Detection and Modeling of Non-Gaussian Apparent Diffusion Coefficient Profiles in Human Brain Data

D.C. Alexander,[1]* G.J. Barker,[2] and S.R. Arridge[1]

This work details the observation of non-Gaussian apparent diffusion coefficient (ADC) profiles in multi-direction, diffusion-weighted MR data acquired with easily achievable imaging parameters ($b \approx 1000$ s/mm^2). A technique is described for modeling the profile of the ADC over the sphere, which can capture non-Gaussian effects that can occur at, for example, intersections of different tissue types or white matter fiber tracts. When these effects are significant, the common diffusion tensor model is inappropriate, since it is based on the assumption of a simple underlying diffusion process, which can be described by a Gaussian probability density function. A sequence of models of increasing complexity is obtained by truncating the spherical harmonic (SH) expansion of the ADC measurements at several orders. Further, a method is described for selection of the most appropriate of these models, in order to describe the data adequately but without overfitting. The combined procedure is used to classify the profile at each voxel as isotropic, anisotropic Gaussian, or non-Gaussian, each with reference to the underlying probability density function of displacement of water molecules. We use it to show that non-Gaussian profiles arise consistently in various regions of the human brain where complex tissue structure is known to exist, and can be observed in data typical of clinical scanners. The performance of the procedure developed is characterized using synthetic data in order to demonstrate that the observed effects are genuine. This characterization validates the use of our method as an indicator of pathology that affects tissue structure, which will tend to reduce the complexity of the selected model. **Magn Reson Med 48:331–340, 2002. © 2002 Wiley-Liss, Inc.**

Key words: diffusion tensor magnetic resonance imaging; human brain; non-Gaussian; spherical harmonic; model selection

Diffusion imaging, particularly diffusion tensor magnetic resonance imaging (DT-MRI) (1) has become popular because of the insight it provides into the structural connectivity of tissue (2,3). Water is a large constituent of biological tissue, and water molecules in tissue constantly undergo random, Brownian motion. Tissue also contains rigid structures, such as the walls of cells, that form barriers to diffusion, and it is this hindrance to diffusion that allows tissue structure to be probed through measurements of water mobility due to diffusion processes. In some types of tissue, such as most gray matter in the brain, the structure has no preferred orientation and so causes approximately the same amount of hindrance to diffusion in all directions. The amount of diffusion or water mobility is thus approximately equal in all directions, i.e., isotropic. Other types of tissue, however, have ordered structure that hinders diffusion to different degrees in different directions, causing diffusion anisotropy. White matter in the brain, for example, consists of bundles of axon fibers, and water is free to diffuse along the axis of the fibers but is hindered in the perpendicular directions.

The diffusion of water molecules in tissue over some time interval, t, can be described by a probability density function, p_t, on the displacement, $\underline{x}$, of water molecules after time t. p_t reflects the underlying tissue microstructure, because it is largest in the directions of least hindrance to diffusion and smaller in other directions. In white matter, for example, p_t is largest in directions parallel to fibers, but is small in perpendicular directions and thus reveals fiber orientations. The goal of diffusion imaging is to obtain information about p_t that leads to meaningful inferences about the microstructure of the material being imaged.

p_t can be shown to relate to the NMR signal attenuation, $S(q)$, measured through a pulsed gradient spin-echo experiment, via a Fourier transform (FT) with respect to $q = \gamma \delta G \underline{\hat{k}}$ (4,5):

$$S(q) = \int p_t(\underline{x}) \exp(-i\underline{q} \cdot \underline{x}) d\underline{x}. \qquad [1]$$

The spin-echo attenuation, $S(q)$, is defined as the normalized diffusion-weighted signal, $s(\underline{q})/s(0)$, where $s(\underline{q})$ is the NMR signal in the presence of a diffusion-weighting gradient of magnitude G and direction $\hat{k}$, and $s(0)$ is the signal in the absence of any such gradient. δ is the length of the gradient pulse, and γ is the magneto-gyric ratio of protons in water. We note that Eq. [1] relies on the fact that δ is negligibly small compared to t. This assumption is rarely justified in practice, but the effect of non-negligible δ is merely to introduce a convolution over a range of diffusion times (6) into the measurements, which generally preserves the large-scale structure and orientation of the inferred p_t.

Given enough measurements of $S(q)$ spread over a suitable range of q, the FT can be inverted to obtain an estimate of p_t. A more common approach, however, is to assume a simple model for p_t the FT of which can be related directly to the spin-echo attenuation, which allows p_t to be inferred from a much sparser set of measurements. The simplest model for p_t that incorporates second-order statistics (anisotropy) is a multivariate, zero-mean Gaussian, which has covariance $2Dt$ at time t:

$$p_t(\underline{x}) = \frac{1}{\sqrt{(4\pi t)^3 |D|}} \exp\left[\frac{-\underline{x}^T D^{-1} \underline{x}}{4t}\right]. \qquad [2]$$

This is the distribution of an initial point concentration at $x = 0$, $t = 0$, diffusing according to the simple anisotropic diffusion equation (7):

$$\frac{\partial p_t}{\partial t} = \nabla(D\nabla p_t).$$ [3]

The FT of p_t is then also Gaussian, which gives rise to a simple relationship between the parameters of p_t (the elements of D) and the spin-echo attenuation:

$$S(q) = \exp[-tq^T Dq] = \exp[-b\underline{\hat{k}}^T D\underline{\hat{k}}].$$ [4]

In Eq. [4], b is the diffusion-weighting factor given by $b = t|q|^2$.

The simplest form of diffusion-weighted (DW) MRI (8) models p_t with a Gaussian in one dimension. A single spin-echo attenuation measurement allows the single parameter of p_t—its scalar variance—to be estimated from Eq. [4]:

$$d(\underline{\hat{k}}) = -\frac{1}{b}\log S(q).$$ [5]

The diffusion coefficient, $d(\underline{\hat{k}})$, is proportional to the variance of the Gaussian model, which describes the root mean squared displacement of water molecules in the direction of the applied DW gradient $\underline{\hat{k}}$. In DW-MRI, the measure of $d(\underline{\hat{k}})$ obtained from Eq. [5] is often called the *apparent* diffusion coefficient (ADC) (1), both because it represents a spatial average of the diffusion coefficient over an image voxel and because it is based on this Gaussian assumption, which may be unjustified.

DT-MRI extends this basic idea to 3D, where the diffusion coefficient is described by a diffusion tensor (DT), D, which is proportional to the covariance matrix of the trivariate Gaussian, as in Eq. [2]. D relates to the 1D diffusion coefficient $d(\underline{\hat{k}})$ in any chosen direction $\underline{\hat{k}}$ as follows:

$$d(\underline{\hat{k}}) = \underline{\hat{k}}^T D\underline{\hat{k}}.$$ [6]

In 3D, D is a symmetric 3×3 matrix and thus has six free parameters. A minimum of six estimates of $d(\underline{\hat{k}})$ in independent directions is thus required to estimate D, which requires six measurements of the spin-echo attenuation (seven MR measurements in total) to be made with the DW gradient applied in these independent directions. The estimate of D obtained from such a set of measurements is referred to as the *apparent* diffusion tensor (ADT) (1), as in the case of the ADC.

We define the "ADC profile" to be the estimate of $d(\underline{\hat{k}})$ over the range of $\underline{\hat{k}}$, which is the unit sphere. When p_t is Gaussian, the ADC profile is described by Eq. [6], and a standard approach (9) to the estimation of D is to acquire DW images in a large number of directions (many more than six) spread evenly over the unit hemisphere. This oversampling of the ADC profile provides a more robust estimate of D. Note that the antipodal symmetry of p_t and hence $d(\underline{\hat{k}})$ is assumed, so that $d(\underline{\hat{k}}) = d(-\underline{\hat{k}})$ and only half of the sphere needs to be sampled. When p_t is not Gaussian

the ADC profile deviates from that described by Eq. [6], and this kind of multiple gradient direction scheme affords the possibility of observing significant deviations should they arise.

It has long been recognized (1,10–14) that the Gaussian, DT model is inappropriate when complex tissue structure is found within a single image voxel. There are alternative models for p_t that can capture certain non-Gaussian effects that occur in these circumstances. A simple alternative is the multi-Gaussian model (10,11). This model is based on the assumption that a voxel contains n separate compartments, each containing a different tissue type in proportion a_i ($\Sigma_i a_i = 1$, $i = 1, \ldots n$) and that the diffusion within each compartment can be described by a Gaussian p_t with DT, D_i. The model assumes further that there is no exchange of molecules between these separate compartments. p_t then becomes a weighted sum of Gaussians and Eq. [4] becomes:

$$S(q) = \sum_{i=1}^{n} a_i \exp(-b\underline{\hat{k}}^T D_i\underline{\hat{k}}).$$ [7]

With this model for p_t, the ADC profile can have shape very different from that described by Eq. [6], which is often modeled poorly by a single DT (10). Figure 1 shows ADC profiles simulated from prolate, oblate, and isotropic Gaussian p_t's, together with profiles obtained from bi-Gaussian p_t's that combine them. Note the characteristic peanut shape of the prolate Gaussian ADC profiles and the "filled bagel" or red blood cell shape of the oblate Gaussian profile, which are typical of Gaussian functions plotted over a sphere. The contours of the corresponding Gaussian functions in 3D have the more familiar ellipsoidal contours: cigar-shaped in the prolate case, and Frisbee-shaped in the oblate case.

Alexander et al. (10) analyzed the behavior of the ADC profile within voxels containing multiple tissue compartments. They showed how the observability of higher-order profiles increases (their non-Gaussian shape becomes more pronounced) with the size of the diffusion-weighting factor, b. They used the instability of the DT and its derived scalar measures, such as the mean diffusivity and fractional anisotropy, to locate regions in the brain where the DT description of the ADC profile is poor. Several regions in data acquired with $b = 3000$ s/mm^2 were highlighted by this approach, including the pons, corpus callosum, cingulum, internal capsule, and arcuate fasciculus.

Frank (11) showed that a 4th-order SH series provides a first approximation to the ADC profile obtained from a multi-Gaussian p_t. Frank fitted a 4th-order SH series to ADC measurements acquired with $b = 3000$ s/mm^2 and showed that significant 4th-order (i.e, non-Gaussian) components arise in locations of the human brain similar to those highlighted in Ref. 10 (see above).

In other related work, Wedeen and Tuch et al. (12,13) used q-space techniques (4) to measure p_t. This technique exploits the Fourier relationship between p_t and the spin-echo attenuation directly by acquiring a large number of measurements of $S(q)$ over a wide range of q in order to obtain enough samples of the FT of p_t to perform a stable
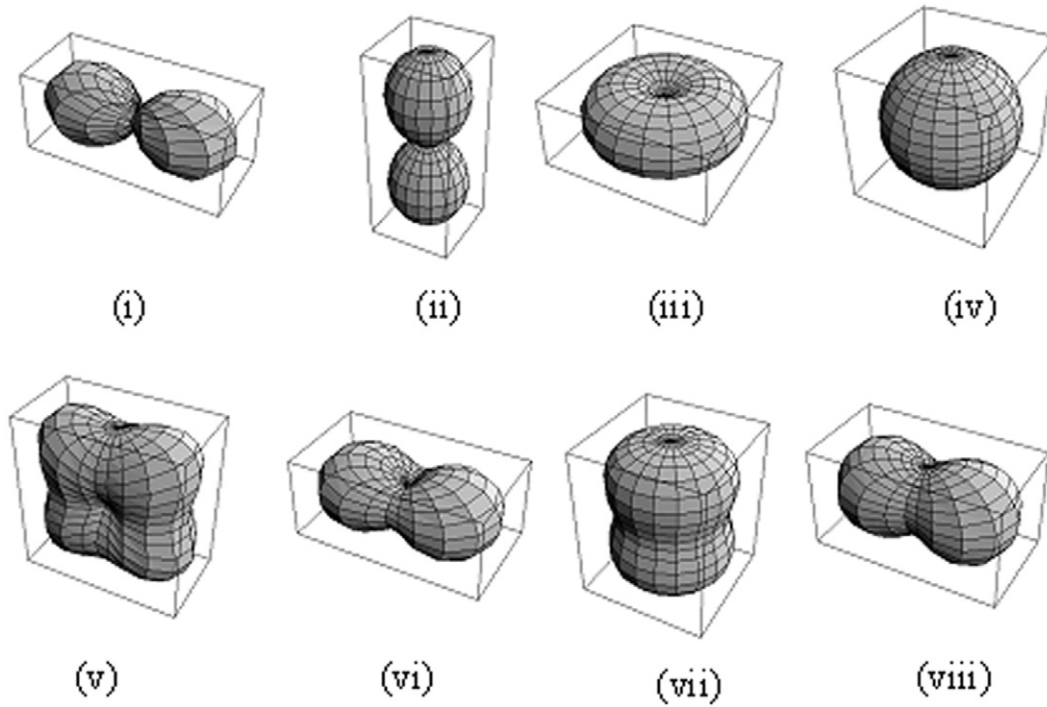
FIG. 1. Simulated examples of ADC profiles. Top row: profiles corresponding to Gaussian diffusion processes: (i) prolate DT oriented along the x-axis (eigenvalues [1700, 200, 200] $\times10^{-6}$ mm^2/s), (ii) prolate DT oriented along the z-axis (eigenvalues [200, 200, 1700] $\times10^{-6}$ mm^2/s), (iii) oblate DT (eigenvalues [950, 950, 200] $\times10^{-6}$ mm^2/s), and (iv) isotropic DT (eigenvalues [700, 700, 700] $\times10^{-6}$ mm^2/s). Bottom row: ADC profiles corresponding to the bi-Gaussian model combining pairs of DTs from the top row in equal proportion with b set to 1000 s/mm^2; (v) combines (i) and (ii), (vi) combines (i) and (iii), (vii) combines (ii) and (iii), and (viii) combines (i) and (iv).

inversion. Distinctly non-Gaussian p_t's have been observed in both the human brain and heart, particularly at locations where anisotropic fibers cross within a single voxel.

In this work, we investigate the observability of non-Gaussian ADC profiles in DW data acquired using acquisition parameters more typical of those used clinically. Our sequence consists of a multiple gradient direction scheme based on the work of Jones et al. (9) using a 1.5T scanner, with a maximum b of approximately 1000 s/mm^2. We use the spherical harmonic (SH) series to provide a hierarchy of models for the ADC profile. In the Methods section, an efficient and robust method for fitting the SH series to DW-MRI data is described together with a method, based on the analysis of variance (ANOVA) test for addition/deletion of variables, for selecting the most appropriate level of truncation of the series. The combined fitting and model selection procedures developed are used to classify the profile in each voxel as arising from isotropic, anisotropic Gaussian, or non-Gaussian p_t, and thus to produce maps of where these different types of behavior occur. Non-Gaussian behavior is most likely to arise and be observed in regions of high tissue complexity containing distributions of fiber orientations with multiple peaks, such as fiber intersections. The maps generated from the model selection procedure provide extra diagnostic information in pathologies involving neuronal loss, degeneration, or demyelination, since non-Gaussian behavior will tend to disappear in the affected areas because the complexity of the tissue structure is reduced. These maps can

also be used to highlight regions in which the ADT and its derived indices are unreliable. We apply the method to both in vivo human brain data and to synthetic data for performance evaluation and validation.

METHODS

Models for the ADC Profile

In this section, we describe how the SH series can be used to model the ADC profile. We define the SH series, describe how its coefficients are computed for a given function or set of sampled data, and discuss the relationship to more familiar models of the ADC.

SH Series

The SH series (15) is analogous to the rectilinear Fourier series, and provides an orthonormal basis of functions on the sphere that can be combined linearly to represent any complex valued spherical function. It consists of a set of functions $Y_{l,\,m}$: S$_2 \rightarrow$ **C**, where S$_2$ is the unit sphere in 3D, which we parametrize by $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$, the angles of colatitude and longitude, respectively; **C** is the set of complex numbers. $l = 0, 1, 2, \ldots$ defines the *order* of the SH and $m \in \{-l, \ldots, 0, \ldots l\}$ indexes the $2l+1$ SH functions of order l.

Any spherical function, f: S$_2 \rightarrow$ **C**, can be written as a linear combination of the SHs:

$$f(\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{l,m} \, Y_{l,m} \, (\theta,\varphi). \qquad [8]$$

The complex coefficients, $c_{l,\,m}$, of the SHs in Eq. [8] are given by:

$$c_{l,m} = \int_{0}^{2\pi} \int_{0}^{\pi} f \, (\theta,\varphi) Y_{l,m}^{*} \, (\theta,\varphi) \sin \theta d\theta d\varphi. \qquad [9]$$

In Eq. [9] and henceforward, * denotes the complex conjugate.

Fitting the Series to Sampled Data

In the discrete case, where we have sampled data, $\underline{F} = \{f(\theta_i, \varphi_i), i = 1, \ldots, n\}$, one way to estimate the $c_{l,\,m}$ is to replace the integral in Eq. [9] by a summation. However, as noted by Brechbuhler et al. (16), this approach often provides poor estimates, and a more robust approach is to compute the least-squares solution. We adopt a similar approach here. First, we define an enumeration of the SH series, so that each SH is indexed by a single, unique integer, $j(l, m) = l^2 + l + m$. We select a maximum order for the series, $l_{\max}$, and then define an $n \times j(l_{\max}, l_{\max})$ complex matrix, X, to be the matrix with elements $X_{i,\,j(l,\,m)} = Y_{l,\,m}(\theta_i, \varphi_i)$. If we define $\underline{C}$ to be the $j(l_{\max}, l_{\max})$ component vector of SH coefficients, then (16):

$$\underline{C} = M \, \underline{F}, \text{ where } M = (X^{*\,T} \, X)^{-1} \, X^{*\,T}. \qquad [10]$$

$l_{\max}$ can be chosen by consideration of the number of free parameters defining the series up to each order, which should be less than or equal to the number of sampled points.

Modeling DW-MR Data

In its most general form, the SH series can represent any complex-valued function of the sphere in 3D. The set of functions required to model ADC profiles, however, is more constrained. In particular, the ADC is real-valued and exhibits antipodal symmetry ($d(\hat{\underline{k}}) = d(-\hat{\underline{k}})$).

The real-valued constraint ensures both that the imaginary part of $c_{l,\,o}$ is zero for all l, and that $c_{l,m} = (-1)^m \, c_{l,-m}^{*}$. The SHs of odd order all define asymmetric functions on the sphere, whereas the even orders are all symmetric. The antipodal symmetry of the ADC profile thus ensures that it can be represented by a series consisting only of even-order SHs.

These constraints dramatically reduce the number of parameters required to describe SH models truncated at each order. The numbers of free parameters required for a SH series including terms up to order 2N for general spherical functions is $2(2N + 1)^2$, which is reduced to $(2N + 1)(N + 1)$ for our real, symmetric functions.

Equation [10] shows that the calculation of the SH coefficients from the set of sampled points on the ADC profile is a linear transformation and so can be computed very

efficiently. Moreover, the set of sampled points at each voxel in the image correspond to the same set of directions, so that $\{(\theta_i, \varphi_i), i = 1, \ldots n\}$ is fixed over the image. The matrix M, of Eq. [10], therefore need only be calculated once in order to compute $\underline{C}$ at each voxel.

Models at Different Orders

If the SH series is truncated at order 0, the series provides an isotropic model, since the single 0th-order SH is constant over the sphere. If the series is truncated at order 2, it provides a model that is completely equivalent to the familiar DT model. An expression for the DT in terms of the 0th- and 2nd-order SH coefficients can be obtained if we express $\hat{\underline{k}}$ in Eq. [6] in terms of θ and φ, as in Eq. [8], and equate the right sides of these two equations. When we include higher-order members of the series, a range of more complex shapes becomes available. In particular, at order 4 we can obtain models with two pairs of peaks that have similar shape to the profiles obtained from the bi-Gaussian model, shown in Fig. 1.

Model Selection

Once we have computed the coefficients of the SH series for a set of measurements, a hierarchy of increasingly complex models is obtained by truncating the series at each order, $l = 0, 2, 4, , l_{\max}$ (the coefficients of the SHs above order l are set to zero). Although higher-order models are more descriptive, in many cases they are not necessary to describe the underlying function from which the measurements were taken. For example, if the underlying function is isotropic, we only need a series up to order 0 to describe it, and higher-order terms will only represent noise added to the data during the imaging process.

We use ANOVA to determine whether the addition of more parameters to the model, i.e., truncating the series at a higher order as opposed to a lower order, significantly changes the fit of the model to the data (17). Given a set of N sampled points, together with a lower order model M_1 with p_1 free parameters and a higher order model M_2 with p_2 free parameters, the appropriate statistic for the *F*-test of the hypothesis that the two models are equivalent, i.e., the lower order model is sufficient, is

$$F(M_1, M_2) = \frac{(N - p_2 - 1)(\text{Var}(M_2) - \text{Var}(M_1))}{(p_2 - p_1)E(M_2)}. \qquad [11]$$

The degrees of freedom are $N - p_2 - 1$ and $p_2 - p_1$ (17). In Eq. [11], Var(M) denotes the variance of model M about its mean value, and E(M) denotes the mean squared error between model M and the N sampled points.

The full algorithm for modeling the ADC profile in one voxel is outlined below:

- Compute the coefficients of the even SH series up to order $l_{\max}$ using Eq. [10].
- Truncate the series at order 0 to obtain model M_ω.
- Set $i = 2$ and $\omega = 0$.
- While ($i \leq l_{\max}$),

  ○ Truncate the series at order i to obtain model M_i.
  ○ Compute $F(M_\omega, M_i)$ using Eq. [11].

- ○ Define the null hypothesis to be that models M_i and M_ω are equivalent.
- ○ Compute the critical value, T, for $F(M_\omega, M_i)$ such that if $F(M_\omega, M_i) <$ T then the probability of the null hypothesis is less than a decision threshold, α_ω.
- ○ If $F(M_\omega, M_i) <$ T (null hypothesis is rejected)
  - Set $\omega = i$.
- ○ Set $i = i + 2$.
- • Select model M_ω.

EXPERIMENTS

The central hypothesis to be tested is that ADC profiles corresponding to non-Gaussian diffusion processes occur in the human brain, and that their effects can be observed in data collected with parameters typical of standard clinical MR scanners, using the methods described in the previous section. In order to verify this hypothesis we apply our methods to a variety of synthetic data and to human brain DW-MR data.

The synthetic data is created by a Monte Carlo simulation of the imaging process. This data is first used to set the parameters of the model selection procedure, and subsequently to characterize its performance in terms of the number and type of misclassifications that occur for noisy profiles of known order. We go on to show the results of applying our procedure to human brain data, which exhibit clusters of non-Gaussian profiles in various well-defined regions. In order to confirm that the higher-order regions observed in the human brain data genuinely correspond to non-Gaussian behavior, we perform a final simulation in which parameters derived from fitting Gaussian (order 2 SH) models in these regions are used to generate synthetic data, which is then reprocessed to show that if the data were truly Gaussian this would have been detected reliably. For brevity, only the most significant results are included, but more comprehensive results can be found in Ref. 23.

Simulation Experiments

We simulate the imaging sequence used to acquire the human brain data described in the next section in order to generate noisy measurements of various ideal ADC profiles. A range of DTs is used to provide models for profiles corresponding to isotropic and anisotropic Gaussian diffusion, and we use a bi-Gaussian model, c.f. Eq. [7], to obtain non-Gaussian ADC profiles corresponding to mixed tissue and fiber crossings.

For a given profile, we simulate a 128×128 2D array of noisy measurements of the profile. Isotropic complex Gaussian noise is added to the simulated data in the time domain, at a level corresponding to that observed in the scanner in the absence of any signal. Noisy magnitude images are reconstructed from this data, which correspond to each of the 60 DW images acquired in our sequence (see next section) together with three unweighted images. Spin-echo attenuation and thence ADC measurements in each of the 60 sampled directions are then calculated, and our procedure is applied.

The first set of experiments performed are designed to provide appropriate settings for the decision thresholds,

α_ω, used in the *F*-test in the algorithm described in the Model Selection section. Lower settings of the α_ω's cause the null hypotheses to be more difficult to reject, which will generally result in an increase in the proportions of lower-order models. Higher values of the α_ω's result in a greater proportion of higher-order models, and in general there is a trade-off between the number of overfitted and underfitted voxels, the occurrence of both of which we would like to minimize. We simulate the following Gaussian ADC profiles for testing using typical values as given in Ref. 3:

1. Gray matter (GM). Approximately isotropic diffusion with eigenvalues [700, 700, 700] ($\times 10^{-6}$ mm^2/s) and the unweighted signal is chosen so that its signal-to-noise ratio (SNR), $\sigma_0 = 35$, which is typical in our scanner data.
2. Prolate white matter (WM). Eigenvalues [1700, 200, 200] ($\times 10^{-6}$ mm^2/s), $\sigma_0 = 35$.
3. Oblate WM. Eigenvalues [950, 950, 200] ($\times 10^{-6}$ mm^2/s), $\sigma_0 = 35$.

We also simulate profiles from orthogonal crossing fibers in equal proportions, where each fiber is represented by a prolate DT with eigenvalues [1700, 200, 200] ($\times 10^{-6}$ mm^2/s) and $\sigma_0 = 35$.

For each profile, α is varied and the number of voxels at which each model order is selected is recorded. For each model order ω, we choose α_ω to be the value for which the sum of the rate of overfitting of profiles of known order ω and the rate of underfitting of profiles of known order $\omega+2$ is minimized.

Once appropriate decision thresholds for our *F*-tests have been chosen in this way, we can characterize the performance of our procedure on a wider range of input data, which enables us to interpret results obtained from scanner data reliably. A series of oblate and prolate Gaussian ADC profiles are tested in which the anisotropy is increased from zero to approximately the highest levels observed in our brain data. The classification rates of our procedure applied to the prolate data are plotted in Fig. 2, and similar results were obtained from the oblate profiles. These results demonstrate that with the selected α_ω's our procedure will identify anisotropic, Gaussian profiles correctly at order 2, with a classification rate over 92%, when the difference ratio of the largest and smallest eigenvalues of the DT is approximately 1.5 or greater. At this lowest level of detected anisotropy, the rate of overfitting at order 4 is around 2%, but the rate increases steadily with increasing anisotropy to about 8% for the most anisotropic DT tested. This is reasonable to expect, because highly anisotropic DTs contain very low ADC measurements in which the noise component is high (19) and hence more likely to cause significant deviation from the Gaussian profile. Misclassification at orders above 4 is negligible. Figure 2 also demonstrates the classification of isotropic GM profiles, which is reliably order 0—the rate of misclassification at order 2 is less than 0.1%. These rates are consistent for other isotropic DTs with larger eigenvalues.

We also simulate bi-Gaussian profiles to emulate the behavior at crossing fibers intersecting at angles of 90°, 67.5°, 45°, and 22.5°, and in proportions 1:1 and 3:1. Each fiber is represented by a prolate DT with eigenvalues
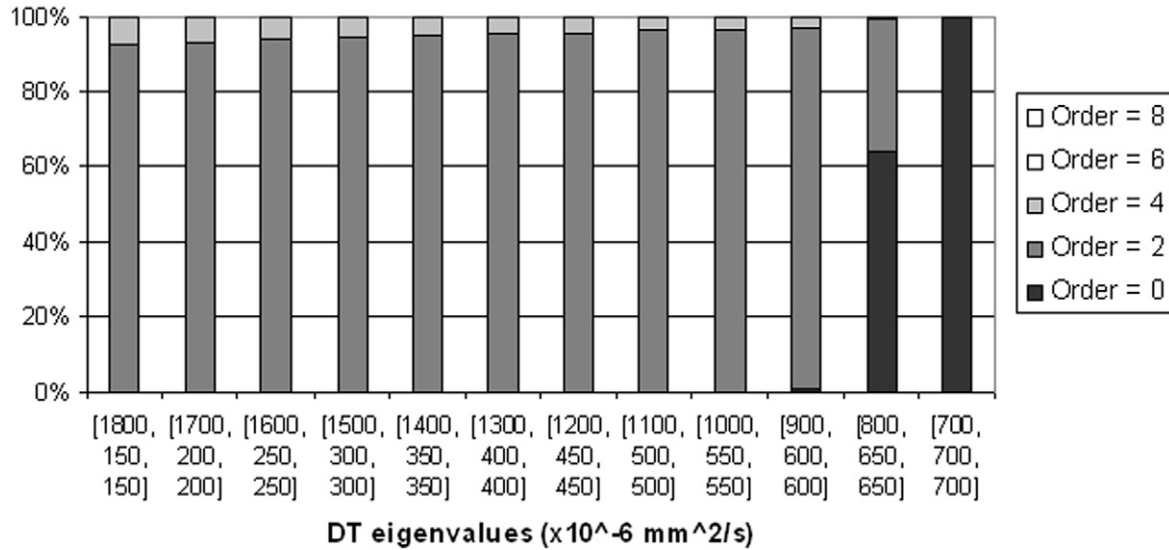
FIG. 2. Proportions of SH series model orders selected for profiles of simulated measurements from prolate DTs with varying levels of anisotropy (increasing right to left).

[1700, 200, 200] ($\times 10^{-6}$ mm^2/s) and $\sigma_0 = 35$. Classification rates from these profiles are plotted in Fig. 3. Figure 3 shows that our simulated fiber intersection profiles are classified as order 4 consistently when the fibers are in equal proportion and are orthogonal, when only 3% of the profiles are underfit as order 2 profiles. However, the rate of misclassification at order 2 increases significantly as the proportion of the two fibers becomes less balanced and as the two fibers become more parallel. This is again reasonable to expect, because both these effects cause the deviation of the profile from Gaussian to decrease. Classification above order 4 for all these bi-Gaussian profiles varies between 0.5% and 1%. Other results (not shown, but see Ref. 23) demonstrate that the deviation from Gaussian behavior that is obtained by mixing isotropic and anisotropic Gaussian diffusion with the bi-Gaussian model is not reliably detected by our procedure with these imaging parame-

ters—only a minor increase in the order 4 classification rate (from 7% to 18%) is observed for these mixed profiles.

Human Data Experiments

DW-MRI data was acquired from four healthy volunteers using a protocol similar to that outlined by Jones et al. (9). All subjects were scanned with the approval of the joint National Hospital and Institute of Neurology ethics committee and gave informed, written consent. Three unweighted ($b = 0$ s/mm^2) images were acquired together with 60 DW images with different gradient directions spread evenly over the hemisphere, with gradient pulse width $\delta = 0.032$ s, pulse separation $\Delta = 0.04$ s, and gradient strength G = 0.022 Tm^{-1}, which gives $b \approx 1000$ s/mm^2 in each case. The reconstructed image array is 128 $\times$ 128 in plane with a field of view (FOV) of 220 mm, and a
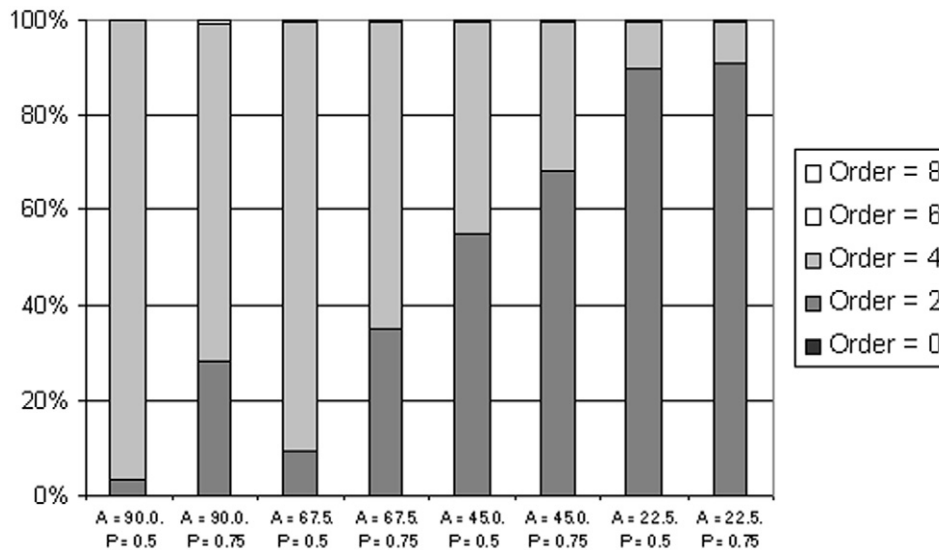

FIG. 3. Proportions of SH series model orders selected for profiles of simulated measurements of a bi-Gaussian ADC profile from a combination of two prolate DTs intersecting at various angles, A. Results are included for cases in which there are equal proportions ($Q = 0.5$) of the two tissue types, and three times more of one type ($Q = 0.75$).
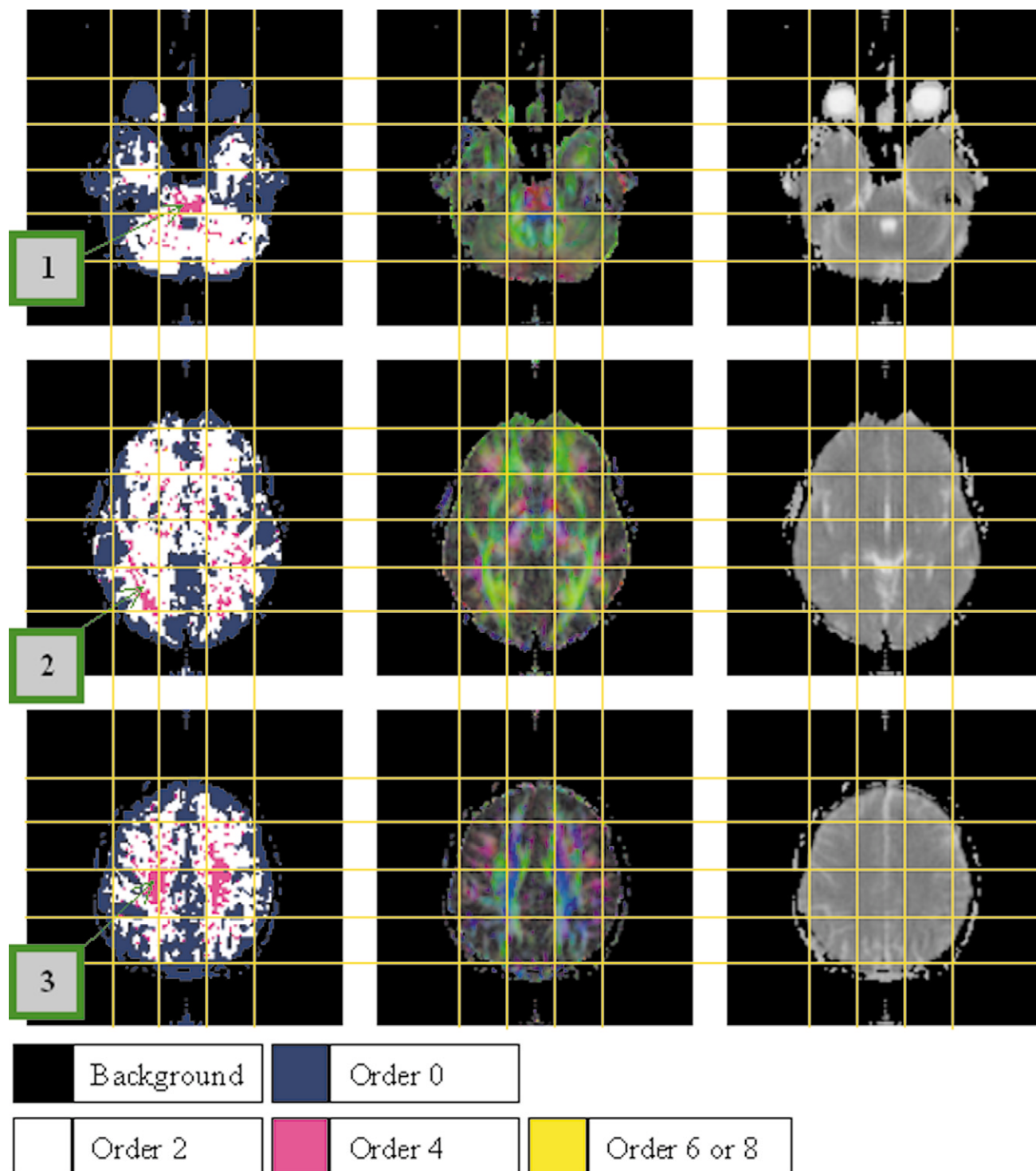
FIG. 4. Order maps from human brain data (left) together with color-coded principal DT direction maps (middle) and mean diffusivity maps (right). The key below the images refers to the order maps in the left column.

total of 42 slices evenly spaced at 2.5 mm intervals were acquired.

The procedures described in Methods were applied to this data. Since 60 samples are available for each ADC profile, the maximum series order that we can fit is 8 (45 parameters). We thus compute the coefficients of all even-order SHs up to and including order 8 and then apply the model selection procedure described above, with the decision thresholds determined from the simulated data, to choose the appropriate level of truncation of the series. Prior to model selection, two thresholds on σ_0 are applied: if $\sigma_0 < 8.5$, the voxel is classified as background and no model is assigned; if $\sigma_0 > 85$, we assume the voxel corresponds to CSF and assign an order 0 model, since flow
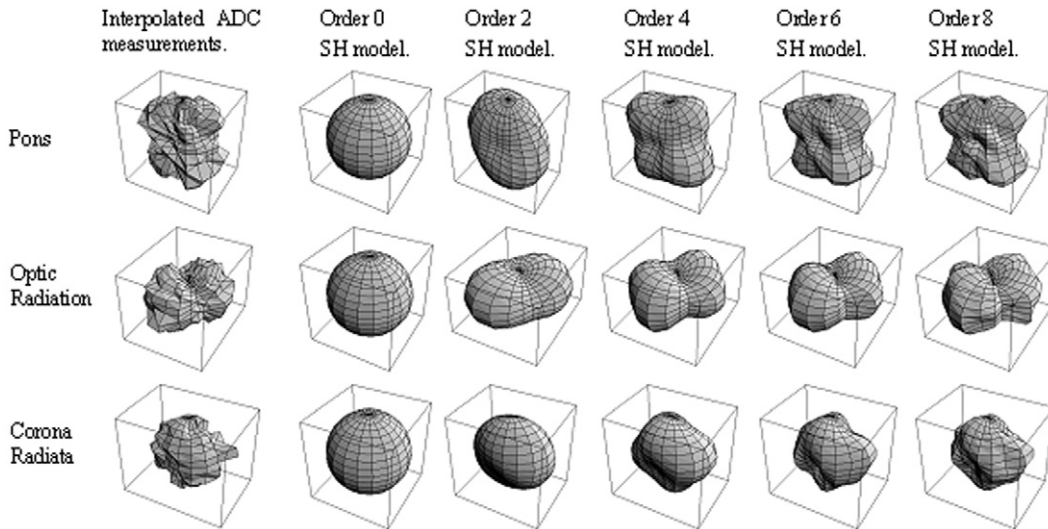
FIG. 5. Typical measurements (left) together with SH models of orders 0, 2, 4, 6, and 8 (second from left to right) from each of the three higher-order clusters labeled in Fig. 4. Top: measurement from the pons (order 4 model selected). Middle: measurement from the optic radiation (order 4 model selected). Bottom: measurement from the corona radiata (order 4 model selected).

artifacts in these regions make the measurements unreliable.

Figure 4 shows maps of the order chosen by our model selection procedure (left) for three axial slices of one of our data sets (results from the other data sets can be found in Ref. 23), together with mean diffusivity (right) and anisotropy-weighted color maps indicating the principal direction of the DT at each point (middle). In the images in the middle column, the red, green, and blue intensities correspond to the size of the x, y, and z components of the DT principal direction unit vector, weighted by the fractional anisotropy (2) as proposed by Pajevic and Pierpaoli (18).

A detailed anatomical analysis of the results is beyond the scope of this work, but we will highlight some interesting features. First, our procedure appears to separate regions of isotropic and anisotropic diffusion successfully. Regions of GM are mostly assigned order zero models, while regions of WM, both dense and peripheral, are mostly assigned order 2. A significant percentage of voxels are assigned higher order models. The spatial distribution of these voxels exhibits distinct clusters in the image, which have clear symmetry about the midline of the brain. The clustering and symmetry strongly suggest that the higher-order models correspond to genuine anatomical effects.

Each slice shown in Fig. 4 contains an anatomic region that contains clusters of higher-order voxels consistently in each of our four data sets. In the top slice, the region corresponding to the pons contains a dense cluster of order 4 voxels (label 1). In this region, the right–left trans-pontine tracts cross the inferior-superior pyramidal tracts, causing partial volume effects between these two orthogonal fibers. In the middle slice, dense clusters of order 4 voxels are found in the optic radiation on both sides of the brain (label 2). These occur precisely at the points where the anterior–posterior tracts of the optic radiation cross the predominantly right–left fibers of the corpus callosum. In the bottom slice, large clusters of order

4 voxels can be observed within the corona radiata (label 3). Again this is reasonable to expect, since the diverging fibers of the corona radiation are crossed by U-fibers in these areas.

Figure 5 shows typical ADC profiles from each of these three regions, together with the order 0, 2, 4, 6, and 8 SH models. In each case, it is clear that there is significant difference between the order 4 and order 2 models, which indicates significant non-Gaussian behavior. The models with order greater than 4 do not appear to change the overall profile shape significantly further, and serve only to incorporate noise effects. The measurement from the pons is particularly striking and is typical of that predicted by the bi-Gaussian model for two orthogonally intersecting WM fibers (see Fig. 1). The measurement from the optic radiation is similar but oriented within the axial plane and somewhat less pronounced, possibly due to a less balanced mix between the two fibers, or a smaller angle of intersection. The profile from the corona radiata is more difficult to interpret, and may be the result of effects that are not modeled naturally by a bi-Gaussian model, such as fiber divergence or compartments with exchange.

Further Synthetic Experiments

Since the set of DTs tested in the synthetic experiments described in the Simulation Experiments section is not exhaustive, it is possible that the regions we observe in the human brain data simply exhibit Gaussian diffusion processes that are particularly prone to overfitting in our procedure. Thus, for completeness, we include an additional set of experiments to test this assertion. In each region, we fit the DT (rather than the full SH series) at each voxel and model the distribution of DT eigenvalues, $\lambda_1 > \lambda_2 > \lambda_3$, together with the value of σ_0. We then use a Monte Carlo method to draw samples from this distribution and test the likelihood that the corresponding Gaussian ADC profiles are overfitted with SH models with order greater

than 2. σ_0 is included in this model, because the noise levels in ADC measurements depend on σ_0 as well as the magnitude of the ADC itself (19).

For each of the regions highlighted in the previous section (the pons, optic radiation, and corona radiata), a subregion of the order 4 cluster was selected. A model of the distribution of DT shapes in each subregion was obtained by fitting a 4D multivariate Gaussian model to the distribution of vectors (λ_1, λ_2, λ_3, and σ_0). Our procedure was then applied to large numbers of samples drawn from each of these distributions, and the rate of classification above order 2 was found to be less than 5% in each case. These results verify that if the underlying diffusion process had been Gaussian, our procedure would have been effective and assigned most profiles in these regions order 2 models. Thus we conclude that we are observing genuine non-Gaussian effects in these regions.

DISCUSSION

We have described methods for modeling and detection of non-Gaussian ADC profiles and shown that such profiles can be observed with scanning parameters typical of standard clinical DW-MRI data. The SH series up to order 8 was fit to samples of the ADC profile in each voxel, which provides a sequence of models of increasing complexity. A series of ANOVA tests was used to find the simplest of these models that adequately describes the data. This latter procedure classifies the profile at each voxel as isotropic, anisotropic Gaussian, or non-Gaussian, and allows maps of these different types of behavior to be produced, as shown in Fig. 4.

Our procedure was applied to human brain data collected with parameters typical of those used in clinical scans, and appeared to classify isotropic (GM) and anisotropic (WM) regions correctly as order 0 and order 2, respectively. On average, 5% of profiles in voxels within the brain were classified as order 4 or above (non-Gaussian). Several regions—in particular, the pons, optic radiation, and corona radiation—were found consistently to contain dense clusters of order 4 models. Validation of our method was performed by characterizing its performance using synthetic data with realistic noise properties, as well as applying it to DT models of data in regions of the brain that were consistently classified as non-Gaussian. Although results from only one data set are shown here, our method was applied to four data sets and other results can be found in Ref. 23. Acquisition of a larger ensemble of data sets is currently underway, which will enable parametric mapping to be performed in order to allow comparisons to be made between the occurrence of non-Gaussian diffusion in different population groups.

The behavior maps produced by our method provide new insights into the complexity of tissue structure in the brain. These maps have a number of practical applications. As mentioned in the Introduction, one aim is to use these maps as a stain for diagnosis of structure-reducing pathology. Furthermore, these maps can be used in postprocessing algorithms, such as tractography algorithms, which need to identify when diffusion is anisotropic and when the principal direction of the DT can be relied upon to describe the orientation of the underlying tissue. Finally,

these maps indicate when a more complex model (for example, a bi-Gaussian model) than the DT needs to be used to describe p_t adequately. Typically, such models are more difficult to fit to data than the DT and nonlinear fitting algorithms must be employed (see for example Ref. 21). Such procedures are computationally expensive, so it is advantageous to be able to identify only when they need to be performed.

It seems likely that most of the non-Gaussian behavior we observe is due to the intersection of WM fibers with different orientation. This kind of tissue structure gives rise to profiles similar to those that are derived from multi-Gaussian p_t's (22), although it is also likely that some exchange of particles between the tissue compartments corresponding to each fiber occurs in the timescale of the diffusion measurement, which will cause p_t to deviate from the multi-Gaussian (22). Other types of non-Gaussian processes almost certainly occur in the brain, caused by restriction due to impermeable barriers (22). However, the deviation from Gaussian that is caused by these effects is less marked than those due to multicompartmentation, so such behavior may not be observed reliably—particularly at low *b*-values such as those used here.

The advantages of the SHs as models for ADC profiles lie both in their generality and in the simplicity and robustness of the fitting procedure. Linearity of the fitting procedure is a significant advantage in terms of computational effort, but also ensures that the fitting procedure is well posed and is not prone to spurious erroneous results. There is little *physical* justification for the use of the SHs in this context, and there may be more appropriate sets of basis functions that better reflect the kind of ADC profiles that are likely to arise given the underlying physical processes. An advantage of this series, however, is its generality: any profile can be represented, and we do not limit ourselves to particular models of the underlying processes. We note for clarity that when p_t is non-Gaussian the shape of the ADC profile does not relate to p_t in a straightforward way, and thus the SH shape does not provide any direct information about the underlying tissue structure. Although it is possible to extract some information of this type from non-Gaussian profiles on the sphere (21), this issue is beyond the scope of this work.

There are many techniques for model selection in the literature. The use of ANOVA and the *F*-test for deletion of variables has proved more successful than most in our application, but there may be others that improve performance to some extent. The test we used is based on an assumption of Gaussian errors in the measurements. Although this is a reasonable first approximation for our data (23), the full analytic form of the errors in ADC measurements is not Gaussian. There are tests that incorporate noise models for the data, which may improve classification performance, particularly in extreme cases such as very anisotropic diffusion. We note that the same technique could be used to produce a finer classification of diffusion profiles; for example, we could distinguish between axisymmetric (two eigenvalues equal) and nonsymmetric (all eigenvalues unequal), anisotropic Gaussian diffusion.

In the present study we used only data acquired with common clinical imaging parameter settings, and one of

our goals was to demonstrate that significant non-Gaussian behavior can be observed at *b*-values as low as 1000 s/mm^2. Recently there has been a trend to move toward higher *b*-values, which can produce profiles richer in information; in particular, non-Gaussian behavior becomes more pronounced (10). Our methods are equally applicable to data acquired with higher *b*-values, and we would expect to observe a higher proportion of non-Gaussian profiles in such data, which may highlight other regions of the brain in which non-Gaussian behavior occurs.

REFERENCES

1. Basser PJ, Matiello J, Le Bihan D. MR diffusion tensor spectroscopy and imaging. Biophys J 1994;66:259−267.
2. Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative diffusion tensor MRI. J Magn Reson B 1996;111:209−219.
3. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. Radiology 1996;201:637−648.
4. Callaghan PT. Principles of magnetic resonance microscopy. Oxford, UK: Oxford Science Publications; 1991.
5. Karger J, Pfeifer H, Heink W. Principles and applications of self diffusion measurements by nuclear magnetic resonance. Adv Magn Reson 1988;12:1−89.
6. Mitra PP, Halperin BI. Effects of finite gradient pulse widths in pulsed field gradient diffusion measurements. J Magn Reson 1995;113:94−101.
7. Crank J. The mathematics of diffusion. Oxford, UK: Oxford University Press; 1975.
8. Stejskal EO, Tanner JE. Spin diffusion measurements: spin echoes in the presence of time-dependent field gradient. J Chem Phys 1965;42:288−292.
9. Jones DK, Horsfield MA, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn Reson Med 1999;42:515−525.
10. Alexander AL, Hasan KM, Lazar M, Tsuruda JS, Parker DL. Analysis of partial volume effects in diffusion-tensor MRI. Magn Reson Med 2001; 45:770−780.
11. Frank LR. Characterization of anisotropy in high angular resolution diffusion weighted MRI. In: Proceedings of the 9th Annual Meeting of ISMRM, Glasgow, Scotland, 2001. p 1531.
12. Tuch DS, Weisskoff RM, Belliveau JW, Wedeen VJ. High angular resolution diffusion imaging of the human brain. In: Proceedings of the 7th Annual Meeting of ISMRM, Philadelphia, 1999. p 321.
13. Wedeen VJ, Reese TG, Tuch DS, Weigel MR, Dou J-G, Weisskoff RM, Chesler D. Mapping fiber orientation spectra in cerebral white matter with Fourier transform diffusion MRI. In: Proceedings of the 8th Annual Meeting of ISMRM, Denver, 2000. p 82.
14. Frank LR. Anisotropy in high angular resolution diffusion-weighted MRI. Magn Reson Med 2001;45:935−939.
15. Nikiforov AF, Uvarov VB. Special functions of mathematical physics. Basel: Birkhauser Verlag; 1988.
16. Brechbuhler CH, Gerig G, Kubler O. Parametrization of closed surfaces for 3D shape description. Comput Vis Image Underst 1995;61:154−170.
17. Armitage P, Berry G. Statistical methods in medical research. Oxford, UK: Blackwell Scientific Publications; 1971.
18. Pajevic S, Pierpaoli C. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. Magn Reson Med 1999;42:526−540.
19. Armitage PA, Bastin ME. Utilizing the diffusion-to-noise ratio to optimize magnetic resonance diffusion tensor acquisition strategies for improving measurements of diffusion anisotropy. Magn Reson Med 2001;45:1056−1065.
20. Alexander DC, Barker GJ, Arridge SR. Detection and modelling of non-Gaussianity in MR diffusion imaging. UCL, Department of Computer Science Research Note, RN/01/35, 2001. http://www.cs.ucl.ac.uk/staff/D.Alexander/Papers/RN35−2001.pdf
21. Alexander DC, Jansons KM. Spin echo attenuation to diffusion displacement density: a general inversion for measurements on a sphere. In: Proceedings of the ISMRM Workshop on Diffusion MRI: Biophysical issues (what can we measure?), St. Malo, France, 2002.
22. Le Bihan D. Molecular diffusion, tissue microdynamics and microstructure. NMR Biomed 1995;8:375−386.
23. Anderson AW. Theoretical analysis of the effects of noise of diffusion tensor imaging. Magn Reson Med 2001;46:1174−1188.