A Normal Distribution for Tensor-Valued Random Variables: Applications to Diffusion Tensor MRI

Peter J. Basser* and Sinisa Pajevic

***Abstract*—Diffusion tensor magnetic resonance imaging (DT-MRI) provides a statistical estimate of a symmetric, second-order diffusion tensor of water, $\boldsymbol{D}$, in each voxel within an imaging volume. We propose a new normal distribution, $p(\boldsymbol{D}) \propto \exp(-1/2\, \boldsymbol{D} : \boldsymbol{A} : \boldsymbol{D})$, which describes the variability of $\boldsymbol{D}$ in an ideal DT-MRI experiment. The scalar invariant, $\boldsymbol{D} : \boldsymbol{A} : \boldsymbol{D}$, is the contraction of a positive definite symmetric, fourth-order precision tensor, $\boldsymbol{A}$, and $\boldsymbol{D}$. A correspondence is established between $\boldsymbol{D} : \boldsymbol{A} : \boldsymbol{D}$ and the elastic strain energy density function in continuum mechanics—specifically between $\boldsymbol{D}$ and the second-order infinitesimal strain tensor, and between $\boldsymbol{A}$ and the fourth-order tensor of elastic coefficients. We show that $\boldsymbol{A}$ can be further classified according to different classical elastic symmetries (i.e., isotropy, transverse isotropy, orthotropy, planar symmetry, and anisotropy). When $\boldsymbol{A}$ is an isotropic fourth-order tensor, we derive an explicit analytic expression for $p(\boldsymbol{D})$, and for the distribution of the three eigenvalues of $\boldsymbol{D}$, $p(\gamma_1, \gamma_2, \gamma_3)$, which are confirmed by Monte Carlo simulations. We show how $\boldsymbol{A}$ can be estimated from either real or synthetic DT-MRI data for any given experimental design. Here we propose a new criterion for an optimal experimental design: that $\boldsymbol{A}$ be an isotropic fourth-order tensor. This condition ensures that the statistical properties of $\boldsymbol{D}$ (and quantities derived from it) are rotationally invariant. We also investigate the degree of isotropy of several DT-MRI experimental designs. Finally, we show that the univariate and multivariate distributions are special cases of the more general tensor-variate normal distribution, and suggest how to generalize $p(\boldsymbol{D})$ to treat normal random tensor variables that are of third– (or higher) order. We expect that this new distribution, $p(\boldsymbol{D})$, should be useful in feature extraction; in developing a hypothesis testing framework for segmenting and classifying noisy, discrete tensor data; and in designing experiments to measure tensor quantities.**

***Index Terms*—Covariance, distribution, experimental design, fourth-order, Gaussian, normal, precision, probability, random variable, second-order, strain-energy, tensor.**

NOMENCLATURE

$\boldsymbol{x}$	Vector random variable.
$\boldsymbol{D}$	Second-order symmetric tensor random variable.
$\boldsymbol{M}$	Precision matrix.
$\boldsymbol{A}$	Fourth-order symmetric precision tensor.
$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{x}$	Quadratic function (*form*) of elements of x.
$\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}$	Quadratic function of elements of $\boldsymbol{D}$.
$p(\boldsymbol{x})$	Normal probability density function (pdf) of $\boldsymbol{x}$.
$p(\boldsymbol{D})$	normal pdf of $\boldsymbol{D}$.
$\boldsymbol{D}^0$	Mean tensor of $\boldsymbol{D}$.
$\tilde{\boldsymbol{D}}$	$\boldsymbol{D}$ (written as a vector).
λ	Lamé constant and scalar parameter used in isotropic $\boldsymbol{A}$.
μ	Shear modulus and scalar parameter used in isotropic $\boldsymbol{A}$.
δ_{ij}	Kroneker delta (3×3), and isotropic second order tensor.
γ_i	ith eigenvalue of $\boldsymbol{D}$.
ε_i	ith eigenvalue of $\boldsymbol{D}(3 \times 1)$.
γ_i	"whitened" ith eigenvalue of $\boldsymbol{D}$.
β_i	ith eigenvalue of precision matrix.
$\boldsymbol{\Sigma}$	Experimental error covariance matrix.
b_{ij}	b-matrix summarizing effects of pulse gradients on nonmagnetic resonance (NMR) signal.
$A(\boldsymbol{b})$	NMR signal intensity for given b-matrix.
$\boldsymbol{B}$	Design matrix for DT-MRI experiment.

Manuscript received October 24, 2002; revised January 11, 2003. Some material in this paper was presented in abstract form at the 2002 IEEE International Symposium on Biomedical Imaging: Macro to Nano, Washington, DC, 2002. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was M. Viergever. *Asterisk indicates corresponding author.*

*P. J. Basser is with the STBB/LIMB/NICHD, National Institutes of Health, Bldg. 13, Rm. 3W16, 13 South Drive, Bethesda, MD 20892-5772 USA (e-mail: pjbasser@helix.nih.gov).

S. Pajevic is with the MSCL/CIT, National Institutes of Health, Bethesda, MD 20892-5772 USA.

Digital Object Identifier 10.1109/TMI.2003.815059

I. INTRODUCTION

DIFFUSION tensor magnetic resonance imaging (DT-MRI) [1] provides a measurement of a symmetric second-order translational diffusion tensor of water, $\boldsymbol{D}$, for each voxel within an imaging volume. Recently, it was shown that in an ideal DT-MRI experiment, noise in the estimate of diffusion tensor element data is distributed according to a multivariate Gaussian distribution [2], [3]. In this analysis, second-order symmetric diffusion tensors were written as 6×1 vector random variables.

However, writing a tensor as a vector fails to preserve certain intrinsic algebraic relationships among its elements and their geometric relationships with the laboratory coordinate system in which the tensor elements are measured. For example, algebraic operations naturally performed on $\boldsymbol{D}$ (e.g., decomposing it into its eigenvalues and eigenvectors), or geometric operations (e.g., projecting it along a particular direction, or applying an affine transformation to it), are unwieldy when $\boldsymbol{D}$ is written as a vector.

Additionally, the vector form of the estimated covariance (or precision) matrix of tensor elements offers no insights into the way noise or features of the experimental design affects their distribution or that of other estimated tensor-derived quantities. The democratic way in which the vector representation treats tensor components makes it difficult to appreciate their unique roles.

The new tensor-variate normal distribution we propose here preserves the algebraic form and geometric structure of the

tensor random variable and, thus, our ability to perform various algebraic and geometric operations on it.

The key idea motivating this work is intuitive. Just as vector-valued data are written in vector form in the exponent of a multivariate normal distribution [4], second- (and higher) order tensors should be written in tensor form in the exponent of a *tensor*-variate normal distribution.

In this paper, we propose the form of a normal distribution for a symmetric second-order tensor random variable, $p(\boldsymbol{D})$, in which we introduce a positive definite symmetric fourth-order precision tensor, $\boldsymbol{A}$, as a parameter. We apply symmetry arguments to simplify the form of $\boldsymbol{A}$, and suggest how to classify $\boldsymbol{A}$ according to different classical elastic symmetries (i.e., isotropy, transverse isotropy, orthotropy, planar symmetry, and anisotropy). The case in which $\boldsymbol{A}$ is an isotropic fourth-order tensor is of particular importance, since it implies that the statistical properties of $p(\boldsymbol{D})$ are independent of the choice of the coordinate system in which the tensor components are measured. For this case, we derive explicit expressions for $p(\boldsymbol{D})$, and for the distribution of the three eigenvalues of $\boldsymbol{D}$, $p(\gamma_1, \gamma_2, \gamma_3)$, which are confirmed by Monte Carlo (MC) simulations. We also propose an expression that can be used to obtain sample estimates of $\boldsymbol{A}$, which one can calculate from $\boldsymbol{D}$ data. Using MC simulated DT-MRI experiments, we generate sample estimates of $\boldsymbol{A}$ for typical values of $\boldsymbol{D}$ found in gray matter, white matter, and cerebrospinal-fluid-filled regions of the brain. Finally, we show how this statistical framework can be used to aid in the design of optimal DT-MRI experiments.

II. Theory

The scalar exponent of a multivariate normal pdf, $p(\boldsymbol{x})$, contains a quadratic form, $\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}$, of an N-dimensional normal random vector, $\boldsymbol{x}$, and the precision (or inverse covariance) matrix, $\boldsymbol{M}$

$$p(\boldsymbol{x}) = \sqrt{\frac{|\boldsymbol{M}|}{(2\pi)^N}}\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}\right) = \sqrt{\frac{|\boldsymbol{M}|}{(2\pi)^N}}\exp\left(-\frac{1}{2}x_iM_{ij}x_j\right) \quad (1)$$

where $|\boldsymbol{M}|$ is the determinant of the matrix, $\boldsymbol{M}$. In tensor parlance, $\boldsymbol{x^T M x}$ is a *scalar contraction*—a linear operation that reduces one or more higher order tensors to a zeroth-order tensor (or scalar). In this case, $x_iM_{ij}x_j$[1] above is a scalar contraction of a second-order precision tensor,[2] $\boldsymbol{M}$, and the first-order tensor, $\boldsymbol{x}$. The result is a linear combination of quadratic functions formed from the products of the elements of $\boldsymbol{x}$, x_ix_j, and the corresponding elements of $\boldsymbol{M}$, M_{ij}.

In generalizing the multivariate normal distribution to a *tensor*-variate normal distribution, we seek a *tensor* analog to the quadratic form $\boldsymbol{x^T M x}$ above containing terms that are products of the elements of $\boldsymbol{D}$, $D_{ij}D_{mn}$. The most general scalar function that contains all possible linear combinations of these tensor elements is

$$D_{ij}A_{ijmn}D_{mn} \quad (2)$$

In this case, $D_{ij}A_{ijmn}D_{mn}$ is a scalar contraction of the fourth-order tensor, $\boldsymbol{A}$, and a second-order tensor, $\boldsymbol{D}$. The result is a linear combination of quadratic functions formed from products of the elements of $\boldsymbol{D}$, $D_{ij}D_{mn}$, each weighted by the corresponding elements of $\boldsymbol{A}$, A_{ijmn}.

We propose the normal distribution for a second-order tensor random variable, $\boldsymbol{D}$, of the form

$$p(\boldsymbol{D}) = c\exp\left(-\frac{1}{2}D_{ij}A_{ijmn}D_{mn}\right) \quad (3)$$

where $\boldsymbol{A}$ is a fourth-order precision tensor and c is the normalization constant to be determined below.

A. Analogies Between $D_{ij}A_{ijmn}D_{mn}$ and the Elastic Strain Energy Density

The exponent in (3), $-1/2\, D_{ij}A_{ijmn}D_{mn}$ has the same form as the strain energy density, W, (e.g., see [5]) that appears in the theory of linear elasticity.[3] Specifically, there is a direct analogy between $\boldsymbol{D}$ and the infinitesimal strain tensor, and between $\boldsymbol{A}$ and the fourth-order tensor of elastic coefficients.

In the theory of elasticity $\boldsymbol{A}$ must be positive definite to ensure that the material is elastically stable, i.e., that stresses developed within the sample always act to return the object to its equilibrium configuration [6]. In this statistical application, the same requirement must apply to ensure that the variances of the components of $\boldsymbol{D}$ are all positive.

The fourth-order precision tensor, $\boldsymbol{A}$, shares other properties with the tensor of elastic coefficients. $\boldsymbol{A}$ also possesses symmetries, which are reflected by its value being unaltered by the exchange of certain pairs of indexes. For example, since the product of two elements of the second-order tensor commutes in $D_{ij}A_{ijmn}D_{mn}$ (i.e., $D_{ij}D_{mn} = D_{mn}D_{ij}$), the corresponding coefficients of $\boldsymbol{A}$ should also be the same (i.e., $A_{ijmn} = A_{mnij}$). Moreover, since $\boldsymbol{D}$ is symmetric (i.e., $D_{ij} = D_{ji}$ and $D_{mn} = D_{nm}$), we require that $A_{ijmn} = A_{jimn}$ and $A_{ijmn} = A_{ijnm}$, respectively.[4] Owing to these symmetry conditions, there are at most 21 independent elements of $\boldsymbol{A}$ that we must specify *a priori* [8], or estimate from sample data.

In the theory of elasticity, these symmetry conditions arise because W should be independent of the coordinate system in which the components of the strain tensor are measured (e.g., see [8]). This requirement applies equally well to $p(\boldsymbol{D})$. The probability that a particular tensor arises is an intrinsic property that should be independent of the coordinate system in which $\boldsymbol{D}$ is written. This requirement, that $\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}$ is a rotationally invariant quantity, also ensures that $\boldsymbol{A}$ is a fourth-order tensor, by a simple application of the Quotient Rule Theorem [9].

The theory of elasticity also provides us with a scheme to classify fourth-order tensors of elastic coefficients according

[1]We use the Einstein summation convention in which indexes that are repeated in the expression are summed over the range of their allowable values. So, for example, $x_iM_{ij}x_j$ means $\sum_{i=1}^{3}\sum_{j=1}^{3}x_iM_{ij}x_j$.

[2]$\boldsymbol{M}$ is usually referred to as a matrix, but it actually transforms as a second-order tensor.

[3]W measures the amount of internal energy stored as a homogeneous elastic body deforms.

[4]It is known that a fourth-order tensor possessing these symmetry properties given above is self-adjoint (e.g., see [7]).

to the number, types, and degrees of symmetries they possess. The most general linear constitutive law of an elastic solid corresponds to *anisotropy* (or *aeolotropy*), requiring all 21 constants to specify the form of the tensor of elastic coefficients [5]. Other models of elastic behavior require fewer constants (e.g., see [5]). These include the cases of *planar symmetry*, requiring 13 elastic coefficients; *orthotropy*, requiring nine elastic coefficients; *transverse isotropy*, requiring five elastic coefficients; and *isotropy*, requiring only two elastic coefficients. Below, we analyze the most tractable and important case in detail, *isotropy*.

B. Relationship Between the Fourth-Order Precision Tensor, $\boldsymbol{A}$, and the 6 × 6 Precision Matrix, $\boldsymbol{M}$

The scalar contraction, $D_{ij}A_{ijmn}D_{mn}$, above can also be recast as a quadratic form, $\tilde{D}_r M_{rs} \tilde{D}_s$, in which the random second-order tensor, $\boldsymbol{D}$, is rewritten as a six-dimensional (6-D) column vector, $\tilde{D}_r = (D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz})^T$, and $\boldsymbol{M}$ is a 6 × 6 symmetric matrix. An important result that is often used in continuum mechanics, and which we also exploit here, is that any fourth-order tensor, $\boldsymbol{A}$, satisfying the symmetry properties given in the previous section, can be mapped to a 6 × 6 symmetric matrix $\boldsymbol{M}$. Both $\boldsymbol{A}$ and $\boldsymbol{M}$ contain the same 21 independent coefficients (e.g., see [5], [7], and [10]). This correspondence allows us to construct a 6 × 6 precision matrix, $\boldsymbol{M}$, from any fourth-order precision tensor, $\boldsymbol{A}$ and, thus, to construct a corresponding multivariate normal distribution directly from a tensor-variate normal distribution. Below, we use this correspondence to calculate the normalization constant for the tensor-variate distribution using the mathematical machinery developed for multivariate distributions.

C. Normalization Constant for the Tensor-Variate Normal Distribution

We obtain the normalization constant of the tensor-variate normal pdf by integrating the distribution over the entire range of all six independent elements of the symmetric tensor, $\boldsymbol{D}$. This integration is carried out in the following way. We require that

$$1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} c \exp\left(-\frac{1}{2} D_{ij} A_{ijmn} D_{mn}\right) dD = c \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}\right) dD. \quad (4)$$

Here, the tensor dot product ":" denotes the contraction of second-order tensors with the fourth-order precision tensor.[5] If the random tensor has a nonzero mean tensor, $\boldsymbol{D}^0$, we can always center the distribution about its mean using $\boldsymbol{Y} = \boldsymbol{D} - \boldsymbol{D}^0$. Then, the distribution becomes

$$\frac{1}{c} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \boldsymbol{Y}:\boldsymbol{A}:\boldsymbol{Y}\right) dY. \quad (5)$$

The exponent in the integrand can be rewritten as a quadratic form, where the coefficients of the quadratic terms in $\boldsymbol{Y}$ are contained in the matrix, $\boldsymbol{M}$

$$\frac{1}{c} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \tilde{Y}^T \boldsymbol{M} \tilde{Y}\right) d\tilde{Y} \quad (6)$$

where $\tilde{Y} = (Y_{xx}, Y_{yy}, Y_{zz}, Y_{xy}, Y_{xz}, Y_{yz})^T$ and (7), shown at the bottom of page, holds. The integral in (6) is known from the theory of multivariate normal distributions (e.g., see [11]); the normalization constant is readily obtained from $\boldsymbol{M}$

$$c = \sqrt{\frac{|\boldsymbol{M}|}{(2\pi)^6}} = \frac{1}{(2\pi)^3} \sqrt{|\boldsymbol{M}|}. \quad (8)$$

By writing $\boldsymbol{M}$ as four (3 × 3) square block matrices, as shown in (7), and by noting that the diagonal block matrices are symmetric (i.e., $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^T$ and $\boldsymbol{\Psi} = \boldsymbol{\Psi}^T$), we can write $|\boldsymbol{M}|$ in (8) simply in terms of these four block matrices

$$|\boldsymbol{M}| = |\boldsymbol{\Gamma} - \boldsymbol{\Xi}\boldsymbol{\Psi}^{-1}\boldsymbol{\Xi}^T||\boldsymbol{\Psi}|. \quad (9)$$

So, the tensor-variate normal distribution with precision tensor, $\boldsymbol{A}$, and mean tensor, $\boldsymbol{D}^0$, is

$$p(\boldsymbol{D}) = \sqrt{\frac{|\boldsymbol{\Gamma} - \boldsymbol{\Xi}\boldsymbol{\Psi}^{-1}\boldsymbol{\Xi}^T||\boldsymbol{\Psi}|}{(2\pi)^3}} \times \exp\left(-\frac{1}{2}\left(D_{ij} - D^0_{ij}\right) A_{ijmn} \left(D_{mn} - D^0_{mn}\right)\right). \quad (10)$$

The distribution in (10) possesses the basic form and properties of a normal distribution. Since $\boldsymbol{A}$ is positive definite, $\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}$ is always nonnegative, and $1 > p(\boldsymbol{D}) \geq 0$. Also, the exponent in (10) is a quadratic function of the random variable (in this case, a tensor random variable) whose mean and precision tensors appear in the exponent in an analogous way to the mean vector and precision matrix of the multivariate normal distribution.

[5]In the case of tensor products between tensors of unequal order, such as $\boldsymbol{D}$ and $\boldsymbol{A}$, we use the definition, $\boldsymbol{D}:\boldsymbol{A} = D_{ij}A_{klmn}\delta_{ik}\delta_{jl} = D_{ij}A_{ijmn}$.

$$M = \begin{pmatrix} A_{xxxx} & A_{xxyy} & A_{xxzz} & 2A_{xxxy} & 2A_{xxxz} & 2A_{xxyz} \\ A_{xxyy} & A_{yyyy} & A_{yyzz} & 2A_{yyxy} & 2A_{xzyy} & 2A_{yyyz} \\ A_{xxzz} & A_{yyzz} & A_{zzzz} & 2A_{xyzz} & 2A_{xzzz} & 2A_{yzzz} \\ 2A_{xxxy} & 2A_{xyyy} & 2A_{xyzz} & 4A_{xyxy} & 4A_{xyxz} & 4A_{xyyz} \\ 2A_{xxxz} & 2A_{xzyy} & 2A_{xzzz} & 4A_{xyxz} & 4A_{xzxz} & 4A_{xzyz} \\ 2A_{xxyz} & 2A_{yyyz} & 2A_{yzzz} & 4A_{xyyz} & 4A_{xzyz} & 4A_{yzyz} \end{pmatrix} = \begin{pmatrix} \Gamma & \Xi \\ \Xi^T & \Psi \end{pmatrix} \quad (7)$$

In fact, we can exploit the formal correspondence between the tensor-variate normal distribution in (10) and the multivariate normal distribution in (1) to obtain many properties of the tensor-variate normal distribution by using mathematical tools and approaches that have already been developed to analyze multivariate distributions (e.g., see [11]).

D. $p(\boldsymbol{D})$ when $\boldsymbol{A}$ is a General Isotropic Fourth-Order Tensor

We now derive the explicit form of $p(\boldsymbol{D})$ for the case in which $\boldsymbol{A}$ is a general isotropic fourth-order tensor, $\boldsymbol{A}^{\text{iso}}$. In this context, isotropy means that the precision tensor is rotationally invariant, i.e., its form is unchanged under any proper rotation, reflection, or inversion of coordinates in which the components of $\boldsymbol{D}$ are measured.

When $\boldsymbol{D}$ is a symmetric tensor, the most general form of $\boldsymbol{A}^{\text{iso}}$ is (e.g., see [8] and [10])

$$A^{\text{iso}}_{ikpm} = \lambda\delta_{ik}\delta_{mp} + \mu(\delta_{im}\delta_{kp} + \delta_{ip}\delta_{km}) \tag{11}$$

where μ and λ are as yet undetermined constants,[6] and $\boldsymbol{\delta}$ is the second-order isotropic tensor. This choice, $\boldsymbol{A} = \boldsymbol{A}^{\text{iso}}$ also corresponds to the tensor of elastic coefficients for a general isotropic linearly elastic solid.

The scalar contraction of the exponent of the tensor-variate normal distribution, $\boldsymbol{D}:\boldsymbol{A}^{\text{iso}}:\boldsymbol{D}$, becomes

$$D_{ik}A^{\text{iso}}_{ikpm}D_{pm} = D_{ik}\left(\lambda\delta_{ik}\delta_{mp} + \mu(\delta_{im}\delta_{kp} + \delta_{ip}\delta_{km})\right)D_{pm}. \tag{12}$$

In Appendix A, we show that this expression reduces to a linear combination of two scalar invariants of $\boldsymbol{D}$, i.e.,

$$\mathrm{D}:\boldsymbol{A}^{\text{iso}}:\boldsymbol{D} = \lambda\left(Trace(\boldsymbol{D})\right)^2 + 2\mu Trace(\boldsymbol{D}^2). \tag{13}$$

The distribution $p(\boldsymbol{D})$ must assume the same form under any proper rotation, reflection or inversion of laboratory coordinates because it depends only on functions of $\boldsymbol{D}$ that are rotationally invariant, $\mathrm{Trace}(\boldsymbol{D})$ and $\mathrm{Trace}(\boldsymbol{D}^2)$. Thus, we find that isotropy of $\boldsymbol{A}$ implies rotational invariance of $p(\boldsymbol{D})$.

If $\boldsymbol{D}$ is a tensor whose mean is $\boldsymbol{D}^0$, it is also easy to show that the tensor contraction in (13) becomes

$$\begin{aligned}(\boldsymbol{D}-\boldsymbol{D}^0):\boldsymbol{A}^{\text{iso}}:(\boldsymbol{D}-\boldsymbol{D}^0) &= \lambda\left(Trace(\boldsymbol{D}-\boldsymbol{D}^0)\right)^2 \\ &\quad +2\mu\, Trace\left((\boldsymbol{D}-\boldsymbol{D}^0)^2\right)\end{aligned} \tag{14}$$

so that $p(\boldsymbol{D}-\boldsymbol{D}^0)$ is also rotationally invariant in this more general case.

To obtain the form of $p(\boldsymbol{D})$ using $\boldsymbol{A}^{\text{iso}}$ in (11), we again write $\boldsymbol{D}$ as a vector, $\tilde{D} = (D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz})^T$, and rewrite the scalar contraction in (13) as a quadratic form, $\tilde{D}^T\boldsymbol{M}\tilde{D}$. Then, the precision matrix, $\boldsymbol{M}$, from (7) becomes

$$\boldsymbol{M} = \begin{pmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 4\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & 4\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 4\mu \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}^T & \boldsymbol{\Psi} \end{pmatrix}. \tag{15}$$

[6]In continuum mechanics, λ and μ correspond to the Lamé constant and shear modulus of the isotropic material, respectively.

Clearly, since the block matrix, $\boldsymbol{\Gamma}$, is not diagonal, the three diagonal elements of $\boldsymbol{D}$ are mutually correlated. However, the structure of $\boldsymbol{\Gamma}$ implies that coupling among D_{xx}, D_{yy}, and D_{zz} is independent of their size and of the particular choice of the x, y, and z axes in the laboratory coordinate frame. Since $\boldsymbol{\Xi} = \boldsymbol{0}$, diagonal elements of $\boldsymbol{D}$ are not correlated with off-diagonal elements of $\boldsymbol{D}$. However, since $\boldsymbol{\Psi} = 4\mu\boldsymbol{I}$, where $\boldsymbol{I}$ is the 3 $\times$ 3 identity matrix, the three off-diagonal elements of $\boldsymbol{D}$ are mutually uncorrelated, and have equal variances.

Using (15), $p(\boldsymbol{D})$ simplifies to

$$\begin{aligned} p(\boldsymbol{D}) &= p(D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}) \\ &= p(D_{xx}, D_{yy}, D_{zz})p(D_{xy})p(D_{xz})p(D_{yz}) \end{aligned} \tag{16}$$

where

$$\begin{aligned} p(D_{xx}, D_{yy}, D_{zz}) &= \sqrt{\frac{4\mu^2(2\mu+3\lambda)}{(2\pi)^3}} \\ &\times \exp\left(-\frac{1}{2}\left(D_{xx}-\overline{D}_{xx}, D_{yy}-\overline{D}_{yy}, D_{zz}-\overline{D}_{zz}\right)\right. \\ &\left.\times \begin{pmatrix} 2\mu+\lambda & \lambda & \lambda \\ \lambda & 2\mu+\lambda & \lambda \\ \lambda & \lambda & 2\mu+\lambda \end{pmatrix} \begin{pmatrix} D_{xx}-\overline{D}_{xx} \\ D_{yy}-\overline{D}_{yy} \\ D_{zz}-\overline{D}_{zz} \end{pmatrix}\right) \end{aligned} \tag{17a}$$

and

$$\begin{aligned} p(D_{xy}) &= \sqrt{\frac{2\mu}{\pi}}\exp\left(-2\mu(D_{xy}-\overline{D}_{xy})^2\right) \\ p(D_{xz}) &= \sqrt{\frac{2\mu}{\pi}}\exp\left(-2\mu(D_{xz}-\overline{D}_{xz})^2\right) \\ p(D_{yz}) &= \sqrt{\frac{2\mu}{\pi}}\exp\left(-2\mu(D_{yz}-\overline{D}_{yz})^2\right) \end{aligned} \tag{17b}$$

where it is assumed that the mean of $\boldsymbol{D}$ is given by $\overline{D}$ in the distributions above.

Thus, a meaningful distinction between the new tensor-variate and multivariate normal distributions is the way in which their covariances are characterized. While $\boldsymbol{A}^{\text{iso}}$, given in (11), is an isotropic fourth-order tensor, the corresponding $\boldsymbol{M}$ matrix for the multivariate distribution, given in (15) has a nonintuitive block form, which is clearly not a 6-D isotropic precision matrix. Only in the special case in which $\lambda = 0$, when all elements of $\boldsymbol{D}$ are independently distributed, is $\boldsymbol{M}$ a diagonal matrix. Even then, all of its diagonal elements are still not equal. Clearly, the relationship between the tensor-variate and multivariate normal distributions is not a trivial one.

In the subsequent sections, we will use the new tensor-variate distribution for $\boldsymbol{A} = \boldsymbol{A}^{\text{iso}}$ to obtain an analytical expression for the distribution of the eigenvalues of $\boldsymbol{D}$, and to design optimal DT-MRI experiments.

E. The Distribution of the Eigenvalues of $\boldsymbol{D}$ for $\boldsymbol{A} = \boldsymbol{A}^{\text{iso}}$

For $\boldsymbol{A} = \boldsymbol{A}^{\text{iso}}$ in (12) we can immediately obtain the joint probability distribution of γ_1, γ_2, and γ_3, the three eigenvalues of $\boldsymbol{D}$. The distribution, $p(\gamma_1, \gamma_2, \gamma_3)$ is a special case of $p(\boldsymbol{D})$ in (17a) and (b), obtained by performing a principal coordinate transformation in which the three diagonal elements of $\boldsymbol{D}$ are mapped to the three eigenvalues of $\boldsymbol{D}$. Integrating over all pos-

sible values of the off-diagonal elements, and substituting γ_1, γ_2, and γ_3 for D_{xx}, D_{yy}, and D_{zz} in the distribution above, we obtain

$$p(\gamma_1,\gamma_2,\gamma_3) = \sqrt{\frac{4\mu^2(2\mu+3\lambda)}{(2\pi)^3}} \times \exp\left(-\frac{1}{2}(\gamma_1-\overline{\gamma}_1,\gamma_2-\overline{\gamma}_2,\gamma_3-\overline{\gamma}_3) \times \begin{pmatrix} 2\mu+\lambda & \lambda & \lambda \\ \lambda & 2\mu+\lambda & \lambda \\ \lambda & \lambda & 2\mu+\lambda \end{pmatrix} \times \begin{pmatrix} \gamma_1-\overline{\gamma}_1 \\ \gamma_2-\overline{\gamma}_2 \\ \gamma_3-\overline{\gamma}_3 \end{pmatrix}\right) \quad (18)$$

where $\overline{\gamma}_1$, $\overline{\gamma}_2$, and $\overline{\gamma}_3$ are the three mean eigenvalues. Equivalently, we can obtain the result in (18) by substituting the expressions, $\mathrm{Trace}(\boldsymbol{D}) = \gamma_1+\gamma_2+\gamma_3$ and $\mathrm{Trace}(\boldsymbol{D}^2) = \gamma_1^2+\gamma_2^2+\gamma_3^2$ into (13), and by collecting terms.[7]

The joint distribution of the eigenvalues of $\boldsymbol{D}$ is characterized by only two parameters, μ and λ. While the eigenvalues are correlated, their coupling is independent of their order or assignment (which is not the case for eigenvalues of random matrices described by a Wishart distribution [4]). This finding follows because $p(\boldsymbol{D})$ with its exponent given in (13), depends only on $\mathrm{Trace}(\boldsymbol{D}^2)$ and $(\mathrm{Trace}(\boldsymbol{D}))^2$, scalar invariants of $\boldsymbol{D}$, which are inherently insensitive to the order of the eigenvalues. Thus, permuting the eigenvalue order will always leave this distribution unchanged.

We can uncorrelate or "whiten" $p(\gamma_1,\gamma_2,\gamma_3)$ by diagonalizing Γ in (15) and (18) using its three eigenvalues, $\beta_1 = 3\lambda+2\mu$, $\beta_2 = 2\mu$, and $\beta_3 = 2\mu$; and its three corresponding normalized eigenvectors, $(1/\sqrt{3})(1,1,1)$, $(1/\sqrt{2})(0,1,-1)$, and $\sqrt{2/3}(1,-1/2,-1/2)$. In the principal frame of Γ, $p(\gamma_1',\gamma_2',\gamma_3')$ is simply the product of three independent univariate normal distributions

$$\begin{aligned} p(\gamma_1') &= \sqrt{\frac{2\mu+3\lambda}{2\pi}}\exp\left(-\left(\frac{2\mu+3\lambda}{2}\right){\gamma_1'}^2\right) \\ p(\gamma_2') &= \sqrt{\frac{\mu}{\pi}}\exp\left(-\mu{\gamma_2'}^2\right) \\ p(\gamma_3') &= \sqrt{\frac{\mu}{\pi}}\exp\left(-\mu{\gamma_3'}^2\right) \end{aligned} \quad (19)$$

when we use the transformed random variables

$$\begin{aligned} \gamma_1' &= \frac{1}{\sqrt{3}}(\gamma_1+\gamma_2+\gamma_3) \\ \gamma_2' &= \frac{1}{\sqrt{2}}(\gamma_2-\gamma_3) \\ \gamma_3' &= \sqrt{\frac{2}{3}}\left(\gamma_1-\frac{(\gamma_2+\gamma_3)}{2}\right) \end{aligned}$$

or

$$\begin{pmatrix}\gamma_1' \\ \gamma_2' \\ \gamma_3'\end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix}\begin{pmatrix}\gamma_1 \\ \gamma_2 \\ \gamma_3\end{pmatrix}. \quad (20)$$

[7]N.B. The theoretical distribution, $p(\gamma_1,\gamma_2,\gamma_3)$, may not always conform to an empirical distribution obtained, e.g. by using MC simulations, because of the well-known sorting artifact that occurs when one orders calculated eigenvalues.

Interestingly, γ_1' is proportional to $\mathrm{Trace}(\boldsymbol{D})$, which measures the average size of the isotropic part of $\boldsymbol{D}$. The other two variables, γ_2' and γ_3', characterize the anisotropic part of $\boldsymbol{D}$. Specifically, γ_3' measures the difference between the predominant eigenvalue and the average of the two remaining eigenvalues, while γ_2' measures the difference between the two latter eigenvalues. Together, γ_1', γ_2', and γ_3' represent novel parameters with which to characterize the size and shape of the probability ellipsoid [8] one can construct from $p(\boldsymbol{D})$.

More importantly, the coefficients μ and λ in the isotropic fourth-order precision tensor can easily be related to the variances of $p(\gamma_1',\gamma_2',\gamma_3')$ in (19)

$$\sigma_T^2 = \frac{1}{2\mu+3\lambda} \quad \text{and} \quad \sigma_S^2 = \frac{1}{2\mu} \quad (21)$$

Above, the variances, σ_T^2 and σ_S^2, correspond roughly to the "Trace" and the "Skewness" of the uncertainty of $\boldsymbol{D}$, respectively, as in (20). Since σ_{T}^2 and σ_{S}^2 can be estimated statistically from sample data, μ and λ can now be expressed in terms of measurable parameters,[9] σ_{T}^2 and σ_{S}^2

$$\mu = \frac{1}{2\sigma_S^2} \quad \text{and} \quad \lambda = \frac{1}{3}\left(\frac{1}{\sigma_T^2}-\frac{1}{\sigma_S^2}\right) \quad (22)$$

This result allows us to write $\boldsymbol{A}^{\mathrm{iso}}$ in (12) explicitly in terms of σ_{T}^2 and σ_{S}^2

$$A_{ikpm}^{\mathrm{iso}} = \frac{1}{3}\left(\frac{1}{\sigma_T^2}-\frac{1}{\sigma_S^2}\right)\delta_{ik}\delta_{mp} + \frac{1}{2\sigma_S^2}\left(\delta_{im}\delta_{kp}+\delta_{ip}\delta_{km}\right). \quad (23)$$

Note that in deriving $p(\gamma_1',\gamma_2',\gamma_3')$, we make no explicit assumption that all eigenvalues are positive, i.e., that $\boldsymbol{D}$ is positive definite. This condition could be added as a constraint to the tensor-variate distribution but the distribution would no longer be Gaussian. As an aside, $p(\boldsymbol{D})$ provides no explicit information about the distribution of eigenvectors of $\boldsymbol{D}$ when $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{iso}}$.

F. MC Simulations of $p(\gamma_1,\gamma_2,\gamma_3)$

In Fig. 1, we plot MC estimates of σ_T and σ_S versus their theoretical values obtained from (21). First, MC estimates of $\boldsymbol{D}$ are generated from a multivariate normal distribution with the precision matrix given in (15). Then, the eigenvalues, γ_1, γ_2, and γ_3 are computed for each $\boldsymbol{D}$, and an empirical distribution, $p(\gamma_1',\gamma_2',\gamma_3')$, is constructed, from which σ_T and σ_S are estimated. Agreement between the analytical distribution in (18) and MC simulated data is excellent. Values of μ and λ are chosen randomly within their allowable range (as described in footnote 9), but so that the distributions of distinct eigenvalues

[8]Surfaces of constant probability are obtained for the distribution of the eigenvalues of $\boldsymbol{D}$ above by setting the exponent to a constant, e.g.,

$$(\gamma_1,\gamma_2,\gamma_3)\begin{pmatrix} 2\mu+\lambda & \lambda & \lambda \\ \lambda & 2\mu+\lambda & \lambda \\ \lambda & \lambda & 2\mu+\lambda \end{pmatrix}\begin{pmatrix}\gamma_1 \\ \gamma_2 \\ \gamma_3\end{pmatrix} = 1.$$

This quadratic form can be represented by a cylindrically symmetric, pancake- or cigar-shaped, three–dimensional ellipsoid (e.g., see [12]) whose three principal axes are σ_T, σ_S, and σ_S.

[9]N.B. In order for the exponent of the tensor-variate cumulative distribution to be unitless, μ and λ must be in units that are the inverse square of the eigenvalues' units, consistent with μ and λ being inverse variances. To ensure positive definiteness of the covariance matrix, we also require that $\mu \geq 0$ and $\lambda \geq -2\mu/3$.

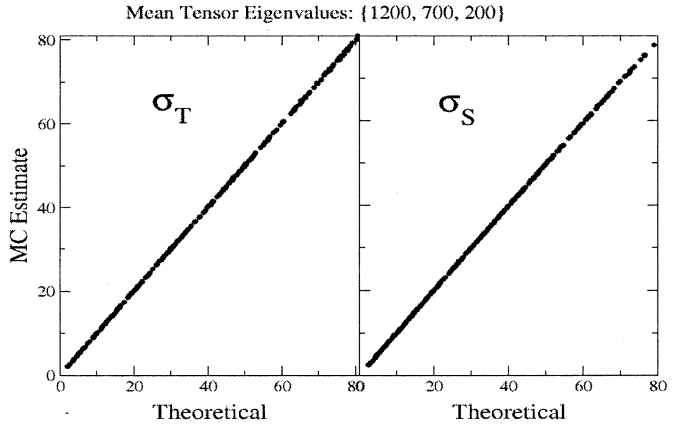

Fig. 1. Three-hundred points from MC simulations of second-order tensor, $\boldsymbol{D}$, with $(\gamma_1, \gamma_2, \gamma_3) = (1200,\ 700,\ 200)$, typical of brain white matter. The precision of MC estimates was 0.2%. This figure indicates that (19) precisely and accurately predicts the uncertainty of the estimated eigenvalues of $\boldsymbol{D}$.

do not overlap, thus avoiding a known "sorting" artifact that would bias the estimates of σ_T and σ_S (see [13]).

G. Optimal Experimental Design—The Rotational Invariance Principle

Several groups have proposed methods for optimally designing DT-MRI experiments in which independent experimental variables, such as the number of gradient acquisitions, the gradient directions, and gradient strengths, are chosen to minimize some objective or performance measure [15]–[19]. Skare *et al.* have proposed minimizing the condition number of the covariance matrix of the estimated diffusion tensor elements [15], while Jones and Papadakis minimize the orientational dependence of the variance of the fractional anisotropy (FA) [16]–[19].

Here, we propose that a *necessary* condition for an optimally designed DT-MRI experiment is that $\boldsymbol{A}$ be an isotropic fourth-order tensor, $\boldsymbol{A}^{\text{iso}}$ of the form given in (11). Since $\boldsymbol{A}$ describes the observed variations of $\boldsymbol{D}$ due to background noise in the measurements, this condition will ensure that $p(\boldsymbol{D})$, and consequently all tensor-derived quantities (e.g., FA, Trace, and the relative anisotropy), have orientationally invariant statistical properties. Certainly, the constraint that $\boldsymbol{A}$ be an isotropic fourth-order tensor can be used in conjunction with other objective functions or performance measures.

To analyze different DT-MRI experimental designs we first consider the log-linear form of the basic model relating the NMR signal to the diffusion tensor [20]

$$\ln\left(\frac{A(\boldsymbol{b})}{A(0)}\right) = -b_{xx}D_{xx} - b_{yy}D_{yy} - b_{zz}D_{yy} - 2b_{xy}D_{xy} - 2b_{xz}D_{xz} - 2b_{yz}D_{xz} \tag{24}$$

where $A(\boldsymbol{b})/A(0)$ is the measured echo intensity, b_{ij}s are the elements of the symmetric b-matrix constructed from all applied gradient waveforms. The predicted form of the precision matrix for this model is given by

$$\boldsymbol{M}^{\text{pred}} = \boldsymbol{B}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{B} \tag{25}$$

where $\boldsymbol{\Sigma}$ is the error covariance matrix, and $\boldsymbol{B}$ is the experimental design matrix whose ith row, $B^{\text{i}} = (\text{b}^{\text{i}}_{\text{xx}}, \text{b}^{\text{i}}_{\text{yy}}, \text{b}^{\text{i}}_{\text{zz}}, 2\text{b}^{\text{i}}_{\text{xy}}, 2\text{b}^{\text{i}}_{\text{xz}}, 2\text{b}^{\text{i}}_{\text{yz}})$, contains the b-matrix elements associated with the ith gradient acquisition [20].

It is reasonable to assume that experimental variances are uncorrelated in the MR experiment, so that Σ is diagonal [20]. However, Batchelor further assumes that all experimental variances are equal.[10] Then, (25) becomes $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, and $\boldsymbol{M}^{\text{pred}} = (1/\sigma^2)\mathbf{B}^{\text{T}}\mathbf{B}$. Moreover, it is sometimes possible to design MR sequences in which $b_{\text{jk}} \propto \text{G}_{\text{j}}\text{G}_{\text{k}}$, where G_{j} and G_{k} represent the peak magnetic field gradients (diffusion gradients) applied along the jth and kth coordinate directions, x_{j} and x_{k}, respectively.[11] In the special case in which all gradients used in an experiment have the same magnitude (i.e., $\text{G}_{\text{j}} = \text{G}_{\text{k}}$) then $b_{\text{jk}} \propto \text{x}_{\text{j}}\text{x}_{\text{k}}$. Under these restricted assumptions, $\boldsymbol{M}^{\text{pred}}$ is proportional to the *mean normal matrix* ($\boldsymbol{MNM}$) used by Batchelor[12] [22]. To compare our predictions with those of Batchelor, we will first consider experimental designs in which all these simplifying assumptions have been applied.

Batchelor proposed that an MR acquisition scheme in which diffusion gradients were oriented at vertices of an icosahedron possessed orientationally (i.e. rotationally) invariant statistical properties of the estimated diffusion tensor by showing that the $\boldsymbol{MNM}$ for this scheme is the same as the $\boldsymbol{MNM}$ obtained when using a gradient sampling scheme with an infinite number of gradient vectors that are uniformly distributed on the surface of a unit sphere [22].

Within the context of the tensor-variate distribution we can understand Batchelor's notion of rotational invariance: his $\boldsymbol{MNM} = \boldsymbol{M}^{\text{pred}}$ has the same form as $\boldsymbol{M}^{\text{iso}}$, the precision matrix associated with $\boldsymbol{A}^{\text{iso}}$ in (11) with $\lambda = \mu$. Choosing an isotropic fourth-order tensor $\boldsymbol{A}^{\text{iso}}$ with $\lambda = \mu = 1/15$, we are able to reproduce Batchelor's $\boldsymbol{MNM}$ exactly for an isocahedral gradient scheme, and for a gradient sampling scheme with an infinite number of gradient vectors that are uniformly distributed on a unit sphere[13] [22]. This is shown in Fig. 2.

In fact, many other gradient schemes can be constructed that satisfy this rotational invariance requirement. The simplest rotationally invariant gradient scheme uses only six gradient directions. It consists of gradient vectors whose coordinates are the noncollinear vertices of an icosahedron. This scheme is identical to one proposed in [23], and is given in Table I. Interestingly, one finds that gradient designs using the ten noncollinear vertices of a dodecahedron (the dual regular polyhedron of the icosahedron), an icosidodecahedron (polyhedron obtained by adding a tetrahedron on each of the faces of the dodecahedron), a Buckminster "Fullerene," as well as other patterns (i.e., those of Jones [16] and Muthupallai [23]) produce rotationally invariant experimental designs with the same values

[10]It is important to note that this assumption was not made in [20], in which a first-order correction was applied to account for the effect of the log-linear transformation on the variance of the measured signal.

[11]This assumes that there are no cross-terms arising from imaging gradients [21].

[12]$\boldsymbol{M}^{\text{pred}} = (1/\sigma^2)\mathbf{B}^{\text{T}}\mathbf{B} \sim N\,\boldsymbol{MNM}$ where N is the number of acquisitions.

[13]Note, the matrix "$\mathbf{A}$" used in Batchelor is a special case of the matrix "$\mathbf{B}$" used in [20] and should not be confused with our use of $\boldsymbol{A}$ as a fourth-order tensor.

$$\boldsymbol{M}^{\mathrm{pred}} \propto \begin{pmatrix} 3 & 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 \\ 1 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Fig. 2. Elements of the predicted precision matrix, $\boldsymbol{M}^{\mathrm{pred}}$, obtained for experimental designs whose gradient vectors lie on the vertices of an icosahedron, dodecahedron (the dual regular polyhedron of the icosahedron), an icosidodecahedron (polyhedron obtained by adding a tetrahedron on each of the faces of the dodecahedron), a Buckminster "Fullerene," and the polygons of Jones [16] and Muthupallai [23]. This $\boldsymbol{M}^{\mathrm{pred}}$ is also proportional to the $\boldsymbol{MNM}$ matrix obtained for the infinite uniform directional gradient sampling scheme described by Batchelor [29].

TABLE I
DIFFUSION GRADIENT VECTORS WRITTEN IN TERMS OF THEIR x, y, AND z COMPONENTS, $\{G_x, G_y, G_z\}$. THE SIX VECTORS ABOVE LIE ON AN ICOSAHEDRON. ALTHOUGH IT HAS TWELVE VERTICES, THE ICOSAHEDRON IS ANTIPODALLY SYMMETRIC, SO ONLY SIX DISTINCT ORIENTATIONS ARE INDEPENDENT. THE QUANTITY f_r IS FIBONACCI'S GOLDEN RATIO, $f_r = (\sqrt{5}-1)/2 \approx 0.618\,03$. THIS ACQUISITION SCHEME IS IDENTICAL TO WHAT WAS PROPOSED PREVIOUSLY IN [23]

G_x	G_y	G_z
1	f_r	0
1	$-f_r$	0
0	1	f_r
0	1	$-f_r$
f_r	0	1
$-f_r$	0	1

of λ and μ. $\boldsymbol{M}^{\mathrm{pred}}$for these designs is shown in Fig. 2. In Fig. 3, we consider the gradient scheme of Papadakis [17]. Interestingly, we find that it is approximately, but not strictly isotropic.

The advantage of using this new tensor-variate distribution framework to design DT-MRI experiments is that we can consider gradient schemes having different numbers of gradient acquisitions, gradient strengths and gradient magnitudes rather than those with uniform gradient strength. We can also use this framework to show that any combination of rotationally invariant experimental designs (with arbitrary rotations and scaling factors) will produce a rotationally invariant experimental design, so that these designs can be concatenated, producing a combined design that is also rotationally invariant. When combining these different designs, the constant of proportionality changes, but the precision matrix with $\lambda = \mu$ remains isotropic in form. Moreover, we are not limited to one particular choice of $\lambda = \mu$ to produce a rotationally invariant DT-MR experiment.

Additionally, our statistical framework also provides a natural way to assess the degree of rotational invariance of any experimental design, or rather, the degree to which an experimental

$$\boldsymbol{M}^{\mathrm{pred}} \propto \begin{pmatrix} 3.02 & 1.01 & 1 & 0 & 0 & 0 \\ 1.01 & 2.98 & 1 & 0 & 0 & 0 \\ 1 & 1 & 3.0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.04 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4.0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4.0 \end{pmatrix}$$

Fig. 3. $\boldsymbol{M}^{\mathrm{pred}}$ obtained for the Papadakis scheme. This scheme has 12 directions but they are not vertices of icosahedron, although it resembles it. The $\boldsymbol{M}^{\mathrm{pred}}$ is approximately isotropic, but not exactly isotropic. It is obtained by minimizing the variance in the elements of $\boldsymbol{D}$ (Papadakis, personal communication).

design deviates from statistical isotropy. One way is to measure the mean-squared deviation between $\boldsymbol{A}$ for a particular experimental design and an isotropic fourth-order precision tensor $\boldsymbol{A}^{\mathrm{iso}}$, but many other such measures can be contemplated.

H. Estimating $\boldsymbol{A}$ From Simulated DT-MRI Data

Note that $\boldsymbol{M}^{\mathrm{pred}}$ formalism assumes a linear relationship between the measured MR signal and the unknown diffusion tensor elements with additive Gaussian noise. In the MR experiment, however, this relationship is nonlinear, and, if the log-linear form as in (24) is used, the noise is not additive [24]. Thus, the actual precision matrix obtained by using the least-square solution will differ from $\boldsymbol{M}^{\mathrm{pred}}$. However, we defer these issues for another paper and here we just report MC simulations of DT-MRI experiments that yield an "isotropic" $\boldsymbol{A}$.

We performed MC simulations [13] to synthesize noisy replicates of diffusion tensors, $\boldsymbol{D}$, typical of those measured in isotropic gray matter regions of the human brain with DT-MRI using experimental parameters provided in [25]. From these MC data, we obtained sample estimates of $\boldsymbol{M}$ and $\boldsymbol{A}$ using formulae described in Appendix B. Estimated precision matrices using (B.2) for simple schemes like the one shown in Table I, do not produce isotropic precision matrices as predicted by (25) due to the log-linear transformation of the MR signal data. However, we found that when using a large number of directions, we can obtain approximately isotropic designs, but with λ not necessarily equal to μ. Fig. 4 shows $\boldsymbol{M}$ displayed as a 6×6 matrix with coefficients organized as in (7). Such a relationship holds rather well in the case of isotropic diffusion and a large number of gradient directions ($\cong$ 50 or more); for the data shown in Fig. 4 the number of directions was 60. The two matrices displayed show results for the cases when no nondiffusion-weighted (non-DW) images were used in the simulation, and where ten non-DW images were used. Although the actual values of μ and λ, and their ratio, depend on the number of non-DW images, the isotropic form appears to hold up to 20 non-DW images. However, in the case of anisotropic diffusion the log-linearization introduces a dependence of $\boldsymbol{M}$ on the mean values of $\tilde{\boldsymbol{D}}$. Investigating this problem and other problems of optimal design will be a subject of another paper.

a)

$$\overline{\boldsymbol{M}}(0)=\begin{pmatrix} 40 & -8 & -8 & -3 & 0 & 0 \\ -8 & 42 & -9 & 0 & 0 & 0 \\ -8 & -9 & 42 & 1 & 0 & 0 \\ -3 & 0 & 1 & 97 & -3 & -3 \\ 0 & 0 & 0 & -3 & 100 & -3 \\ 0 & 0 & 0 & -3 & -3 & 100 \end{pmatrix}$$

b)

$$\overline{\boldsymbol{M}}(10)=\begin{pmatrix} 56 & 8 & 8 & -4 & -1 & -1 \\ 8 & 59 & 8 & -1 & -1 & -1 \\ 8 & 8 & 59 & 0 & -1 & -1 \\ -4 & -1 & 0 & 97 & -3 & -3 \\ -1 & -1 & -1 & -3 & 100 & -3 \\ -1 & -1 & -1 & -3 & -3 & 100 \end{pmatrix}$$

Fig. 4. Elements of the estimated precision matrix $\overline{\boldsymbol{M}}$, organized as in (7), and obtained using (B.2) on MC replicates of a simple acquisition scheme consisting of three-fold repetitions of the gradients, which are the vertices of a dodecahedron. In (a) no nondiffusion-weighted images were used whereas in (b) ten nondiffusion-weighted images were used in the experimental design. The elements in the lower diagonal positions equal 4 μ. It is easy to see that $\mu \cong 25$ in both cases, and $\lambda(0) \cong -8$, $\lambda(10) \cong 8$; hence, $\lambda + 2\mu$ should assume values 42 and 58, respectively, which is close to the actual values of the three diagonal elements in the upper diagonal matrices.

Note that the formulae given in Appendix B can also be used to obtain estimates of $\boldsymbol{A}$ from empirically estimated $\boldsymbol{D}$ data on a voxel by voxel basis, using Bootstrap methods to resample the set of acquired diffusion weighted images [3].

III. Discussion

DT-MRI applications require a normal distribution for a tensor-valued random variable. While the multivariate normal distribution per se can describe the variability of individual elements of $\boldsymbol{D}$, it does not naturally yield other useful information [2]. Specifically, we would like to predict how the distribution of $\boldsymbol{D}$ would change if the laboratory coordinate system were rotated, or if a general affine transformation were applied to $\boldsymbol{D}$, for example, by applying shearing or dilatation operations required in image warping and registration applications [26]. It is also of interest to know how the first and higher moments of the apparent diffusion coefficient behave. This quantity is obtained by projecting the diffusion tensor along a particular direction. It is also of interest to know how the principal diffusivities (eigenvalues) and principal directions (eigenvectors) of $\boldsymbol{D}$ are distributed. Moreover, we would also like to know the distribution of scalar invariants of $\boldsymbol{D}$ (e.g., $\mathrm{Trace}(\boldsymbol{D})^2$ and $\mathrm{Trace}(\boldsymbol{D}^2)$) that characterize the type and degree of anisotropic diffusion.

On a more fundamental level, a tensor-variate distribution is needed because, at present no statistical model describes variability of second and higher-order tensors, which would be useful in prediction, estimation, filtering, and hypothesis testing applications of tensor data, and in improving the ability to design and interpret experiments involving tensor data.

A key attribute of using a fourth-order tensor, $\boldsymbol{A}$, to characterize the covariance structure of the tensor-variate distribution—rather than rewriting it as a vector—is that it preserves the form of the tensor random variable, $\boldsymbol{D}$, and our ability to perform admissible algebraic operations on it.[14] Unlike the multivariate distribution—whose only natural coordinate system is that of the covariance matrix, $\boldsymbol{\Sigma}$—the tensor-variate distribution refers the components of $\boldsymbol{D}$ and $\boldsymbol{A}$ explicitly to the reference or laboratory coordinate system.

However, because we have also shown how to convert between vector- and tensor-variate Gaussian distributions, we can employ all of the mathematical and statistical machinery developed for multivariate Gaussian distributions (e.g., see [11]) to analyze tensor data without having to rederive these findings and results.

The tensor formalism also allows us to view univariate and multivariate normal distributions as special cases of the more general tensor-variate distribution. The univariate distribution results from the contraction of two zeroth-order tensor random variables and a zeroth-order precision tensor; the multivariate distribution results from the contraction of two first-order random variables and a second-order precision tensor. In general, an nth-order tensor-variate normal distribution can be constructed by contracting two nth-order random tensors and a $2n$th-order precision tensor. In this way, we are able to generate distributions for random variables that are tensors of second and higher order.

There are a number of disciplines to which this new statistical methodology could be applied. In imaging sciences and signal processing, the most obvious application is to diffusion tensor MRI data [1], [20]. This new framework will help us estimate moments of the tensor-variate distribution, and perform numerous hypothesis tests for diffusion tensor-derived quantities in clinical, biological, and materials sciences applications. In the physical sciences, quantities such as the moment of inertia tensors, rotational or spin-diffusion tensors, and elastic coefficient tensors of elastic media, nematics, and crystals [27] are routinely measured. In some cases, they may conform to a normal tensor-variate distribution, especially if they are measured using regression methods (e.g., as in [20]). In the physics of continuous media, and in materials engineering, tensor quantities arise in constitutive equations that are used to describe charge, mass, momentum, and energy transport. These include the translational diffusion tensor, the particle dispersion tensor, the fabric tensor, the electrical conductivity tensor, the thermal conductivity tensor, and the hydraulic permeability tensor. These quantities are measured using a variety of

[14]Algebraists say that two vector spaces of the same dimension are "isomorphic," but that the isomorphism is not "canonic." in the sense that the isomorphism is not uniquely prescribed. Such is the case with the vector and tensor-variate distributions.

methods, and in some cases, their individual components may also conform to a normal distribution that could be described using the formalism above. Finally, many input/output matrix models used in engineering and in the social sciences may have coefficients that are also described by this new distribution.

IV. Concluding Remarks

The idea of using the tensor contraction operation (in this case, applied to fourth- and second-order tensors) in the exponent of a normal distribution appears to be novel to the theory of statistical distributions, and significantly extends the scope and applicability of the normal distribution to accommodate many types of high dimensional data.

In the near term, this new tensor-variate distribution should improve our ability to estimate $\boldsymbol{D}$ and quantities derived from it in DT-MRI studies. It should also lead to the development of hypothesis tests with which to analyze *in vivo* DT-MRI data. Finally, it should lead to improvements in the experimental design of DT-MRI studies, providing a unifying framework for understanding the effect of changing independent experimental parameters.

Appendix A
$\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}$ for an Isotropic $\boldsymbol{A}$ Tensor

When $\boldsymbol{A}$ is an isotropic fourth-order tensor, the contraction $\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}$ can be written as

$$D_{ik}A_{ikpm}D_{pm} = \lambda D_{ik}\delta_{ik}\delta_{mp}D_{pm} + \mu(D_{ik}\delta_{im}\delta_{kp}D_{pm} + D_{ik}\delta_{ip}\delta_{km}D_{pm}). \quad \text{(A.1)}$$

With a little algebra, this simplifies to[15]

$$D_{ik}A_{ikpm}D_{pm} = \lambda D_{ii}D_{pp} + \mu(D_{km}D_{km} + D_{pk}D_{pk}). \quad \text{(A.2)}$$

Now, using the definitions (see [28])

$$Trace(\boldsymbol{D}) = D_{ii} \quad \text{and} \quad Trace(\boldsymbol{D}^2) = D_{ij}D_{jk}\delta_{ik} = D_{ij}D_{ji} = D_{ij}D_{ij} \quad \text{(A.3)}$$

we can rewrite the tensor contraction in (A.1) as

$$\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D} = \lambda\,(Trace(\boldsymbol{D}))^2 + 2\mu\,Trace(\boldsymbol{D}^2) \quad \text{(A.4)}$$

which is the result we set out to show.

Appendix B
Formulae for Obtaining Sample Estimates of $\overline{\boldsymbol{D}}$ and $\overline{\boldsymbol{A}}$

Sample estimates of the mean and precision tensors for the tensor-variate distribution are readily obtained from simulated DT-MR data.

The sample mean tensor for a sample of size N is

$$\overline{\boldsymbol{D}} = \frac{1}{N}\sum_{m=1}^{N}\boldsymbol{D}^m. \quad \text{(B.1)}$$

[15]If $\boldsymbol{D}$ is interpreted as an infinitesimal strain tensor, and its parameters are elastic constants, then this expression is exactly of the form of the stress of an isotropic solid (e.g. see [8] and [28]). Moreover, the term $\boldsymbol{D}:\boldsymbol{A}:\boldsymbol{D}$ can be interpreted as a strain energy function for such a material.

The estimate of the fourth-order precision tensor can be obtained through its relationship to the precision matrix, $\boldsymbol{M}$, [see (7)]. The estimate of the precision matrix is the inverse of the unbiased estimate of the covariance matrix of the diffusion (column) vector, $\tilde{D}$

$$\overline{\boldsymbol{M}} = \left(\frac{1}{N-1}\sum_{m=1}^{N}\left(\tilde{D}_m - \overline{\tilde{D}}\right)\left(\tilde{D}_m - \overline{\tilde{D}}\right)^T\right)^{-1}. \quad \text{(B.2)}$$

The resulting $\overline{\boldsymbol{M}}$ can then be used to obtain the elements of $\overline{\boldsymbol{A}}$ using (7).

Acknowledgment

The authors would like to acknowledge the late E. Turan Onat of Yale University who counseled them about the properties of fourth-order tensors and the various ways in which they can be represented. They would like to thank P. Munson, W. Jarisch, and A. Aldroubi who made thoughtful comments and substantive editorial suggestions. They also acknowledge D. Jones who made many substantive suggestions that improved the clarity, readability, and correctness of this work. Finally, they would like to thank L. Salak who edited this manuscript.

References

[1] P. J. Basser, J. Mattiello, and D. Le Bihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, pp. 259–267, 1994.

[2] S. Pajevic and P. J. Basser, "Parametric description of noise in diffusion tensor MRI," presented at the 8th Annu. Meeting ISMRM, Philadelphia, PA, 1999.

[3] ——, "Parametric and non-parametric statistical analysis of DT-MRI data," *J. Magn. Reson.*, vol. 161, no. 1, pp. 1–14, 2003.

[4] K. Fukunaga, "Introduction to statistical pattern recognition," in *Electrical Sciences*, N. DeClaris, Ed. New York: Academic, 1972.

[5] A. E. Green and W. Zerna, *Theoretical Elasticity*, 2nd ed. London, U.K.: Oxford Univ. Press, 1954.

[6] L. E. Malvern, *Introduction to the Mechanics of a Continuous Medium*. Englewood Cliffs, NJ: Prentice-Hall, 1969.

[7] E. T. Onat, J. P. Boehler, and J. A. A. Kirillov, "On the polynomial invariants of the elasticity tensor," *J. Elasticity*, pp. 97–110, 1994.

[8] H. Jeffreys, *Cartesian Tensors*. Cambridge, U.K.: Cambridge Univ. Press, 1931.

[9] G. B. Arfken and H.-J. Weber, *Mathematical Methods for Physicists*, 5th ed. New York: Harcourt/Academic, 2000.

[10] Q. S. Zheng, "A note on representation for isotropic functions of 4th-order tensors in 2-dimensional space," *Z. angew. Math. Mech.*, vol. 74, pp. 357–359, 1994.

[11] T. W. Anderson, *An Introduction to Multivariate Statistics*, 2nd ed. New York: Wiley, 1984.

[12] G. Strang, *Introduction to Applied Mathematics*. Wellesley, MA: Wellesley-Cambridge Press, 1986.

[13] C. Pierpaoli and P. J. Basser, "Toward a quantitative assessment of diffusion anisotropy," *Magn. Reson. Med.*, vol. 36, pp. 893–906, 1996.

[14] ——, "Erratum to 'Toward a quantitative assessment of diffusion anisotropy'," *Magn. Reson. Med.*, vol. 37, no. 6, p. 972, 1997.

[15] S. Skare, M. Hedehus, M. E. Moseley, and T. Q. Li, "Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI," *J. Magn. Reson.*, vol. 147, pp. 340–352, 2000.

[16] D. K. Jones, M. A. Horsfield, and A. Simmons, "Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging," *Magn. Reson. Med.*, vol. 42, pp. 515–525, 1999.

[17] N. G. Papadakis, D. Xing, C. L. Huang, L. D. Hall, and T. A. Carpenter, "A comparative study of acquisition schemes for diffusion tensor imaging using MRI," *J. Magn. Reson.*, vol. 137, pp. 67–82, 1999.

[18] N. G. Papadakis, D. Xing, G. C. Houston, J. M. Smith, M. I. Smith, M. F. James, A. A. Parsons, C. L. Huang, L. D. Hall, and T. A. Carpenter, "A study of rotationally invariant and symmetric indices of diffusion anisotropy," *Magn. Reson. Imag.*, vol. 17, pp. 881–892, 1999.

[19] N. G. Papadakis, C. D. Murrills, L. D. Hall, C. L. Huang, and T. A.T. Adrian Carpenter, "Minimal gradient encoding for robust estimation of diffusion anisotropy," *Magn. Reson. Imag.*, vol. 18, pp. 671–679, 2000.

[20] P. J. Basser, J. Mattiello, and D. Le Bihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *J. Magn. Reson.*, vol. B 103, pp. 247–254, 1994.

[21] J. Mattiello, P. J. Basser, and D. Le Bihan, "Analytical expression for the b matrix in NMR diffusion imaging and spectroscopy," *J. Magn. Reson.*, vol. A 108, pp. 131–141, 1994.

[22] P. Batchelor, "Optimization of direction schemes for tensor imaging," presented at the Workshop on Diffusion MRI: Biophysical Issues (What Can We Measure?), St. Malo, France, 2002.

[23] R. Muthupallai, C. A. Holder, A. W. Song, and W. T. Dixon, "Navigator aided, multishot EPI diffusion images of brain with complete orientation and anisotropy information," presented at the 8th Annu. Meeting ISMRM, Philadelphia, PA, 1999.

[24] R. M. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Med. Phys.*, vol. 12, pp. 232–233, 1985.

[25] A. Virta, A. Barnett, and C. Pierpaoli, "Visualizing and characterizing white matter fiber structure and architecture in the human pyramidal tract using diffusion tensor MRI," *Magn. Reson. Imag.*, vol. 17, pp. 1121–1133, 1999.

[26] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee, "Spatial transformations of diffusion tensor magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1131–1139, Nov. 2001.

[27] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 3rd ed. Oxford, U.K.: Pergamon, 1986, vol. 7.

[28] P. M. Morse and H. Feschbach, *Methods of Theoretical Physics*. New York: McGraw-Hill, 1953, vol. 1.

[29] P. G. Batchelor, D. Atkinson, D. L. G. Hill, F. Calamante, and A. Connelly, "Optimal choice of directions for diffusion tensor imaging," presented at the *10th ISMRM Conf.*, Honolulu, HI, 2002.