Published in final edited form as:
Neuroimage. 2011 June 1; 56(3): 1181–1201. doi:10.1016/j.neuroimage.2011.01.053.

A Nonparametric Riemannian Framework for Processing High Angular Resolution Diffusion Images and its Applications to ODF-based Morphometry

Alvina Goh[a,d,*], **Christophe Lenglet**[b], **Paul M. Thompson**[c], and **René Vidal**[d]

[a] Department of Mathematics, National University of Singapore, Singapore

[b] Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota Medical School, Minneapolis MN, USA

[c] Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles CA, USA

[d] Center for Imaging Science, Department of Biomedical Engineering, Johns Hopkins University, Baltimore MD, USA

Abstract

High angular resolution diffusion imaging (HARDI) has become an important technique for imaging complex oriented structures in the brain and other anatomical tissues. This has motivated the recent development of several methods for computing the orientation probability density function (PDF) at each voxel. However, much less work has been done on developing techniques for filtering, interpolation, averaging and principal geodesic analysis of orientation PDF fields. In this paper, we present a Riemannian framework for performing such operations. The proposed framework does not require that the orientation PDFs be represented by any fixed parameterization, such as a mixture of von Mises-Fisher distributions or a spherical harmonic expansion. Instead, we use a nonparametric representation of the orientation PDF. We exploit the fact that under the square-root re-parameterization, the space of orientation PDFs forms a Riemannian manifold: the positive orthant of the unit Hilbert sphere. We show that various orientation PDF processing operations, such as filtering, interpolation, averaging and principal geodesic analysis, may be posed as optimization problems on the Hilbert sphere, and can be solved using Riemannian gradient descent. We illustrate these concepts with numerous experiments on synthetic, phantom and real datasets. We show their application to studying left/right brain asymmetries.

Keywords

Diffusion weighted MRI; High Angular Resolution Diffusion Imaging

[*]Correspondence to: Alvina Goh, Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, Block S17, Singapore 119076. Tel: +65 65162738. Fax: +65 67795452. agoh@nus.edu.sg (Alvina Goh).

1. Introduction

Diffusion weighted magnetic resonance imaging (DW-MRI) is an imaging technique that produces *in vivo* images of biological tissues by measuring the constrained diffusion of water molecules. DW-MRI exploits the fact that water diffusion is hindered less along fibrous structures because the molecules are less constrained by tissue barriers when they move in that direction. The signal intensity at each position depends on the local microstructure in which the water molecules diffuse, and the strength and direction of the applied magnetic field gradients. Multiple MR scans with varying gradient directions and diffusion weightings are used to estimate the complete diffusion profile of a tissue. As the directions of maximum diffusion indicate the structural anisotropy of the medium, DW-MRI may be used to infer the organization and orientation of tissue components. Many experiments have shown that water diffusion is anisotropic in organized tissues such as the spinal cord, muscles, heart or brain white matter. This has generated great enthusiasm, because DW-MRI is presently the only available approach to non-invasively study the three-dimensional architecture of human fiber pathways, and quantify their physical and geometrical properties.

Several techniques may be used to reconstruct the local orientation of the tissue from DW-MRI. One classical technique, diffusion tensor imaging (DTI) (Basser et al., 1994), characterizes the diffusivity profile by a single oriented 3D Gaussian probability distribution function (PDF). Water diffusion is represented mathematically with a symmetric positive definite (SPD) tensor field $\mathbf{D}$: $\mathbb{R}^3 \rightarrow \mathrm{SPD}(3) \subset \mathbb{R}^{3\times 3}$ that measures the extent of diffusion in any direction $\boldsymbol{v} \in \mathbb{R}^3$ as $\boldsymbol{v}^\top \mathbf{D}\boldsymbol{v}$. DTI assumes a relatively simple diffusion model, but it has been successfully applied in regions of the brain and spinal cord with strong white-matter coherence, and has enabled the mapping of anatomical connections in the central nervous system (Basser et al., 2000; Hagmann et al., 2003). However, DTI models the diffusion with a single tensor, so it can only reveal a single dominant fiber orientation within each voxel. When fibers cross or bifurcate at a voxel, which is common in complex brain structures, DTI fails to provide accurate insights into the diffusivity profile or fiber geometry at that voxel.

Recent advances in diffusion MRI address this well-known limitation of DTI. In particular, high angular resolution diffusion imaging (HARDI) (Tuch, 2002) measures diffusion along n uniformly distributed directions on the sphere, and can characterize more complex fiber geometries. Several reconstruction techniques can be used to characterize diffusion based on the HARDI signals. One class is based on higher-order tensors (Barmpoutis et al., 2009; Ghosh et al., 2008), which model diffusivity using higher order polynomials and leverage prior work on DTI which used polynomials of degree two. Another approach reconstructs the orientation distribution function (ODF) from the HARDI measures. The ODF is the angular profile of the diffusion PDF of water molecules. Much current HARDI research focuses on computing the ODF from HARDI signals. One of the earliest methods, Q-ball Imaging (QBI), uses the Funk-Radon transform to estimate ODFs (Tuch, 2004). ODFs have also been approximated using different sets of basis functions such as spherical harmonics (SH) (Özarslan and Mareci, 2003; Descoteaux et al., 2007; Frank, 2002; Hess et al., 2006; Aganj et al., 2010). Such methods are relatively fast to implement because the ODF is computed analytically. Most of these methods model the ODF as a linear combination of basis functions. Consequently, the model parameters may be computed using linear least-squares optimization, rather than using expectation maximization. Besides spherical harmonics, ODFs may be approximated using the Poussin kernel (Rathi et al., 2009), where the optimal linear combination from a large function dictionary is found using matching pursuit. These methods normally do not assume a fixed upper bound on the maximum number of fiber crossings allowed at a voxel.

While many current efforts focus on computing ODFs accurately, much less work has been done on developing methods for processing these high-dimensional datasets, especially across subjects. That is, given a set of ODFs, how does one perform statistical analysis (e.g., principal components analysis), or basic operations such as interpolation, filtering and convolution? For DTI, the geometry of SPD(3) is well-studied and several metrics for comparing tensors have been proposed, including affine-invariant and log-Euclidean metrics (Arsigny et al., 2006; Kindlmann et al., 2007; Lenglet et al., 2006b; Pennec et al., 2006). Based on these metrics, several frameworks for processing tensor fields have been developed and applied to various topics in DTI. These studies have involved tensor estimation (Niethammer et al., 2006; Tschumperle and Deriche, 2003; Wang et al., 2004), processing (Arsigny et al., 2006; Batchelor et al., 2005; Fletcher and Joshi, 2007; Lenglet et al., 2006b; Pennec et al., 2006), segmentation and fiber tractography (Basser et al., 2000; Goh and Vidal, 2008a; Awate et al., 2007; Brun et al., 2004, 2003; Hagmann et al., 2003; Jonasson et al., 2004, 2007, 2005; Lenglet et al., 2006a; Savadjiev et al., 2006; Vemuri et al., 2001; Wang and Vemuri, 2005; Zalesky, 2008; Zhukov et al., 2003; Ziyan et al., 2006) and registration (Cao et al., 2005; Goh and Vidal, 2006; Guimond et al., 2002; Ruiz-Alzola et al., 2002; Zhang et al., 2006), as well as statistical tests such as voxel-based analysis of diffusion tensors (Buchel et al., 2004; Wang et al., 2006; Hua et al., 2008; Jahanshad et al., 2010). This work provided a better understanding of brain connectivity and differences in brain microstructure between populations of subjects. However, despite extensive research on DTI, the question of which metric is more appropriate for analyzing DTI data remains open. (Arsigny et al., 2006; Fletcher and Joshi, 2007; Lenglet et al., 2006b; Pennec et al., 2006) show that the proposed affine-invariant and log-Euclidean metrics avoid the “swelling effect” whereas the Euclidean metric allows swelling to occur. However, (Pasternak et al., 2010) found that the Euclidean metric is consistent with the expected statistical properties of tensor distributions, whereas the affine-invariant metric is not.

For HARDI, several metrics have been proposed for comparing different ODFs have been proposed, but metric selection remains challenging. Existing metrics include the Kullback-Leibler divergence (KLD) (Descoteaux and Deriche, 2009), the Jensen-Shannon divergence (JSD) (Chiang et al., 2008) and the Riemannian distance between a mixture of von Mises-Fisher distributions (McGraw et al., 2006). However, the KLD is difficult to compute for spherical data with high angular resolution and it also has to be symmetrized. Both the symmetrized version of the KLD (SKLD) and of the JSD do not obey the triangle inequality, so they are not proper metrics.

To our knowledge, there are very few existing frameworks for processing HARDI data. McGraw et al. (2006) represent the precomputed ODF with a mixture of von Mises-Fisher distributions whose parameters are computed by expectation-maximization. The authors compute distances and various Riemannian operations between different mixtures of von Mises-Fisher distributions using closed form formulae. However, this framework suffers from two major issues. First, the number of components in the mixture model is arbitrarily chosen. To find the optimal number of components, the difficult question of model selection must be addressed. Second, a nonlinear least-squares technique is used to compute the unknown parameters. This technique may be unstable when the complexity of the tissue increases, due to an increase in the number of parameters to be estimated. Our recent work (Goh et al., 2009b) addresses these issues by using a nonparametric Riemannian framework for HARDI processing. Instead of representing the ODF with a parametric model, we use the fact that, under the square-root re-parametrization, the space of ODFs becomes a convex subset of the unit Hilbert sphere. As such, various Riemannian operations are easily and efficiently computable. This allows one to develop several techniques for processing ODFs, such as filtering, interpolation and principal component analysis. After the publication of our initial conference paper (Goh et al., 2009b), a related framework was published

independently in (Cheng et al., 2009). In this work, the authors approximate the square root of the ODF as a linear combination of orthonormal basis functions. Since the coefficients of this expansion live in a finite-dimensional sphere, processing operations can be performed in the space of coefficients with reduced computational complexity.

Paper Contributions and Outline

This paper presents a Riemannian framework for processing ODFs. We operate directly on the precomputed ODF and do not require that the ODF be represented by any parametric model, such as a mixture of von Mises-Fisher distributions. This framework allows us to do population studies - more precisely, ODF-based morphometry - a generalization of voxel-based morphometry. Experimentally, we found that the Riemannian ODF-based morphometry is more sensitive in capturing the variations of the data compared to ODF morphometry based on the Euclidean metric. This paper proceeds as follows. In §2.1 we illustrate the Riemannian structure of the space of ODFs. We exploit the fact that ODFs are PDFs defined on the 2-sphere $\mathbb{S}^2$ and that, under the square-root re-parametrization, the space of ODFs forms a Riemannian manifold, namely the unit Hilbert sphere. Therefore, various Riemannian operations such as the exponential map, logarithm map and geodesics are not only available in closed form, but also easily and efficiently computable. §2.2 shows how to calculate the (weighted) mean of an ODF field. §2.3–2.4 show that spatial interpolation and convolution reduce to weighed averaging in the space of ODFs. §2.3 illustrates how to spatially interpolate ODFs on regular and non-regular grids. §2.4 shows how to perform convolution and (isotropic) filtering on an ODF field, followed by anisotropic filtering. §2.5 shows how to calculate the covariance and principal components of an ODF field and generalizes the Hotelling's T-squared statistic for performing multivariate hypothesis testing on two populations of ODFs. §3 presents experimental results for our processing framework. Experiments are shown for synthetic, phantom, and real HARDI brain data, with illustrative tests of brain asymmetry. §4 is the conclusion.

2. Material and methods

2.1. The Riemannian Manifold of Orientation Distribution Functions

As mentioned previously, HARDI measurements can be used to reconstruct the orientation distribution function (ODF) - the angular profile of the diffusion PDF of water molecules. The ODF is actually a PDF on the 2-sphere and therefore, the manifold structure of the ODFs should be considered. We refer the reader to Appendix A which shows how to convert HARDI signals to ODFs. In this section, we study the Riemannian geometry of the space of PDFs on the 2-sphere $\mathbb{S}^2$,

$$\mathcal{P}=\{\boldsymbol{p}:\mathbb{S}^2 \rightarrow \mathbb{R}^+ | \forall \boldsymbol{s} \in \mathbb{S}^2, \boldsymbol{p}(\boldsymbol{s}) \geq 0; \int_{\boldsymbol{s}\in\mathbb{S}^2} \boldsymbol{p}(\boldsymbol{s})d\boldsymbol{s}=1\},$$

and show that, under the square-root representation, various Riemannian operations can be computed in closed form. To study probability and information theory from a differential geometric point of view, Rao (1945) introduced the Riemannian structure formed by the statistical manifold whose elements are probability density functions. Rao (1945) also showed that the *Fisher-Rao* metric defined below, determines a Riemannian metric,

$$\langle \boldsymbol{q}_j, \boldsymbol{q}_k \rangle_{\boldsymbol{p}_i} = \int_{\boldsymbol{s}\in\mathbb{S}^2} \boldsymbol{q}_j(\boldsymbol{s})\boldsymbol{q}_k(\boldsymbol{s}) \frac{1}{\boldsymbol{p}_i(\boldsymbol{s})} d\boldsymbol{s}, \quad (1)$$

where $\boldsymbol{q}_j, \boldsymbol{q}_k \in T_{\mathbf{p}_i}\mathcal{P}$ are tangent functions to $\mathcal{P}$ at the point $\boldsymbol{p}_i$. As shown in (Cencov, 1982), this metric is the *unique intrinsic metric* on the statistical manifold $\mathcal{P}$, therefore it is invariant to re-parameterizations (essentially coordinate transforms) of the functions. However, this representation is difficult to work with as the geodesic between two elements is hard to compute (Srivastava et al., 2007).

Even though the space $\mathcal{P}$ is difficult to work with, it is not the only possible representation for PDFs. In this paper, our choice of the parameterization is the square-root representation, used in (Goh et al., 2009b; Cheng et al., 2009; Goh and Vidal, 2008b; Srivastava et al., 2007). The main reason for this choice is that the resulting manifold is a unit sphere in a Hilbert space, and the Fisher-Rao metric is the $\mathbb{L}^2$ metric (using the usual inner product between functions). Therefore, the various Riemannian operations such as geodesics, exponential maps, logarithm maps are available in closed form.

The square-root density function is defined as

$$\boldsymbol{\psi}(\boldsymbol{s}) = \sqrt{p(\boldsymbol{s})}, \tag{2}$$

where $\boldsymbol{\psi}(\boldsymbol{s})$ is assumed to be non-negative to ensure uniqueness. The space of such functions is defined as:

$$\boldsymbol{\Psi} = \{\boldsymbol{\psi} : \mathbb{S}^2 \rightarrow \mathbb{R}^+ | \forall \boldsymbol{s} \in \mathbb{S}^2, \boldsymbol{\psi}(\boldsymbol{s}) \geq 0; \int_{\boldsymbol{s}\in\mathbb{S}^2} \boldsymbol{\psi}^2(\boldsymbol{s}) d\boldsymbol{s} = 1\}. \tag{3}$$

From Eqn. (3), the functions $\boldsymbol{\psi}$ lie on the positive orthant of a unit Hilbert sphere, as shown in Fig. 1[1]. Making use of the representation in Eqn. (3) and the fact that Riemannian metrics are determined up to a constant scaling factor, we can rewrite Eqn. (1) and obtain the Fisher-Rao metric as the $\mathbb{L}^2$ metric

$$\langle \boldsymbol{\varphi}_j, \boldsymbol{\varphi}_k \rangle_{\boldsymbol{\psi}_i} = \int_{\boldsymbol{s}\in\mathbb{S}^2} \boldsymbol{\varphi}_j(\boldsymbol{s}) \boldsymbol{\varphi}_k(\boldsymbol{s}) d\boldsymbol{s},$$

where $\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_k \in T_{\boldsymbol{\psi}_i}\boldsymbol{\Psi}$ are tangent vectors. The *tangent space* of $\boldsymbol{\Psi}$ at a point $\psi \in \boldsymbol{\Psi}$, denoted as $T_{\boldsymbol{\psi}}\boldsymbol{\Psi}$, is defined as the span of the tangent vectors for all the possible smooth curves γ, $\gamma(t): \mathbb{R} \rightarrow \boldsymbol{\Psi}$, passing through $\boldsymbol{\psi}$. A curve between $\boldsymbol{\psi}_i$ and $\boldsymbol{\psi}_j$ with minimum length is called a geodesic. The advantage of choosing the square-root density becomes obvious, as many of the Riemannian operations for the unit Hilbert sphere are well-known and computable in closed-form as we show next.

We define the *exponential map* $\exp_{\boldsymbol{\psi}_i}: T_{\boldsymbol{\psi}_i}\boldsymbol{\Psi} \rightarrow \boldsymbol{\Psi}$, which maps each tangent vector $t\boldsymbol{\varphi} \in T_{\boldsymbol{\psi}_i}\boldsymbol{\Psi}$ to the point $\gamma(t) \in \boldsymbol{\Psi}$ obtained by following the geodesic $\gamma(t)$ (parametrized with arc-length) passing through $\boldsymbol{\psi}_i$ with direction $\boldsymbol{\varphi}$ for a distance t. For the sphere, the exponential map has the closed-form formula,

[1]The use of the word "sphere" might be confusing. ODFs are defined on the 2-sphere $\mathbb{S}^2$, whereas in Eqn. (3), each $\boldsymbol{\psi}$ is a point on the unit Hilbert sphere, which is an infinite dimensional manifold.

$$\exp_{\psi_i}(\varphi)=\cos(\|\varphi\|_{\psi_i})\psi_i+\sin(\|\varphi\|_{\psi_i})\frac{\varphi}{\|\varphi\|_{\psi_i}}, \tag{4}$$

where $\varphi \in T_{\psi_i}\mathbf{\Psi}$ is a tangent vector at ψ_i and

$$\|\varphi\|_{\psi_i}=\sqrt{\langle\varphi,\varphi\rangle_{\psi_i}}.$$

By restricting $\|\varphi\|_{\psi_i}$ such that the geodesic $\gamma(t)$ lies in the positive orthant, we ensure that the exponential map is bijective. Note that the range of $\|\varphi\|_{\psi_i}$ depends on ψ_i.

Next, we define the set $C_{\psi_i} \subset \mathbf{\Psi}$ as the set of all the points $\psi_j = \exp_{\psi_i}(t_0\varphi)$ such that $\gamma = \exp_{\psi_i}(t\varphi)$ is a length-minimizing geodesic for $t \in [0, t_0]$. The boundary of C_{ψ_i} is called the *cut locus*, and, intuitively, it is the set of points for which the distance from ψ_i stops to be a differentiable function of t. The exponential map is therefore invertible for all the points in the interior of C_{ψ_i} and we can define the *logarithm map* $\log_\psi: C_\psi \rightarrow T_\psi\mathbf{\Psi}$ as $\log_\psi=\exp_\psi^{-1}$. The logarithm map from ψ_i to ψ_j has the closed-form formula,

$$\overrightarrow{\psi_i\psi_j}=\log_{\psi_i}(\psi_j)=\frac{\psi_j-\langle\psi_i,\psi_j\rangle\psi_i}{\sqrt{1-\langle\psi_i,\psi_j\rangle^2}}\cos^{-1}\langle\psi_i,\psi_j\rangle. \tag{5}$$

Now, we show how to compute the various Riemannian operations in $\mathbf{\Psi}$, to evolve (or "march") along its geodesic curves. For any two functions $\psi_i, \psi_j \in \mathbf{\Psi}$, the geodesic distance between these two points on a unit Hilbert sphere is simply the angle between them,

$$\mathrm{dist}(\psi_i,\psi_j)=\left\|\log_{\psi_i}(\psi_j)\right\|_{\psi_i}=\cos^{-1}\langle\psi_i,\psi_j\rangle=\cos^{-1}\left(\int_{s\in\mathbb{S}^2}\psi_i(s)\psi_j(s)ds\right), \tag{6}$$

where $\langle\cdot,\cdot\rangle$ is the normal dot product between points in the sphere under the $\mathbb{L}^2$ metric. For any two points $\psi_i, \psi_j \in \mathbf{\Psi}$, the geodesic curve γ from ψ_i to ψ_j is contained in $\mathbf{\Psi}$. Therefore, from the definition of convexity for Riemannian manifolds (Berger, 2003), $\mathbf{\Psi}$ forms a *convex subset* of the unit Hilbert sphere.

2.2. Computing the mean of a set of ODFs

As we show later, many ODF processing operations, such as filtering and convolution, reduce to averaging a set of ODFs. Moreover, the average of a set of ODFs is also important for statistical analysis of group differences. Next, we show how to compute a weighted mean of a set of ODFs.

As defined by Frechet (1948), for a Riemannian manifold, the intrinsic mean $\bar{\psi}$ of a set of n points $\{\psi_i\}_{i=1}^n$ is the solution to the following minimization problem

$$\bar{\boldsymbol{\psi}} = \arg\min_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} \frac{1}{n}\sum_{i=1}^{n} \mathrm{dist}(\boldsymbol{\psi}, \boldsymbol{\psi}_i)^2. \tag{7}$$

Similarly to Eqn. (7), the weighted Karcher mean is defined as

$$\bar{\boldsymbol{\psi}} = \arg\min_{\boldsymbol{\psi}\in\boldsymbol{\Psi}} \frac{1}{2}\sum_{i=1}^{n} w_i \mathrm{dist}(\boldsymbol{\psi}, \boldsymbol{\psi}_i)^2, \tag{8}$$

where the weights are chosen such that $w_i \geq 0$ and $\sum_{i=1}^{n} w_i = 1$. Cheng et al. (2009) proved that the weighted Riemannian mean *is unique and exists for* $\boldsymbol{\Psi}$. We refer the reader to Appendix B for more details. Notice that when $w_i = \frac{1}{n}$, $\bar{\boldsymbol{\psi}}$ becomes the intrinsic mean. The gradient of the cost function is given as $\mathrm{grad}\frac{1}{2}\sum_{i=1}^{n} w_i \mathrm{dist}(\boldsymbol{\psi}, \boldsymbol{\psi}_i)^2 = \sum_{i=1}^{n} w_i \log_{\bar{\boldsymbol{\psi}}}(\boldsymbol{\psi}_i)$. In addition, the (weighted) mean is characterized by $\mathrm{grad}\frac{1}{2}\sum_{i=1}^{n} w_i \mathrm{dist}(\boldsymbol{\psi}, \boldsymbol{\psi}_i)^2 = \mathbf{0}$, that is, $\bar{\boldsymbol{\psi}}$ is the unique solution to

$$\sum_{i=1}^{n} w_i \log_{\bar{\boldsymbol{\psi}}}(\boldsymbol{\psi}_i) = \mathbf{0}. \tag{9}$$

Since the exponential and logarithm maps are known (Eqn. (4) and (5)), it is possible to compute $\bar{\boldsymbol{\psi}}$ using geodesic marching, as detailed in Algorithm 1. Algorithm 1, which iteratively searches for a point $\bar{\boldsymbol{\psi}}$ which satisfies the condition in Eqn. (9), is equivalent to doing a simple gradient descent, and has a linear rate of convergence. Other algorithms to compute the mean with better convergence rates have also been proposed (Buss and Fillmore, 2001;Absil et al., 2008). In the next section, we show that interpolation and linear convolution of ODF fields reduce to weighted Karcher averaging.

Before proceeding further, notice that when computing of the Riemannian weighted mean $\bar{\boldsymbol{\psi}}$ of continuous PDFs on the sphere is cumbersome, as one needs to compute the dot products between $\boldsymbol{\psi}$'s and this involves an integral over the 2-sphere $\mathbb{S}^2$. One way to deal with this computational complexity, as proposed by Goh et al. (2009b), is to discretize the PDF and use the square root of histograms. In this case, the Hilbert sphere becomes a finite dimensional sphere and the dot products become summations. The complexity is now a function of the number of bins. Another method, proposed by Cheng et al. (2009), is to approximate the square-root ODF $\boldsymbol{\psi}$ as a linear combination of the first N orthonormal basis functions, i.e., $\boldsymbol{\psi}(\boldsymbol{s}) \approx \sum_{i=1}^{N} c_i B_i(\boldsymbol{s})$, where $\{B_i(\boldsymbol{s})\}_{i=1}^{\infty}$ is any orthonormal basis set and $\{c_i\}$ are the coefficients. Existing literature (Cheng et al., 2009; Descoteaux et al., 2007) propose to used SH basis with orders = 4, 6, 8 with corresponding N = 15, 28, 45 to represent ODFs. The (truncated) spherical harmonics basis is chosen in (Cheng et al., 2009), where the order of the basis is set at a predetermined value of N = 15, and the coefficients are calculated in the least-squares sense. Cheng et al. (2009) noted that since the spherical harmonics functions are orthogonal, in this case, the Hilbert sphere also reduces to a finite-dimensional sphere and the theoretical results of the Riemannian framework, such as the uniqueness of the mean, hold true. Therefore, instead of working with $\boldsymbol{\psi}$, one can work with $\{c_i\}$. The advantage of using such a representation is that it is more efficient and computationally

simpler, as the complexity is a function of the number of coefficients considered. Computationally, it may be more efficient to deal with the second alternative, as the number of coefficients is typically smaller than the number of bins. If one works with the discrete histograms and due to the high dimension, the iterative Riemannian operations will require more computational power. However, both approximations incur some errors relative to the continuous domain, and it is not clear which approximation is better. The coefficients would still need to be computed from histograms, so there is also an additional cost of computing the coefficients. It is not clear which method is more efficient or more accurate. In this paper, we choose to work with the discrete square-root of histograms directly.

2.3. Interpolation of ODF Fields

Interpolation is one of the most important operations in image processing. In general, almost any geometric transformation such as translation, rotation, scaling, or warping, requires interpolation of an image or volume. Here we show how to interpolate a given ODF field. As reported in (Chiang et al., 2008), spatial interpolation of ODFs is a new issue.

While trilinear interpolation has been previously used in (Chiang et al., 2008), in this paper, we show how to do spatial interpolation on regular and non-regular grids. We concentrate on linear, bilinear and trilinear interpolation of ODF fields. However, the same analysis can be extended to more sophisticated forms of interpolation such as polynomial interpolation, spline interpolation, radial basis function interpolation and kriging.

We first examine bilinear, trilinear and multi-dimensional interpolation. We assume that we have known data at the m points in an n-dimensional grid. In the standard Euclidean case, multi-dimensional interpolation of a signal $\boldsymbol{\psi}$ at a point $\boldsymbol{x} = [x_1, \ldots, x_n]$ is defined as

$$\boldsymbol{\psi}(\boldsymbol{x}) = \sum_{j=1}^{m} w_j(\boldsymbol{x}) \boldsymbol{\psi}(\boldsymbol{x}_j), \tag{10}$$

where $\boldsymbol{x}_j = [(x_j)_1, \ldots, (x_j)_n]$ are the coordinates of the endpoints, $\boldsymbol{\psi}(\boldsymbol{x}_j)$ are the corresponding signal values and $w_j(\boldsymbol{x})$ are the interpolation weights. One can show that

$$\boldsymbol{\psi}(\boldsymbol{x}) = \arg\min_{\boldsymbol{\psi} \in \boldsymbol{\Psi}} \sum_{j=1}^{2^n} w_j(\boldsymbol{x}) \left\| \boldsymbol{\psi}(\boldsymbol{x}_j) - \boldsymbol{\psi} \right\|^2. \tag{11}$$

Therefore, we can define multi-dimensional interpolation $\boldsymbol{\psi}(\boldsymbol{x})$ for a Riemannian manifold $\boldsymbol{\Psi}$ as

$$\boldsymbol{\psi}(\boldsymbol{x}) = \arg\min_{\boldsymbol{\psi} \in \boldsymbol{\Psi}} \sum_{j=1}^{2^n} w_j(\boldsymbol{x}) \operatorname{dist}(\boldsymbol{\psi}, \boldsymbol{\psi}(\boldsymbol{x}_j))^2.$$

Notice that $\boldsymbol{\psi}$ is the Riemannian weighted mean of $\{\boldsymbol{\psi}(\boldsymbol{x}_j)\}$, which can be computed using Algorithm 1.

In the case of a regular grid $\cup_{i=1}^{n}[x_{i,1}, x_{i,2}]$, $m = 2^n$ and the weights are simply given by $w_j(\boldsymbol{x})=\prod_{i=1}^{n}\left(1-\frac{|x_i-(x_j)_i|}{x_{i,2}-x_{i,1}}\right)$. In the case of a non-regular grid, we can either use simple nearest-neighbor interpolation where $\boldsymbol{\psi}(\boldsymbol{x})=\boldsymbol{\psi}(\boldsymbol{x}_j^*)$ and $\boldsymbol{x}_j^*$ is the nearest neighbor of $\boldsymbol{x}$. Alternatively, we can estimate the unknown values via the inverse distance weighting (IDW) method (Shepard, 1968), which is based on the assumption that observations further away should have their contributions diminished according to how far away they are. Specifically,

Shepard (1968) defines the interpolation weights as $w_j(\boldsymbol{x})=\frac{K(\boldsymbol{x},\boldsymbol{x}_j)^{-p}}{\sum_{i=1}^{m}K(\boldsymbol{x},\boldsymbol{x}_j)^{-p}}$, where K is a given metric operator (distance) from $\boldsymbol{x}_j$ to $\boldsymbol{x}$, and p is a positive real number, called the power parameter. An example of K is $K(\boldsymbol{x}, \boldsymbol{x}_j) = \text{dist}(\boldsymbol{x}, \boldsymbol{x}_j)$.

2.4. Convolution and Filtering of ODF Fields

Convolution is used for many image processing tasks such as denoising, smoothing, edge detection, image sharpening and filtering. In this section, we will show how to apply these operations to HARDI data. The motivation for filtering diffusion data is that the signal-to-noise ratio in HARDI data is generally low, especially with the high gradient strengths necessary for imaging and resolving crossing white matter tracts and other regions of complex tissue architecture.

2.4.1. Convolution and Isotropic Filtering—Traditionally, discrete convolution of an image $\boldsymbol{\psi}$ with the filter $\mathbf{g}$ is defined as

$$\boldsymbol{\xi}(\boldsymbol{x})=\boldsymbol{\psi}(\boldsymbol{x}) * \mathbf{g}(\boldsymbol{x})=\sum_{\mathbf{u}}\mathbf{g}(\mathbf{u})\boldsymbol{\psi}(\boldsymbol{x}+\mathbf{u}). \tag{12}$$

Assuming that $\Sigma_{\mathbf{u}}\,\mathbf{g}(\mathbf{u}) = 1$, we can rewrite $\boldsymbol{\xi}(\boldsymbol{x})$ as

$$\boldsymbol{\xi}(\boldsymbol{x})=\arg\min_{\boldsymbol{\xi}}\sum_{\mathbf{u}}\mathbf{g}(\mathbf{u})(\boldsymbol{\xi}-\boldsymbol{\psi}(\boldsymbol{x}+\mathbf{u}))^2. \tag{13}$$

For HARDI data, $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ and $\boldsymbol{\xi} \in \boldsymbol{\Psi}$, we can define the equivalent minimization problem for convolution as

$$\boldsymbol{\xi}(\boldsymbol{x})=\arg\min_{\boldsymbol{\xi}\in\boldsymbol{\Psi}}\sum_{\mathbf{u}}\mathbf{g}(\mathbf{u})\text{dist}(\boldsymbol{\xi},\boldsymbol{\psi}(\boldsymbol{x}+\mathbf{u}))^2. \tag{14}$$

Again, this is a Riemannian weighted average problem. When $\mathbf{g}(\mathbf{u}) \geq 0$, the optimization problem in Eqn. (14) remains convex and the convergence of Algorithm 1 to the global minimum is again guaranteed. Moreover, the solution $\boldsymbol{\xi}(\boldsymbol{x})$ is obtained using Algorithm 1. When $\mathbf{g}(\mathbf{u}) < 0$, convergence is guaranteed when the Hessian of the function in Eqn. (14) is positive definite.

2.4.2. Anisotropic Filtering—In the case of scalar images, linear filtering can blur and mislocate discontinuities in the data. To that end, nonlinear diffusion filtering techniques have been proposed; these smooth the data while preserving discontinuities, have been proposed (Perona and Malik, 1990; Weickert, 1997; Weickert and Brox, 2002). The idea behind nonlinear diffusion filters in image processing comes from the use of the Gaussian

filter in multi-scale image analysis. More precisely, one considers the following linear diffusion equation

$$\delta_t \boldsymbol{\xi}(\boldsymbol{x}, t) = \Delta \boldsymbol{\xi}(\boldsymbol{x}, t) = \nabla \cdot (\nabla \boldsymbol{\xi}(\boldsymbol{x}, t)), \tag{15}$$

$$\boldsymbol{\xi}(\boldsymbol{x}, 0) = \boldsymbol{\psi}(\boldsymbol{x}), \tag{16}$$

where $\Delta\boldsymbol{\xi}(\boldsymbol{x}, t)$ is the Laplacian of $\boldsymbol{\xi}(\boldsymbol{x}, t)$, $\boldsymbol{\psi}$ is the original noisy image and $\boldsymbol{\xi}$ is the filtered image. The solution to this diffusion equation is obtained by convolving $\boldsymbol{\psi}$ with a Gaussian filter $\mathbf{g}(\boldsymbol{x}, \sigma) = \frac{1}{2\pi\sigma^2} \exp(-\frac{\left\|\boldsymbol{x}\right\|^2}{2\sigma^2})$. Specifically, $\boldsymbol{\xi}(\boldsymbol{x}, t) = g(\boldsymbol{x}, \sqrt{2t}) * \boldsymbol{\psi}(\boldsymbol{x})$.

The main idea behind anisotropic filtering is to use a non-homogeneous diffusivity function $c(\boldsymbol{x}, \cdot)$ that depends on the data at $\boldsymbol{x}$. More generally, the following nonlinear diffusion equation is used,

$$\delta_t \boldsymbol{\xi}(\boldsymbol{x}, t) = \nabla \cdot (c(\boldsymbol{x}, t) \nabla \boldsymbol{\xi}(\boldsymbol{x}, t)), \tag{17}$$

$$\boldsymbol{\xi}(\boldsymbol{x}, 0) = \boldsymbol{\psi}(\boldsymbol{x}), \tag{18}$$

where $c(\boldsymbol{x}, t)$ is the diffusivity function. When $c(\boldsymbol{x}, t)$ is a decreasing function that depends on the derivative of the image $\boldsymbol{\xi}(\boldsymbol{x}, t)$, it prevents diffusion across the edges, while allowing diffusion along the edge.

When the data lives in d dimensions and if $c(\boldsymbol{x}, t) = \exp(-\frac{\left\|\nabla \boldsymbol{\xi}(\boldsymbol{x}, t)\right\|^2}{\sigma^2})$, we obtain the following equation

$$\delta_t \boldsymbol{\xi}(\boldsymbol{x}, t) \propto \sum_{j=1}^{d} \exp(-\frac{\left\|\delta_j \boldsymbol{\xi}(\boldsymbol{x}, t)\right\|^2}{\sigma^2}) \delta_j^2 \boldsymbol{\xi}(\boldsymbol{x}, t). \tag{19}$$

In this case, the evolution equation is:

$$\boldsymbol{\xi}(\boldsymbol{x}, t + \Delta t) = \boldsymbol{\xi}(\boldsymbol{x}, t) + \Delta t \delta_t \boldsymbol{\xi}(\boldsymbol{x}, t). \tag{20}$$

Now, unlike in the Euclidean case where we make use of the Laplace operator, for a field of ODFs, we will first approximate the Laplace-Beltrami operator on the manifold as the sum of all the second partial derivatives in the spatial coordinates

$$\sum_{i,j} \delta_i \delta_j \boldsymbol{\xi}(\boldsymbol{x}, t),$$

and we will get the following equation

$$\delta_t \boldsymbol{\xi}(\boldsymbol{x},t) \propto \sum_{i,j}^{d} \exp\left(-\frac{\left\|\delta_{ij}\boldsymbol{\xi}(\boldsymbol{x},t)\right\|^2}{\sigma^2}\right)\delta_{ij}\boldsymbol{\xi}(\boldsymbol{x},t).$$

The evolution equation is again

$$\boldsymbol{\xi}(\boldsymbol{x},t+\Delta t)=\boldsymbol{\xi}(\boldsymbol{x},t)+\Delta t\,\delta_t\boldsymbol{\xi}(\boldsymbol{x},t). \tag{21}$$

Finally, we will show how to approximate the operator $\delta_i\delta_j$ for a given field of ODFs $\boldsymbol{\xi}$ by using the finite central difference method. To this end, recall that we have the following expressions for the Euclidean derivatives,

$$\delta_i\boldsymbol{\xi}(\boldsymbol{x}) \approx \frac{1}{\varepsilon}\left(\boldsymbol{\xi}\left(\boldsymbol{x}+\frac{\varepsilon\mathbf{e}_i}{2}\right)-\boldsymbol{\xi}(\boldsymbol{x})+\boldsymbol{\xi}(\boldsymbol{x})-\boldsymbol{\xi}\left(\boldsymbol{x}-\frac{\varepsilon\mathbf{e}_i}{2}\right)\right), \tag{22}$$

$$\delta_i^2\boldsymbol{\xi}(\boldsymbol{x}) \approx \frac{1}{\varepsilon^2}\left(\boldsymbol{\xi}(\boldsymbol{x}+\varepsilon\mathbf{e}_i)-2\boldsymbol{\xi}(\boldsymbol{x})+\boldsymbol{\xi}(\boldsymbol{x}-\varepsilon\mathbf{e}_i)\right), \tag{23}$$

$$\delta_i\delta_j\boldsymbol{\xi}(\boldsymbol{x}) \approx \frac{1}{\varepsilon^2}\left(\boldsymbol{\xi}\left(\boldsymbol{x}+\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}+\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)-\boldsymbol{\xi}\left(\boldsymbol{x}+\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}-\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)-\boldsymbol{\xi}\left(\boldsymbol{x}-\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}+\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)+\boldsymbol{\xi}\left(\boldsymbol{x}-\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}-\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)\right), \tag{24}$$

where $\{\mathbf{e}_i\}$ are the orthonormal vectors of the canonical basis for $\mathbb{R}^3$ and $\varepsilon > 0$. The derivation of the various approximations of the derivatives of $\boldsymbol{\xi}$ proceeds as in DTI case(Krajsek et al., 2008; Pennec et al., 2006; Castano-Moraga et al., 2007). Therefore, for the first and second order derivatives in Riemannian space $\boldsymbol{\xi}$, we have the following approximations,

$$\delta_i\boldsymbol{\xi}(\boldsymbol{x}) \approx \frac{1}{\varepsilon}\left(\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}\left(\boldsymbol{x}+\frac{\varepsilon\mathbf{e}_i}{2}\right)}-\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}\left(\boldsymbol{x}-\frac{\varepsilon\mathbf{e}_i}{2}\right)}\right), \tag{25}$$

$$\delta_i^2\boldsymbol{\xi}(\boldsymbol{x}) \approx \frac{1}{\varepsilon^2}\left(\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}(\boldsymbol{x}+\varepsilon\mathbf{e}_i)}+\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}(\boldsymbol{x}-\varepsilon\mathbf{e}_i)}\right), \tag{26}$$

$$\delta_i\delta_j\boldsymbol{\xi}(\boldsymbol{x}) \approx \frac{1}{\varepsilon^2}\left(\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}\left(\boldsymbol{x}+\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}+\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)}-\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}\left(\boldsymbol{x}+\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}-\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)}+\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}\left(\boldsymbol{x}-\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}-\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)}-\overrightarrow{\boldsymbol{\xi}(\boldsymbol{x})\boldsymbol{\xi}\left(\boldsymbol{x}-\frac{\varepsilon\mathbf{e}_i}{\sqrt{2}}+\frac{\varepsilon\mathbf{e}_j}{\sqrt{2}}\right)}\right). \tag{27}$$

2.5. Statistics on the Space of ODFs

Here we show how to calculate the covariance and principal components of a set of points lying on $\boldsymbol{\Psi}$. Once the first and second order statistics for ODFs are defined, we generalize the Hotelling's T-squared statistic for two populations of ODFs, and extend the commonly used neuroimaging analysis technique known as voxel-based morphometry (Ashburner and Friston, 2000) to what we will call ODF-based morphometry.

2.5.1. Covariance and Principal Geodesic Analysis of ODFs—Principal component analysis (PCA) (Jolliffe, 1986) is the method for finding the optimal subspace that minimizes the sum of squared distances between $\boldsymbol{x}_i$ and its projection $\pi_{\boldsymbol{\Phi}}(\boldsymbol{x}_i)$ onto the subspace $\boldsymbol{\Phi}$. This is equivalent to maximizing the projected variance along the subspace. Specifically, given a multivariate random variable $\boldsymbol{x} \in \mathbb{R}^D$, the d principal components of $\boldsymbol{x}$ are defined as the uncorrelated linear components $y_i = \mathbf{u}_i^\top(\boldsymbol{x} - \mu) \in \mathbb{R}$, $\mathbf{u}_i \in \mathbb{R}^D$, $i = 1, \ldots, d$, where $\mu = E(\boldsymbol{x})$, such that the variance of y_i is maximized subject to $\mathbf{u}_i^\top \mathbf{u}_i = 1$ and $\mathbf{var}(y_1) \geq \ldots \geq \mathbf{var}(y_d)$. The solution is given as $\{\mathbf{u}_i\}_{i=1}^d$, which are the d leading eigenvectors of the population covariance matrix $E[\boldsymbol{x}\boldsymbol{x}^\top]$ and the first d principal components of $\boldsymbol{x}$ are $y_i = \mathbf{u}_i^\top(\boldsymbol{x} - \mu)$.

Now, by the definition of ODFs $\int \boldsymbol{p}(\boldsymbol{s})d\boldsymbol{s} = 1$, it lies in a hyperplane. While in principle, it may appear as if the original ODF representation is Euclidean, this is actually not the case. The reason is that under the Fisher-Rao metric, the hyperplane is turned into a non-Euclidean manifold. Therefore, while one can work with the hyperplane representation, means cannot be computed as usual, because the Fisher-Rao metric in this space does not coincide with the $\mathbb{L}^2$ metric. The whole point of taking the square-root and working in the Hilbert sphere is that the Fisher-Rao metric becomes the $\mathbb{L}^2$ metric, hence we know how to compute the Riemannian operations. There may be other valid choices for the metric but the Fisher-Rao metric makes it theoretically and computationally easy to deal with this kind of data.

For a Riemannian manifold, the equivalent task of finding the principal components is to fit an optimal low-dimensional geodesic submanifold $\boldsymbol{\Phi}$ to the data. This is known as Principal Geodesic Analysis (PGA), first proposed in (Fletcher et al., 2004) for the analysis of shape spaces, and applied to DTI in (Fletcher and Joshi, 2007; Pennec et al., 2006). PGA attempts to find the geodesic submanifold $\boldsymbol{\Phi}$ such that the squared Riemannian distance from the data to that submanifold is minimized. Denoting the geodesic submanifold at the mean $\bar{\boldsymbol{\psi}}$ by $\boldsymbol{\Phi}$, the projection of a point $\boldsymbol{\psi}$ is defined as

$$\pi_{\boldsymbol{\Phi}}(\boldsymbol{\psi}) = \arg\min_{\boldsymbol{\varphi} \in \boldsymbol{\Phi}} \mathrm{dist}(\boldsymbol{\varphi}, \bar{\boldsymbol{\psi}})^2. \tag{28}$$

Here, the mean $\bar{\boldsymbol{\psi}}$ of the data can be found by using the Karcher mean algorithm described in §2.2 (see Algorithm 1). For the space of ODFs, the projection that is solved via a minimization problem is guaranteed to exist and to be unique. This is generally not true for arbitrary manifolds. Unlike in Euclidean space, where one could easily find the projection of each data point onto the subspace, finding the equivalent projection onto the Riemannian submanifold is not an easy problem to solve. Fletcher and Joshi (2007) assume that the projection onto the geodesic submanifold may be approximated linearly onto the tangent space of $\boldsymbol{\Phi}$. That is,

$$\log_{\bar{\psi}}(\pi_{\Phi}(\psi)) \approx \arg\min_{v \in T_{\bar{\psi}}\Phi} \left\| \log_{\bar{\psi}}(\psi) - v \right\|^2 = \sum_{i=1}^{k} \langle v_i, \log_{\bar{\psi}}(\psi) \rangle v_i, \tag{29}$$

where $v_1, \ldots, v_k$ forms an orthonormal basis for $T_{\bar{\psi}}\Phi$. By making such an assumption, in (Fletcher and Joshi, 2007) the tangent vectors that span the principal geodesic submanifolds at $\bar{\psi}$ are found by performing an eigendecomposition of the covariance matrix

$$\mathrm{cov}(\psi) = \frac{1}{n-1} \sum_{i=1}^{n} \overrightarrow{\bar{\psi}\psi_i} \overrightarrow{\bar{\psi}\psi_i}^{\top} = \frac{1}{n-1} \sum_{i=1}^{n} \log_{\bar{\psi}}\psi_i \log_{\bar{\psi}}\psi_i^{\top}. \tag{30}$$

Pennec et al. (2006) first note that the covariance matrix for Euclidean vectors is simply given as $\frac{1}{n-1}\sum_{i=1}^{n} \overrightarrow{\bar{\psi}\psi_i} \overrightarrow{\bar{\psi}\psi_i}^{\top}$ and give the equivalent covariance matrix that measures the amount of dispersion in the data with respect to the mean as that defined in Eqn. 30. Once the covariance matrix is defined, we can find the *d* principal components as before. We refer the reader to Appendix C for how to calculate the infinite-dimensional covariance matrix when φ_i is a continuous function. Now in practice, the ODF is not a continuous PDF; rather it is a histogram over *M* different gradient directions. We refer the reader to Appendix D for practical issues relating to the discrete representation of the ODFs. Therefore, we can calculate the sample covariance matrix $\mathrm{cov}(\psi) = \frac{1}{n-1}\sum_{i=1}^{n} \varphi_i \varphi_i^{\top} \in \mathbb{R}^{M \times M}$ exactly. This algorithm is summarized in Algorithm 2.

In Euclidean space, it is easy to prove that the problem of minimizing the residual sum of squared distances between the data point and its projection reduces to the problem of maximizing the projected variance along the subspace, but there is no corresponding simple proof for Riemannian manifolds. Fletcher et al. (2004) shows that this tangent space approximation is only valid when the data lie in a small neighborhood. Recent works by Sommer et al. (2010a,b) presents a comparison between the two definitions of PGA in addition to the exact computation of geodesic submanifolds on a wide class of manifolds. In this paper, as in (Fletcher et al., 2004; Fletcher and Joshi, 2007; Pennec et al., 2006), we will use the linear approximation onto the tangent space of Φ.

Finally, as mentioned in §2.2, Cheng et al. (2009) assumes that the square-root ODF ψ can be approximated by a linear combination of the first *N* orthonormal basis functions, i.e., $\psi(s) \approx \sum_{i=1}^{N} c_i B_i(s)$, where $\{B_i(s)\}_{i=1}^{\infty}$ is any orthonormal basis set and $\{c_i\}$ are the coefficients. Under such a representation, the computation of PGA becomes significantly more efficient as the sample covariance matrix, rather than being infinite-dimensional, is a $N \times N$ matrix instead. However, as mentioned previously, it is not clear which method (ψ versus $\{c_i\}$) is more efficient and which method is more accurate. Finally, there is a possibility of missing a PC if the function is not included in the initial set using $\{c_i\}_{i=1}^{N}$.

2.5.2. ODF based Morphometry—Voxel based morphometry (VBM) (Ashburner and Friston, 2000) is a neuroimaging analysis technique to investigate regional differences in brain anatomy. VBM-type methods have made a significant contribution to the understanding of brain changes in several disorders and their relation to clinical measures (Whitwell, 2009). For example, VBM has been used to study a number of different disorders, including neurodegenerative diseases (Whitwell and Jack, 2005), multiple

sclerosis (Sepulcre et al., 2006), schizophrenia (Williams, 2008), movement disorders (Whitwell and Josephs, 2007) and epilepsy (Keller and Roberts, 2008). Appendix E gives an overview of VBM for multivariate testing. Voxel-based analysis may be applied to any vector-valued data, using nonparametric permutation tests. Therefore, one could treat an ODF as a Euclidean vector, and apply this technique directly to the original ODFs, or to the SH coefficients (via decomposing the ODF is as a linear function of the SH basis) by using the Euclidean T-squared statistic defined in Eqn. E.1. Such an approach is fast as computation for the Euclidean mean is closed-form. However, treating the ODFs or the SH coefficients of the ODF as Euclidean vectors is not theoretically correct, as it does not respect the geometry of the statistical manifold as the Riemannian distance is not equal to the Euclidean distance. We will next show how to generalize the Hotelling's T-squared statistic that accounts for the Fisher-Rao metric for PDFs.

We first assume that the ODF volumes of different subjects are already registered to a common template (HARDI registration remains an open research question, beyond the scope of this paper). Here, we obtain the registration of the ODF fields by registering the *T*2 images as we will show in §3. Having registered the ODF volumes, we next compute the mean and covariance matrices of the ODF-valued data, as described in §2.2 and §2.5.1, respectively. Let $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{n_\psi}$ and $\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_{n_\varphi}$ be the ODFs drawn from two populations and let $\bar{\boldsymbol{\psi}}$ and $\bar{\boldsymbol{\varphi}}$ be the sample means of the two populations computed using Algorithm 1. We generalize the Hotelling's T-squared statistic and VBM to ODFs using a similar strategy in an analogous fashion to what was developed for DTI in (Lepore et al., 2007; Hua et al., 2008) and we get the following equation for computing the T-squared statistic for ODFs,

$$T^2 = \frac{n_\psi n_\varphi}{n_\psi + n_\varphi} \overrightarrow{\bar{\boldsymbol{\psi}}\bar{\boldsymbol{\varphi}}}^\top \mathbf{W}^{-1} \overrightarrow{\bar{\boldsymbol{\psi}}\bar{\boldsymbol{\varphi}}} = \frac{n_\psi n_\varphi}{n_\psi + n_\varphi} \log_{\bar{\boldsymbol{\psi}}} \bar{\boldsymbol{\varphi}}^\top \mathbf{W}^{-1} \log_{\bar{\boldsymbol{\psi}}} \bar{\boldsymbol{\varphi}},$$

where

$$\begin{aligned} \mathbf{W} &= \frac{\sum_{i=1}^{n_\psi} \overrightarrow{\bar{\boldsymbol{\psi}}\boldsymbol{\psi}_i} \overrightarrow{\bar{\boldsymbol{\psi}}\boldsymbol{\psi}_i}^\top + \sum_{i=1}^{n_\varphi} \overrightarrow{\bar{\boldsymbol{\varphi}}\boldsymbol{\varphi}_i} \overrightarrow{\bar{\boldsymbol{\varphi}}\boldsymbol{\varphi}_i}^\top}{n_\psi + n_\varphi - 2} \\ &= \frac{\sum_{i=1}^{n_\psi} \log_{\bar{\boldsymbol{\psi}}} \boldsymbol{\psi}_i \log_{\bar{\boldsymbol{\psi}}} \boldsymbol{\psi}_i^\top + \sum_{i=1}^{n_\varphi} \log_{\bar{\boldsymbol{\varphi}}} \boldsymbol{\varphi}_i \log_{\bar{\boldsymbol{\varphi}}} \boldsymbol{\varphi}_i^\top}{n_\psi + n_\varphi - 2}. \end{aligned} \tag{31}$$

Having formulated the equivalent of the Hotelling's T-squared test for ODFs, we will perform a statistical group comparison of ODFs using multivariate hypothesis testing via the permutation tests. Alternatively, one could make use of the representation in (Cheng et al., 2009) which decomposes the square-root ODF $\boldsymbol{\psi}(\boldsymbol{s})$ as a linear function of the SH basis, and calculate the Riemannian T-squared statistic in Eqn. 31 based on the SH coefficients $\{c_i\}_{i=1}^N$. As in the case for the mean and PGA, this representation is computationally less intensive than using $\boldsymbol{\psi}(\boldsymbol{s})$. However, while using continuous $\boldsymbol{\psi}(\boldsymbol{s})$ guarantees that the T-squared statistic calculated in Eqn. 31 correctly determines whether the population means differ, this may not be true of the equivalent statistic calculated using $\{c_i\}_{i=1}^N$, or the discretized square-root ODFs (histograms with *M* directions), or the SH coefficients of the discretized square-root ODFs. Eqn. 31 is only applicable to the square-root ODFs, or its SH expansion, but not to the ODFs or the SH expansion of the ODFs.

3. Results and Discussion

In this section, we show how our proposed algorithms perform on synthetic data, phantom data, and real human brain HARDI data

3.1. Synthetic Data

In the first part of this section, we will assume that the ODFs are precomputed and we do not have raw HARDI signals. We first illustrate the various operations previously introduced, such as interpolation, filtering, and computation of the mean and principal components (PC) of the synthetic ODF fields in Fig. 2. Fig. 2(a) shows the bilinear Riemannian interpolation of the four endpoints where the original ODFs are shaded in orange. The top right and bottom left voxels contain ODFs of 1 fiber, the bottom right ODF of 2 fibers and the top left ODF of 3 fibers, with 1 fiber pointing out of the plane. Figs. 2(b)–2(d) show the Riemannian mean and the first two principal components of the ODF field in Fig. 2(a) with eigenvalues 0.104 and 0.041, respectively. The first PC is oriented from left to right, whereas the second PC is oriented going out of the plane. We apply the original PCA to ODFs, based on the Euclidean metric. Fig. 2(e) shows the bilinear Euclidean interpolation. Figs. 2(f)–2(h) show the Euclidean mean and the first two principal components of the ODF field in Fig. 2(e) with eigenvalues 0.1378 and 0.0641, respectively. The Riemannian and Euclidean means are similar, but the principal components computed by the two frameworks are very different. While the Riemannian framework is able to capture primarily, the direction going out of the plane as the second PC, the Euclidean framework does not exhibit the same sensitivity.

We also examine the interpolation when the rotational discrepancy, in degrees, between two ODFs varies. Essentially, the two ODFs have the same diffusion behavior and are related to each other via a rotation. Figs. 3(a)–3(c) show the results. When the angular difference is around 38 degrees, the interpolated ODFs start to deviate from the diffusion properties of each interpolant. The proposed interpolation scheme allows one to march along the manifold freely and hence does not restrict the interpolated ODFs to be related by a rotation. Since the diffusion properties do not remain the same using our interpolation scheme, the next natural question to ask is whether interpolation for ODFs will result in the so-called swelling effect. For DTI, Euclidean interpolation suffers from the so-called swelling effect, whereas Riemannian interpolation can overcome it. Fig. 4 shows the result of four different sets of interpolation. Set 1 shows interpolation of a single fiber with increased diffusivity in the principal direction, set 2 shows interpolation from a voxel with one fiber to an isotropic voxel, set 3 shows interpolation from a voxel with one fiber to one with two fibers, and finally, set 4 shows interpolation from a voxel with one fiber to another with one fiber rotated 90 degrees. We use the Shannon entropy of the ODF to measure the diffusivity of the ODF. As shown in Fig. 4, both frameworks suffer from the problem of the interpolants having a greater Shannon entropy than the original data points.

Next, we perform anisotropic filtering (AF) on an ODF field with a sharp discontinuity, as shown in Fig. 5(a). The function $c(x)=\exp(-\frac{x^2}{\kappa})$ is used in AF, and 30 iterations are performed. We add noise to this ODF field by applying a random tangent vector to each ODF, as shown in Fig. 5(b). We vary the parameter κ, for $\kappa = 0.1, 0.5, 1$, and 10. Figs. 5(c), 5(d), 5(e), and 5(f) show the results of Euclidean Anisotropic Filtering (EAF), where instead of taking into account the Riemannian structure of the ODFs, we simply treat each sampled ODF as a vector in Euclidean space, and perform anisotropic filtering in the classical Euclidean way. Figs. 5(g), 5(h), 5(i), and 5(j) show the results of Riemannian Anisotropic Filtering (RAF). Next, we examine the differences in performance between RAF and EAF over 100 trials for varying κ. For both types of AFs, we measure the amount of error from the ground-truth ODF field using two measures, Riemannian $\sum_{x}\left\|\log_{\psi_{truth}(x)}(\psi_{filtered}(x))\right\|_{x}^{2}$

and Euclidean $\Sigma_{\mathbf{x}}||\boldsymbol{\psi}_{truth}(\mathbf{x}) - \boldsymbol{\psi}_{filtered}(\mathbf{x})||^2$ dissimilarity. In each trial, we iterate AF 30 times. Table 1 shows the ratio of error for RAF to EAF under the Euclidean and Riemannian measures. RAF gives a smaller error than EAF, except when $\kappa = 0.1$, where the error of RAF is larger. This is because the Riemannian distance between two ODFs is always larger than or equal to the Euclidean distance, hence for the same value of κ, the rate of diffusion is faster for EAF than for RAF. This is why for large values of κ, EAF will blur the edges earlier than RAF, whereas for small values of κ, RAF is unable to filter away the noise.

We next evaluate our method's performance on synthetic data generated using the tensor method in (Descoteaux et al., 2007). Here instead of applying random tangent vectors to the clean ODFs, noise is added to the raw HARDI signals and the ODFs are calculated from the noisy HARDI signals. The HARDI signal $S(\mathbf{g})$ in the direction $\mathbf{g}$, where $\mathbf{g}$ is a unit vector, is generated by

$$S(\mathbf{g}) = \exp(-b\mathbf{g}^\top \mathbf{D}\mathbf{g}), \tag{32}$$

where the b-value is given and $\mathbf{D}$ is the diffusion tensor profile. Noisy versions of $S(\mathbf{g})$ are generated by adding complex Gaussian noise with zero mean and standard deviation $\sigma = \frac{1}{\zeta}$, where ζ is the signal-to-noise ratio (SNR). Again, we construct an ODF field with a sharp discontinuity, as shown in Fig. 6(a) using the tensor method. After noise is added to the HARDI signals, we obtain a noisy ODF field and an example is shown in Fig. 6(b). We apply Riemannian Gaussian Filtering (RGF) and RAF to the noisy ODF field and calculate the mean error between the denoised ODF field and the noiseless ODF field. Fig. 6 shows an example when the SNR = 20 and the b-value is 3000 mm/s^2. The voxels at the edge where the discontinuity occurs are marked in green. Fig. 7 shows the errors before and after applying RGF and RAF for various b-values and SNR ranging from 10 to 50. Fig. 7(a) shows the mean error across the entire ODF field. Notice that when the b-values increase, the errors between the noiseless ODF field and the noisy versions decreases. In addition, Fig. 7(a) shows clearly that both RAF and RGF give ODF fields that are significantly closer to the true ODF. Fig. 7(b) compares the errors in the edge voxels. Notice that the error of RGF is significantly more than the error of the noisy ODF field and the error of RAF. This is because RGF blurs the edge.

3.2. Phantom Data

We will now show qualitative results on the biological phantom created from two excised Sprague-Dawley rat spinal cords embedded in 2% agar (Campbell et al., 2005). A diffusion-weighted dataset was acquired using this phantom, with 90 diffusion encoding directions, with b-value of 3000 s/mm^2 (Campbell et al., 2005). Fig. 8(a) shows the Generalized Fractional Anisotropy (GFA) image for this dataset, which is defined as $\mathrm{GFA}(x) = \frac{\mathrm{std}(\boldsymbol{p}(\boldsymbol{x}))}{\mathrm{rms}(\boldsymbol{p}(\boldsymbol{x}))}$, where std is the standard deviation and rms is the root mean square value of the ODF $\boldsymbol{p}(\boldsymbol{x})$ at voxel $\boldsymbol{x}$. Figs. 8(b), 8(e) and 8(h) show three zoomed-in subregions of the ODF field in manually marked areas with different diffusion profiles. Two regions, shown in red and blue, contain a single fiber and one region, shown in green, contains two crossing fibers. Figs. 8(c), 8(f) and 8(i) show the ODF field after applying RGF and Figs. 8(d), 8(g) and 8(j) show the ODF field after applying RAF. While RAF and RGF reduce the noise in the ODF field, RGF tends to mislocate the structure of thin fibers, as shown in the box in red.

Next, we present experiments on a phantom dataset provided at http://www.lnao.fr/spip.php?rubrique79. This MR phantom contains a plethora of crossing, kissing, splitting and bending fiber configurations. Technical details of the phantom

construction are provided in (Poupon et al., 2008). Fig. 9(a) shows the ground truth fibers used in the phantom dataset. Diffusion-weighted MR images are obtained using the following imaging parameters. The image size is 64×64×3 with resolution of 3 × 3 × 3 mm. Two repetitions of 65 images are acquired, one image with no diffusion sensitization and $N = 64$ diffusion weighted images at $b = 1500$ s/mm^2, corresponding to the echo times $TE = 94$ ms.

Figs. 9(c) and 9(h) show the reconstructed ODF, Figs. 9(d) and 9(i) show the ODF field after applying Euclidean Gaussian Filtering (EGF), Figs. 9(e) and 9(j) show the ODF field after applying EAF, Figs. 9(f) and 9(k) show the ODF field after applying RGF, and Figs. 9(g) and 9(l) show the ODF field after applying RAF. The effects of filtering are obvious in the region marked in blue, as shown in Figs. 9(k) and 9(l), where the ODFs in the fiber bundle become oriented mostly in the northeast direction. This is correct according to the ground truth given in Fig. 9(a). However, both EGF and RGF merge into a single bundle two fiber bundles that are close to each other - as shown in the region that is marked in green in Figs. 9(j) and 9(k).

3.3. HARDI Human Data

Finally, we apply our Riemannian framework to a real HARDI human brain dataset (from Chiang et al. (2008)). Diffusion-weighted MR images were obtained at 4 Tesla using the following imaging parameters: 55 contiguous slices (2 mm thick), TR/TE= 8250/92.3ms, FOV= 230 × 230 and 1.8 × 1.8 mm^2 in-plane resolution. A total of 105 images were acquired: 11 with no diffusion sensitization and 94 diffusion-weighted images with $b = 1132$ s/mm^2 (Chiang et al., 2008). We apply our Riemannian filtering method to this ODF field, shown in Fig. 10(a). Figs. 10(b) and 10(c) show the results of RGF and RAF, respectively. Figs. 10(d)–10(l) shows the zoomed-in regions of the different ODF fields in manually marked areas of different diffusion. The function $c(\cdot)$ used in RAF is a Gaussian function, and 30 iterations are done. RAF is able to smooth the ODF field sufficiently but still preserves the discontinuities between different bundles. In addition, both RAF and RGF give a GFA map that is significantly cleaner than that of the unfiltered ODF field and the different regions of the brain can be seen more clearly.

Next, we use our framework to compute multi-subject statistics on ODFs in a database of 40 right-handed healthy young adults, scanned with the imaging parameters given before (Chiang et al., 2009). In order to bring all the ODF volumes of the subjects to the same reference frame, there is a need to register the ODF volume of each subject to a common ODF template. This requires solving the HARDI registration problem, which is an open research question outside the scope of this paper. We circumvent HARDI registration by applying an approximate algorithm, as shown in Fig. 11. First of all, the non-diffusion-weighted images (also called $b = 0$ images) from all subjects compute a group "average-shape" brain, or minimal deformation template (MDT). The $b = 0$ images are then nonlinearly registered to this template. Since nonlinear registration is used, the deformation field and the corresponding Jacobian of the transformation for each subject differs depending on the voxel coordinates as well. Therefore, based on this registration, we denote the deformation field $T_{i,\boldsymbol{x}}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and the corresponding Jacobian field $J_{i,\boldsymbol{x}}: \mathbb{R}^3 \rightarrow \mathbb{R}^{3\times 3}$ of each subject, where $i = 1,\ldots, N_{subjects}$ and $\boldsymbol{x}$ is the voxel coordinates. The deformation field $T_{i,\boldsymbol{x}}$ of each subject is then applied to the diffusion weighed images of the same subject. We extract the rotational component $R_{i,\boldsymbol{x}}$ of the Jacobian of the transformation, where $J_{i,\boldsymbol{x}} = R_{i,\boldsymbol{x}}S_{i,\boldsymbol{x}}$ with $S_{i,\boldsymbol{x}}$ is a shear matrix, and apply it to each gradient direction $\tilde{\mathbf{g}}_{j,\boldsymbol{x}} = R_{i,\boldsymbol{x}}\mathbf{g}_j$ in order to reorient the ODF according to the spatial transformation. Note that since the Jacobian $J_{i,\boldsymbol{x}}$ of the transformation and the extracted rotational component $R_{i,\boldsymbol{x}}$ are different at different voxels, and therefore the reorientated gradient is dependent on $\boldsymbol{x}$.

The ODFs are then constructed from the transformed diffusion weighed images at the transformed gradient directions. Fig. 12 shows different axial slices of the calculated mean ODF. Notice that the mean ODF image preserves the features in the corpus callosum, cingulum, and corticospinal tract bundles, in which the directions of the fiber tracts are seen clearly. We also show two different ROIs in Fig. 13. The first ROI shown in Figs. 13(a)–13(d) is the region where the corticospinal tract and the corpus callosum cross and intermix. Notice that in the mean ODF field, we still see the presence of crossing fibers, as shown in green. In the second ROI shown in Figs. 13(e)–13(h), the cingulum and the corpus callosum do not cross each other; however, we expect to see some voxels where there is partial volume averaging of the cingulum and the corpus callosum. Notice that in the mean ODF field, we still see the effect of partial volume averaging, as shown in blue.

Finally, we show how ODF-based morphometry may be used to study brain asymmetry, as described in §2.5.2. While little is known about white matter connectivity asymmetry (Dubois et al., 2009; Buchel et al., 2004), several recent studies have examined left-right differences from DTI-derived measures in large human populations ($N = 374$; (Jahanshad et al., 2010)). ODF-based morphometry allows us to extend the study of hemispheric asymmetry to a microscopic scale by examining differences in fiber characteristics across hemispheres rather than just differences in structure shapes and volumes. In order to study brain asymmetry, we reflect each subject's ODF field with respect to the mid-sagittal plane. We compute voxel-wise maps of significant asymmetries of the differences between the ODFs in the left and right hemispheres, using the Hotelling's T-squared test to measure differences in the left and right mean ODFs.

Statistically significant asymmetries ($p < 0.05$) are shown in Fig. 14. ODF based morphometry reveals a number of white matter microstructural asymmetries. First of all, the planum temporale, a highly lateralized brain structure involved with language processing, is clearly detected. This localization is consistent with the literature, as the language comprehension areas are among the most highly functionally lateralized areas of the brain. Secondly, the posterior limb of the internal capsule and the corticospinal or pyramidal tract are also detected. These areas are major components of the primary sensorimotor pathways. This result agrees well with previous studies reporting the asymmetry of these fiber tracts (Ardekani et al., 2007;Herve et al., 2006;Rademacher et al., 2001;Tensaouti et al., 2009;Westerhausen et al., 2007), further validating this approach. The main regions of the brain that are deemed asymmetric by the two tests are generally in agreement with each other. The suprathreshold cluster test generally gives contiguous regions, at the cost of reduced localizing power since only large regions are detected, but the single threshold approach does not.

Finally, one can perform morphometry directly on the ODFs by treating them as vectors, as mentioned in §2.5.2. Here, we perform the much simpler Euclidean equivalent and compare the results to those obtained from the Riemannian version. Statistically significant asymmetries ($p < 0.05$) are found by the Euclidean metric (Fig. 15). From Figs. 15(a)–15(b), we see that the single threshold test does not detect the asymmetric region of the planum temporale. In the suprathreshold cluster test, if we threshold the regions at the same value as in the Riemannian test, we see that the planum temporale is not detected as well, as shown in Figs. 15(c)–15(d). As the two metrics are not equivalent, with the Riemannian distance being greater than or equal to the Euclidean distance, one could lower this threshold to see if similar results can be obtained. The results of the suprathreshold cluster test threshold at a lower value is shown in Figs. 15(e)–15(f). Such a strategy will reduce the localizing power of the test, as almost everywhere is now deemed asymmetric.

4. Conclusion

We presented a Riemannian framework for processing ODFs. Our approach does not require that the ODF be represented by any specific parametric model. Our framework also offers desirable properties that the other metrics do not have. The various Riemannian operations are in closed form and easily and efficiently computed. Choosing the correct metric for ODFs remains an open question and the question of how we can add in extra constraints into the current framework that respect the appropriate diffusion properties of the ODF field remains unsolved. Results on synthetic, phantom and real data quantify the advantage of working with our proposed framework versus the Euclidean approach. In addition, we also performed ODF-based morphometry to study asymmetries in the microstructure of the brain in a population of healthy subjects. This led to results that are in general agreement with prior reports, further validating our method and showing it is applicable to studies of structural brain asymmetry. Future research will apply the same kind of analysis to study differences between the brains of healthy and diseased subjects to better understand how different disorders affect the brain.

Acknowledgments

We would like to thank Neda Jahanshad of UCLA for providing us with the registered datasets and Jennifer Campbell of the McConnell Brain Imaging Centre, McGill University for providing us with the biological phantom data. In addition, the first author would like to thank Professor Guillermo Sapiro of the University of Minnesota for inviting her for an extremely fruitful visit to his lab where the ODF based morphometry work was conceptualized. This work was funded by startup funds from JHU, by grants NSF CAREER IIS-0447739, NIH R01 HD050735, NIH R01 EB007813, NIH R01 EB008432, NIH P41 RR008079, NIH P30 NS057091, ONR N00014-05-10836 and ONR N00014-09-1-0084, and the University of Minnesota Institute for Translational Neuroscience.

References

Absil, P-A.; Mahony, R.; Sepulchre, R. Optimization Algorithms on Matrix Manifolds. Princeton University Press; Princeton, NJ: 2008.

Aganj I, Lenglet C, Sapiro G, Yacoub E, Ugurbil K, Harel N. Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. Magnetic Resonance in Medicine. 2010; 64:554–566. [PubMed: 20535807]

Ardekani S, Kumar A, Bartzokis G, Sinha U. Exploratory voxel-based analysis of diffusion indices and hemispheric asymmetry in normal aging. Magnetic Resonance Imaging. 2007; 25 (2):154 – 167. [PubMed: 17275609]

Arsigny V, Fillard P, Pennec X, Ayache N. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magnetic Resonance in Medicine. 2006; 56:411–421. [PubMed: 16788917]

Ashburner J, Friston KJ. Voxel-based morphometry–the methods. NeuroImage. Jun; 2000 11 (6):805–821. [PubMed: 10860804]

Awate S, Hui Z, Gee J. A fuzzy, nonparametric segmentation framework for DTI and MRI analysis: With applications to DTI-tract extraction. IEEE Trans on Medical Imaging. 2007; 26 (11):1525–1536.

Barmpoutis A, Hwang M, Howland D, Forder J, Vemuri B. Regularized positive-definite fourth order tensor field estimation from DW-MRI. NeuroImage. 2009; 45 (1, Supplement 1):S153–S162. [PubMed: 19063978]

Basser P, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. Journal of Magnetic Resonance B. 1994; 103:247–254.

Basser P, Pajevic S, Pierpaoli C, Duda J, Aldroubi A. In vivo fiber tractography using DT-MRI data. Magnetic Resonance in Medicine. 2000; 44 (1):625–632. [PubMed: 11025519]

Batchelor PG, Moakher M, Atkinson D, Calamante F, Connelly A. A rigorous framework for diffusion tensor calculus. Magnetic Resonance in Medicine. 2005; 53 (1):221–225. [PubMed: 15690523]

Berger, M. A Panoramic View of Riemannian Geometry. Springer; 2003.

Brun, A.; Knutsson, H.; Park, HJ.; Shenton, ME.; Westin, C-F. Medical Image Computing and Computer-Assisted Intervention. 2004. Clustering fiber tracts using normalized cuts; p. 368-375.

Brun A, Park H-J, Knutsson H, Westin C-F. Coloring of DT-MRI fiber traces using Laplacian eigenmaps. Computer Aided Systems Theory. 2003:564–572.

Buchel C, Raedler T, Sommer M, Sach M, Weiller C, Koch M. White Matter Asymmetry in the Human Brain: A Diffusion Tensor MRI Study. Cereb Cortex. 2004; 14 (9):945–951. [PubMed: 15115737]

Buss SR, Fillmore JP. Spherical averages and applications to spherical splines and interpolation. ACM Trans Graph. 2001; 20 (2):95–126.

Campbell JS, Siddiqi K, Rymar VV, Sadikot AF, Pike GB. Flow-based fiber tracking with diffusion tensor and q-ball data: Validation and comparison to principal diffusion direction techniques. NeuroImage. 2005; 27 (4):725–736. [PubMed: 16111897]

Cao Y, Miller M, Winslow R, Younes L. Large deformation diffeomorphic metric mapping of vector fields. IEEE Trans on Medical Imaging. Sept; 2005 24 (9):1216–1230.

Castano-Moraga C, Lenglet C, Deriche R, Ruiz-Alzola J. A riemannian approach to anisotropic filtering of tensor fields. Signal Processing. 2007; 87 (2):263–276.

Cencov, NN. Translations of Mathematical Monographs. Vol. 53. AMS; 1982. Statistical decision rules and optimal inference.

Cheng, J.; Ghosh, A.; Jiang, T.; Deriche, R. A Riemannian framework for orientation distribution function computing. LNCS; Medical Image Computing and Computer-Assisted Intervention; Springer-Verlag; 2009. p. 911-918.

Chiang, M-C.; Barysheva, M.; Lee, AD.; Madsen, SK.; Klunder, AD.; Toga, AW.; McMahon, K.; de Zubicaray, GI.; Meredith, M.; Wright, MJ.; Srivastava, A.; Balov, N.; Thompson, PM. Medical Image Computing and Computer-Assisted Intervention. 2008. Brain fiber architecture, genetics, and intelligence: A high angular resolution diffusion imaging (HARDI) study; p. 1060-1067.

Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C, Klunder AD, Toga AW, McMahon KL, de Zubicaray GI, Wright MJ, Srivastava A, Balov N, Thompson PM. Genetics of Brain Fiber Architecture and Intellectual Performance. J Neurosci. 2009; 29 (7):2212–2224. [PubMed: 19228974]

Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Regularized, fast and robust analytical Q-ball imaging. Magnetic Resonance in Medicine. 2007; 58:497–510. [PubMed: 17763358]

Descoteaux M, Deriche R. High angular resolution diffusion MRI segmentation using region-based statistical surface evolution. Journal of Mathematical Imaging and Vision. 2009; 33 (2):239–252.

Dubois J, Hertz-Pannier L, Cachia A, Mangin J, Le Bihan D, Dehaene-Lambertz G. Structural Asymmetries in the Infant Language and Sensori-Motor Networks. Cereb Cortex. 2009; 19 (2): 414–423. [PubMed: 18562332]

Fletcher PT, Joshi S. Riemannian geometry for the statistical analysis of diffusion tensor data. Signal Processing. 2007; 87(2)

Fletcher PT, Lu C, Pizer SM, Joshi S. Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans on Medical Imaging. 2004; 23 (8):995–1005.

Frank LR. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. Magnetic Resonance in Medicine. 2002; 47 (6):1083–1099. [PubMed: 12111955]

Frechet M. Les elements aleatoires de nature quelconque dans un espace distancie. Annales De L'Institut Henri Poincare. 1948; 10:235–310.

Ghosh, A.; Descoteaux, M.; Deriche, R. Medical Image Computing and Computer-Assisted Intervention. 2008. Riemannian framework for estimating symmetric positive definite 4th order diffusion tensors; p. 858-865.

Goh, A.; Lenglet, C.; Thompson, P.; Vidal, R. Medical Image Computing and Computer-Assisted Intervention. LNCS 5761. Springer-Verlag; 2009a. Estimating orientation distribution functions with probability density constraints and spatial regularity; p. 877-885.

Goh, A.; Lenglet, C.; Thompson, P.; Vidal, R. A nonparametric Riemannian framework for processing High Angular Resolution Diffusion Images (HARDI). Conference on Computer Vision and Pattern Recognition; 2009b. p. 2496-2503.

Goh, A.; Vidal, R. Algebraic methods for direct and feature based registration of diffusion tensor images. European Conference on Computer Vision; 2006. p. 514-525.

Goh, A.; Vidal, R. Segmenting fiber bundles in diffusion tensor images. European Conference on Computer Vision; 2008a. p. 238-250.

Goh, A.; Vidal, R. Unsupervised Riemannian clustering of probability density functions. European Conference on Machine Learning; 2008b. p. 377-392.

Good, PI. Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer Series in Statistics). Springer-Verlag New York, Inc; 2004.

Guimond, A.; Guttmann, CRG.; Warfield, SK.; Westin, C-F. Deformable registration of DT-MRI data based on transformation invariant tensor characteristics. IEEE Int. Symposium on Biomedical Imaging; 2002. p. 761-764.

Hagmann P, Thiran J, Jonasson L, Vandergheynst P, Clarke S, Maeder P, Meuli R. DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. Neuroimage. 2003; 19 (3):545–554. [PubMed: 12880786]

Herve PY, Crivello F, Perchey G, Mazoyer B, Tzourio-Mazoyer N. Handedness and cerebral anatomical asymmetries in young adult males. NeuroImage. 2006; 29 (4):1066–1079. [PubMed: 16198126]

Hess CP, Mukherjee P, Han ET, Xu D, Vigneron DB. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magnetic Resonance in Medicine. 2006; 56 (1): 104–117. [PubMed: 16755539]

Hua X, Leow AD, Parikshak N, Lee S, Chiang MC, Toga AW Jr, CRJ, Weiner MW, Thompson PM. Tensor-based morphometry as a neuroimaging biomarker for alzheimer's disease: An MRI study of 676 ad, mci, and normal subjects. NeuroImage. 2008; 43 (3):458–469. [PubMed: 18691658]

Jahanshad N, Lee AD, Barysheva M, McMahon KL, de Zubicaray GI. Genetic influences on brain asymmetry: A dti study of 374 twins and siblings. NeuroImage. 2010; 52 (2):455 – 469. [PubMed: 20430102]

Jansons KM, Alexander DC. Persistent angular structure: new insights from diffusion magnetic resonance imaging data. Inverse Problems. 2003; 19 (5):1031–1046.

Jolliffe, I. Principal Component Analysis. Springer-Verlag; New York: 1986.

Jonasson L, Bresson X, Hagmann P, Cuisenaire O, Meuli R, Thiran J. White matter fiber tract segmentation in DT-MRI using geometric flows. Medical Image Analysis. October.2004

Jonasson L, Hagmann P, Pollo C, Bresson X, Richero Wilson C, Meuli R, Thiran J. A level set method for segmentation of the thalamus and its nuclei in DT-MRI. Signal Processing. 2007; 87 (2):309–321.

Jonasson, L.; Hagmann, P.; Thiran, J.; Wedeen, V. Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering. Proceedings of 13th Annual Meeting ISMRM; Miami. ISCAS. SPIE. 2005. p. 1310

Keller SS, Roberts N. Voxel-based morphometry of temporal lobe epilepsy: An introduction and review of the literature. Epilepsia. 2008; 49 (5):741–757. [PubMed: 18177358]

Kindlmann, G.; Estepar, RSJ.; Niethammer, M.; Haker, S.; Westin, C-F. Medical Image Computing and Computer-Assisted Intervention. 2007. Geodesic-loxodromes for diffusion tensor interpolation and difference measurement; p. 1-9.

Krajsek, K.; Menzel, M.; Zwanger, M.; Scharr, H. Riemannian anisotropic diffusion for tensor valued images. European Conference on Computer Vision; 2008. p. 326-339.

Lenglet C, Rousson M, Deriche R. DTI segmentation by statistical surface evolution. IEEE Trans Med Imaging. 2006a; 25 (6):685–700. [PubMed: 16768234]

Lenglet C, Rousson M, Deriche R, Faugeras O. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor mri processing. J Math Imaging Vis. 2006b; 25 (3):423–444.

Leow AD, Zhu S, Zhan L, McMahon K, de Zubicaray GI, Meredith M, Wright MJ, Toga AW, Thompson PM. The tensor distribution function. Magnetic Resonance in Medicine. 2009; 61:205–214. [PubMed: 19097208]

Lepore, N.; Brun, CA.; Pennec, X.; Chou, Y-Y.; Lopez, OL.; Aizenstein, H.; Becker, JT.; Toga, AW.; Thompson, PM. Medical Image Computing and Computer-Assisted Intervention. 2007. Mean template for tensor-based morphometry using deformation tensors; p. 826-833.

McGraw, T.; Vemuri, B.; Yezierski, R.; Mareci, T. Segmentation of high angular resolution diffusion MRI modeled as a field of von Mises-Fisher mixtures. European Conference on Computer Vision; 2006. p. 463-475.

Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Human Brain Mapping. January; 2002 15 (1):1–25. [PubMed: 11747097]

Niethammer, M.; Estepar, RSJ.; Bouix, S.; Shenton, M.; Westin, C-F. On Diffusion Tensor Estimation. International Conference of the IEEE Engineering in Medicine and Biology Society; 2006. p. 2622-2625.

Özarslan E, Mareci T. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. Magnetic Resonance in Medicine. 2003; 50:955–965. [PubMed: 14587006]

Özarslan E, Shepherd T, Vemuri B, Blackband S, Mareci T. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). NeuroImage. 2006; 31:1086–1103. [PubMed: 16546404]

Pasternak O, Sochen N, Basser PJ. The effect of metric selection on the analysis of diffusion tensor MRI data. NeuroImage. 2010; 49 (3):2190 –2204. [PubMed: 19879947]

Pennec X, Fillard P, Ayache N. A Riemannian framework for tensor computing. Int Journal of Computer Vision. 2006; 66 (1):41–46.

Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. IEEE Trans on Pattern Analysis and Machine Intelligence. Jul; 1990 12 (7):629–639.

Poupon C, Rieul B, Kezele I, Perrin M, Poupon F, Mangin JF. New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. Magnetic Resonance in Medicine. 2008; 60 (6):1276–1283. [PubMed: 19030160]

Rademacher J, Burgel U, Geyer S, Schormann T, Schleicher A, Freund HJ, Zilles K. Variability and asymmetry in the human precentral motor system: A cytoarchitectonic and myeloarchitectonic brain mapping study. Brain. 2001; 124 (11):2232–2258. [PubMed: 11673325]

Rao CR. Information and accuracy attainable in the estimation of statistical parameters. Bull Calcutta Math Soc. 1945; 37:81–89.

Rathi Y, Michailovich O, Shenton ME, Bouix S. Directional functions for orientation distribution estimation. Medical Image Analysis. 2009; 13 (3):432–444. [PubMed: 19269242]

Ruiz-Alzola J, Westin CF, Warfield SK, Alberola C, Maier SE, Kikinis R. Nonrigid registration of 3D tensor medical data. Medical Image Analysis. 2002; 6 (2):143–161. [PubMed: 12045001]

Savadjiev P, Campbell J, Pike G, Siddiqi K. 3D curve inference for diffusion MRI regularization and fibre tractography. Medical Image Analysis. 2006; 10:799–813. [PubMed: 16919994]

Schölkopf, B.; Smola, A. Learning with Kernels. MIT Press; Cambridge, MA: 2002.

Sepulcre J, Sastre-Garriga J, Cercignani M, Ingle GT, Miller DH, Thompson AJ. Regional gray matter atrophy in early primary progressive multiple sclerosis: A voxel-based morphometry study. Arch Neurol. 2006; 63 (8):1175–1180. [PubMed: 16908748]

Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 23rd ACM national conference; New York, NY, USA: ACM; 1968. p. 517-524.

Sommer, S.; Lauze, F.; Hauberg, S.; Nielsen, M. Manifold valued statistics, exact principal geodesic analysis and the effect of linear approximations. European Conference on Computer Vision; 2010a. p. 43-56.

Sommer, S.; Lauze, F.; Nielsen, M. CoRR. 2010b. The differential of the exponential map, jacobi fields and exact principal geodesic analysis. abs/1008.1902

Srivastava, A.; Jermyn, I.; Joshi, S. Riemannian analysis of probability density functions with applications in vision. IEEE Conference on Computer Vision and Pattern Recognition; 2007. p. 1-8.

Stejskal EO, Tanner JE. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. The Journal of Chemical Physics. 1965; 42 (1):288–292.

Tensaouti, F.; Lotterie, JA.; Clarisse, P.; Berry, I. Assessment of asymmetry in pyramidal tract by using fiber tracking. 2009. p. 1341-1346.

Tristan-Vega A, Westin CF, Aja-Fernandez S. Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. NeuroImage. 2009; 47 (2):638–650. [PubMed: 19393321]

Tschumperle D, Deriche R. DT-MRI images: estimation, regularization, and applications. Computer Aided Systems Theory - EUROCAST 2003. 2003; 2809:530–541.

Tuch DS. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magnetic Resonance in Medicine. 2002; 48:577–582. [PubMed: 12353272]

Tuch DS. Q-ball imaging. Magnetic Resonance in Medicine. 2004; 52 (6):1358–1372. [PubMed: 15562495]

Vemuri, B.; Chen, Y.; Rao, M.; McGraw, T.; Wang, Z.; Mareci, T. Fiber tract mapping from diffusion tensor MRI. IEEE Workshop on Variational and Level Set Methods in Computer Vision; 2001. p. 81-88.

Wang S, Poptani H, Bilello M, Wu X, Woo J, Elman L, McCluskey L, Krejza J, Melhem E. Diffusion tensor imaging in amyotrophic lateral sclerosis: Volumetric analysis of the corticospinal tract. American Journal of Neuroradiology. 2006; 27:1234–1238. [PubMed: 16775271]

Wang Z, Vemuri BC. DTI segmentation using an information theoretic tensor dissimilarity measure. IEEE Trans on Medical Imaging. October; 2005 24 (10):1267–1277.

Wang Z, Vemuri BC, Chen Y, Mareci TH. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. IEEE Trans on Medical Imaging. 2004; 23:930–939.

Wedeen VJ, Hagmann P, Tseng WYI, Reese TG, Weisskoff RM. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn Reson Med. 2005; 54 (6):1377–1386. [PubMed: 16247738]

Weickert, J. A review of nonlinear diffusion filtering. SCALE-SPACE ’97: Proceedings of the First International Conference on Scale-Space Theory in Computer Vision; 1997. p. 3-28.

Weickert, J.; Brox, T. Vol. 313 of Contemporary Mathematics. AMS, Ch. Diffusion and regularization of vector- and matrix-valued images. 2002. Inverse Problems, Image Analysis and Medical Imaging; p. 251-268.

Westerhausen R, Huster RJ, Kreuder F, Wittling W, Schweiger E. Corticospinal tract asymmetries at the level of the internal capsule: Is there an association with handedness? NeuroImage. 2007; 37 (2):379–386. [PubMed: 17601751]

Whitwell JL. Voxel-Based Morphometry: An Automated Technique for Assessing Structural Changes in the Brain. J Neurosci. 2009; 29 (31):9661–9664. [PubMed: 19657018]

Whitwell JL, Jack CRJ. Comparisons between alzheimer disease, frontotemporal lobar degeneration, and normal aging with brain mapping. Topics in Magnetic Resonance Imaging. 2005; 16 (6):409–425. [PubMed: 17088691]

Whitwell, JL.; Josephs, KA. Voxel-based morphometry and its application to movement disorders. Parkinsonism & Related Disorders; proceedings of the XVII WFN World Congress on Parkinson’s Disease and Related Disorders.; 2007. p. S406 -S416.

Williams LM. Voxel-based morphometry in schizophrenia: implications for neurodevelopmental connectivity models, cognition and affect. Expert Review of Neurotherapeutics. 2008; 8 (7):1049–1065. [PubMed: 18590476]

Zalesky A. DT-MRI fiber tracking: A shortest path approach. IEEE Trans on Medical Imaging. 2008; 27 (10):1458–1471.

Zhang H, Yushkevich PA, Alexander DC, Gee JC. Deformable registration of diffusion tensor mr images with explicit orientation optimization. Medical Image Analysis. 2006; 10 (5):764 – 785. [PubMed: 16899392]

Zhukov L, Museth K, Breen D, Whitaker R, Barr A. Level set segmentation and modeling of DT-MRI human brain data. Journal of Electronic Imaging. 2003:125–133.

Ziyan U, Tuch D, Westin C-F. Segmentation of thalamic nuclei from DTI using spectral clustering. Medical Image Computing and Computer-Assisted Intervention. 2006:807–814. [PubMed: 17354847]

Appendix A. Extracting Orientation from HARDI Signals

In this section, we review what orientation PDFs are and how they can be extracted from HARDI signals. Briefly, DW-MRI measures the molecular water motion in the direction of a given diffusion gradient **g**. For this purpose, the Stejskal-Tanner imaging sequence (Stejskal and Tanner, 1965) is commonly used. In this case, the MR signal attenuation $S(\mathbf{q}, \τ)$ is related to the diffusion PDF of the net displacement vector **R**, $\mathbf{P}(\mathbf{R}|\τ)$, by

$$\frac{S(q,\tau)}{S_0}=\int_{\mathbb{R}^3}\mathbf{P}(\mathbf{R}|\tau)e^{-2\pi i q^{\top}\mathbf{R}}d\mathbf{R}=\mathcal{F}[\mathbf{P}(\mathbf{R}|\tau)], \tag{A.1}$$

where S_0 is a reference signal acquired with no diffusion gradient, τ is the molecular diffusion time, $\mathbf{q} = \gamma\delta\mathbf{g}/2\pi$ is the displacement reciprocal vector, γ is the gyromagnetic ratio of water protons and δ is the duration of the diffusion gradients.

The signal $S(\mathbf{q}, \tau)$ is thus expressed as the 3D Fourier transform $\mathcal{F}$ of the diffusion PDF, also known as the ensemble average propagator (EAP), which is the function we want to reconstruct in DW-MRI. To do this accurately, one could sample the *q*-space at many gradient directions and then compute the inverse Fourier transform. However, this would be a computationally intensive and would require acquiring too many images. A more efficient scheme is to sample from a single spherical shell ($||\mathbf{q}|| = q_0$) and estimate the orientation PDFs. Examples include Q-ball imaging QBI (Tuch, 2004), Persistent Angular Structure (Jansons and Alexander, 2003) and Diffusion Orientation Transform (Özarslan et al., 2006).

While the framework proposed in this paper can be used for any definition of orientation PDFs, in this paper, we use the definition of orientation distribution function (ODF) proposed in (Wedeen et al., 2005; Aganj et al., 2010; Tristan-Vega et al., 2009). The ODF is defined as

$$p(\theta,\varphi) \doteq p(\boldsymbol{r})=\int_0^{\infty}\mathbf{P}(R\boldsymbol{r})R^2dR, \tag{A.2}$$

where $\boldsymbol{r} = \boldsymbol{R}/R$ and $\boldsymbol{R} = R(\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta))$, where $R \geq 0$, $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi]$. Notice that the ODF is the angular profile of the diffusion PDF of water molecules, so it is actually a PDF on the 2-sphere. Therefore, the manifold structure of the ODFs should be considered as well.

Appendix B. Existence and Uniqueness of the Mean of a Set of Points on the Sphere

Unlike the Euclidean case, in general there is no closed form formula for the intrinsic mean on a Riemannian manifold. Moreover, there is no guarantee that it exists or is unique. The following theorem of Buss and Fillmore (2001) gives the conditions for the existence and uniqueness of the mean of a set of points on the sphere.

Theorem B.1. (Buss and Fillmore, 2001)

Suppose the data points $\boldsymbol{\psi}_1,\ldots, \boldsymbol{\psi}_n$ all lie in a hemisphere with at least one point $\boldsymbol{\psi}_i$ lying in the interior of the hemisphere. Then there exists only one global minimizer to the cost function defined in Eqn. (7).

As a direct consequence of Theorem B.1, Cheng et al. (2009) showed that the mean and the weighted Riemannian mean of a set of points in the space of ODFs *is unique and exists for* $\boldsymbol{\Psi}$.

Appendix C. Calculation of Infinite Dimensional Covariance Matrix for Continuous ODFs

In the case of ODFs, each data point $\boldsymbol{\varphi}_i$ is a continuous function. Therefore, the sample covariance matrix becomes an infinite dimensional operator. Therefore, we need to compute eigenfunctions rather than the eigenvectors. To find the eigenfunctions used in PGA, we will now make use of the kernel trick, commonly used in kernel PCA (Schölkopf and Smola, 2002). More precisely, we seek the non-zero eigenvalues λ and non-zero eigenfunction $\mathbf{u}$ of $\mathrm{cov}(\boldsymbol{\psi})$ i.e., $\mathrm{cov}(\boldsymbol{\psi})\mathbf{u} = \lambda\mathbf{u}$. Let $\mathcal{L}: \mathbb{R}^n \to \boldsymbol{\psi}$ be the operator $\mathcal{L}(v)=\sum_{i=1}^{n}\boldsymbol{\psi}_i v_i$. Then $\mathrm{cov}(\boldsymbol{\psi}) = \mathcal{L}\mathcal{L}^*$, where $\mathcal{L}^*: \boldsymbol{\psi} \to \mathbb{R}^n$ is the adjoint operator. Therefore

$$\lambda\mathbf{u}=\mathcal{L}\mathcal{L}^*\mathbf{u} \iff \mathbf{u}=\mathcal{L}(\lambda^{-1}\mathcal{L}^*\mathbf{u}).$$

Notice that $\mathbf{u}$ is always in the span of the operator $\mathcal{L}$. Now if we define $\mathbf{v} = \lambda^{-1}\mathcal{L}^*\mathbf{u}$, we see that

$$\mathcal{L}^*\mathcal{L}\boldsymbol{v}=\lambda^{-1}\mathcal{L}^*\mathcal{L}\mathcal{L}^*\mathbf{u}=\mathcal{L}^*\mathbf{u}=\lambda\boldsymbol{v}.$$

That is, $\mathbf{v}$ is an eigenfunction of $\mathcal{L}^*\mathcal{L}$ with eigenvalue λ. While $\mathcal{L}\mathcal{L}^*$ is infinite dimensional, $\mathcal{L}^*\mathcal{L}$ is an $n \times n$ matrix. Therefore, to compute $\mathbf{u}$, one would first compute $\mathbf{v}$, an eigenvector of $\mathcal{L}^*\mathcal{L}$, and recover the corresponding $\mathbf{u}$ as $\mathbf{u} = \mathcal{L}\mathbf{v}$. Finally, if we use the same terminology as in kernel PCA, the kernel matrix $\mathcal{K} \in \mathbb{R}^{n\times n}$ used here makes use of the Riemannian distance and is

$$\mathcal{K}_{ij}=\langle\varphi_i,\varphi_j\rangle_{\overline{\psi}} \in \mathbb{R}.$$

In this case, we are using the Fisher-Rao distance to build the kernel matrix and proceeding as in done in kernel PCA. To summarize, we have shown how to calculate the principal components for a set of probability density functions on $\mathbb{S}^2$. In practice, the ODF is not a continuous PDF; rather it is a histogram over M different gradient directions. Therefore, instead of using the kernel trick, we can calculate the sample covariance matrix $\mathrm{cov}(\boldsymbol{\psi})=\frac{1}{n-1}\sum_{i=1}^{n}\boldsymbol{\varphi}_i\boldsymbol{\varphi}_i^{\top} \in \mathbb{R}^{M\times M}$ exactly.

Appendix D. Discrete Representation of ODF in Practice

When ODFs are approximated using different sets of basis functions such as spherical harmonics (SH) (Özarslan and Mareci, 2003; Descoteaux et al., 2007; Frank, 2002; Hess et al., 2006; Aganj et al., 2010), they are continuous functions defined on $\mathbb{S}^2$. In practice, however, the ODF is sometimes discretized to be a histogram over M different gradient directions. This is precisely the approach that we adopt in this paper. How would one go about selecting the size and the number of bins of the histogram is an important question. So far, existing literature has proposed to sample directions equally spaced on the sphere by the

icosahedral tessellation of the unit hemisphere, ranging from the 2nd order with $M = 162$ directions (Barmpoutis et al., 2009), the 3rd order with $M = 642$ directions (Leow et al., 2009), to the 4th order with $M = 2562$ (Zalesky, 2008). In this paper, we choose $M = 162$ as experimentally, we found that it is a good tradeoff between computational overhead and accuracy. The surface resolution is controlled by the number of bins; when the number of bins is small, the surface is a crude representation of the continuous ODF; when the number of bins is big, the surface is a good representation. Our proposed framework directly works with the histogram. Therefore, when the number of bins is large, the increased surface resolution comes at the expense of increased computational processing cost.

In (J. Cheng and R. Deriche, personal communications, 18 January 2011), the issue of normalization is discussed and we will give a summary here. In theory, if the ODF $\boldsymbol{p}(\boldsymbol{s})$ is a continuous probability density function defined on $\mathbb{S}^2$, we have $\int_{\boldsymbol{s}\in\mathbb{S}^2}\boldsymbol{p}(\boldsymbol{s})d\boldsymbol{s} = 1$. In the discrete case, we have $1=\int_{s\in\mathbb{S}^2} p(\boldsymbol{s})d\boldsymbol{s} \approx \sum_{i=1}^{M} p(\boldsymbol{s}_i)\Delta_i$ where Δ_i is the surface area at the direction $\boldsymbol{s}_i$. Note that in the discrete case, $\sum_{i=1}^{M} p(\boldsymbol{s}_i) \neq 1$. To show this, consider when M equally spaced samples are used, we have $\Delta=\Delta_i=\frac{4\pi}{M}$ and hence $\sum_{i=1}^{M} p(\boldsymbol{s}_i)=\frac{M}{4\pi}$. Therefore, if the estimated discrete ODF does not satisfy the PDF condition, then it should be normalized by a constant Z as done in (Tuch, 2004) such that $\frac{1}{Z}\sum_{i=1}^{M} p(\boldsymbol{s}_i)=1$. An alternative method is to do normalization in the continuous case. When spherical harmonics are used to represent the ODF, i.e., $p(\boldsymbol{s})=\sum_{i=1}^{N} d_i B_i(\boldsymbol{s})$, then the ODF should be normalized such that $d_1=\frac{1}{\sqrt{4\pi}}$ in order for $\int_{\boldsymbol{s}\in\mathbb{S}^2}\boldsymbol{p}(\boldsymbol{s})d\boldsymbol{s} = 1$. If the estimation of the ODF is done with solid angle constraints as done in (Aganj et al., 2010; Goh et al., 2009a), then d_1 is automatically normalized to $\frac{1}{\sqrt{4\pi}}$. Cheng et al. (2009) then estimates the coefficients for the square-root samples $\boldsymbol{\psi}(\boldsymbol{s})=\sum_{i=1}^{N} c_i B_i(\boldsymbol{s})$, where $\boldsymbol{\psi}^2(\boldsymbol{s}) = \boldsymbol{p}(\boldsymbol{s})$. In this case, if M equally spaced samples are used, notice that

$$1=\int_{\mathbb{S}^2} p(\boldsymbol{s})d\boldsymbol{s}=\int_{\mathbb{S}^2}\sum_{i=1}^{N} c_i B_i(\boldsymbol{s})\sum_{j=1}^{N} c_j B_j(\boldsymbol{s})d\boldsymbol{s}$$
$$\sum_{i,j} c_i c_j \int_{\mathbb{S}^2} B_i(\boldsymbol{s})B_j(\boldsymbol{s})d\boldsymbol{s} \approx \sum_{i,j} c_i c_j \sum_{k=1}^{M} B_i(\boldsymbol{s}_k)B_j(\boldsymbol{s}_k)\Delta. \quad \text{(D.1)}$$

Since $\sum_{k=1}^{M} B_i(\boldsymbol{s}_k)B_j(\boldsymbol{s}_k)\Delta \approx \int_{\mathbb{S}^2} B_i(\boldsymbol{s})B_j(\boldsymbol{s})d\boldsymbol{s}=\delta_{i,j}$, and hence, $1 \approx \sum_{i=1}^{N} c_i^2$. Experimentally, $\sum_{i=1}^{N} c_i^2$ is always in [0.99, 1.01] if $d_1=\frac{1}{\sqrt{4\pi}}$.

Appendix E. Voxel Based Morphometry for Multivariate Testing

Let $\mathbf{x}_1,\ldots,\mathbf{x}_{n_x}$ and $\mathbf{y}_1,\ldots,\mathbf{y}_{n_y}$ be the (Euclidean) feature vectors drawn from two populations, $\bar{\mathbf{x}}=\frac{1}{n_x}\sum_{i=1}^{n_x}\mathbf{x}_i$ the sample mean of the first population, and $\bar{\mathbf{y}}=\frac{1}{n_y}\sum_{i=1}^{n_y}\mathbf{y}_i$ the sample mean of the second population. To test the hypothesis that $\bar{\mathbf{x}}=\bar{\mathbf{y}}$, a commonly used statistic is the Hotelling's T-squared statistic - a generalization of the Student's t-statistic commonly used in multivariate hypothesis testing. The Hotelling's T-square statistic is defined as

$$T^2=\frac{n_x n_y}{n_x+n_y}(\bar{\mathbf{x}}-\bar{\mathbf{y}})^\top \mathbf{W}^{-1}(\bar{\mathbf{x}}-\bar{\mathbf{y}})=\frac{n_x n_y}{n_x+n_y}\overrightarrow{\bar{\mathbf{y}}\bar{\mathbf{x}}}^\top \mathbf{W}^{-1}\overrightarrow{\bar{\mathbf{y}}\bar{\mathbf{x}}}, \tag{E.1}$$

where **W** is the unbiased pooled covariance matrix estimate defined as

$$\begin{aligned}\mathbf{W}&=\frac{\sum_{i=1}^{n_x}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})^\top+\sum_{i=1}^{n_y}(\mathbf{y}_i-\bar{\mathbf{y}})(\mathbf{y}_i-\bar{\mathbf{y}})^\top}{n_x+n_y-2}\\&=\frac{\sum_{i=1}^{n_x}\overrightarrow{\bar{\mathbf{x}}\mathbf{x}_i}\overrightarrow{\bar{\mathbf{x}}\mathbf{x}_i}^\top+\sum_{i=1}^{n_y}\overrightarrow{\bar{\mathbf{y}}\mathbf{y}_i}\overrightarrow{\bar{\mathbf{y}}\mathbf{y}_i}^\top}{n_x+n_y-2}.\end{aligned} \tag{E.2}$$

If $\mathbf{x}_1,\ldots,\mathbf{x}_{n_x}$ and $\mathbf{y}_1,\ldots,\mathbf{y}_{n_y}$ are drawn from two independent multivariate normal distributions with the same mean and covariance, one can show that the Hotelling's T-squared statistic is related to the F-distribution. As we do not assume that the observations come from two multivariate normal distributions, we use nonparametric permutation tests (Nichols and Holmes, 2002; Good, 2004) instead. As all voxels in the volume need to be considered simultaneously, we address the multiple comparisons and give corrected *p*-values based on two different methods, proposed in (Nichols and Holmes, 2002), that are commonly used for multiple comparisons correction in neuroimaging:

1. the single threshold method, to assess the significance of voxels whose statistical values exceed the critical threshold, whose value is dependent on the fixed significant level; and
2. a suprathreshold cluster test, to assess the significance of clusters of connected voxels that are all above a predetermined threshold.

The reader is referred to (Nichols and Holmes, 2002) for details of the two methods. Note that voxel-based analysis may be applied to any vector-valued data, using the nonparametric permutation tests.

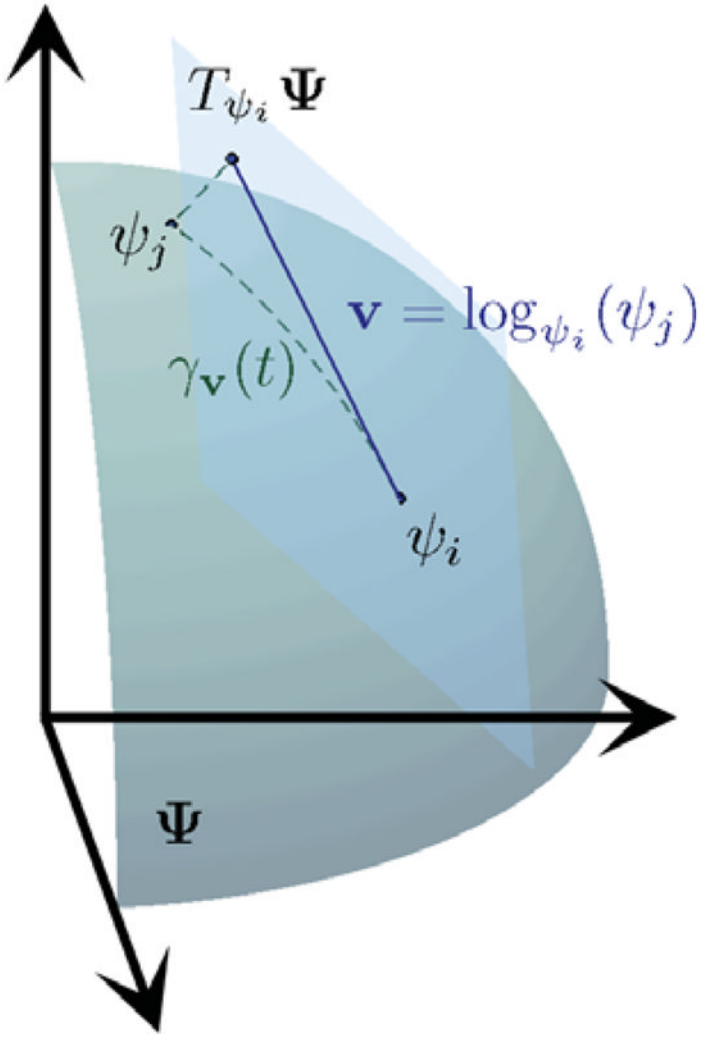
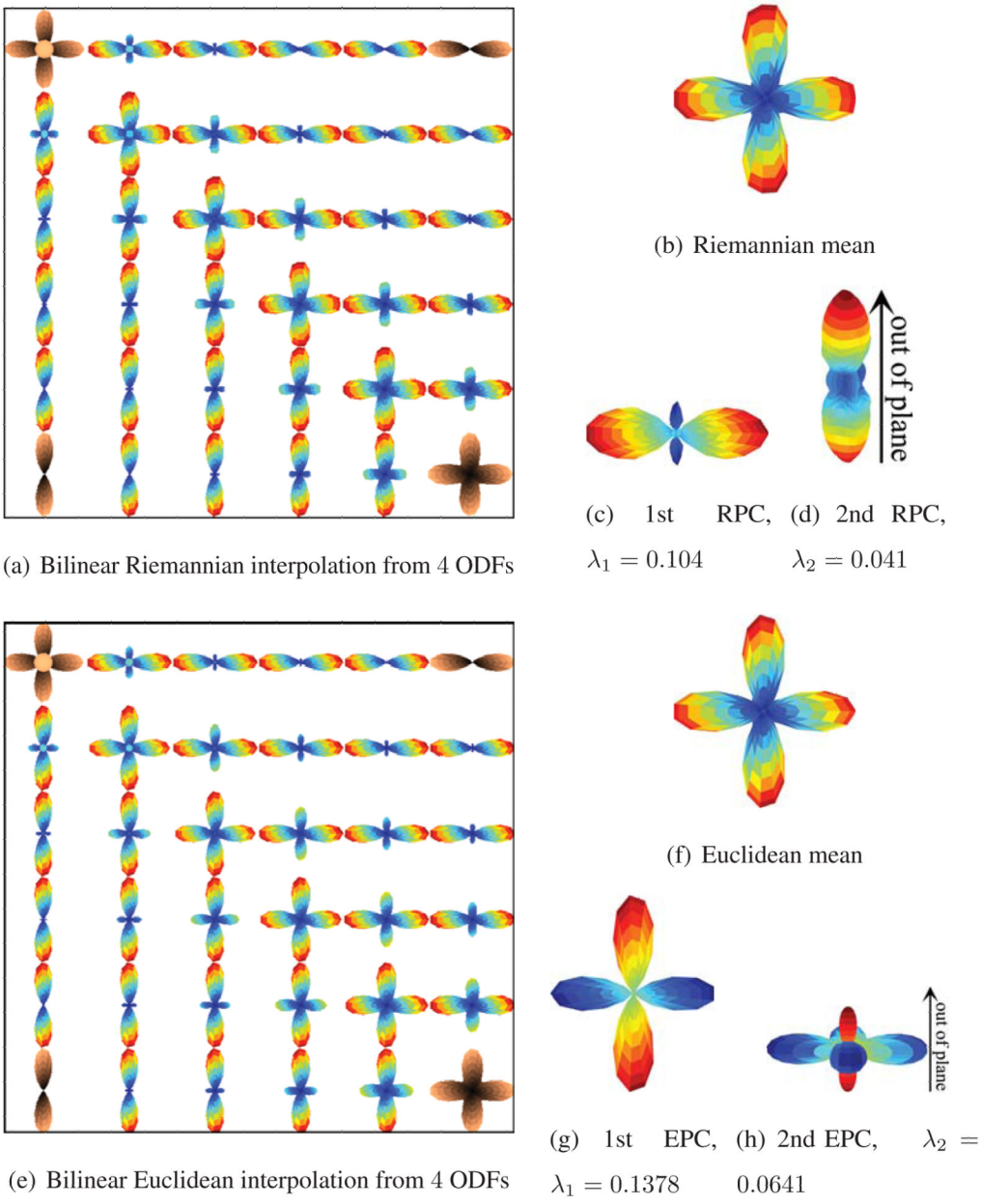


Figure 1.
Square root representation of ODFs. The tangent plane at ψ_i, together with the exponential and logarithm maps relating ψ_i and ψ_j, are also shown.



(a) Bilinear Riemannian interpolation from 4 ODFs

(b) Riemannian mean

(c) 1st RPC, $\lambda_1 = 0.104$

(d) 2nd RPC, $\lambda_2 = 0.041$

(e) Bilinear Euclidean interpolation from 4 ODFs

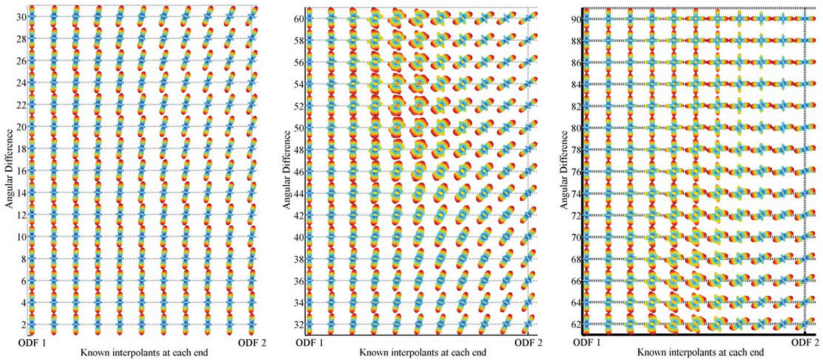
(f) Euclidean mean

(g) 1st EPC, $\lambda_1 = 0.1378$

(h) 2nd EPC, $\lambda_2 = 0.0641$

Figure 2.

Interpolation and statistical analysis of synthetic ODF fields. Fig. 2(a) shows the bilinear Riemannian interpolation of the four ODFs shaded in orange. Figs. 2(b)–2(d) show the Riemannian mean and the first two principal components. Fig. 2(e) shows the bilinear Euclidean interpolation of the four ODFs shaded in orange. Figs. 2(f)–2(h) show the Euclidean mean and the first two principal components.



(a) 2 to 30 degrees (b) 32 to 60 degrees (c) 62 to 90 degrees

Figure 3.

Interpolation between two single-fiber ODFs, whose main orientations are related to each other by a rotation of 2 to 90 degrees. In each row, the first and last ODFs are the known ODFs.

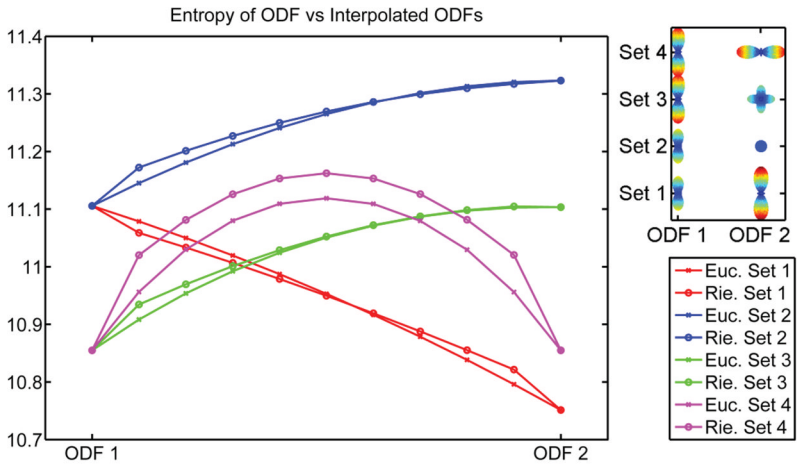


Figure 4.
Study of swelling effects for Riemannian and Euclidean interpolation, using the entropy measure. In each row, the first and last ODFs are known.

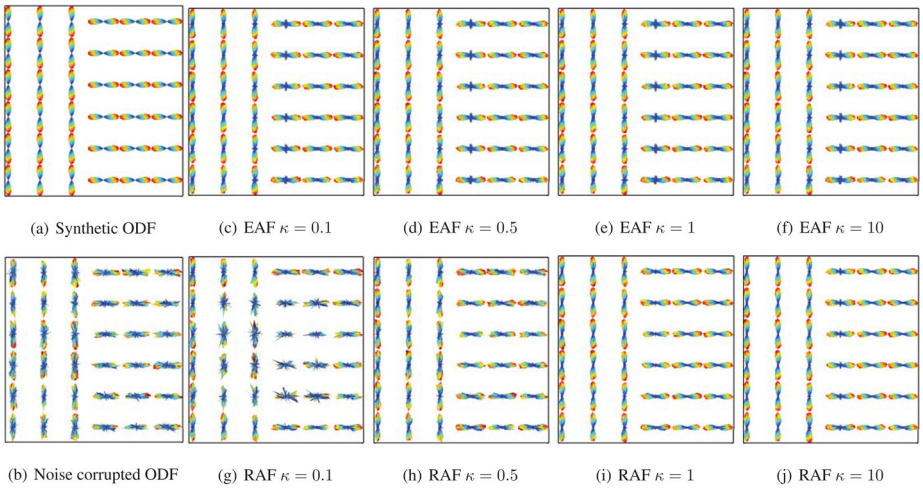


Figure 5.
Euclidean and Riemannian Anisotropic Filtering of synthetic ODF fields, with varying κ.

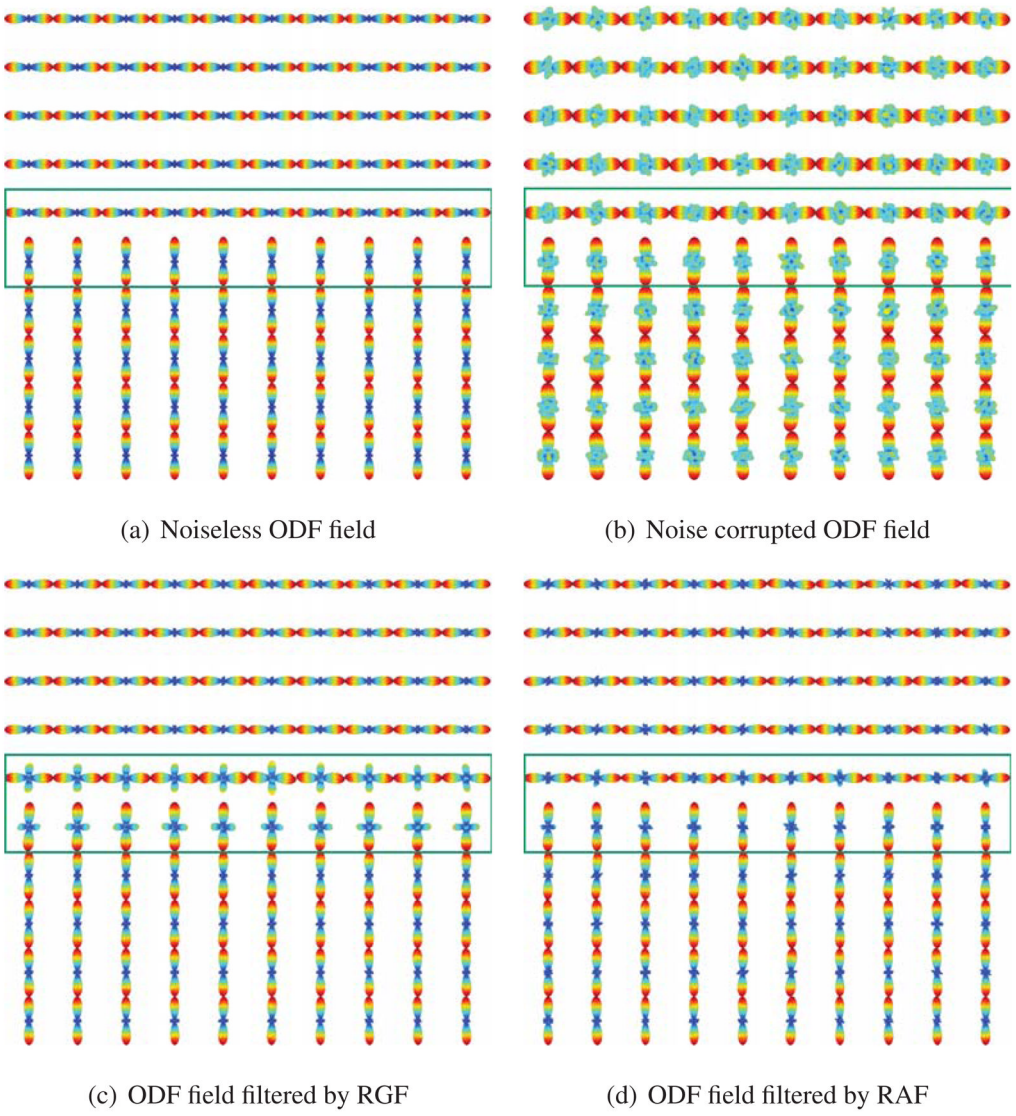
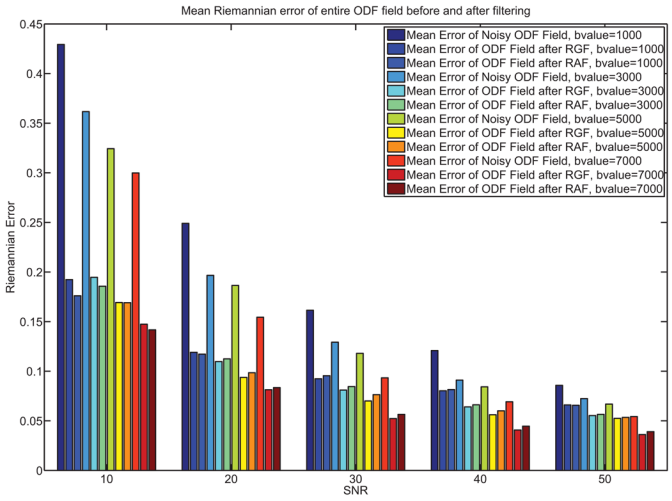
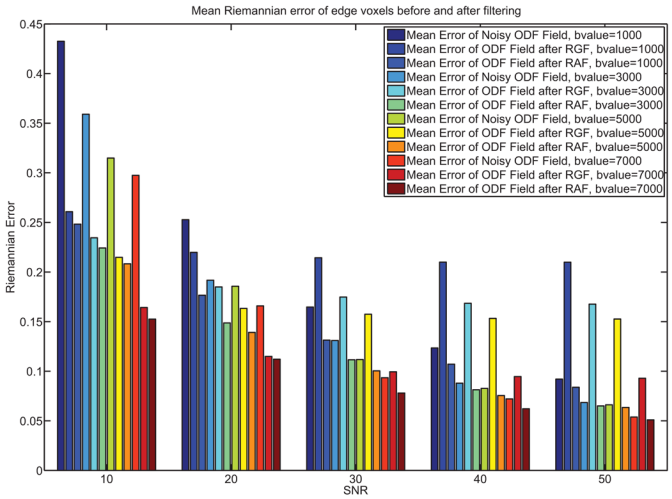


Figure 6.

Riemannian Gaussian filtering and anisotropic filtering of a synthetic ODF field at SNR 20 and *b*-value 3000 mm/s^2. Both RAF and RGF give ODF fields that are significantly closer to the true ODF. Notice also that RGF blurs the discontinuity.

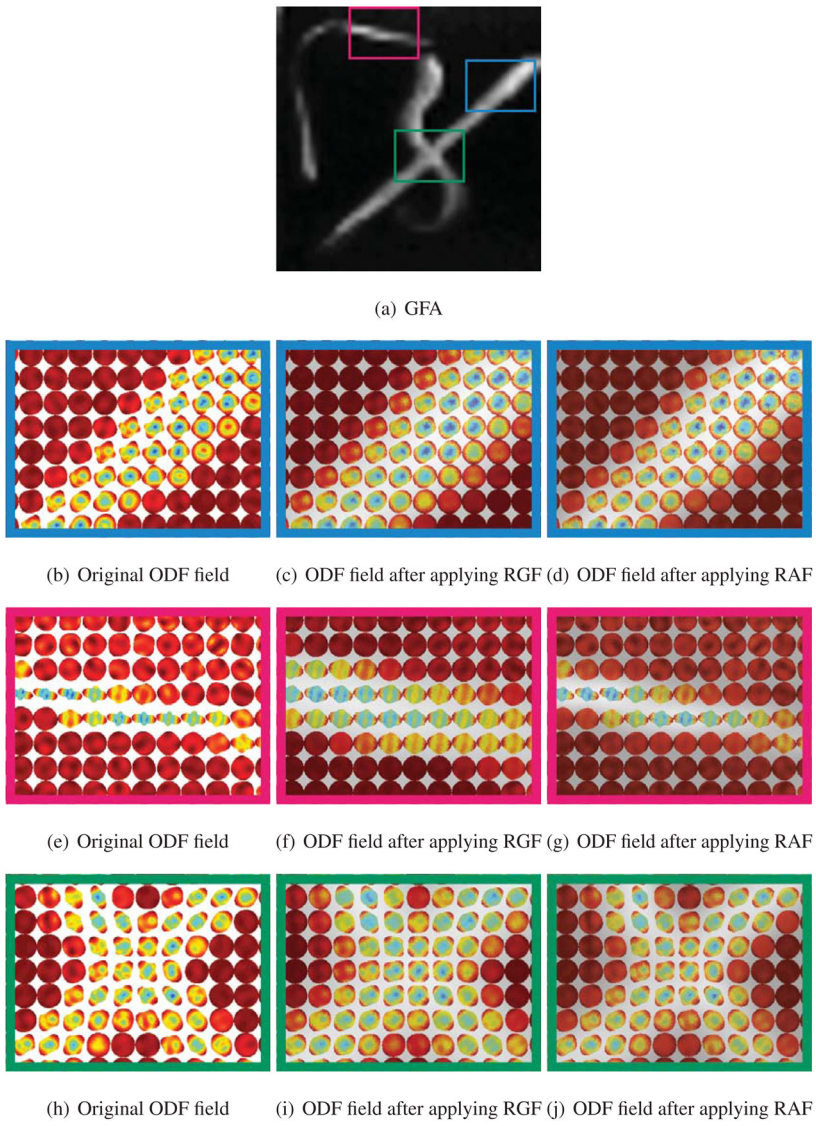


(a) Mean Riemannian error for the entire ODF field



(b) Mean Riemannian error for edge voxels, marked in green in Fig. 6

Figure 7.
Errors before and after filtering of the synthetic ODF field shown in Fig. 6.



(a) GFA

(b) Original ODF field (c) ODF field after applying RGF (d) ODF field after applying RAF

(e) Original ODF field (f) ODF field after applying RGF (g) ODF field after applying RAF

(h) Original ODF field (i) ODF field after applying RGF (j) ODF field after applying RAF

Figure 8.
McGill rat phantom data.

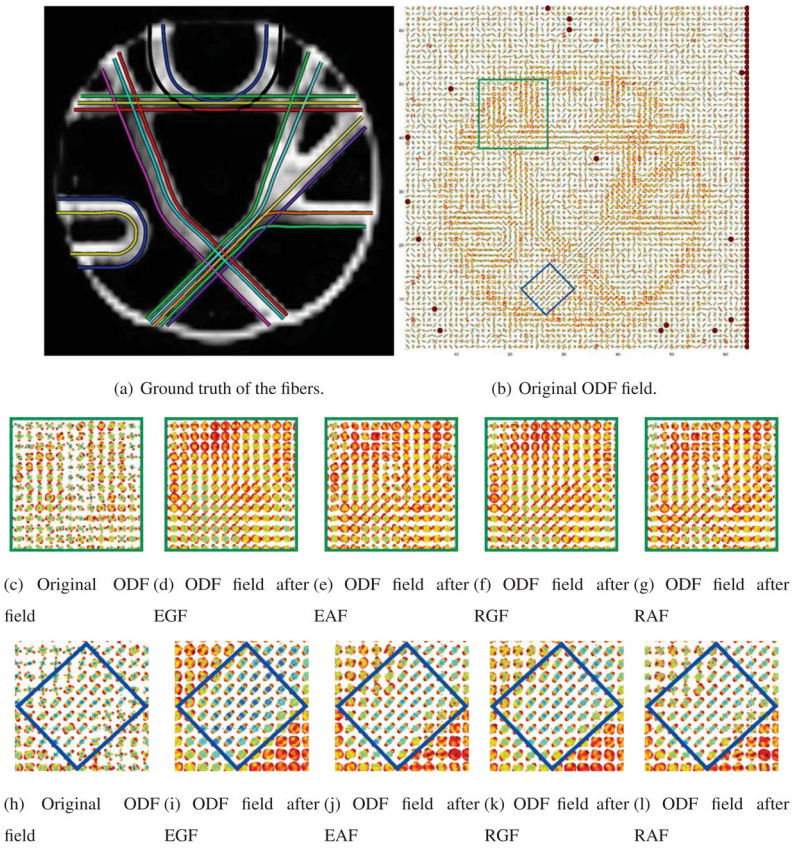


Figure 9.
Phantom data from Neurospin.

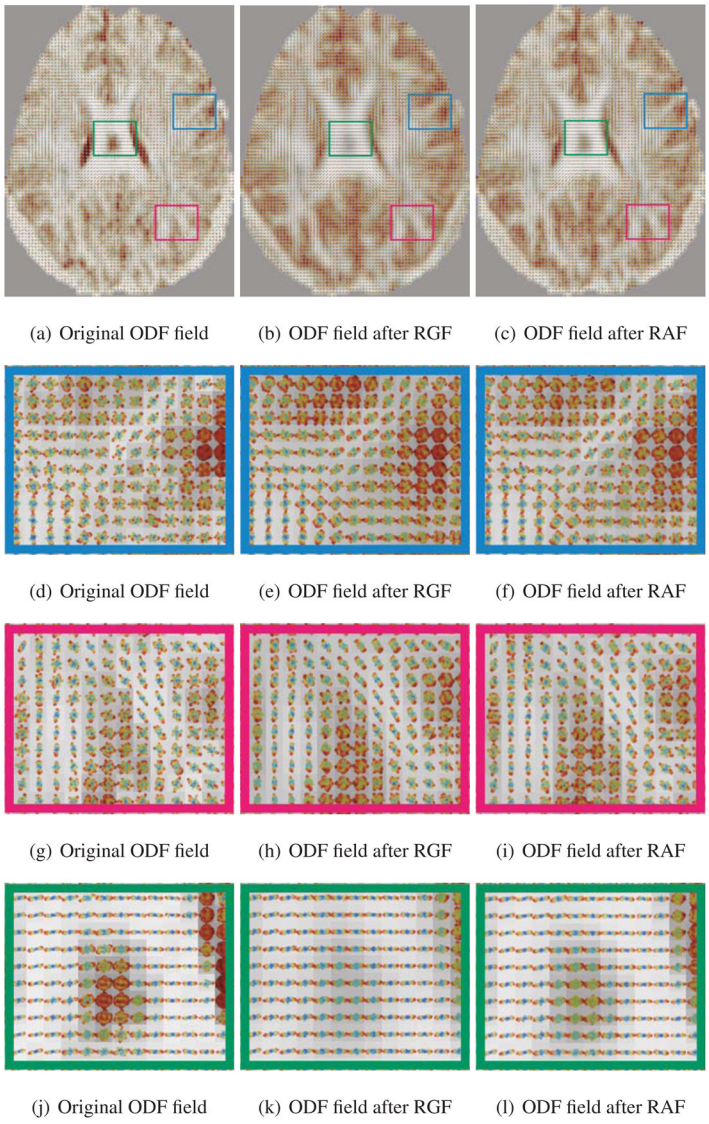


Figure 10.
ODF field of a real HARDI brain dataset. The ODFs are superimposed on top of the generalized fractional anisotropy maps.

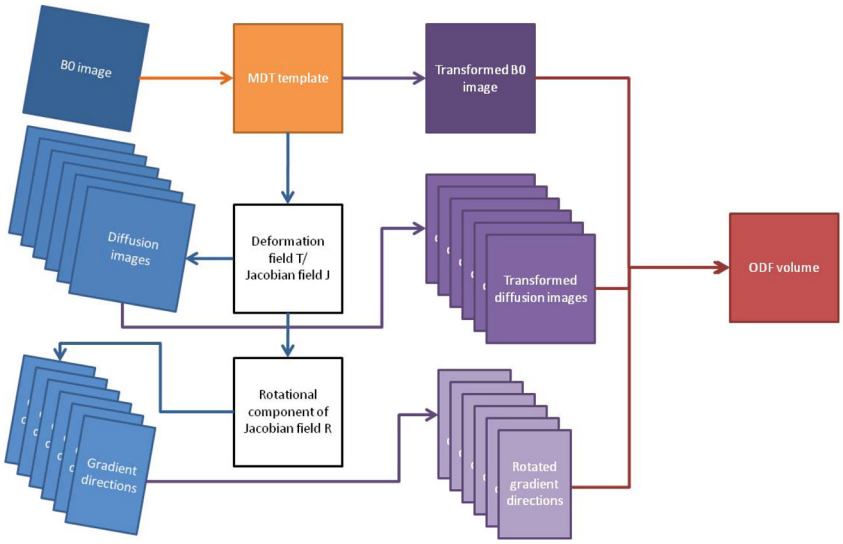
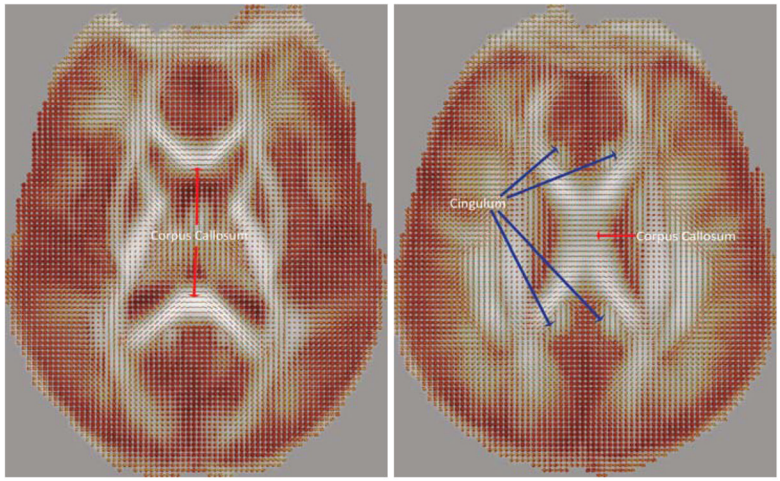
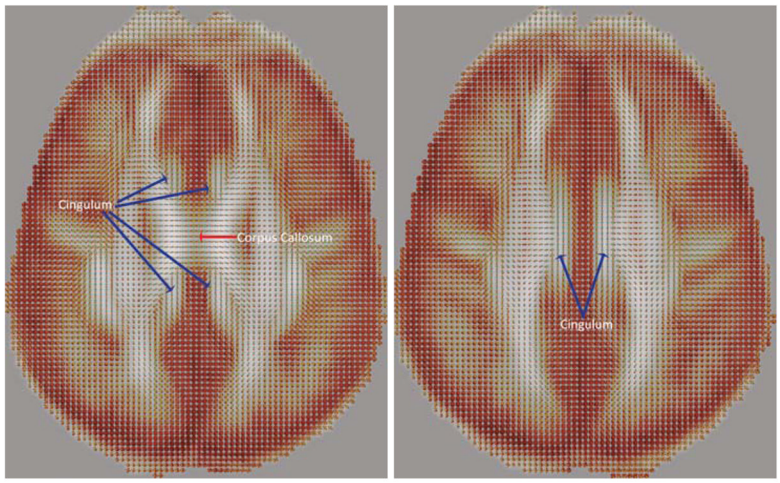


Figure 11.
Flowchart of steps used to reconstruct the ODF volume of each subject. Each ODF volume is registered to the minimal deformation template (MDT).



(a) Axial Slice 25 (b) Axial Slice 31



(c) Axial Slice 33 (d) Axial Slice 35

Figure 12.
Mean ODF field of a real HARDI brain dataset of 40 subjects. The ODFs are superimposed on top of the GFA maps.

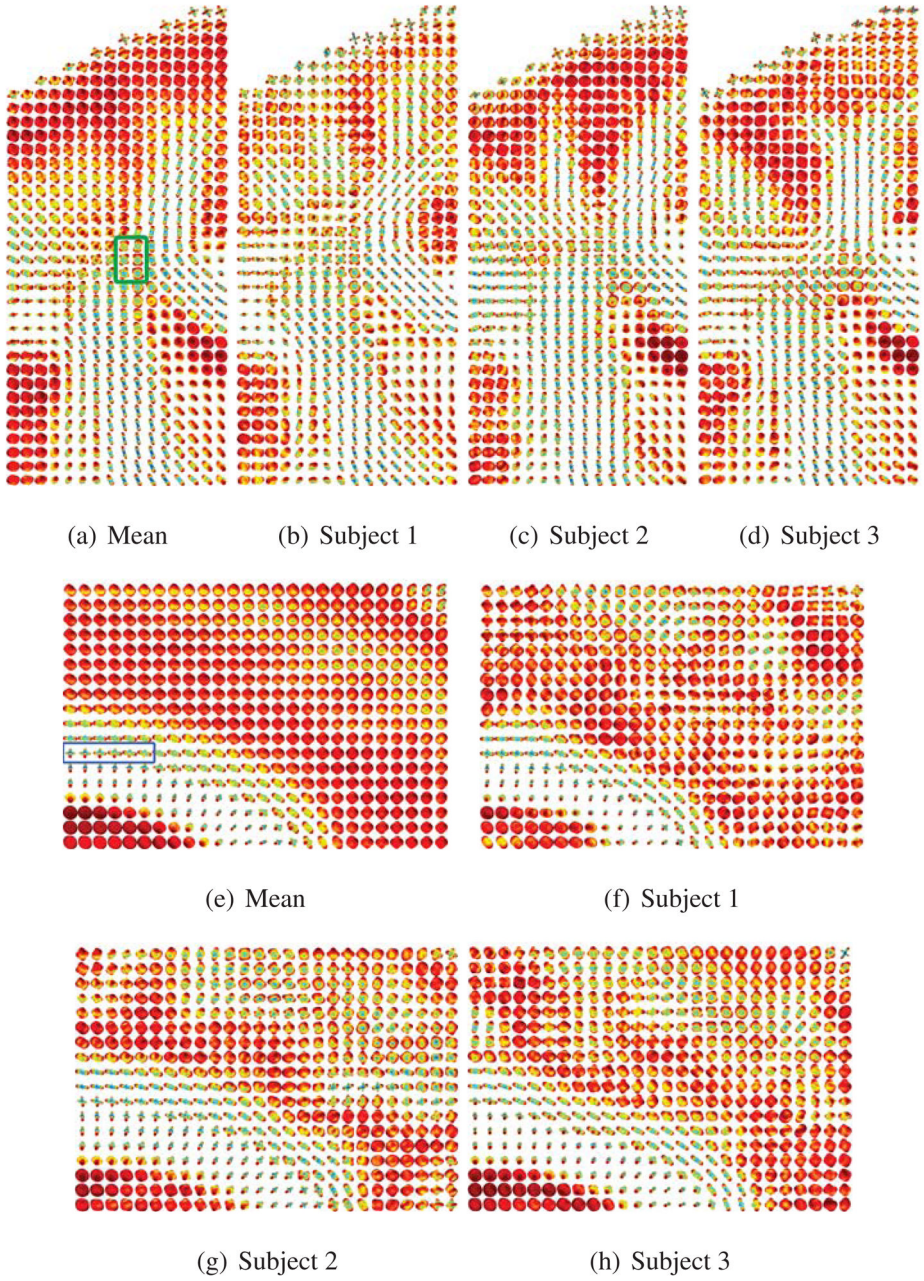
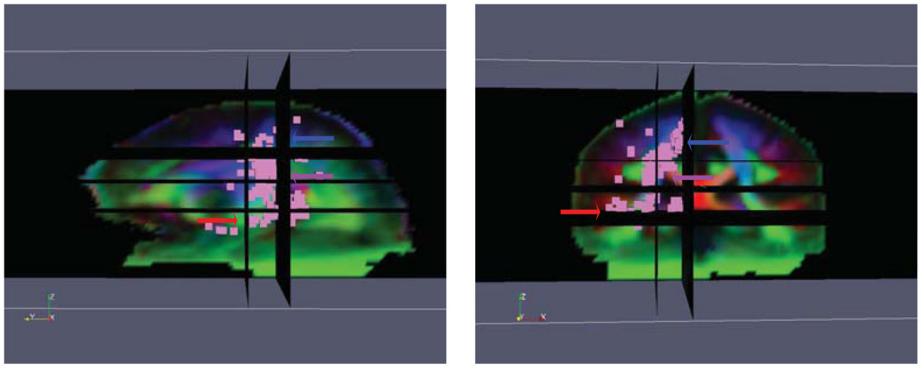
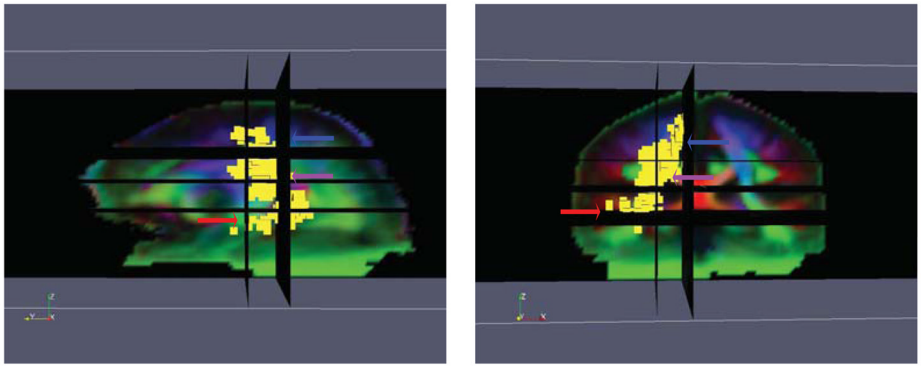


Figure 13.

Comparing the mean ODF to the ODF of individual subjects. Figs. 13(a)–13(d) correspond to the ROI of the cortical spinal tract and corpus callosum. The ROI in green corresponds to the two fiber tracts crossing. Figs. 13(e)–13(h) correspond to the ROI of the cingulum and the corpus callosum. The ROI in blue corresponds to the voxels exhibiting partial volume averaging of the two fiber tracts.



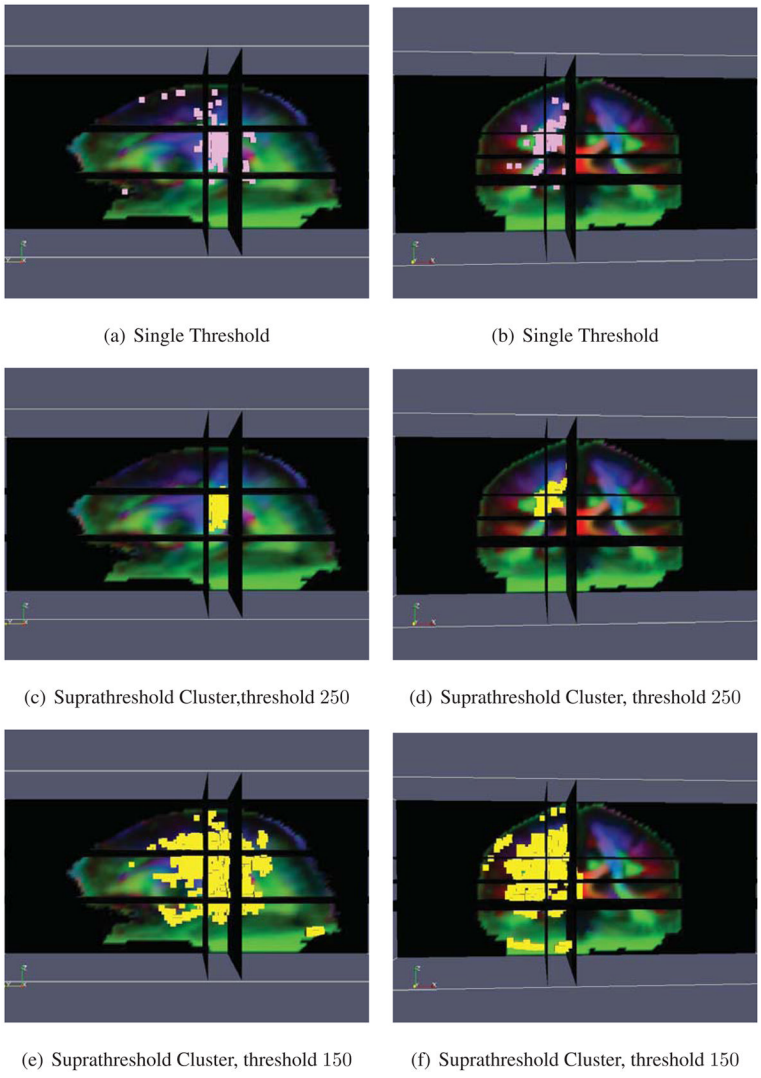
(a) Single Threshold (b) Single Threshold



(c) Suprathreshold Cluster, threshold 250 (d) Suprathreshold Cluster, threshold 250

Figure 14.

Results of performing the ODF Riemannian statistical test on brain asymmetry. Figs. 14(a) and 14(c) show the sagittal view and Figs. 14(b) and 14(d) the coronal view. The pink voxels in Figs. 14(a) and 14(b) are deemed significant by the single threshold test and the yellow voxels in Figs. 14(c) and 14(d) are deemed significant by the suprathreshold cluster test. The red arrow shows the planum temporale, the purple arrow the internal capsule, and the blue arrow the corticospinal tract.



(a) Single Threshold (b) Single Threshold

(c) Suprathreshold Cluster,threshold 250 (d) Suprathreshold Cluster, threshold 250

(e) Suprathreshold Cluster, threshold 150 (f) Suprathreshold Cluster, threshold 150

Figure 15.

Results of performing the ODF Euclidean statistical test on brain asymmetry. Figs. 15(a), 15(c) and 15(e) show the sagittal view and Figs. 15(b), 15(d) and 15(f) the coronal view. The pink voxels in Figs. 15(a) and 15(b) are deemed significant by the single threshold test and the yellow voxels in Figs. 15(c)–15(f) are deemed significant by the suprathreshold cluster test.

Table 1

Mean of (Amount of error after RAF)/(Amount of error after EAF) in 100 trials under two different measures.

κ	0.1	0.5	1	10	100
Euclidean measure	1.197	0.767	0.480	0.468	0.468
Riemannian measure	1.108	0.945	0.801	0.787	0.787

Algorithm 1

Weighted Karcher Mean

Given data points $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_n \in \boldsymbol{\Psi}$, weights $w_1, \ldots, w_n$, a predefined threshold ε, and a maximum number of iterations T,

1 Initialize $t = 1$, $\bar{\boldsymbol{\psi}}_1 = \boldsymbol{x}_i$ for a random i and $\boldsymbol{\varphi} = \mathbf{0}$.

2 While $t \leq T$ or $\|\boldsymbol{\varphi}\|_{\bar{\boldsymbol{\psi}}} \geq \varepsilon$,

a. Compute tangent vector $\boldsymbol{\varphi} = \sum_{i=1}^{n} w_i \log_{\bar{\boldsymbol{\psi}}_t}(\boldsymbol{\psi}_i)$,

b. Set $\bar{\boldsymbol{\psi}}_{t+1} = \exp_{\bar{\boldsymbol{\psi}}_t}(\lambda \boldsymbol{\varphi})$, where λ is the learning rate. Buss and Fillmore(2001) proved that $\lambda = 1$ is sufficient for convergence.

Algorithm 2

Principal Geodesic Analysis

Given data points $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_n \in \boldsymbol{\Psi}$,

1. Compute intrinsic mean $\bar{\boldsymbol{\psi}}$ as in Algorithm 1.
2. Calculate the tangent vectors $\boldsymbol{\varphi}_i = \log_{\bar{\psi}}(\boldsymbol{\psi}_i)$ about $\bar{\boldsymbol{\psi}}$.
3. Construct the sample covariance matrix $\mathrm{cov}(\boldsymbol{\psi}) = \frac{1}{n-1}\sum_{i=1}^{n} \boldsymbol{\varphi}_i \boldsymbol{\varphi}_i^{\top}$.
4. Perform eigenanalysis of the matrix cov($\boldsymbol{\psi}$), with the eigenvectors $\{\mathbf{u}_i\}_{i=1}^{d}$ giving the principal directions. $\{\mathbf{u}_i\}_{i=1}^{d}$ forms an orthonormal basis for $T_{\bar{\psi}}\boldsymbol{\Psi}$.