculated by SVD. This corresponds to an unweighted linear least-squares fit of the logarithms of the signal intensities.

Weighted Linear Least-Squares Fit with the B Matrix

In the unweighted linear least-squares fits with either the $\mathbf{H}$ matrix or the $\mathbf{B}$ matrix, all data points are treated as if they were equally accurate. This is generally true for the original signal intensities, so it is not true for their logarithms. If the signal intensities have the same variance, then the uncertainty in $\ln(S_i)$ is proportional to $1/S_i$. Thus, the least-squares fit should give more weight (importance) to the $\ln(S_i)$ of high S_i values and less weight to $\ln(S_i)$ of low S_i values. Mathematically, this is accomplished by modifying Eq. [49] as follows (*1*):

$$\alpha = (\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{B})^{-1}(\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1})\mathbf{x} \quad [51]$$

where $\boldsymbol{\Sigma}^{-1}$ is a diagonal $N \times N$ matrix whose elements are ${S_i}^2/{\sigma_i}^2$:

$$\Sigma^{-1} = \mathrm{diag}(S_i^2/\sigma_i^2) \quad [52]$$

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \dfrac{S_1^2}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \dfrac{S_2^2}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{S_N^2}{\sigma_N^2} \end{pmatrix} \quad [53]$$

$$\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} -b_{xx1}\dfrac{S_1^2}{\sigma_1^2} & -b_{xx2}\dfrac{S_2^2}{\sigma_2^2} & \cdots & -b_{xxN}\dfrac{S_N^2}{\sigma_N^2} \\ -b_{yy1}\dfrac{S_1^2}{\sigma_1^2} & -b_{yy2}\dfrac{S_2^2}{\sigma_2^2} & \cdots & -b_{yyN}\dfrac{S_N^2}{\sigma_N^2} \\ -b_{zz1}\dfrac{S_1^2}{\sigma_1^2} & -b_{zz2}\dfrac{S_2^2}{\sigma_2^2} & \cdots & -b_{zzN}\dfrac{S_N^2}{\sigma_N^2} \\ -2b_{xy1}\dfrac{S_1^2}{\sigma_1^2} & -b_{xy2}\dfrac{S_2^2}{\sigma_2^2} & \cdots & -b_{xyN}\dfrac{S_N^2}{\sigma_N^2} \\ -2b_{xz1}\dfrac{S_1^2}{\sigma_1^2} & -b_{xz2}\dfrac{S_2^2}{\sigma_2^2} & \cdots & -b_{xzN}\dfrac{S_N^2}{\sigma_N^2} \\ -2b_{yz1}\dfrac{S_1^2}{\sigma_1^2} & -b_{yz2}\dfrac{S_2^2}{\sigma_2^2} & \cdots & -b_{yzN}\dfrac{S_N^2}{\sigma_N^2} \\ \dfrac{S_1^2}{\sigma_1^2} & \dfrac{S_2^2}{\sigma_2^2} & \cdots & \dfrac{S_N^2}{\sigma_N^2} \end{pmatrix} \quad [54]$$

$$\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{B} = \begin{pmatrix} \sum_{i=1}^{N} b_{xxi}^2 \frac{S_i^2}{\sigma_i^2} & \sum_{i=1}^{N} b_{xxi}b_{yyi} \frac{S_i^2}{\sigma_i^2} & \cdots & 2\sum_{i=1}^{N} b_{xxi}b_{yzi} \frac{S_i^2}{\sigma_i^2} & -\sum_{i=1}^{N} b_{xxi} \frac{S_i^2}{\sigma_i^2} \\ \sum_{i=1}^{N} b_{yyi}b_{xxi} \frac{S_i^2}{\sigma_i^2} & \sum_{i=1}^{N} b_{yyi}^2 \frac{S_i^2}{\sigma_i^2} & \cdots & 2\sum_{i=1}^{N} b_{yyi}b_{yzi} \frac{S_i^2}{\sigma_i^2} & -\sum_{i=1}^{N} b_{yyi} \frac{S_i^2}{\sigma_i^2} \\ \sum_{i=1}^{N} b_{zzi}b_{xxi} \frac{S_i^2}{\sigma_i^2} & \sum_{i=1}^{N} b_{zzi}b_{yyi} \frac{S_i^2}{\sigma_i^2} & \cdots & 2\sum_{i=1}^{N} b_{zzi}b_{yzi} \frac{S_i^2}{\sigma_i^2} & -\sum_{i=1}^{N} b_{zzi} \frac{S_i^2}{\sigma_i^2} \\ 2\sum_{i=1}^{N} b_{xyi}b_{xxi} \frac{S_i^2}{\sigma_i^2} & 2\sum_{i=1}^{N} b_{xyi}b_{yyi} \frac{S_i^2}{\sigma_i^2} & \cdots & 4\sum_{i=1}^{N} b_{xyi}b_{yzi} \frac{S_i^2}{\sigma_i^2} & -2\sum_{i=1}^{N} b_{xyi} \frac{S_i^2}{\sigma_i^2} \\ 2\sum_{i=1}^{N} b_{xzi}b_{xxi} \frac{S_i^2}{\sigma_i^2} & 2\sum_{i=1}^{N} b_{xzi}b_{yyi} \frac{S_i^2}{\sigma_i^2} & \cdots & 4\sum_{i=1}^{N} b_{xzi}b_{yzi} \frac{S_i^2}{\sigma_i^2} & -2\sum_{i=1}^{N} b_{xzi} \frac{S_i^2}{\sigma_i^2} \\ 2\sum_{i=1}^{N} b_{yzi}b_{xxi} \frac{S_i^2}{\sigma_i^2} & 2\sum_{i=1}^{N} b_{yzi}b_{yyi} \frac{S_i^2}{\sigma_i^2} & \cdots & 4\sum_{i=1}^{N} b_{yzi}^2 \frac{S_i^2}{\sigma_i^2} & -2\sum_{i=1}^{N} b_{yzi} \frac{S_i^2}{\sigma_i^2} \\ -\sum_{i=1}^{N} b_{xxi} \frac{S_i^2}{\sigma_i^2} & -\sum_{i=1}^{N} b_{yyi} \frac{S_i^2}{\sigma_i^2} & \cdots & -2\sum_{i=1}^{N} b_{yzi} \frac{S_i^2}{\sigma_i^2} & -\sum_{i=1}^{N} \frac{S_i^2}{\sigma_i^2} \end{pmatrix} \quad [55]$$

If each original image has the same noise level, then σ_i^2 is the same in each original image. However, the σ_i^2 in $\boldsymbol{\Sigma}^{-1}$ can be modified to reflect different amounts of signal averaging for individual directions or for the $b = 0$ images. Replacing $\boldsymbol{\Sigma}^{-1}$ with the identity matrix in Eq. [51] yields the unweighted linear least-squares fit. The ($\mathbf{B}^{\mathrm{T}}\ \boldsymbol{\Sigma}^{-1}\ \mathbf{B}$) matrix is called the covariance matrix.

The S_i values in Eqs. [53–55] should be the true (noiseless) values. Because these are not known, the observed (noisy) S_i values usually are used. After a least-squares fit is calculated, it is possible to repeat the calculation with the fitted S_i values, which should be closer to the true S_i values than the original noisy S_i values were.

Comparison of the H and B Matrix Approaches

In principle, the **B** matrix approach has several potential advantages over the **H** matrix approach. First, the **B** matrix allows the use of two or more different b factors in addition to $b = 0$. Second, an estimate of $\ln(S_0)$ is produced in the least-squares fit with the **B** matrix approach, while $\ln(S_0)$ from the average $b = 0$ signal intensities is incorporated into the **Y** vector in the **H** matrix approach. Third, the individual $b = 0$ images can be fitted separately in the **B** matrix approach. Fourth, a weighted linear least-squares fit is possible with the **B** matrix approach. Fifth, the **B** matrix approach yields the covariance matrix ($\mathbf{B}^{\mathrm{T}}\ \boldsymbol{\Sigma}^{-1}\ \mathbf{B}$). The covariance matrix can be used for propagation-of-error calculations and optimization of DTI parameters, which are discussed in a future section.

In practice, a single b factor is almost always used in DTI because of nonmonoexponential signal decay. Furthermore, an estimate of $\ln(S_0)$ from the $b = 0$ images is usually adequate. In addition, it is better to average the $b = 0$ images than to fit them individually. Either approach can usually provide a good starting point for a nonlinear least-squares fit of the signal intensities. The availability of the weighted linear least-squares fit and the covariance matrix can be desirable in some circumstances.

Nonlinear Least-Squares Fit

Even the weighted linear least-squares fit is not ideal with high noise levels because symmetrically distributed noise in S_i is not symmetrically distributed in $\ln(S_i)$. For example, variation of S_i by ± 5 changes $\ln(100 \pm 5)$ by -0.051 to $+0.049$ units, whereas $\ln(20 \pm 5)$ changes by -0.288 to $+0.223$ units, and $\ln(10 \pm 5)$ changes by -0.693 to $+0.405$ units. In the

linear least-squares fit, if the fitted value is $S_i = 10$, then $S_i = 5$ and $S_i = 20$ are considered equally good fits, even though $S_i = 20$ is twice as far from the fitted value. To correct for this, a linear least-squares fit can serve as a starting point for a nonlinear least-squares fit of the signal intensities (*8*). Because the noise level usually is the same in each original image, the nonlinear fit does not need weighting factors. In contrast to the analytic linear least-squares solution, the nonlinear least-squares solution requires an iterative numerical approach, usually the Levenberg-Marquardt algorithm (*7*). Although unweighted linear, weighted linear, and nonlinear fits have been compared in several contexts (*5, 9*), they do not appear to have been compared directly in tensor calculations with noisy data in the absence of outliers. More advanced processing methods are required for reproducible results in the presence of outliers, data points that deviate significantly from the expected value due to patient motion or cardiac pulsation (*5*).

Important Points in "Calculating the Diffusion Tensor from DWI Data"

Estimating the tensor elements from noisy data by a linear least-squares fit can be calculated with a matrix approach. The **B** matrix approach (Eqs. [1–9]) is more flexible than the **H** matrix approach (Eqs. [11–16]) and allows a weighted fit to be performed. Nonlinear least-squares fits require an iterative fit, which takes more time. With exactly six gradient directions, if repetitions are averaged before the tensor is calculated so that there are only six data points to be fitted, all the fitting methods yield the same result.

NOISE AND SIMULATIONS

The purposes of this section are (1) to describe the effects of noise on MRI signal intensities when the signal-to-noise ratio (*SNR*) is low, (2) to explain basic propagation-of-error formulas, (3) to show how to simulate DTI measurements, and (4) to explain that noise introduces a bias (systematic deviation) into DTI data in addition to the expected scatter (random deviation).

Noise

Random noise in raw data produces a scatter in the raw data so that each individual measurement differs from the true value by a variable amount. This noise propagates into errors or uncertainty in parameters derived from the raw data. In addition to this scatter in the data, noise in DTI data produces a systematic shift (bias) in some parameters, including eigenvalues (*10–13*), DAIs (*13–16*), and the standard deviations (*SD*s) of DAIs (*17*).

Another type of bias can be introduced by the use of magnitude data in MR imaging when the *SNR* is low. Because of this bias, the effects of noise in DTI data cannot be completely described by conventional propagation-of-error formulas. However, such formulas are still useful in certain circumstances.

This section discusses the following topics: 1) the effects of noise in MR images when *SNR* is low; 2) applications of propagation-of-error formulas in DTI; 3) simulating DTI measurements; and 4) the bias in measured DTI parameters caused by noise. For a more complete discussion of noise and statistics, the reader is referred to standard references (*6*).

Definitions of Noise Terms

Noise is the deviation of a measured variable, such as voltage, from the true value. Noise is considered to have a mean of zero, that is, there are equal probabilities of negative and positive deviations from the true value μ. In addition, the noise is considered to have a Gaussian distribution with standard deviation σ, so that the relative probability for making a single observation x_i within an interval dx_i is given by (*6*)

$$dP_i = \exp\{(-1/2)[(x_i - \mu)/\sigma]^2\}/[\sigma(2\pi)^{1/2}]dx_i \quad [56]$$

Round-off error occurs when an analog variable, such as voltage, is digitized. In MRI this error is usually much smaller than the noise, and is commonly ignored in simulations. Round-off error is considered to be uniformly distributed with a mean of zero.

Scatter refers to the range of observed values of a parameter. With a single true value, this scatter can arise from noise, round-off error, and propagation of errors from raw data into calculated parameters. When sample heterogeneity is present, as in a biological sample, the heterogeneity also can contribute to the observed scatter. Scatter is a measure of the precision of a measurement.

The *variance*, σ^2, is the mean square deviation of observed values from the mean value, μ:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \quad [57]$$

Note that this is the "N variance," not the "$N - 1$ variance" that is commonly used in statistical computations.

The *standard deviation* (*SD*), σ, is the square root of the variance.

The *expectation value* of S, $\langle S \rangle$, is the mean value that would be observed after an infinite number of measurements. The notation $\langle S \rangle$ is also used to indicate the mean value of a repeated measurement or calculation. In contrast, the mean of a number of measurements of different S values is indicated by $\bar{S}$.

Bias is a systematic deviation of observed values from the true value, so that $\langle S \rangle$ is not equal to the true value, μ. This may be caused by physical factors (faulty equipment, inaccurate calibration, or incorrect use of equipment) or by formulas used in data analysis. Bias decreases the accuracy of a measurement, and it may also increase or decrease the precision of the measurement.

Error refers to the difference between an observed value and the true value, including contributions from noise, round-off error, and bias.

Uncertainty refers to the reliability of a measurement, and typically is indicated by the *SD*.

Effects of Noise in MR Images with Low *SNR*

The effects of noise in MR images with low *SNR* has been the subject of several reports (*18–26*). MRI signal intensities are usually measured with quadrature detection, so that two orthogonal signal channels (Re and Im, for real and imaginary) have independent noise. (The two quadrature channels are sometimes called A and D for absorption and dispersion signals. That notation will not be used here, to avoid confusion with the *ADC* value, *D*, and the anisotropy index *A*. The references cited here use *A* for the true signal, which is called *S* in the present work.) The magnitude signal is calculated as (*18, 22*)

$$S = (Re^2 + Im^2)^{1/2} \qquad [58]$$

The effects of noise in magnitude images were discussed in an early article by S. O. Rice (*19, 23, 25, 27*), and the resulting data are referred to as the "Rician distribution" for this reason (*19, 25*). For an actual signal amplitude S, the measured noisy signal, M, is (*18, 22*)

$$M = [(Re + \sigma_R)^2 + (Im + \sigma_I)^2]^{1/2} \qquad [59]$$

$$M = [Re^2 + Im^2 + 2Re\sigma_R + 2Im\sigma_I + \sigma_R^2 + \sigma_I^2]^{1/2} \qquad [60]$$

The noise is assumed to have the same properties in the Re and Im channels, so that

$$\langle \sigma_R \rangle = \langle \sigma_I \rangle = \langle \sigma \rangle \qquad [61]$$

but the noise is independent in each channel. If $\langle \sigma_R \rangle$ is zero, then the expectation value of M^2 (*18, 19, 22, 23*) is

$$\langle M^2 \rangle = S^2 + 2\sigma^2 \qquad [62]$$

The expectation value of M (*19, 23*) is

$$\langle M \rangle = \sigma(\pi/2)^{1/2}\exp(-K)[(1 + 2K)I_0(K) + 2KI_1(K)] \qquad [63]$$

$$K = S^2/4\sigma^2 \qquad [64]$$

where I_0 is the zero-order modified Bessel function, and I_1 is the first-order modified Bessel function,

$$I_0(K) = \sum_{n=0}^{\infty} \frac{K^{2n}}{n!n!2^{2n}} = 1 + \frac{K^2}{1!1!2^2} + \frac{K^4}{2!2!2^4} + \cdots \qquad [65]$$

$$I_1(K) = \sum_{n=0}^{\infty} \frac{K^{2n+1}}{n!(n+1)!2^{2n+1}} = \frac{K}{0!1!2} + \frac{K^3}{1!2!2^3} + \frac{K^5}{2!3!2^5} + \cdots \qquad [66]$$

For $S > 2\sigma$, $\langle M \rangle$ can be approximated as (*22, 23*)

$$\langle M \rangle \approx (S^2 + \sigma^2)^{1/2} \approx (S + \sigma^2/2S), \ (S > 2\sigma) \qquad [67]$$

Thus, for $SNR > 2$, the fractional increase in the measured signal compared to the true signal is approximately

$$(\langle M \rangle - S)/S \approx 1/(2SNR^2), \ (S > 2\sigma) \qquad [68]$$

When $S = 0$, the noise distribution is called a Rayleigh distribution (*19–21, 25*). The expectation value of the background signal and its variance can be calculated with Eqs. [62] and [63] (*20, 25*):

$$\langle M \rangle = \sigma(\pi/2)^{1/2}, \ (S = 0) \qquad [69]$$

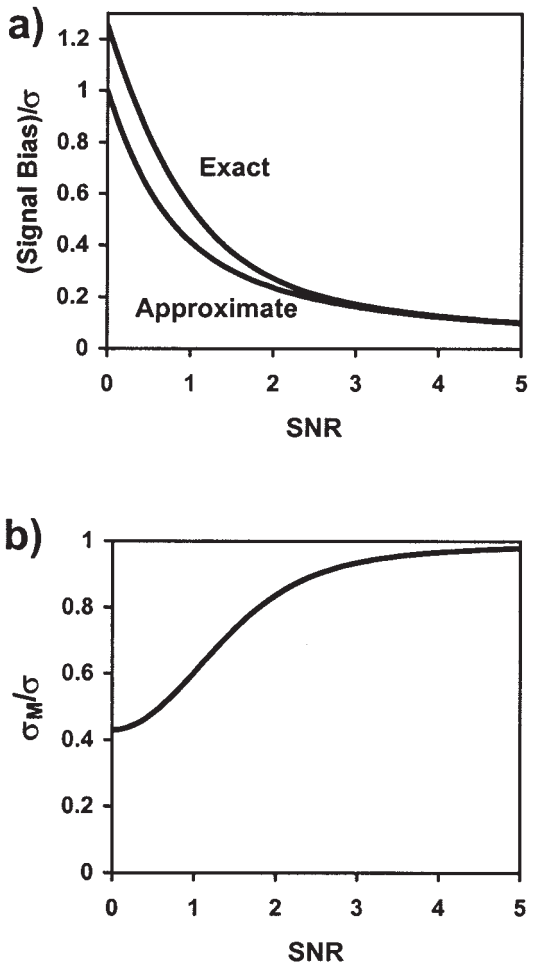

Figure 1 Noise-induced bias in (a) the signal intensity and (b) measured *SD* of a magnitude MRI signal, both expressed in units of σ, the true signal standard deviation (*SD*) in each quadrature receiver channel. *S* is the true signal intensity, *M* is the measured signal intensity, and σ_M is the *SD* measured in the magnitude signal. In (a), the exact formula is $(\langle M\rangle - S)/\sigma$, with $\langle M\rangle$ shown in Eq. [63]. The approximate formula is $(S^2 + \sigma^2)^{1/2}/\sigma$ (see Eq. [67]). In (b), the ratio σ_M/σ is shown. The measured σ_M is calculated from Eq. [70], using the expressions in Eqs. [62] and [63].

$$\sigma_M^2 = \langle M^2\rangle - \langle M\rangle^2 = \sigma^2(2 - \pi/2),\ (S = 0) \qquad [70]$$

Equations [69] and [70] are useful for calculating the noise in an image from the mean and *SD* of the background signal. Equation [67] is useful for estimating the average bias in a signal intensity. Plots of the actual signal bias ($\langle M\rangle - S)/\sigma$, from Eq. [63]), the bias estimated from Eq. [67] ($[(S^2 + \sigma^2)^{1/2} - S]/\sigma$), and σ_M/σ are shown as a function of S/σ (*SNR*) in Fig. 1. Henkelman showed a good plot of $\langle M\rangle$, *S*, and σ_M as a function of S/σ (*22*). Gudbjartsson and Patz showed a plot of the M/σ distribution for integer S/σ, including the mean values, and compared some formulas that attempt to reduce the bias (*25*).

Propagation-of-Error Formulas

When noisy data are used for the calculation of other parameters, the noise in the data propagates into the other parameters in predictable ways (*6*). The commonly used propagation-of-error formulas assume a first-order linear approximation of a function. As the noise level increases, this approximation breaks down. However, it is a good approximation for many applications and can be extended to higher-order derivatives if necessary.

The noise in a calculated parameter can be estimated from the noise in the original data by standard propagation-of-error formulas, which are of the form

$$\sigma^2[f(x, y, z)] = \sigma_x^2\left(\frac{\partial f}{\partial x}\right)^2 + \sigma_y^2\left(\frac{\partial f}{\partial y}\right)^2 + \sigma_z^2\left(\frac{\partial f}{\partial z}\right)^2 \qquad [71]$$

Therefore, when one parameter can be expressed as a function of another parameter, for example $f(x)$, Eq. [71] shows how the error or uncertainty in x propagates into $f(x)$. This formula works as long as there is a direct analytic expression relating the two parameters.

The formula does not apply in certain situations. First, the formula does not apply when there is not an exact expression relating the two parameters. However, in this case there may be an approximate function relating the two parameters near a certain value. In this case, if the noise level is small enough, Eq. [71] may provide a reasonably good estimate of the error propagation near that value, though different $f(x)$ formulas may be needed at other values. Second, the formula does not apply when the variance in different parameters is correlated (covariance is nonzero). The formula can be modified to include the effect of covariance (*6*).

The formula must be applied carefully in two-stage calculations. Consider what happens to $f(x, y)$ if x and y are functions of s and t, $x(s, t)$ and $y(s, t)$. Propagation of error from s and t to into f can be calculated correctly by

$$\sigma_f^2 = \sigma_s^2\left(\frac{\partial f}{\partial s}\right)^2 + \sigma_t^2\left(\frac{\partial f}{\partial t}\right)^2 \qquad [72]$$

$$\left(\frac{\partial f}{\partial s}\right)^2 = \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}\right)^2 \qquad [73]$$

$$\left(\frac{\partial f}{\partial s}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2\left(\frac{\partial x}{\partial s}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\left(\frac{\partial y}{\partial s}\right)^2 + 2\,\frac{\partial f}{\partial x}\frac{\partial x}{\partial s}\frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \quad [74]$$

$$\left(\frac{\partial f}{\partial t}\right)^2 = \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}\right)^2 \quad [75]$$

$$\left(\frac{\partial f}{\partial t}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2\left(\frac{\partial x}{\partial t}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\left(\frac{\partial y}{\partial t}\right)^2 + 2\,\frac{\partial f}{\partial x}\frac{\partial x}{\partial t}\frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \quad [76]$$

The last terms in Eqs. [74] and [76] are covariance terms and are essential for the correct propagation of errors from s and t into f. A common mistake (*28, 29*) is to attempt to propagate errors from s and t into x and y, and then separately from x and y into f. Propagation of errors from s and t into x and y yields

$$\sigma_x^2 = \sigma_s^2\left(\frac{\partial x}{\partial s}\right)^2 + \sigma_t^2\left(\frac{\partial x}{\partial t}\right)^2 \quad [77]$$

$$\sigma_y^2 = \sigma_s^2\left(\frac{\partial y}{\partial s}\right)^2 + \sigma_t^2\left(\frac{\partial y}{\partial t}\right)^2 \quad [78]$$

Propagation of errors from x and y into f yields

$$\sigma_f^2 = \sigma_x^2\left(\frac{\partial f}{\partial x}\right)^2 + \sigma_y^2\left(\frac{\partial f}{\partial y}\right)^2 \quad [79]$$

$$= \sigma_s^2\left(\frac{\partial f}{\partial x}\right)^2\left(\frac{\partial x}{\partial s}\right)^2 + \sigma_t^2\left(\frac{\partial f}{\partial x}\right)^2\left(\frac{\partial x}{\partial t}\right)^2 + \sigma_s^2\left(\frac{\partial f}{\partial y}\right)^2\left(\frac{\partial y}{\partial s}\right)^2 + \sigma_t^2\left(\frac{\partial f}{\partial y}\right)^2\left(\frac{\partial y}{\partial t}\right)^2 \quad [80]$$

Notice that Eq. [80] lacks the covariance terms found in Eqs. [74] and [76] and therefore will generally produce an incorrect result.

As an example, consider the calculation of D_{av} from six measurements with gradient pairs (scheme 6p in Table II-11). The measured *ADC*s are shown in Eqs. [19–24] with $u = 1$, and D_{av} is calculated in Eq. [25]. The tensor elements can be calculated as in Eqs. [28–33] with $u = 1$ and $h = ½$. Propagation of the errors in D_i (σ_i) to D_{av} (σ_{av}) can be calculated from Eq. [71] and the first half of Eq. [25]:

$$\sigma_{av}^2 = \frac{1}{6}\sum_{i=1}^{6}\sigma_i^2 \quad [81]$$

Propagation of the D_i errors to D_{xx}, D_{yy}, and D_{zz} can also be calculated from Eq. [71] and Eqs. [28–30]:

$$\sigma_{xx}^2 = \sigma_{yy}^2 = \sigma_{zz}^2 = \frac{1}{2}\sum_{i=1}^{6}\sigma_i^2 \quad [82]$$

If σ_{av} is calculated from Eq. [71] and the second half of Eq. [25],

$$\sigma_{av}^2 = \frac{1}{3}3\frac{1}{2}\sum_{i=1}^{6}\sigma_i^2 = \frac{1}{2}\sum_{i=1}^{6}\sigma_i^2 \quad [83]$$

Thus, the variance calculated by a two-step propagation of error (Eq. [83]) is three times the actual variance (Eq. [81]).

Simulating DTI Measurements

Many individual steps in the processing and analysis of DTI data can be described by analytic formulas. However, there is no analytic formula to describe the entire DTI process in the general case. Therefore, attempts to optimize DTI data acquisition, and to understand possible problems, often rely on Monte Carlo simulations. Another reason to perform simulations is that analytic formulas do not consider the bias introduced by the Rician distribution with low *SNR*, as discussed previously. Ideally, simulations would always be performed with high *SNR* to show the intrinsic properties of a process, and low *SNR* to show other features that may appear with the biases introduced by low *SNR*. Unfortunately, this combination of simulations is rarely performed.

In Monte Carlo simulations, theoretical signal intensities (*SI*s) are calculated, then random Gaussian-distributed noise is added to each signal intensity. These noisy *SI*s are then analyzed as DTI data would be analyzed. This process is repeated many times, and the means and standard deviations of the resulting variables are calculated and compared with the known noise-free values. Each of the following steps will be described in detail. Steps 1–3 can be performed in any order:

1. Selection of a diffusion tensor
2. Selection of a b factor and gradient sampling scheme

3. Selection of a signal-to-noise ratio (*SNR*)
4. Calculation of *b*:*D* for each gradient direction
5. Calculation of the theoretical *SI* for $b = 0$ and for each gradient direction
6. Addition of noise to each theoretical *SI*
7. Calculation of the observed *bD* or *D* for each gradient direction
8. Calculation of the six observed tensor elements from *D* measured in six or more noncollinear directions
9. Calculation of parameters derived from the tensor, such as the trace ($I_1 = 3D_{av}$), eigenvalues, eigenvectors, and DAIs
10. Repeat steps 6–9 many times
11. Statistical analysis
12. Repeat steps 1–11 while one or more parameters change

1. *Selection of a diffusion tensor* includes choosing the eigenvalues and the tensor orientation. It is also possible to model two or more tensors, each with different anisotropy and orientation. For each tensor, a diagonal tensor is formed from the three eigenvalues. It is common to model diffusion as cylindrically symmetric, with eigenvalues *D*1, *D*2, and *D*2, for two reasons. First, this appears to be a reasonably good approximation of diffusion in brain white matter. Second, there is one less parameter to vary in simulations, so results can be presented more concisely. However, in some brain regions the intermediate *ADC* component may have a preferential direction, suggesting that $\lambda_2 > \lambda_3$ even in the absence of noise (*30*).

Typical values of D_{av} at 37°C are about 2,400 $\mu m^2/s$ in water and 700–800 $\mu m^2/s$ in human brain, where *sRA* can reach levels of 0.7 in the splenium of the corpus callosum, corresponding to $D1 = 1{,}920$ $\mu m^2/s$ and $D2 = 240$ $\mu m^2/s$ if $D_{av} = 800$ $\mu m^2/s$ (*31*). Each tensor can be rotated by specific angles about specific axes, or it can be rotated so that the eigenvectors are in desired directions, as described in the tensor rotation section in part I.

2. *Selection of a b factor and gradient sampling scheme*. Calculation of an *ADC* requires measurements with at least two *b* factors. Usually one *b* factor is chosen as $b = 0$ in simulations, though in practice the imaging gradients usually provide a *b* factor around 1–5 s/mm^2. The second *b* factor is typically set to about 1,000 s/mm^2, but this is one of the parameters that is commonly varied in simulations, and values up to 3,000 s/mm^2 or even higher may be compared. Two or more high *b* factors were sometimes used in early DTI work, and this required a least-squares fit to calculate the *ADC* in each direction, or else a more complicated least-squares fit of the tensor elements including data with different *b* factors and different gradient directions. It is now generally accepted that a single *b* factor provides optimum results for a single *ADC*, so most applications and simulations use a single *b* factor. Because biexponential signal decay has been observed in normal brain tissue (*32–41*), no single *ADC* fully characterizes diffusion in each direction. The use of a single nonzero *b* factor in vivo provides an "effective *ADC*" for that *b* factor, whereas the use of two or more *b* factors would be complicated by combining data with two or more effective *ADC*s into a single tensor. With anisotropic diffusion, the use of two or more *b* factors may be better than a single *b* factor (*42*), but this would introduce the complication of different effective *ADC*s with biexponential signal decay.

Different gradient sampling schemes are often compared in simulations. Each gradient can be represented by a *b* factor and three directional cosines g_x, g_y, and g_z, with $g_x^2 + g_y^2 + g_z^2 = 1$. The six unique elements of the *b* matrix can then be calculated from the gradients with the proportionality factor *b* (Eqs. [II-134] and [II-135]). Although the imaging gradients and background gradients usually are ignored in simulations, they can be included when the *b* matrix is calculated, as described in the "Measuring Apparent Diffusion Coefficients" section of part II.

3. *Selection of a signal-to-noise ratio (SNR)*. The *SNR* of a DTI simulation usually refers to the *SNR* of the $b = 0$ data. Data with $b > 0$ would have lower *SNR*, depending on the value of **b**:**D**. Typical *SNR* for a single DTI acquisition is about 15:1 to 30:1. This can be increased by signal averaging. With low *SNR* and high *b* factors, the *SNR* in individual images can easily approach 3:1 or even lower. In this case it is important to consider the Rician statistics bias (see "Effects of Noise in MR Images with Low *SNR*" in a previous section). Because low *SNR* can introduce artifacts into the data, it may be useful to include simulations with high *SNR* (50:1 or higher) to reveal intrinsic properties and trends, as well as lower *SNR* to reveal systematic deviations that may occur in actual applications.

4. *Calculation of* **b**:**D** *for each gradient direction* for each tensor requires forming the inner product between the tensor and each **b** matrix (Eq. [II-130]). Tensor selection was described in step 1. The **b** matrices were selected in step 2.

5. *Calculation of the theoretical SI for b = 0 and for each gradient direction* requires calculation of Eq. [II-131] for each **b**:**D** product in step 4. Biexponential relaxation requires the use of an equation like

$$S = S_0[f\exp(-\mathbf{b}:\mathbf{D}_1) + (1-f)\exp(-\mathbf{b}:\mathbf{D}_2)] \quad [84]$$

where

$$0 \le f \le 1 \quad [85]$$

6. *Addition of noise to each theoretical SI*. Generation of random numbers with a mean of zero and a Gaussian distribution with standard deviation σ is possible with common software packages, including Excel, IDL®, and Matlab®. Because MRI signals are calculated as the magnitude signal from two quadrature channels (Eq. [58]), it is standard practice to add noise with $\sigma = S_0(b = 0)/SNR$ to each channel. The signal can be entirely in one channel, or it can be divided between the two channels as long as Eq. [58] is satisfied. The resulting noisy signal intensity is given in Eqs. [59] and [60]. The σ_R^2 and σ_I^2 terms produce a slight positive bias in the calculated *SI*, as explained in the section, "Effects of Noise in MR Images with Low *SNR*" (Eqs. [62] and [67]). The data can be left in floating-point format or they can be rounded off to the nearest integer to simulate round-off errors.

7. *Calculation of the observed bD and D for each gradient direction* is performed with Eq. [II-102], which is reproduced here in a slightly different form:

$$bD = -\ln[M/M(b=0)] = -\ln[M(b=0)] - \ln(M) \quad [86]$$

where M is the noisy signal intensity. Division by b then yields D, the effective *ADC* in this direction. For two tensors this equation must be modified to

$$bD_{\text{eff},n} = -\ln[M/M(b=0)] = \ln[f\exp(-\mathbf{b}:\mathbf{D}_1) + (1-f)\exp(-\mathbf{b}:\mathbf{D}_2)] \quad [87]$$

where $D_{\text{eff},n}$ is the effective *ADC* in the n-th direction, defined as $-\ln[M/M(b=0)]/b$.

8. *Calculation of the six observed tensor elements from D measured in six or more noncollinear directions* can be performed analytically for six directions and requires a fitting procedure for more than six directions. This was discussed in the previous section titled "Calculating the Diffusion Tensor from DWI Data." For two tensors, the 12 tensor elements are calculated from D measured in 12 or more directions, or in six or more directions with two or more b factors.

9. *Calculation of parameters derived from the tensor, such as the trace ($I_1 = 3D_{av}$), eigenvalues, eigenvectors, and DAIs*. Several parameters can be derived from the tensor, including the trace or D_{av} (see Tables I-1 through I-4), eigenvalues and eigenvectors ("Calculation of Eigenvalues and Eigenvectors" in part I), and DAIs ("Diffusion Anisotropy Indices" in part II, and Tables II-1 through II-4). With certain gradient sampling schemes, the trace and certain DAIs can be calculated directly from the diffusion measurements without calculating the tensor elements ("Selecting Gradient Directions" in part II, especially Table II-15).

10. *Repeat steps 6–9 many times*. Steps 6–9 usually are repeated at least 1,000 times, and often 10,000 to 50,000 times. More repetitions are needed with low *SNR*s so that the effects of noise will be averaged out, and the observed mean values will be close to the expectation values (the values that would be observed with infinite repetitions). For example, if a parameter has an expectation value 10 and $SD = 1$, then single observations will have a mean of 10 and *SD* of 1. If 100 repetitions are averaged, then the *SD* of the average decreases to $1/(100)^{1/2} = 0.1$, and if 10,000 repetitions are averaged, the *SD* of the average decreases to $1/(10{,}000)^{1/2} = 0.01$.

11. *Statistical analysis* may include calculation of the mean, *SD*, and bias (deviation of the expectation value from the true value) of many parameters, including the noisy *SI* (step 6), *ADC* in each direction (step 7), tensor elements (step 8), and parameters calculated from the tensor (step 9).

12. *Repeat steps 1–11 while one or more parameters change*. The entire simulation process usually is repeated, with one or more parameters varying in discrete steps. The variable could be tensor eigenvalues (D_{av} and anisotropy) or eigenvectors (orientation), *SNR*, or the gradient magnitude (b) or sampling scheme (the number of directions, M, or their orientations, or both, and the number of repetitions with $b = 0$).

Effects of Noise on DTI Data

Noise in the MRI signal intensities (*SI*s) causes scatter in parameters derived from the tensor, including the mean *ADC* (or trace), DAIs, eigenvalues, and eigenvectors (*10, 13, 14*). In addition to the increased scatter, simulations have shown that noise also induces a bias (systematic error) in DAIs (*13-16*) and eigenvalues (*10–13*). There has been an elegant explanation, based on perturbation theory, about why the eigenvalues and DAIs should be biased (*11*). The

explanations provided here are more intuitive and less formal.

It is easy to understand how noise causes apparent anisotropy when diffusion is really isotropic. In the presence of noise, it is unlikely that all three eigenvalues would have identical measured values or calculated values. Simply adding noise to the diagonal tensor elements would create unequal eigenvalues. However, this does not explain why anisotropic diffusion appears more anisotropic, and it does not completely explain the magnitude of the DAI bias with isotropic diffusion. Three mathematical explanations for the positive bias in scaled relative anisotropy (*sRA*) and fractional anisotropy (*FA*) are considered here. Because *sRA* and *FA* bias must be reflected in the eigenvalues, these biased eigenvalues would also produce a positive bias in the other DAIs. Each explanation involves one of the following considerations:

1. The *sRA* formula for six icosahedral gradient directions
2. Bias introduced by noise in the *sRA* and *FA* formulas in Table II-2
3. Off-diagonal noise in the 2D case, which can be extrapolated to 3D

None of these explanations rigorously proves that the DAIs must have a positive bias, but they help to understand why the bias exists. The first two explanations are based on the idea that if a positive number can be expressed as a fraction, and if the numerator has a greater percentage increase than the denominator, then the resulting fraction is greater than the original fraction. The percentage increase in the numerator or denominator is determined by adding a noise term (σ) to each *ADC* or tensor element and substituting the noisy values into the appropriate formula. The result will include terms without any noise components, corresponding to the true value; terms with linear noise components, σ or σ_1 σ_2; and terms with quadratic noise components, σ^2. If $\langle\sigma\rangle = 0$, the σ and σ_1 σ_2 terms do not change the expectation value of the result. The σ^2 terms are always positive and produce a bias in the expectation value. The third explanation takes a different approach, considering tensor rotation and eigenvalues instead of *sRA* and *FA* formulas.

1. *The sRA formula for six icosahedral gradient directions.* With six icosahedral gradient directions, *sRA* and *FA* can be calculated directly from the individual *ADC*s without calculating the tensor elements. Equation [II-178] can be rearranged to

$$sRA^2 = 1.25\overline{D^2}/D_{\text{av}}^2 - 1.25 \quad [88]$$

The biases in $\overline{D^2}$ and D_{av}^2 equal the noisy expectation values minus the true values. Because noise in the individual measurements is independent, terms that are first order in σ_i have expectation values of zero:

$$\langle\sigma_i\rangle = 0 \quad [89]$$

$$\langle\sigma_i D_i\rangle = 0 \quad [90]$$

$$\langle\sigma_i\sigma_j\rangle = 0 \quad [91]$$

In the presence of noise, the expectation values of $\overline{D^2}$ and D_{av}^2 for $N = 6$ measurements with $b > 0$ become

$$\langle\overline{D^2}\rangle = \frac{1}{N}\sum_{i=1}^{N}(D_i + \sigma_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(D_i^2 + 2\sigma_i D_i + \sigma_i^2) \quad [92]$$

$$\langle\overline{D^2}\rangle = \overline{D^2} + \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \quad [93]$$

$$\langle D_{\text{av}}\rangle = \frac{1}{N}\sum_{i=1}^{N}(D_i + \sigma_i) = \frac{1}{N}\left(ND_{\text{av}} + \sum_{i=1}^{N}\sigma_i\right) \quad [94]$$

$$\langle D_{\text{av}}^2\rangle = \frac{1}{N^2}\left(N^2 D_{\text{av}}^2 + \sum_{i=1}^{N}\sigma_i^2\right) \quad [95]$$

After subtracting the true values from the noisy values,

$$\langle\overline{D^2}\rangle - \overline{D^2} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \quad [96]$$

$$\langle D_{\text{av}}^2\rangle - D_{\text{av}}^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2 \quad [97]$$

Thus, the bias in the numerator of $\overline{D^2}/D_{\text{av}}^2$ is N times the bias in the denominator, with $N = 6$. Because

$$0 \le sRA^2 \le 1 \quad [98]$$

inserting this value into Eq. [88] and rearranging shows that

$$1 \leq \overline{D^2}/D_{\text{av}}^2 \leq 1.8 \quad [99]$$

Therefore, the percentage bias in the numerator is at least 6/1.8 = 3.3 times greater than the percentage bias in the denominator. As a result, the expectation value of *sRA* is greater than the true *sRA*. A similar analysis could be performed for *FA* in Eq. [II-179]. Although this analysis is strictly valid only for six icosahedral gradient directions, it does not require calculation of tensor elements. Tensor element calculations introduce correlations among the noise terms, complicating the propagation-of-error formulas.

2. *Bias introduced by noise in the sRA and FA formulas in Table II-2*. An approach similar to that used in the icosahedral example can be applied to the *sRA* and *FA* formulas in Table II-2. The scale factors are considered to be in the denominators, and we consider sRA^2 and FA^2. The numerator of both sRA^2 and FA^2 is $\mathbf{D}_{\text{an}}$:$\mathbf{D}_{\text{an}}$, where $\mathbf{D}_{\text{an}}$ is the anisotropic part of $\mathbf{D}$ (Eq. [I-67]). The bias in $\mathbf{D}_{\text{an}}$:$\mathbf{D}_{\text{an}}$ can be calculated from the identity in Eq. [I-71], keeping in mind Eqs. [89–91]:

$$\langle \mathbf{D}_{\text{an}}:\mathbf{D}_{\text{an}} \rangle - \mathbf{D}_{\text{an}}:\mathbf{D}_{\text{an}} = (2/3)(\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2) + 2(\sigma_{xy}^2 + \sigma_{xz}^2 + \sigma_{yz}^2) \quad [100]$$

The denominator biases can be calculated by a similar approach:

$$\langle 6D_{\text{av}}^2 \rangle - 6D_{\text{av}}^2 = (2/3)(\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2) \quad [101]$$

$$\langle (\mathbf{D}:\mathbf{D})/1.5 \rangle - (\mathbf{D}:\mathbf{D})/1.5 = (2/3)[\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 + 2(\sigma_{xy}^2 + \sigma_{xz}^2 + \sigma_{yz}^2)] \quad [102]$$

Each of these denominator biases is less than the numerator bias. Because $sRA \leq 1$ and $FA \leq 1$, the numerators of both *sRA* and *FA* are smaller than the denominators, and they change by a greater amount. Therefore, the percentage increase in the numerator is greater than the percentage increase in the denominator, so $\langle sRA \rangle$ and $\langle FA \rangle$ increase in the presence of noise. Qualitatively, it is expected that the bias would be greater at lower anisotropy levels, and this is consistent with published reports (*10, 13, 14, 16, 17*).

3. *Off-diagonal noise in the 2D case*. Consider the three 2D tensors

$$\mathbf{T}_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad [103]$$

$$\mathbf{T}_2 = \begin{pmatrix} \lambda_1 & D_{xy} \\ D_{xy} & \lambda_2 \end{pmatrix} \quad [104]$$

$$\mathbf{T}_3 = \begin{pmatrix} \lambda_3 & 0 \\ 0 & \lambda_4 \end{pmatrix} \quad [105]$$

where $\mathbf{T}_1$ is the true tensor, $\mathbf{T}_2$ is $\mathbf{T}_1$ with noise in the off-diagonal element D_{xy}, and $\mathbf{T}_3$ is the diagonal form of $\mathbf{T}_2$. In this example, the noisy tensor $\mathbf{T}_2$ appears to be a rotated version of the more anisotropic tensor $\mathbf{T}_3$, rather than a noisy version of the less anisotropic tensor $\mathbf{T}_1$. The apparent rotation angle of the reference frame, $-\theta$, can be calculated from Eq. [I-39].

$$-\theta_{\text{rot}} = \{\arctan[2D_{xy}/(\lambda_2 - \lambda_1)]\}/2 \quad [106]$$

According to comment 7 after Eq. [I-42], a rotation that decreases the magnitude of D_{xy} to zero must increase the difference between the two diagonal elements. Thus the expectation value of the larger eigenvalue increases and the expectation value of the smaller eigenvalue decreases. This leads to an increase in the calculated anisotropy. This is true whether D_{xy} is positive or negative, so that the apparent rotation angle θ is negative or positive. Thus, repeated DAI measurements with noisy data will not decrease $\langle \text{DAI} \rangle$. The DAI bias can only be reduced by increasing the *SNR* of the original MR images, or by signal averaging before calculation of the tensor or DAI.

The 3D case could be analyzed in a similar way, with three single-axis rotations. The result is that, in general, the largest eigenvalue appears too large, the smallest eigenvalue appears too small, and the effect on the middle eigenvalue depends on the tensor details.

For a more complete understanding of eigenvalue bias, a brief discussion of the effects of noise on eigenvectors is in order because in some cases the eigenvalue bias is linked to the eigenvectors. Because the eigenvectors must be orthogonal, an error in one eigenvector is accompanied by an error in at least one other eigenvector. With small angular deviations and well-separated eigenvalues, each noisy eigenvector can easily be associated with a true eigenvector. When two eigenvalues are identical, their eigenvectors are not unique, because any linear combination of the eigenvectors is also an eigenvector. When cylindrical symmetry is assumed for simulations, two of the original eigenvectors are not uniquely defined. Therefore, the noisy eigenvectors for these eigenvalues cannot be uniquely associated with one true eigenvector. However, the deviation from the plane of the true eigenvectors can be determined. When two or three eigenvalues are identical or similar, it is not always clear how to associate each noisy eigenvector/

eigenvalue pair with a true eigenvector/eigenvalue pair.

If the eigenvalues are sorted by magnitude, noise introduces a bias into the sorted eigenvalues. With three distinct eigenvalues, the expectation value of the largest eigenvalue is greater than the true value,

$$\langle \lambda_1 \rangle > \lambda_1 \qquad [107]$$

the expectation value of the smallest eigenvalue is smaller than the true value,

$$\langle \lambda_3 \rangle < \lambda_3 \qquad [108]$$

and the expectation value of the intermediate eigenvalue is usually near the true value (*10, 11, 14*). Of course, in any individual measurement or simulation, these relationships may not apply. With two identical eigenvalues, magnitude sorting results in one identical eigenvalue being too high and one being too low, while the unique eigenvalue is biased away from the two identical eigenvalues (*10, 11, 14*). With isotropic diffusion, magnitude sorting yields three apparently distinct eigenvalues with statistically different values (*10–12*).

In addition to the eigenvalue and DAI bias discussed previously, an additional bias occurs because of the noise-induced bias in the MRI signal intensities, as discussed in "Effects of Noise in MR Images with Low *SNR*." With $\langle M \rangle$ from Eq. [63] (or from the approximation in Eq. [67]), this bias in the expectation value increases from about 0.5% (0.5% from Eq. [67]) at $SNR = 10$ to 5.8% (5.6%) at $SNR = 3$ and 13.6% (12.5%) at $SNR = 2$. Even with reasonable *SNR* at $b = 0$, diffusion-weighted *SNR* can easily be 5–10-fold lower with high anisotropy levels if $bD_{av} \geq 1$. Thus, the *SNR* for measurement of a high *ADC* value may be very low, resulting in a large positive bias (overestimation) in *M*. When the *ADC* is calculated from Eq. [II-102] or [86], the *ADC* is underestimated (*9, 28*). The underestimate in D_{av} has been reported in an abstract, but without an explanation (*43*). The amount of bias in the largest eigenvalue and in D_{av} depends on the tensor orientation when only six gradient orientations are used. For example, the bias is greatest if the λ_1 eigenvector is parallel to one of the six gradient directions, and less if the eigenvector is midway between two gradient directions. If time permits more data acquisition, signal averaging of the raw *k*-space data or of phase-coherent images before calculating the magnitude would increase the effective *SNR*. Alternatively, the use of 12 or more directions ensures more uniform spatial coverage and a decreased influence of a single poor measurement (*8, 44, 45*).

The underestimation of λ_1 caused by MRI signal overestimation can offset the positive DAI bias, so that the final calculated DAI is closer to the true value than it would be with either effect alone. That is, if the *SNR* of the original MR images is high enough to prevent significant *SI* bias, the DAI and λ_1 would be overestimated. If the *SNR* decreases, the DAI and λ_1 biases could decrease somewhat. With six gradient directions and a suitable tensor orientation, a λ_1 underestimation could become significant. As a result, λ_1 and the DAI may begin to approach their true values, whereas D_{av} would become more underestimated (negative bias). This counterbalancing effect would be negligible with isotropic diffusion and could become significant at high anisotropy levels. A reported decrease in *FA* with increasing noise is probably caused by this effect (*9, 28*). One moral of this story is that in simulations, calculating a result near the true value does not prove the lack of a bias—a positive bias and a negative bias could counteract each other.

So far, the discussion of the effects of noise on DTI data has focused on the bias in the measured *ADC*s, D_{av}, eigenvalues, and DAIs. There is some qualitative similarity between the DAI bias and the MRI *SI* bias. Indeed, both values involve taking the sum of the squares of some numbers, so that the final result is never negative. Although the mathematical details are different, the qualitative trends are similar. It is therefore reasonable to expect a negative bias in the *SD* of DAIs, similar to the reduced *SD* of MRI *SI*s. Although the positive bias in the DAI values has been known for some time (*13–16*), the negative bias in the *SD* of DAIs appears to have been recognized only recently (*17*). Notice that the decreased *SD*s in MRI *SI*s, in their corresponding *ADC*s, and in the DAIs are accompanied by a bias in the associated values. Thus, obtaining a lower *SD* in simulations does not automatically mean that the measurement is more accurate, less biased, or better in any other sense. This *SD* bias also has implications for comparing the contrast-to-noise ratio (*CNR*) of DAIs, as discussed in "Contrast-to-Noise Ratios of Diffusion Anisotropy Indices" in part II.

Bootstrap Analysis

The variance in parameters derived from experimental DTI data can be estimated by bootstrap analysis (*4, 46, 47*). This consists of acquiring multiple replicates of data in each direction and randomly selecting (with replacement) one of the replicates in each direction

for calculation of the tensor. After many repetitions of this calculation, the mean and *SD* of various parameters can be calculated.

Important Points in "Noise and Simulations"

Several terms used in the discussion of noise were defined. Noise causes a positive bias in the measured MR signal intensity when *SNR* is low (Eq. [67]). Propagation-of-error formulas were presented in Eqs. [71–76]. The need for caution in applying propagation-of-error formulas to multiple-step propagations was demonstrated (Eqs. [81] and [83]). A 12-step approach to simulating DTI data was explained in detail. Noise in DTI data causes a positive bias in measured anisotropy, and three ways of understanding this bias were presented.

OPTIMUM *b* FACTORS

The purposes of this section are to explain 1) the criteria for choosing optimum acquisition parameters (*b* factor, and number of $b = 0$ images) for DTI measurements, and how certain factors influence the optimum *b* factor; 2) how some optimum *b* factors can be calculated analytically; and 3) how much the *b* factor and the number of $b = 0$ images can be changed without seriously affecting the quality of the results.

Criteria for Optimum *b* Factors

Because DTI is highly sensitive to noise, and the *SNR* in each MR image is low because of the long echo time (*TE*) required for diffusion weighting, it is important to optimize the collection and analysis of DTI data. Suggestions for optimally spaced gradient directions were presented in the section on selecting gradient directions in part II. The equivalence of many DAIs with respect to *CNR* was discussed in the section on contrast-to-noise ratios of diffusion anisotropy indices in part II. This section considers what *b* factor to use and how many $b = 0$ images to acquire for optimum results in various applications. If T_2 effects can be neglected, the product *bD* is the important factor. Therefore, the optimum *bD* product will often be derived. The effect of the *b* factor on T_2 losses can then be calculated for a specific T_2 and maximum gradient strength, as discussed in part II.

In discussions of *SNR* and *CNR*, the word "signal" refers to the parameter of interest (e.g., MRI signal intensity, *ADC*, or DAI value). The *SNR* of the *ADC* has been called the diffusion-to-noise ratio, or *DNR* (*48*). The optimum *b* factor is generally the one that provides the highest *SNR* for the parameter of interest in a single tissue, or the highest *CNR* between two tissues of interest. Contrast is the signal difference between two tissues, *A* and *B*:

$$\Delta S = S_B - S_A \qquad [109]$$

For *SNR*, noise is the *SD* of the single measurement, σ. For *CNR*, noise is the square root of the sum of the noise variances in the two measurements:

$$\text{Noise} = \sigma = (\sigma_A^2 + \sigma_B^2)^{1/2}. \qquad [110]$$

Thus, *SNR* and *CNR* can be calculated as

$$SNR = S_A/\sigma_A \qquad [111]$$

$$CNR = \Delta S/(\sigma_A^2 + \sigma_B^2)^{1/2} = (S_B - S_A)/(\sigma_A^2 + \sigma_B^2)^{1/2} \qquad [112]$$

The choice of an optimum *b* factor, b_{opt}, depends on the parameter being measured—diffusion-weighted signal intensity (e.g., contrast between ischemic and normal tissue), *ADC*, DAI, or eigenvector. The b_{opt} may depend on many factors, including the mean *ADC*s in the regions of interest, the amount of anisotropy, the tensor orientation, the presence or absence of differences in spin density or T_2, and whether *TE* changes when the *b* factor changes. The following sections calculate optimum *b* factors for several situations. Precision (σ) is generally within 10% of the optimum for nearly a twofold range of *b* factors, from about 70% to 130% of the optimum value (*49*). Further, because a higher *b* factor requires a longer *TE*, b_{opt} when *TE* changes is generally 10–20% less than when *TE* is constant for all *b* factors (*28, 49*). The exact decline in b_{opt} depends on the maximum gradient strength, the pulse sequence, the T_2 value(s) of the tissue(s), and the parameter being measured. For the present calculations, the *TE* is assumed to be constant, because this allows analytic expressions to be derived. The optimum *b* factors calculated here should be considered approximate starting values, not the absolute truth.

One recent publication used a numerical approach to estimate near-optimal ranges of *b* and n_2/n_1 (the ratio of the number of $b > 0$ images to the number of $b = 0$ images) for measuring D_{av}, *FA*, and the principal eigenvector direction (*50*). They considered both one-fiber and two-fiber cases. Although that work involved extensive Monte Carlo simulations, and con-

sidered some factors often ignored, it did not produce analytic formulas to allow easy extension to other situations. The discussion that follows focuses on analytic approaches to optimization. Several publications on optimizing the calculation of D_{av} from three measurements with isotropic diffusion are discussed. In addition, recent abstracts have considered analytical approaches to optimizing measurements of D_{av} (*51, 52*), anisotropy (*51*), and eigenvectors (*52*), as well as a numerical approach for eigenvectors (*51*).

Number of *b* Factors

An exponential decay curve is described by three parameters: a starting point, an end point, and a rate constant for the transition from the starting point to the end point. The considerations that apply to *ADC* measurements are similar to those that apply to T_1 relaxation (*53, 54*) and T_2 relaxation (*54*). Lessons learned about measuring any one of these processes can be applied to the other processes, with appropriate modifications.

The end point for DWI signal loss is zero, so only two parameters need to be determined experimentally. The problem of a slightly nonzero end point due to noise in magnitude images has been dealt with elsewhere (*18, 19, 21–26*) and is omitted from the present discussion.

Clearly, one data point should have the lowest possible *b* factor so that the signal and *SNR* are as large as possible. In early measurements of T_1, T_2, and *ADC*, several points on the exponential decay curves were measured in an attempt to improve the precision of the measurement. If signal decay may be biexponential, several data points may be required to define both decay rates adequately. If a wide range of decay rates is present, several data points may be required to cover the entire range adequately. However, if a single decay constant is present, and a narrow range of decay constants, precision is optimum if a single optimum data point is measured in addition to measuring the starting and end points (*48, 54, 55*). If time allows more measurements, the optimum point should be repeated, rather than adding suboptimal data points (*48, 54, 55*). It is now generally accepted that a single optimum nonzero *b* factor provides the most reliable results for a single *ADC*. For multiple *ADC*s, optimization depends on what is being measured (e.g., the individual *ADC*s, D_{av}, a DAI, eigenvalues, or eigenvectors) and on the criterion being used to evaluate the results (e.g., average over a given range, or worst case within a range) (*15, 42*).

Another reason to use a single *b* factor is that diffusion-weighted signal decay in biological tissues appears to be biexponential. The use of several data points to fit a monoexponential curve to biexponential data will yield different results for different choices of data points. Although precision may be improved by using at least two nonzero *b* factors for DWI or DTI measurements in very anisotropic systems, use of a single optimum *b* factor avoids some of the problems caused by biexponential signal decay.

Although the low *b* factor is not quite zero, it usually is chosen to be as near zero as possible. The formulas calculated below assume that the low *b* factor is zero. This will simplify some expressions by allowing the use of *b* instead of $b_2 - b_1$. If cross-terms between imaging gradients and diffusion gradients can be ignored, the small *b* factor from the imaging gradients is present with b $\sim$ 0 and $b > 0$, and will cancel when *D* is calculated from Eq. [II-102] or [86].

MRI Signal Contrast with Isotropic Diffusion

With isotropic diffusion, diffusion can be measured with a single $b = 0$ measurement and a single $b > 0$ measurement in any desired direction. In this case, the DWI contrast between two objects depends on each object's spin density, T_2, and *ADC*. The diffusion-weighted signal intensity of an object is given by (*49*)

$$S = P\ \exp(-TE/T_2)\exp(-bD) = S_0\exp(-bD) \quad [113]$$

where *P* is a function of the proton density and S_0 is the signal intensity without diffusion-sensitizing gradients ($b = 0$). The contrast between two objects with $D_A > D_B$ is given by (*49*)

$$\Delta S = S_B - S_A = S_{0B}\exp(-bD_B) - S_{0A}\exp(-bD_A) \quad [114]$$

If *TE* is constant so that S_{0A} and S_{0B} do not change with *b*, then setting $d\Delta S/db = 0$ yields the *b* factor for the maximum signal intensity difference ΔS_{max}:

$$d\Delta S/db = -D_B S_{0B}\exp(-bD_B) + D_A S_{0A}\exp(-bD_A) = 0 \quad [115]$$

$$b_{\Delta S\,max} = \frac{\ln\left(\frac{D_A}{D_B}\right) - \ln\left(\frac{S_{0B}}{S_{0A}}\right)}{D_A - D_B} \quad [116]$$

If $S_{0A} = S_{0B}$, this reduces to

$$b_{\Delta S\,\max} = \frac{\ln(D_A) - \ln(D_B)}{D_A - D_B} \quad [117]$$

In the limit as D_B approaches D_A, application of L'Hôpital's rule yields

$$b_{\Delta S\max} D = 1 \quad [118]$$

MRI Contrast-to-Noise Ratio with Anisotropic Diffusion

With anisotropic diffusion, no single D value completely describes the system. The intensity of a diffusion-weighted image depends on the direction of the diffusion-weighting gradient. The directional dependence of D and of the diffusion-weighted signal intensity can be eliminated by measuring $D_{\rm av}$ with any of the gradient sampling schemes in Tables II-6 and II-8 through II-15. The signal corresponding to $D_{\rm av}$ is $S_{\rm av}$. With M separate *ADC* measurements,

$$S_i = S_0 \exp(-\mathbf{b}_i{:}\mathbf{D}) = S_0 \exp(-bD_i) \quad [119]$$

$$D_i = \ln(S_0/S_i)/b = [\ln(S_0) - \ln(S_i)]/b \quad [120]$$

$$D_{\rm av} = \frac{1}{M}\sum_{i=1}^{M} D_i = \frac{1}{b}\left(\ln(S_0) - \frac{1}{M}\sum_{i=1}^{M}\ln(S_i)\right) \quad [121]$$

$$S_{\rm av} = S_0 \exp(-bD_{\rm av}) = \left(\prod_{i=1}^{M} S_i\right)^{1/M} \quad [122]$$

Note that $D_{\rm av}$ is the arithmetic mean of the individual D_i, whereas $S_{\rm av}$ is the geometric mean of the individual S_i, not the arithmetic mean.

When M images are combined to produce a final image (Eq. [122]), the *CNR* and the optimum b factor depend on how the noise in each initial image affects the final image. Previous work has generally assumed that the optimum b factor for anisotropic diffusion was the same as for isotropic diffusion (*28*). Based on the following calculations, it can be shown that the optimum b factor for anisotropic diffusion is always less than for isotropic diffusion (Fig. 2 in (*49*)).

With Gaussian noise, the noise variance in the final DW image is found by applying Eq. [71] to Eq. [122], yielding

$$\sigma^2(S_{\rm av}) = \left(\frac{\sigma_{\rm m} S_{\rm av}}{M}\right)^2 \sum_{i=1}^{M} S_i^{-2} \quad [123]$$

where $\sigma_{\rm m}^2$ is the noise variance in each original MR image. With isotropic diffusion each $\langle S_i\rangle = S_{\rm av}$, so

$$\sigma^2(S_{\rm av}) = \sigma_{\rm m}^2/M \text{ (isotropic)} \quad [124]$$

The individual signal intensities can be calculated from Eq. [119], contrast from Eq. [114], and the noise variance in the final diffusion-weighted signal intensity from Eq. [123]. After applying Eq. [123] to the two different tissues, substituting for each S_i from Eq. [119], and substituting for S_A/S_{0A} and S_B/S_{0B} from Eq. [113], the resulting noise variances are

$$\sigma_A^2 = \left(\frac{\sigma_{\rm m}}{M}\right)^2 \sum_{i=1}^{M} \exp[2b(D_{A,i} - D_{A,\rm av})] \quad [125]$$

$$\sigma_B^2 = \left(\frac{\sigma_{\rm m}}{M}\right)^2 \sum_{i=1}^{M} \exp[2b(D_{B,i} - D_{B,\rm av})] \quad [126]$$

The *CNR* can be calculated from Eq. [112] with the variances shown in Eqs. [125] and [126], and with ΔS calculated from Eq. [114] with D_B and D_A replaced by $D_{B,\rm av}$ and $D_{A,\rm av}$, respectively. With isotropic diffusion, the *CNR* becomes

$$CNR = (M/2)^{1/2}\Delta S/\sigma_{\rm m} \text{ (isotropic)} \quad [127]$$

Thus, isotropic *CNR* depends only on the contrast, ΔS, the noise in the individual MR images, $\sigma_{\rm m}$, and the number of measurements, M. If T_2 relaxation is ignored, the *SNR* of the $b = 0$ image is not affected by the b factor, and isotropic *CNR* is optimized by optimizing the contrast, just as it was for a single DWI measurement. This generally is not true with anisotropic diffusion, unless each individual measurement yields the same *ADC*. This could occur in two special cases. First, for three orthogonal measurements along the x, y, and z axes (scheme 3x), a cylindrically symmetric ellipsoid can be oriented with its $D1$ eigenvector along one of the directions (±1, ±1, ±1), so that it makes equal projections along the x, y, and z axes. Second, for four tetrahedral measurements (scheme 4t), a cylindrically symmetric ellipsoid can be oriented with its $D1$ eigenvector along the x, y, or z axis, so that it makes equal projections along the (±1, ±1, ±1) axes.

The *CNR* as a function of anisotropy and b factor, for a cylindrically symmetric diffusion ellipsoid with its three principal axes along the three diffusion gradient directions, was shown in (*49*).

ADC Measurements (*DNR*)

An *ADC* can be calculated for a single direction, and D_{av} can be calculated as the mean of the individual *ADC*s for several sampling schemes, including those in Tables II-6 and II-8 through II-15. The optimization described here is valid for all these cases. Although it is not strictly proven for other sampling schemes, the results are likely to be valid for any well-spaced gradient sampling scheme. It is assumed that the low b factor is 0 s/mm^2 in the following calculations. If this is not true, then b should be replaced by $b_2 - b_1$. As explained above, the small b factor from the imaging gradients will cancel when D is calculated from Eq. [II-102] or [86].

The *SNR* of the measured *ADC* has been called the diffusion-to-noise ratio, or *DNR* (*48*). Several parameters must be selected for a *DNR* measurement, including the high b factor (or the bD or bD_{av} product), the total number of measurements at $b = 0$ (n_1) and at the high b factor ($n_2 = N - n_1$), and the number of directions at the high b factor, M. Thus, the number of measurements at the high b factor in each direction is

$$n_2/M = (N - n_1)/M \qquad [128]$$

Optimization involves finding the minimum possible σ_{av}, the *SD* of D_{av} (*28, 48, 55*).

Each signal intensity is calculated from Eq. [119], the *ADC* is calculated from Eq. [120], and D_{av} is calculated from Eq. [121]. For a given noise level (*SD*) in the original images of σ_m, the *SNR*s are

$$SNR_0 = S_0/\sigma_m \qquad [129]$$

$$SNR_i = S_i/\sigma_m = S_0\exp(-bD_i)/\sigma_m = SNR_0\exp(-bD_i) \qquad [130]$$

The *SD*s of each calculated $\ln(S)$, σ_0 and σ_i, can be calculated from the propagation-of-error formula, Eq. [71]:

$$\sigma_0 = \sigma[\ln(S_0)] = \sigma_m d[\ln(S_0)]/dS_0 = \sigma_m/S_0 = 1/SNR_0 \qquad [131]$$

$$\sigma_i = \sigma[\ln(S_i)] = \sigma_m/S_i = 1/SNR_i = \exp(bD_i)/SNR_0 \qquad [132]$$

With multiple measurements of an individual intensity, if the geometric mean of each signal intensity is calculated as in Eq. [122], then the *SD*s of $\ln(S_0)$ and $\ln(S_i)$ become

$$\sigma_0 = 1/(n_1^{1/2}SNR_0) \qquad [133]$$

$$\sigma_i = \exp(bD_i)/[(n_2/M)^{1/2}SNR_0] \qquad [134]$$

The variance in each D_i can be calculated by applying Eq. [71] to Eq. [120], and the variance in D_{av} is

$$\sigma_{av}^2 = \frac{1}{b^2}\left(\frac{\sigma_0^2}{n_1} + \frac{\frac{1}{M^2}\sum_{i=1}^{M}\sigma_i^2}{(N-n_1)/M}\right) = \left(\frac{1}{SNR_0^2 b^2}\right) \times \left(\frac{1}{n_1} + \frac{1}{M(N-n_1)}\sum_{i=1}^{M}\exp(2bD_i)\right) \qquad [135]$$

As discussed above, the minimum *TE* increases along with b, so the choice of b indirectly affects SNR_0. As a result, the optimum b factor is typically 10–20% lower than that calculated for a constant *TE* at all b factors (*28, 49*).

The value of σ_{av}^2 can be minimized numerically by changing b and n_1, yielding the optimum values of b and n_1 for a given N, M, and set of D_i. For notational simplicity let

$$\Sigma_i = \frac{1}{M}\sum_{i=1}^{M}\exp(2bD_i) \qquad [136]$$

$$\Sigma_D = \frac{1}{M}\sum_{i=1}^{M}D_i\exp(2bD_i) \qquad [137]$$

When σ_{av}^2 is a minimum, the derivatives of σ_{av}^2 with respect to b and with respect to n_1 should equal 0. Setting $d\sigma_{av}^2/db = 0$ yields

$$n_2/n_1 = b\Sigma_D - \Sigma_i \qquad [138]$$

Setting $d\sigma_{av}^2/dn_1 = 0$ yields

$$n_2/n_1 = \Sigma_i^{1/2} \qquad [139]$$

When σ_{av}^2 is minimized numerically, Eqs. [138] and [139] can be used to check the validity of the solution. When both n_1 and b are varied, both equations should be satisfied. If n_1 is held constant while b is optimized, then Eq. [138] is satisfied, whereas Eq. [139] may not be satisfied. If b is held constant while n_1 is optimized, then Eq. [139] is satisfied while Eq. [138] may not be satisfied. For isotropic diffusion, Eqs. [136–139] simplify to

$$\Sigma_i = \exp(2bD) \text{ (isotropic)} \quad [140]$$

$$\Sigma_D = D \exp(2bD) \text{ (isotropic)} \quad [141]$$

$$n_2/n_1 = (bD - 1)\exp(2bD) \text{ (isotropic)} \quad [142]$$

$$n_2/n_1 = \exp(bD) \text{ (isotropic)} \quad [143]$$

For comparison with previously published work, the sensitivity factor for *ADC* measurements defined by Xing et al. (*48*), κ_D, can be modified for anisotropic diffusion:

$$DNR = D_{av}/\sigma_{av} = SNR_0\kappa_D \quad [144]$$

$$\kappa_D = \frac{bD_{av}}{\sqrt{\frac{1}{n_1} + \frac{\Sigma_i}{N - n_1}}} \quad [145]$$

Minimizing σ_{av}^2 yields the same result as maximizing κ_D. For isotropic diffusion, numerical optimization yields

$$n_2/n_1 = 3.591 \text{ (optimum)} \quad [146]$$

$$bD = 1.278 \text{ (optimum)} \quad [147]$$

Because n_1 and n_2 must be integers, the actual experimental parameters must be slightly different than these optimum values. That is, Eq. [147] can always be satisfied exactly, whereas Eq. [146] will only be satisfied approximately. For each N, the two n_2/n_1 ratios that are just above and just below 3.591 can be selected. For each n_2/n_1 ratio, bD can be optimized numerically (or from Eq. [143] for isotropic diffusion), σ_{av}^2 can be calculated from Eq. [135], and κ_D can be calculated from Eq. [145]. The n_2/n_1 ratio with the higher κ_D or the lower σ_{av}^2 provides the optimum n_1 and n_2 values for that N. The results shown in Table 1 are identical to published values for $N = 2$ to 15 (*42*) and $N = 2$ to 7 plus infinity (*54*). Similar results have been shown graphically for $N = 2$ to 16 (*55*), and slightly rounded results have been tabulated for $N = 2$ to 15 (*48*). The optimum bD product increases from 1.11 for $N = 2$ to 1.19 for $N = 3$, is in the range 1.22 to 1.34 for $N = 4$ to 16, and settles at 1.28 for large N.

When a range of anisotropies and tensor orientations is present, as in human brain, the optimum choices of bD and n_2/n_1 are not as clear. The optimum bD product and n_2/n_1 ratio for three orthogonal measurements, parallel to the three eigenvectors, can be calculated for any value of A. Compared with isotropic diffusion, the optimum b factor and n_2/n_1 ratio decrease with anisotropy. Because brain tissue contains a range of tensor orientations and a range of anisotropy levels up to $sRA > 0.7$ (*17, 31*), one approach is to optimize for $sRA \approx 0.35$ with the extreme orientation possibilities, corresponding to $A = 0$ and $A = 0.35$. Thus, the approximate optimum values are those for $A = 0.2$, or $bD = 1.09$ and $n_2/n_1 = 3.31$. Another approach is to calculate the variances for the extreme values of A, and to select bD and n_2/n_1 so that the worst case measurement is as good as possible. With six gradient directions, this worst-case approach should also include the fact that the measured DTI parameters are more uncertain with some tensor orientations than with other orientations (*44, 45*), and the optimum orientation seems to be different for eigenvectors than for D_{av} and anisotropy (unpublished observations). With 20 or more gradient directions, the orientational dependence is negligible.

Table 1 Optimum *bD* Products for Isotropic Diffusion with Several Different Values of *N*

N	n_1	n_2	bD for κ_{max}	n_2/n_1
2	1	1	1.11	1
3	1	2	1.19	2
4	1	3	1.25	3
5	1	4	1.30	4
6	1	5	1.34	5
7	2	5	1.22	2.5
8	2	6	1.25	3
9	2	7	1.27	3.5
10	2	8	1.3	4
11	2	9	1.32	4.5
12	3	9	1.25	3
13	3	10	1.27	3.33
14	3	11	1.28	3.67
15	3	12	1.30	4
∞	n_1	3.591 n_1	1.278	3.591

The range of useful bD products and n_2/n_1 ratios can be estimated by considering the range where σ_{av} and κ_D are within 10% of their optimum values. This is approximately a factor of 3 in n_2/n_1 and at least $\pm 30\%$ in bD. For the typical case of $A = 0.2$, the results are

$$0.75 < bD < 1.51 \text{ (optimum)} \quad [148]$$

$$1.11 < n_2/n_1 < 9.87 \text{ (optimum)} \quad [149]$$

If the other parameter is adjusted slightly when one parameter reaches its extreme point, these ranges could be extended slightly. Thus, it is not important to have the exact optimum values for good results, but

knowledge of the optimum range will facilitate the optimization of data acquisition.

As discussed previously, the minimum *TE* increases along with *b*, so the choice of *b* indirectly affects SNR_0. As a result, the optimum *b* factor is typically 10–20% lower than that calculated for a constant *TE* at all *b* factors (*28, 49*). The optimum n_2/n_1 ratio also decreases along with the optimum *b* factor (*28*).

Diffusion Anisotropy Indices

The optimization of acquisition parameters for measurements of diffusion anisotropy has received little attention (*28*). Although several DAIs have been suggested, *sRA* and *FA* have gained the most widespread use. Because *sRA* and *FA* have identical *CNR*, and their variances are related by Eqs. [II-36] and [II-85], optimization of one will also optimize the other. Therefore, this discussion focuses on *sRA* optimization.

As discussed previously, with icosahedral gradient encoding schemes, *sRA* can be calculated without calculating the entire tensor. The following derivation is based on such icosahedral encoding schemes. Although the derivation is not strictly valid for other encoding schemes, the range of icosahedral schemes for which it is valid (including combinations of icosahedral schemes with various rotations) suggests that it is approximately valid for any set of well-spaced gradient directions. According to Eq. [II-178],

$$sRA = \frac{\sqrt{5}\ \sigma_D}{2D_{\mathrm{av}}} = \frac{\sqrt{5(\overline{D_i^2} - D_{\mathrm{av}}^2)}}{2D_{\mathrm{av}}} \qquad [150]$$

Propagation of the error in the MR images to *sRA* can be calculated by calculating the derivative of *sRA* with respect to each original image intensity, S_0 or S_i. From Eqs. [72–76] with $f = sRA$, $x = D_{\mathrm{av}}$, $y = \overline{D_i^2}$, $s = S_0$, and $t = S_i$:

$$\frac{\partial sRA}{\partial S_0} = \frac{\partial sRA}{\partial D_{\mathrm{av}}}\frac{\partial D_{\mathrm{av}}}{\partial S_0} + \frac{\partial sRA}{\partial \overline{D_i^2}}\frac{\partial \overline{D_i^2}}{\partial S_0} \qquad [151]$$

$$\frac{\partial sRA}{\partial S_i} = \frac{\partial sRA}{\partial D_{\mathrm{av}}}\frac{\partial D_{\mathrm{av}}}{\partial S_i} + \frac{\partial sRA}{\partial \overline{D_i^2}}\frac{\partial \overline{D_i^2}}{\partial S_i} \qquad [152]$$

The formula for D_{av} is shown in Eq. [121], and $\overline{D_i^2}$ is given by

$$\overline{D_i^2} = \frac{1}{M}\sum_{i=1}^{M} D_i^2 = \frac{1}{Mb^2}\sum_{i=1}^{M} (\ln S_0 - \ln S_i)^2$$

$$= \frac{1}{Mb^2}\sum_{i=1}^{M} (\ln^2 S_0 - 2\ln S_0 \ln S_i + \ln^2 S_i) \qquad [153]$$

The derivatives in Eqs. [151] and [152] can then be calculated:

$$\frac{\partial sRA}{\partial D_{\mathrm{av}}} = \frac{-\sqrt{5}\ \overline{D_i^2}}{2D_{\mathrm{av}}^2\sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}} \qquad [154]$$

$$\frac{\partial sRA}{\partial \overline{D_i^2}} = \frac{\sqrt{5}}{4D_{\mathrm{av}}\sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}} \qquad [155]$$

$$\partial D_{\mathrm{av}}/\partial S_0 = 1/bS_0 \qquad [156]$$

$$\partial D_{\mathrm{av}}/\partial S_i = -\exp(bD_i)/MbS_0 \qquad [157]$$

$$\frac{\partial \overline{D_i^2}}{\partial S_0} = \frac{2D_{\mathrm{av}}}{bS_0} \qquad [158]$$

$$\frac{\partial \overline{D_i^2}}{\partial S_i} = \frac{-2D_i\exp(bD_i)}{MbS_0} \qquad [159]$$

Substituting these derivatives into Eqs. [151] and [152] yields

$$\frac{\partial sRA}{\partial S_0} = \frac{-\sqrt{5}\ \overline{D_i^2}}{2bS_0D_{\mathrm{av}}^2\sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}} + \frac{2\sqrt{5}\ D_{\mathrm{av}}}{4bS_0D_{\mathrm{av}}\sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}} \qquad [160]$$

$$\frac{\partial sRA}{\partial S_i} = \frac{\sqrt{5}\ \overline{D_i^2}\exp(bD_i)}{2MbD_{\mathrm{av}}^2S_0\sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}} + \frac{-2\sqrt{5}\ D_i\exp(bD_i)}{4MbD_{\mathrm{av}}S_0\sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}} \qquad [161]$$

which simplify to

$$\frac{\partial sRA}{\partial S_0} = \left(\frac{-\sqrt{5}\ \sqrt{\overline{D_i^2} - D_{\mathrm{av}}^2}}{2bD_{\mathrm{av}}^2S_0}\right) = -sRA/bD_{\mathrm{av}}S_0 \qquad [162]$$

$$\frac{\partial sRA}{\partial S_i} = \left(\frac{sRA}{bD_{av}S_0}\right)\left(\frac{\exp(bD_i)(\overline{D_i^2} - D_iD_{av})}{M(\overline{D_i^2} - D_{av}^2)}\right) \quad [163]$$

The noise in each S_0 or S_i image is σ_m, SNR_0 is defined in Eq. [129], there are n_1 acquisitions at $b = 0$ and $n_2 = N - n_1$ acquisitions at the high b factor, or $(N - n_1)/M$ acquisitions at the high b factor in each direction. After considering that each noise variance decreases as $1/n_1$ or $1/(n_2/M) = M/(N - n_1)$, the variance in *sRA* can be calculated by application of Eq. [71] to Eqs. [162] and [163].

$$\sigma_{sRA}^2 = \left(\frac{sRA}{bD_{av}SNR_0}\right)^2\left[\frac{1}{n_1} + \frac{1}{M(N - n_1)} \times \sum_{i=1}^{M} \exp(2bD_i)\left(\frac{\overline{D_i^2} - D_iD_{av}}{\overline{D_i^2} - D_{av}^2}\right)^2\right] \quad [164]$$

For completeness, the following derivatives and variances of $\overline{D_i^2}$ and of $\sigma_\lambda = 2.5^{1/2}\sigma_D$ (Eq. [II-177]) are included.

$$\partial\sigma_\lambda/\partial S_0 = 0 \quad [165]$$

$$\frac{\partial\sigma_\lambda}{\partial S_i} = \frac{\sqrt{2.5}\,(D_{av} - D_i)e^{bD_i}}{M(N - n_1)bSNR_0[\overline{D_i^2} - D_{av}^2]} \quad [166]$$

$$\sigma^2(\sigma_\lambda) = \left(\frac{1}{bSNR_0}\right)^2\left[\frac{2.5}{M(N - n_1)[\overline{D_i^2} - D_{av}^2]} \times \sum_{i=1}^{M}(D_{av} - D_i)^2e^{2bD_i}\right] \quad [167]$$

$$\sigma^2(\overline{D_i^2}) = \left(\frac{2}{bSNR_0}\right)^2\left[\frac{D_{av}^2}{n_1} + \frac{1}{M(N - n_1)}\sum_{i=1}^{M} D_i^2e^{2bD_i}\right] \quad [168]$$

Equations [135] and [164] are valid for 2D with uniformly spaced gradient directions.

Eigenvectors

Optimization of DTI parameters for eigenvector calculation has been performed with Monte Carlo simulations (*50, 51*), and an analytic calculation has appeared in abstract form (*52*). The optimum b factor seems to be similar to that for *sRA* (*51*). However, it is not necessary to acquire $b = 0$ data for eigenvector calculations, so the optimum n_2/n_1 ratio is as high as possible. Recent abstracts have discussed an apparent bias in the principal eigenvector direction (*56, 57*), but the factors involved with this bias (gradient sampling scheme, tensor orientation relative to gradient scheme, b factor, tensor fitting method) have not been explored.

Important Points in "Optimum *b* Factors"

The reason for choosing a single b factor was explained. The optimum b factor differs slightly, depending on the parameter being measured—contrast between tissues in diffusion-weighted images, *ADC*, anisotropy (DAIs), eigenvalues, or eigenvectors. If the effect of the b factor on TE_{min} is ignored, with certain gradient sampling schemes analytic expressions are available for *ADC* and *sRA* or *FA*, and these expressions appear to be good approximations for other well-spaced gradient schemes. The optimum b factor for eigenvectors has been estimated by simulations, and an analytical approach may be published soon. Results are generally good (*SD* within 10% of the optimum) for ~30% deviations from the optimum value of b, and for almost a threefold change in n_2/n_1. The optimum number of $b = 0$ acquisitions is highest for *ADC*, intermediate for DAIs, and zero for eigenvectors.

BEYOND THE TENSOR

Diffusion tensor imaging has been useful for detecting anisotropy differences in different brain regions, at different ages, and in certain diseases. It is easy to visualize the diffusion ellipsoid associated with the diffusion tensor. Furthermore, tractography with DTI has provided some promising results (*58–67*). Clearly, DTI has been a significant advance over simple DWI. However, the limits of DTI should be understood.

There are certain brain regions where the diffusion tensor does not adequately describe the diffusion process. Such regions include areas where multiple fibers meet and either cross or diverge ("kiss"). If DWI data are collected in enough directions, the data can be analyzed by more advanced methods (*68–74*). One such method is to fit the observed diffusivity profile with spherical or circular harmonics (*72–75*). Another approach is to use higher-order tensors—rank 4, 6, and even higher—in addition to the rank-2 tensor that is commonly called the diffusion tensor and corre-

sponds to the diffusion ellipsoid (*69, 70*). A third approach is to model the diffusivity as two rank-2 tensors (*50*).

Even when a single rank-2 tensor appears adequate, the signal's decay with the *b* factor's increase may not be strictly exponential, especially for the diffusion in biological tissues (*32–41*). In these cases, use of different *b* factors in addition to many directions allows more complex processing methods, including *q*-space imaging and its variations (*34, 76–80*). Such techniques may provide more complicated information, such as probability density function (PDF) or orientation density function (ODF) of the underlying diffusion process. Clearly, there is much more to be learned about the self-diffusion process in brain and other organs. In all these advanced approaches, little is known about optimization of the acquisition parameters, and the optimal values may be very different from the optimal values for simple DTI of single fibers (*50*). A solid understanding of the diffusion tensor model and DTI mathematics will provide a starting point for understanding those more advanced models and analyses.

ACKNOWLEDGMENTS

This work was supported, in part, by an award from the Bachmann-Strauss Dystonia and Parkinson Foundation and by NIH grants R01-NS047668, R03-MH64554, R01-MH64823, R01-NS043448, and K01-MH65580. The author is grateful to W. Gordon Monahan for many useful discussions and suggestions during the preparation of this manuscript, to Jianhui Zhong for helpful comments on the manuscripts, and to the reviewers for several helpful suggestions.

REFERENCES

1. Basser PJ, Mattiello J, LeBihan D. 1994. Estimation of the effective self-diffusion *tensor* from the NMR spin echo. J Magn Reson B 103:247–254.
2. Hasan KM, Narayana PA. 2003. Computation of the fractional anisotropy and mean diffusivity maps without tensor decoding and diagonalization: Theoretical analysis and validation. Magn Reson Med 50:589–598.
3. Hasan KM, Parker DL, Alexander AL. 2001. Comparison of gradient encoding schemes for diffusion-tensor MRI. J Magn Reson Imaging 13:769–780.
4. Hasan KM, Parker DL, Alexander AL. 2002. Magnetic resonance water self-diffusion tensor encoding optimization methods for full brain acquisition. Image Anal Stereol 21:87–96.
5. Chang LC, Jones DK, Pierpaoli C. 2005. RESTORE: Robust estimation of tensors by outlier rejection. Magn Reson Med 53:1088–1095.
6. Bevington PR. 1969. Data reduction and error analysis for the physical sciences. New York: McGraw-Hill.
7. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1997. Numerical recipes in C. Cambridge: Cambridge University Press.
8. Papadakis NG, Murrills CD, Hall LD, Huang CLH, Carpenter TA. 2000. Minimal gradient encoding for robust estimation of diffusion anisotropy. Magn Reson Imaging 18:671–679.
9. Jones DK, Basser PJ. 2004. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. Magn Reson Med 52:979–993.
10. Basser PJ, Pajevic S. 2000. Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. Magn Reson Med 44:41–50.
11. Anderson AW. 2001. Theoretical analysis of the effects of noise on diffusion tensor imaging. Magn Reson Med 46:1174–1188.
12. Martin KM, Papadakis NG, Huang CLH, Hall LD, Carpenter TA. 1999. The reduction of the sorting bias in the eigenvalues of the diffusion tensor. Magn Reson Imaging 17:893–901.
13. Pierpaoli C, Basser PJ. 1996. Toward a quantitative assessment of diffusion anisotropy. Magn Reson Med 36:893–906. [Published erratum: Magn Reson Med 1997;37:972].
14. Bastin ME, Armitage PA, Marshall I. 1998. A theoretical study of the effect of experimental noise on the measurement of anisotropy in diffusion imaging. Magn Reson Imaging 16:773–785.
15. Armitage PA, Bastin ME. 2001. Utilizing the diffusion-to-noise ratio to optimize magnetic resonance diffusion tensor acquisition strategies for improving measurements of diffusion anisotropy. Magn Reson Med 45: 1056–1065.
16. Armitage PA, Bastin ME. 2000. Selecting an appropriate anisotropy index for displaying diffusion tensor imaging data with improved contrast and sensitivity. Magn Reson Med 44:117–121.
17. Kingsley PB, Monahan WG. 2005. Contrast-to-noise ratios of diffusion anisotropy indices. Magn Reson Med 53:911–918. [Erratum: Magn Reson Med 2005;54:251]
18. Miller AJ, Joseph PM. 1993. The use of power images to perform quantitative analysis on low SNR MR images. Magn Reson Imaging 11:1051–1056.
19. Dietrich O, Heiland S, Sartor K. 2001. Noise correction for the exact determination of apparent diffusion coefficients at low SNR. Magn Reson Med 45:448–453.
20. Edelstein WA, Bottomley PA, Pfeifer LM. 1984. A signal-to-noise calibration procedure for NMR imaging systems. Med Phys 11:180–185.
21. van der Weerd L, Vergeldt FJ, de Jager PA, Van As H. 2000. Evaluation of algorithms for analysis of NMR relaxation decay curves. Magn Reson Imaging 18:1151–1157.
22. Henkelman RM. 1985. Measurement of signal intensi-

ties in the presence of noise in MR images. Med Phys 12:232–233. [Erratum: Med Phys 1986;13:544].
23. Bernstein MA, Thomasson DM, Perman WH. 1989. Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction. Med Phys 16:813–817.
24. McGibney G, Smith MR. 1993. An unbiased signal-to-noise ratio measure for magnetic resonance images. Med Phys 20:1077–1078.
25. Gudbjartsson H, Patz S. 1995. The Rician distribution of noisy MRI data. Magn Reson Med 34:910–914. [Letter, response, and erratum: Magn Reson Med 1996; 36:331–333].
26. Bonny JM, Zanca M, Boire JY, Veyre A. 1996. T_2 maximum likelihood estimation from multiple spin-echo magnitude images. Magn Reson Med 36:287–293.
27. Rice SO. 1944. Mathematical analysis of random noise. Bell System Tech J 23:282–332.
28. Jones DK, Horsfield MA, Simmons A. 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn Reson Med 42:515–525.
29. Poonawalla AH, Zhou XJ. 2004. Analytical error propagation in diffusion anisotropy calculations. J Magn Reson Imaging 19:489–498.
30. Lazar M, Lee JH, Alexander AL. 2005. Axial asymmetry of water diffusion in brain white matter. Magn Reson Med 54:860–867.
31. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. 1996. Diffusion tensor MR imaging of the human brain. Radiology 201:637–648.
32. Niendorf T, Dijkhuizen RM, Norris DG, van Lookeren Campagne M, Nicolay K. 1996. Biexponential diffusion attenuation in various states of brain tissue: implications for diffusion-weighted imaging. Magn Reson Med 36:847–857.
33. Beaulieu C. 2002. The basis of anisotropic water diffusion in the nervous system—a technical review. NMR Biomed 15:435–455.
34. Cohen Y, Assaf Y. 2002. High *b*-value q-space analyzed diffusion-weighted MRS and MRI in neuronal tissues—a technical review. NMR Biomed 15:516–542.
35. Schwarcz A, Bogner P, Meric P, Correze JL, Berente Z, Pál J, et al. 2004. The existence of biexponential signal decay in magnetic resonance diffusion-weighted imaging appears to be independent of compartmentalization. Magn Reson Med 51:278–285.
36. Mulkern RV, Vajapeyam S, Robertson RL, Caruso PA, Rivkin MJ, Maier SE. 2001. Biexponential apparent diffusion coefficient parametrization in adult vs newborn brain. Magn Reson Imaging 19:659–668.
37. Mulkern RV, Zengingonul HP, Robertson RL, Bogner P, Zou KH, Gudbjartsson H, et al. 2000. Multi-component apparent diffusion coefficients in human brain: relationship to spin-lattice relaxation. Magn Reson Med 44:292–300.
38. Mulkern RV, Gudbjartsson H, Westin CF, Zengingonul HP, Gartner W, Guttmann CRG, et al. 1999. Multi-component apparent diffusion coefficients in human brain. NMR Biomed 12:51–62.
39. Maier SE, Vajapeyam S, Mamata H, Westin CF, Jolesz FA, Mulkern RV. 2004. Biexponential diffusion tensor analysis of human brain diffusion data. Magn Reson Med 51:321–330.
40. Maier SE, Bogner P, Bajzik G, Mamata H, Mamata Y, Repa I, et al. 2001. Normal brain and brain tumor: Multicomponent apparent diffusion coefficient line scan imaging. Radiology 219:842–849.
41. Clark CA, Le Bihan D. 2000. Water diffusion compartmentation and anisotropy at high b values in the human brain. Magn Reson Med 44:852–859.
42. Brihuega-Moreno O, Heese FP, Hall LD. 2003. Optimization of diffusion measurements using Cramer-Rao lower bound theory and its application to articular cartilage. Magn Reson Med 50:1069–1076.
43. Basser PJ, Pierpaoli C. 1997. Estimating the principal diffusivities (eigenvalues) of the effective diffusion tensor. ISMRM 5th Scientific Meeting, p 1739 (abstract).
44. Batchelor PG, Atkinson D, Hill DLG, Calamante F, Connelly A. 2003. Anisotropic noise propagation in diffusion tensor MRI sampling schemes. Magn Reson Med 49:1143–1151.
45. Jones DK. 2004. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. Magn Reson Med 51:807–815.
46. Hasan KM, Alexander AL, Narayana PA. 2004. Does fractional anisotropy have better noise immunity characteristics than relative anisotropy in diffusion tensor MRI? An analytical approach. Magn Reson Med 51: 413–417.
47. Pajevic S, Basser PJ. 2003. Parametric and non-parametric statistical analysis of DT-MRI data. J Magn Reson 161:1–14.
48. Xing D, Papadakis NG, Huang CLH, Lee VM, Carpenter TA, Hall LD. 1997. Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain. Magn Reson Imaging 15:771–784.
49. Kingsley PB, Monahan WG. 2004. Selection of the optimum *b* factor for diffusion-weighted magnetic resonance imaging assessment of ischemic stroke. Magn Reson Med 51:996–1001.
50. Alexander DC, Barker GJ. 2005. Optimal imaging parameters for fiber-orientation estimation in diffusion MRI. Neuroimage 27:357–367.
51. Kingsley PB. 2005. Optimization of DTI acquisition parameters. ISMRM 13th Scientific Meeting, p 1294 (abstract).
52. Chang LC. 2005. Analytic expressions for the uncertainly of DTI-derived parameters and their validation using Monte Carlo methods. ISMRM Workshop on Methods for Quantitative Diffusion MRI of Human Brain, p 46 (abstract).
53. Kingsley PB. 1999. Methods of measuring spin-lattice

(T_1) relaxation times: An annotated bibliography. Concepts Magn Reson 11:243–276.
54. Jones JA, Hodgkinson P, Barker AL, Hore PJ. 1996. Optimal sampling strategies for the measurement of spin-spin relaxation times. J Magn Reson B 113:25–34.
55. Bito Y, Hirata S, Yamamoto E. 1995. Optimum gradient factors for apparent diffusion coefficient measurements. Proc Soc Magn Reson 3rd Scientific Meeting, Nice, France, p 913 (abstract).
56. Lu Y, Ding Z, Qi J, Gore JC, Anderson AW. 2005. Probabilistic connectivity mapping using the asymmetric uncertainty of diffusion tensor imaging. ISMRM 13th Scientific Meeting, p 1312 (abstract).
57. Jeong H-K, Lu Y, Ding Z, Anderson AW. 2005. Characterizing cone of uncertainty in diffusion tensor MRI. ISMRM 13th Scientific Meeting, p 1317 (abstract).
58. Mori S, Wakana S, Nagae-Poetscher LM, van Zijl PCM. 2005. MRI atlas of human white matter. Amsterdam: Elsevier.
59. Masutani Y, Aoki S, Abe O, Hayashi N, Otomo K. 2003. MR diffusion tensor imaging: recent advance and new techniques for diffusion tensor visualization. Eur J Radiol 46:53–66.
60. Basser PJ. 2002. Diffusion and diffusion tensor MR imaging. In: Attard JJ, editor. Magnetic resonance imaging of the brain and spine. Philadelphia: Lippincott Williams & Wilkins. p 197–214.
61. Mori S, van Zijl PCM. 2002. Fiber tracking: principles and strategies—a technical review. NMR Biomed 15: 468–480.
62. Mangin JF, Poupon C, Cointepas Y, Rivière D, Papadopoulos-Orfanos D, Clark CA, et al. 2002. A framework based on spin glass models for the inference of anatomical connectivity from diffusion-weighted MR data—a technical review. NMR Biomed 15:481–492.
63. Lori NF, Akbudak E, Shimony JS, Cull TS, Snyder AZ, Guillory RK, Conturo TE. 2002. Diffusion tensor fiber tracking of human brain connectivity: Acquisition methods, reliability analysis and biological results. NMR Biomed 15:494–515.
64. Bammer R, Acar B, Moseley ME. 2003. In vivo MR tractography using diffusion imaging. Eur J Radiol 45: 223–234.
65. Melhem ER, Mori S, Mukundan G, Kraut MA, Pomper MG, van Zijl PCM. 2002. Diffusion tensor MR imaging of the brain and white matter tractography. AJR Am J Roentgenol 178:3–16.
66. Leemans A, Sijbers J, Verhoye M, Van der Linden A, Van Dyck D. 2005. Mathematical framework for simulating diffusion tensor MR neural fiber bundles. Magn Reson Med 53:944–953.
67. Schwartzman A, Dougherty RF, Taylor JE. 2005. Cross-subject comparison of principal diffusion direction maps. Magn Reson Med 53:1423–1431.
68. Frank LR. 2001. Anisotropy in high angular resolution diffusion-weighted MRI. Magn Reson Med 45:935–939.
69. Özarslan E, Vemuri BC, Mareci TH. 2005. Generalized scalar measures for diffusion MRI using trace, variance, and entropy. Magn Reson Med 53:866–876.
70. Özarslan E, Mareci TH. 2003. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. Magn Reson Med 50:955–965.
71. von dem Hagen EAH, Henkelman RM. 2002. Orientational diffusion reflects fiber structure within a voxel. Magn Reson Med 48:454–459.
72. Alexander DC, Barker GJ, Arridge SR. 2002. Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. Magn Reson Med 48:331–340.
73. Zhan W, Stein EA, Yang Y. 2004. Mapping the orientation of intravoxel crossing fibers based on the phase information of diffusion circular spectrum. Neuroimage 23:1358–1369.
74. Zhan W, Gu H, Xu S, Silbersweig DA, Stern E, Yang Y. 2003. Circular spectrum mapping for intravoxel fiber structures based on high angular resolution apparent diffusion coefficients. Magn Reson Med 49:1077–1088.
75. Frank LR. 2002. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. Magn Reson Med 47:1083–1099.
76. Callaghan PT, Codd SL, Seymour JD. 1999. Spatial coherence phenomena arising from translational spin motion in gradient spin echo experiments. Concepts Magn Reson 11:181–202.
77. Basser PJ. 2002. Relationships between diffusion tensor and *q*-space MRI. Magn Reson Med 47:392–397.
78. Assaf Y, Freidlin RZ, Rohde GK, Basser PJ. 2004. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. Magn Reson Med 52:965–978.
79. Tuch DS. 2004. Q-ball imaging. Magn Reson Med 52:1358–1372.
80. Tuch DS, Reese TG, Wiegell MR, Wedeen VJ. 2003. Diffusion MRI of complex neural architecture. Neuron 40:885–895.

BIOGRAPHY

Peter B. Kingsley, MRI Research Physicist at North Shore University Hospital in Manhasset, NY, became fascinated with NMR during his undergraduate studies at Dartmouth College (B.A., 1974). At Cornell University (Ph.D., 1980) he extended the concept of "deuterated solvent" by preparing perdeuterated phospholipids for ^{1}H NMR studies of lipid-soluble compounds. After some ^{31}P NMR studies of perfused heart metabolism at the University of Minnesota (1985–1991), he helped establish a human MR spectroscopy program at St. Jude Children's Research Hospital in Memphis, Tennessee, before moving to Long Island in 1996. His current research and clinical interests include diffusion tensor imaging, MR spectroscopy, and T_1 and T_2 mapping.