

# Fluid Registration of Diffusion Tensor Images Using Information Theory

Ming-Chang Chiang, Alex D. Leow, Andrea D. Klunder, Rebecca A. Dutton, Marina Barysheva, Stephen E. Rose, Katie L. McMahon, Greig I. de Zubicaray, Arthur W. Toga, and Paul M. Thompson\*

**Abstract**—We apply an information-theoretic cost metric, the symmetrized Kullback-Leibler (sKL) divergence, or  $J$ -divergence, to fluid registration of diffusion tensor images. The difference between diffusion tensors is quantified based on the sKL-divergence of their associated probability density functions (PDFs). Three-dimensional DTI data from 34 subjects were fluidly registered to an optimized target image. To allow large image deformations but preserve image topology, we regularized the flow with a large-deformation diffeomorphic mapping based on the kinematics of a Navier-Stokes fluid. A driving force was developed to minimize the  $J$ -divergence between the deforming source and target diffusion functions, while reorienting the flowing tensors to preserve fiber topography. In initial experiments, we showed that the sKL-divergence based on full diffusion PDFs is adaptable to higher-order diffusion models, such as high angular resolution diffusion imaging (HARDI). The sKL-divergence was sensitive to subtle differences between two diffusivity profiles, showing promise for nonlinear registration applications and multisubject statistical analysis of HARDI data.

**Index Terms**—Diffusion tensor imaging (DTI), fluid registration, high angular resolution diffusion imaging (HARDI), Kullback-Leibler divergence.

## I. INTRODUCTION

**D**IFFUSION tensor imaging (DTI) is becoming increasingly popular for *in vivo* studies of fiber architecture in the brain (tractography), investigations of anatomical and functional connectivity, and clinical studies of white matter (WM) integrity [1], [2]. Because DTI is sensitive to subtle differences in WM fiber orientation and diffusion anisotropy, it provides a powerful new approach to reveal 3-D deficit patterns

in population studies of neurological diseases, e.g., Alzheimer's disease, multiple sclerosis, or schizophrenia [2].

Multisubject studies and group comparisons of DTI require a scheme to correlate or compare data across subjects, and to perform statistical analysis of DTI signals in a standard coordinate space. DTI data is inherently multidimensional (with effectively six independent parameters defining a positive-definite symmetric matrix at each voxel) and non-Euclidean (in that Euclidean interpolation for diffusion tensors may lead to loss of anisotropy, or tensor swelling [3]). DTI, therefore, presents unique challenges for image analysis. Extending work on the differential geometry of manifolds [4], researchers at INRIA in Sophia-Antipolis, France [3], [5] developed an extensive calculus for performing computations on tensors. They derived a log-Euclidean metric for comparing and averaging diffusion tensors on a Riemannian manifold by transforming each diffusion tensor into its matrix logarithm counterpart. The processing of diffusion tensors then becomes Euclidean in the logarithmic domain, and standard formulae can be applied to compute means and variances. Fletcher and Joshi [6] further developed the notion of principal geodesic analysis (PGA), which extends principal component analysis to manifolds, by using an affine-invariant metric on the space of symmetric positive-definite tensors. Work treating diffusion tensors as elements of a Lie group with a non-Euclidean metric is being pursued on many fronts, for DTI smoothing and anisotropic filtering [7], interpolation, segmentation and fiber tracking, and estimation of probabilistic connectivity matrices [8]–[11]. Khurd *et al.* [12], for example, use isometric mapping and manifold learning techniques (eigendecomposition of the distance matrix) to directly fit a manifold to the tensors, compute its dimensionality, and distinguish groups using Hotelling's  $T^2$  statistics. Taking a different conceptual approach, diffusion tensors, which are positive-definite, may instead be interpreted as first-order summaries of 3-D probability density functions (representing diffusion processes) and these distributions can also be processed using information theory [13]. Tsuda *et al.* [14] introduced the Kullback-Leibler (KL) divergence as the distance between two Gaussian probability density functions (PDFs), parameterizing the PDFs by using a positive-definite matrix or tensor as their covariance matrices. Wang and Vemuri [15], as well as Lenglet *et al.* [16], pioneered the use of the symmetrized KL-divergence in the context of DTI segmentation, in which the symmetrized KL-divergence served as a natural measure of dissimilarity among symmetric positive-definite (SPD) tensors, and used it in a region-based active contour model. The KL-divergence was part of a cost function for segmentation, which was minimized in a level-set evolution

Manuscript received June 5, 2007; revised August 13, 2007. This work was supported in part by the National Institute on Aging, in part by the National Library of Medicine, in part by the National Institute for Biomedical Imaging and Bioengineering, and in part by the National Center for Research Resources (NCRR). The work of P. M. Thompson was supported by the National Institute for Child Health and Development under Grant AG016570, Grant LM05639, Grant EB01651, Grant RR019771, and Grant HD050735. The work of M.-C. Chiang was supported by a Taiwan government fellowship. The work of A. W. Toga was supported by the NCRR under Grant RR13642. Asterisk indicates corresponding author.

M.-C. Chiang, A. D. Leow, A. D. Klunder, R. A. Dutton, M. Barysheva, and A. W. Toga are with the Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, CA 90095 USA.

S. E. Rose, K. L. McMahon, and G. I. de Zubicaray are with Centre for Magnetic Resonance, University of Queensland, 4072 Brisbane, Australia.

\*P. M. Thompson is with the Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, 710 Westwood Plaza, Los Angeles, CA 90095 USA (e-mail: thompson@loni.ucla.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2007.907326

scheme. However, to our knowledge, the symmetrized KL-divergence has not to date been used for DTI registration. This is the focus of our efforts in this paper.

For nonlinear registration of DT images, Guimond *et al.* [17] proposed a multidimensional “Demons” method (building on work on demons by Thirion *et al.* [18]), which treats the DTI as multiple-channel data. Ruiz-Alzola *et al.* [19] proposed a multiresolution scheme in which areas with local detail in the tensor data were matched at successively refined resolution levels; in this approach, deformation fields were interpolated using a Kriging estimator. Alexander and Gee [20] compared the performance of different DTI matching approaches using various DT similarity measures in a multiresolution elastic registration scheme, pioneering the concept of tensor reorientation, which, given an affine transform  $\mathbf{A}$ , correctly applies only the rotational component of  $\mathbf{A}$  to the tensors during warping, preserving the underlying fiber connectivity. More recently, Zhang *et al.* [21] included a “finite strain” strategy for tensor reorientation [22] in the cost function, and approximated a nonlinear deformation using a piecewise affine algorithm. Cao *et al.* [23], [24] have worked on a problem that is perhaps most closely related to the topic of this paper. They registered two DT images by minimizing the Euclidean distance between the principal eigenvectors [23] or between diffusion tensors themselves (using the Frobenius norm) [24], with images deformed using diffeomorphic flows.

In this paper, we match diffusion tensors based on variationally minimizing the symmetrized KL-divergence between Gaussian PDFs whose covariance matrices are given by the diffusion tensors. The registration is computed in a fluid regularization scheme, in which diffeomorphic and inverse-consistent mappings are guaranteed. Moreover, we extend the symmetrized KL-divergence to model higher-order water diffusivity profiles based on the concept of spherical harmonics [25], [26]. Our work is most related to that presented by Cao *et al.* [23], [24].

The remainder of the paper is organized as follows. Section II explains the variational scheme to minimize the cost function based on the symmetrized KL-divergence, the extension of the symmetrized KL-divergence to higher-order diffusion models, the concept of reorientation of diffusivity profiles, and the theory of the inverse-consistent fluid model. Section III details the experimental procedures in which DT images are registered. Section IV provides the measures by which the registration performance is evaluated. Section V presents the results on the registration of DTI. We also compared the symmetrized KL-divergence to other cost functions in DTI, in synthetic two-fiber diffusivity models, and in real high angular resolution diffusion imaging (HARDI) data. Section VI compares our algorithm with other approaches for DTI registration.

## II. REGISTRATION SCHEME FOR DIFFUSION TENSORS

### A. DTI Registration Using the Symmetrized KL-Divergence

In a diffusion tensor-magnetic resonance (DT-MR) image, the random diffusion of water molecules with respect to a specific image point  $\mathbf{x}$  can be modeled, to the first order, using a 3-D Gaussian probability density function (PDF), whose spatial

covariance can be approximated by a positive-definite second-order symmetric tensor  $\mathbf{T}(\mathbf{x})$  [27]

$$p_{\mathbf{x}}(\mathbf{r}) = \frac{1}{\sqrt{(2\pi)^3 |\mathbf{T}(\mathbf{x})|}} \exp\left(\frac{-\mathbf{r}^T \mathbf{T}^{-1}(\mathbf{x}) \mathbf{r}}{4\tau}\right) \quad (1)$$

where  $\mathbf{r}$  is the displacement of water molecules relative to  $\mathbf{x}$ , and  $\tau \geq 0$  is the diffusion time for a molecule initially located at  $\mathbf{x}$  at time  $\tau = 0$ . The superscript  $T$  denotes transpose. Without loss of generality, we can set  $\mathbf{D}(\mathbf{x}) = 2\mathbf{T}(\mathbf{x})\tau$  to absorb the diffusion time. Equation (1) then becomes

$$p_{\mathbf{x}}(\mathbf{r}) = \frac{1}{\sqrt{(2\pi)^3 |\mathbf{D}(\mathbf{x})|}} \exp\left(\frac{-\mathbf{r}^T \mathbf{D}^{-1}(\mathbf{x}) \mathbf{r}}{2}\right). \quad (2)$$

Throughout this paper, we denote tensors and matrices by uppercase bold fonts, vectors by lowercase bold fonts, and scalars by nonbold fonts.

Given two continuous PDFs  $p(\mathbf{x})$  and  $q(\mathbf{x})$ , the KL-divergence (also known as the *information gain*, or *relative entropy*) of  $q(\mathbf{x})$  from  $p(\mathbf{x})$  is defined as [15]

$$\text{KL}(p||q) = \int p(\mathbf{x}) \log \frac{p(\mathbf{x})}{q(\mathbf{x})} d\mathbf{x}. \quad (3)$$

Because this measure is not symmetric, instead we use the symmetrized KL-divergence, (also known as the *J-divergence*), as proposed in [15], which is given by

$$\begin{aligned} \text{sKL}(p, q) &= \frac{1}{2} \{ \text{KL}(p||q) + \text{KL}(q||p) \} \\ &= \frac{1}{2} \left\{ \int p(\mathbf{x}) \log \frac{p(\mathbf{x})}{q(\mathbf{x})} d\mathbf{x} \right. \\ &\quad \left. + \int q(\mathbf{x}) \log \frac{q(\mathbf{x})}{p(\mathbf{x})} d\mathbf{x} \right\}. \end{aligned} \quad (4)$$

Let  $\mathbf{S}$  and  $\mathbf{T}$  denote tensor-valued images, with diffusion tensors  $\mathbf{S}(\mathbf{x})$  and  $\mathbf{T}(\mathbf{x})$ , respectively, at each voxel. As we model  $\mathbf{S}(\mathbf{x})$  and  $\mathbf{T}(\mathbf{x})$  as the covariance matrix of the associated Gaussian PDFs, they were replaced by positive-definite counterparts [28], if any of their eigenvalues was negative due to noise [29].

Denoting by  $\mathbf{y} = \mathbf{x} - \mathbf{u}$  the mapping of  $\mathbf{S}$  at voxel  $\mathbf{y}$  to  $\mathbf{T}$  at voxel  $\mathbf{x}$  with the deformation vector  $\mathbf{u}$ , taken together with (2) and (4), the symmetrized KL-divergence of the Gaussian PDFs  $p_{\mathbf{S}, \mathbf{x}-\mathbf{u}}(\mathbf{r})$  and  $p_{\mathbf{T}, \mathbf{x}}(\mathbf{r})$  has a closed form given by [15]

$$\text{sKL}(p_{\mathbf{S}, \mathbf{x}-\mathbf{u}}, p_{\mathbf{T}, \mathbf{x}}) = \frac{1}{4} \{ \text{tr}(\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u})) + \text{tr}(\mathbf{S}(\mathbf{x} - \mathbf{u})\mathbf{T}^{-1}(\mathbf{x})) - 6 \} \quad (5)$$

where  $\text{tr}(\cdot)$  represents the trace of a square matrix.  $\text{sKL}(p_{\mathbf{S}, \mathbf{x}-\mathbf{u}}, p_{\mathbf{T}, \mathbf{x}})$  has been proved to be affine-invariant [15]. Vemuri *et al.* [15] pioneered the use of this functional, and have used it extensively for DT segmentation, but not registration.

We adopted the variational framework proposed in [30] to derive a driving force for registering the tensor-valued image

$\mathbf{S}(\mathbf{x})$  to  $\mathbf{T}(\mathbf{x})$  by minimizing the symmetrized KL-divergence of the two images

$$E(\mathbf{u}, \mathbf{S}, \mathbf{T}) = \frac{1}{4} \int_{\Omega} \left\{ \text{tr}(\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u})) + \text{tr}(\mathbf{S}(\mathbf{x} - \mathbf{u})\mathbf{T}^{-1}(\mathbf{x})) - 6 \right\} d\mathbf{x} \quad (6)$$

where  $\Omega$  is a bounded domain in which the two images are defined. If  $\mathbf{w}$  is a small perturbation of the deformation field  $\mathbf{u}$  that registers the images, then taking the definition of the first variation of  $E(\mathbf{u}, \mathbf{S}, \mathbf{T})$ , we have

$$\begin{aligned} & \left. \frac{\partial E(\mathbf{u} + \varepsilon \mathbf{w}, \mathbf{S}, \mathbf{T})}{\partial \varepsilon} \right|_{\varepsilon=0} \\ &= \frac{1}{4} \int_{\Omega} \left\{ \left. \frac{\partial [\text{tr}(\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - (\mathbf{u} + \varepsilon \mathbf{w})))]}{\partial \varepsilon} \right|_{\varepsilon=0} + \left. \frac{\partial [\text{tr}(\mathbf{S}(\mathbf{x} - (\mathbf{u} + \varepsilon \mathbf{w}))\mathbf{T}^{-1}(\mathbf{x}))]}{\partial \varepsilon} \right|_{\varepsilon=0} \right\} d\mathbf{x} \\ &= \frac{1}{4} \int_{\Omega} [(\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u})\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u}) - \mathbf{T}^{-1}(\mathbf{x}))_c]^T \cdot \nabla \mathbf{S}(\mathbf{x} - \mathbf{u})\mathbf{w}(\mathbf{x}) d\mathbf{x} \end{aligned} \quad (7)$$

where  $(\mathbf{X})_c$  denotes the  $pq \times 1$  column vector formed by concatenating the columns of a  $p \times q$  tensor  $\mathbf{X}$ , and  $\nabla \mathbf{S}(\mathbf{x} - \mathbf{u}) = [(\partial(\mathbf{S}(\mathbf{x}))_c/\partial x_1)(\partial(\mathbf{S}(\mathbf{x}))_c/\partial x_2) \dots (\partial(\mathbf{S}(\mathbf{x}))_c/\partial x_n)]$  taking the value at  $\mathbf{x} - \mathbf{u}$ . See Appendix A for details of this derivation. Therefore, the driving force  $\mathbf{F}(\mathbf{x}, \mathbf{u})$  can be defined as

$$\begin{aligned} \mathbf{F}(\mathbf{x}, \mathbf{u}) &= \nabla_{\mathbf{u}} E \\ &= \frac{1}{4} \left\{ [(\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u})\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u}) - \mathbf{T}^{-1}(\mathbf{x}))_c]^T \right. \\ &\quad \times \nabla \mathbf{S}(\mathbf{x} - \mathbf{u}) \left. \right\}^T \\ &= \frac{1}{4} [\nabla \mathbf{S}(\mathbf{x} - \mathbf{u})]^T \\ &\quad \times (\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u})\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u}) - \mathbf{T}^{-1}(\mathbf{x}))_c. \end{aligned} \quad (8)$$

### B. Applicability of the Symmetrized KL-Divergence to Higher-Order Diffusion Models

The Gaussian PDF in (1) is simply a first-order approximation of the diffusion behavior of water molecules within a voxel, assuming that the water diffusion at each voxel can be fully described by the covariance matrix of the Gaussian PDF. This is not the case in regions where fiber tracts are crossing within a single voxel, and a higher-order approximation is required, as illustrated in Fig. 1. Below we show that the symmetrized KL-divergence is extensible beyond this first-order diffusion model, if higher-order spherical harmonic terms are available, e.g., using diffusion spectrum imaging (DSI) [31],  $q$ -space reconstruction methods [32], or high angular resolution diffusion imaging (HARDI) [33]. We give the definition of symmetrized KL-divergence for higher-order diffusion profiles and compare it to other standard cost functions in HARDI data. The registration framework for HARDI data using the symmetrized KL-divergence is very similar to that for DTI, which we will implement and validate more comprehensively in a separate paper.

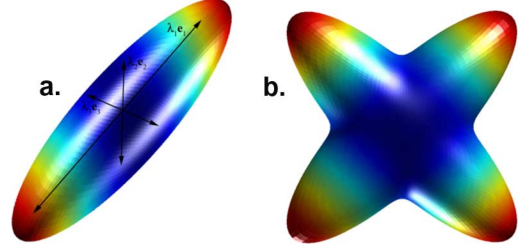


Fig. 1. (a) To a first-order approximation, water diffusion is modeled by a single Gaussian PDF. The covariance matrix or tensor, with eigenvalue-eigenvector pairs  $(\lambda_i, \mathbf{e}_i)$ , characterizes the diffusion properties, and can be visualized as an ellipsoid in 3-D space. (b) However, a single Gaussian model is unable to capture multidirectional water diffusion in brain regions where intravoxel fiber crossing is significant. In this case, the probability distribution for water diffusion, which may have a shape such as that displayed in (b), can be estimated from the MR signals, as in diffusion spectrum or  $q$ -ball imaging.

For higher-order diffusion profiles, the symmetrized KL-divergence can no longer be defined using a second-rank tensor model. Inspired by [34], we model the scalar diffusivity function (apparent diffusion coefficient)  $D$  as a probability density function. The diffusivity function  $D$  is given by the signal attenuation in a specific diffusion-encoded gradient direction  $\mathbf{g}$  based on the Stejskal-Tanner equation [35]

$$S(\mathbf{g}) = S_0 \exp(-bD(\mathbf{g})) \quad (9)$$

where  $b$  is the diffusion weighting factor, and  $\mathbf{g} = \mathbf{g}(\theta, \phi) = [\sin \theta \cos \phi \sin \theta \sin \phi \cos \theta]^T$ :  $\theta$  and  $\phi$  are the polar and azimuthal angles. The PDF is then defined by normalizing the integral of  $D$  over the spherical angle  $\Omega$  to 1

$$\begin{aligned} p(\theta, \phi) &= D(\theta, \phi) / \text{gtr}(D) \\ \text{gtr}(D) &= \int_{\Omega} D(\theta, \phi) d\Omega \end{aligned} \quad (10)$$

where  $\text{gtr}(D)$  is the generalized trace of  $D$  [34]. For two diffusivity functions  $D_p$  and  $D_q$ , similar to (4), we define the symmetrized KL-divergence on the unit sphere based on the corresponding PDFs  $p(\theta, \phi)$  and  $q(\theta, \phi)$

$$\begin{aligned} \text{sKL}(p, q) &= \frac{1}{2} \left\{ \int_{\Omega} p \log \left( \frac{p}{q} \right) d\Omega + \int_{\Omega} q \log \left( \frac{q}{p} \right) d\Omega \right\} \\ &= \frac{1}{2} \left\{ \int_{\Omega} p \log(p) d\Omega - \int_{\Omega} p \log(q) d\Omega \right. \\ &\quad \left. + \int_{\Omega} q \log(q) d\Omega - \int_{\Omega} q \log(p) d\Omega \right\}. \end{aligned} \quad (11)$$

Applying (10) to the integrals in (11), for example

$$\begin{aligned} \int_{\Omega} p \log(q) d\Omega &= \int_{\Omega} \frac{D_p}{\text{gtr}(D_p)} [\log(D_q) - \log(\text{gtr}(D_q))] d\Omega \\ &= \left[ \frac{1}{\text{gtr}(D_p)} \int_{\Omega} D_p \log(D_q) d\Omega \right] \\ &\quad - \log(\text{gtr}(D_q)) \end{aligned} \quad (12)$$

we have

$$\text{sKL}(p, q) = \frac{1}{2} \left\{ \frac{1}{\text{gtr}(D_p)} \left[ \int_{\Omega} D_p \log(D_p) d\Omega - \int_{\Omega} D_p \log(D_q) d\Omega \right] + \frac{1}{\text{gtr}(D_q)} \left[ \int_{\Omega} D_q \log(D_q) d\Omega - \int_{\Omega} D_q \log(D_p) d\Omega \right] \right\}. \quad (13)$$

Direct estimation of sKL in (13) is computationally expensive, but is faster if we expand the diffusivity functions  $D(\theta, \phi)$  in terms of a spherical harmonic (SH) series. See Appendix B for derivations.

### C. Reorientation of Diffusion Profiles

Unlike in the case of warping a scalar-valued image, adjustment of the direction of the diffusivity profiles is required when warping a tensor-valued image or HARDI, in order to keep the direction of diffusivity profiles oriented with the direction of the underlying fibers. For DTI, the tensor reorientation step has been well described in [22] and [36]. In this paper, tensor orientations were adjusted at each iteration of the fluid registration using the preservation of principal direction (PPD) method [22]. When image  $\mathbf{S}$  is deformed to match image  $\mathbf{T}$  with the mapping  $\mathbf{h} : \mathbf{T} \rightarrow \mathbf{S}$ , the eigenvectors of the tensor (ordered as  $\mathbf{e}_1$ ,  $\mathbf{e}_2$ , and  $\mathbf{e}_3$ ) at the lattice point  $\mathbf{x}$  of  $\mathbf{S}$  are distorted according to the Jacobian matrix  $\mathbf{J} = \nabla \mathbf{h}^{-1}(\mathbf{x})$ . Using the PPD procedure, the rotation matrix by which  $\mathbf{e}_1$  is rotated onto  $\mathbf{J}\mathbf{e}_1/|\mathbf{J}\mathbf{e}_1|$  and  $\mathbf{e}_2$  onto the plane spanned by  $\mathbf{J}\mathbf{e}_1/|\mathbf{J}\mathbf{e}_1|$  and  $\mathbf{J}\mathbf{e}_2/|\mathbf{J}\mathbf{e}_2|$ , was applied to the whole tensor.

In HARDI, however, a single diffusivity function may have multiple local maxima of diffusivity, with respect to spherical angle, that are computationally expensive to determine [37], and thus impractical to perform at every iteration during image warping. Here, we propose a fast algorithm to determine the principal direction of the diffusivity function, based on principal component analysis (PCA) of its shape [38]. At each direction sampled by the diffusion gradient  $\mathbf{g}(\theta_i, \phi_i)$ ,  $0 \leq i < n_s$ , where  $n_s$  is the number of gradient directions, we define a point with distance to the origin proportional to  $D(\theta_i, \phi_i)$ ,  $\mathbf{d}_i = D(\theta_i, \phi_i) \cdot \mathbf{g}(\theta_i, \phi_i) = (d_{i0}, d_{i1}, d_{i2})$ , where the last term is the Cartesian coordinates of  $\mathbf{d}_i$ , and the diffusivity function is assumed to be centered at the origin. When the diffusivity functions are sampled at radially symmetric angles,  $\mathbf{d}_i$  at opposing locations cancel, and therefore the mean vector  $\boldsymbol{\mu}$  of the point set  $\{\mathbf{d}_i\}$  is a zero vector. The covariance matrix  $\boldsymbol{\Sigma}$  of  $\{\mathbf{d}_i\}$  is given by

$$\boldsymbol{\Sigma} = \frac{1}{n_s} \sum_{i=0}^{n_s-1} \mathbf{d}_i \mathbf{d}_i^T. \quad (14)$$

The principal direction of the diffusivity function is determined by the first eigenvector of  $\boldsymbol{\Sigma}$ . The rotation matrix  $\mathbf{R}$ , which adjusts the direction of the diffusivity function, is then obtained using the PPD procedure, and the diffusivity is sampled discretely at the new gradient directions  $\mathbf{g}(\theta'_i, \phi'_i) = \mathbf{R} \cdot \mathbf{g}(\theta_i, \phi_i)$ .

It is advantageous to compute values of the reoriented diffusivity function in the *original* directions  $(\theta_i, \phi_i)$ , rather than the new ones  $(\theta'_i, \phi'_i)$  for a couple of reasons. First, we do not need to keep track of new reoriented gradient directions at each iteration of image warping. Second, the symmetrized KL-divergence (or other cost functions) must be computed from diffusivity functions that are sampled at the same gradient directions for the target image (which is fixed) and the source (which is moving and in which the diffusivity functions are being reoriented). Because the gradient directions are sampled discretely, the reoriented diffusivity functions are constructed by “pushing” the values sampled at the original directions to the new directions, so the values of the reoriented function at the original directions are not known. These new values can be computed from the SH series, as the basis functions  $Y_{lm}$  are continuous and defined at all spherical angles. If  $(\Delta\theta_i, \Delta\phi_i) = (\theta'_i - \theta_i, \phi'_i - \phi_i)$ , then the values of the reoriented diffusivity functions  $D'$  in the original directions are given by

$$D'(\theta_i, \phi_i) = D(\theta_i - \Delta\theta_i, \phi_i - \Delta\phi_i) = \sum_{l=0, \text{even}}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\theta_i - \Delta\theta_i, \phi_i - \Delta\phi_i) \quad (15)$$

where  $c_{lm}$  are the SH coefficients of the original diffusivity function  $D$  [see (B2)].

### D. Inverse-Consistent Fluid Regularization

We adopted the inverse-consistent regularization scheme proposed in [39], and extended it from an elastic to a fluid model to deal with large local deformations while maintaining one-to-one topology. We refer to [6], [40]–[42] for background on these fluid models, where essentially mappings regularized by linear elastic or Laplacian penalties are not guaranteed to be diffeomorphic (i.e., differentiable one-to-one mappings with smooth inverses). Because of this disadvantage of small-deformation models, smoothness can be guaranteed by moving to a fluid or other large-deformation models (e.g., LDDMM [43] or geodesic shooting [44]). Let  $\mathbf{h}(\mathbf{x}) = \mathbf{x} - \mathbf{u}_1(\mathbf{x})$  be the displacement field mapping image  $\mathbf{S}$  to image  $\mathbf{T}$  (in the forward direction), and  $\mathbf{h}^{-1}(\mathbf{x}) = \mathbf{x} - \mathbf{u}_2(\mathbf{x})$  be the mapping from  $\mathbf{T}$  to  $\mathbf{S}$  (in the backward direction). The fluid regularization scheme is then formulated by minimizing the following matching cost functional, regarding forward and backward velocities  $\mathbf{v}_1(\mathbf{x})$  and  $\mathbf{v}_2(\mathbf{x})$ , respectively

$$E_{\text{total}}(\mathbf{S}, \mathbf{T}) = \underbrace{E(\mathbf{u}_1, \mathbf{S}, \mathbf{T}) + \int_0^1 \int_{\Omega} \|\mathbf{L}\mathbf{v}_1(\mathbf{x})\|^2 d\mathbf{x} dt}_{E_1} + \underbrace{E(\mathbf{u}_2, \mathbf{T}, \mathbf{S}) + \int_0^1 \int_{\Omega} \|\mathbf{L}\mathbf{v}_2(\mathbf{x})\|^2 d\mathbf{x} dt}_{E_2} \quad (16)$$

where  $L = a\nabla^2 + b\nabla\nabla \cdot + cI$ ;  $I$  is the identity linear operator. We set  $a = \mu$  and  $b = \mu + \lambda$ , and  $c = 0$ , where  $\mu$  and  $\lambda$  are the viscosity constants ( $c$  is an inertial term that vanishes for very low Reynolds number flow, see [40]);  $E(\mathbf{u}_1, \mathbf{S}, \mathbf{T})$  and  $E(\mathbf{u}_2, \mathbf{T}, \mathbf{S})$  are given in (6). This cost functional is invariant to the order of  $\mathbf{S}$  and  $\mathbf{T}$ , i.e.,  $E_{\text{total}}(\mathbf{S}, \mathbf{T}) = E_{\text{total}}(\mathbf{T}, \mathbf{S})$ .

We adopted the “greedy descent” algorithm, which updates the mapping by perturbing the displacement fields from the previous time step. This type of approach does not update the whole time-dependent velocity field at each iteration, but accumulates velocity fields from the gradient at each iteration. The iterative process is an extension of that in [39] for elastic registration, and detailed in Appendix D.

### III. EXPERIMENTAL METHODS

#### A. DTI Acquisition

DTI datasets from 34 elderly adults (14F/20M, age  $73.6 \pm 9.0$  SD years) were analyzed. These images were acquired as part of a previous study where subject demographics are detailed [45]. DTI was performed with a 1.5 T Siemens Sonata MRI scanner using an optimized diffusion tensor sequence [46]. The imaging parameters were 43 axial slices, FOV = 23 cm, TR/TE 6000/106 ms, 2.5 mm slice thickness with a 0.25-mm gap, with an acquisition matrix of  $128 \times 128$  and 60 images acquired at each location consisting of 16 low ( $b = 0$ ) and 44 high diffusion-weighted images in which the encoding gradient vectors are uniformly radially distributed in space ( $b = 1100 \text{ s mm}^{-2}$ ) using the electrostatic approach described elsewhere [46]. The reconstruction matrix was  $256 \times 256$ , resulting in an in-plane resolution of  $0.89 \times 0.89 \text{ mm}^2$ . The total scan time was 8 min. A fractional anisotropy (FA) map was constructed from each DT image using the equation shown at the bottom of the page, where  $\lambda_1, \lambda_2, \lambda_3$  are eigenvalues [1].

#### B. Preprocessing of DTI

Each subject’s FA image was manually stripped of nonbrain tissues, yielding a binary brain-extraction mask (cerebellum included) for the DT images. The masked images were then coregistered with global scaling (affine nine-parameter transformation) to the upsampled FA image (dimension:  $256 \times 256 \times 132$  voxels, resolution:  $0.89 \times 0.89 \times 0.89 \text{ mm}^3$ ) from a randomly selected control subject. The resulting transformation parameters were used to rotationally reorient the tensor at each voxel [21], [22], and then affine align the DT images based on trilinear interpolation of the log-transformed tensors. As noted in [3], [5], [6], and in other Lie group work, transforming diffusion tensors using the matrix logarithm transformation and performing subsequent arithmetic operations in the logarithmic domain is beneficial as it makes the processing Euclidean. The interpolation of log-transformed tensors causes less “swelling,”

or loss of anisotropy, than direct interpolation of the original tensors. We used this interpolation method in the following fluid registration experiments.

#### C. Target Selection

As in several other studies, e.g., in [47] and [48], we elected to register each subject’s data to a typical single-subject target image included in the data set. This choice was motivated by the fact that, unlike a multisubject average intensity atlas target, a single-subject target typically has comparable contrast-to-noise, and spatial frequency content to the other images. However, to reduce the bias toward the particular geometry of a single subject, we adopted the “best individual target” (BIT) method proposed by Kochunov *et al.* [49]. (We note that arguments can be made in favor of using other templates, such as minimal deformation templates [50], or geodesic means [48], and the best approach may depend on the application). For each FA image serving as the target, a target quality score (TQS) [49] was computed, defined as the average voxel-wise product of the mean and the standard deviation of the deformation fields of all other subjects fluidly mapped onto the target by maximizing the Jensen-Rényi divergence [51]. Although our proposed fluid registration algorithm, which minimizes the symmetrized KL-divergence between two DT images, can also be used for target selection here, we decided to use an established scalar method for this step. The subject with the lowest TQS was then defined as the best individual target.

#### D. Fluid Registration of Diffusion Tensor Images

The DT images were registered respectively by minimizing the symmetrized KL-divergence under inverse-consistent fluid regularization. The time step  $\varepsilon$  was set to 0.02. In all of the following experiments,  $\mu$  and  $\lambda$  were set to 0.9 and 6.0 based on our earlier experiences fluidly registering scalar-valued MR images [51]. We have found that a relatively high setting for  $\lambda$  is useful to penalize extreme volume distortions in the deforming template, and to keep any such distortions spatially smooth. To facilitate convergence, we modified the hierarchical multiresolution scheme used in [52] to solve [in (D2)] with a 1/8 sampling of data along each axis ( $32 \times 32 \times 32$  voxels) for 2000 iterations, 1/4 subsampling ( $64 \times 64 \times 64$  voxels) for 1000 iterations, and 1/2 resolution ( $128 \times 128 \times 128$  voxels) for 50 iterations. At each resolution level, initial displacement vectors were obtained from the previous level using a cubic B-spline interpolation [53]. Iterative registration was stopped (or proceeded to the next resolution level) when the value of the symmetrized KL-divergence (either forward or backward direction of registration) was no longer monotonically decreasing. Fig. 2 shows the symmetrized KL-divergence plotted against the iteration number. The values of the symmetrized KL-divergence are consistent in both the forward and backward directions, and decrease smoothly. The

$$\text{FA} = \sqrt{((\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2) / (2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2))}$$

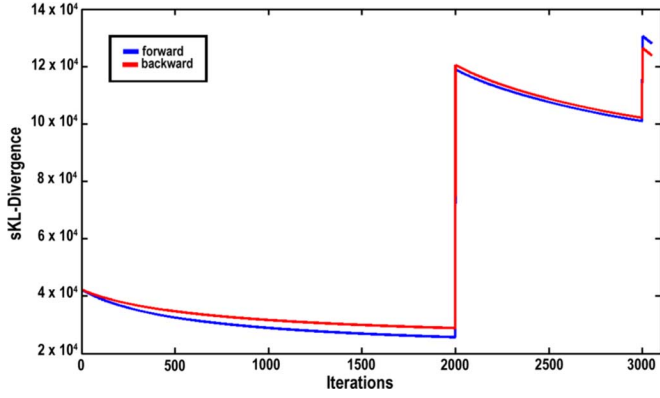


Fig. 2. The sKL-divergence for both the forward and backward fluid registration mappings decreases with increased iterations. The abrupt increases in the sKL-divergence are due to upsampling of the DTI to a higher resolution level in the multiresolution scheme.

final displacement vector fields were interpolated to the full resolution of the image. The registration processes were executed in parallel by submitting one process to one compute node of a 306-node, dual-processor SUN Microsystems V20z cluster. Each compute node has a dual 64-bit 2.4 GHz AMD Opteron CPU. The average computation time is about 5 h by a single processor, excluding preprocessing steps such as target selection.

#### IV. EVALUATION OF THE REGISTRATION ALGORITHM

##### A. Evaluation of Registration Accuracy

There is no widely-accepted gold standard for validation of nonlinear registration algorithms, so we evaluated the performance of our fluid mapping method on the basis of two criteria: 1) the similarity between the source and the target image is maximized (in terms of data agreement and the regularity of the solution), and 2) the variability of a population of deformed source images is minimized, when a mapping is optimized. To assess the accuracy of the mapping, we defined several overlap measures that were not explicitly optimized in the cost function, such as the inverse consistency (IC) on a test sample and several measures of agreement for the eigenstructures of the DTI before and after alignment. Clearly, each of these represents different desirable characteristics in the mappings that are arguably necessary but not sufficient for good registration. We used the tensor overlap index (OVL) [54] as a measure for the first criterion and compared the dyadic coherence of principal eigenvectors [21], [54] and the variance of FA, before and after fluid registration, to evaluate the second criterion.

*Measurement of Tensor Overlap:* We adopted the measure of overlap between eigenvalue–eigenvector pairs proposed by Basser and Pajevic [54] to define the matching of two diffusion tensors

$$\text{OVL} = \frac{\sum_{i=1}^3 \lambda_i \lambda'_i (\mathbf{e}_i \cdot \mathbf{e}'_i)^2}{\sum_{i=1}^3 \lambda_i \lambda'_i} \quad (17)$$

where  $(\lambda_i, \mathbf{e}_i)$  is the eigenvalue–eigenvector pairs. For each voxel,  $\text{OVL} = 0$  indicates no overlap and 1 indicates complete overlap of the three principal axes of diffusion tensors. The average of the OVL (aOVL) across all the voxels in the registered DTI within a region of interest (ROI) was used as an overall measure of the tensor alignment. There is a risk of circularity in defining a data fidelity term to assess how well the tensors are registered, as tensor agreement is formulated operationally using the registration cost function. We selected the eigenvalue–eigenvector metric as it is not identical to the cost function being used to drive the registration (other possible choices might include the inner products of the principal eigenvectors, after normalization). As noted earlier, our registration cost function is not based on the eigenvectors directly, but rather on the symmetrized KL-divergence of the diffusion PDFs (here simplified using Gaussian covariance functions to represent each PDF). As such it is a slightly more general model, given that the PDF approach is readily extensible to nonparametric or higher-order representations of the diffusion process, and to cases of fiber-crossing where fiber topography is not well represented by the eigenvectors of a first-order model.

*Dyadic Coherence of Principal Eigenvectors:* The variability within a group of tensors can be approximated by measuring the dispersion (or coherence) of individual principal eigenvectors from their mean vector. (Clearly, this is one of several measures of variance that can be defined for tensor-valued data.) This is achieved by analyzing the shape of the mean dyadic tensor of principal eigenvectors, given by [21], [54]

$$\langle \mathbf{e}_1 \mathbf{e}_1^T \rangle = \frac{1}{N} \sum_{i=1}^N \mathbf{e}_1^i \mathbf{e}_1^{iT} \quad (18)$$

where  $\mathbf{e}_1^i$  is the principal eigenvector of subject  $i$ . Then the dyadic coherence  $\kappa$  is defined as

$$\kappa = 1 - \sqrt{\frac{\beta_2 + \beta_3}{2\beta_1}} \quad (19)$$

where  $\beta_1, \beta_2$ , and  $\beta_3$  are eigenvalues of  $\langle \mathbf{e}_1 \mathbf{e}_1^T \rangle$  in descending order.  $\kappa = 0$  is the minimum, indicating that the principal eigenvectors are randomly scattered, and the mean dyadic tensor  $\langle \mathbf{e}_1 \mathbf{e}_1^T \rangle$  represents a sphere ( $\beta_1 = \beta_2 = \beta_3$ ). When the principal eigenvectors are perfectly aligned,  $\kappa$  reaches a maximum value of 1.

*Variance of FA:* We compared the differences in the intersubject variance of FA values between affine-registered and fluidly registered DTI (by affine-registered, we mean data aligned with only a nine-parameter linear spatial transformation, which adjusts for differences in overall brain scale). Although this measure does not invoke an independent source of ground truth, its use is based on the premise that any misregistration on nonhomologous fiber tracts will result in increases in the FA voxel-level variance in a group. As such it makes sense to see if this FA variance can be reduced using diffeomorphic flows that preserve the overall FA values in the image (as only local rotations are ever applied to the tensors—they are not scaled or



sheared). These comparisons help distinguish specific brain regions where cross-subject mismatches in fiber geometry are reduced by nonlinear registration. At each voxel, an  $F$ -statistic  $f = \sigma_1^2/\sigma_2^2$  was obtained based on the intersubject variance of FA under two different conditions (fluid registered versus affine-registered only), and was converted to a Student's  $t$ -statistic with  $(n-2)$  degrees-of-freedom, for paired comparison of data processed in the two different ways [55]

$$t = \frac{(f-1)\sqrt{n-2}}{\sqrt{4f(1-r^2)}}. \quad (20)$$

Here,  $n$  denotes number of subjects;  $r$  is the Pearson correlation coefficient. Values of  $t < 0$  indicate that the variance of the numerator (fluid-registered) is less than that of the denominator (affine-registered). We used color-coded maps to visualize  $t$ -values obtained from the last equation. Overall significance, correcting for the multiple spatial comparisons implicit in computing an image of statistics, was assessed by pFDR methods [56] (see Appendix E for details) for strong control over the likelihood of false rejections of the null hypotheses with multiple comparisons.

### B. IC Error of the Deformation Fields

If  $\mathbf{h}_{ST}(\mathbf{x})$  and  $\mathbf{h}_{TS}(\mathbf{x})$  are the forward and backward displacement vector fields computed for registering two images  $\mathbf{S}$  and  $\mathbf{T}$ , then the average and maximum IC errors within a ROI  $\Omega$  are defined as [52]

$$\begin{aligned} \text{Average IC error} &= \frac{1}{|\Omega|} \int_{\Omega} \|\mathbf{h}_{ST}[\mathbf{h}_{TS}(\mathbf{x})] - \mathbf{x}\| d\mathbf{x} \\ \text{Maximum IC error} &= \max_{\mathbf{x} \in \Omega} \|\mathbf{h}_{ST}[\mathbf{h}_{TS}(\mathbf{x})] - \mathbf{x}\|. \end{aligned} \quad (21)$$

Equation (21) was discretized for implementation, and trilinear interpolation was used when  $\mathbf{h}_{TS}(\mathbf{x})$  did not fall on a lattice point. This measure was used, following [52] to assess the internal consistency of the deformation fields, assuming that an ideal registration algorithm should set up correspondences that do not depend on the order of the images.

## V. RESULTS

### A. Inverse-Consistent Fluid Registration for DTI

Fig. 3 displays the forward (source to target) and backward (target to source) deformations of DTI based on the inverse-consistent fluid algorithm. The Jacobian determinant of the forward and backward deformation fields was confirmed to be greater than 0.6 everywhere in all subjects examined, confirming that the mappings were indeed diffeomorphic. The Jacobian tensor field, along with its determinant and rotation angle, is visualized on Fig. 4. Fiber tracts surrounding the lateral and third ventricles, e.g., the corpus callosum and cingulum, experienced the greatest magnitude of deformation. This is also shown by the tensor glyphs in Fig. 5, in which the diffusion tensors in the source image are displaced and reoriented to map the target. The average registered DTI (source to target) in Fig. 6 shows increased image sharpness (assessed visually) after fluid registration following affine registration, relative to the sharpness

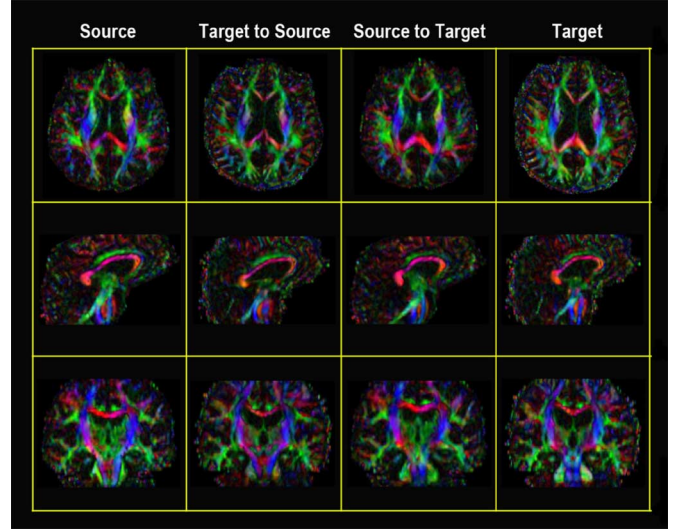


Fig. 3. FA-weighted principal eigenvector maps showing fluid mappings of two DT images by minimizing the symmetrized KL-divergence. Eigenvectors are shown, as is conventional, using an RGB color code to represent the orientation of the normalized principal eigenvector relative to the medial-lateral (coded in red), anterior-posterior (coded in green), and superior-inferior (coded in blue) axes of the anatomical reference frame. The intensity of these colors is further weighted using the FA. The forward (source to target) and backward (target to source) registration endpoints were simultaneously achieved, based on the inverse-consistent property of our algorithm.

obtained after using affine registration only. Moreover, the variance of FA across the subjects is reduced after fluid registration (pFDR = 0.0061), especially in major WM fiber tracts with high diffusion anisotropy. These suggest that fluid registration further reduces the intersubject morphological variability. Fig. 7 depicts how registration performance depends on the fluid viscosity constants  $\mu$  and  $\lambda$  by measuring the aOVL between the deformed source and the target DTI and the average IC error of the deformation fields. A high setting of  $\mu$  and  $\lambda$  (e.g.,  $\mu = 0.9$  and  $\lambda = 6.0$ ) is favorable as it results in high aOVL and low IC error. Fig. 8 shows that our algorithm on average achieves subvoxel error in terms of IC when applied to noisy images and allowing large deformations (average IC error =  $0.09 \pm 0.02$  voxel; maximum IC error =  $1.24 \pm 0.39$  voxel). The IC error maps indicate that greater errors occur in the regions where greater magnitudes of deformation are required to map the target, such as the periventricular areas (see Figs. 3 and 4). Nevertheless, IC error is small (typically less than a voxel) in these areas too. At each voxel, the average IC error increases with the magnitude of the deformation vector [Fig. 8(b) and (c)]. This implies that in some cases there may be a trade-off between IC and registration accuracy. For instance, if the registration does nothing (i.e., the displacement field is zero everywhere) then the maximum and average IC errors will be zero too. Therefore, while low IC error is evidence of correct implementation of a registration strategy, it does not necessarily indicate high registration accuracy.

### B. Evaluation of Registration Performance

The OVL maps in Fig. 9 show that the major fiber structures are well matched in the cerebrum (e.g., the corpus callosum,

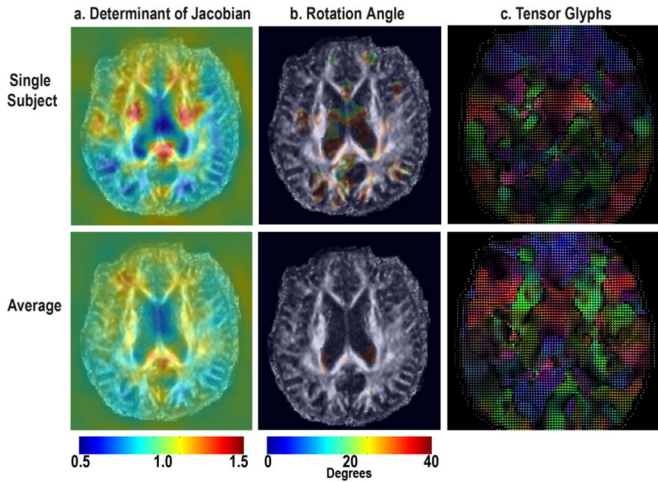


Fig. 4. Visualization of the Jacobian tensor of the registration deformation field (for the forward direction of fluid registration) in one subject (upper row), and the average tensor across all subjects (lower row). Because the Jacobian tensor is not generally symmetric, we computed the corresponding strain tensor, defined as  $\Sigma = \mathbf{J}^T \mathbf{J} = \mathbf{R} \mathbf{S}^2 \mathbf{R}^T$ . The average tensor was obtained by averaging the strain tensors of all subjects in the log-Euclidean space, and exponentiating the result. (a) The determinant of the Jacobian tensor is  $\det(\Sigma^{1/2}) = \det(\mathbf{S})$ . (b) The rotation angle was estimated by finding the smallest angle between the principal eigenvector of  $\Sigma$  and the  $x$ -,  $y$ -, and  $z$ - axes. If the deformation does not contain any local rotations, the principal eigenvector of  $\Sigma$  will be aligned with the  $x$ -,  $y$ -, and  $z$ - axes. The principal eigenvector of  $\Sigma$  is poorly defined when the Jacobian tensor is approximately isotropic, so we displayed the rotation angles only in regions where the Jacobian tensor was somewhat anisotropic, which we defined as when the fractional anisotropy of the Jacobian (strain) tensor  $> 0.3$ . (Note that the fractional anisotropy of the Jacobian tensor is different from the fractional anisotropy of the diffusion tensors). (c) To better visualize the Jacobian (strain) tensor, we modeled the tensor as an ellipsoid, using the visualization software “BrainSuite” (<http://www.loni.ucla.edu/Software/>) [62]. The RGB color coding is the same as in Fig. 3. Compared with the Jacobian determinant and rotational matrix maps from the single subject, the determinant of the *average* Jacobian (strain) tensor is closer to 1, the rotation angle is smaller, and the average tensor is more isotropic. Recent work has used this average strain map as a criterion to optimize a population average brain imaging template [63], and the accompanying average strain maps can be used, in general, as an index of how much a registration target deviates from the population.

cingulum, and internal capsule), as is the WM of the brainstem where the values of OVL are close to 1. The aOVL measure of correspondence in the DTI data increases after fluid registration based on minimization of the symmetrized KL-divergence (aOVL before/after fluid registration =  $0.448 \pm 0.015/0.494 \pm 0.012$ ,  $P < 0.0001$ ; paired- $t$  tests). Moreover, the OVL maps demonstrate that the improvements in tensor alignment between linear and nonlinear registration occur mostly in the WM regions (defined as  $\text{FA} > 0.3$ , see [57]), with aOVL (before/after fluid registration) =  $0.567 \pm 0.032/0.672 \pm 0.021$  ( $P < 0.0001$  based on paired- $t$  tests).

Fig. 10 shows that fluid registration increases the dyadic coherence of the principal eigenvectors. These findings are complementary to the OVL findings in Fig. 9.

We further compared the registration accuracy based on symmetrized KL-divergence versus other two cost functions, the inner product of diffusivity functions in (B3) with and without the linear ( $l = 0$ ) term [21], [58], denoted by  $\text{IP}_L$  (inner product with the linear term) and  $\text{IP}_D$  (inner product with the linear term deleted).  $\text{IP}_D$  is designed to compare only the anisotropic part of the diffusivity functions [21]. Under the assumption of Gaussian

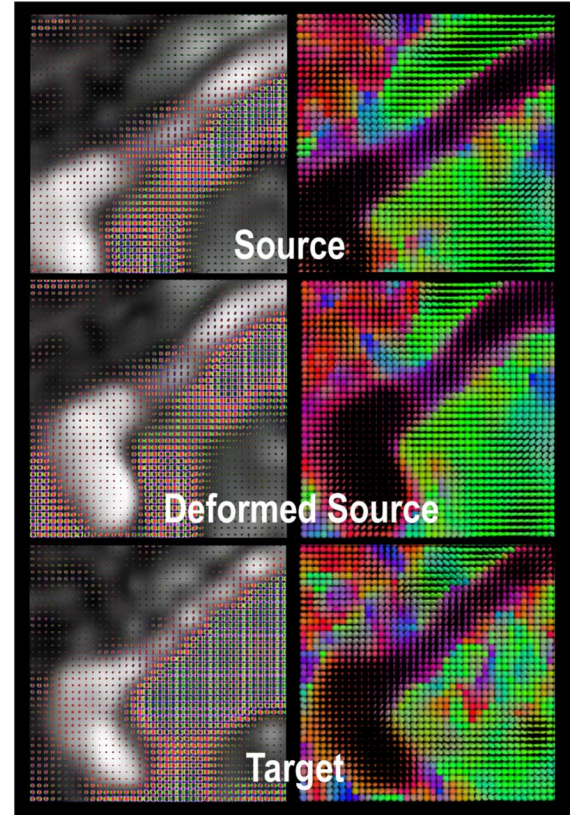


Fig. 5. Ellipsoidal tensor glyphs visualize fluid registration in tensor fields. The color of a tensor ellipsoid was determined by projections of the principal eigenvector onto the medial-lateral (coded in red), anterior-posterior (coded in green), and superior-inferior (coded in blue) axes of the anatomical reference frame, such that the color of the ellipsoid is independent of the view. We show tensor glyphs overlaid on the corresponding FA image, using the DTI visualization software, Camino, developed at University College, London [64] (left column), and also the ones that are fitted as colored ellipsoids (right column), as the former ones made using Camino present more information by overlaying tensor glyphs on the corresponding FA images, while the latter ones make it easier to see consistent patterns of fiber directions. The source and the target DT images are the same as in Fig. 3. Close-up views of the posterior corpus callosum for source, deformed source, and target DT images demonstrate that the tensor ellipsoids are displaced during the shape deformation.

diffusion,  $\text{IP}_L$  and  $\text{IP}_D$  have a closed form in terms of the diffusion tensors [21]. See Appendix F for the derivation of the driving force for fluid registration. Fig. 11 shows that the aOVL measure of correspondence in DTI data after fluid registration is higher using symmetrized KL-divergence as the cost functional than using  $\text{IP}_D$  (symmetrized KL-divergence:  $0.494 \pm 0.012$ ,  $\text{IP}_D$ :  $0.489 \pm 0.012$ ,  $P < 0.0001$ ; paired- $t$  tests). This difference is greater in the WM regions (defined as voxels where  $\text{FA} > 0.3$ ), with aOVL (symmetrized KL-divergence/ $\text{IP}_D$ ) =  $0.672 \pm 0.021/0.658 \pm 0.022$  ( $P < 0.0001$  based on paired- $t$  tests). For the registration experiments using  $\text{IP}_L$  as the cost function, the gradient field was not robust enough to keep  $\text{IP}_L$  constantly increasing.

### C. Application of the Symmetrized KL-Divergence to Higher-Order Diffusivity Models

*Synthetic examples:* We constructed a two-fiber diffusivity function using two orthogonal Gaussian tensors,  $\mathbf{T}_0$  and  $\mathbf{T}_1$ , with typical eigenvalues for WM fibers [25], [26]. We set



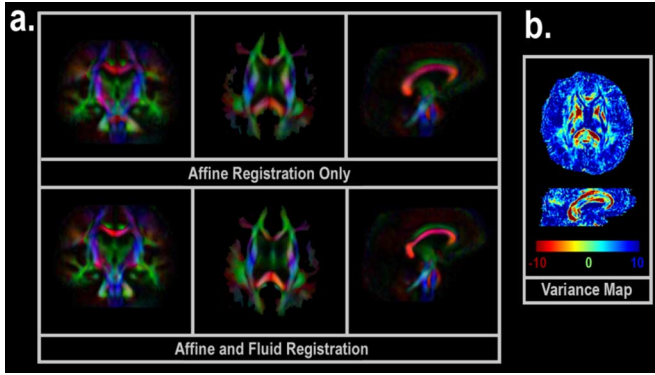


Fig. 6. (a) Average DTI obtained from registering all the subjects (the subject who served as the target was not included) to the target DTI displayed in Fig. 3, based on affine alignment (nine-parameter) only (top row), or affine followed by fluid registration (bottom row). Increased sharpness of the average DTI after fluid registration suggests that the fluid algorithm helps to eliminate local nonlinear morphological variability across subjects. (b) Color-coded  $t$ -maps obtained from comparing the variance of FA before and after fluid registration. Although a  $t$ -value greater than 4 already indicates a highly significant reduction in variance (this corresponds to light-blue areas, with  $P < 0.0005$  in one-tailed paired comparisons; the degrees-of-freedom = number of subjects  $- 2 = 31$ , and the subject serving as the target is not included), the color bar of the  $t$ -values is exaggerated to show that fluid registration reduces the variance of FA to a greater degree in the periventricular areas.

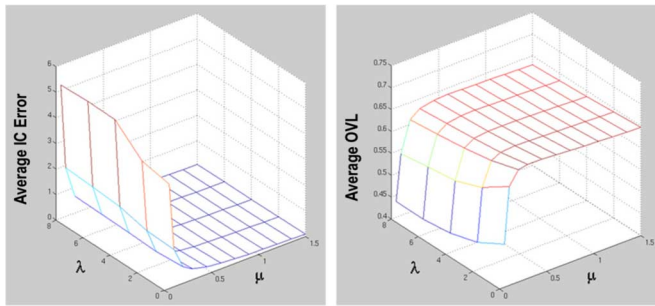


Fig. 7. Average IC error and the aOVL at different settings of fluid viscosity constants  $\mu$  and  $\lambda$ . The performance of the fluid algorithm is stable, in terms of high aOVL and low IC error, when  $\mu$  is greater than 0.5.

$\mathbf{T}_0 = \text{diag}(200, 200, 1700) \times 10^{-6}$  (mm<sup>2</sup>/s);  $\mathbf{T}_1$  was obtained by rotating  $\mathbf{T}_0$  90° around the  $y$ -axis. We also generated an isotropic gray-matter (GM) diffusivity function by setting  $\mathbf{T}_{\text{iso}} = \text{diag}(700, 700, 700) \times 10^{-6}$  (mm<sup>2</sup>/s). Then the diffusivity function  $D$  is given by

$$D = -\frac{1}{b} \sum_{i=0}^{n-1} p_i \exp(-b \mathbf{g}^T \mathbf{T}_i \mathbf{g}) \quad (22)$$

where  $n$  is the number of fibers.  $p_0 = p_1 = 0.5$  for two-fiber structures and  $p_0 = 1$  for the isotropic diffusivity function. The gradient direction vector  $\mathbf{g}$  consisted of 162 sampled points determined using an electrostatic approach [46]. Different  $b$  values (500, 1500, 3000) s/mm<sup>2</sup> were tested. The order of the SH series ( $l_m$ ) was set to 8. We compared the symmetrized KL-divergence with the inner product of diffusivity functions,  $\text{IP}_L$  (with the linear term) and  $\text{IP}_D$  (without the linear term), on synthetic examples that were noise-free, or with Rician noise added to MR signals  $S$  such that the signal-to-noise ratio (SNR) was 35

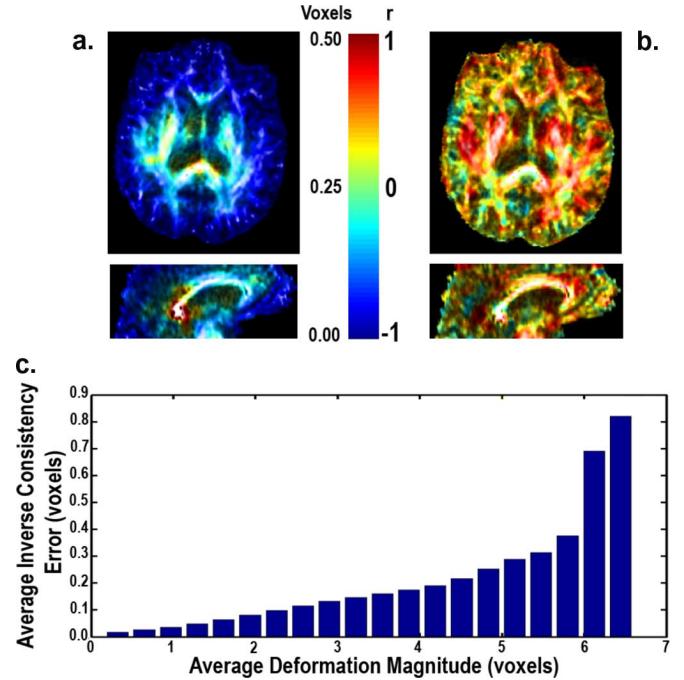


Fig. 8. (a) Color-coded maps are intended to display the spatial distribution of the IC error (upper panel: transverse view; lower panel: sagittal view). The greatest IC error is found in the splenium of the corpus callosum where a large warping of the underlying anatomy is required to map the target (see Fig. 3). This suggests, as might be naturally expected, that the IC error increases with the magnitude of the required deformation. This trend is more clearly demonstrated in the color-coded maps for the correlation coefficient  $r$  between the IC error and the magnitude of deformation in (b), and when the average IC error is plotted against the magnitude of deformation in (c)

or 10 [59]. Two identical two-fiber synthetic diffusivity functions served as the source and the target objects, with the source object rotating from 0 (complete overlap) to 90°. The three metrics were normalized for comparisons (see Fig. 12 for details).

Fig. 12 shows that at different noise levels, the symmetrized KL-divergence detects angular discrepancies in diffusivity functions more sensitively than  $\text{IP}_L$  and  $\text{IP}_D$ , in low  $b$ -settings. The symmetrized KL-divergence is also still comparable in performance with  $\text{IP}_D$ , at a high  $b$ -value. The performance of the symmetrized KL-divergence is therefore relatively stable under various diffusion weightings, and is applicable under ordinary MRI/DTI acquisition settings, though a high  $b$ -value can better detect higher-order angular structures in WM fibers at the cost of a decreased SNR [33].

**HARDI Data:** We further compared the three cost functions, which were summed over all voxels, in the 3-D HARDI data. The HARDI data was acquired from a healthy 22-year-old man imaged as part of a research study of twins on a 4 T Bruker Medspec MRI scanner using an optimized diffusion tensor sequence [46]. Imaging parameters were: 21 axial slices (5 mm thick), FOV = 24 cm, TR/TE 6090/104.5 ms, 0.5-mm gap, with a  $128 \times 100$  acquisition matrix and 30 images acquired at each location: 3 low ( $b = 0$ ) and 27 high diffusion-weighted images in which the encoding gradient vectors were uniformly radially distributed in space ( $b = 1100$  s/mm<sup>2</sup>) using the electrostatic approach in [46]. The reconstruction matrix was  $128 \times 128$ , yielding a  $1.875 \times 1.875$  mm<sup>2</sup> in-plane resolution. The total scan

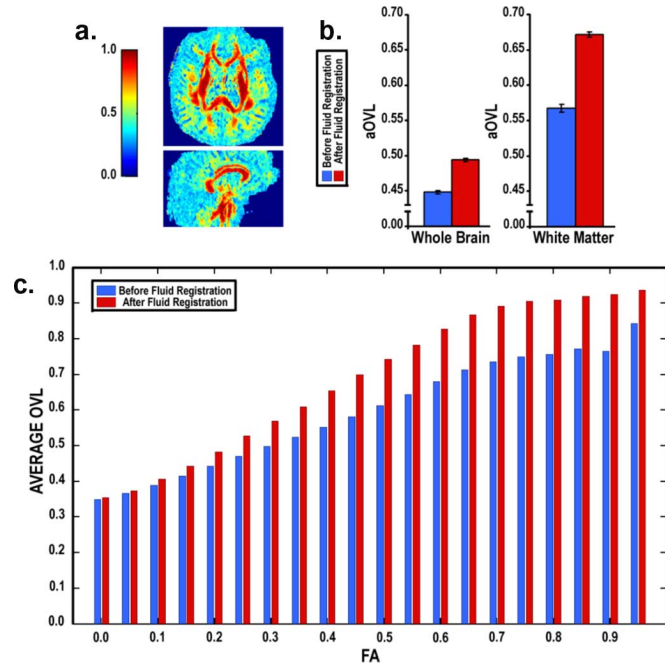


Fig. 9. (a) OVL maps with OVL averaged across subjects fluidly registered to the target DTI, color-coded to visualize the spatial distribution of the matching between diffusion tensors. Red-colored areas (OVL close to 1) outline the major fiber structures in the cerebral WM and brainstem, indicating that the diffusion tensors are well matched in these areas. (b) Bar graphs show the change of aOVL before and after fluid registration. Fluid registration increases the aOVL to a greater degree in the WM regions (which were defined, arbitrarily, as regions where  $FA > 0.3$  [57]). This is more clearly demonstrated when the average OVL (averaged across all voxels with a given FA value) is plotted against the values of FA in (c).

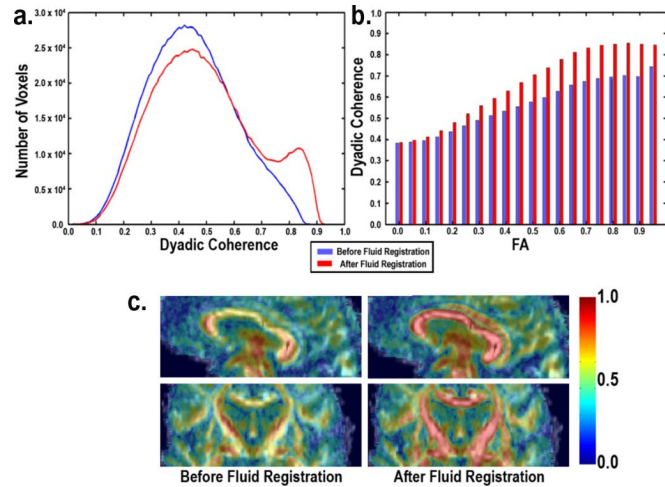


Fig. 10. (a) The histogram of the dyadic coherence between principal eigenvectors of the DT images shows that the dyadic coherence increases, or the variability of the diffusion tensors decreases after fluid registration, especially in highly anisotropic regions where  $FA > 0.3$ , as displayed in (b). (c) This is better visualized using the color-coded maps, with red-colored areas indicating that the dyadic coherence is close to 1.

time was 3.09 min. We set  $l_m = 4$  for the spherical harmonic analysis. The HARDI scan was compared with the same scan but rotated by angle  $\varphi$  ( $\varphi$  starts from  $0^\circ$ , which means complete overlap, up to  $\pm 20^\circ$ ), with the symmetrized KL-divergence and the inner products ( $IP_L$  and  $IP_D$ ) computed at every  $2^\circ$ .

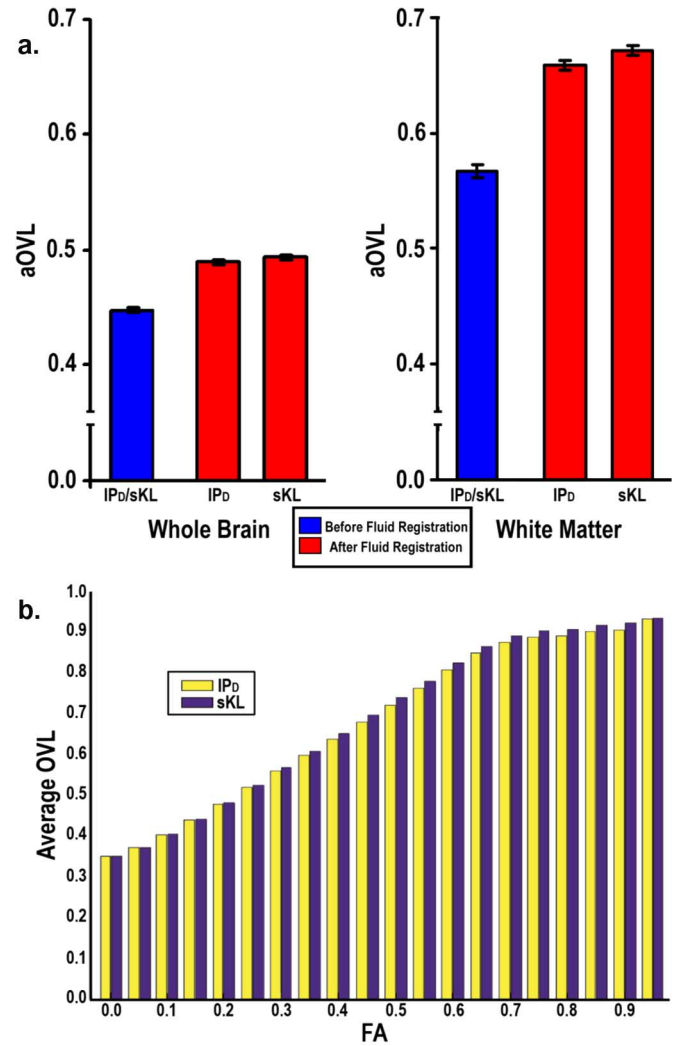


Fig. 11. Comparisons of tensor overlap after fluid registration between using symmetrized KL-divergence and using  $IP_D$  as the cost function. (a) The bar graphs show that the aOVL is higher using symmetrized KL-divergence than  $IP_D$  ( $0.494 \pm 0.012$  versus  $0.489 \pm 0.012$ ,  $P < 0.0001$ ), especially in the WM regions, defined as regions where  $FA > 0.3$  ( $0.672 \pm 0.021$  versus  $0.658 \pm 0.022$ ,  $P < 0.0001$ ). (b) This difference in the average OVL (averaged across all voxels with a given FA value) is plotted against the values of FA.

Fig. 13 shows that the symmetrized KL-divergence has a very sharp gradient near the optimal solution, and is sensitive enough to detect a  $2^\circ$  deviation in the images. The symmetrized KL-divergence is, therefore, a good candidate cost function for HARDI registration.

## VI. DISCUSSION

In this paper, we proposed an inverse-consistent fluid registration algorithm, based on minimizing the symmetrized KL-divergence of the Gaussian PDFs between two DT-MR images, in which the diffusion tensor at each voxel was modeled as a spatial covariance matrix. Unlike the traditional tensor distance based on the Frobenius norm, this cost function based on the symmetrized KL-divergence (5) is affine invariant [15]. This invariance is helpful as it makes the computed correspondences between images robust to errors in estimating the linear (global alignment) term. Fiber structures between source and target DT images were well matched, especially in highly anisotropic areas, where the DTI signals convey the greatest information.

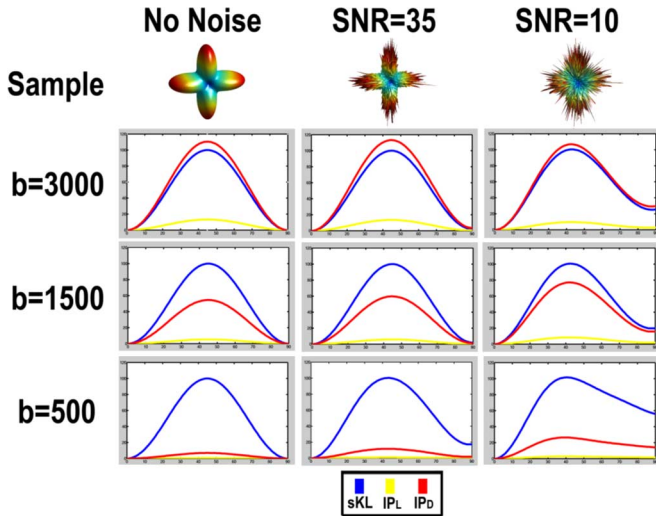


Fig. 12. Two identical synthetic diffusivity samples (no noise or with Rician noise added) were initially overlapped and rotated by  $\varphi = 0-90^\circ$ . We compared the differences between the rotated and nonrotated samples with the sKL and IP,  $IP_L$  (with  $l = 0$  term), and  $IP_D$  (without  $l = 0$  term). In noise-free samples,  $\varphi = 45^\circ$  gives the maximum sKL and minimum IP values. To facilitate comparisons, sKL and IP values have been normalized such that the normalized  $sKL(\varphi) = 100 \times sKL(\varphi)/sKL(\varphi = 45^\circ)$ , and normalized  $IP(\varphi) = 100 \times [1 - IP(\varphi)/IP(\varphi = 0^\circ)]$ .

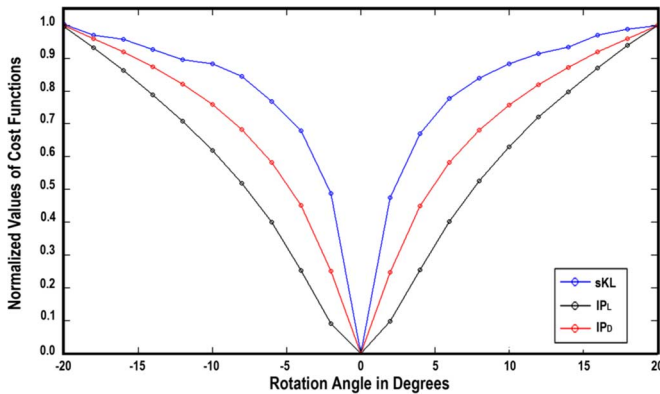


Fig. 13. Comparisons of the changes in sKL and IP ( $IP_L$  and  $IP_D$ ) at different rotation angles  $\varphi$  (from  $-20^\circ$  to  $+20^\circ$ , in increments of  $2^\circ$ ) for two identical HARDI scans. sKL and IP values have been normalized, with normalized  $f(\varphi) = \text{abs}[(f(\varphi) - f(\varphi = 0^\circ))/(f(\varphi = 20^\circ) - f(\varphi = 0^\circ))]$ . Angular profile of sKL is very sharp, and can detect rotational deviations of the image, with a magnitude as small as  $2^\circ$ .

Relative to approaches that aim to align the eigenvectors of the local diffusion process, we show that our approach is more generalizable, as the symmetrized KL-divergence is extendable to model higher-order water diffusivity profiles, by defining it as a PDF on the unit sphere and expanding the PDF using spherical harmonics. The symmetrized KL-divergence performs slightly but significantly better than  $IP_D$  (inner product of diffusivity functions with isotropic components deleted) as a cost functional for DTI registration in terms of tensor overlap. The OVL measure favors  $IP_D$ , but  $IP_D$  and OVL are somewhat correlated in gauging the alignment of eigenvalue-eigenvector pairs. We found that using  $IP_L$ , i.e., inner product with the linear term, did not provide a good gradient to maximize overlap between two DT images, which may indicate that introducing isotropic components of the diffusion

tensors into a cost functional makes the cost functional less sensitive for accurate matching of DTI [21]. For HARDI, the symmetrized KL-divergence is more robust than standard inner product measures for detecting small rotational deviations between HARDI data at various diffusion weights and noise levels, making it an attractive measure for HARDI registration. In HARDI, we used the PPD procedure to adjust the direction of the diffusivity profiles, which is arguably more accurate than other methods (e.g., finite strain [21], [22]) given that it takes the originally estimated fiber directions into account. PCA determines one eigendirection, so it is appropriate both for diffusivity functions with a single global maximum, and for those with a dominant local maximum relative to other small local maxima. However, diffusivity functions may have multiple local maxima, e.g., in regions where fibers cross (e.g., the synthetic samples in Fig. 12), or when there is noise-free purely isotropic diffusion. In such cases, the principal direction estimated by PCA becomes arbitrary, and the simple PPD procedure may not be applicable in these regions. At this point, the finite strain method is advantageous as it extracts the rotational component from the Jacobian of the deformation fields, and the extracted rotational matrix can be used to reorient the diffusivity functions by resampling the SH series in (15). Nevertheless, the finite strain method discards some deformation components (e.g., shearing and stretching), and is accurate only when the image deformations are small, or when large deformations are decomposed into small ones (e.g., the piecewise affine algorithm in [21]).

Our algorithm is similar to that of Cao *et al.* [23], [24] for nonlinear warping of DTI data, using image deformations represented by flows that are guaranteed to be diffeomorphic (i.e., smooth mappings with smooth inverses). Cao *et al.*'s approaches have provided the key ingredient of adapting a scalar registration method to a tensor registration method. The cost function in our work is based on information theory, and we adopted an inverse-consistent fluid model to guarantee that our transformations are diffeomorphic. This work forms part of a general improvement in nonlinear registration methods that have recently been modified to allow 1) large diffeomorphic deformations (e.g., large deformation diffeomorphic metric mapping (LDDMM) [43] or geodesic shooting [44]), and 2) to achieve inverse-consistent mappings (i.e., maps where correspondences do not depend on the order of the two images being matched [39], [52]). The symmetrized KL-divergence, as a cost metric, can also be combined into the LDDMM scheme for DTI registration [24]. Nevertheless, our algorithm differs fundamentally from those by Cao *et al.* in that 1) the cost functional in our work is based on information theory rather than aligning principal fiber directions [23] or minimizing the Frobenius distance between two tensors [24]; 2) we showed that the symmetrized KL-divergence is extendable to any representation of the PDF that models higher-order water diffusion, whereas Cao *et al.*'s approaches assume that water diffusion in each voxel can be described by the dominant orientation or the three eigenvalue-eigenvector pairs of a rank-2 tensor; and 3) in Cao's methods the tensor reorientation is explicitly optimized based on the PPD procedure, as eigenvectors of the tensors are part of the cost functional [21]. Because our approach can deal with generalized diffusion PDFs, it is likely to provide more reasonable assessments of signal matching in WM regions that may not have a single, well-defined eigenstructure. Ultimately,

the best cost function should be the one that decreases error variance in the specific signal that is analyzed after the registration step, which differs in different applications of DTI (e.g., FA mapping versus statistics on fiber orientation parameters). In computing the symmetrized KL-divergence, it may help to incorporate tensor reorientation into the cost functional, e.g., by incorporating the dyadic expression for the diffusion tensor,  $\mathbf{D} = \sum_{i=1}^3 \lambda_i \mathbf{e}_i \mathbf{e}_i^T$ , into (6). This explicit reorientation, as used in [21] and [24], optimizes metrics comparing tensors as a whole, in which the gradients of the cost function incorporate the effect of tensor reorientation. Nevertheless, the symmetrized KL-divergence in (4) is a generalized metric and can be applied to higher-order diffusion imaging data (e.g., HARDI data), in which the closed form for explicit tensor reorientation may not exist. Moreover, in Fig. 2, we show that our implicit tensor reorientation scheme (where tensor reorientation is followed by each iteration of the fluid deformation) still ensures that the values of the KL-divergence for both the forward and backward deformations are smoothly decreasing, although we note that, compared with the explicit reorientation where the cost functional is minimized under the effect of reorientation, this implicit tensor reorientation scheme does not preclude the registration process from being trapped in local minima.

Finally, we should note some conceptual limitations of matching DTI. As with matching cortical gyri in standard structural images, there may not be a homology across subjects, between fibers or even major tracts at the gyrus by gyrus level, but some pragmatic normalization is necessary for multisubject mapping studies. Moreover, it might be problematic to match isotropic tensors, or those with very low anisotropy, e.g., in the gray matter and CSF, as good tract correspondence across subjects may not exist or may not be inferable from the available data [57]. In these cases, whether or not registration is successful depends on the goals and sensitivity of the analysis. Nevertheless, it is possible to quantify practical improvements in signal detection for clinically significant alterations in DTI parameters, without requiring a gold standard for matching fibers at the individual level.

## APPENDIX A

### Derivation of (7)

If matrix  $\mathbf{Y}$  is the function of matrix  $\mathbf{X}$ , i.e.,  $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$ , and  $\mathbf{A}$  is a constant matrix, from linear algebra we have  $(d(\text{tr}(\mathbf{A}\mathbf{Y}^{-1}))/d\mathbf{X}) = -((\mathbf{Y}^{-T}\mathbf{A}^T\mathbf{Y}^{-T})_c)^T(d\mathbf{Y}/d\mathbf{X})$ , and  $(d(\text{tr}(\mathbf{Y}\mathbf{A}))/d\mathbf{X}) = ((\mathbf{A}^T)_c)^T(d\mathbf{Y}/d\mathbf{X})$ . Using the above equations and the chain rule, we obtain

$$\begin{aligned} & \left. \frac{\partial [\text{tr}(\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - (\mathbf{u} + \varepsilon\mathbf{w})))]}{\partial \varepsilon} \right|_{\varepsilon=0} \\ &= \left. \frac{d[\text{tr}(\mathbf{T}\mathbf{S}^{-1}(\mathbf{z}))]}{d\mathbf{z}} \right|_{\mathbf{z}=\mathbf{x}-(\mathbf{u}+\varepsilon\mathbf{w})} \left( \left. \frac{\partial \mathbf{z}(\varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0} \right) \\ &= -((\mathbf{S}^{-T}(\mathbf{x} - \mathbf{u})\mathbf{T}(\mathbf{x})\mathbf{S}^{-T}(\mathbf{x} - \mathbf{u}))_c)^T \\ &\quad \times \nabla \mathbf{S}(\mathbf{x} - \mathbf{u})(-\mathbf{w}(\mathbf{x})) \\ &= ((\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u})\mathbf{T}(\mathbf{x})\mathbf{S}^{-1}(\mathbf{x} - \mathbf{u}))_c)^T \\ &\quad \times \nabla \mathbf{S}(\mathbf{x} - \mathbf{u})\mathbf{w}(\mathbf{x}). \end{aligned} \quad (\text{A1})$$

and

$$\begin{aligned} & \left. \frac{\partial [\text{tr}(\mathbf{S}(\mathbf{x} - (\mathbf{u} + \varepsilon\mathbf{w}))\mathbf{T}^{-1}(\mathbf{x}))]}{\partial \varepsilon} \right|_{\varepsilon=0} \\ &= \left. \frac{d[\text{tr}(\mathbf{S}(\mathbf{z})\mathbf{T}^{-1})]}{d\mathbf{z}} \right|_{\mathbf{z}=\mathbf{x}-(\mathbf{u}+\varepsilon\mathbf{w})} \left( \left. \frac{\partial \mathbf{z}(\varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0} \right) \\ &= ((\mathbf{T}^{-T}(\mathbf{x}))_c)^T \nabla \mathbf{S}(\mathbf{x} - \mathbf{u})(-\mathbf{w}(\mathbf{x})) \\ &= -((\mathbf{T}^{-1}(\mathbf{x}))_c)^T \nabla \mathbf{S}(\mathbf{x} - \mathbf{u})\mathbf{w}(\mathbf{x}). \end{aligned} \quad (\text{A2})$$

Here we use the property  $(\mathbf{S}^{-1})^T = (\mathbf{S}^T)^{-1} = \mathbf{S}^{-1}$  for a symmetric matrix  $\mathbf{S}$ . Equation (7) is then derived by taking (A1) and (A2) together.

## APPENDIX B

### Computing $s\text{KL}$ in terms of Spherical Harmonic Series

The diffusivity function  $D(\theta, \phi)$  can be expanded in terms of a spherical harmonic (SH) series [25], [26]

$$\begin{aligned} D(\theta, \phi) &= \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi), \\ Y_l^m(\theta, \phi) &= \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos \theta) e^{im\phi}. \end{aligned} \quad (\text{B1})$$

Here,  $P_l^m(\cos \theta)$  are the associated Legendre polynomials.  $D(\theta, \phi)$  is real and radially symmetric, so it is sufficient to adopt a real basis function set  $Y_{lm}$  while retaining the orthonormality of  $Y_{lm}$  [26]

$$\begin{aligned} D(\theta, \phi) &= \sum_{l=0, \text{even}}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\theta, \phi) \\ Y_{lm} &= \begin{cases} \frac{\sqrt{2}}{2} ((-1)^m Y_l^m + Y_l^{-m}), & \text{if } -l \leq m < 0 \\ Y_l^0, & \text{if } m = 0 \\ \frac{\sqrt{2}i}{2} ((-1)^{m+1} Y_l^m + Y_l^{-m}), & \text{if } 0 < m \leq l \end{cases} \\ &\quad l = 0, 2, 4, \dots \end{aligned} \quad (\text{B2})$$

As the  $Y_{lm}$  are orthonormal, the inner product of the real functions  $D_1$  and  $D_2$  can be expressed in terms of their SH coefficients ( $c_{lm}^{D_1}$  and  $c_{lm}^{D_2}$ )

$$\int_{\Omega} D_1(\theta, \phi) D_2(\theta, \phi) d\Omega = \sum_{l=0, \text{even}}^{\infty} \sum_{m=-l}^l c_{lm}^{D_1} c_{lm}^{D_2}. \quad (\text{B3})$$

Moreover, it can be shown that (see Appendix C for derivations)

$$\int_{\Omega} D(\theta, \phi) d\Omega = 2\sqrt{\pi} c_{00}. \quad (\text{B4})$$



Given  $D_j = \sum_{l=0,\text{even}}^{\infty} \sum_{m=-l}^l c_{lm}^j Y_{lm}$  and  $\log(D_j) = \sum_{l=0,\text{even}}^{\infty} \sum_{m=-l}^l d_{lm}^j Y_{lm}$ , where  $j \in \{p, q\}$ , sKL in (13) can be expanded in terms of the SH series, as follows:

$$\text{sKL}(p, q) = \frac{1}{2} \left\{ \frac{1}{c_{00}^p} \sum_{l=0,\text{even}}^{\infty} \sum_{m=-l}^l [c_{lm}^p d_{lm}^p - c_{lm}^p d_{lm}^q] + \frac{1}{c_{00}^q} \sum_{l=0,\text{even}}^{\infty} \sum_{m=-l}^l [c_{lm}^q d_{lm}^q - c_{lm}^q d_{lm}^p] \right\}. \quad (\text{B5})$$

For numerical implementation, we use a truncated SH series with  $l \leq l_m$ , where  $l_m$  is a positive even integer; the total length of the SH series is  $n_b = (l_m + 1)(l_m + 2)/2$ . We follow the least-squares method [25], [26] to solve (B2), yielding the SH coefficients

$$\mathbf{c} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{d}. \quad (\text{B6})$$

Here,  $\mathbf{c} = [c_0 \ c_1 \ \dots \ c_{n_b-1}]^T$ , and  $\mathbf{d} = [D(\theta_0, \phi_0) D(\theta_1, \phi_1) \dots D(\theta_{n_s-1}, \phi_{n_s-1})]^T$  which represents the diffusivity function measured in  $n_s$  gradient directions, and  $\mathbf{B}$  is the matrix of basis functions, with elements  $B_{ij} = Y_j(\theta_i, \phi_i)$ . Here, we map  $(c_{lm}, Y_{lm})$  in (B2) to  $(c_j, Y_j)$ , using the relationship  $j = [(l^2 + l)/2] + m$ . Usually  $n_b \ll n_s$ , so it is more cost-effective to compute sKL using the SH method (B5) rather than using (13) directly.

#### APPENDIX C

##### Derivation of (B4)

Given the orthonormality property of  $Y_{lm}$

$$\int_{\Omega} Y_l^m(\theta, \phi) \bar{Y}_{l'}^{m'}(\theta, \phi) d\Omega = \delta_{ll'} \delta_{mm'} \quad (\text{C1})$$

where the overbar denotes complex conjugation and  $\delta_{ll'}$  is the Kronecker delta, we have

$$\int_{\Omega} Y_{lm} \bar{Y}_{00} d\Omega = \begin{cases} \frac{\sqrt{2}}{2} \int_{\Omega} ((-1)^m Y_l^m + Y_l^{-m}) \times \bar{Y}_0^0 d\Omega = 0, & \text{if } -l \leq m < 0 \\ \int_{\Omega} Y_l^0 \bar{Y}_0^0 d\Omega = \delta_{l0}, & \text{if } m = 0 \\ \frac{\sqrt{2}i}{2} \int_{\Omega} ((-1)^{m+1} Y_l^m + Y_l^{-m}) \times \bar{Y}_0^0 d\Omega = 0, & \text{if } 0 < m \leq l. \end{cases} \quad (\text{C2})$$

Here, we use the definition of  $Y_{lm}$  in (B2). Therefore

$$\begin{aligned} \int_{\Omega} D(\theta, \phi) d\Omega &= \int_{\Omega} \sum_{l=0,\text{even}}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\theta, \phi) d\Omega \\ &= \sum_{l=0,\text{even}}^{\infty} \sum_{m=-l}^l c_{lm} \end{aligned}$$

$$\begin{aligned} &\times \int_{\Omega} Y_{lm}(\theta, \phi) (\bar{Y}_{00}(\theta, \phi) \cdot 2\sqrt{\pi}) d\Omega \\ &= 2\sqrt{\pi} \sum_{l=0,\text{even}}^{\infty} c_{l0} \delta_{l0} \\ &= 2\sqrt{\pi} c_{00}. \end{aligned} \quad (\text{C3})$$

Here, we note that  $Y_{00} = Y_0^0 = 1/(2\sqrt{\pi})$  [see (B1)].

#### APPENDIX D

##### Iterative Computations for Forward and Backward Displacement Fields

With the mapping initialized as the identical mapping at  $t = 0$ , the forward and the backward displacement fields at the time step  $n$ ,  $\mathbf{h}_n(\mathbf{x})$  and  $\mathbf{h}_n^{-1}(\mathbf{x})$ , are given by

$$\begin{aligned} \mathbf{h}_n(\mathbf{x}) &= \mathbf{h}_{n-1}(\mathbf{x}) + \varepsilon \boldsymbol{\eta}_1(\mathbf{x}) + \varepsilon \boldsymbol{\eta}_2(\mathbf{x}) \\ \mathbf{h}_n^{-1}(\mathbf{x}) &= \mathbf{h}_{n-1}^{-1}(\mathbf{x}) + \varepsilon \boldsymbol{\xi}_1(\mathbf{x}) + \varepsilon \boldsymbol{\xi}_2(\mathbf{x}). \end{aligned} \quad (\text{D1})$$

Here,  $\varepsilon$  is an infinitesimally small positive time step.  $\boldsymbol{\eta}_1(\mathbf{x})$  and  $\boldsymbol{\xi}_1(\mathbf{x})$  are the gradient directions that arise from the minimization of  $E_1$  in (16) in the forward and backward mappings, respectively, and similarly  $\boldsymbol{\eta}_2(\mathbf{x})$  and  $\boldsymbol{\xi}_2(\mathbf{x})$  are defined in terms of  $E_2$ .  $\mathbf{h}_0(\mathbf{x}) = \mathbf{h}_0^{-1}(\mathbf{x}) = \mathbf{x}$  are the initial mappings (and these are initially set to the identity, or zero displacement).  $E_1$  and  $E_2$  were minimized by solving for the velocity fields  $\mathbf{v}_1(\mathbf{x})$  and  $\mathbf{v}_2(\mathbf{x})$  in the following Navier-Stokes equations [40], with the driving force  $\mathbf{F}$  given in (8)

$$\begin{aligned} \nabla E_{1,2} &= L \mathbf{v}_{1,2} + \mathbf{F}(\mathbf{x}, \mathbf{u}_{1,2}) \\ &= \mu \nabla^2 \mathbf{v}_{1,2} + (\lambda + \mu) \nabla (\nabla \cdot \mathbf{v}_{1,2}) + \mathbf{F}(\mathbf{x}, \mathbf{u}_{1,2}) \\ &= 0 \end{aligned} \quad (\text{D2})$$

corresponding to  $\mathbf{v}_1(\mathbf{x})$  and  $\mathbf{v}_2(\mathbf{x})$ , respectively. This velocity field formulation, based on the kinematics of a compressible fluid with negligible inertia, was first developed by Christensen *et al.* [40]. By constructing the deformation as the integral of a time-varying velocity field, this approach is advantageous as it always guarantees diffeomorphic mappings (i.e., differentiable maps with differentiable inverses), even when large deformations are required; simpler formulations, such as those based on linear elasticity, thin-plate splines, or Laplacian penalties can lead to folding or tearing of the image when large deformations are necessary to match the images [60]. Equation (D2) was solved iteratively in time by convolving the force field with a filter kernel of size  $D \times D \times D$ . This filter is essentially a discretized approximation of the Green's function of the linear operator  $L$ , derived using its eigenfunctions [61]. In the Eulerian reference frame, as used in [40],  $\boldsymbol{\eta}_1(\mathbf{x})$  can be obtained from  $\mathbf{v}_1(\mathbf{x})$  by

$$\begin{aligned} \boldsymbol{\eta}_1(\mathbf{x}) &= \frac{\partial \mathbf{h}(\mathbf{x})}{\partial t} = -\frac{\partial \mathbf{u}(\mathbf{x})}{\partial t} \\ &= -\left( \mathbf{v}_1(\mathbf{x}) - \sum_{i=1}^3 v_{1,i}(\mathbf{x}) \frac{\partial \mathbf{u}(x)}{\partial x_i} \right) \\ &= -(\nabla \mathbf{h}(\mathbf{x})) \mathbf{v}_1(\mathbf{x}) \end{aligned} \quad (\text{D3})$$

where  $\mathbf{h} = \mathbf{x} - \mathbf{u}$  is the displacement vector at time step  $n$ . Similarly,  $\xi_2(\mathbf{x}) = -(\nabla \mathbf{h}^{-1}(\mathbf{x})) \mathbf{v}_2(\mathbf{x})$ . By concatenating  $\mathbf{h}(\mathbf{x})$  and  $\mathbf{h}^{-1}(\mathbf{x})$  using  $(\mathbf{h}^{-1} \circ \mathbf{h})(\mathbf{x}) = \mathbf{x}$  and following the derivations in [39], we obtain

$$\begin{aligned}\eta_2(\mathbf{x}) &= -(\nabla \mathbf{h}(\mathbf{x})) \xi_2(\mathbf{h}(\mathbf{x})) \\ &= -(\nabla \mathbf{h}(\mathbf{x})) [ -(\nabla \mathbf{h}^{-1}(\mathbf{h}(\mathbf{x}))) \mathbf{v}_2(\mathbf{h}(\mathbf{x})) ] \\ &= \mathbf{v}_2(\mathbf{h}(\mathbf{x})) \\ \xi_1(\mathbf{x}) &= -(\nabla \mathbf{h}^{-1}(\mathbf{x})) \eta_1(\mathbf{h}^{-1}(\mathbf{x})) \\ &= -(\nabla \mathbf{h}^{-1}(\mathbf{x})) [ -(\nabla \mathbf{h}(\mathbf{h}^{-1}(\mathbf{x}))) \mathbf{v}_1(\mathbf{h}^{-1}(\mathbf{x})) ] \\ &= \mathbf{v}_1(\mathbf{h}^{-1}(\mathbf{x})).\end{aligned}\quad (\text{D4})$$

Here, we used the properties  $\mathbf{I} = \nabla(\mathbf{h}^{-1} \circ \mathbf{h})(\mathbf{x}) = [\nabla \mathbf{h}^{-1}(\mathbf{h}(\mathbf{x}))] \nabla \mathbf{h}(\mathbf{x})$  and  $\mathbf{I} = \nabla(\mathbf{h} \circ \mathbf{h}^{-1})(\mathbf{x}) = [\nabla \mathbf{h}(\mathbf{h}^{-1}(\mathbf{x}))] \nabla \mathbf{h}^{-1}(\mathbf{x})$ .

#### APPENDIX E

##### *The pFDR Method for Testing Overall Significance of Voxel Statistics*

The pFDR method [56] estimates the probability that the null hypothesis is true, from the empirical distribution of observed  $p$ -values. Briefly, pFDR is the false discovery rate conditioning on the event that positive findings, rejecting the null hypothesis, have occurred, and is given by

$$\text{pFDR}(\gamma) = \frac{\pi_0 \Pr(P \leq \gamma | H = 0)}{\Pr(P \leq \gamma)} \quad (\text{E1})$$

where  $\pi_0 = \Pr(H = 0)$  is the probability that the null hypothesis is true, and  $\gamma$  is the rejection threshold for the individual hypothesis, which was set to 0.01 in our experiments. We refer readers to [56] for the details of the estimation procedures to obtain pFDR for statistical maps. By convention, a statistical map with  $\text{pFDR} < 0.05$ , i.e., the false discovery rate is less than 5%, was considered to be significant.

#### APPENDIX F

##### *Derivation of the Driving Force Based on $IP_L$ and $IP_D$*

Given two diffusivity functions,  $D_S$  and  $D_T$ , under the assumption of Gaussian diffusion,  $IP_L$  and  $IP_D$  of  $D_S$  and  $D_T$  can be extended in terms of their corresponding diffusion tensors  $\mathbf{S}$  and  $\mathbf{T}$ , given by [21]

$$IP = \frac{8\pi}{15} [tr(\mathbf{S}\mathbf{T}) + \alpha \cdot tr(\mathbf{S})tr(\mathbf{T})]. \quad (\text{F1})$$

$IP = IP_L$ , if  $\alpha = 1/2$ ;  $IP = IP_D$ , if  $\alpha = -1/3$ . Let  $\mathbf{h}(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x})$  be the displacement field mapping image  $\mathbf{S}$  to image  $\mathbf{T}$ , the cost functional based on  $IP_L$  or  $IP_D$  is

$$\begin{aligned}E(\mathbf{u}, \mathbf{S}, \mathbf{T}) &= - \int_{\Omega} \frac{8\pi}{15} \{ tr(\mathbf{S}(\mathbf{x} - \mathbf{u})\mathbf{T}(\mathbf{x})) \\ &\quad + \alpha \cdot tr(\mathbf{S}(\mathbf{x} - \mathbf{u})) tr(\mathbf{T}(\mathbf{x})) \} d\mathbf{x}. \quad (\text{F2})\end{aligned}$$

The negative sign before the cost functional integral is to transform a maximization problem into a minimization. Taking the

gradient of  $E(\mathbf{u}, \mathbf{S}, \mathbf{T})$ , as in (7) and (8) and (A1) and (A2), the driving force  $\mathbf{F}(\mathbf{x}, \mathbf{u})$  can be easily derived

$$\mathbf{F}(\mathbf{x}, \mathbf{u}) = \frac{8\pi}{15} [\nabla \mathbf{S}(\mathbf{x} - \mathbf{u})]^T (\mathbf{T}(\mathbf{x}) + \alpha \cdot tr(\mathbf{T}(\mathbf{x})) \cdot \mathbf{I})_c \quad (\text{F3})$$

where  $\mathbf{I}$  is the identity tensor.

#### REFERENCES

- [1] S. Mori and J. Zhang, "Principles of diffusion tensor imaging and its applications to basic neuroscience research," *Neuron*, vol. 51, pp. 527–539, 2006.
- [2] M. Catani, "Diffusion tensor magnetic resonance imaging tractography in cognitive disorders," *Curr. Opin. Neurol.*, vol. 19, pp. 599–606, 2006.
- [3] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Fast and simple calculus on tensors in the log-Euclidean framework," in *Int. Conf. Med. Image Comput. Comput. Assist. Intervention (MICCAI)*, 2005, vol. 8, pp. 115–122.
- [4] M. P. do Carmo, *Riemannian Geometry*. Boston, MA: Birkhauser, 1992.
- [5] X. Pennec, P. Fillard, and N. Ayache, "A Riemannian framework for tensor computing," *Int. J. Comput. Vis.*, vol. 66, pp. 41–66, 2006.
- [6] P. T. Fletcher and S. Joshi, "Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors," in *Comput. Vis. Math. Methods Med. Biomed. Image Anal.: ECCV 2004 Workshops CVAMIA MMBIA*, Prague, Czech Republic, 2004, pp. 87–98.
- [7] M. K. Chung, M. Lazar, A. L. Alexander, Y. Lu, and R. Davidson, Probabilistic connectivity measure in diffusion tensor imaging via anisotropic kernel smoothing. Dept. Statist., Univ. Wisconsin, Tech. Rep. 1081, 2003.
- [8] C. Lenglet, R. Deriche, and O. Faugeras, "Inferring white matter geometry from diffusion tensor MRI: Application to connectivity mapping," in *Eur. Conf. Comput. Vis. (ECCV)*, Prague, Czech Republic, 2004, pp. 127–140.
- [9] C. Chef'd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, "Constrained flows of matrix-valued functions: Application to diffusion tensor regularization," in *Int. Conf. Med. Image Comput. Comput. Assist. Intervention (MICCAI)*, 2002, pp. 251–265.
- [10] L. O'Donnell, S. Haker, and C.-F. Westin, "New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and Geodesics in a tensor-warped space," in *Med. Image Comput. Computer-Assisted Intervention (MICCAI)*, Tokyo, Japan, 2002, pp. 459–466.
- [11] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, "Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics," presented at the IEEE Int. Symp. Biomed. Imag.: From Nano to Macro (ISBI), Arlington, VA, 2006.
- [12] P. Khurd, R. Verma, and C. Davatzikos, "On characterizing and analyzing diffusion tensor images by learning their underlying manifold structure," in *2006 Conf. Comput. Vis. Pattern Recognition Workshop*, 2006, pp. 61–68.
- [13] S. I. Amari, "Information geometry on hierarchy of probability distributions," *Information Theory, IEEE Trans.*, vol. 47, pp. 1701–1711, 2001.
- [14] K. Tsuda, S. Akaho, and K. Asai, "The EM algorithm for kernel matrix completion with auxiliary data," *J. Mach. Learn. Res.*, vol. 4, pp. 67–81, 2003.
- [15] Z. Wang and B. C. Vemuri, "DTI segmentation using an information theoretic tensor dissimilarity measure," *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1267–1277, Oct. 2005.
- [16] C. Lenglet, M. Rousson, and R. Deriche, "Segmentation of 3-D probability density fields by surface evolution: Application to diffusion MRI," presented at the Int. Conf. Med. Image Comput. Comput. Assist. Intervention (MICCAI), Saint-Malo, France, 2004.
- [17] A. Guimond, C. R. G. Guttman, S. K. Warfield, and C. F. Westin, "Deformable registration of DT-MRI data based on transformation invariant tensor characteristics," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2002, pp. 761–764.
- [18] J.-P. Thirion, "Non-rigid matching using demons," in *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition (CVPR'96)*, 1996, pp. 245–251.
- [19] J. Ruiz-Alzola, C. F. Westin, S. K. Warfield, C. Alberola, S. Maier, and R. Kikinis, "Nonrigid registration of 3-D tensor medical data," *Med. Image Anal.*, vol. 6, pp. 143–161, 2002.

- [20] D. C. Alexander and J. C. Gee, "Elastic matching of diffusion tensor images," *Comput. Vis. Image Understand.*, vol. 77, pp. 233–250, 2000.
- [21] H. Zhang, P. A. Yushkevich, D. C. Alexander, and J. C. Gee, "Deformable registration of diffusion tensor MR images with explicit orientation optimization," *Med. Image Anal.*, vol. 10, pp. 764–785, 2006.
- [22] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee, "Spatial transformations of diffusion tensor magnetic resonance," *IEEE Trans. Med. Imag.*, vol. 20, no. 11, pp. 1131–1139, Nov. 2001.
- [23] Y. Cao, M. I. Miller, R. L. Winslow, and L. Younes, "Large deformation diffeomorphic metric mapping of vector fields," *IEEE Trans. Med. Imag.*, vol. 24, no. 9, pp. 1216–1230, Sep. 2005.
- [24] Y. Cao, M. I. Miller, S. Mori, R. L. Winslow, and L. Younes, "Diffeomorphic matching of diffusion tensor images," in *2006 Conf. Comput. Vis. Pattern Recognition Workshop*, 2006, pp. 67–74.
- [25] D. C. Alexander, G. J. Barker, and S. R. Arridge, "Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data," *Magn. Reson. Med.*, vol. 48, pp. 331–340, 2002.
- [26] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, Apparent diffusion coefficients from high angular resolution diffusion images: Estimation and applications INRIA, Sophia Antipolis, France, Res. Rep. 5681, 2005.
- [27] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, pp. 259–267, 1994.
- [28] N. J. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra Appl.*, vol. 103, pp. 103–118, 1988.
- [29] G. J. Parker, J. A. Schnabel, M. R. Symms, D. J. Werring, and G. J. Barker, "Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging," *J. Magn. Reson. Imag.*, vol. 11, pp. 702–710, 2000.
- [30] G. Hermosillo, "Variational methods for multimodal image matching," Ph.D. dissertation, Univ. Nice (INRIA-ROBOTVIS), Sophia Antipolis, France, 2002.
- [31] D. S. Tuch, T. G. Reese, M. R. Wiegell, and V. J. Wedeen, "Diffusion MRI of complex neural architecture," *Neuron*, vol. 40, pp. 885–895, 2003.
- [32] D. S. Tuch, "Q-ball imaging," *Magn. Reson. Med.*, vol. 52, pp. 1358–1372, 2004.
- [33] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magn. Reson. Med.*, vol. 48, pp. 577–582, 2002.
- [34] E. Ozarslan, B. C. Vemuri, and T. H. Mareci, "Generalized scalar measures for diffusion MRI using trace, variance, and entropy," *Magn. Reson. Med.*, vol. 53, pp. 866–876, 2005.
- [35] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *J. Chem. Phys.*, vol. 42, pp. 288–292, 1965.
- [36] D. Xu, S. Mori, D. Shen, P. C. van Zijl, and C. Davatzikos, "Spatial normalization of diffusion tensor fields," *Magn. Reson. Med.*, vol. 50, pp. 175–182, 2003.
- [37] D. C. Alexander, "Maximum entropy spherical deconvolution for diffusion MRI," in *Inf. Process. Med. Imag.*, Glenwood Springs, CO, 2005, pp. 76–87.
- [38] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. Boston, MA: Academic, 1990.
- [39] A. Leow, S. Huang, A. Geng, J. Becker, S. Davis, A. Toga, and P. Thompson, "Inverse consistent mapping in 3-D deformable image registration: Its construction and statistical properties," presented at the Inf. Process. Med. Imag. (IPMI), Glenwood Springs, CO, 2005.
- [40] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Image Process.*, vol. 5, no. 10, pp. 1435–1447, Oct. 1996.
- [41] M. I. Miller, "Computational anatomy: Shape, growth, and atrophy comparison via diffeomorphisms," *NeuroImage*, vol. 23, pp. S19–S33, 2004.
- [42] B. Avants and J. C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interpolation," *NeuroImage*, vol. 23, pp. S139–S150, 2004.
- [43] M. F. Beg, M. I. Miller, A. Trounev, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vis.*, vol. 61, pp. 139–157, 2005.
- [44] M. Vaillant, M. I. Miller, L. Younes, and A. Trounev, "Statistics on diffeomorphisms via tangent space representations," *NeuroImage*, vol. 23, pp. S161–S169, 2004.
- [45] S. E. Rose, K. L. McMahon, A. L. Janke, B. O'Dowd, G. de Zubicaray, M. W. Strudwick, and J. B. Chalk, "Diffusion indices on magnetic resonance imaging and neuropsychological performance in amnesic mild cognitive impairment," *J. Neurol. Neurosurg. Psychiatry*, vol. 77, pp. 1122–1128, 2006.
- [46] D. K. Jones, M. A. Horsfield, and A. Simmons, "Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging," *Magn. Reson. Med.*, vol. 42, pp. 515–525, 1999.
- [47] C. Studholme, V. Cardenas, N. Schuff, H. Rosen, B. Miller, and M. Weiner, "Detecting spatially consistent structural differences in Alzheimer's and fronto temporal dementia using deformation morphometry," in *Int. Conf. Med. Image Comput. Comput. Assist. Intervention (MICCAI)*, 2001, pp. 41–48.
- [48] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, pp. S151–S160, 2004.
- [49] P. Kochunov, J. L. Lancaster, P. Thompson, R. Woods, J. Mazziotta, J. Hardies, and P. Fox, "Regional spatial normalization: Toward an optimal target," *J. Comput. Assist. Tomogr.*, vol. 25, pp. 805–816, 2001.
- [50] P. Kochunov, J. Lancaster, P. Thompson, A. W. Toga, P. Brewer, J. Hardies, and P. Fox, "An optimized individual target brain in the Talairach coordinate system," *NeuroImage*, vol. 17, pp. 922–927, 2002.
- [51] M.-C. Chiang, R. A. Dutton, K. M. Hayashi, O. L. Lopez, H. J. Aizenstein, A. W. Toga, J. T. Becker, and P. M. Thompson, "3-D pattern of brain atrophy in HIV/AIDS visualized using tensor-based morphometry," *NeuroImage*, vol. 34, pp. 44–60, 2007.
- [52] G. E. Christensen and H. J. Johnson, "Invertibility and transitivity analysis for nonrigid image registration," *J. Electron. Imag.*, vol. 12, pp. 106–117, 2003.
- [53] P. Thevenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Process.*, vol. 7, pp. 27–41, 1998.
- [54] P. J. Basser and S. Pajevic, "Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise," *Magn. Reson. Med.*, vol. 44, pp. 41–50, 2000.
- [55] D. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, 3rd ed. Boca Raton, FL: Chapman Hall/CRC, 2004.
- [56] J. D. Storey, "A direct approach to false discovery rates," *J. R. Stat. Soc.: Series B (Stat. Methodol.)*, vol. 64, pp. 479–498, 2002.
- [57] S. M. Smith, M. Jenkinson, H. Johansen-Berg, D. Rueckert, T. E. Nichols, C. E. Mackay, K. E. Watkins, O. Ciccarelli, M. Z. Cader, P. M. Matthews, and T. E. Behrens, "Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data," *NeuroImage*, vol. 31, pp. 1487–1505, 2006.
- [58] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, "A fast and robust ODF estimation algorithm in Q-ball imaging," presented at the IEEE Int. Symp. Biomed. Imag.: From Nano to Macro (ISBI), Arlington, VA, 2006.
- [59] J. Sijbers, A. J. den Dekker, J. Van Audekerke, M. Verhoye, and D. Van Dyck, "Estimation of the noise in magnitude MR images," *Magn. Reson. Imag.*, vol. 16, pp. 87–90, 1998.
- [60] G. E. Christensen, R. D. Rabbitt, M. I. Miller, S. C. Joshi, U. Grenander, T. A. Coogan, and D. C. V. Essen, "Topological properties of smooth anatomic maps," presented at the 14th Conf. Inf. Process. Med. Imag., Ile de Berder, France, 1995.
- [61] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Proc. 4th Int. Conf. Vis. Biomed. Comput.*, 1996, pp. 267–276.
- [62] D. W. Shattuck and R. M. Leahy, "BrainSuite: An automated cortical surface identification tool," *Med. Image Anal.*, vol. 6, pp. 129–142, 2002.
- [63] N. Leporé, C. A. Brun, X. Pennec, Y. Y. Chou, O. L. Lopez, H. J. Aizenstein, J. T. Becker, A. W. Toga, and P. M. Thompson, "Mean template for tensor-based morphometry using deformation tensors," in *Int. Conf. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, Brisbane, Australia, 2007.
- [64] P. A. Cook, Y. Bai, S. Nedjati-Gilani, K. K. Seunarine, M. G. Hall, G. J. Parker, and D. C. Alexander, "Camino: Open-source diffusion-MRI reconstruction and processing," in *14th Sci. Meeting Int. Soc. Magn. Reson. Med.*, Seattle, WA, 2006, p. 2759.