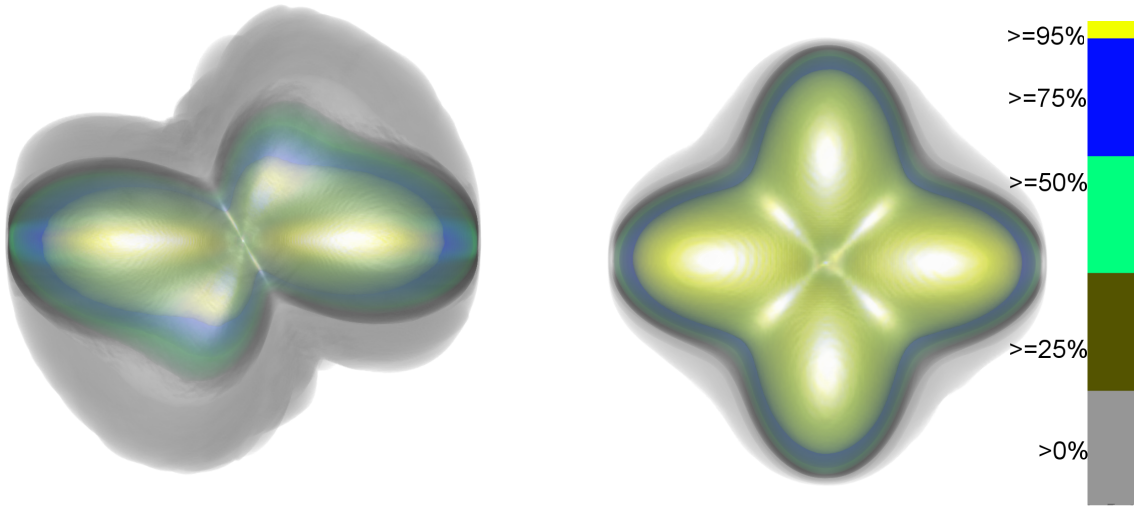
Uncertainty Visualization in HARDI based on Ensembles of ODFs

Fangxiang Jiao[12] Jeff M. Phillips[2] Yaniv Gur[12] Chris R. Johnson[12]
[1]SCI Institute [2]University of Utah *



(a) b-value 7000 $\mathrm{s/mm^2}$, SNR=10, $(0.6, 0.4)$ 60° (b) b-value 1000 $\mathrm{s/mm^2}$, SNR=20, $(0.5, 0.5)$ 90°

Figure 1: Visualization of the uncertainty in two diffusion shapes. (a) Two fibers crossing at 60 degrees with relative weight of 0.6:0.4 and SNR of 10. (b) Two fibers crossing at 90 degrees with equal weight and SNR of 20 (with much less uncertainty).

ABSTRACT

In this paper, we propose a new and accurate technique for uncertainty analysis and uncertainty visualization based on fiber orientation distribution function (ODF) glyphs, associated with high angular resolution diffusion imaging (HARDI). Our visualization applies volume rendering techniques to an ensemble of 3D ODF glyphs, which we call *SIP functions* of *diffusion shapes*, to capture their variability due to underlying uncertainty. This rendering elucidates the complex heteroscedastic structural variation in these shapes. Furthermore, we quantify the extent of this variation by measuring the fraction of the volume of these shapes, which is consistent across all noise levels, the *certain volume ratio*. Our uncertainty analysis and visualization framework is then applied to synthetic data, as well as to HARDI human-brain data, to study the impact of various image acquisition parameters and background noise levels on the diffusion shapes.

Keywords: DT-MRI, Uncertainty, HARDI, Rank-k tensor decomp.

1 INTRODUCTION

Diffusion-weighted magnetic resonance (DW-MR) is an imaging technique that enables one to measure diffusion priorities of water molecules in a fibrous tissue [3]. It is primarily used to reveal the white-matter fibers structure of the brain, as well as structures of muscle fibers. Under the assumption of a Gaussian diffusion, the connection between the acquired DW signal and the diffusion rate is described by the Stejskal-Tanner equation:

$$S(\mathbf{g}, b) = S_0 e^{-bD(\mathbf{g})}, \quad (1)$$

*Supported by NIH/NCRR Center for Integrative Biomedical Computing, 2P41-RR12553-12, Award KUS-C1-016-04, by KAUST, and DOE SciDAC VACET and DOE NETL, by subaward to the Univ. Utah under NSF award 1019343 to CRA, and by NIH Autism Center of Excellence grant (NIMH and NICHD #HD055741).

where $S(\mathbf{g}, b)$ measures the signal decay in a direction $\mathbf{g}$ due to the application of a magnetic field gradient, and S_0 is the non-weighted diffusion signal (the signal measured in the presence of a constant magnetic field). Given a vector $\mathbf{S}$ of measurements in n gradient directions, an optimization problem is solved for $D(\mathbf{g})$, the diffusivity function. This problem is solved for each brain voxel where $D(\mathbf{g})$ describes the apparent diffusion coefficient (ADC) in direction $\mathbf{g}$. The parameter b, which we denote here as *the sensing parameter*, is determined by the physical properties of the measurement (e.g., magnetic field strength, pulse duration).

In diffusion-tensor imaging (DTI), the diffusivity function is represented by a symmetric positive-definite second-order tensor known as the diffusion tensor (DT). Diffusion tensors can only model one dominant diffusion direction per voxel, and thus, are incapable of describing complex fiber structures. To overcome this limitation, new imaging techniques, which are known as high angular resolution diffusion imaging (HARDI), have emerged. There are various image reconstruction techniques associated with HARDI (e.g., [49, 46, 11]). In most of these reconstruction techniques an orientation distribution function (ODF) is used to describe the diffusion profile. Unlike a DT, an ODF can have multiple maxima, thus, it can capture complex fiber structures such as crossing fibers.

The estimation accuracy of an ODF (or a DT) is influenced by a number of parameters such as the b-value, number of gradient directions and signal-to-noise ratio. These different parameters introduce a level of uncertainty to the estimated diffusion profiles, and hence, uncertainty analysis and visualization become crucial to obtaining a better understanding of the fiber structure.

Over the last decade, various works have been devoted to the uncertainty study in the DTI model (e.g., [35, 27, 53]). However, only a handful of papers study the uncertainty in the ODFs associated with the HARDI model [8, 12, 55, 15, 42, 46].

An ODF in DW-MR is often presented as a symmetric spherical glyph that provides information on the three-dimensional diffusion profile within a brain voxel. We refer to these glyphs and their many variants [31, 21] as *diffusion shapes* since they provide a tan-

gible shape-based representation of the dominant diffusion directions. The distance from the origin to the boundary of the diffusion shape is typically defined as being proportional to the value of the ODF in that direction; we use this definition. While an ODF is a function that assigns a scalar value to each point in S^2, a diffusion shape assigns either 0 or 1 to each point in $\mathbb{R}^3$, where a point inside the diffusion shape is assigned 1, and a point outside is assigned 0. This binary representation is crucial in our ensemble structure. However, diffusion shapes alone do not provide any information on ODF variabilities due to uncertainty caused by the different parameters. The main contribution of this paper is an uncertainty analysis and visualization tool based on *ensembles* of diffusion shapes.

Our approach to analyze ensembles follows the framework of Löffler and Phillips [34] to construct a shape inclusion probability function (or SIP function), applied here to diffusion shapes. That is, we construct a function on a three-dimensional domain where each point indicates the probability of inclusion in the diffusion shape. When there is no uncertainty in the data, this function is again binary where the values 1 or 0 indicate whether a point belongs to the diffusion shape, or not, respectively. However, in real scenarios, the uncertainty results in a region of points that take SIP values in the range $[0,1]$. Studying and visualizing this heterogeneous region provides insight into the effects of different parameters describing the underlying uncertainty. We focus on one, we believe, particularly effective visualization of this SIP function. It is rendered using a 1D transfer function, which results in a 3D multiple iso-level visualization that approximates the inclusion probability at several levels simultaneously.

Following the approach of Löffler and Phillips to construct the SIP function, we first instantiate a number (1000) of DW-MR data sets drawn from the probability distribution modeling the uncertainty. For each instantiation we construct the diffusion shape, and at each voxel we report the fraction of diffusion shapes which include that point. This provides a sample estimate, but accompanying analysis provides accuracy bounds that depend only on the number of samples and complexity of the shapes, indicating that a large enough sample (such as 1000) is sufficient to accurately visualize the SIP for the diffusion shape. This entire approach produces a *quantifiable visualization* where the key structure is clearly presented *and* the underlying data values can be accurately recovered.

Outline. The paper is organized as follows: In Section 2 we give an overview of the related work. In Section 3 we present the method to construct a SIP function from an ensemble of estimated diffusion shapes. Section 4 describes how the data is generated in the synthetic and real case. Results are presented in Section 5 where we employ this technique to study the impact of the background noise (Rice distributed) and b-value on the uncertainty in diffusion modeling. Finally, concluding remarks are drawn in Section 6.

2 Related Work

Several approaches for analyzing the uncertainty inherent in Diffusion Tensor Imaging (DTI) have been proposed. One approach studies the uncertainty in diffusion anisotropy measures derived from the components of the tensor [3], such as the FA (Fractional Anisotropy), RA (Rational Anisotropy), and principal diffusion directions [35, 27, 29, 33, 52, 9]. Often bootstrapping methods are employed to simulate the distribution governing the random noise on real data sets. A different approach studies DTI (and HARDI) data sets, where several measurements in multiple imaging directions are available. These multiple readings allow one to construct a distribution of possible values, and an estimation of noise related to the choice of b value [13, 46, 12, 17, 38], the registration [41, 23], the model fitting [1, 39, 30, 51, 24], the maximum finding, decomposition, or enhancement [25, 43, 22, 38, 42], or the fiber tracking algorithms [18, 10, 14, 36, 54, 26, 12].

Most of the uncertainty analysis and visualization focuses on the eigenvector directional uncertainty caused by the background noise [27, 19, 7] or the choice of a b value [12]. Only a few studies [55, 4, 1, 15, 53] address the uncertainty of the diffusion shape, which can describe complex white-matter structures. Schwartzman [45] provides a simplified representation for diffusion tensors as ellipsoids and presents a 2D-quantile visualization along restricted slices. PCA analysis was also employed to perform the uncertainty analysis in 2nd order tensor field [53] and Q-ball imaging [15]. Recently, Zhan et al. [55] studied the influence of the number of diffusion weighted directions on the diffusion shape and the resulting parameters. They showed that as the number of sensing directions increased, the signal to noise ratio also increased when measured with respect to fractional anisotropy (FA), rational anisotropy (RA), geodesic anisotropy (GA) and hyperbolic tangent of geodesic anisotropy (tGA) for an ROI in corpus callosum. They also showed that the diffusion shape converges toward the ground truth shape when the number of sensing directions is increased. However, they did not study or visualize the more nuanced structural uncertainty with respect to the sensing parameter b, nor with respect to the parameters governing the true diffusion shape such as the angle between the dominant diffusion directions and the corresponding weights. Also, Tournier et.al. [46] visualize uncertainties of the ODFs using semitransparent glyphs; but are only able to represent the mean and standard deviation of variation in diffusion shapes. This lacks the ability to capture the full fiber orientation variabilities described by a large set of noise realizations. Therefore, a better understanding about uncertainty analysis and visualization in HARDI is needed.

3 Methods

We propose a general method that can be applied to visualize uncertainty in ODFs associated with various sources of data. In this paper, we demonstrate our proposed method on ensembles of diffusion shapes derived from ODFs that are estimated from DW-MR data sets. First, we revisit how to determine ODFs. Our framework is agnostic to which technique is used, but we describe the specific variant that our visualization is built on. Second, we discuss our proposed method to construct the SIP functions to visualize the uncertainty represented by a large sample estimate of the diffusion shapes according to the uncertainty in the data due to noise.

3.1 The Diffusion Shape Estimation Technique

Classical DTI constructs a diffusion tensor representing the major diffusion directions by solving the best fit (second-order) tensor in (1) using a least-squares approach. A single diffusion tensor is associated with a second-order homogeneous polynomial which is incapable of describing complex diffusion patterns, such as crossing fibers. In HARDI, however, the estimation process results in an ODF that is associated with an even-order homogeneous polynomial of a higher order. Unlike classical DT, an ODF has more degrees of freedom and, hence, can represent 2 or more crossing fibers simultaneously and their corresponding weights (e.g., [49, 46, 11]).

Although our framework is applicable to ODFs associated with polynomials of any order, we focus here on fourth-order polynomials as they provide adequate trade-off between angular resolution and over-fitting due to noise [44]. The diffusivity function is

$$D(\mathbf{g}) = \sum_{a=0}^{4} \sum_{b=0}^{4-a} c_{ab} g_1^a g_2^b g_3^{4-a-b}, \tag{2}$$

where c_{ab} are the tensor coefficients and each vector $\mathbf{g} = [g_1, g_2, g_3]$ denotes a measurement direction. Figure 2 shows an example of different noise realizations of fourth-order diffusion shapes reconstructed from synthetically generated dataset that simulates two crossing fibers at 60 degrees at a b-value of 2000 and an SNR of

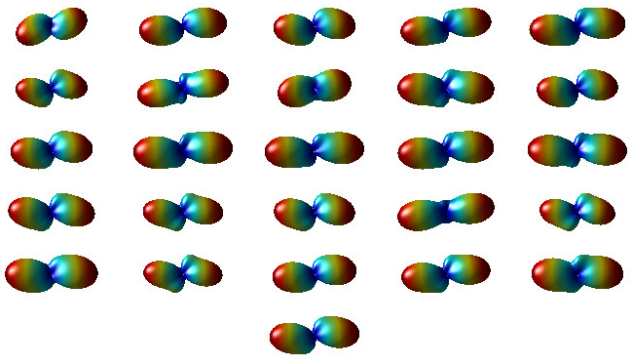


Figure 2: 25 noisy diffusion shapes at the crossing of two fibers at 60° and relative weights $0.7:0.3$. The SNR is 10 and b-value is 2000 $\mathrm{s/mm^2}$. The original, noiseless diffusion shape is shown below.

10. The diversity of the reconstructed shapes illustrates variation of diffusion patterns under noise due to modest SNR values, and motivates the study of visualization of this inherent uncertainty.

There are various HARDI techniques to estimate ODFs (e.g., [46, 11, 49, 25]), and all are applicable within our framework. We choose the recent HARDI variant based on spherical deconvolution to perform all further experiments because it enables one to extract easily the diffusion information (e.g., [47, 25, 51]). These approaches ensure that the ODF is a positive definite entity. Following [25], we solve a minimization problem of the form

$$\min_D \frac{1}{2}\sum_{i=1}^{n}\left\| S(\mathbf{g}_i,b) - S_0\int_{S^2} D(\mathbf{v})K(\mathbf{g}_i,\mathbf{v},b)d\mathbf{v}\right\|^2, \qquad (3)$$

subject to a set of constraints that guarantees the positive-definiteness of the ODF in m reconstruction directions. The optimization problem is solved for the coefficients of the homogeneous polynomial, c_{ab}, described in (2). The coefficients are estimated for a given set of diffusion-weighted measurements, S, and a convolution kernel, K. For a fourth-order polynomial we estimate a set of 15 coefficients that describe it uniquely. Once the coefficients have been estimated, the associated diffusion shape is reconstructed according to (2). Importantly, the restriction that the ODF is described by 15 parameters prevents over-fitting, and will be relevant in the forthcoming accuracy analysis.

3.2 Diffusion Shape as a 3D Volume

To better understand how the diffusion shape varies with respect to the different acquisition parameters and background noise, we propose to change the widely used (single) triangular mesh visualization of diffusion shapes [2] (many instances are shown in Figure 2) to a probabilistic 3D volume, where volume rendering and isosurfacing techniques can be used to visualize these variations. Figure 1 shows the volume rendering of two tensors with four approximate iso-levels of the SIP function (defining 5 layers), indicating the probability (governed by uncertainty) that a diffusion shape occupies that part of the volume. Boundaries between layers demarcate the approximate iso-levels at the 25% quantile (inside grey region, outside dark grey region), the 50% quantile (inside dark grey region, outside green region), the 75% quantile (inside green region, outside blue region), and the 95% quantile (outside yellow region, inside blue region). That is, all points in the yellow region are included in the diffusion shape with probability at least 0.95 and represent the *certain* part of the diffusion shape.

The diffusion shape is centered in a $200\times200\times200$ cube and normalized so the furthest point is at the boundary of the cube. This is justified by the fact that the outcome of the spherical deconvolution is a fiber-ODF, which only provides relative measurements of fiber orientations; hence, the scale is insignificant. The SIP function is evaluated at the center of each voxel. Its value is based on diffusion shapes estimated from 1000 different noise realizations of the raw data (as described in Section 4). Specifically, the SIP measures the fraction of diffusion shapes, each constructed using (2) from an instantiated data set, that contains the voxel center.

As the SIP function is calculated via a random process, it is subject to some error itself, but this process is unbiased and the error can be quantified [34]. In particular, via Theorem 4 in [34], the worst case error of a SIP function at any voxel is inversely proportional to the square root of the number of random instantiations m, and proportional to the square root of the complexity of the diffusion shape ν (i.e. the VC-dimension [50] dual to the shape, called ν' within proof of Theorem 4[34], which can be bounded by the largest degree of a polynomial describing a shape). Specifically, the maximum error ε of the SIP function over all voxels is approximately $\sqrt{(1/2m)(\nu+\ln(1/\delta))}$ with probability $1-\delta$.

Thus, in our case, with $\delta=.01$, $m=1000$, and $\nu=15$, then on 99% of all simulations the maximum error for any voxel is at most 10%. Most voxels will have much less error. We can interpret this as being able to reliably distinguish between iso-levels about every 10% (i.e. $>90\%, >80\%, >70\%, \ldots, >10\%$).

However, our visualization only chooses to highlight four different iso-levels (25%, 50%, 75%, 95%) and 5 probability regions ($<25\%$, $25\%-50\%$, $50\%-75\%$, $75\%-95\%$, $>95\%$) because of the challenge of rendering a larger number of layers in 3D in an effective way. We use an analogous color map [40]. We choose one dominant color (we found a bright yellow visually appealing) to highlight the certain, innermost region. And then the subsequent, intermediate layers were shown as blues and greens which are adjacent on the color wheel. Neutral grays were chosen to represent the outermost layers with most uncertainty.

By decreasing the opacity of the outer layers, we are able to easily recognize all 5 regions on the 3D image, even through the outer layers completely occlude the inner layers. We found using more than 5 regions detracted from the visual aesthetics of the images without adding significant additional information. Most analogous colormaps derived to distinguish several iso-levels (for instance as provided by colorbrewer : `colorbrewer2.org`) contrast alternating levels by varying brightness. However, when using these colormaps in our 3D visualizer, we view the inner layers through

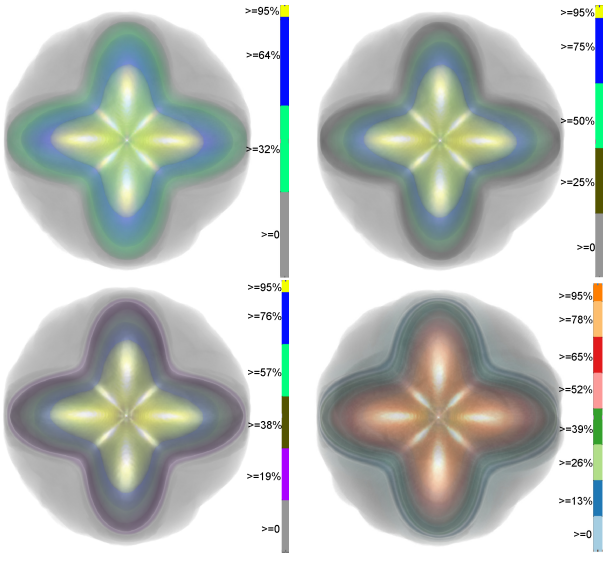


Figure 3: SIP function of diffusion shapes with $b=1000$ and SNR of 5 illustrated with colormaps highlighting 4 (upper left), 5 (upper right), 6 (lower left), or 8 (lower right) layers.

the decreased-opacity outer layers; this combines these outer layers and muddles the view of the inner layers. As seen in Figure 3 which shows alternate color maps with between 4 and 8 layers, each layer is recognizable (on close inspection), but they appear increasingly drab as the number of layers increase, despite vibrant color maps.

This choice of only 4 iso-levels (and 5 regions) can be used to reinterpret the analysis in two ways. Either we can say with 1000 samples we achieve 25% accuracy on the SIP function with probability greater than $1-1/10^{10}$. Or we reach 10% accuracy, meaning each iso-level visualized (say the .5 iso-level) represents the true value with error at most 10% (its value is between .4 and .6) in the entire SIP function with probability greater than 0.99.

The visualization technique that we propose here is applicable to various diffusion shape representations, reconstructed from DTI or HARDI data. Figure 4 illustrates our technique for computing SIP glyphs when using several different algorithms to generate the base diffusion shape, including classic DTI [37, 32], super quadratics [31] and fourth-order homogeneous polynomial [25] (as a representative visualization of any ODF-based HARDI techniques such as Qball [49]). As expected, the two techniques based on classical DTI (left two) fail to represent the two crossing fibers. Also, they exhibit more ambiguous uncertainty (note the larger, more uniform light grey region), presumably as the DTI technique is unable to consistently fit a single dominant direction on two crossing fiber directions. Hence, we focus here on the HARDI approaches.

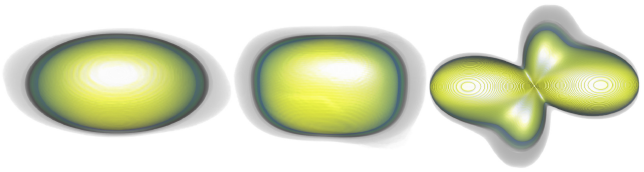


Figure 4: SIP functions for different ODF representations: Left to right: ellipsoids (DTI), super quadratic (DTI), fourth-order homogeneous polynomial (HARDI). All use the same data with SNR of 20, b-value = 3000, two crossing fibers at 75° and relative weights $0.6:0.4$.

4 Synthetic and Human Brain Data

We apply our method to synthetic and real, in-vivo, human brain DW data.

4.1 Synthetic Data Design

In the synthetic data setting two crossing fibers are simulated according to the multi-compartment model [48]

$$S(\mathbf{g},b) = S_0 \sum_{i=1}^{2} w_i e^{-b\mathbf{g}\mathrm{D}_i\mathbf{g}^T}. \tag{4}$$

For both tensors we assume a prolate tensor model with eigenvalues $\lambda_1 = 1.9\cdot 10^{-4}\mathrm{mm}^2/\mathrm{s}$, $\lambda_2=\lambda_3 = 1\cdot 1.0^{-5}\mathrm{mm}^2/\mathrm{s}$ and b-values varying from $1000\,\mathrm{s/mm}^2$ to $8000\,\mathrm{s/mm}^2$ in $1000\,\mathrm{s/mm}^2$ intervals. The principal eigenvectors of the tensors were chosen to simulate crossing fibers at angles of 30, 60, and 90 degrees. To test the influence of the fiber weights data sets are generated with fibers weight equally or at ratio $0.7:0.3$. The signal was simulated using 64 gradient directions. The clean data sets $S(\mathbf{g}_i,b)$ were corrupted by Rician distributed noise to create six levels of noisy data sets with SNR_0[1]: 50, 40, 30, 20, 10 and 5.

For each noise level the process was repeated 1000 times. In each repetition, for each set of parameters (i.e., noise level, b-value, crossing angle and weight) a diffusion shape was estimated according to the method described in Section 3.1. Then, the 1000 different instantiated diffusion shape samples are used to construct the SIP function as described in Section 3.2. See Figure 5.

[1] $\mathrm{SNR}_0 = S_0/\sigma$, where σ in the noise standard deviation.

4.2 The Data Acquisition and Random Effects Simulation of Human Brain

The human brain data was acquired on a 3 Tesla Siemens Tim Trio scanner with a single-shot spin-echo EPI sequence with a resolution of $2\times 2\times 2\mathrm{mm}^3$. One B0 image and 64 diffusion weighted images, with image resolution of $106\times 106\times 76$, were acquired. To test the effects of the b-value, three different b-values were used to acquire the data: 1000, 2000 and 3000 $\mathrm{s/mm}^2$. Then, eddy current correction and mutual information based affine registration was applied.

There are two predominant approaches to simulate noise in real data sets for DW-MR: directly adding Rician noise, and wild bootstrapping. Both techniques start by first solving for a single ODF or DT. For each sensing direction $\mathbf{g}_i$ let $S_i = S(\mathbf{g}_i)$ be the sensed value. Let $\mathcal{S} = \{S_1,\ldots,S_n\}$ be this data set and $\hat{D}$ be the estimated ODF according to the method discussed in Section 3. Then, for each direction $\mathbf{g}_i$, we compute $\hat{S}(\mathbf{g}_i)$ as a function of $\hat{D}$ using (3). The residual is given by $r_i = \hat{S}(\mathbf{g}_i) - S_i$.

The technique of adding Rician noise estimates the standard deviation σ of a complex Gaussian using the r_i values. It instantiates uncertain data sets $\mathcal{S}_j = \{S_{1,j}, S_{2,j},\ldots,S_{n,j}\}$ from the perceived noise levels. To replicate each sensed value it sets $S_{i,j} = \hat{S}(\mathbf{g}_i)\oplus R(\sigma)$, where $\oplus R(\sigma)$ signifies the perturbation of $S(\mathbf{g}_i)$ by Rician noise with standard deviation of σ. It has been observed that the noise in these DTI measurements closely follows a Rice distribution [16], so this technique has been accepted as a realistic way of simulating noise [5].

Wild bootstrapping [9, 28] makes an effort to preserve the heteroscedasticity of the noise (that is, certain areas have larger noise distributions than others). But, as consequence, sacrifices the strict adherence to the Rice noise model. Again starting from each $\hat{S}(\mathbf{g}_i)$ we instantiate a data set from the noise as $\mathcal{S}_j = \{S_{1,j}, S_{2,j},\ldots,S_{n,j}\}$. We then set $S_{i,j} = \hat{S}(\mathbf{g}_i) \pm r_i$, where the choice of adding or subtracting the residual is chosen independently at random for each simulated measurement and instantiated data set (i.e. each j and i).

A third approach estimates the posterior probability density functions of the model parameters under certain noise models [6].

To simulate noise on real data, any of these conventional methods could be used. They would be repeated 1000 times (i.e. for $j=1,\ldots,1000$), to instantiate 1000 data sets representing the distribution of possible data sets as prescribed by the residual noise in the sensed data. In our experiments we choose to follow the wild bootstrapping method since we believe it will highlight our algorithm's ability to visualize complicated heteroscedastic noise.

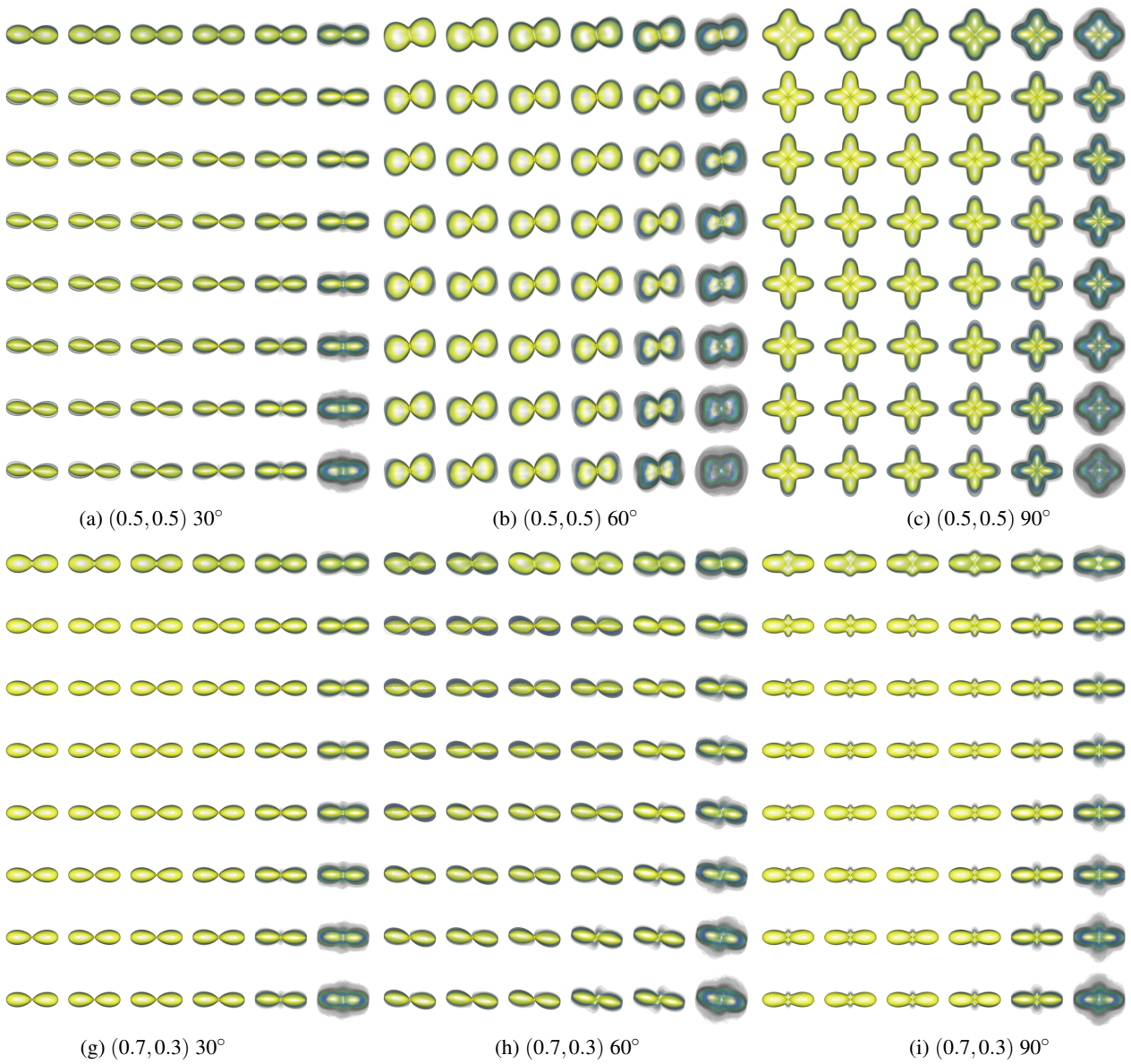
5 Results

5.1 Uncertainty in diffusion shapes generated from synthetic data

Using synthetic data, we demonstrate the usefulness of our diffusion uncertainty visualization by comparing other approaches, and show variation of SIPs of diffusion shapes for varying parameters.

Comparison to alternate visualizations. The most standard visualization of an ODF is hiding the uncertainty and showing the one best-fit diffusion shape. Not only does this not illustrate the uncertainty, but it gives the illusion that there is no uncertainty, which as demonstrated by other figures is a dangerous assumption.

Another approach is to treat each uncertain shape in the ensemble of uncertain shapes as a vector of data elements. Each element is represented as either each discrete diffusion direction, or as each voxel's binary value determined by inclusion in the diffusion shape. Then, given a set of 1000 diffusion shapes, we have a matrix of data, with each diffusion shape vector as one column. We can run PCA on this matrix and determine the most dominant modes of variation. From the visualization perspective, there are two clear disadvantages of this approach. First, it is difficult to visualize more



(a) $(0.5, 0.5)$ 30° (b) $(0.5, 0.5)$ 60° (c) $(0.5, 0.5)$ 90°

(g) $(0.7, 0.3)$ 30° (h) $(0.7, 0.3)$ 60° (i) $(0.7, 0.3)$ 90°

Figure 5: SIP functions of diffusion shapes for two fibers crossing at 30 (left), 60 (middle), and 90 (right) degrees. In (a), (b), and (c) the fibers have equal weight, and in (d), (e), and (f) they have relative weights of 0.7 and 0.3. Six levels of noise (SNR = $\{5, 10, 20, 30, 40, 50\}$ from right to left) each with 8 different b-values (= $\{1, 2, 3, 4, 5, 6, 7, 8\} \times 1000\text{s}/\text{mm}^2$ from top to bottom) were used in each subfigure.

than one principal component in a single glyph. The typical approach is to show the variation along the top several principal components in separate visuals of the same object, but in the context of diffusion shapes this is not an option since there are typically a collection of them (see for instance Figure 10). Additionally, displaying the variation along the top component is usually done by overlaying two (or more) images offsetting the mean shape $\hat{D}$ by 2 standard deviations in positive and negative direction of variation. These generated shapes may not actually be physically releasable configurations. For instance, if all of the variation comes from two modes, one standard deviation from the mean shape in each direction, then no configuration ever reaches two standard deviations away. Second, the first few principal components (let alone the top one) does not capture all of the complex variations of the diffusion shapes. Figure 6 shows the cumulative variation explained by the top $k = \{1, 2, \ldots, \}$ principal components. Only about 55% of the variation is shown by the top component, and it takes approximately six components to explain 90% of the variation. Intuitively, since the diffusion shape is described by $\nu = 15$ parameters, it is reasonable to expect 15 different components to capture the main sources of the variation. This partially agrees with one of the findings in [15], which reflects the complexity of the meaning of the first several major principal axises.

A third attempt at visualizing the diffusion uncertainty on a single glyph is to choose a single representative diffusion shape and use a color map to annotate the amount of uncertainty along each direction. A variation of this paradigm is present in the thesis of Hlawitschka [20]. This approach does not have any dramatic shortcomings, and we produce a variant of our own. We measure the variability of the diffusion shape between the 50% and 95% iso-levels of the ensemble of diffusion shapes and display this using a colormap on the 50% iso-level. This leaves the glyph a bit less cluttered, and provides a sense of how stable the 50% iso-level is with respect to the 95% iso-level. However, this approach (as well

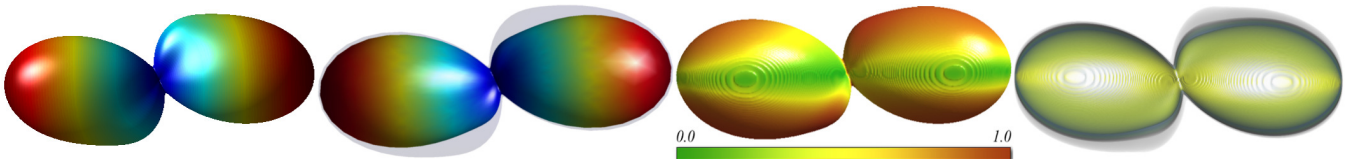


Figure 7: Different visualizations of diffusion uncertainty in a single glyph. Left to right: the single best-fit diffusion shape; two-standard deviation of top principal component in either direction superimposed; local variation color-mapped onto the 50% iso-level of diffusion shape; and SIP function of diffusion shapes with iso-levels at 25%, 50%, 75%, and 95%. The colormap for the left two images signifies ODF values, not uncertainty.

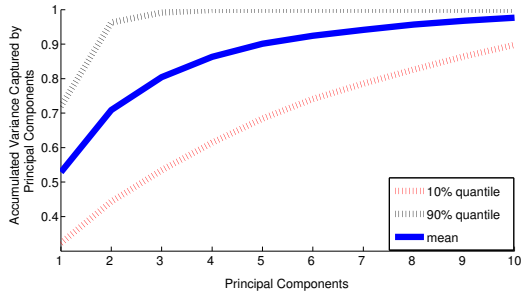


Figure 6: Distribution of cumulative variation explained by up to the first 10 principal components for cases shown in Figure 5.

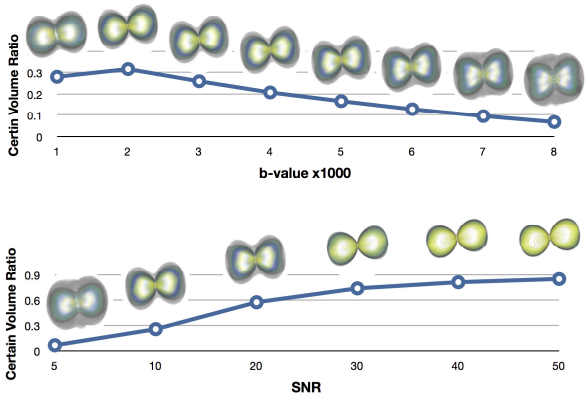


Figure 8: Plots of certain volume ratio as a function of b-value (top) and SNR (bottom). All experiments on synthetic example of two fibers crossing at 60° with equal relative weights and an SNR of 10 (top) or b-value of 3000 (bottom).

as other variants) fail to capture how the uncertainty changes (and robustness of iso-levels change) among all iso-levels. Just showing the standard deviation, or any local variability, does not capture the full distribution of variation in the same way our visualization does.

To compare these glyphs, we present them on the same data set in Figure 7. This simulates two crossing fibers at 60°, with relative weights of 30%, 70% using b-value of 3000 and SNR of 20. Much more nuanced and heteroscedastic uncertainty is presented in our visualization using a SIP function of diffusion shapes on far right.

Variation in parameters. To document the inferences available from the use of SIP functions of diffusion shapes, we show a variety of glyphs generated by varying the parameters to the synthetic data. We show either two crossing fibers at angles of either 30, 60, or 90 degrees, and with relative weights of either 50% and 50%, or of 30% and 70%. We then vary the manner in which the data is gathered by letting the b-value be from $\{1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000\}$, and the SNR be from $\{5, 10, 20, 30, 40, 50\}$; see Figure 5. For visualization and comparison purposes the uncertain diffusion shapes have been individually normalized so they are all approximately the same size. Figure 1 shows two enlarged SIP functions for diffusion shapes.

Immediately two possible conclusions could be drawn from this data, although further study should be performed before making any definitive claims. First, as the b-value increases, the variability and hence uncertainty increases and then decreases. It could be inferred that the best b-value to minimize uncertainty is around 2000 or 3000; or any value in that range. Second, as expected, as the SNR increases, the uncertainty decreases. Interestingly, not much is gained as the SNR goes above 40 (or even above 30 in some cases). This might suggests that DW images with an SNR of 40 are sufficient for successful HARDI analysis and visualization.

A more quantitative way to analyze the amount of uncertainty in the diffusion, based on the variation of input parameters, is to use a measure we call the *certain volume ratio*, which computes the ratio of the volume of the region within the 95% iso-level (the certain volume) to the area inside the 50% iso-level (a representation of the typical volume). Both parts of the ratio are robust and thus not sensitive to outliers in the generation of the diffusion shapes. The larger the ratio, the closer the 50% iso-level is to the 95% iso-level, indicating that most shapes are all very similar. Figures 8 and Figure 9 plot certain volume ratio as a function of b-value and SNR, respectively, and support our visual observations about the b-value and SNR. The certain volume ratio is plotted for the glyphs corresponding to those highlighted in Figure 5.

5.2 Uncertainty in diffusion shapes generated from human brain data

The human brain data was acquired using the parameters provided in Section 4.2. The SNR_0 values of the $b = 1000$, 2000 and 3000 s/mm^2 acquisitions, were estimated as 21.06, 19.76, and 19.48, respectively. These estimates were derived from the mean intensity of five different regions of interest (ROI) in the white matter and grey matter regions, divided by the variance of the pure background intensity of the $S0$ image, one ROI is shown in Figure 10. For the spherical deconvolution we have used the single fiber response kernel (e.g., [51]), where the kernel parameter was estimated from the data to compensate for the different b-values, as described in [47].

Both single triangular mesh diffusion shapes as well as the the SIP functions of diffusion shapes are shown. The extent of the uncertainty is shown clearly in the latter.

Our analysis based on the certain volume ratio (as shown in Figure 9) recommends a b-value of 2000. By visual inspection of the uncertain shapes in Figure 10, the data sensed with b-value 2000 seems to have the least uncertainty; a b-value of 1000 also seems adequate. Note that the regions with $b = 3000$ have been scaled small since the gray region with low probability is noticeably larger. This is supported by the certain volume ratios of 0.343, 0.367, and 0.283 for b-values of 1000, 2000, and 3000, respectively. Thus, our guidelines recommending b-value of about 2000 s/mm^2 for this noise level were observed to be useful.

Also notice the variability of the single best fit diffusion shape on the top of Figure 10 as the b-value changes. These, in principle, should represent the same diffusion functions, but due to real sensing errors, they have non-trivial variance. Our presented SIP

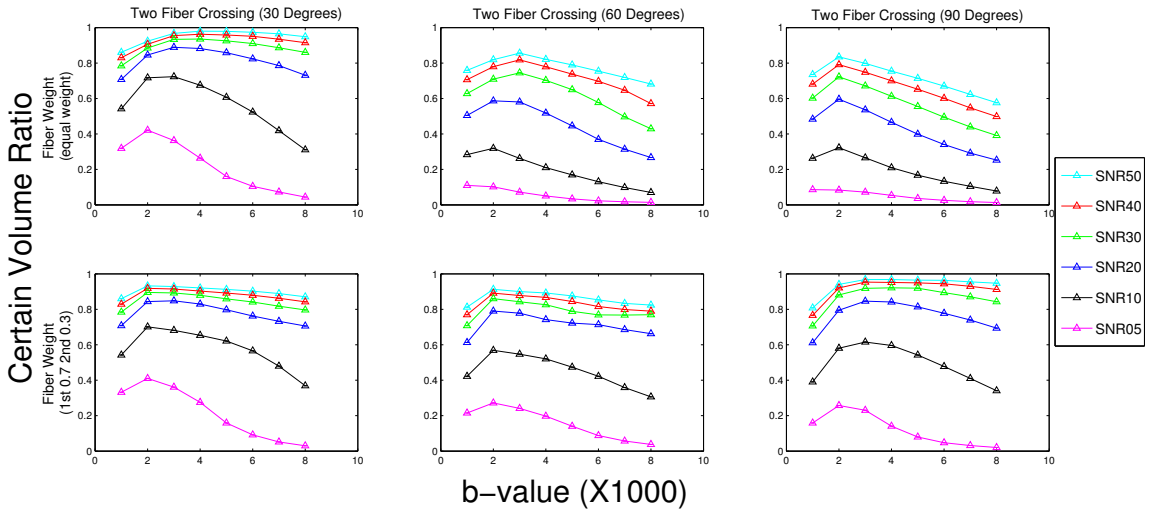


Figure 9: The certain volume ratio for two fiber crossing at 30 (left hand side), 60 (middle), and 90 (right hand side) degrees. Two groups of fiber weights are shown: 0.5 and 0.5 (top row) and 0.7 and 0.3 (bottom row). Six levels of noise with the $SNR_0$of 50,40,30,20,10, and 5 were applied.

functions of diffusion shapes below faithfully represent a realistic amount of uncertainty, preventing a user from drawing too strong a conclusion about the true diffusion patterns from this data.

6 Conclusions

Representing uncertainty in ODFs is a real and important challenge. In this work, we provide a general and elegant technique to visualize and quantify uncertainty in DW imaging. The key object of our technique is a SIP function generated from a large number of diffusion shapes, associated with different noise realizations of a given DW data. The SIP function is constructed by transforming a 3D ODF to a general volume rendering problem using a 1D transfer function to generate 5 layers separated by approximate iso-levels. We found that the SIP function provides a simple, clear and informative way to study uncertainty, especially in comparison with other techniques.

As future work we plan to elaborate this technique by using 3D glyphs that combine both diffusion profile information, as well as the fiber orientations and weights, accompanied by appropriate quantification measures. We hope to increase the admittedly slow construction time using a GPU, and plan to incorporate into an interactive visualization system. We hope this will help to understand better the impact of the different acquisition parameters, and consequently, to design successful HARDI acquisition protocols.

Acknoledgements. Thanks Theresa-Marie Thyne (color scheme analysis expertise), Kristin Potter and Joel Daniels (producing Fig 7(third-from-left)), Thomas Fogal and Liang Zhou (helpful discussions).

References

[1] H. E. Assemlal, D. Tschumperléa, and L. Brun. Efficient and robust computation of PDF features from diffusion MR signal. *Medical Image Analysis*, 13:715–729, 2009.

[2] A. Barmpoutis and B. C. Vemuri. A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. In *ISBI*, pages 1385–1388, 2010.

[3] P. Basser, J. Mattiello, and D. Lebihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B*, 103:247–254, 1994.

[4] P. Basser and S. Pajevic. Spectral decomposition of a 4th-order covariance tensor: Applications to diffusion tensor MRI. *Signal Processing*, 87:220–236, 2007.

[5] S. Basu, P. T. Fletcher, and R. Whitaker. Rician noise removal in diffusion tensor MRI. In *MICCAI*, LNCS, pages 117–125, 2006.

[6] T. E. J. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magn Reson Med*, 50(5):1077–1088, 2003.

[7] J. Berman, S. Chung, P. Mukherjee, C. Hess, E. Han, and R. Henry. Probabilistic streamline q-ball tractography using the residual bootstrap. *NeuroImage*, 35:215–222, 2008.

[8] K.-H. Cho, C. H. Yeh, J. D. Tournier, Y. P. Chao, J. H. Chen, and C. P. Lin. Evaluation of the accuracy and angular resolution of q-ball imaging. *NeuroImage*, 42(1):262–271, 2008.

[9] S. Chung, Y. Lu, and R. G. Henry. Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. *NeuroImage*, 33:531–541, 2006.

[10] C. Demiralp and D. Laidlaw. Tract-based probability densities of diffusivity measures in DT-MRI. In *MICCAI*, pages 542–549, Beijing, China, 2010.

[11] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical Q-ball imaging. *Magn. Res. Med.*, 58(3):497–510, 2007.

[12] H. Fonteijn, F. Verstraten, and D. Norris. Probabilistic inference on q-ball imaging data. *IEEE Trans. Med. Imaging*, 26:1515–1524, 2007.

[13] L. Frank. Anisotropy in high angular resolution diffusion-weighted MRI. *Magn. Reson. Med.*, 45:935–939, 2001.

[14] O. Friman and C. F. Westin. Uncertainty in white matter fiber tractography. In *MICCAI*, pages 107–114, 2005.

[15] J. E. I. Gonzalez, P. M. Thompson, A. Zhao, and Z. Tu. Modeling diffusion-weighted MRI as a spatially variant gaussian mixture: application to image denoising. *Medical Physics*, 38:4350–4364, 2011.

[16] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34(6):910–914, 1995.

[17] E. V. D. Hagen and R. Henkelman. Orientational diffusion reflects fiber structure within a voxel. *Magnetic Resonance in Medicine*, 48:454–459, 2002.

[18] H. K. Hahn, J. K. amd C. Nimsky, J. Rexilius, and H. Peitgen. Uncertainty in diffusion tensor based fibre tracking. *Medical Technologies in Neurosurgery*, 98:33–41, 2006.

[19] H. Haroon, D. Morris, K. Embleton, D. Alexander, and G. Parker. Using the model-based residual bootstrap to quantify uncertainty in fiber orientations from q-ball analysis. *IEEE Trans. Med. Imaging.*, 28:535–550, 2009.

[20] M. Hlawitschka. *Efficient Visualization of Tensor Fields with Application to Magnetic Resonance Data*. PhD thesis, Leipzig University, 2008.

[21] M. Hlawitschka and G. Scheuermann. Hot-lines: Tracking lines in higher order tensor fields. pages 27–34, 2005.

[22] T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn Reson Med*, 54:1480–1489, 2005.

[23] M. Irfanoglu, C. G. Koay, S. Pajevic, R. Machiraju, and P. Basser. Dif-

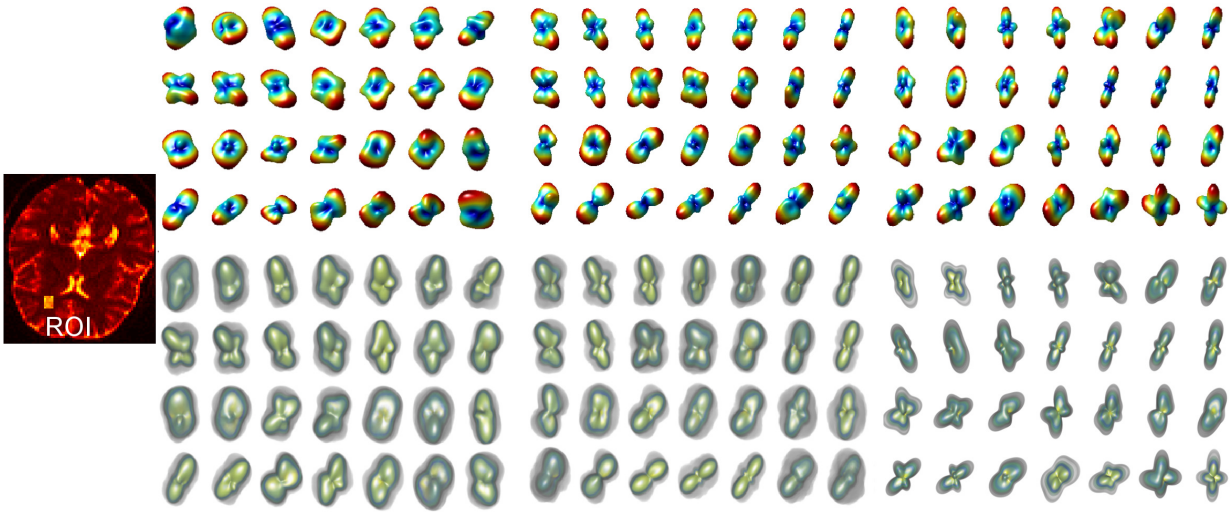


Figure 10: Left: 7 by 4 region of interest from the white matter. Right three columns: human data with a b-value of 1000, 2000, and 3000 s/mm2. Each column shows 28 diffusion shapes as a single triangular mesh (top) and a SIP function using 1000 instantiations of each diffusion shape.

fusion tensor field registration in the presence of uncertainty. *MICCAI*, 12:181–189, 2009.

[24] B. Jian, B. C. Vemuri, and E. Özarslan. A mixture of wisharts (mow) model for multifiber reconstruction. *Visualization and Processing of Tensor Fields*, 1:39–56, 2009.

[25] F. Jiao, Y. Gur, S. Joshi, and C. R. Johnson. Detection of crossing white matter fibers with high-order tensors and rank-k decompositions. *Information processing in medical imaging : proceedings of the conference*, 2011.

[26] F. Jiao, J. M. Phillips, J. Stinstra, J. Kruger, R. V. Kommaraju, E. Hsu, J. Korenberg, and C. R. Johnson. Metrics for uncertainty analysis and visualization of diffusion tensor images. In *MIAR*, pages 179–190, Bejing, China, 2010.

[27] D. Jones. Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49(1):7–12, 2003.

[28] D. K. Jones. Tractography gone wild: Probabilistic fibre tracking using the wild bootstrap with diffusion tensor MRI. *IEEE Trans. Med. Imaging*, 27(9):1268–1274, 2008.

[29] D. K. Jones and C. Pierpaoli. Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach. *Magn. Reson. Med.*, 53:1143–1149, 2005.

[30] D.-J. Kim, I.-Y. Kim, S.-O. Jeong, , and H.-J. Park. Evaluation of Bayesian tensor estimation using tensor coherence. *Physics in Medicine and Biology*, 54:3785–3802, 2009.

[31] G. Kindlmann. Superquadric tensor glyphs. In *Proceedings of IEEE TVCG/EG Symposium on Visualization*, pages 147–154, May 2004.

[32] D. Laidlaw, E. Ahrens, D. Kremers, M. Avalos, C. Readhead, and R. Jacobs. Visualizing diffusion tensor images of the mouse spinal cord. In *IEEE Visualization*, pages 127–134, October 1998.

[33] M. Lazara and A. L. Alexandera. Bootstrap white matter tractography (boot-trac). *NeuroImage*, 24:524–532, 2005.

[34] M. Löffler and J. M. Phillips. Shape fitting of point sets with probability distributions. In *Proceedings of 17th Annual European Symposium on Algorithms*, 2009.

[35] S. Pajevic and P. Basser. Parametric and non-parametric statistical analysis of DT-MRI data. *J. Magn. Reson.*, 161:1–14, 2003.

[36] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivière, A. Constantinesco, D. L. Bihan, and J. Mangin. Fiber tracking in q-ball fields using regularized particle trajectories. 19:52–63, 2005.

[37] C. Pierpaoli and P. Basser. Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine*, 36:893–906, 1996.

[38] V. Prckovska, A. F. Roebroeck, W. Pullens, A. Vilanova, and B. M. T. H. Romeny. Optimal acquisition schemes in high angular resolution diffusion weighted imaging. *LNCS*, 5242(vol. 2):9–17, 2008.

[39] L. Qi, G. Yu, and E. Wu. Higher order positive semidefinite diffusion tensor imaging. *SIAM J. Imaging Sci.*, 3:416–433, 2010.

[40] T.-M. Rhyne. Exploring visualization theory. *IEEE Computer Graphics and Applications*, 31(3):6–7, 2011.

[41] P. Risholm, S. Pieper, E. Samset, and W. M. W. 3rd. Summarizing and visualizing uncertainty in non-rigid registration. In *MICCAI*, pages 554–561, Beijing, China, 2010.

[42] T. Schultz and G. Kindlmann. A maximum enhancing higher-order tensor glyph. *Computer Graphics Forum*, 29(3):1143–1152, 2010.

[43] T. Schultz and H. P. Seidel. Estimating crossing fibers: A tensor decomposition approach. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1635–1642, 2008.

[44] T. Schultz, C.-F. Westin, and G. Kindlmann. Multi-diffusion-tensor fitting via spherical deconvolution: A unifying framework. In T. Jiang, N. Navab, J. Pluim, and M. Viergever, editors, *MICCAI*, volume 6361 of *LNCS*, pages 673–680, 2010.

[45] A. Schwartzman. Random ellipsoids and false discovery rates: Statistics for diffusion tensor imaging data. June 2006. PhD dissertation.

[46] J. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23:1176–1185, 2004.

[47] J.-D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472, 2007.

[48] D. Tuch. *Diffusion MRI of Complex Tissue Structure*, January 2002.

[49] D. S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52:1358–1372, 2004.

[50] V. Vapnik and A. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *The. of Prob. App.*, 16:264–280, 1971.

[51] Y. Weldeselassie, A. Barmpoutis, and M. Atkins. Symmetric positive-definite Cartesian tensor orientation distribution functions (CT-ODF). In *MICCAI*, pages 582–589, Beijing, China, 2010.

[52] B. Whitcher, D. S. Tuch, and L. Wang. The wild bootstrap to quantify variability in diffusion tensor MRI. In *13th Annual Meeting of ISMRM*, page 1333, Miami, 2005.

[53] Y. Xie, B. C. Vemuri, and J. Ho. Statistical analysis of tensor fields. volume 6361 of *LNCS*, pages 682–689, 2010.

[54] A. Zalesky. DT-MRI fiber tracking: a shortest paths approach. *IEEE Trans Med Imaging*, 27:1458–1471, 2008.

[55] L. Zhan, A. D. Leow, N. Jahanshad, M. C. Chiang, M. Barysheva, A. D. Lee, A. Toga, K. McMahon, G. de Zubicaray, M. Wright, and P. M. Thompson. How does angular resolution affect diffusion imaging measures. *NeuroImage*, 49:1357–1371, 2010.