VARIATIONAL DENOISING OF DIFFUSION WEIGHTED MRI

TIM MCGRAW

West Virginia University
Morgantown, WV 26506, USA

BABA VEMURI

University of Florida
Gainesville, FL 32601, USA

EVREN ÖZARSLAN

National Institutes of Health
Bethesda, MD 20892, USA

YUNMEI CHEN

University of Florida
Gainesville, FL 32601, USA

THOMAS MARECI

University of Florida
Gainesville, FL 32601, USA

(Communicated by the associate editor name)

ABSTRACT. In this paper, we present a novel variational formulation for restoring high angular resolution diffusion imaging (HARDI) data. The restoration formulation involves smoothing signal measurements over the spherical domain and across the 3D image lattice. The regularization across the lattice is achieved using a total variation (TV) norm based scheme, while the finite element method (FEM) was employed to smooth the data on the sphere at each lattice point using first and second order smoothness constraints. Examples are presented to show the performance of the HARDI data restoration scheme and its effect on fiber direction computation on synthetic data, as well as on real data sets collected from excised rat brain and spinal cord.

1. **Introduction.** Observing the directional dependence of water diffusion in the nervous system can allow us to infer structural information about the surrounding tissue. Axonal membranes and myelin sheath present a barrier to molecules diffusing in directions perpendicular to the white matter fiber bundles whereas in directions parallel to the fibers, the diffusion process is less restricted [10]. This results in anisotropic diffusion that can be observed using magnetic resonance (MR) measurements by the utilization of magnetic field gradients [47]. In general, the acquired MR signal depends on the strength and the direction of these diffusion sensitizing gradients. Repeated measurements of water diffusion in tissue with varying

---

gradient directions provide a means to quantify the level of anisotropy as well as to determine the local fiber orientation within the tissue.

In a series of publications, Basser and colleagues [6, 7, 8] have formulated an imaging modality called "diffusion tensor MRI (DT-MRI or DTI)" that employs a second order, positive definite, symmetric diffusion tensor to represent the local tissue structure. They have proposed several rotationally invariant scalar indices that quantify different aspects of water diffusion observed in tissue, similar to different "stains" used in histological studies [4]. Under the hypothesis that the preferred orientations of water diffusion will coincide with the fiber directions, one can determine the directionality of neuronal fiber bundles. This fact has been exploited to generate fiber-tract maps that yield information on structural connections in human [8, 34, 38, 22] as well as rat brains [42, 63, 55, 40, 39] and spinal cords [54].
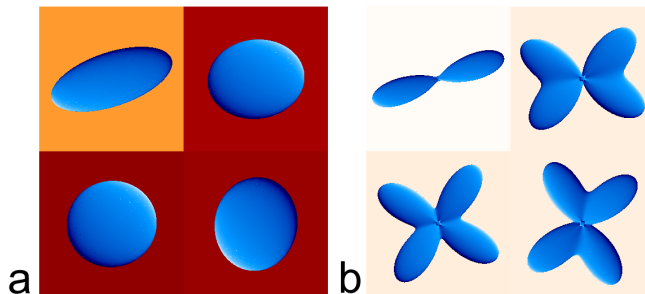

FIGURE 1. The effect of fiber orientation heterogeneity on diffusion MR measurements. (a) Isosurfaces of the Gaussian probability maps assumed by DTI overlaid on fractional anisotropy maps computed from the diffusion tensors. (b) Probability profiles computed using the diffusion orientation transform (DOT) from HARDI data overlaid on generalized anisotropy (GA) maps. Both schemes perform well when there is only one orientation (top left portions of both panels). HARDI based method is able to resolve fiber crossings whereas DTI yields an averaged profile.

Despite its apparent success, DT-MRI has significant shortcomings when the tissue of interest has a complicated geometry. This is due to the relatively simple tensor model that assumes a unidirectional —if not isotropic— local structure. In the case of orientational heterogeneity, DT-MRI technique is likely to yield incorrect fiber directions, and artificially low anisotropy values. This is due to the Gaussian model implicit in DTI that allows only one preferred direction for water diffusion. In order to overcome these difficulties several approaches have been taken. Q-space imaging, a technique commonly used to examine porous structures [13], has been suggested as a possible solution [59]. However this scheme requires strong gradient strengths and long acquisition times [5], or significant reduction in the resolution of the images. Q-space imaging requires many images to be acquired since the space of diffusion encoding gradients is sampled on a 3D lattice. As a more viable alternative Tuch et al. have proposed to do the acquisition such that the diffusion sensitizing gradients sample the surface of a sphere [53, 52]. In this high angular resolution diffusion imaging (HARDI) method, one does not have to be restricted to the tensor model and instead, it is possible to calculate diffusion

coefficients along many directions. This method *does not* require more powerful hardware systems than those required by DT-MRI. Several groups have already performed HARDI acquisitions in clinical settings and have reported 43 to 126 different diffusion weighted images acquired in 20 to 40 minutes of total scanning time [30, 52, 33] indicating the feasibility of the high angular resolution scheme as a clinical diagnostic tool.

In Figure 1, we present a matrix of simulated voxels showing renderings of DTI-based estimates of orientation and HARDI-based orientation estimates computed using the scheme we present in Section 2.1. The orientation heterogeneity is evident from the HARDI-based renderings at each voxel since HARDI measurements can resolve multiple dominant directions of molecular diffusion in a voxel, a lacking feature of DTI. Since the HARDI data acquisition is very nascent, not many techniques of processing the HARDI data have been reported in literature. In the following section we will review the recently reported techniques of HARDI data denoising, which may be done prior to further analysis or visualization.

1.1. **Review.** We will first briefly describe the physics of acquisition and then point to various recent restoration techniques followed by methods for computing anisotropy measures from HARDI. This will be followed by an overview of our method.

1.1.1. *Physics of Diffusion MR and HARDI Acquisition.* The random process of diffusion of water molecules is described by the diffusion displacement PDF $p_t(\mathbf{r})$. This is the probability that a given molecule has a diffusion displacement of $\mathbf{r}$ after time t. The relation between the measured MR signal, and the diffusion displacement PDF is given by [13]

$$p_t(\mathbf{r}) = \int \frac{S(\mathbf{q})}{S_0} \, \exp(-2\pi \, i \, \mathbf{q} \cdot \mathbf{r}) \, d\mathbf{q} \, , \tag{1}$$

where $S(\mathbf{q})$ is the MR signal when a diffusion gradient pulse of strength $\mathbf{G}$ and duration δ is applied yielding the wave vector $\mathbf{q} = \gamma \delta \mathbf{G}$ where γ is the gyromagnetic ratio for protons. S_0 is the image acquired with no diffusion encoding gradient applied. The above formula indicates that water displacement probabilities are simply the Fourier transform of $S(\mathbf{q})/S_0$. It is the orientational modes of $p_t(\mathbf{r})$ that are taken to be the underlying fiber directions.

The HARDI processing proceeds by acquiring diffusion weighted images with many diffusion encoding gradient directions, effectively sampling a spherical shell of the q-space (the space of diffusion encoding gradients) as described by Tuch [51]. It is desired that this sampling minimize the average angle between gradient directions so that the diffusion signal may be accurately reconstructed. The gradient direction for each image has been chosen to correspond to the vertices of an icosahedron which has been repeatedly subdivided. Our data sets include diffusion-weighted images acquired with the application of diffusion gradients along 81 or 46 directions in addition to one image with no diffusion weighting. Since the process of diffusion is known to have antipodal symmetry [30], we need to sample only one of the hemispheres in q-space.

1.1.2. *Restoration.* Processing of HARDI data sets has received increased attention lately and a few researchers have reported their results in literature. The use of spherical harmonic expansions have been quite popular in this context since the HARDI data primarily consists of scalar signal measurements on a sphere located

at each lattice point on a 3D image grid. Tuch et al. [53, 52] developed the HARDI acquisition and processing and later Frank [30] showed that it is possible to use the spherical harmonics expansion of the HARDI data to characterize the local geometry of the diffusivity profiles. Although elimination of odd-ordered terms and the truncation of the Laplace series provide some level of smoothing, there is no discussion of smoothing the data across the lattice points. Chen et al. [20] find a regularized spherical harmonic expansion by solving a constrained minimization problem. However the expansion is a truncated spherical harmonic expansion of order four, restricting the level of complexity that can be modeled using this approach. In [33], Jansons and Alexander described a new statistic, persistent angular structure, which was computed from the samples of a 3D function. In this case, the function described displacement of water molecules in each direction. The goal in their work was to resolve voxels containing one or more fibers. However, there was no discussion on how to restore the noisy HARDI data prior to resolution of the fiber paths. More recently, Descoteaux et al., [23, 24], proposed an analytical solution to the reconstruction of the diffusion orientation distribution function (ODF). They model the signal using a spherical harmonic function of order eight and fit this model to the noisy data using a regularization constraint involving the Laplace-Beltrami operator for smoothing the HARDI data over the sphere of directions at each voxel. Their analytic form for the ODF reconstruction requires a numerical solution to a linear system and they do not consider regularization across the 3D lattice which can be important in order to obtain a piecewise smooth representation of the given HARDI data. Wiest-Daessle et al. [62, 61] described several variants of non-local mean denoising applied to diffusion MRI. The approach which is applicable to HARDI involved considering the dataset as a vector-valued image, however this approach does not respect the directional relationship among the images. Assemlal et al. also employ only spatial regularization approaches to robustly determine the diffusion ODF [1] and PDF [2] fields. Savadjiev et al. [46] formulate a novel spatial regularization in terms of the underlying 3D curves which represent neuronal fibers.

In contrast to HARDI denoising, DT-MRI denoising has been more popular and numerous techniques exist in literature. For sampling of the techniques used to denoise DT-MRI, we refer the reader to [50, 17, 60, 57, 58, 19, 27, 9, 28, 3, 32]. Most of these works use a linearized Stejskal-Tanner equation [47] describing the MR signal decay with the exception of Wang et al., [57, 58]. Using the Stejskal-Tanner equation as is, is quite important in preserving the accuracy of the restored data and this was shown in the experiments in [58]. Another important constraint in the DTI restoration is the positive definiteness of the tensors, in this context, work in [18] introduced an elegant differential geometric framework to achieve the solution. The work in [57, 58] and [60] chose alternative methods to impose the positive definiteness of the restored tensor fields namely, a linear algebraic and a PDE-based method respectively. Approaches to filtering based on the Riemannian geometry of the manifold of symmetric positive-definite matrices have been reported [31, 14].

1.2. **Overview of Our Modeling Scheme.** In this section we present a novel and effective variational formulation that will directly estimate a smooth signal $S(\theta, \phi)$ and the probability distribution of the water molecule displacement over all

directions $p(\theta, \phi)$, given the noisy measurement

$$\hat{S}(\theta, \phi) = S_0 \exp(-bD(\theta, \phi)) + \eta(\theta, \phi) \,, \qquad (2)$$

where $\hat{S}$ is the signal measurement taken on a sphere of constant gradient magnitude over all (θ, ϕ), b is the diffusion weighting factor, $D(\theta, \phi)$ is the apparent diffusivity as a function of the direction expressed by the polar and azimuthal angles on the sphere and $\eta(\theta, \phi)$ is Rician noise. The noise is due to additive Gaussian noise corrupting the complex-valued k-space measurements. However, for high signal-to-noise ratios we may consider η to be Gaussian distributed. A variational formulation for denoising using a data constraint based on the Rician likelihood was given by Basu et al. [9]. However, this leads to a highly nonlinear evolution equation since it involves the ratio of two Bessel functions. A modification to the non-local means algorithm which can handle Rician noise was presented by Descoteaux et al. [25]. However, neither of these approaches address smoothing over the spherical domain. In contrast, Clarke et al. [21] propose a robust method for estimating fiber orientation distributions in the presence of Rician noise, but they do not consider smoothness constraints over the voxel lattice. In this work we will assume a high SNR so that the Gaussian additive noise is a good assumption. Since we are performing high-field ex-vivo experiments, we can acquire many images and use averaging to increase the SNR so that this assumption is valid.

The variational principle involves smoothing S values over the sphere and across the 3D image lattice. The key factor that complicates this problem is that the domain of the data at each voxel in the lattice is a sphere. One may use the level-set techniques developed by Tang et al., [48] to achieve this smoothing. However, when data sets are large, it becomes computationally impractical to apply the level-set technique at each voxel independently to restore these scalar-valued measurements on the sphere. Alternative approaches to solving variational problems over non-planar domains have been described in recent literature. Cecil et al. [15] propose several numerical approaches to dealing with discontinuous derivatives due to periodic boundaries encountered when solving problems on S^1 and S^2. Liu et al. [37] proceed by finding a conformal mapping from the surface to the plane, then solving the problem in the 2D parameter space. Bogdanova et al. [12] presented explicit formulations of differential operators on parametric surfaces in terms of the Riemannian metric. Since our input data are sparsely distributed over a triangulated sphere (gradient directions are computed by subdividing an icosahedron, we simply use the spherical triangles as our computational domain. We arrive at a computationally efficient solution to this problem by using the finite element method (FEM) on the sphere and choosing local basis functions for the data restoration. Unlike the reported work on spherical harmonic basis expansion of the diffusivity function on the sphere [29, 43, 20], the FEM basis functions have local support and are more stable to perturbations due to noise in the data. From the denoised data we will compute a probability, $p_t(\theta, \phi)$, of molecular diffusion over a sphere of directions.

The rest of this paper is organized as follows: Section (2) contains a variational formulation of the HARDI denoising problem including smoothing the scalar signal over a sphere of directions at each 3D lattice point and across lattice points, computation of probability of water molecular diffusion over the sphere of directions and several measures of anisotropy computed from the field of probability densities. In section (3), we present several experimental results depicting the performance of our algorithms on synthetically generated and real data sets. Finally, we conclude

in section ([4](#)). Appendices A and B contain the details of the finite element basis used and the element as well as the global equations.

2. **Formulation of the HARDI restoration.** Normally, the diffusion weighted images are quite noisy especially when acquired using large field gradients. One can reduce some amount of noise by signal averaging for each gradient direction used. However, this by itself does not preserve the details in the data. We now present a variational formulation for effective denoising of the HARDI data.

2.1. **Variational Smoothing.** We propose a membrane-spline deformation energy minimization for smoothing the measured image $\hat{\mathbf{S}}(\mathbf{x}, \theta, \phi)$. The variational principle for estimating a smooth $\mathbf{S}(\mathbf{x}, \theta, \phi)$ is given by

$$\min_{\mathbf{S}} E(\mathbf{S}) = \frac{\mu}{2} \int_{\Omega} \int_{S^2} |\mathbf{S}(\mathbf{x}, \theta, \phi) - \hat{\mathbf{S}}(\mathbf{x}, \theta, \phi)|^2 dS\, d\mathbf{x}$$

$$+ \int_{S^2} \|\nabla_{(\theta,\phi)}\mathbf{S}\|^2 dS\, d\mathbf{x} + \int_{\Omega} g(\mathbf{x})\|\nabla_x \mathbf{S}\| d\mathbf{x} \tag{3}$$

where Ω is the domain of the image lattice and S^2 is the sphere on which the signal measurements are specified at each voxel. The first term of Equation ([3](#)) is a data fidelity term which makes the solution to be close to the given data. The degree of data fidelity can be controlled by the input parameter μ. The second term is a regularization constraint enforcing smoothness of the data over the spherical domain at each voxel. The minimizer of this energy term is a membrane spline over the sphere which is in Sobolev space $H^1(S^2)$ [49]. The third term is another regularization term which causes the solution to be piecewise smooth over the spatial domain (the 3D voxel lattice). The minimizer of this TV norm is in the space $BV(R^3)$, functions of bounded variation [35]. $g(\mathbf{x})$ inhibits smoothing across discontinuities in S over the lattice. More on this in section ([2.3](#)). The choice of membrane spline smoothness over S^2 is motivated by the partial volume effect in MRI. The signal at each voxel is the average over a volume much larger than a single axonal fiber. Within this volume there may be fibers of varying orientation and regions of isotropy. Though the diffusivity function may be nearly discontinuous over S^2 at a point near a fiber bundle, it is highly unlikely for the volume average to be so. For this reason, we do not use TV norm minimization over the spherical domain.

2.2. **Finite Element Method based smoothing of $\mathbf{S}(\theta, \phi)$.** We will consider a deformation energy functional which is a weighted combination of the thin-plate spline energy and the membrane spline energy, which is commonly used in computer vision literature for smoothing scalar-valued data in $\Re^3$ (see McInerney et al., [41], Lai et al., [36]). In our case, the data at each voxel is an image on the sphere, $\mathbf{S}(\theta, \phi)$, so the problem is inherently 2 dimensional.

The diffusion-encoding gradient directions are taken as the vertices of a subdivided icosahedron, to achieve a nearly uniform sampling of gradient directions over the sphere. We map this piecewise planar approximation of the sphere to the global FEM coordinate system (u, v) by setting $(u = \theta, v = \phi)$ for each gradient direction. This domain is triangulated so each face of the subdivided icosahedron will have a corresponding triangle in the (u, v) domain. A periodic boundary condition is imposed so that $S(2\pi, v) = S(0, v)$. The area element in the (u, v) domain is $du\, dv = \sin\phi d\theta d\phi$. A similar mapping was used by McInerney & Terzopoulos [41] and Vemuri and Guo [56] for finite elements over a spherical domain.

Note that, after mapping, the data can be seen as a height field over the (u, v) plane. The smoothness of the height function, $z(u, v)$, will be enforced by the smoothing functional

$$E_p = \int \int_\Omega (\alpha(|z_u|^2 + |z_v|^2) + \beta(|z_{uu}|^2 + 2|z_{uv}|^2 + |z_{vv}|^2))du\, dv. \qquad (4)$$

The weight on the membrane term is α and the weight on the thin-plate term is β. Once we have computed a smooth $z(u, v)$, the result will then be mapped back to the image on the sphere, $S(\theta, \phi)$.

The data energy due to virtual work of the data forces, f, and virtual displacement, z, is

$$E_d = -\int \int_\Omega z(u, v) f(u, v) du\, dv. \qquad (5)$$

By the principal of virtual work, the spline system is in equilibrium when the total work done by all forces is zero for all virtual displacements.

The restoration of $S(\theta, \phi)$ at each voxel is formulated as the energy minimization

$$\min_S E(S) = \min_S (E_p(S) + E_d(S)), \qquad (6)$$

with $\nabla E_p(S) = -\nabla E_d(S)$ defining the equilibrium condition of the system.

We use polynomial shape functions, N_i, as a basis for the unknown smooth approximation, z, of the data over the u, v plane. We may write z as

$$z(u, v) = \sum_{i=1}^n q_i N_i(u, v) = \mathbf{N}\mathbf{q} \qquad (7)$$

where $\mathbf{N}$ is a $(1 \times n)$ row vector, and $\mathbf{q}$ is a column vector of nodal variables.

The domain, Ω, is partitioned into triangular elements, Ω_j, each with their own local shape functions. The shape functions, in terms of local (barycentric) coordinates are given in Appendix A. For each element j, we have,

$$z(u, v) = \mathbf{N}^j(u, v)\mathbf{q}^j \qquad (8)$$

for $(u, v) \in \Omega_j$. In the rest of this section we will derive linear equations for the element potential energies, E_p^j, and data energies, E_d^j, in terms of the coefficients $\mathbf{q}^j$. Finally, we will assemble a global linear system, and solve for $\mathbf{q}$. This will allow us to evaluate $z(u, v)$ using Equation 7.

The global potential energy is the sum of the energies of each finite element,

$$E_p = \sum_j E_p^j \qquad (9)$$

where the local potential energy function for each element is given by

$$E_p^j = \int \int_{\Omega_j} (\alpha|z_u^j|^2 + \alpha|z_v^j|^2 + \beta|z_{uu}^j|^2 + 2\beta|z_{uv}^j|^2 + \beta|z_{vv}^j|^2)du\, dv. \qquad (10)$$

The element strain vector (given by Dhatt and Touzot [26]) is

$$\epsilon^j = \begin{bmatrix} z_u^j \\ z_v^j \\ z_{uu}^j \\ z_{uv}^j \\ z_{vv}^j \end{bmatrix} \qquad (11)$$
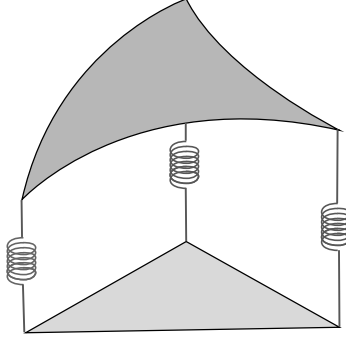
FIGURE 2. Data forces are applied at each vertex in the triangulated domain.

which may be rewritten as

$$\epsilon^j = \begin{bmatrix} (N_1)_u & \cdots & (N_n)_u \\ (N_1)_v & \cdots & (N_n)_v \\ (N_1)_{uu} & \cdots & (N_n)_{uu} \\ (N_1)_{uv} & \cdots & (N_n)_{uv} \\ (N_1)_{vv} & \cdots & (N_n)_{vv} \end{bmatrix} \mathbf{q}^j = \mathbf{B}\mathbf{q}^j \tag{12}$$

where we have defined $\mathbf{B}$ as the $(5 \times n)$ matrix of derivatives of the nodal basis functions. We can then rewrite the element potential (strain) energy as

$$E_p^j = \int\int_{\Omega_j} \epsilon^{jT} \mathbf{D} \epsilon^j \, du \, dv \tag{13}$$

where we define

$$\mathbf{D} = \begin{bmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 2\beta & 0 \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix}, \tag{14}$$

the diagonal matrix containing the membrane and thin-plate spline weighting factors. We have the option of finding solutions in the space H^1 by setting $\beta = 0$, or in H^2 by making $\beta > 0$. In general, the parameter values depend on the angular resolution of the underlying signal. Making the values too high may smooth out salient details, and setting the values too low may result in fitting the spline to the noise. In practice we determine the values empirically by processing synthetic datasets.

Since $\mathbf{q}^j$ is constant over each element we can derive the element stiffness matrix, $\mathbf{K}$, in terms of $\mathbf{D}$ and $\mathbf{B}$ giving us the element strain energy as,

$$E_p^j = \int\int_{\Omega_j} \mathbf{q}^{jT} \mathbf{B}^{jT} \mathbf{D} \mathbf{B}^j \mathbf{q}^j \, du \, dv = \mathbf{q}^{jT} \mathbf{K}^j \mathbf{q}^j. \tag{15}$$

We will model the data constraint as springs pulling $z(u,v)$ toward the measured values $z_0(u,v)$, as illustrated in Figure(2). The force at each point will obey $f = k(z - z_0)$, where k is the spring constant. For small displacements the spring constant, $k = \frac{\mu}{2}$ where μ is the data constraint coefficient from Equation (3).

The element deformation energy due to virtual displacement $z(u, v)$ is given by

$$E_d^j = -\int\int_{\Omega_j} \mathbf{N}^j \mathbf{q}^j k(\mathbf{N}^j \mathbf{q}^j - z_0)\, du\, dv. \tag{16}$$

We can split the deformation energy into two terms : $E_d^j = \mathbf{q}^{jT}(\mathbf{F}^j \mathbf{q}^j - \mathbf{f}^j)$ by defining

$$\mathbf{F}^j = -k\int\int_{\Omega_j} \mathbf{N}^{jT}\mathbf{N}^j\, du\, dv \tag{17}$$

and

$$\mathbf{f}^j = k\int\int_{\Omega_j} \mathbf{N}^{jT} z_0\, du\, dv. \tag{18}$$

We may now balance the deformation energy and data energy by solving the following linear system:

$$(\mathbf{K}^j + \mathbf{F}^j)\mathbf{q}^j = \mathbf{f}^j. \tag{19}$$

The global linear system for smoothing the entire mesh may be obtained by appropriately summing the local element matrices, as detailed in Appendix B. The global system is symmetric, and has a sparse banded structure with 18 nonzero diagonal bands. Since the global matrix is positive-definite, an efficient solution to $\mathbf{q}$ is obtained via Cholesky factorization.

2.3. **Spatial smoothing of $S(x)$.** We are now ready to describe the smoothing of the data across the 3D lattice. There are many existing methods that one can apply to this problem as discussed earlier. Smoothing the raw vector-valued data, $S(x)$, is posed as a variational principle involving a first order smoothness constraint on the solution to the smoothing problem. Note that the data at each voxel are m measurements of S over a sphere of directions and can be assembled into a vector after the smoothing on the spherical coordinate domain has been accomplished. We propose a weighted TV-norm minimization for smoothing this vector-valued image $\mathbf{S}$. This smoothing scheme reduces the effect of inter-region blurring, a drawback Gaussian convolution and isotropic diffusion suffer. Our method is a modified version of the work in Blomgren et. al., [11]. *The novelty here lies in the choice of the weighting i.e., the coupling term between the channels.* The variational principle for estimating a smooth $\mathbf{S}(\mathbf{x})$ is given by

$$\min_{\mathbf{S}} E(\mathbf{S}) = \int_\Omega (g(\mathbf{x})\sum_{i=1}^m \|\nabla S_i\| + \frac{\mu}{2}\sum_{i=1}^m (S_i - \hat{S}_i)^2)d\mathbf{x} \tag{20}$$

where, Ω is the image domain, μ is a regularization factor and m is the number of images. The first term here is the regularization constraint on the solution to have a certain degree of smoothness. The second term in the variational principle makes the solution faithful to the data in the L_2 sense. We have used the coupling function $g(\mathbf{x}) = 1/(1 + \|\nabla GA(\mathbf{x})\|^2)$ for smoothing HARDI, where GA is the generalized anisotropy index defined in Özarslan et al., [45] and is computed from the variance of normalized diffusivity. For a more detailed discussion on GA, we refer the reader to [45]. This selection criterion preserves edges in anisotropy while smoothing the rest of the data. This anisotropy measure is chosen since it can be computed without explicitly computing the ODF, and it is our goal to smooth the data prior to ODF computation. An image of the coupling term for a typical slice is shown in Figure (3).
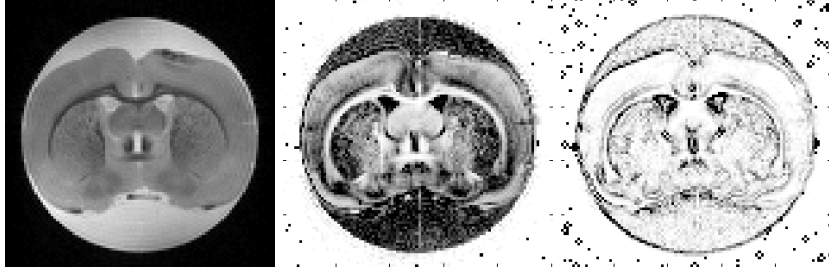
FIGURE 3. S_0 (left), GA (center), and coupling factor g (right).

Here we have used a different TV-norm than the one used by Blomgren and Chan [11]. The $TV_{n,m}$ norm is an L_2 norm of the vector of $TV_{n,1}$ norms $(\int_\Omega \|\nabla S_i(x)\|^2 dx)$ for each channel. We use the L_1 norm instead, which is known to have better discontinuity preservation properties.

The gradient descent form of the above minimization is given by

$$\frac{\partial S_i}{\partial t} = \text{div}\left(\frac{g\nabla S_i}{\|\nabla S_i\|}\right) - \mu(S_i - \hat{S}_i) \quad i = 1, ..., m$$

$$\frac{\partial S_i}{\partial n}|_{\partial\Omega \times \mathbb{R}^+} = 0 \ \ and \ \ \mathbf{S}(\mathbf{x}, t = 0) = \hat{\mathbf{S}}(\mathbf{x}) \tag{21}$$

The use of a modified TV-norm in equation (20) results in a looser coupling between channels than when using the $TV_{n,m}$ norm. This reduces the numerical complexity of Equation (21) and makes solution for large data sets feasible.

The gradient descent of the vector-valued image smoothing using the $TV_{n,m}$-norm $TV_{n,m}(\mathbf{S}(\mathbf{x})) = \sqrt{\sum_{i=1}^{m}[TV_{n,1}(S_i)]^2}$ presented in [11] is given by,

$$\frac{\partial S_i(x,t)}{\partial t} = \frac{TV_{n,1}(S_i)}{TV_{n,m}(\mathbf{S})}\nabla \cdot \left(\frac{\nabla S_i}{\|\nabla S_i\|}\right)$$

$$\mathbf{S}(x, 0) = \mathbf{S}_0(x). \tag{22}$$

Note that the $TV_{n,m}$ norm appears in the gradient descent solution of the vector-valued minimization problem. Considering that our data sets consist of up to 82 images, corresponding to (magnetic field) gradient directions, calculating the $TV_{n,m}$ norm by numerically integrating over the 3-dimensional data set at each step of an iterative process would be prohibitively expensive. In contrast using our modified TV-norm described earlier leads to a more efficient solution. We are now ready to present the numerical solution to equation (21).

2.3.1. *Fixed-Point Lagged-Diffusivity.* Since the m Equations(21) are coupled only through the function g, we can drop the subscript on S with no ambiguity (later the subscript will refer to spatial coordinates.) In this section we will discuss the numerical solution for each channel, S, of the vector-valued image $\mathbf{S}$. Equation (21) is nonlinear due to the presence of $\|\nabla S_i\|$ in the denominator of the first term. We linearize Equation (21) by using the method of "lagged-diffusivity" presented by Chan and Mulet [16]. By considering $\|\nabla S\|$ to be a constant for each iteration, and using the value from the previous iteration we can instead solve

$$-\frac{1}{\|\nabla S^t\|}(\nabla g \cdot \nabla S^t + g\nabla^2 S^{t+1}) + \mu(S^{t+1} - S^0) = 0 \tag{23}$$

Here the superscript denotes iteration number. Equation (23) can be recast in the form

$$-\nabla^2 S^{t+1} + \frac{\mu \|\nabla S^t\|}{g} S^{t+1} = \frac{\mu \|\nabla S^t\| S^0 + \nabla g \cdot \nabla S^t}{g}. \tag{24}$$

We now discretize the above equation in the following.

2.3.2. *Discretized Equations.* To write Equation (24) as a linear system ($A\mathbf{S}^{t+1} = f^t$), we discretize the Laplacian and gradient terms. Using central differences for the Laplacian we have

$$\begin{aligned} \nabla^2 S^{t+1} &= S^{t+1}_{x-1,y,z} + S^{t+1}_{x,y-1,z} + S^{t+1}_{x,y,z-1} \\ &- 6S^{t+1}_{x,y,z} + S^{t+1}_{x+1,y,z} + S^{t+1}_{x,y+1,z} + S^{t+1}_{x,y,z+1} \end{aligned} \tag{25}$$

We define the standard central differences to be

$$\Delta_x S = \frac{1}{2}(S_{x+1,y,z} - S_{x-1,y,z})$$

$$\Delta_y S = \frac{1}{2}(S_{x,y+1,z} - S_{x,y-1,z})$$

$$\Delta_z S = \frac{1}{2}(S_{x,y,z+1} - S_{x,y,z-1}) \tag{26}$$

We can rewrite Equation (24) in discrete form using the definitions in Equation (26)

$$\begin{aligned} &-S_{x-1,y,z} - S_{x,y-1,z} - S_{x,y,z-1} \\ &+(6 + \frac{\mu\sqrt{(\Delta_x S^t)^2 + (\Delta_y S^t)^2 + (\Delta_z S^t)^2}}{g})S_{x,y,z} \\ &-S_{x+1,y,z} - S_{x,y+1,z} - S_{x,y,z+1} \\ &= \frac{1}{g}(\mu S^0 \sqrt{(\Delta_x S^t)^2 + (\Delta_y S^t)^2 + (\Delta_z S^t)^2} \\ &+\Delta_x g \Delta_x S^t + \Delta_y g \Delta_y S^t + \Delta_z g \Delta_z S^t) \end{aligned} \tag{27}$$

This results in a sparse linear system. The matrix of coefficients of S^{t+1} has 7 nonzero bands, and is given by

$$\begin{bmatrix} 6 + \frac{\mu\|\nabla S^t\|_0}{g_0} & -1 & \cdots & -1 & \cdots & -1 & \cdots \\ -1 & 6 + \frac{\mu\|\nabla S^t\|_1}{g_1} & -1 & \cdots & -1 & \cdots & -1 \\ 0 & -1 & 6 + \frac{\mu\|\nabla S^t\|_2}{g_2} & -1 & \cdots & -1 & \cdots \\ \cdots & \ddots & \ddots & \ddots & \ddots & \cdots \ddots & \end{bmatrix}. \tag{28}$$

The matrix in Equation (28) is symmetric and diagonally dominant. We employ the conjugate gradient descent to solve this system. The solution of Equation (28) represents one fixed-point iteration. This iteration is continued until $|S^t - S^{t+1}| < c$, where c is a small prespecified tolerance.

2.4. **Computing Probabilities.** The probability in Equation (1) can now be evaluated by computing the quantity $S(q)/S_0$ and performing the FFT. If the signal, S, is assumed to decay mono-exponentially from the origin of q-space (where $S(0) = S_0$), one can interpolate the signal values for arbitrary q. It is then possible to extrapolate (using the monoexponential decay model) from the spherical coordinate locations to grid points in cartesian space and then perform the FFT on this extrapolated data. The result is a probability of water molecule displacement over a small time constant. Since the quantity of interest is primarily the direction of water displacement, one can integrate out the radial component of $p_t(\mathbf{r})$ to

get $p_t(\theta, \phi)$. This is commonly referred to as the diffusion orientation distribution function or simply ODF. Computing the ODF with this method is computationally expensive since it requires a 3D FFT at each voxel, and then a numerical integration for each direction. For the sake of efficiency, we will compute a probability profile (not the ODF), which will make processing large datasets feasible. This probability profile, written as $p_t(r, \theta, \phi)$ quantifies the probability that a water molecule diffuses through a sphere of fixed radius, r. A more detailed treatment by Özarslan et al. can be found in [44]. This scheme provides a fast way to calculate the orientation profiles. In our implementation we have evaluated the series given in [44] up to $l = 6$ terms since the reconstructed surfaces have very simple shapes which can be accurately represented using a truncated spherical harmonics series, and r_0 was set to $17.5\mu m$. An alternative approach is to use the Funk-Radon Transform proposed by Tuch [51], however this introduces smoothing due to a spherical convolution step which would make evaluation of our denoising algorithm more difficult.

To enhance the visual impact of the probability profiles we apply a sharpening transform to the distribution by subtracting a uniform distribution (sphere) from each profile, as shown in Figure (4). The radius of the sphere is the minimum of the probability over all directions. By performing this operation the direction of maximum probability becomes more apparent.

FIGURE 4. Original probability profile (left), Minimum probability sphere (center), and sharpened probability profile (right).

3. **Experimental Results.** The denoising and rendering techniques described in the previous section were first applied to a synthetic HARDI dataset. This dataset was generated using the technique described by Özarslan et al. in [45]. The dataset was designed to depict a region of curving fibers, a region of straight fibers, and a crossing between the two. A total of 81 acquisition directions are simulated with b $= 1500 \ s/mm^2$.

A small sample of the probability surfaces $p(\theta, \phi)$ computed from the synthetic data, taken from near the crossing region, is shown in Figure (5a). The real-valued synthetic data was corrupted with Gaussian noise of zero mean, and variance $\sigma^2 = 0.005$. $p(\theta, \phi)$ surfaces computed from the noisy data (without any denoising) are shown in Figure (5b). The same voxels are shown – after smoothing over the spherical manifold at each voxel independently – in Figure (5c), after smoothing over the image lattice, in Figure (5d) and after both techniques have been used, in Figure (5e). The parameter values used for the restoration were $\mu = 0.97$, $\alpha = 0.40$, $\beta = 0.22$, $k = 100$. The right-hand side plate in each figure shows the sharpened profiles computed from the S values depicted on the left-hand side. Note that the probability surfaces in figure 5e) depict better smoothing than those in either of 5c) or 5d), visually indicating that one needs to perform smoothing on the sphere and across the lattice and not just one or the other.

TABLE 1. Error between ground-truth probabilities $\hat{p}$ and probabilities computed from restored synthetic data when SNR = 14.

Method	$\mu(d(\hat{p}, p))$	$\sigma(d(\hat{p}, p))$
p = No Restoration	0.9409	0.2516
p = FEM Restoration	0.5540	0.1997
p = TV Restoration	0.2840	0.2129
p = FEM + TV Restoration	0.1889	0.1748
p = TV + FEM Restoration	0.2128	0.1631

TABLE 2. Error between ground-truth probabilities $\hat{p}$ and probabilities computed from restored synthetic data when SNR = 5.

Method	$\mu(d(\hat{p}, p))$	$\sigma(d(\hat{p}, p))$
p = No Restoration	2.7461	0.3432
p = FEM Restoration	1.1848	0.2424
p = TV Restoration	0.9175	0.1903
p = TV + FEM Restoration	0.4970	0.2046
p = FEM + TV Restoration	0.6552	0.1984

From Figure (5b), it can be seen that the noise has a large influence on the smoothness of the distribution. As expected from the variational formulation, the spikes of noise present in the raw data have been smoothed while preserving the overall shape of the S profile. This smoothness is evident in the computed probability profiles as well.

A quantitative evaluation can be obtained by comparing the distributions computed from the smoothed data with the ground-truth by using the square root of J-divergence (symmetrized KL-divergence) as a measure. This divergence is defined as

$$d(p, q) = \sqrt{J(p, q)} \tag{29}$$

where

$$J(p, q) = \frac{1}{2} \sum_{i=1}^{n} p(\theta_i, \phi_i) \log \frac{p(\theta_i, \phi_i)}{q(\theta_i, \phi_i)} + q(\theta_i, \phi_i) \log \frac{q(\theta_i, \phi_i)}{p(\theta_i, \phi_i)} \tag{30}$$

In Table (1) we compare the distances, $d(\hat{p}, p)$, between the densities computed from the original synthetic data, $(\hat{p})$, and the unrestored data, the data restored only using the FEM method, the data restored using only the TV-norm minimization, and the data restored using both techniques. For each technique, the mean distance, $\mu(d(\hat{p}, p))$, between the densities in corresponding voxels and the standard deviation, $\sigma(d(\hat{p}, p))$, is presented. As evident from Table (1), the TV restoration has superior performance over the FEM technique in terms of the mean error. The combination of techniques has a lower mean error and standard deviation of the error than either the L_2-norm based or the TV-norm based restoration. Note also that the error achieved by applying smoothing over the sphere prior to smoothing over the voxel lattice is lower than when the order is reversed. Since TV-norm minimization can be seen as a nonlinear diffusion process, performing the denoising in this order propagates smoothed intensities within homogeneous regions. In subsequent experiments we perform the denoising in this order.

The denoising algorithm was applied to a dataset consisting of one non-diffusion weighted image and 46 diffusion weighted images of a rat spinal cord. Our data were acquired using a 14.1 Tesla (600 MHz) Bruker Avance Imaging spectrometer system with a diffusion weighted spin echo pulse sequence. Imaging parameters were : TR = 1400 ms, TE = 25 ms, Delta = 17.5 ms, delta = 1.5 ms, $b_{high} = 1500s/mm^2$, $b_{low} = 0s/mm^2$, diffusion gradient strengths = 0 mT/m with 28 averages were measured for $b = 0s/mm^2$ and diffusion gradient strengths = $733s/mm^2$ with 7 averages were measured for each of the 46 diffusion weighting-gradient directions. The 46 directions were derived from the tessellation of a hemisphere. The image field of view was $4.3 \times 4.3 \times 12mm^3$, acquisition matrix was $72 \times 72 \times 40$. The approximate SNR for the S_0 and diffusion weighted images were 58 and 50 respectively. The parameter values used for the restoration were $\mu = 0.97$, $\alpha = 0.02$, $\beta = 0.0$, $k = 1.0$.

Axial slices before and after denoising are shown for the non-diffusion weighted image in Figure (6) and one diffusion weighted image in Figure (7). The ringing artifacts visible near the sample boundary in Figure (6) have been noticeably decreased. Note that the edges in the image have been well preserved.

Figures (8) and (9) show restored probability profiles from rat brain and spinal cord datasets. The brain data were acquired using a 17.6 Tesla (750 MHz) Bruker Avance Imaging spectrometer system with a diffusion weighted spin echo pulse sequence. Imaging parameters were : TR = 2000 ms, TE = 28 ms, Delta = 17.8 ms, delta = 2.2 ms, $b_{high} = 1500s/mm^2$, $b_{low} = 0s/mm^2$, 6 averages for each of the 81 diffusion weighting-gradient directions. The 81 directions were derived from the tessellation of a hemisphere. The image field of view was $150 \times 150 \times 300\mu m^3$, acquisition matrix was $100 \times 100 \times 60$. The approximate SNR for the S_0 and diffusion weighted images were 206 and 177 respectively.

Figure (8b) shows a detail from the rat hippocampus. The piecewise smoothing behavior of the algorithm is evident within the anisotropic hippocampus region. This region has been smoothed independently of the more isotropic surrounding regions. The spherical smoothing term has also suppressed some peaks of the distribution which were probably due to noise in the acquired data. Figure (8c) shows a detail from the rat corpus callosum. The data dependent coupling term in the restoration algorithm has permitted intraregion smoothing within the corpus callosum while preventing interregion smoothing. Note that the fiber directions within the corpus callosum have been well preserved.

Figure (9) shows details from the rat spinal cord dataset. Here the noise reduction can be seen to enhance the coherence of structures in the inner core of grey matter.

The data were processed by a MATLAB implementation of the algorithm running on a system with Intel Quad Core QX6700 2.66 GHz CPU and 4 GB RAM. The computation times for the finite element smoothing over the sphere depends on the number of diffusion-encoding gradient directions in the image acquisition. For the spinal cord data with 46 directions the time was 0.018 seconds per voxel, and for the brain dataset with 81 directions the time was 0.038 seconds per voxel. The computation time for the TV-norm minimization problem for each diffusion weighted image depends on the size of the acquisition matrix. For the spinal cord the resolution was $72 \times 72 \times 40$ and the computation time per image was 28.3 seconds. For the brain dataset the resolution was $100 \times 100 \times 60$ and computation required 82.4 seconds per image.

TABLE 3. Gauss-Radau weights

i	r_i	wi_i	ξ_i	a_i
1	0.0469100770	0.1184634425	0.0398098571	0.1007941926
2	0.2307653449	0.2393143353	0.1980134179	0.2084506672
3	0.5	0.2844444444	0.4379748102	0.2604633916
4	0.7692346551	0.2393143353	0.6954642734	0.2426935942
5	0.9530899230	0.1184634425	0.9014649142	0.1598203766

4. **Conclusion.** In this paper, we presented a new variational formulation for restoring HARDI data and an FEM technique for implementing the restoration. Our formulation of the HARDI restoration involves two types of smoothness constraints. The first is smoothness over the spherical domain of acquisition directions, and the second is smoothness between neighboring voxels in the Cartesian domain. The smoothing technique is capable of preserving discontinuities in the data. This was demonstrated on synthetic and real anatomical data. By using J-divergence as a measure of distance between distribution, we were able to show quantitatively that the combination of restoration techniques performs better than either technique alone.

Appendix A.

Local Element Coordinates. We now present the coordinate system for the local elements. For local elements, triangular elements are used with a barycentric coordinate system (γ, ξ, η). Each coordinate is in the range $[0, 1]$ and $\gamma + \xi + \eta = 1$ for points on the triangle.

The global coordinates, (u, v), can be computed from the local coordinates by

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_1 - u_0 & u_2 - u_0 \\ v_1 - v_0 & v_2 - v_0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}. \tag{31}$$

The Jacobian, $\mathbf{J}$, of the transformation between coordinate systems is defined by

$$\begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial \xi} & \frac{\partial u}{\partial \eta} \\ \frac{\partial v}{\partial \xi} & \frac{\partial v}{\partial \eta} \end{bmatrix} \begin{bmatrix} d\xi \\ d\eta \end{bmatrix} = \mathbf{J} \begin{bmatrix} d\xi \\ d\eta \end{bmatrix}. \tag{32}$$

Integrals over the (u, v) domain to be converted to integrals over the local (ξ, η) domain by

$$\iint_{\Omega_j} f(u, v) du\, dv = \iint_{\Omega_j} f(u(\xi, \eta), v(\xi, \eta)) \det(\mathbf{J}) d\xi\, d\eta. \tag{33}$$

Using the Gauss-Radau quadrature rules given in [26], we can approximate the integral in Equation (33) by the summation

$$\sum_{i=1}^{5} \sum_{j=1}^{5} wi_i wj_j f(u(\xi_j, \eta_{i,j}), v(\xi_j, \eta_{i,j})) \det(\mathbf{J}) \tag{34}$$

where $\eta_{i,j} = r_i(1 - \xi_j)$, $wj_j = a_j(1 - \xi_j)$, ξ_j, and wi_i are given in Table 3.

Derivatives over (u, v) can be written in terms of local coordinates by applying the chain rule:

$$\frac{\partial N}{\partial u} = \frac{\partial N}{\partial \xi}\frac{\partial \xi}{\partial u} + \frac{\partial N}{\partial \eta}\frac{\partial \eta}{\partial u}$$
$$\frac{\partial N}{\partial v} = \frac{\partial N}{\partial \xi}\frac{\partial \xi}{\partial v} + \frac{\partial N}{\partial \eta}\frac{\partial \eta}{\partial v}. \tag{35}$$

The partial derivatives of ξ and η with respect to u and v can be computed by inverting the Jacobian

$$\begin{bmatrix} d\xi \\ d\eta \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi}{\partial u} & \frac{\partial \xi}{\partial v} \\ \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v} \end{bmatrix}\begin{bmatrix} du \\ dv \end{bmatrix} = \mathbf{J}^{-1}\begin{bmatrix} du \\ dv \end{bmatrix}. \tag{36}$$

The inverse of $\mathbf{J}$ is given by

$$\mathbf{J}^{-1} = \frac{1}{\det(\mathbf{J})}\begin{bmatrix} v_2 - v_0 & -(u_2 - u_0) \\ -(v_1 - v_0) & u_1 - u_0 \end{bmatrix}. \tag{37}$$

We use the fifth order element shape functions given by Dhatt and Touzot [26]. *This element guarantees C^1 (surface normal) continuity across triangles.* The quintic basis functions are given by

$$
\begin{aligned}
N_1 &= \lambda^2(10\lambda - 15\lambda^2 + 6\lambda^3 + 30\xi\eta(\xi + \eta)) \\
N_2 &= \xi\lambda^2(3 - 2\lambda - 3\xi^2 + 6\xi\eta) \\
N_3 &= \eta\lambda^2(3 - 2\lambda - 3\eta^2 + 6\xi\eta) \\
N_4 &= \frac{1}{2}\xi^2\lambda^2(1 - \xi + 2\eta) \\
N_5 &= \xi\eta\lambda^2 \\
N_6 &= \frac{1}{2}\eta^2\lambda^2(1 + 2\xi - \eta) \\
N_7 &= \xi^2(10\xi - 15\xi^2 + 6\xi^3 + 15\eta^2\lambda) \\
N_8 &= \frac{1}{2}\xi^2(-8\xi + 14\xi^2 - 6\xi^3 - 15\eta^2\lambda) \\
N_9 &= \frac{1}{2}\xi^2\eta(6 - 4\xi - 3\eta - 3\eta^2 + 3\xi\eta) \\
N_{10} &= \frac{1}{4}\eta^2(2\xi(1 - \xi)^2 + 5\eta^2\lambda) \\
N_{11} &= \frac{1}{2}\xi^2\eta(-2 + 2\xi + \eta + \eta^2 - \xi\eta) \\
N_{12} &= \frac{1}{4}\xi^2\eta^2\lambda + \frac{1}{2}\xi^3\eta^2 \\
N_{13} &= \eta^2(10\eta - 15\eta^2 + 6\eta^3 + 15\xi^2\lambda) \\
N_{14} &= \frac{1}{2}\xi\eta^2(6 - 3\xi - 4\eta - 3\xi^2 + 3\xi\eta) \\
N_{15} &= \frac{1}{2}\eta^2(-8\eta + 14\eta^2 - 6\eta^3 - 15\xi^2\lambda) \\
N_{16} &= \frac{1}{4}\xi^2\eta^2\lambda + \frac{1}{2}\xi^2\eta^3 \\
N_{17} &= \frac{1}{2}\xi\eta^2(-2 + \xi + 2\eta + \xi^2 - \xi\eta) \\
N_{18} &= \frac{1}{4}\eta^2(2\eta(1 - \eta)^2 + 5\xi^2\lambda).
\end{aligned}
$$
$$\tag{38}$$

The quintic shape functions have nodal variables which can be written in terms of local or global coordinates as,

$$
\mathbf{q}_{\xi,\eta} = \begin{bmatrix} z \\ z_\xi \\ z_\eta \\ z_{\xi\xi} \\ z_{\xi\eta} \\ z_{\eta\eta} \end{bmatrix}, \mathbf{q}_{u,v} = \begin{bmatrix} z \\ z_u \\ z_v \\ z_{uu} \\ z_{uv} \\ z_{vv} \end{bmatrix}.
\tag{39}
$$

The local and global nodal variables are related to each other by

$$
\mathbf{q}_{u,v} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & \xi_u & \eta_u & 0 & 0 & 0 \\
0 & \xi_v & \eta_v & 0 & 0 & 0 \\
0 & 0 & 0 & \xi_u^2 & 2\xi_u\eta_u & \eta_u^2 \\
0 & 0 & 0 & \xi_u\xi_v & (\xi_u\eta_v + \eta_u\xi_v) & \eta_u\eta_v \\
0 & 0 & 0 & \xi_v^2 & 2\xi_v\eta_v & \eta_v^2
\end{bmatrix} \mathbf{q}_{\xi,\eta}.
\tag{40}
$$

Appendix B.

Global Matrices. We wish to construct global matrices so that the energy balance over the entire FEM mesh is given by the linear system

$$
\mathbf{Kq} = \mathbf{f}
\tag{41}
$$

where $\mathbf{K}$ is a $(6n \times 6n)$ matrix since we have 6 variables per node.

We will consider the simple case of 2 elements. Expanding the element Equation(19) in terms of nodal variables for element 0 we get

$$
\begin{bmatrix}
\mathbf{K}_{0,0}^0 & \mathbf{K}_{0,1}^0 & \mathbf{K}_{0,2}^0 \\
\mathbf{K}_{1,0}^0 & \mathbf{K}_{1,1}^0 & \mathbf{K}_{1,2}^0 \\
\mathbf{K}_{2,0}^0 & \mathbf{K}_{2,1}^0 & \mathbf{K}_{2,2}^0
\end{bmatrix}
\begin{bmatrix} \mathbf{q}_0^0 \\ \mathbf{q}_1^0 \\ \mathbf{q}_2^0 \end{bmatrix} =
\begin{bmatrix} \mathbf{f}_0^0 \\ \mathbf{f}_1^0 \\ \mathbf{f}_2^0 \end{bmatrix},
\tag{42}
$$

and for element 1 we have

$$
\begin{bmatrix}
\mathbf{K}_{3,3}^1 & \mathbf{K}_{3,2}^1 & \mathbf{K}_{3,1}^1 \\
\mathbf{K}_{2,3}^1 & \mathbf{K}_{2,2}^1 & \mathbf{K}_{2,1}^1 \\
\mathbf{K}_{1,3}^1 & \mathbf{K}_{1,2}^1 & \mathbf{K}_{1,1}^1
\end{bmatrix}
\begin{bmatrix} \mathbf{q}_3^1 \\ \mathbf{q}_2^1 \\ \mathbf{q}_1^1 \end{bmatrix} =
\begin{bmatrix} \mathbf{f}_3^1 \\ \mathbf{f}_2^1 \\ \mathbf{f}_1^1 \end{bmatrix}.
\tag{43}
$$

where each $\mathbf{q}_l^j$ is a (6×1) column vector of nodal variables. We expand each $\mathbf{K}^j$ to be $(6n \times 6n)$ by inserting rows and columns of zeros corresponding to each node of the mesh. Also expand $\mathbf{f}^j$ to $(6n \times 1)$. The global $\mathbf{K}$ and $\mathbf{q}$ are obtained by summing the expanded matrices from each element in the mesh. For our 2 element example we have

$$
\begin{bmatrix}
\mathbf{K}_{0,0}^0 & \mathbf{K}_{0,1}^0 & \mathbf{K}_{0,2}^0 & 0 \\
\mathbf{K}_{1,0}^0 & \mathbf{K}_{1,1}^0 + \mathbf{K}_{1,1}^1 & \mathbf{K}_{1,2}^0 + \mathbf{K}_{1,2}^1 & \mathbf{K}_{1,3}^1 \\
\mathbf{K}_{2,0}^0 & \mathbf{K}_{2,1}^0 + \mathbf{K}_{2,1}^1 & \mathbf{K}_{2,2}^0 + \mathbf{K}_{2,2}^1 & \mathbf{K}_{2,3}^1 \\
0 & \mathbf{K}_{3,1}^1 & \mathbf{K}_{3,2}^1 & \mathbf{K}_{3,3}^1
\end{bmatrix}
\begin{bmatrix} \mathbf{q}_0^0 \\ \mathbf{q}_1^0 \\ \mathbf{q}_2^0 \\ \mathbf{q}_3^0 \end{bmatrix} =
\begin{bmatrix} \mathbf{f}_0^0 \\ \mathbf{f}_1^0 + \mathbf{f}_1^1 \\ \mathbf{f}_2^0 + \mathbf{f}_2^1 \\ \mathbf{f}_3^0 \end{bmatrix}.
\tag{44}
$$

REFERENCES

[1] H-E. Assemlal, D. Tschumperlé, and L. Brun, *Fiber tracking on HARDI data using robust ODF fields*, in "IEEE International Conference on Image Processing" (2007), 344–351.

[2] H.-E. Assemlal, D. Tschumperlé, and L. Brun, *Robust variational estimation of pdf functions from diffusion MR signal*, in "Computational Diffusion MRI Workshop held at the Medical Image Computing and Computer-Assisted Intervention" (2008), 1–8.

[3] S.P. Awate and R.T. Whitaker, *Feature-preserving MRI denoising: A nonparametric empirical bayes approach*, IEEE Transactions on Medical Imaging **26** (2007), 1242–1255.

[4] P. J. Basser, *New histological and physiological stains derived from diffusion-tensor MR images*, Ann. N Y Acad. Sci. **820** (1997), 123–138.

[5] P. J. Basser, *Relationships between diffusion tensor and q-space MRI*, Magn. Reson. Med. **47** (2002), 392–397.

[6] P. J. Basser, J. Mattiello, and D. Le Bihan, *MR diffusion tensor spectroscopy and imaging*, Biophys. J. **66** (1994), 259–267.

[7] P. J. Basser, J. Mattiello, and D.Le Bihan, *Estimate of the effective self-diffusion tensor from the NMR spin echo*, J. Magn. Reson. **B** (1994), 247–254.

[8] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, *In vivo fiber tractography using DT-MRI data*, Magn. Reson. Med. **44** (2000), 625–632.

[9] S. Basu, T. Fletcher, and R. Whitaker, *Rician noise removal in diffusion tensor MRI.*, in "Medical Image Computing and Computer-Assisted Intervention", (2006), 117–125.

[10] C. Beaulieu and P. S. Allen, *Determinants of anisotropic water diffusion in nerves*, Magn Reson Med **31** (1994), 394–400.

[11] P. Blomgren and T. F. Chan, *Color TV: total variation methods for restoration of vector-valued images*, IEEE Transactions on Image Processing **7** (1998), 304–309.

[12] I. Bogdanova, X. Bresson, J.P. Thiran, and P. Vandergheynst, *Scale space analysis and active contours for omnidirectional images*, IEEE Transactions On Image Processing **16** (2007), 1888–1901.

[13] P.T. Callaghan, "Principles of Nuclear Magnetic Resonance Microscopy", Clarendon Press, Oxford, 1991.

[14] C.A. Castańo-Moraga, C. Lenglet, R. Deriche, and J. Ruiz-Alzola, *A Riemannian approach to anisotropic filtering of tensor fields*, Signal Processing **87** (2007), 263–276.

[15] T. Cecil, S. Osher, and L. Vese, *Numerical methods for minimization problems constrained to S1 and S2*, Journal of Computational Physics **198** (2004), 567–579.

[16] T. Chan and P. Mulet, *On the convergence of the lagged diffusivity fixed point method in total variation image restoration*, SIAM Journal on Numerical Analysis **36** (1999), 354–367.

[17] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, *Regularizing flows for constrained matrix-valued images*, Journal of Mathematical Imaging and Vision **20** (2004), 147–162.

[18] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. D. Faugeras, *Constrained flows of matrix-valued functions: Application to diffusion tensor regularization*, European Conference on Computer Vision **1** (2002), 251–265.

[19] B. Chen and E.W. Hsu, *Noise Removal in Magnetic Resonance Diffusion Tensor Imaging*, Magnetic Resonance In Medicine **54** (2005), 393–401.

[20] Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B. C. Vemuri, and Y. Liu, *Estimation, smoothing, and characterization of apparent diffusion coefficient profiles from high angular resolution DWI.*, in "Computer Vision and Pattern Recognition", (2004), 588–593.

[21] R. A. Clarke, P. Scifo, G. Rizzo, F. DellAcqua, G. Scotti, and F. Fazio, *Noise correction on Rician distributed data fibre orientation estimators*, IEEE Transactions Medical Imaging **27** (2008), 1242–1251.

[22] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle, *Tracking neuronal fiber pathways in the living human brain*, Proc. Natl. Acad. Sci. **96** (1999), 10422–10427.

[23] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, *A fast and robust odf estimation algorithm in q-ball imaging*, in "International Symposium on Biomedical Imaging: From Nano to Macro", (2006), 81–84.

[24] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, *Regularized, fast, and robust analytical Q-ball imaging*, Magnetic Resonance In Medicine **58** (2007), 497–510.

[25] M. Descoteaux, N. Wiest-Daesslé, S. Prima, C. Barillot, and R. Deriche, *Impact of rician adapted non-local means filtering on HARDI*, in "Medical Image Computing and Computer-Assisted Intervention", (2008), 122–130.

[26] G. Dhatt and G. Touzot, "The Finite Element Method Displayed", Wiley, New York, 1984.

[27] C. Feddern, J. Weickert, B. Burgeth, and M. Welk, *Curvature-Driven PDE Methods for Matrix-Valued Images*, International Journal of Computer Vision **69** (2006), 93–107.

[28] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, *Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics*, IEEE Transactions on Medical Imaging **26** (2007), 1472–1482.

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

[30] L. R. Frank, *Characterization of anisotropy in high angular resolution diffusion-weighted MRI*, Magn Reson Med **47** (2002), 1083–1099.

[31] Y. Gur and N. Sochen, *Denoising tensors via lie group flows*, Variational, Geometric, and Level Set Methods in Computer Vision **3752** (2005), 13–24.

[32] G. Hamarneh and J. Hradsky, *Bilateral Filtering of Diffusion Tensor Magnetic Resonance Images*, IEEE Transactions On Image Processing **16** (2007), 2463–2475.

[33] K. M. Jansons and D. C. Alexander, *Persistent angular structure: new insights from diffusion magnetic resonance imaging data*, Inverse problems **19** (2003), 1031–1046.

[34] D. K. Jones, A. Simmons, S. C. R. Williams, and M. A. Horsfield, *Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI*, Magn. Reson. Med. **42** (1999), 37–41.

[35] E. Fatemi L. Rudin, S. Osher, *Nonlinear total variation based noise reduction algorithms*, Physica D **60** (1992), 259–268.

[36] S.-H. Lai and B. C. Vemuri, *Physically based adaptive preconditioning for early vision.*, IEEE Trans. Pattern Anal. Mach. Intell. **19** (1997), 594–607.

[37] L. Lui, Y. Wang, and TF Chan, *Solving PDEs on manifold using global conformal parameterization*, in "Variational, Geometric, and Level Set Methods in Computer Vision: Third International Workshop", (2005), 307–319.

[38] N. Makris, A. J. Worth, A. G. Sorensen, G. M. Papadimitriou, O. Wu, T. G. Reese, V. J. Wedeen, and *et al*, *Morphometry of in vivo human white matter association pathways with diffusion-weighted magnetic resonance imaging*, Ann. Neurol. **42** (1999), 951–962.

[39] T. McGraw, B. C. Vemuri, Y. Chen, M. Rao, and T. Mareci, *DT-MRI denoising and neuronal fiber tracking*, Medical Image Analysis **8** (2004), 95–111.

[40] T. McGraw, B.C. Vemuri, Z. Wang, Y. Chen, M. Rao, and T. Mareci, *Line integral convolution for visualization of fiber tract maps from DTI*, in "Medical Image Computing and Computer-Assisted Intervention", (2002), 615–622.

[41] T. McInerney and D. Terzopoulos, *A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis*, Comp. Med. Imag. Graph. **19** (1995), 69–83.

[42] S. Mori, B. J. Crain, V. P. Chacko, and P. C. M. van Zijl, *Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging*, Ann. Neurol. **45** (1999), 265–269.

[43] E. Özarslan and T. H. Mareci, *Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging*, Magn Reson Med **50** (2003), 955–965.

[44] E. Özarslan, T. M. Shepherd, B. C. Vemuri, S. J. Blackband, and T. H. Mareci, *Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT)*, NeuroImage **31** (2006), 1086–1103.

[45] E. Özarslan, B. C. Vemuri, and T. H. Mareci, *Generalized scalar measures for diffusion MRI using trace, variance and entropy*, Magn Reson Med **53** (2005), 866–876.

[46] P. Savadjiev, J. S. W. Campbell, G. B. Pike, and K. Siddiqi, *3d curve inference for diffusion MRI regularization and fibre tractography*, Medical Image Analysis **10** (2006), 799–813.

[47] E. O. Stejskal and J. E. Tanner, *Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient*, J. Chem. Phys. **42** (1965), 288–292.

[48] B. Tang, G. Sapiro, and V. Caselles, *Diffusion of general data on non-flat manifolds via harmonic maps theory: The direction diffusion case*, International Journal of Computer Vision **36** (2000), 149–161.

[49] D. Terzopoulos, *Regularization of inverse visual problems involving discontinuities*, IEEE Transactions on Pattern Analysis and Machine Intelligence **8** (1986), 413–242.

[50] D. Tschumperlé and R. Deriche, *Variational frameworks for DT-MRI estimation, regularization and visualization*, International Conference on Computer Vision (2003), 116–121.

[51] D. S. Tuch, *Q-ball imaging*, Magn Reson Med. **52** (2004), 1358–1372.

[52] D. S. Tuch, T. G. Reese, M. R. Wiegell, and V. J. Wedeen, *Diffusion MRI of complex neural architecture*, Neuron **40** (2003), 885–895.

[53] D. S. Tuch, R. M. Weisskoff, J. W. Belliveau, and V. J. Wedeen, *High angular resolution diffusion imaging of the human brain*, Proc. of the 7th Annual Meeting of ISMRM (Philadelphia, PA), 1999, p. 321.

[54] B. C. Vemuri, Y. Chen, M. Rao, T. McGraw, Z. Wang, and T. Mareci, *Fiber tract mapping from diffusion tensor MRI*, in "Proceedings of the IEEE Workshop on Variational and Level Set Methods", (2001), 81–88.

[55] B. C. Vemuri, Y. Chen, M. Rao, Z. Wang, T. McGraw, T. Mareci, S. J. Blackband, and P. Reier, *Automatic fiber tractography from DTI and its validation*, in "IEEE Intl. Symp. on Biomedical Imaging", (2002), 501–504.

[56] B. C. Vemuri and Y. Guo, *Snake pedals: Compact and versatile geometric models with physics-based control*, IEEE Trans. Pattern Anal. Mach. Intell. **22** (2000), 445–459.

[57] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci, *A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI.*, in "Information Processing in Medical Imaging", (2003), 660–671.

[58] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci, *A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI*, IEEE Transactions on Medical Imaging **23** (2004), 930–939.

[59] V. J. Wedeen, T. G. Reese, D. S. Tuch, M. R. Weigel, J. G. Dou, R.M. Weiskoff, and D. Chessler, *Mapping fiber orientation spectra in cerebral white matter with Fourier transform diffusion MRI*, in "Proc. of the 8th Annual Meeting of ISMRM", 2000, 82.

[60] J. Weickert and T. Brox, *Diffusion and regularization of vector- and matrix-valued images*, Contemporary Mathematics **313** (2002), 251–268.

[61] N. Wiest-Daesslé, S. Prima, P. Coupé, S. P. Morrissey, and C. Barillot, *Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI*, in "Medical Image Computing and Computer-Assisted Intervention", (2008), 171–179.

[62] N. Wiest-Daessle, S. Prima, P. Coupe, S.P. Morrissey, and C. Barillot, *Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI*, Medical Image Computing and Computer-Assisted Intervention **4792** (2007), 344.

[63] R. Xue, P. C. M. van Zijl, B. J. Crain, H. Solaiyappan, and S. Mori, *In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging*, Magn. Reson. Med. **42** (1999), 1123–1127.
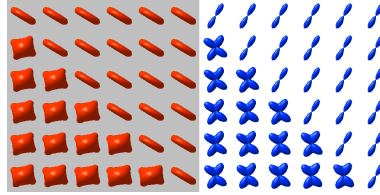
E-mail address: `tim.mcgraw@mail.wvu.edu`
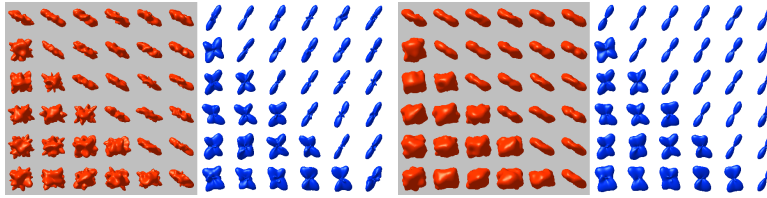*E-mail address*: `vemuri@cise.ufl.edu`
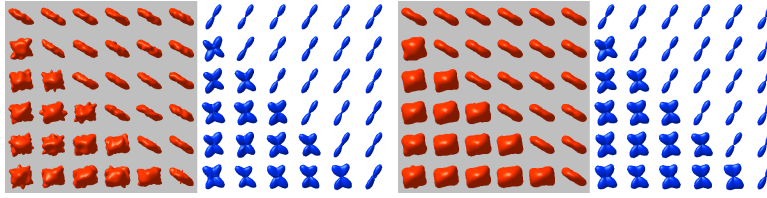*E-mail address*: `evren@helix.nih.gov`
E-mail address: `yun@math.ufl.edu`
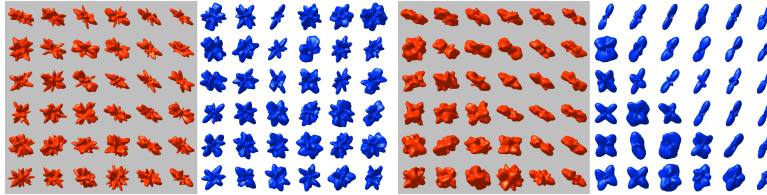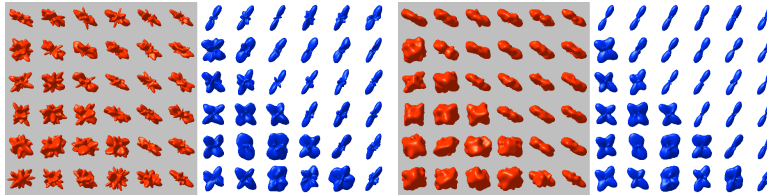*E-mail address*: `thmareci@ufl.edu`

(a) Original synthetic data

(b) Added Gaussian noise, $\sigma^2 = 0.005$, (c) Manifold smoothing with $\alpha = 0.01$
SNR $= 14$

(d) Lattice smoothing with $\mu = 0.95$ (e) Manifold smoothing followed by Lattice smoothing

(f) Added Gaussian noise, $\sigma^2 = 0.04$, (g) Manifold smoothing with $\alpha = 0.01$
SNR $= 5$

(h) Lattice smoothing with $\mu = 0.95$ (i) Manifold smoothing followed by Lattice smoothing

FIGURE 5. Simulations from a 64×64 synthetic high angular resolution diffusion image, subsampled for display purposes. The signal profiles, $S(\theta, \phi)$, are shown on the left side of each panel whereas the probability profiles are provided on the right. Results from manifold smoothing using FEM, lattice smoothing using TV-norm minimization, and a combination of the two techniques are shown for noisy data.
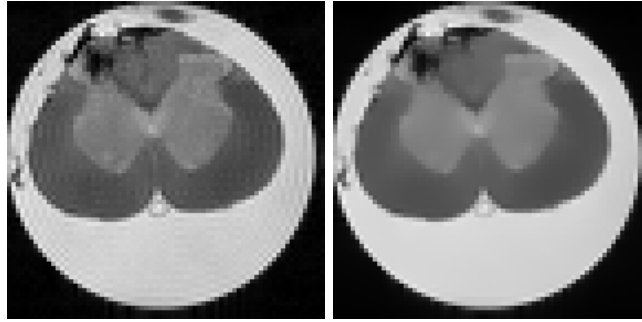
FIGURE 6. Original S_0 (non-diffusion weighted) image (left), and denoised (right) from spinal cord data.
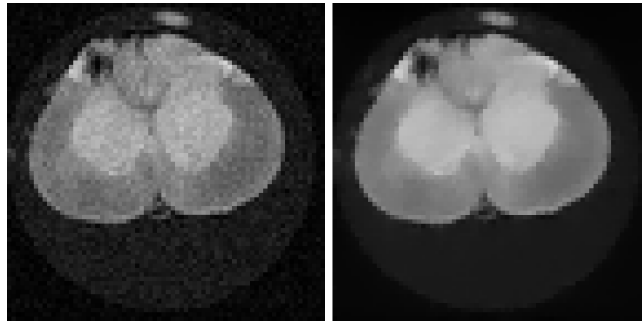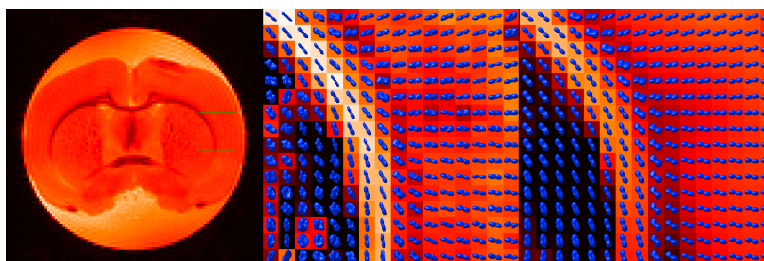

FIGURE 7. A slice from the original 3D diffusion weighted image (left), and corresponding slice from the denoised (right) spinal cord data.
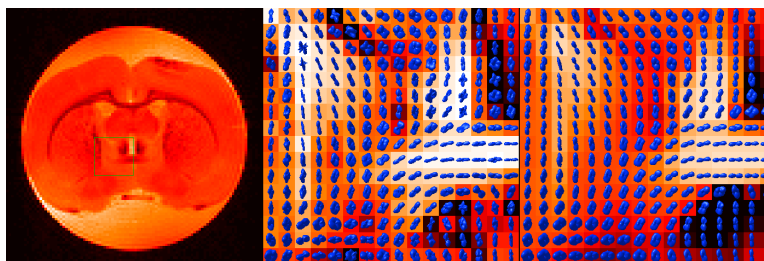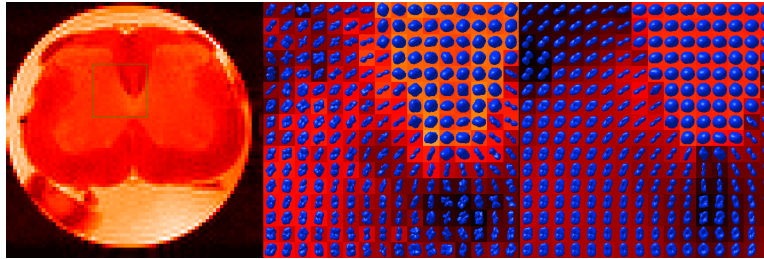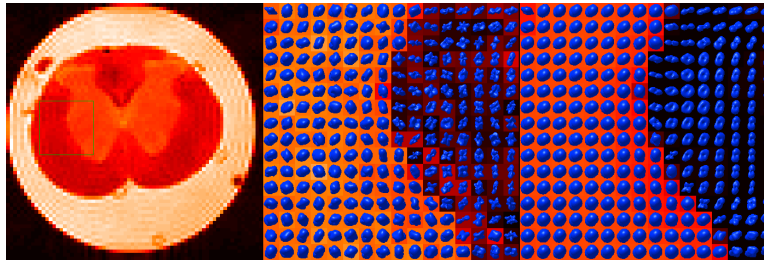
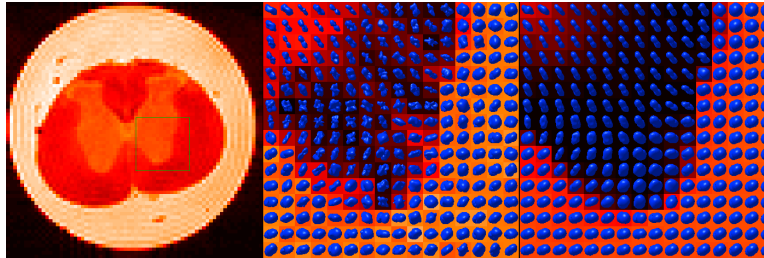(a) Anterior commissure
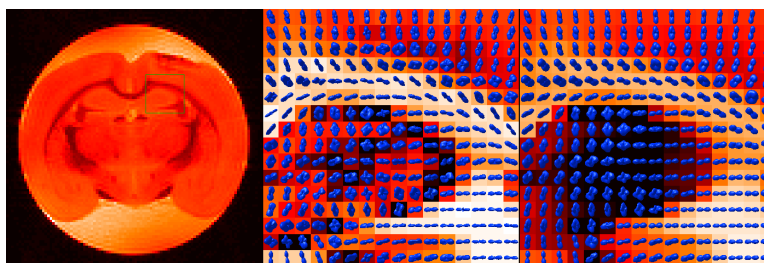
(b) Hippocampus

(c) Corpus callosum

FIGURE 8. Restored probability profiles from rat brain data overlaid on anisotropy images. Representative profiles from the region of interest (left) are shown for the original (middle) and denoised (right) data.

(a) Pyramidal tract

(b) Lumbar region

(c) Thoracic region

FIGURE 9. Restored probability profiles from rat spinal cord data overlaid on anisotropy images. Representative profiles from the region of interest (left) are shown for the original (middle) and denoised (right) data.
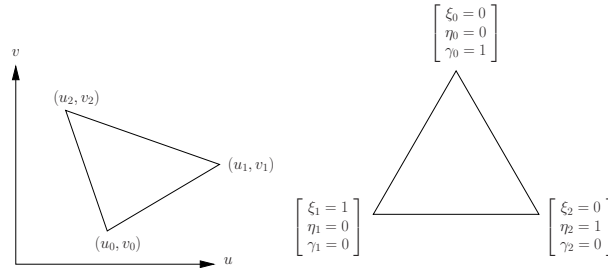

FIGURE 10. Mapping to barycentric coordinates