

A STATISTICAL FRAMEWORK FOR HIGH ANGULAR RESOLUTION DIFFUSION IMAGING IN Q -SPACE

BY SOFIA C. OLHEDE^{*,†} AND BRANDON WHITCHER[‡]

University College London[†] and GlaxoSmithKline[‡]

High angular resolution diffusion imaging data is the observed characteristic function for the local diffusion of water molecules in tissue. These data are used to infer the organization of white matter fiber bundles in structural brain imaging. Non-parametric scalar measures are proposed to summarize such data, and to locally characterize spatial features of the diffusion probability density function (PDF), relying on the geometry of the characteristic function. Summary statistics are defined so that their distributions are, to first order, both independent of nuisance parameters and analytically tractable. The dominant direction of the diffusion at a spatial location (voxel) is determined, and a new set of axes are introduced in Fourier space. Variation quantified in these axes determines the local spatial properties of the diffusion density. Non-parametric hypothesis tests for determining whether the diffusion is unimodal, isotropic or multi-modal are proposed. More subtle characteristics of white-matter microstructure, such as the degree of anisotropy of the diffusion PDF and symmetry compared with a variety of asymmetric PDF alternatives, may be ascertained directly in the Fourier domain without parametric assumptions on the form of the diffusion PDF. We simulate a set of diffusion processes and characterize their local properties using the newly introduced summaries. We show how complex white-matter microstructure that spans multiple voxels exhibits clear ellipsoidal and asymmetric structure in simulation, and assess the performance of the statistics in clinically-acquired MRI data.

1. Introduction. Many applications in brain imaging are based on calculating local statistics that are later combined to infer global properties of spatial links or connections. In this paper we focus on the local analysis of high angular resolution diffusion imaging (HARDI) data, a special type of magnetic resonance imaging (MRI). HARDI observations correspond to the local (voxel-wise) measurement of the Fourier transform of the local water molecule diffusion probability density function (PDF) ([Callaghan, 1993](#)) in a number of different orientations over a spherical shell of fixed radius.

A HARDI acquisition scheme permits the characterization of directional

^{*}Sofia Olhede is with the Departments of Computer Science and Statistical Science.

Keywords and phrases: anisotropy, asymmetry, magnetic resonance imaging, diffusion weighted imaging, non-parametric

spatial properties of the diffusion PDF. The local structure of brain tissue can be inferred from such measurements [Basser \(2002\)](#); [Basser, Mattiello and Le Bihan \(1994\)](#). Once local statistics have been formed, it is of interest to combine information across voxels (spatial locations), to track nerve fiber bundles (e.g. connect local directions of estimated diffusion PDFs to recognize nerve connections) or to infer local fiber structure from the estimated diffusions ([Mori and van Zijl, 2002](#)), and/or to use other locally-defined summaries in inference procedures ([Jensen *et al.*, 2005](#)).

Different orientational sampling designs can be used at each location (or voxel) and if a simple parametric model is used for the diffusion PDF, then rather sparse sampling will be sufficient to recover the parameters of the model. Traditional analysis of HARDI measurements is based on modelling the diffusion PDF parametrically as a (zero-mean) Gaussian, and estimating a diffusion tensor (the covariance matrix of the Gaussian PDF), a procedure which corresponds to diffusion tensor imaging (DTI) ([Basser, Mattiello and Le Bihan, 1994](#)). Such methods have drawbacks, for example not describing more complex fiber structures very well, and their usage trades a small variance for potentially large biases. While the diffusion tensor model has both theoretical justification and has been extremely popular, it prohibits one from describing more complicated (white-matter voxel) structure, such as crossing, kissing and forking fibers ([Mori and van Zijl, 2002](#)).

It is believed that intravoxel orientational heterogeneity affects as many as one third of all imaged white-matter voxels ([Behrens *et al.*, 2007](#)), and so addressing such structure is important. With more time-intensive sampling schemes (such as HARDI ([Tuch *et al.*, 2002](#)) or diffusion spectrum imaging) various more complicated estimators can be used, e.g. multi-tensor modelling ([Alexander, 2005](#)), and non-parametric alternatives such as persistent angular structure MRI ([Jansons and Alexander, 2003](#)), Q-ball imaging ([Tuch, 2004](#)), the diffusion orientation transform ([Özarslan *et al.*, 2006](#)), and spherical deconvolution ([Tournier *et al.*, 2004](#)). While a non-parametric approach may remove bias, usage of such non-parametric methods is challenging because the diffusion is measured in the Fourier domain (q -space¹), and the characteristic function has been considerably undersampled for scanning times that may be used in the clinic. This challenges the stable inversion of

¹ q -space is the Fourier domain representation of the local diffusion and is the space where measurements are made. The global image Fourier representation is usually inverted to a spatial representation, but the local Fourier transform is not inverted as part of the acquisition, leaving the spatial domain observations associated with a measurement of local diffusion in a Fourier domain orientation.

information (the local characteristic function) to local spatial structure.

This paper develops a statistical framework for characterising HARDI data directly in q -space (Tuch *et al.*, 2002) without local inversion. This avoids calculating non-linear transformations of the data, whose usage usually leads to intractability of the distributions of calculated data summaries. The approximate distributions of the proposed estimators in this paper are derived and these are defined so that to first order they are free of any nuisance parameters. The proposed statistics are a first step towards automated detection of subtle characteristics of white-matter microstructure that go beyond the typical measurements; i.e., scalene diffusions or asymmetry in decay in a fixed axes. Both properties, scalene diffusion and asymmetry, have been found in a forking fiber structure, and may be important summaries to feed into fiber-tracking algorithms (Mori and van Zijl, 2002). The derived methods also serve as a warning when interpreting multi-tensor models in clinically-feasible acquisition schemes, as similar characteristics can be obtained from more complex single peaked structures.

Global features like bi- or multi-modality of the diffusion PDF are described reasonably well by many methods over a range of signal-to-noise values, with the small caveat that the various implicit assumptions inherent to any of the given methods must be satisfied. Parametric models introduce bias when they are not appropriate, whereas using a nonparametric method increases the variance in the estimation. Using a moderate number of directions in the HARDI sampling scheme restricts the possibility of determining smaller scale structure of the diffusion PDF. Strong parametric assumptions increase the power of any proposed statistic to detect multiple diffusion directions, with the consequence that any deviation from the prescribed structure in the parametric model may be used to reject null hypotheses such as unimodality.

In the method proposed here to determine the properties of the diffusion PDF, the directional PDFs (i.e., diffusion PDFs that are strongly elongated in a given direction) are separated from the isotropic PDFs (non-preferential directional structure), using a test based on a comparison of relative magnitudes in q -space. Subsequently, multimodal distributions are sifted from the isotropic and unidirectional. The unidirectional diffusion is associated with a great circle in q -space (Tuch, 2004), and we call this the *dominant great circle*. The strongest direction defines an important spatial summary of the diffusion PDF, and specifies the major axis of the diffusion in q -space. The perpendicular to the major direction in space defines a set of points lying on a great circle in q -space, which exactly corresponds to the dominant great circle.

If a given voxel has been diagnosed as unidirectional (or if there is a dominant great circle in q -space) then we seek to characterize its main unidirectional structure in more detail. A scalar measure resembling the popular fractional anisotropy² statistic is defined as the *anisotropy statistic*, by comparing the magnitude of the q -space diffusion on the dominant great circle with its two perpendicular point(s). This measure determines the degree of anisotropy of the diffusion PDF. Further investigation of unidirectional voxels causes us to focus on quantifying the uniformity of decay in the minor axes of the diffusion PDF, or the perpendicular to the dominant great circle, to describe further detailed structure of the characteristic function.

Ellipsoidal diffusions are an important class of diffusion processes and the scalene structure of the diffusion PDF is particularly important when combining voxel-wise information (Seunarine *et al.*, 2007). The aforementioned work showed that the scalene structure of the peak is related to the peak anisotropy in space and important for treating bending and fanning fibers. Seunarine *et al.* (2007) used the peak Hessian to improve PICO (Parker and Alexander, 2005), a fiber tracking method, but such information can feed into any tracking algorithm of choice. For diffusions with ellipsoidal decay, their minor axes are well defined by this (scalene) decay structure, while for non-ellipsoidal diffusions the minor axes correspond to a set of axes in the plane of the dominant great circle, parameterising locations on the dominant great circle. We examine the scalene structure of the PDF, which is quantified by the difference in decay in the two spatial minor axes, defined as such also for non-ellipsoid diffusions. A statistical test for uniformity on the great circle is developed that can be related to the spatial decay of the diffusion PDF in the minor axes. Another feature of interest in the PDF is asymmetry in the decay in a fixed direction perpendicular to the dominant great circle. This heuristic may be visualized in space as a diffusion PDF that appears ellipsoidal but the peak is in one of the foci rather than the center of the ellipse. We introduce a test statistic for asymmetry based on this understanding.

The methodology presented here improves our understanding of the diffusion PDF by not relying on parametric assumptions when analysing the MRI measurements, yet still relating q -space structure directly to spatial properties. Non-parametric statistical summaries are defined directly in q -space to increase power of the proposed hypothesis tests and theoretical critical values for the statistics are provided. Understanding the inherent limitations of HARDI measurements can be obtained directly from our discussion of

²The fractional anisotropy (FA) is a measure of uniformity of the eigenvalues of a Gaussian covariance matrix (Basser and Pierpaoli, 1996).

simulated diffusions, thus increasing our understanding of the parametric assumptions that are necessary to derive more complicated structures from the diffusion PDF.

2. Statistical Models for HARDI Data.

2.1. Observational Model. We denote the sampling of the observations by the set $\mathcal{Q}_0 = \{\tilde{\mathbf{q}}_i\}_{i=1}^n$. At each $\tilde{\mathbf{q}}_i = (\tilde{q}_{i1}, \tilde{q}_{i2}, \tilde{q}_{i3})$ on the unit sphere $\|\tilde{\mathbf{q}}\| = (\tilde{q}_1^2 + \tilde{q}_2^2 + \tilde{q}_3^2)^{1/2} = 1$ we obtain an observed measurement $\tilde{A}(\tilde{\mathbf{q}}_i) \geq 0$, corresponding to the magnitude of a complex-valued observation (the noisy characteristic function of local diffusion³). Furthermore we take n_0 observations at $\mathbf{q} = \mathbf{0}$, denoted by $\tilde{A}_k(\mathbf{0})$ for $k = 1, \dots, n_0$. We distinguish here between the measured apparent diffusion at $\tilde{\mathbf{q}}_i$, namely $\tilde{A}(\tilde{\mathbf{q}}_i)$, and the theoretical diffusion value, $\mathcal{A}(\tilde{\mathbf{q}}_i)$. Note that the expected value of $\tilde{A}(\tilde{\mathbf{q}}_i)$ is *not* equivalent to $\mathcal{A}(\tilde{\mathbf{q}}_i)$, for two reasons. Firstly because the observations are magnitudes, with the noise contributing in the expectation, and secondly we need to renormalize the observed diffusion to have unit volume, as noted by [Alexander \(2005\)](#). As the PDF is a density it has to satisfy the normalization of

$$(2.1) \quad \int \int \int a(\mathbf{x}) d^3\mathbf{x} = 1 \Rightarrow \mathcal{A}(\mathbf{0}) = 1,$$

where $a(\mathbf{x})$ is the diffusion probability density function (PDF). We apply a biased estimator of a simple average to estimate the inverse of the normalizing constant by $\overline{A}(\mathbf{0}) = \sum_{k=1}^{n_0} \tilde{A}_k(\mathbf{0})/n_0$. We renormalize the observed diffusion such that $A(\tilde{\mathbf{q}}_i) = \tilde{A}(\tilde{\mathbf{q}}_i)/\overline{A}(\mathbf{0})$. $A(\tilde{\mathbf{q}}_i)$ has (approximately) a Rician distribution with parameters $\mathcal{A}(\tilde{\mathbf{q}}_i)$ and σ^2 ([Gudbjartsson and Patz, 1995](#)). As the SNR will be large at $\mathbf{q} = \mathbf{0}$, the noise floor of the Rician distribution will have limited impact in the estimation of the normalization.

The normalized diffusion measurements $\mathcal{A}(\tilde{\mathbf{q}}_i)$ should exhibit symmetry as the PDF is real-valued, symmetric and indeed positive; i.e., $\mathcal{A}(-\tilde{\mathbf{q}}_i) = \mathcal{A}(\tilde{\mathbf{q}}_i)$ ([Wedeen et al., 2005](#)). To fully exploit the Hermitian symmetry we shall reflect the observations to the augmented set $\mathcal{Q} = \{\mathbf{q} : \mathbf{q} \in \mathcal{Q}_0\} \cup \{\mathbf{q} : -\mathbf{q} \in \mathcal{Q}_0\}$, and set $A(-\tilde{\mathbf{q}}_i) = A(\tilde{\mathbf{q}}_i)$ as was also done by [Jansons and Alexander \(2003\)](#).

We assume that a non-parametric estimator of the diffusion in q -space is constructed. We chose to use a variable bandwidth estimator ([Olhede and Whitcher, 2008a,b](#)), but the outlined methodology is applicable to other linear estimators, and can be extended to for example radial basis functions and/or

³Note that this is different from the empirical characteristic function.

spherical harmonics, if with some straightforward alteration of the statistical properties (second order structure) of the estimator.

2.2. Great Circles in q -Space. Spatial properties of the diffusion PDF may be described directly in q -space. The advantage of such an operation is that we avoid the need to invert the PDF to the spatial domain for analysis, allowing us to employ a broad range of modelling approaches. A basic building block of our analysis is an *ellipsoid density*. We refer to a density $a_E(\mathbf{x})$ as an ellipsoid density if its FT takes the form

$$(2.2) \quad \mathcal{A}_E(\mathbf{q}; \mathbf{\Lambda}, \mathbf{\Upsilon}) = B \left(\sqrt{\sum_{j=1}^3 \lambda_j |\mathbf{v}_j^T \mathbf{q}|^2} \right),$$

where $\lambda_j \geq 0$, $j = 1, 2, 3$, $\{\mathbf{v}_j\}$ constitutes a basis for \mathbb{R}^3 and $B(\cdot)$ is a monotonically decreasing function. For example, commonly used is the Gaussian characteristic function, with $B(q) = e^{-2(\pi q)^2}$. We collect the eigenvalues in the matrix $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$, and define

$$\mathbf{\Upsilon}^T = \{[v_{11} \ v_{12} \ v_{13}]; [v_{21} \ v_{22} \ v_{23}]; [v_{31} \ v_{32} \ v_{33}]\},$$

to model the axis of any orientational structure. Ellipsoid densities are natural building blocks, just like the special case of the DTI model, but do not (for example) include multimodal densities. If the q -space density takes this form, then the spatial PDF is given by inverting the FT

$$(2.3) \quad a_E(\mathbf{x}; \mathbf{\Lambda}, \mathbf{\Upsilon}) = \int \int \int_{\mathbb{R}^3} \mathcal{A}_E(\mathbf{q}; \mathbf{\Lambda}, \mathbf{\Upsilon}) e^{i2\pi \mathbf{q}^T \mathbf{x}} d^3 \mathbf{q}$$

(Callaghan, 1993). We note for $\mathbf{x} \in \mathbb{R}^3$, with $x = \|\mathbf{x}\|$ and $q = \|\mathbf{q}\|$, that $a_E(\mathbf{x}; \mathbf{\Lambda}, \mathbf{\Upsilon})$ takes the form

$$(2.4) \quad a_E(\mathbf{x}; \mathbf{\Lambda}, \mathbf{\Upsilon}) = |\mathbf{\Lambda}|^{1/2} b \left(\left\| \mathbf{\Lambda}^{-1/2} \mathbf{\Upsilon} \mathbf{x} \right\| \right),$$

$$(2.5) \quad b(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(q) e^{i2\pi \mathbf{q}^T \mathbf{x}} d^3 \mathbf{q}$$

$$(2.6) \quad \begin{aligned} &= \frac{1}{2\pi^2 x} \int_0^{\infty} B\left(\frac{q'}{2\pi}\right) \sin(xq') q' dq' \\ &= \frac{2}{x} \int_0^{\infty} B(q) \sin(2\pi xq) q dq, \end{aligned}$$

which follows from Gradshteyn and Ryzhik (2000, p. 1112). The meaning of “ellipsoid density” becomes clear from this expression, since whenever

$\|\mathbf{\Lambda}^{-1/2}\mathbf{\Upsilon}\mathbf{x}\| = R$, where $R \geq 0$ is a constant, the function $a_E(\mathbf{x})$ takes the same value in space. As long as all the eigenvalues are positive, $a_E(\cdot)$ will therefore map out ellipsoid contours of equal function value in space. The Gaussian DTI model fits into this class of densities with $b(x) = e^{-x^2/2}/(2\pi)^{3/2}$ as do, for example, the Matérn family with the spatial variable exchanged with the spatial frequency variable (Matérn, 1960). The model proposed by Kaden, Knösche and Anwender (2007) is also related to such densities. Typical shapes are found in Figure 1 where we show a prolate ellipsoid distribution, and its Fourier transform ((a) and (e)), a scalene ellipsoid PDF ((b) and (f)), a mixture of ellipsoids ((c) and (g)), and a non-ellipsoid ((d) and (h)). The values of $\mathbf{\Upsilon}$ specify the orientation of the PDF, while $\mathbf{\Lambda}$ gives its qualitative appearance, when coupled with $B(\cdot)$. Unfortunately, looking directly at Figure 1 we do not get the clearest feeling for the local structure near the peak, and so additional ways of characterizing the PDF will be formulated.

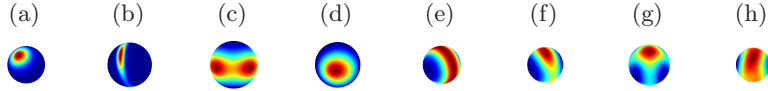


FIG 1. Unimodal densities ((a), (b) and (d)) in space and ((e), (f) and (h)) frequency, as well as a mixture of ellipsoids (c) in space, and (g) in frequency.

2.3. The Orientation Distribution Function. An important tool in understanding HARDI data is via the Orientational Distribution Function (ODF). The ODF quantifies the directional structure of the diffusion PDF *in space*. A popular object of study, it corresponds to several *different* functions. Several authors define the ODF to be

$$(2.7) \quad \text{ODF}_T(\theta, \phi) = \frac{1}{Z} \int_0^\infty a(r\mathbf{u}) dr,$$

where $\mathbf{x} = r\mathbf{u}$, $\|\mathbf{u}\| = 1$ and Z is a normalizing constant (Descoteaux *et al.*, 2007; Hess *et al.*, 2006; Tuch, 2004). A non-linear transformation is necessary for the ODF to have a more peaked and clear directional structure. Because this is not a true marginalization of a PDF (the increment needs a weighting by r^2), and weighs very low scales heavily, the diffuse directional structure of the large scale structure smooths the marginal PDF of orientations, giving it a “blunted” appearance. Wedeen *et al.* (2005) define the ODF as the truly marginalized PDF over all spatial radii

$$\text{ODF}_W(\theta, \phi) = \int_0^\infty r^2 a(r\mathbf{u}) dr.$$

An alternative version may be found in [Jansons and Alexander \(2003\)](#), where the orientational structure associated with a single radius is fitted to the observed data (i.e. the Persistent Angular Structure (PAS-MRI) algorithm). It is useful to note that the observed data is not associated purely with a single radius, and for this to be a mathematically correct procedure the observed HARDI measurements should be convolved with a suitable kernel prior to estimation. Despite this fact PAS-MRI usually gives good results in practice. All three of these orientational summaries are measuring *different* properties of the directional structure of the data, and only $\text{ODF}_W(\cdot, \cdot)$ is a true marginal PDF.

Another directional representation of diffusion data corresponds to the spherical convolution model ([Tournier *et al.*, 2004](#)). In this model, q -space observations are modelled as convolved fiber ODFs, and fiber populations are estimated using deconvolution methods. The magnitudes are then not comparable with previously defined estimators of ODFs. Extensions to these methods have been proposed: by modelling the ODF as a mixture of Bingham distributions ([Kaden, Knösche and Anwender, 2007](#)), and by regularizing the deconvolution problem by proposing the use of constrained optimization methods ([Jian and Vemuri, 2007](#)). The solution in [Kaden, Knösche and Anwender \(2007\)](#) is parametric and the theoretical assumptions necessary to apply the regularized methods are (in general) violated ([Jian and Vemuri, 2007](#)).

The ellipsoid PDF model in (2.4) can be extended into a larger class of arbitrarily peaked deformed PDFs by taking $\mathbf{\Lambda}(\mathbf{x}) = \text{diag}(\lambda_{11}(\mathbf{x}), \lambda_{22}(\mathbf{x}), \lambda_{33}(\mathbf{x}))$, $\lambda_{jj}(\mathbf{x}) \geq 0$ for all \mathbf{x} , with C a normalizing constant,

$$(2.8) \quad a_{DE}(\mathbf{x}) = C \sqrt{|\mathbf{\Lambda}(\mathbf{\Upsilon}\mathbf{x})|} b \left(\left\| \mathbf{\Lambda}(\mathbf{\Upsilon}\mathbf{x})^{-1/2} \mathbf{\Upsilon}\mathbf{x} \right\| \right),$$

$$(2.9) \quad a_{DE}(\mathbf{\Upsilon}^T \mathbf{x}) = C \sqrt{|\mathbf{\Lambda}(\mathbf{x})|} b \left(\left\| \mathbf{\Lambda}(\mathbf{x})^{-1/2} \mathbf{x} \right\| \right).$$

Because $\mathbf{\Lambda}(\mathbf{x})$ is a diagonal matrix $a_{DE}(\mathbf{\Upsilon}^T \mathbf{x})$ exhibits the axes (1, 0, 0), (0, 1, 0) and (0, 0, 1). Implementing a Fourier transform directly with the change of variables, the Fourier transform is mixed over the strengths in $\mathbf{\Lambda}(\mathbf{x})$, but exhibits the same orientational axes, if the ordering in magnitude of the eigenvalues does not switch over \mathbf{x} . We have the model of

$$(2.10) \quad \mathcal{A}_{DE}(\mathbf{q}) = C \int \int \int_{\mathbb{R}^3} |\mathbf{\Lambda}(\mathbf{x})|^{1/2} b \left(\left\| \mathbf{\Lambda}(\mathbf{x})^{-1/2} \mathbf{x} \right\| \right) e^{-i2\pi(\mathbf{\Upsilon}\mathbf{q})^T \mathbf{x}} d^3 \mathbf{x}.$$

This function can take the appearance of a deformed ellipsoid in space, and may then exhibit a different pattern of decay to the left and right of the dominant great circle in q -space. For the regular ellipsoid distribution $a_E(\mathbf{x})$ if one eigenvalue is larger than the two others (say $\lambda_1 > \lambda_2 \geq \lambda_3$)

then the ellipsoid density (or equally in the case of the deformed density if $\inf_{\mathbf{x}} \lambda_1(\mathbf{x}) > \sup_{\mathbf{x}} \lambda_2(\mathbf{x})$) will observe a maximum at the values (compare with (2.2))

(2.11)

$$\mathbf{q}(\beta) = \begin{cases} \beta \mathbf{v}_2 + \sqrt{1 - \beta^2} \mathbf{v}_3 & \text{if } \beta \in [-1, 1]; \\ \text{sgn}(\beta) (|\beta| - 1) \mathbf{v}_2 - \sqrt{1 - (|\beta| - 1)^2} \mathbf{v}_3 & \text{if } \beta \in [-2, 2] \setminus [-1, 1], \end{cases}$$

noting the match to Figure 1(e), (f) and (h). This maximum great circle in q -space corresponds to the perpendicular vector $\pm \mathbf{v}_1$ in space, where the diffusion PDF exhibits a maximum. The structure near the peak ($\mathbf{x} = \pm \mathbf{v}_1$) is mapped to a structure contiguous to the great-circle; i.e., $\mathbf{q} \approx \mathbf{q}(\beta)$, and now compare Figure 1(a) with (b) and (d), with the corresponding frequency descriptions of 1(e) with (f) and (h). For this reason it is desirable to investigate the structure of the PDF near the great circle of points $\{\mathbf{q}(\beta)\}$, in terms of distances from the great circle to characterize structure in the decay from the main peak. To obtain consistency in notation, we define the set of points (or the great circle perpendicular to \mathbf{v}) via $\mathcal{G}(\mathbf{v}) = \{\mathbf{q} : \mathbf{v}^T \mathbf{q} = 0, \|\mathbf{q}\| = q\}$ and $\mathcal{G}(\mathbf{v}_1) \equiv \{\mathbf{q}(\beta)\}$. It is convenient to keep both sets of notation for ease of exposition.

3. Scalar Summaries and Test Statistics.

3.1. Axes of Symmetry. Before we can define appropriate scalar summaries in q -space, additional axes to the β axis (2.11) are required. For any fixed vector $\mathbf{q}(\beta) \in \mathcal{G}(\mathbf{v}_1)$ we traverse a great circle using the vectors

$$(3.1) \quad \mathbf{q}_\perp(\alpha, \beta) = \alpha \mathbf{v}_1 \pm \sqrt{1 - \alpha^2} \mathbf{q}(\beta), \quad \alpha \in [-1, 1],$$

where for $\alpha \in [-2, 2] \setminus [-1, 1]$, the corresponding expression may be formed as in (2.11). Such a great circle for a fixed value of β will be referred to as a *perpendicular great circle*.

An important component in the definition of our non-parametric summaries is the *dominant great circle* $\mathcal{G}(\mathbf{x}_{\max})$ with \mathbf{x}_{\max} given by

$$(3.2) \quad \mathbf{x}_{\max} = \arg \max_{\mathbf{v}} \left\{ \oint_{\mathbf{q} \in \mathcal{G}(\mathbf{v})} \mathcal{A}(\mathbf{q}) d\mathbf{q} \right\}.$$

If $\mathcal{A}(\mathbf{q})$ is an isotropic diffusion process, then \mathbf{x}_{\max} is any vector in \mathbb{R}^3 with a fixed norm. Alternatively, if $\mathcal{A}(\mathbf{q})$ is ellipsoid with $\lambda_1 > \lambda_2 \geq \lambda_3$ then $\mathbf{x}_{\max} = \mathbf{v}_1$. If there are two fibers, then with relative weights of a_1 and a_2 of fiber populations with $\mathbf{\Lambda}^{(1)}$ and $\mathbf{\Lambda}^{(2)}$ the individual eigenvalues

(3.3)

$$\mathbf{x}_{\max} = \arg \max_{\mathbf{v}} \left\{ \left[a_1 \oint_{\mathbf{q} \in \mathcal{G}(\mathbf{v})} \mathcal{A}_E(\mathbf{q}; \mathbf{\Lambda}^{(1)}, \mathbf{\Upsilon}^{(1)}) + a_2 \oint_{\mathbf{q} \in \mathcal{G}(\mathbf{v})} \mathcal{A}_E(\mathbf{q}; \mathbf{\Lambda}^{(2)}, \mathbf{\Upsilon}^{(2)}) \right] d\mathbf{q} \right\}.$$

For example if $a_1 \gg a_2$ then $\mathbf{x}_{\max} \approx \mathbf{v}_1^{(1)}$, or if $a_1 = a_2 = 1/2$ and the great circles do not separate, then \mathbf{x}_{\max} will lie precisely between the two maxima of the two PDFs. Once the great circles start to separate the maximum will go with one of the two.

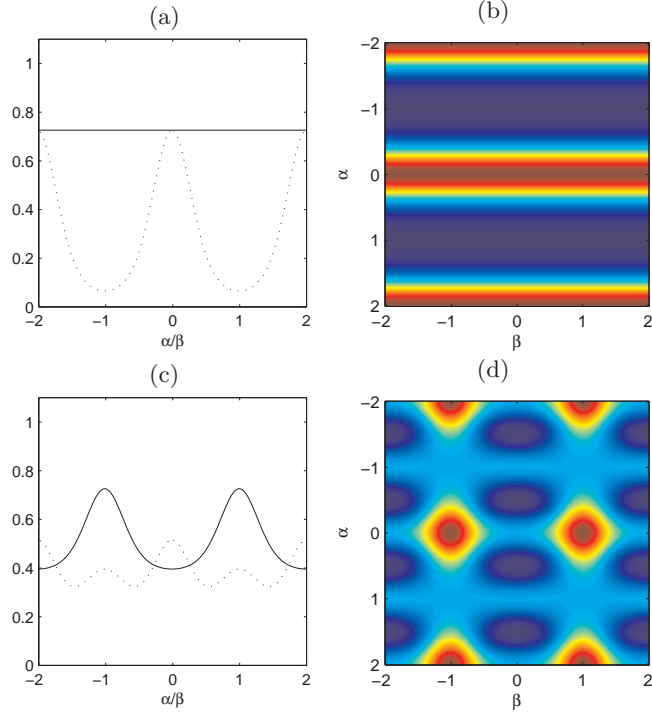


FIG 2. One- and two-dimensional summaries of Gaussian diffusion processes in q -space. **a,b:** Prolate diffusion process ($\lambda_1 \gg \lambda_2 = \lambda_3$). **c,d:** Mixture of two Gaussian diffusion processes. The dominant great circle is the solid line in the one-dimensional summaries **a,c**, while the dotted line is the diffusion from a single perpendicular great circle for **(a)** and the average perpendicular diffusion for **(c)**. In the two-dimensional summaries **(b)** and **(d)** all great circles perpendicular to the dominant great circle are plotted on the y -axis to form the (α, β) plane.

3.2. Degree of Non-Uniformity. We represent a unidirectional Gaussian diffusion by plotting the value of $\mathcal{A}(\mathbf{q}(\beta))$ (solid line) for $\beta \in [-2, 2]$ in Fig. 2a. The magnitude on the dominant great circle is constant over different values of β since $\lambda_2 = \lambda_3$. To illustrate the difference in variation across the dominant and perpendicular great circles we also plot the value of $\mathcal{A}(\mathbf{q}_\perp(\alpha, \beta))$ as a function of α for a fixed β (dotted line). This line perfectly overlaps $\mathcal{A}(\mathbf{q}(\beta))$ at two locations, as it collides with the dominant

great circle when it wraps around the sphere, and decays symmetrically from $\mathbf{q}(\beta)$.

We define a new coordinate system (α, β) , where we expect consistent variability in α and β , using our parameterization of great circles (3.1). We plot the unidirectional Gaussian diffusion $\mathcal{A}(\mathbf{q}_\perp(\alpha, \beta))$ for all perpendicular great circles in the plane (Fig. 2b). This prolate diffusion exhibits variation only in α , which is variation perpendicular to the dominant great circle. For the prolate diffusion example we can therefore reduce the variance by averaging across β and by considering the function strictly in terms of α . We also plot one-dimensional great-circle summaries for a mixture of two Gaussian diffusions in Fig. 2c, where the dominant great circle exhibits a large dynamic range relative to the perpendicular great circles. In fact, one can determine the number of peaks of the diffusion PDF by comparing the dynamic range of the diffusion between the dominant and perpendicular great circles. For a complete picture we also represent the multi-modal diffusion in the (α, β) plane in Fig. 2d, where variation is appreciable in both the α and β axes.

To overcome the need to compare the variation along the dominant great circle with all perpendicular great circles individually, we define the *average perpendicular diffusion* via

$$(3.4) \quad \mathcal{A}_\perp(\alpha) = \frac{1}{4} \int_{-2}^2 \mathcal{A}(\mathbf{q}_\perp(\alpha, \beta)) d\beta.$$

One may also define the average perpendicular diffusion over a half circle by pre-specifying a fixed location on the dominant great circle and integrating in a window size ± 1 around this location. This will prevent certain features being masked by the Hermitian symmetry of the q -space measurements. If $\mathcal{A}(\mathbf{q})$ satisfies (2.2), then we have

$$(3.5) \quad \mathcal{A}_\perp(\alpha) = \frac{1}{4} \int_{-2}^2 B \left(\sqrt{\lambda_1 \alpha^2 + (1 - \alpha^2) [\lambda_2 \|\mathbf{q}(\beta)^T \mathbf{v}_2\|^2 + \lambda_3 \|\mathbf{q}(\beta)^T \mathbf{v}_3\|^2]} \right) d\beta.$$

Thus, we are averaging the density function over small circles parallel to the dominant great circle and $\mathcal{A}_\perp(\alpha)$ measures the average diffusion at a given value of α . In the special case of $\lambda_2 = \lambda_3$, then

$$(3.6) \quad \mathcal{A}_\perp(\alpha) = \frac{1}{4} \int_{-2}^2 B \left(\sqrt{\lambda_1 \alpha^2 + \lambda_2 [1 - \alpha^2]} \right) d\beta = B \left(\sqrt{\lambda_1 \alpha^2 + \lambda_2 [1 - \alpha^2]} \right).$$

The average perpendicular diffusion $\mathcal{A}_\perp(\alpha)$ provides a useful summary of variation perpendicular to the dominant great circle. We define a summary

of the PDF via

$$(3.7) \quad \tau = \left[\frac{\max_{\alpha} \{\mathcal{A}_{\perp}(\alpha)\}}{\min_{\alpha} \{\mathcal{A}_{\perp}(\alpha)\}} \right] \bigg/ \left[\frac{\max_{\beta} \{\mathcal{A}(\mathbf{q}_{\perp}(0, \beta))\}}{\min_{\beta} \{\mathcal{A}(\mathbf{q}_{\perp}(0, \beta))\}} \right] - 1.$$

If the diffusion is isotropic we find that $\lambda_1 = \lambda_2 = \lambda_3$. In this case we have $\mathcal{A}_{\perp}(\alpha_{\max}) = \mathcal{A}_{\perp}(\alpha_{\min}) = B(\sqrt{\lambda_1})$ and $\mathcal{A}(\mathbf{q}(0, \beta_{\max})) = \mathcal{A}(\mathbf{q}(0, \beta_{\min})) = B(\sqrt{\lambda_1})$, and $\tau = 0$. If the diffusion is ellipsoidal and $\lambda_2 = \lambda_3$ then $\tau = B(\sqrt{\lambda_2})/B(\sqrt{\lambda_1}) - 1 > 0$. If we adopt the mixture model with multiple peaks then it is possible to get $\tau \gg 0$ even if we do not have a single diffusion PDF, and hence we define

$$(3.8) \quad \tilde{\tau} = \min_{\beta} \max_{\alpha_1, \alpha_2} \left\{ \frac{\mathcal{A}(\mathbf{q}_{\perp}(\alpha_1, \beta))}{\mathcal{A}(\mathbf{q}_{\perp}(\alpha_2, \beta))} \right\} \bigg/ \left[\frac{\mathcal{A}(\mathbf{q}_{\perp}(0, \beta_{\max}))}{\mathcal{A}(\mathbf{q}_{\perp}(0, \beta_{\min}))} \right] - 1.$$

We note that under isotropy $\tilde{\tau} \equiv 0$ while if we have a single ellipsoid diffusion $\tilde{\tau} \equiv \tau > 0$. For a double tensor $\tilde{\tau}$ is more robust and will (in general) take on a lower value than that τ takes. In contrast with τ and $\tilde{\tau}$ we could also study the variability in the q -space density directly in terms of the ODF. [Tuch \(2004\)](#) for examples defines the generalized fractional anisotropy (GFA) by

$$(3.9) \quad \text{GFA} = \left\{ \frac{n \sum_{i=1}^n (\text{ODF}_W(\theta_i, \phi_i) - \frac{1}{n})^2}{(n-1) \sum_{i=1}^n \text{ODF}_W^2(\theta_i, \phi_i)} \right\}^{1/2},$$

and this measures the non-uniformity of the spatial distribution, as do also the normalized entropy and the nematic order parameter [Tuch \(2004\)](#). While the GFA quantifies the lack of uniformity in the ODF, if there are more fibers than one, determining its statistical properties is non-trivial, unlike the case for τ and $\tilde{\tau}$. Another such measure, generalized anisotropy is defined in terms of the generalized trace of the tensor representation of the mean diffusivity ([Özarslan, Vemuri and Mareci, 2005](#)).

3.3. Measures of Anisotropy. To determine the importance of the identified dominant great circle (or orientation) we can, with a model of (2.2), compare $B(\sqrt{\lambda_1})$ to $B(\sqrt{\lambda_2})$ and $B(\sqrt{\lambda_3})$. We define the following *anisotropy statistic* to perform such a comparison

$$(3.10) \quad \xi = \frac{\log[\mathcal{A}_{\perp}(0)]}{\log[\mathcal{A}_{\perp}(1)]} = \frac{\log[B(\sqrt{\lambda_2})]}{\log[B(\sqrt{\lambda_1})]},$$

where the last equality follows if $\lambda_3 = \lambda_2$. This statistic measures the degree of anisotropy over the q -space shell by comparing the peak-to-trough

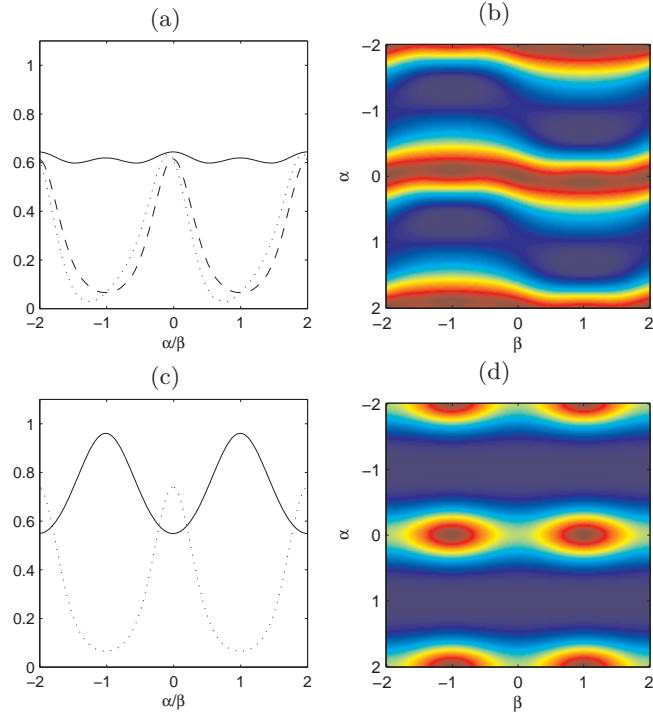


FIG 3. One- and two-dimensional summaries of Gaussian diffusion processes in q -space. An asymmetric diffusion process is displayed in **a** and **b** (this is apparent by the asymmetric decay in great circles perpendicular to the dominant great circle in the (α, β) plane). A scalene diffusion is displayed in **c** and **d**. The dominant great circle is the solid line in the one-dimensional summaries **a, c**, while the dotted line is the average perpendicular diffusion over $\beta \in [-2, 0]$ for **a** and all β 's for **c**. The dashed line in (a) gives the average over all β 's. In the two-dimensional summaries all great circles perpendicular to the dominant great circle are plotted on the y -axis to form the (α, β) plane.

values (i.e. the value at the maximum great circle, compared to the value at the single point perpendicular to that maximum). As an example Fig. 2a displays the difference between the maximum and minimum for an average perpendicular great circle.

The *decay ratio statistic* quantifies the variability of the diffusion over the dominant great circle

$$(3.11) \quad \zeta = \max_{\beta} \frac{\log[\mathcal{A}(\mathbf{q}(\beta))]}{\log[\mathcal{A}(\mathbf{q}(\beta+1))]}.$$

When the two smaller eigenvalues (λ_2 and λ_3) are approximately equal then $\zeta \approx 1$, otherwise $\zeta \gg 1$. The scalene diffusion in Figs. 3c and 3d exhibits such structure; i.e., $\zeta \gg 1$.

An indication of forking in white matter would correspond to an asymmetric decay of the diffusion PDF associated with different decays depending on the parity of the deviation. In this case we may no longer model the diffusion PDF as ellipsoidal. For example, in Figs. 3a and 3b we see that while there is still a strong orientation from the dominant great circle, the PDF no longer exhibits symmetric decay away from the dominant great circle. However, note that the decay is symmetric in α when averaged over the full sphere to produce $\mathcal{A}_{\perp}(\alpha)$.

Averaging over β is no longer appropriate if we want to detect asymmetry since a symmetric distribution will be obtained from the Hermitian symmetry of the HARDI measurements when averaging over a full great circle. Instead we define a suitable *asymmetry statistic* to measure potential asymmetry and we define

$$(3.12) \quad \kappa(\beta) = \frac{2 \int_0^1 [\mathcal{A}(\mathbf{q}_{\perp}(\alpha, \beta)) - \mathcal{A}(\mathbf{q}_{\perp}(-\alpha, \beta))] d\alpha}{\frac{1}{4} \int_0^4 \mathcal{A}(\mathbf{q}_{\perp}(\alpha, \beta)) d\alpha},$$

$$(3.13) \quad \beta_{\max} = \arg \max \kappa(\beta),$$

$$(3.14) \quad \kappa = \int_{\beta_{\max}-1/2}^{\beta_{\max}+1/2} \kappa(\beta) d\beta.$$

The definition of κ is motivated by the wish to both obtain a test statistic with sufficient power and also to reduce its variance. The discrete approximation to κ will have smaller variance than $\kappa(\beta_{\max})$ has. Asymmetry in the decay from the main peak may occur when the PDF is a mixture of diffusions with varying strengths. If the two populations are sufficiently separated and equivalent in magnitude then this will be indicated by τ and/or $\tilde{\tau}$, and the diffusion will be recognized as a “crossing fiber”. If the mixture of diffusions contains two different strengths, then the dominating PDF will be

recognized when determining \mathbf{x}_{\max} . The remaining structure will not be fully consistent with a single tensor and will (in general) appear to be asymmetric compared to the dominant great circle.

TABLE 1
The structure of the proposed models. Key to abbreviation N-P/A (Non-Preference/Anisotropy), C/E (Circular/Ellipsoidal), S/A (Symmetric/Asymmetric), I/M (Isotropic/Multimodal), M/U (Multimodal/Unimodal).

Hypothesis	Statistic	isotropic	prolate	scalene	mixture	heterogeneous
N-P/A	τ	small	large	large	small	large
M/U	$\tilde{\tau}$	small	large	large	small	large
I/M	ξ	one	small	small	small	small
C/E	ζ	-	one	large	-	large
S/A	κ	-	zero	zero	-	large

Let us discuss models that will lead to different structure in the proposed summaries. We refer to table 1 to summarize the properties of the summaries, and different PDFs lead to different structure. It may seem insufficient to consider only an isotropic PDF, a single peak, a double peak, or something more heterogeneous. However, even with a fully parametric model with a Gaussian DTI framework, a two tensors model has 13 (identifiable) parameters and a three tensor model has 19. If one considers acquiring 60 gradient encoding directions, then one is forced to fit a quite saturated model resulting in very noisy estimates – especially at higher b -values where the orientational heterogeneity can be well resolved. Pushing much beyond a small number of parameters or features of interest is not advisable with such sampling.

4. Estimation.

4.1. *Parameterizing the (α, β) axes.* Having proposed various summaries of the population of PDFs at a particular voxel, these must now be estimated from a set of diffusion measurements. The dominant direction is estimated by

(4.1)

$$\hat{v}_1 = \mathbf{x}_{\max} = \arg \max_{\mathbf{v}, \|\mathbf{v}\|=1} \left\{ \int_{\mathbf{q} \in \mathcal{G}(\mathbf{v})} \hat{\mathcal{A}}(\mathbf{q}) \, d\mathbf{q} \right\} \equiv \arg \max_{\mathbf{x}} \text{FRT}\{\hat{\mathcal{A}}\}(\mathbf{x}),$$

where $\text{FRT}\{\cdot\}$ denotes the Funk-Radon Transform (FRT) as utilized in Tuch (2004). Note that $\hat{\mathcal{A}}(\mathbf{q})$ refers to the multiresolution-based estimator in Olhede and Whitcher (2008a). It may be replaced by another appropriate estimator. We assume the availability of $(\hat{\sigma}^*)^2$ an estimator of the variance of the error in $A(\mathbf{q}_k)$, namely σ^2 . The variance of $\hat{\mathcal{A}}(\mathbf{q}_k)$ is assumed

to be $\tilde{\sigma}^2 \leq \sigma^2$ and the variance of an interpolated value of the PDF is $\tilde{\sigma}^2 \leq \tilde{\sigma}^2 \leq \sigma^2$. The integral may be approximated numerically by interpolating the observed HARDI measurements at evenly-spaced points along several great circles, each perpendicular to some \mathbf{x}_i .

The effects of using different numerical methods for this step is a trade-off between increasing numerical accuracy and decreasing variance. Interpolating using spherical harmonics reduces variance, but can smooth out details depending on the choice of regularization, see also the discussion in Descoteaux *et al.* (2007) and Hess *et al.* (2006). We instead use a locally linear interpolation on the polar representation of the data, specifically enforcing the periodicity of the data. Simple structures in terms of the observed points, can mix over several spherical harmonics, and so the magnitude of individual spherical harmonic coefficients may not be large, even if the local coefficient is, thus making the representation inappropriate for using the smoothing methods of previous authors. The choice of interpolation procedure should be considered in terms of which statistic one is using, as the variance and bias must be balanced specifically for *this* purpose. We also note that spherical harmonics do not possess the same properties as Fourier vectors, and that an infinite number of harmonics are required for perfect reconstruction of a surface on a sphere, and so any reconstruction from the continuous basis will be inaccurate.

The spatial maximum is determined from $\{\text{FRT}\{\mathcal{A}\}(\mathbf{x}_i)\}_i$. The spatial location \mathbf{x}_{\max} is an *estimator* of \mathbf{v}_1 and we estimate a vector in the linear subspace spanned by \mathbf{v}_2 and \mathbf{v}_3 from the eigensystem of $\mathbf{I} - \mathbf{x}_{\max}\mathbf{x}_{\max}^T$, this yielding $\hat{\mathbf{v}}_2$ and $\hat{\mathbf{v}}_3$, that maximise the difference in decay in the two axes. For numerical implementation we sample the estimated dominant great circle by discretizing α to $\{\alpha_j\}_{j=1}^N$ and β to $\{\beta_k\}_{k=1}^N$, for an even integer N . A discretized version of (2.11) is then given by

$$(4.2) \quad \hat{\mathbf{q}}_k = \begin{cases} \beta_k \hat{\mathbf{v}}_2 + \sqrt{1 - \beta_k^2} \hat{\mathbf{v}}_3, & \beta_k = 4k/N - 1 \quad \text{for} \quad k = 1, \dots, N/2; \\ \beta_k \hat{\mathbf{v}}_2 - \sqrt{1 - \beta_k^2} \hat{\mathbf{v}}_3, & \beta_k = 4k/N - 3 \quad \text{for} \quad k = N/2 + 1, \dots, N. \end{cases}$$

This then allows us to discretize the estimated dominant great circle via $\{\hat{\mathbf{q}}_k\}$. Once \mathbf{x}_{\max} has been determined we characterize the diffusion directly in q -space. We introduce additional notation by defining the sampled great circle vectors for $\{\hat{\mathbf{q}}_k\}$ in (4.2) as

$$(4.3) \quad \hat{\mathbf{q}}_{\perp jk} = \begin{cases} \alpha_j \hat{\mathbf{v}}_1 + \sqrt{1 - \alpha_j^2} \hat{\mathbf{q}}_k, & \alpha_j = 4j/N - 1 \quad \text{for} \quad j = 1, \dots, N/2; \\ \alpha_j \hat{\mathbf{v}}_1 - \sqrt{1 - \alpha_j^2} \hat{\mathbf{q}}_k, & \alpha_j = 4j/N - 3 \quad \text{for} \quad j = N/2 + 1, \dots, N. \end{cases}$$

A discretized version of the average perpendicular diffusion (3.4) is given by $\hat{\mathcal{A}}_{\perp}(\alpha_j) = \frac{1}{N} \sum_k \hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp jk})$.

4.2. Diagnosing Non-Uniformity. In order to test large-scale properties of the diffusion directly in q -space, we consider the following statistical hypothesis $H_0 : \mathcal{A}(\mathbf{q}) = \mathcal{A} \forall \mathbf{q}$ versus $H_1 : \mathcal{A}(\mathbf{q}) = \mathcal{A}_E(\mathbf{q})$. Our test statistic is based on a discretized version of (3.7), given by

$$(4.4) \quad T = \left[\frac{\max_j \{\hat{\mathcal{A}}_{\perp}(\alpha_j)\}}{\min_j \{\mathcal{A}_{\perp}(\alpha_j)\}} \right] \bigg/ \left[\frac{\max_k \{\mathcal{A}(\hat{\mathbf{q}}_k)\}}{\min_k \{\mathcal{A}(\hat{\mathbf{q}}_k)\}} \right] - 1.$$

The distribution of this test statistic is derived in Appendix A.1. If the observations are isotropic, then the properties along the dominant great circle will be equivalent to the properties on the perpendicular great circle (barring random/discretization errors). We therefore use, with $\hat{\mathbf{q}}_k$ defined by (4.2), as estimators for \mathcal{A} and $\tilde{\sigma}$ under the null of $\hat{\mathcal{A}}(\hat{\mathbf{q}}_k) \cong \mathcal{A} + \tilde{\sigma}\epsilon$,
(4.5)

$$\overline{\mathcal{A}}_N = \frac{1}{N} \sum_{k=1}^N \hat{\mathcal{A}}(\hat{\mathbf{q}}_k) \quad \text{and} \quad \hat{\sigma}_{\mathcal{A}} = \sqrt{\rho} \text{MAD} \left\{ \hat{\mathcal{A}}(\hat{\mathbf{q}}_k) - \overline{\mathcal{A}}_N : k = 1, \dots, N \right\},$$

where $0 < \rho \leq 3$. These equations provide estimators of the mean value of the isotropic diffusion and the standard deviation of $\hat{\mathcal{A}}(\mathbf{q})$ at the observed measurements. The parameter ρ is a constant depending on the linear interpolation method used for the implementation. Taking a value of $\rho = 3$ is suitable for our choice of numerical interpolation and we define

$$(4.6) \quad U = T \frac{\overline{\mathcal{A}}_N}{\hat{\sigma}_{\mathcal{A}}}.$$

We can compute the critical value u_{α} using the fact that $F_U(u_{\alpha}) = 1 - \alpha$, where $F_U(\cdot)$ is given by (A.5). We report two critical values here, $u_{0.05} = 0.1185$ for the m which is consistent with our sampling scheme, and $u_{0.05}^{(\text{con})} = 1.9637$ with a conservative distribution approximation.

We also develop a new test based on a null of a multi-modal diffusion, where we define multi-modal in terms of $(\tilde{\mathcal{A}}_{\max} \mathcal{A}_{\min}) / (\tilde{\mathcal{A}}_{\min} \mathcal{A}_{\max}) < c = 2$, say, where $\tilde{\mathcal{A}}_{\max}$ and $\tilde{\mathcal{A}}_{\min}$ are the maximum and minimum on the perpendicular great circle minimizing (3.8) in β , respectively, while \mathcal{A}_{\max} and \mathcal{A}_{\min} are the maximum and minimum on the dominant great circle. The level c is fairly arbitrary, but to develop a powerful method of separating the clearly unimodal from the multi-modal some level must be chosen, based on the

deterministic structure of the sampled PDF. To separate the unimodal from the multimodal we start from $\tilde{\tau}$ and define

$$(4.7) \quad \tilde{T} = \min_k \left\{ \max_{j_1, j_2} \left\{ \frac{\hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp j_1 k})}{\hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp j_2 k})} \right\} \right\} / \left[\frac{\max_k \{\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)\}}{\min_k \{\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)\}} \right] - 1.$$

We shall now choose to distinguish the multimodal from the unimodal, and so normalize by $\tilde{U} = (\tilde{T} - [c - 1])\hat{\mathcal{A}}_{\min}/(\hat{\sigma}_{\mathcal{A}}\sqrt{2c^2 + 2})$, where $\hat{\mathcal{A}}_{\min} = \hat{\mathcal{A}}(\hat{\mathbf{v}}_1)$. The distribution of this test statistic is derived in Appendix A.1, under the specified null hypothesis.

If on the other hand we have failed to reject the null hypothesis “ $\mathcal{A}(\mathbf{q}_{\perp}(\alpha, \beta))$ equally variable in β for $\alpha = 0$ as it is in α for a fixed β ,” then based on the T -statistic we need to distinguish voxels that indicate two fiber populations versus isotropic voxels. We define a discrete version of (3.10) to be

$$(4.8) \quad X = \frac{\log[\hat{\mathcal{A}}_{\perp}(0)]}{\log[\hat{\mathcal{A}}_{\perp}(1)]}.$$

We can interpret ξ , and the sample version X , as the degree of anisotropy from the average perpendicular great circle. The ADC is defined at a set of defined \mathbf{q} vectors $\tilde{\mathbf{q}}_j$ as $\hat{C}(\tilde{\mathbf{q}}_j) = -b^{-1} \log[A(\tilde{\mathbf{q}}_j)]$ (Alexander, Barker and Arridge, 2002). We recognize that the statistic X is comparing the average ADC on the great circle to the average ADC perpendicular to the great circle, or that (4.8) may be re-written in terms of the ADC at a fixed value of b via

$$(4.9) \quad X = \sum_k \hat{C}(\hat{\mathbf{q}}_k) / \sum_k \hat{C}(\hat{\mathbf{q}}_{\perp N/4k}).$$

With an assumption of ellipsoidal structure, cf (2.2), we have averaged the ADC to reduce variance when estimating ξ without accruing bias. We define X_k as the sample anisotropy calculated using only the k th perpendicular great circle, or $X_k = \log[\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)] / \log[\hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp N/4k})]$, and refer to (4.3) for the definition of $\hat{\mathbf{q}}_{\perp jk}$. Under moderate-to-high SNR we may approximate this as a Gaussian random variable. We quantify uncertainty, when there are potentially several peaks, using $\hat{\sigma}_2 = \min\{\hat{\sigma}_{\mathcal{A}}, \hat{\sigma}^*\}$, where $\hat{\sigma}_{\mathcal{A}}$ is defined in (4.5) and $\hat{\sigma}^*$ is the available estimator for σ . By using the minimum we ensure that the variance is never estimated as greater than before smoothing.

For those voxels where we do not reject isotropy, we may now distinguish between isotropy and a multiple-tensor model using the X . We define

$$(4.10) \quad Q = \frac{\rho(X - 1)}{\hat{\sigma}_2} |\overline{\mathcal{A}}_N \log \overline{\mathcal{A}}_N|$$

as the test statistic for multimodality. So we consider the test $H_0 : \mathcal{A}(\mathbf{q}) = \mathcal{A} \forall \mathbf{q}$ versus $H_1 : \max_k \{\mathcal{A}(\mathbf{q}_k)\} \gg \min_k \{\mathcal{A}(\mathbf{q}_k)\}$ (multiple peaks) and use Q as the test statistic, whose distribution under the null is found in Appendix A.2. These three tests allow us to at a single voxel diagnose the structure of the PDF, where U is used to separate anisotropic PDFs from isotropic PDFs, \tilde{U} is used to separate ellipsoid PDFs from multimodal PDFs and Q is used to separate multimodal from isotropic PDFs.

4.3. Diagnosing Asymmetry. Having established methodology to discriminate the number of peaks in the diffusion PDF at a single voxel, we now provide additional methodology to characterize the diffusion PDF as scalene versus other forms of asymmetry; e.g., to observe an indication of forking. We start by determining

$$(4.11) \quad k_{\max} = \arg \max_{1 \leq k \leq N/4} \frac{\log \hat{\mathcal{A}}(\hat{\mathbf{q}}_k)}{\log \hat{\mathcal{A}}(\hat{\mathbf{q}}_{k+N/4})}.$$

We define the parameters $m < m' < 2m$ (recall that m is the effective degrees of freedom) for robustness so that

$$(4.12) \quad Z = \frac{\log \hat{\mathcal{A}}(\hat{\mathbf{q}}_{k_{\max}+N/(2m')})}{\log \hat{\mathcal{A}}(\hat{\mathbf{q}}_{k_{\max}+N/(2m')+N/4})}.$$

We remark that Z is related in some sense to X_k (refer to Figs. 2 and 3); X_k compares the value of the diffusion on location k on the dominant great circle ($\alpha = 0$) to the value at the perpendicular to the dominant great circle ($\alpha \neq 0$), whereas Z in contrast looks at the difference in values of the diffusion on the great circle itself ($\alpha = 0$ and β varies). Under the null hypothesis $\mathcal{A}(\cdot)$ is constant on the great circle, and if the medium and minor eigenvalues are approximately equal then $E\{Z\} = \zeta \approx 1$, otherwise $\zeta \gg 1$. We define a normalized version of the decay ratio statistic (4.12), given by

$$(4.13) \quad V = \frac{(Z - 1) |\overline{\mathcal{A}}_N \log \overline{\mathcal{A}}_N|}{\hat{\sigma}_2}.$$

A suitable threshold for this statistic may be found in Appendix A.3. The statistics Q and V used to test different hypotheses of non-isotropic decay, have similar forms.

The summary κ allows us to diagnose structure inconsistent with a single ellipsoid diffusion. Departures from a single ellipsoid diffusion structure could be attributable to partial volume effects, or a heterogeneous population of fibers. For such a model, (2.2) is no longer appropriate and we would

rather fit a mixture model with unequal populations; or possibly $\mathcal{A}_{DE}(\cdot)$. In such circumstances one cannot use the average perpendicular great circle to uncover asymmetry since averaging over all possible β 's will produce a symmetric distribution regardless of the underlying fiber characteristics. Taking $\check{k}_{\max} = \arg \max_k P_k$ we define the *asymmetry statistic* by

$$(4.14) \quad K = \frac{1}{N/4 + 1} \sum_{k=\check{k}_{\max}-N/8}^{\check{k}_{\max}+N/8} P_k, \quad P_k = \frac{8 \sum_{j=1}^{N/4-1} [\hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp jk}) - \hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp (j+N/4)k})]}{\sum_{j=1}^N \hat{\mathcal{A}}(\hat{\mathbf{q}}_{\perp jk})},$$

Full details on the distribution of this test statistic may be found in Appendix A.4. We have chosen $N/8$ to improve the power – averaging decreases the variance, but the asymmetry is greatest near the maximum (compare with Figure 3(b)). For tests at a specific voxel we perform the hypothesis test $H_0 : \kappa = 0$ versus $H_1 : \kappa \neq 0$, using quantiles from the standard Gaussian PDF $\phi(\cdot)$. This identifies diffusion PDFs that are non-Gaussian in terms of the parity structure in the principal axes. However, it does not compare the maximum and minimum of a perpendicular great circle, rather it finds a set of perpendicular great circles for which the decay around the dominant great circle is asymmetric and estimates this average asymmetry; e.g., Figure 3(b).

The usage of the statistics is now combined at a voxel level. The most important characteristic is to classify the voxel as isotropic, single peak, or multiple peak. With this information the local structure of the peaks can be further characterised at a local level: when comparing the PDFs between voxels for tracking first the number of voxels is important, and after the mixture components is matched to the most appropriate component from its local characteristics, we can anticipate varying asymmetry values before forking structure; as shall be further discussed in the applications section.

4.4. Example: Crossing and Forking Fibers. We consider two typical heterogeneous white-matter structures, a crossing fiber and a forking fiber in Fig. 4. The spatial representation of the forking fiber is the first row, enumerated by (i), while the q -space representation of the forking fiber is on the second row, enumerated by (ii). In the spatial representation, we see a single fiber population in voxel (i,a) and as we traverse from left-to-right the two populations become more apparent until a second fiber appears in voxel (i,g). The q -space version of these two populations shows a scalene distribution developing in voxels (ii,b)–(ii,e). As the forking progresses from left-to-right it appears both highly warped and scalene until the distribution clearly displays multiple fibers in voxel (iii,g). These fibers were generated

by aggregating two densities or

(4.15)

$$\mathcal{A}(\mathbf{q}) = a_1(t)\mathcal{A}_g(\mathbf{q}; \mathbf{\Lambda}^{(1)}(t), \mathbf{\Upsilon}^{(1)}(t)) + (1 - a_1(t))\mathcal{A}_g(\mathbf{q}; \mathbf{\Lambda}^{(2)}(t), \mathbf{\Upsilon}^{(2)}(t)).$$

In the case of the forking fiber $a_1(t) = 1 - t$ and the angle between the principal direction of $\mathbf{\Upsilon}^{(1)}(t)$ and $\mathbf{\Upsilon}^{(2)}(t)$ $\pi t/2$, with the individual tensors taking values similar to $\mathcal{A}_1(\mathbf{q})$ and $t \in [0, 1]$. The spatial representation of the crossing fiber is the third row (iii), with its corresponding q -space representation in the fourth row (iv). The ellipsoid appears prolate in voxel (iii,a), then two fiber populations are present in voxel (iii,d) and eventually the fiber population returns to a prolate shape. With respect to the parameterization of the crossing fiber in (4.15), $a_1(t) \in \{0, 0.25, 0.5\}$ and the two fibers cross at 90 degrees with parameters similar to $\mathcal{A}_1(\mathbf{q})$.

The description of the crossing fiber is in many ways simpler than the forking. Table 2 lists the summary statistics $(\tau, \tilde{\tau}, \xi, \zeta, \kappa)$ for the crossing and forking fiber examples in Fig. 4. Note that these deterministic summaries have not been normalized, unlike the statistics in Section 4 (as there is in this case no noise variance to compare with). The mixture of populations of unequal strength in the forking fiber shows a number of characteristics not found in the crossing fiber. For example, the forking fiber is clearly diagnosed as a single population until voxels (i,e), (i,f) and (i,g), where there is increasing heterogeneity in the fiber population. This is exhibited by increasing values for the decay ratio ζ -statistic, and the asymmetry κ -statistic. For the crossing fiber, we clearly detect the multiple-fiber population in voxel (iii,d) using either the τ and $\tilde{\tau}$ statistics. The multiple-fiber characteristics in the forking example are more complex, where the second fiber population is initially dominated by the first. If we examine the crossing fiber more closely there is little apparent asymmetry and we can compare the asymmetry statistic, where $\kappa \approx 0$ versus $0.15 \leq \kappa \leq 0.45$ for voxels (i,c)–(i,e). To distinguish the clear multi-tensors from uniformity we observe that $\xi < 1$, which is the expected value under the hypothesis of isotropy.

5. Simulation Study. We illustrate the properties of the proposed q -space summary statistics for the diffusion PDF on a variety of simulated diffusions processes. The following six models attempt to cover common, and not so common, diffusion processes that include both unimodal and

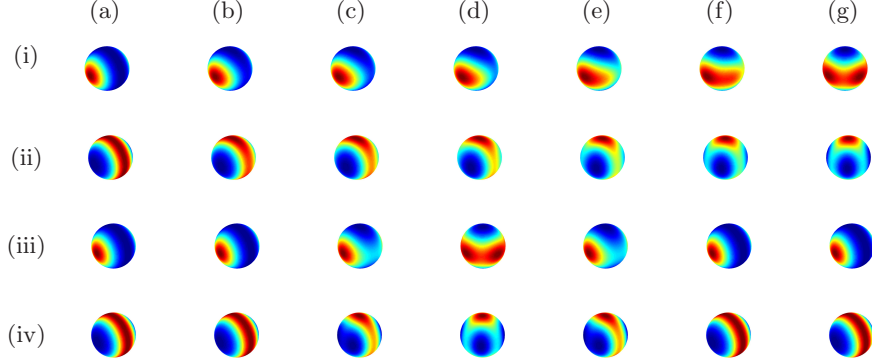


FIG 4. An illustration of the evolution of a diffusion PDF through a number of adjacent voxels. The first and second rows are the spatial and q -space evolution, respectively, of the diffusion PDF for a forking fiber. The third and fourth rows are the spatial and q -space evolution, respectively, of the diffusion PDF for a crossing fiber.

TABLE 2

Discretized statistics based on non-parametric measures of symmetry for a simulated forking and crossing fiber, compare with Table 1.

Statistic	Forking fiber						
	(i,a)	(i,b)	(i,c)	(i,d)	(i,e)	(i,f)	(i,g)
τ	10.18	5.12	3.35	1.86	0.82	0.07	-0.14
$\tilde{\tau}$	8.98	4.94	3.36	2.13	1.19	0.33	-0.07
ξ	0.12	0.18	0.21	0.27	0.35	0.53	0.69
ζ	1.03	1.51	1.72	1.91	2.06	2.21	2.67
κ	-0.01	0.03	0.17	0.35	0.45	0.54	0.30
	Crossing fiber						
	(iii,a)	(iii,b)	(iii,c)	(iii,d)	(iii,e)	(iii,f)	
τ	9.19	9.19	1.15	-0.14	1.15	9.19	
$\tilde{\tau}$	8.98	8.98	1.26	-0.07	1.26	8.98	
ξ	0.12	0.12	0.32	0.69	0.32	0.12	
ζ	1.04	1.04	1.74	2.67	1.74	1.04	
κ	0.00	0.00	0.09	0.30	0.09	0.00	

multiple tensors

$$\begin{aligned}
(5.1) \quad \mathcal{A}_1(\mathbf{q}) &= \exp(-t\mathbf{q}^T \tilde{\mathbf{D}}_i \mathbf{q}), \quad i = 1, 2, 3, \\
\tilde{\mathbf{D}}_1 &= 68\tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T + 8\tilde{\mathbf{e}}_2\tilde{\mathbf{e}}_2^T + 8\tilde{\mathbf{e}}_3\tilde{\mathbf{e}}_3^T, \\
\tilde{\mathbf{D}}_2 &= 68\tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T + 15\tilde{\mathbf{e}}_2\tilde{\mathbf{e}}_2^T + \tilde{\mathbf{e}}_3\tilde{\mathbf{e}}_3^T, \\
\tilde{\mathbf{D}}_3 &= 28\tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T + 28\tilde{\mathbf{e}}_2\tilde{\mathbf{e}}_2^T + 28\tilde{\mathbf{e}}_3\tilde{\mathbf{e}}_3^T, \\
(5.2) \quad \mathcal{A}_4(\mathbf{q}) &= 0.5 \exp(-t\mathbf{q}^T \tilde{\mathbf{D}}_1 \mathbf{q}) + 0.5 \exp(-t\mathbf{q}^T \tilde{\mathbf{D}}_4 \mathbf{q}), \\
\tilde{\mathbf{D}}_4 &= 68\tilde{\mathbf{e}}_2\tilde{\mathbf{e}}_2^T + 8\tilde{\mathbf{e}}_1\tilde{\mathbf{e}}_1^T + 8\tilde{\mathbf{e}}_3\tilde{\mathbf{e}}_3^T, \\
(5.3) \quad \mathcal{A}_5(\mathbf{q}) &= \exp(-11t|\mathbf{q}^T \tilde{\mathbf{e}}_2|^2) \left| \exp(-68t|\mathbf{q}^T \tilde{\mathbf{e}}_1|^2) \right. \\
&\quad \times [\exp(-0.2t|\mathbf{q}^T \tilde{\mathbf{e}}_3|^2) + \exp(-35t|\mathbf{q}^T \tilde{\mathbf{e}}_3|^2)] \\
&\quad \left. + \frac{4}{\pi} D\left(\sqrt{68t}\mathbf{q}^T \tilde{\mathbf{e}}_1\right) \left[D\left(\sqrt{35t}\mathbf{q}^T \tilde{\mathbf{e}}_3\right) - D\left(\sqrt{0.2t}\mathbf{q}^T \tilde{\mathbf{e}}_3\right) \right] \right|, \\
(5.4) \quad \mathcal{A}_6(\mathbf{q}) &= 0.3 \exp(-t\mathbf{q}^T \tilde{\mathbf{D}}_1 \mathbf{q}) + 0.7 \exp(-t\mathbf{q}^T \tilde{\mathbf{D}}_5 \mathbf{q}), \\
\tilde{\mathbf{D}}_5 &= 42.5\check{\mathbf{e}}_1\check{\mathbf{e}}_1^T + 14\check{\mathbf{e}}_2\check{\mathbf{e}}_2^T + 20\check{\mathbf{e}}_3\check{\mathbf{e}}_3^T,
\end{aligned}$$

where $D(x) = \exp(-x^2) \int_0^x \exp(t^2) dt$ is the Dawson function ([Abramowitz and Stegun, 1972](#)). We define $\tilde{\mathbf{e}}_j = \mathcal{R}\mathbf{e}_j$, where the matrix \mathcal{R} rotates the axes $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ to a new coordinate system $(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3)$. This extra step is added to protect against systematic bias in our estimation procedure due to the diffusion PDF coinciding with the sampling grid. In $\mathcal{A}_6(\mathbf{q})$ this rotation is not implemented, but $(\check{\mathbf{e}}_1, \check{\mathbf{e}}_2, \check{\mathbf{e}}_3)$ has been rotated with respect to $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ to produce an asymmetric diffusion in the multiple-tensor model.

We have chosen to define $\tilde{\mathbf{D}}_i = 4 \times 10^{10} \mathbf{D}_i$, $i = 1, \dots, 4$, and normalized $\|\mathbf{q}\| = 1$. With $t = 0.04$ this corresponds to $b = 4t \times 10^{10} = 1600 \text{ s/mm}^2$ ([Alexander, 2005](#)) and the trace of the first three non-normalized matrices \mathbf{D}_i as $2.1 \times 10^{-9} \text{ m}^2/\text{s}$. The function $\mathcal{A}_5(\mathbf{q})$ is obtained from the magnitude of the FT of an asymmetrically decaying diffusion process in space. We illustrate a range of behavior for the scalar statistics defined in q -space using these test functions, providing only a subset in order to compare and contrast their performance. We simulate 1000 realizations for each test function and add Gaussian noise with standard deviation of $\mathcal{A}(0)/2$, $\mathcal{A}(0)/10$, $\mathcal{A}(0)/20$ and $\mathcal{A}(0)/30$ to both the real and imaginary channels using 60 gradient directions.

Results, provided in Table 3, are consistent with varying degrees of the SNR. The prolate diffusion \mathcal{A}_1 is clearly detectable, down to an SNR = 1/10, despite using non-parametric methods via the U -statistic. Detecting the scalene diffusion depends on the SNR, while the isotropic diffusion is clearly distinguishable from its alternatives under the full range of SNR using the U -statistics. The multi-tensor diffusion \mathcal{A}_4 is difficult to classify

TABLE 3

Hypothesis tests for the six diffusion processes $\{\mathcal{A}_i\}_{i=1}^6$. The nominal size of the tests is 5% for U , V and Q , while the nominal size is 10% for K and \tilde{U} . The tests have been carried out at different SNRs. Key to abbreviation N-P/A (Non-Preference/Anisotropy), C/E (Circular/Ellipsoidal), S/A (Symmetric/Asymmetric), I/M (Isotropic/Multimodal), M/U (Multimodal/Unimodal). The SNR decreases the further down the entries are in the table, ranging from $SNR=\frac{1}{30}$, $SNR=\frac{1}{20}$, $SNR=\frac{1}{10}$ to $SNR=\frac{1}{2}$.

Hypothesis	Statistic	\mathcal{A}_1	\mathcal{A}_2	\mathcal{A}_3	\mathcal{A}_4	\mathcal{A}_5	\mathcal{A}_6
H_0/H_1 ,		prolate	scalene	isotropic	crossing	asym.	asym.
N-P/A	U	1000	988	26	492	1000	802
C/E	V	146/1000	924/988	0/26	382/492	120/1000	495/802
S/A	K	191/1000	108/988	10/26	258/492	491 /1000	208/802
I/M	Q	0/0	12/12	21/974	420/508	0/0	195/198
M/U	\tilde{U}	991/1000	136/988	0/26	239/492	996/1000	19/802
H_0/H_1 ,							
N-P/A	U	1000	855	26	484	1000	727
C/E	V	153/1000	794/855	0/26	338/484	148/1000	267/727
S/	K	148/1000	38/855	10/26	199/484	331/1000	136/727
I/M	Q	0/0	0/145	23/974	259/516	0/0	201/273
M/U	\tilde{U}	945/1000	46/805	0/23	151/484	942/1000	8/727
H_0/H_1 ,							
N-P/A	U	1000	239	34	441	998	457
C/E	V	225/1000	214/239	0/34	192/441	194/998	58/457
S/A	K	89/1000	1/239	10/34	99/441	139/998	74/457
I/M	Q	0/0	449/761	20/966	20/539	1/2	18/543
M/U	\tilde{U}	239/1000	7/239	0/34	27/441	174/998	4/457
H_0/H_1 ,							
N-P/A	U	45	21	25	31	37	24
C/E	V	2/45	2/21	5/25	6/31	4/37	4/24
S/A	K	5/45	4/21	2/25	1/31	4/37	3/24
I/M	Q	1/955	3/979	2/975	2/969	4/963	5/976
M/U	\tilde{U}	2/45	0/21	1/25	0/31	5/37	1/24

using the U -statistics and its correct classification depends on how well the location of the dominant peak is estimated. If the dominant peak is well determined then the U -statistic clearly recognizes the density as anisotropic, if not the q -space measurements are characterized as non-Gaussian instead of multimodal. If one was only concerned with empirically separating prolate diffusion PDFs from multimodal diffusion PDFs, rather than performing a hypothesis test, then this would be relatively straightforward; e.g., retaining 95% of the unimodal Gaussian with the $\text{SNR} = 1/20$ leads to rejecting all but 11% of the multi-tensor realizations (see the \tilde{U} -statistic). Since we are interested in detecting ellipsoidal decay around a single direction, the variation over the dominant great circle will be large for anisotropic voxels with ellipsoidal decay as well as for multi-modal diffusion PDFs. At an $\text{SNR} = 1/20$ the \tilde{U} -statistic provides complimentary information by strongly separating the prolate (94% rejected) from the multi-tensor model (15.1% rejected, near the nominal value of 10%), but fails to distinguish between the scalene and the multi-tensor models (Table 3). The highly scalene density is instead seen as multimodal, and such structure may be approximated using two tensors, especially when sparsely sampled on the sphere.

The two distributions with constant behavior on the dominant great circle are not diagnosed with asymmetric decay, while the null hypothesis is rejected for \mathcal{A}_2 in a substantial number of cases in Table 3. The misdiagnosed multimodal diffusion PDF \mathcal{A}_4 also has the null hypothesis of multimodality rejected for a substantial number of cases. This is to be expected since the observed diffusion will experience considerable variation over the dominant great circle, consistent with observing a diffusion process with a single dominant direction and ellipsoidal decay.

We fail to reject the null hypothesis of symmetry for the two diffusion processes that are symmetric (\mathcal{A}_1 and \mathcal{A}_2) in most cases, while we reject a larger proportion for \mathcal{A}_5 . There is unfortunately a lack of power in this test which is due to sampling 60 directions; limiting the performance of the test statistic. For \mathcal{A}_6 and keeping the $\text{SNR} = 1/20$, we reject the null hypothesis 38.9% with 60 directions. For \mathcal{A}_5 we reject the null 51.6% of the time using 245 directions at $\text{SNR} = 1/20$ – a clear increase from 35.2% with 60 directions. Increasing the SNR also increases our power to detect such asymmetry, as shown in Table 3. The power of the test improves as the number of directions increase or the amount of asymmetry (better characterised with better spherical sampling) increases.

6. Analysis of Clinical Data. HARDI data were acquired from one normal subject (30 year old, male Caucasian) in a Siemens TIM Trio 3.0 Tesla

scanner using a 32-channel head coil. Regions of interest (ROIs) from two slices of the clinical data are provided to illustrate the statistical summaries developed in this paper. Slice 1 contains an ROI that is dominated by single-fiber voxels containing structures such as the corpus callosum and cingulum. Figure 5(a) shows the voxels using the common color-coding convention (i.e., RGB for the (x, y, z) coordinates) weighted by the estimated fractional anisotropy (FA) at each voxel. The FA for the ROI is reproduced in Figure 5(b) along with the p -values for the anisotropy and ellipsoidal-ity statistics in Figures 5(c) and 5(d), respectively. We select a very liberal threshold ($p = 0.15$) for the purpose of exploratory data analysis, not confirmatory data analysis. We observe very few voxels that indicate asymmetry at specific voxels, while the ellipsoidality p -values indicate quite a few voxels that exhibit ellipsoidal structure. These voxels are located at the borders of strongly directional structures such as the corpus callosum and cingulum, and reaffirms the results obtained in the simulation studies. Additional information about the structure is obtained by plotting the test statistic for unimodality and the p -values from the multi-modality test statistic in Figures 5(e) and 5(f), respectively. The corpus callosum, and to a lesser extent the cingulum, produce large values in the unimodality test statistic as to be expected from those structures. Multimodality is detected in voxels with reduced FA and/or on the edges of prominent white-matter structures. The pattern of multimodal voxels identified in Figure 5(f) in general do not appear to overlap with those voxels that were identified using the ellipsoidality statistic, providing evidence that this methodology is detecting distinct features in the white-matter microstructure.

The ROI selected in slice 2 captures more complicated interactions between white-matter structures such as the corticopontine tract, anterior thalamic radiation and corpus callosum; Figure 6(a). The FA for the ROI is reproduced in Figure 6(b) along with the p -values for the anisotropy and ellipsoidality statistics in Figures 6(c) and 6(d), respectively. Asymmetry is difficult to detect in these data, but ellipsoidality is quite apparent along the boundaries of the corpus callosum and around the projections into grey matter. The test statistic for unimodality in Figure 6(e) complements the ellipsoidality results quite well, picking out dominant unimodal structures (e.g., the voxels dominated by the corpus callosum and to a lesser degree the cingulum) around which the ellipsoidality measure is finding more complex voxels. Finally, the test statistic for multimodality in Figure 6(f) clearly identifies voxels where the three dominant white-matter structures in this ROI converge, and all other statistics fail to detect any specific structure. The statistical summaries developed here provide complementary informa-

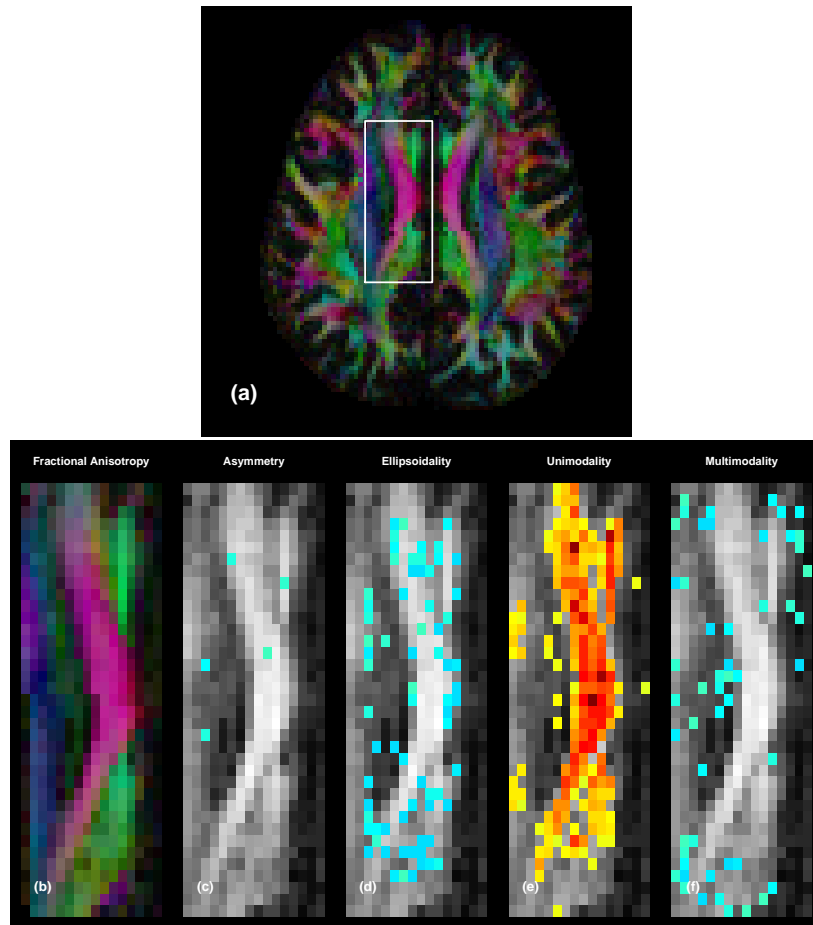


FIG 5. Axial slice from clinical HARDI acquisition. Color-coded fractional anisotropy (FA) for the whole slice is displayed in (a) along with the boundaries for the ROI. For the zoomed-in ROI: color-coded FA (b), anisotropy p -values (c), ellipsoidality p -values (d), unimodality test statistic (e) and multimodality p -values (f).

tion about white-matter microstructure in clinically acquired data.

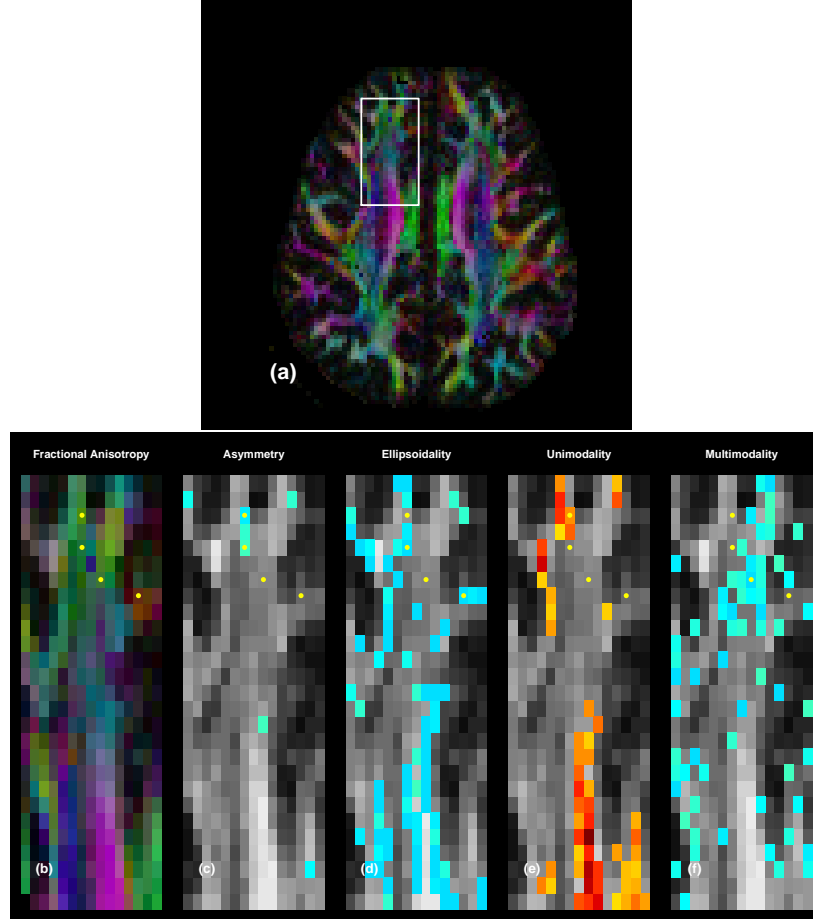


FIG 6. Axial slice from clinical HARDI acquisition. Color-coded fractional anisotropy (FA) for the whole slice is displayed in (a) along with the boundaries for the ROI. For the zoomed-in ROI: color-coded FA (b), anisotropy p -values (c), ellipsoidality p -values (d), unimodality test statistic (e) and multimodality p -values (f).

We focus on a few specific voxels in Figure 6 using the Funk Radon Transform without smoothing. As recommended by Tuch, we have taken the standardized raw FRT to the power five when plotting in order to emphasize structure. Figures 7(a) and 7(b) show the two most anterior voxels that are plotted in 6 (indicated by yellow dots). This tract appears to be “bending” as we move from anterior to posterior, indicated by the shift in direction of the dominant direction seen in the FRTs. The statistics quantify this behaviour; the p -values for asymmetry are 0.14 and 0.02 respectively (indicating that

the posterior-most voxel is bending more). The unimodality of the anterior voxel is seen from the large unimodality statistic in Figure 6(e). We then look at a voxel in a more heterogeneous area, where the major fiber tracts appear to merge: the statistics here indicate multi-modality dominates as is seen in 7(c) and backed up by Figures 6(c)-(f). We observe the most central voxel has summary statistics that are ellipsoidal but not asymmetric, clearly observed in Figure 7(d). The clinical data have provided evidence at a voxel level, backed up by statistical hypothesis testing and observed in the FRT visualizations, that interesting white-matter microstructure may be detected and characterized using the methodology proposed here.

7. Discussion. We have introduced a new set of tools for characterizing orientational structure from HARDI measurements directly in q -space. This methodology is unique when compared with existing methods that rely on reconstructing the spatial information from q -space by different methods of marginalizing the spatial distribution; i.e. from calculating a spatial ODF. An ODF has a different meaning if calculated directly from a Gaussian model, from the non-parametric FRT (average orientational distribution over all radii without using the correct volume increment for a marginal PDF) or using PAS-MRI (orientational distribution associated with a *single spatial radius* or scale). In general the magnitude associated with an ODF is not comparable between methods, neither is the distribution of noise artifacts. Our methodology is technically linked to the FRT, but unlike the FRT we are not constrained to scalar measures calculated from averages on great circles in q -space, and our methods do not depend on appropriate marginalization to produce summaries. The interpretation of our statistical summaries is straightforward but we note that improvements in data acquisition, such as increased sampling of directions and the addition of several shells, will improve estimation of features from the diffusion PDF even further (Khachaturian, Wisco and Tuch, 2007).

Most established methods for characterizing features in white-matter microstructure have focused on the problem of determining the number and orientation of peaks in the diffusion PDF. None of the “magnitude” information of these solutions is comparable or indeed interpretable apart from DTI based models. Savadjiev *et al.* (2006) have already commented on the unsuitability of such magnitudes as quantitative measures. The problem with this fact, and the non-linear transformation often employed for representing q -ball estimates, is that the coherent treatment of noise artifacts becomes much more difficult. The advantage of our theoretical framework developed for summary statistics is that we may perform hypothesis tests using critical

values that are *not* functions of unknown parameters. We stress that simulation studies for features of diffusion PDFs are in general misleading unless the proposed summaries are true statistics; i.e., their distributions under null hypotheses are parameter independent. For example, critical values determined from Monte Carlo studies for a given diffusion PDF will not (in general) be applicable to other diffusion processes than the simulated process since these critical values are parameter dependent. This can be compared to calculating a simple mean rather than a t -statistic. If we try to elicit the distribution of the sample mean using simulations at fixed variances, then these critical values are only useful for variables with the *same* variance.

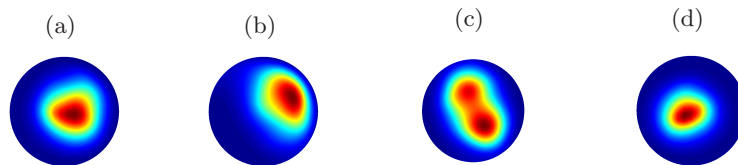


FIG 7. The raw FRT of a set of chosen voxels indicated by yellow dots in Figure 6. These are plotted in order of decreasing x_2 -coordinate (or going from the top of the image to the bottom).

Various non-parametric procedures have been proposed to summarize HARDI data using more than its estimated orientation; e.g., by investigating the model order of spherical harmonic decomposition (Alexander, Barker and Arridge, 2002; Descoteaux *et al.*, 2006; Frank, 2002). Chen *et al.* (2005) modeled the ADC using a product of a truncated spherical harmonics series. In general an *infinite* order of spherical harmonic terms should be taken to approximate an arbitrary Gaussian mixture, but they argued that a crossing fiber should be sufficiently reproduced by such a truncated representation, and expressed its complexity using the normalized terms in the spherical expansion. Other representations include expressing the ADC in terms of higher order tensors and spherical harmonics (Descoteaux *et al.*, 2006), or just via a spherical harmonic representation (Frank, 2002). Second-order terms in a spherical harmonic decomposition contribute to describing a single-tensor fiber, but in general more complicated structure needs to be described in terms of corresponding PDF spatial properties directly, rather than the fourth- and higher-order terms, which give too much freedom in structure to be a precise tool for the description of fine spatial features. Other measures of the entropy of the PDF have been proposed (Rao *et al.*, 2004).

Rather than solely focusing only on the number of peaks in the diffusion PDF, we have characterized white-matter microstructure through the

diffusion PDF directly in q -space without parametric assumptions or imposing smoothness constraints, as we use a variable bandwidth estimator rather than employing a fixed bandwidth smoother (Olhede and Whitcher, 2008a). The tissue microstructure is identified as variation in summary statistics that deviate from a simple, symmetric model for the diffusion PDF and is characterized in behavior relative to the identified dominant great circle in q -space. Ellipsoidal diffusion PDFs (2.2) are simple in structure and imply the existence of a dominant great circle. The deformed ellipsoid class is less stringent in structure, and permits asymmetric decay in minor axes – for example, (2.10) – while still conforming to the existence of a dominant great circle. We describe the precursor to forking by either a deformed ellipse or a mixture model, to capture further asymmetric structure. We differentiate between different white-matter microstructure by examining variation over that great circle, or variation perpendicular to the great circle. Allowing for a greater variety of structure in a unidirectional diffusion PDF implies that the power to detect multi-modal diffusion is necessarily reduced compared to using a parametric multi-model model, if the proposed parametric model is correct. We characterized single peak densities by additional summaries, such as the anisotropy statistic, the decay ratio statistic and the asymmetry statistic. The synthetic forking fiber in Fig. 4 shows an evolution of such measures as we go between a single fiber, and a forking fiber. The synthetic crossing fiber in Fig. 4 does not exhibit the same asymmetries.

If one enforces a strict Gaussian (single diffusion tensor) model, then all variation away from symmetry around the dominant direction will be interpreted as evidence for a multi-modal diffusion (Behrens *et al.*, 2007; Hosey, Williams and Ansorge, 2005; Parker and Alexander, 2005). Modelling using non-Gaussian PDFs allows us to fit asymmetric structure, rather than just the model indicating a lack of fit of a single peak. However, using such models leads to a loss of power if a Gaussian mixture model is appropriate. Caution should be exercised in order to protect against over-interpreting fitted models. With a model that only includes a family of mixtures of Gaussian diffusion processes, one is constrained to estimate a Gaussian mixture, however for a small number of sampled directions there are inevitably issues with identifiability. The same realizations may in some cases equivalently be derived from a unimodal diffusion PDF with asymmetric structure or a Gaussian mixture model. If one chooses to select one model rather than the other (i.e., choose an asymmetric and scalene PDF or multiple-tensor), then this decision is based more on the underlying assumptions of the model rather than on the evidence directly provided by the observed data. A large (possibly infinite) collection of Gaussian diffusion processes may be used to

approximate an observed set of measurements to an arbitrary accuracy, but one has to consider the possibility that the information being fitted is noise instead of signal. We believe the rule of parsimony should be exercised at all times, and that summaries of orientational structure can be estimated and interpreted in q -space rather than using (potentially) over-parameterized models.

One potential application of these q -space summaries would be to improve fiber-tracking algorithms, similar to the use of the Hessian of a local peak to improve probabilistic tractography models (Seunarine *et al.*, 2007). These summaries would be used in addition to directions, to allow more careful tracking through forking and fanning structures (*cf* Figure 6), and distinguish local structure more consistently with crossing from such features using both the asymmetry and ellipsoidality measures.

APPENDIX A: DISTRIBUTIONS FOR TEST STATISTICS

A.1. The T Statistic. By performing a Taylor expansion we have

$$(A.1) \quad T \stackrel{d|H_0}{=} \frac{\tilde{\sigma}}{\mathcal{A}} \left[m^{-1/2}(\varepsilon_1 - \varepsilon_2) + (\varepsilon_4 - \varepsilon_3) \right] + O(\tilde{\sigma}^2),$$

where ε_1 and ε_3 (ε_2 and ε_4) are distributed as the maxima (the minima) of m Gaussian random variables. Assume for a given acquisition scheme that the great circle passes near m independent observations. We find that

$$(A.2) \quad \max_k \{\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)\} \approx \mathcal{A} + \tilde{\sigma}\varepsilon_3 \quad \text{and} \quad \min_k \{\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)\} \approx \mathcal{A} + \tilde{\sigma}\varepsilon_4,$$

where ε_3 and ε_4 are random variables. Then the distribution of ε_3 may be approximated as the maximum of m Gaussian random variables, and likewise with ε_4 as the minimum. As we linearly interpolate the value the extremum will, in the worst case, behave like the extremum of the m variables. Thus

$$(A.3) \quad \max_j \{\hat{\mathcal{A}}(\alpha_j)\} = \mathcal{A} + \frac{\tilde{\sigma}}{\sqrt{m}} \varepsilon_1 \quad \text{and} \quad \min_j \{\hat{\mathcal{A}}(\alpha_j)\} = \mathcal{A} + \frac{\tilde{\sigma}}{\sqrt{m}} \varepsilon_2,$$

where ε_1 and ε_3 (ε_2 and ε_4) are maximum (minimum) order statistics over m Gaussian random variables. Note that maximum and minimum order statistics (Walsh, 1970) can be approximated as independent even for small values of n , and that extrema in the perpendicular great circles will be under H_0 independent of extrema over the dominant great circle.

Assume that $0 < \tilde{\sigma} < \sigma$ is the variance at the observation points after constructing the multiresolution-based estimator $\hat{\mathcal{A}}(\tilde{\mathbf{q}}_j)$ and an estimator of σ is provided by the multiresolution algorithm, and will be denoted $\hat{\sigma}^*$

(Olhede and Whitcher, 2008a). We must now estimate \mathcal{A} and $\tilde{\sigma}$, as indicated in (4.5). We have approximated the interpolated observations as Gaussian random variables, where their variance will vary from observation to observation due to the estimation scheme and the linear interpolation. To be able to determine a suitable threshold we must determine the distribution of the test statistic U

$$(A.4) \quad P(U \leq u) \stackrel{d|H_0}{=} P\left(m^{-1/2}(\varepsilon_1 - \varepsilon_2) + (\varepsilon_4 - \varepsilon_3) \leq u\right) + O\left(\frac{\tilde{\sigma}}{\mathcal{A}}\right).$$

Thus, the PDF of U is approximately given by

$$(A.5) \quad \begin{aligned} f_1(e) &= f_3(e) = f_2(-e) = f_4(-e) = m[\Phi(e)]^{m-1}\phi(e) \\ f_U(u) &= mf_1(e\sqrt{m}) * f_1(e\sqrt{m}) * f_1(e) * f_1(e) \\ F_U(u) &= \int_{-\infty}^u f_U(v) dv. \end{aligned}$$

Under the alternative hypothesis that the “the density is unimodal and prolate,” the expected diffusion on the dominant great circle taking the value \mathcal{A} and on the perpendicular point taking the value \mathcal{A}_{\min} , we have with $\check{\sigma}$ the smoothed variance and $\tilde{\sigma}$ the interpolated variance

$$(A.6) \quad \begin{aligned} T &= \left[\frac{\max_j \{\hat{\mathcal{A}}_{\perp}(\alpha_j)\}}{\min_j \{\hat{\mathcal{A}}_{\perp}(\alpha_j)\}} \right] \left[\frac{\max_k \{\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)\}}{\min_k \{\hat{\mathcal{A}}(\hat{\mathbf{q}}_k)\}} \right]^{-1} - 1 \\ &\stackrel{d|H_1}{=} \left[\frac{\mathcal{A} + \check{\sigma}\eta_1}{\mathcal{A}_{\min} + \check{\sigma}\eta_2} \right] \left[\frac{\mathcal{A} + \tilde{\sigma}\eta_3}{\mathcal{A} + \tilde{\sigma}\eta_4} \right]^{-1} - 1 \\ &= \frac{\mathcal{A}}{\mathcal{A}_{\min}} - 1 + \frac{1}{\mathcal{A}_{\min}} \left[\check{\sigma}\eta_1 - \tilde{\sigma}\eta_3 + \tilde{\sigma}\eta_4 - \frac{\check{\sigma}\mathcal{A}}{\mathcal{A}_{\min}}\eta_2 \right] + \dots \end{aligned}$$

η_3 and η_4 are here maximum and minimum order statistics, respectively, and η_1 and η_2 are Gaussian random variables. We have good estimators for $\check{\sigma}$, $\tilde{\sigma}$ and \mathcal{A} distinct from those used in T , but we are lacking a good independent estimator for $\mathcal{A}/\mathcal{A}_{\min}$. This means we cannot start from a null hypothesis of prolate and unimodal. Let us consider the same object under multi-modality, with $\mathcal{A}_{\perp\max}$ and $\mathcal{A}_{\perp\min}$ the maximum or minimum of the the perpendicular great circle:

$$(A.7) \quad \begin{aligned} T &\stackrel{d|H'_0}{=} \left[\frac{\mathcal{A}_{\perp\max} + \check{\sigma}\eta_5}{\mathcal{A}_{\perp\min} + \check{\sigma}\eta_6} \right] \left[\frac{\mathcal{A}_{\max} + \tilde{\sigma}\eta_7}{\mathcal{A}_{\min} + \tilde{\sigma}\eta_8} \right]^{-1} - 1 \\ &= \check{\sigma} \left(\frac{\eta_8}{\mathcal{A}_{\min}} + \frac{\eta_5}{\mathcal{A}_{\perp\max}} - \frac{\eta_6}{\mathcal{A}_{\perp\min}} - \frac{\eta_7}{\mathcal{A}_{\max}} \right) \dots, \end{aligned}$$

where $\{\eta_i\}$ is a set of Gaussian variables. The distribution of T , under the assumption that

$$(A.8) \quad \frac{\mathcal{A}_{\perp \max} \mathcal{A}_{\min}}{\mathcal{A}_{\perp \min} \mathcal{A}_{\max}} \approx 1,$$

which we shall refer to as H'_0 , is therefore given approximately by the above expression. The test statistic \tilde{U} is calculated under a H'_0 of multimodality, thus $(\tilde{\mathcal{A}}_{\max} \mathcal{A}_{\min})/(\tilde{\mathcal{A}}_{\min} \mathcal{A}_{\max}) \leq c$ for some specified constant such as $c = 2$. Then

$$(\tilde{A}.9) \stackrel{d|H'_0}{=} [c - 1] + \frac{\check{\sigma}}{\mathcal{A}_{\min}} \left(c\tilde{\eta}_8 + \frac{c\mathcal{A}_{\min}}{\tilde{\mathcal{A}}_{\max}} \tilde{\eta}_5 - \frac{c\mathcal{A}_{\min}}{\tilde{\mathcal{A}}_{\min}} \tilde{\eta}_6 - \frac{c\mathcal{A}_{\min}}{\mathcal{A}_{\max}} \tilde{\eta}_7 \right) + \dots,$$

where $\{\tilde{\eta}_l\}_{l=5}^8$ will be Gaussian random variables. The alternative hypothesis will have $(\tilde{\mathcal{A}}_{\max} \mathcal{A}_{\min})/(\tilde{\mathcal{A}}_{\min} \mathcal{A}_{\max}) > c$, and so this is a one-sided test. We assume that $\mathcal{A}_{\min} \approx \tilde{\mathcal{A}}_{\min}$ and that $c\mathcal{A}_{\min}/\tilde{\mathcal{A}}_{\max} \approx c\mathcal{A}_{\min}/\mathcal{A}_{\max} \approx 1$ (these could be replaced by inserting estimators, and most frequently correspond to conservative choices) to produce

$$(A.10) \quad \tilde{T} - [c - 1] \stackrel{d|H'_0}{=} N \left(0, \frac{\check{\sigma}^2}{\mathcal{A}_{\min}^2} (2c^2 + 2) \right) + \dots$$

We estimate \mathcal{A}_{\min} by $\hat{\mathcal{A}}_{\min} = \hat{\mathcal{A}}(\hat{v}_1)$ to get \tilde{U} , and replace $\check{\sigma}$ by the (conservative) estimate of $\hat{\sigma}_{\mathcal{A}}$.

A.2. The X Statistic. Under a null hypothesis of isotropy we find that

$$(A.11) \quad X - 1 \stackrel{d|H_0}{=} \frac{\tilde{\sigma}\eta_9 - \check{\sigma}\eta_{10}}{\mathcal{A} \log \mathcal{A}} + O(\tilde{\sigma}^2),$$

where η_9 is the maximum order statistic and η_{10} may be approximated as a Gaussian random variable. We can approximate the density of Q via

$$(A.12) \quad f_{Q|H_0}(q|H_0) = \rho^{-1/2} \phi \left(\frac{q}{\sqrt{\rho}} \right) * \left[m \Phi^{m-1} \left(-\frac{q}{\rho} \right) \rho^{-1} \phi \left(-\frac{q}{\rho} \right) \right],$$

for $-\infty < q < \infty$, where $*$ denotes the convolution operator, $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard Gaussian CDF and PDF, respectively.

A.3. The Z Statistic. To derive the Z statistic we note that with ε_5 an aggregation of a maximum over m Gaussian random variables, and an independent Gaussian observation, and ε_6 is a Gaussian random variable

$$(A.13) \quad Z = \frac{\log \mathcal{A}(\gamma) + (\tilde{\sigma}\varepsilon_5)^{-1}\mathcal{A}(\gamma)}{\log \mathcal{A}(\iota) + (\check{\sigma}\varepsilon_6)^{-1}\mathcal{A}(\iota)}$$

$$(A.14) \quad = \frac{\log \mathcal{A}(\gamma)}{\log \mathcal{A}(\iota)} + \frac{\tilde{\sigma}\varepsilon_5}{\mathcal{A}(\gamma)\log \mathcal{A}(\iota)} - \frac{\check{\sigma}\varepsilon_6 \log \mathcal{A}(\gamma)}{\mathcal{A}(\iota)\log^2 \mathcal{A}(\iota)},$$

where $\gamma = \mathbf{q}_{k_{\max}+M}$, $\iota = \mathbf{q}_{k_{\max}+M+N/4}$ and $M = N/(2m)$. Under the null $\mathcal{A}(\gamma) = \mathcal{A}(\iota)$ and so

$$(A.15) \quad Z \stackrel{d|H_0}{=} 1 + \frac{\tilde{\sigma}\varepsilon_5}{\mathcal{A}(\iota)\log \mathcal{A}(\iota)} - \frac{\check{\sigma}\varepsilon_6}{\mathcal{A}(\iota)\log \mathcal{A}(\iota)}$$

$$(A.16) \quad V = \frac{(Z-1)\overline{\mathcal{A}}_N \log \overline{\mathcal{A}}_N}{\hat{\sigma}_2} = \varepsilon_5 - \rho^{-1/2}\varepsilon_6 = \varepsilon_7,$$

where ε_5 and ε_6 are due to properties of the interpolation. The random variables ε_5 and ε_6 may be approximated as independent. Note that the PDF of ε_7 is approximated by the following density

$$(A.17) \quad F_{\varepsilon_7}(e) = P(\varepsilon_7 \leq e), \quad g(e) = \frac{25m}{2\sqrt{2}} \left[\Phi^{m-1}(5e') \phi(5e') * \phi\left(\frac{5}{2\sqrt{2}}e'\right) \right](e),$$

and $f_{\varepsilon_7}(e) = (g*g)(e)$. Depending on the value of m' we find that the approximation performs well.

A.4. The K Statistic. Under the null hypothesis K will be approximately distributed as the absolute value of a Gaussian random variable with mean zero and variance given by $\text{var}\{K\} \approx 4(1.35)^2 m^{-1} \text{var}\{P_k\}$, where the factor multiplying the variance of P_k is chosen to reflect the degrees of freedom of the sampled sphere. We note that the variance of P_k is given by approximated by

$$(A.18) \quad \text{var}\{P_k\} \approx \frac{32\tilde{\sigma}^2}{m \left[\sum_{j=1}^N \mathcal{A}_{\perp}(\alpha_j)/N \right]^2} \Rightarrow \text{var}\{K\} = \frac{2^7 \tilde{\sigma}^2 (1.35)^2}{m^2 \left[\sum_{j=1}^N \mathcal{A}_{\perp}(\alpha_j)/N \right]^2}.$$

The normalized statistic is thus defined by

$$(A.19) \quad R = \frac{mK \left[\sum_{j=1}^N \hat{\mathcal{A}}_{\perp}(\alpha_j)/N \right]}{2^{7/2}(1.35)\hat{\sigma}_2} = \frac{R_2}{2^{3/2}(1.35)} \sim 2\phi(r), \quad r > 0.$$

REFERENCES

- ABRAMOWITZ, M. and STEGUN, I. A. (1972). *Handbook of Mathematical Functions*, 10th ed. NBS.
- ALEXANDER, D. C. (2005). Multiple-fibre reconstruction algorithms for diffusion MRI. *Annals of the New York Academy of Sciences* **1046** 113–133.
- ALEXANDER, D. C., BARKER, G. J. and ARRIDGE, S. R. (2002). Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine* **48** 331–340.
- BASSER, P. J. (2002). Relationships between diffusion tensor and q -space MRI. *Magnetic Resonance in Medicine* **47** 392–397.
- BASSER, P. J., MATTIELLO, J. and LE BIHAN, D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin-echo. *J. of Magnetic Resonance B* **103** 247–254.
- BASSER, P. J. and PIERPAOLI, C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. of Magnetic Resonance B* **111** 209–219.
- BEHRENS, T. E. J., JOHANSEN-BERG, H., JBABDI, S., RUSHWORTH, M. F. S. and WOOLRICH, M. W. (2007). Probabilistic tractography with multiple fibre orientations: What can we gain? *NeuroImage* **34** 1077–1088.
- CALLAGHAN, P. T. (1993). *Principles of Nuclear Magnetic Resonance Microscopy*. Clarendon Press, Oxford, UK.
- CHEN, Y., GUO, W., ZHENG, Q., RAO, M. and LIU, Y. (2005). Apparent Diffusion Coefficient Approximation and Diffusion Anisotropy Characterization in DWI. In *IPMI 2005* 246–257.
- DESCOTEAUX, M., ANGELINO, E., FITZGIBBONS, S. and DERICHE, R. (2006). Apparent Diffusion Coefficients from High Angular Resolution Diffusion Imaging: Estimation and Applications. *Magnetic Resonance in Medicine* **56** 395–410.
- DESCOTEAUX, M., ANGELINO, E., FITZGIBBONS, S. and DERICHE, R. (2007). Regularized, Fast and Robust Analytical Q -Ball Imaging. *Magnetic Resonance in Medicine* **58** 497–510.
- FRANK, L. R. (2002). Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine* **47** 1083–1099.
- GRADSHTEYN, I. S. and RYZHIK, I. M. (2000). *Tables of Integrals, series and Products*, 6th ed. Academic Press, New York, USA.
- GUDBJARTSSON, H. and PATZ, S. (1995). The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine* **34** 910–914.
- HESS, C. P., MUKHERJEE, P., HAN, E. T., XU, D. and VIGNERON, D. B. (2006). Q -Ball reconstruction of Multimodal Fibre Orientations Using the Spherical Harmonic Basis. *Magnetic Resonance in Medicine* **56** 104–117.
- HOSEY, T., WILLIAMS, G. and ANSORGE, R. (2005). Inference of multiple fiber orientations in high angular resolution diffusing imaging. *Magnetic Resonance in Medicine* **54** 1480–1489.
- JANSON, K. M. and ALEXANDER, D. C. (2003). Persistent angular structure: New insights from diffusion magnetic resonance imaging data. *Inverse Problems* **19** 1031–1046.
- JENSEN, J. H., HELPERN, J. A., RAMANI, A., LU, H. and KACZYNSKI, K. (2005). Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine* **53** 1432–1440.
- JIAN, B. and VEMURI, B. C. (2007). A Unified Computational Framework for Deconvolution to Reconstruct Multiple Fibres From Diffusion Weighted MRI. *IEEE Trans. on Medical Imaging* **26** 1464–1471.

- KADEN, E., KNÖSCHE, T. R. and ANWANDER, A. (2007). Parametric Spherical Deconvolution: Inferring anatomical connectivity using diffusion MR imaging. *NeuroImage* **37** 474–488.
- KHACHATURIAN, M. H., WISCO, J. J. and TUCH, D. S. (2007). Boosting the sampling efficiency of q -ball imaging using multiple wavevector fusion. *Magnetic Resonance in Medicine* **57** 289–296.
- MATÉRN, B. (1960). Spatial Variation. In *Stochastic Models and Their Application to some problems in forest surveys and other sampling investigations* 39. Statens Skogsforskningsinstitut, Stockholm, Sweden.
- MORI, S. and VAN ZIJL, C. M. (2002). Fiber tracking: principles and strategies – a technical review. *NMR in Biomedicine* **15** 468–480.
- OLHEDE, S. C. and WHITCHER, B. (2008a). HARDI wavelet and Non-Parametric Estimation. University College, London, United Kingdom.
- OLHEDE, S. C. and WHITCHER, B. (2008b). A Statistical Framework to characterise Microstructure in High Angular Resolution Diffusion Imaging. In *5th IEEE International Symposium on Biomedical Imaging* 899–902.
- ÖZARSLAN, E., VEMURI, B. C. and MARECI, T. M. (2005). Generalized Scalar measures for Diffusion MRI Using Trace, Variance and Entropy. *Magnetic Resonance in Medicine* **53** 866–876.
- ÖZARSLAN, E., SHEPHARD, T. M., VEMURI, B. C., BLACKBAND, S. J. and MARECI, T. M. (2006). Resolution of Complex Tissue Microarchitecture using the diffusion orientation transform. *NeuroImage* **31** 1086–1103.
- PARKER, G. J. M. and ALEXANDER, D. C. (2005). Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Trans. of the Royal Society of London B* **360** 893–902.
- RAO, M., CHEN, Y., VEMURI, B. C. and WANG, F. (2004). Cumulative Residual Entropy: A New Measure of information. *IEEE Trans. on Information Theory* **50** 1220–1228.
- SAVADJIEV, P., CAMPBELL, J. S. W., PIKE, G. B. and SIDDIQI, K. (2006). 3D curve inference for diffusion MRI regularization and fibre tractography. *Medical Image Analysis* **10** 799–813.
- SEUNARINE, K., COOK, P. A., HALL, M. G., EMBLETON, K. V., PARKER, G. J. M. and ALEXANDER, D. C. (2007). Exploiting peak anisotropy for tracking through complex structure. In *Mathematical Methods in Biomedical Image Analysis* 1. IEEE Computer Society, Rio de Janeiro.
- TOURNIER, J., CALAMANTE, F., GADIAN, D. and CONNELLY, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* **23** 1176–1185.
- TUCH, D. S. (2004). Q-ball Imaging. *Magnetic Resonance in Medicine* **52** 1358–1372.
- TUCH, D. S., REESE, T. G., WIEGELL, M. R., MAKRIS, N., BELLIVEAU, J. W. and WEDEEN, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine* **48** 577–582.
- WALSH, J. E. (1970). Sample sizes for appropriate independence of largest and smallest order statistics. *J. of the American Statistical Association* **65** 860–863.
- WEDEEN, V. J., HAGMANN, P., TSENG, W.-Y., REESE, T. G. and WEISSKOFF, R. M. (2005). Mapping Complex Tissue Architecture with Diffusion Spectrum Magnetic Resonance Imaging. *Magnetic Resonance in Medicine* **54** 1377–1386.

DEPARTMENT OF STATISTICAL SCIENCE
UNIVERSITY COLLEGE LONDON
GOWER STREET
LONDON WC1 E6BT
UNITED KINGDOM
E-MAIL: s.olhede@ucl.ac.uk

CLINICAL IMAGING CENTRE
GLAXOSMITHKLINE
HAMMERSMITH HOSPITAL
IMPERIAL COLLEGE LONDON
DU CANE ROAD
LONDON W12 0HS
UNITED KINGDOM
E-MAIL: brandon.j.whitcher@gsk.com