# High-Dimensional Spatial Normalization of Diffusion Tensor Images Improves the Detection of White Matter Differences: An Example Study Using Amyotrophic Lateral Sclerosis

Hui Zhang, Brian B Avants, Paul A Yushkevich, John H Woo, Sumei Wang, Leo F McCluskey, Lauren B Elman, Elias R Melhem, James C Gee

Abstract-Spatial normalization of diffusion tensor images plays a key role in voxel-based analysis of white matter (WM) group differences. Currently, it has been achieved using lowdimensional registration methods in the large majority of clinical studies. This paper aims to motivate the use of high-dimensional normalization approaches by generating evidence of their impact on the findings of such studies. Using an ongoing amyotrophic lateral sclerosis (ALS) study, we evaluated three normalization methods representing the current range of available approaches: low-dimensional normalization using the fractional anisotropy (FA), high-dimensional normalization using the FA and highdimensional normalization using full tensor information. Each method was assessed in terms of its ability to detect significant differences between ALS patients and controls. Our findings suggest that inadequate normalization with low-dimensional approaches can result in insufficient removal of shape differences which in turn can confound FA differences in a complex manner, and that utilizing high-dimensional normalization can both significantly minimize the confounding effect of shape differences to FA differences and provide a more complete description of WM differences in terms of both size and tissue architecture differences. We also found that high-dimensional approaches, by leveraging full tensor features instead of tensor-derived indices, can further improve the alignment of WM tracts.

Index Terms-diffusion tensor images, spatial normalization

## I. INTRODUCTION

Diffusion tensor MRI is a unique imaging technique that probes microscopic tissue properties by measuring local diffusion of water molecules [1]. Its demonstrated ability to depict *in vivo* the intricate architecture of white matter (WM) [2] has made it an invaluable tool for furthering our understanding of WM both in normal populations and in populations with brain disorders. Diffusion tensor imaging has been applied to study the variations in normal brain neuronal circuitry [3] and its left-right asymmetry [4]–[7], and to follow its changes over the lifespan both through its maturation and aging [8]– [11]. The disorders to which diffusion tensor imaging has been applied range from various WM diseases [12], such as multiple sclerosis [13], [14], amyotrophic lateral sclerosis (ALS) [15]– [17] and Krabbe disease [18], to diseases characterized by cognitive deficits or behavorial disorders with suspected WM involvement [19], such as schizophrenia [20], [21], Turner syndrome [22] and chromosome 22q11.2 deletion syndrome [23], and to brain injuries including both ischemic and traumatic types [24], [25].

1

To anatomically localize WM differences across populations using diffusion tensor imaging, the approach most commonly taken is voxel-based whole brain analysis [3], [16], [20], [22], [23], [26], [27]. This approach works by first spatially normalizing all subjects, *i.e.*, placing them all in correspondence with one another, thereby removing inter-subject variations in shape or shape confounds. Any remaining differences between subjects then correspond to differences in appearance features, properties of diffusion in this case. They are subsequently evaluated and correlated with other factors on a voxel-by-voxel basis.

The key element of voxel-based analysis is spatial normalization of the diffusion tensor images. The quality of spatial normalization determines the extent to which the shared anatomy, in this case WM tracts, are aligned. Therefore, it has direct impact on the successful removal of shape confounds and consequently on the validity, specificity and sensitivity of the subsequent statistical inferences of group differences. Currently, the large majority of clinical studies have chosen to employ the spatial normalization approach of aligning the diffusion tensor images using low-dimensional image registration algorithms via their corresponding structural images, i.e., T1- or T2-weighted images [3], [16], [20], [22], [23], [26], [27], or via their fractional anisotropy (FA) images [28]. This normalization strategy allows researchers to take advantage of available registration tools, with the nonlinear normalization algorithm within SPM2 [29] being the most commonly used. However, the limited spatial normalization quality of lowdimensional approaches has made the interpretation of their findings challenging, as discussed in [30]. The registration algorithms underlying the low-dimensional appraches use lowdimensional representations of spatial transformation that can not adequately model the complex brain morphological differences often seen across subjects. The high-dimensional normalization approaches improve the quality of normalization by employing registration algorithms that employ high-

All authors are at the University of Pennsylvania. L F McCluskey and L B Elman are in the Department of Neurology. The other authors are from the Department of Radiology, and H Zhang, B B Avants, P A Yushkevich and J C Gee belong to the Penn Image Computing and Science Laboratory. Copyright (c) 2007 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

dimensional representations of spatial transformation. Spatial normalization strategies based on high-dimensional registration methods have so far found few clinical applications in WM studies, with the Park et al. analysis of WM asymmetry [31] being one notable exception. The reluctance to adopt high-dimensional registration can be attributed to these methods being less widely available and more resource-intensive than their low-dimensional counterparts. Furthermore, in the context of WM studies, the impact of improved normalization on the statistical findings has not yet been clearly demonstrated.

The aim of our paper is to motivate the use of highdimensional normalization approaches in WM studies by generating evidence of their impact on the outcomes of multisubject analysis. A secondary aim is to examine the utility of using full tensor information in comparison with tensorderived indices, specifically, the FA. Our evaluation was based on the general task-driven approach, which has been successfully used in the literature to determine the impact of novel normalization techniques on the sensitivity of population-level fMRI analysis ([32]-[34]). For our purposes, we chose to use an ongoing ALS study as our example. ALS is a progressive neurodegenerative disorder that affects both upper and lower motor neurons. Using diffusion tensor imaging, a reduction of FA in the primary motor pathway affected by upper neuron degeneration has been shown consistently by both region-ofinterest [15], [17] and voxel-based analyses [16], [28], making ALS the ideal candidate for this evaluation. Using the diffusion tensor imaging data from the ALS study, we evaluated three normalization methods that represent the current range of available approaches: tensor-based high-dimensional registration, high-dimensional registration of FA images, and lowdimensional FA registration. The performance of each method was assessed in terms of its ability to detect statistically significant differences between ALS patients and controls. We found that WM differences were influenced by shape confounds in a complex manner. Only when the statistical findings were analyzed in conjunction with an evaluation of the quality of spatial normalization, could their interaction be fully understood. Our findings suggest that high-dimensional normalization approaches should play an important role in WM population studies and that they can realize their full potential when full tensor features are used to drive the alignment of WM tracts.

The rest of the paper is organized as follows. In Sec. II, we review the basics of diffusion tensor imaging and related work in diffusion tensor image analysis. Sec. III briefly describes each method in comparison and gives details of the study design, which is followed by a detailed description of our evaluation strategy in Sec. IV. The results of our evaluation are presented in Sec. V and then discussed in Sec. VI.

## II. BACKGROUND

Diffusion tensor MRI leverages molecular diffusion, water diffusion in particular, that can reveal certain microscopic features of the underlying tissue, particularly the presence of fibrous structures. For instance, in WM, which consists of packed axon fibers that constitute neuronal pathways, diffusion appears anisotropic due to restricted movement of water perpendicular to the axon fibers; the direction along which water diffuses most freely coincides with the orientation of the fibers [35]. Diffusion tensor MRI assumes a Gaussian diffusion model for water molecules and captures their diffusion properties in each voxel of a MRI volume by determining an apparent diffusion tensor, a symmetric and positive-definite (SPD) 3-by-3 matrix. The diffusion tensor provides estimates of the mean diffusivity, anisotropy and dominant orientation of diffusion. Within WM of normal brain, the mean diffusivity has been found to be close to constant, whereas the diffusion anisotropy and the dominant orientation of diffusion have been found to vary greatly [2]. The diffusion anisotropy measures the variability of the diffusion along different spatial directions. Although its variation can not be accounted for by any single microstructural factor or a combination of them, the diffusion anisotropy appears to be highly sensitive to differences in WM architecture [2]. The most commonly used diffusion anisotropy measure is fractional anisotropy (FA), which is defined as the normalized standard deviation of the eigenvalues of a diffusion tensor and ranges from 0 to 1: 0 for isotropic diffusion and 1 for perfectly anisotropic diffusion [36]. The dominant direction of diffusion, on the other hand, captures the orientation of the underlying WM fiber bundles. Its variability is a reflection of the complexity of our neuronal network. By following the dominant directions of diffusion within WM using a technique known as fiber tractography (See [37] for a review), major WM tracts have been able to be successfully localized [38]-[40].

The rich description of diffusion properties using the diffusion tensors suggests that optimal spatial normalization can be achieved by utilizing registration methods that leverage full tensor features. However, there are unique challenges for the development of such methods, as discussed in [41]. The most challenging aspect of all is arguably the procedure known as tensor reorientation: the need to correct tensor orientation during image warping such that it remains consistent with the underlying WM organization [42]. Earlier diffusion tensor image registration techniques circumvent tensor reorientation by registering scalar images derived from diffusion tensor images [43], [44], thus discarding the orientation component of the data. Some other methods register actual tensor images but not reorienting the tensors during registration [45], [46], thus introducing inaccuracies in image matching. Later, Park et al [47] showed that using diffusion tensors as a whole improved the quality of registration by better matching the diffusion tensors orientation information; but their method only applied tensor reorientation iteratively and tensor reorientation was not explicitly optimized. Curran et al [48] then demonstrated that explicitly optimizing tensor reorientation during affine registration of synthetic images improved image matching. However, their method [49] is not derivative-based and their registration optimization tends to have difficulties with local minima. We have proposed an affine registration algorithm that both explicitly optimizes tensor reorientation and has a novel derivative-based formulation in [50], [51]. Our synthetic examples show that our derivative-based method is faster and reaches global minima more consistently than the

method not using derivatives. More recently, high-dimensional methods have been developed that optimize tensor orientation explicitly. Cao *et al* [52] developed a large deformation diffeomorphic registration algorithm for vector fields. The algorithm was applied to register diffusion tensor images by matching their corresponding principal eigenvectors. We have proposed a dense piecewise affine deformable algorithm for diffusion tensor images that explicitly optimizes tensor orientation [53]. Most recently, in a preliminary work, Cao *et al* have extended their large deformation diffeomorphic registration algorithm for vector fields to diffusion tensor images [54].

Besides voxel-based analysis, the other commonly used approach is region-of-interest (ROI) based analysis. The ROIbased analysis is often appropriate when study hypotheses relate to specific regions [6], [12], [15], [17]. However, this approach is often labor intensive and its effectiveness is limited by the consistency of the ROI delineations across subjects. The development of fiber tractography algorithms, which offer a semi-automatic and anatomy-based means of segmenting major WM tracts, has led to an improved ROI-based analysis strategy in which the ROIs are derived using tractography such that the subsequent analysis is confined to the fiber bundles [55]-[59]. Most recently, tract-specific analysis has been extended to allow corresponding locations along a fiber tract to be evaluated across subjects [60]-[62]. Despite its numerous advantages, this approach should be taken with caution because tractography algorithms have not yet been thoroughly validated. There has been evidence that suggests tractography underestimates the size of fiber bundle at least in pathological condition [63]. Very recently, a hybrid approach known as tract-based spatial statistics (TBSS) was proposed, which attempts to combine the best of voxel-based and tractbased analyses [30].

#### **III. MATERIALS AND METHODS**

## A. Subjects and Data Acquisition

The subjects used in this evaluation study were recruited from the community served by the University of Pennsylvania Health System (UPHS) as part of an ongoing clinical investigation into WM changes in ALS using magnetic resonance imaging. Out of a total of 29 subjects scanned, only 16 were acquired with the same diffusion imaging protocol and they were chosen for the present study. Among them were 8 ALS patients (age 42-77, mean age and standard deviation  $60\pm11$ ; 6 male, 2 female) and 8 healthy controls (age 40-56, mean age and standard deviation  $46\pm6$ ; 6 male, 2 female). All subjects provided informed consent, following procedures approved by the local Institutional Review Board of the UPHS. Diffusion tensor imaging was performed using a single-shot, spin-echo, diffusion-weighted echo-planar imaging (EPI) sequence on a 3.0-T Siemens Trio scanner (Siemens Medical Solutions, Erlangen, Germany). The diffusion sampling scheme consisted of one image without diffusion gradients ( $b = 0 \text{ s/mm}^2$ ), followed by 12 images measured with 12 non-collinear and noncoplanar diffusion encoding directions isotropically distributed in space ( $b = 1000 \text{ s/mm}^2$ ). Additional imaging parameters for the diffusion-weighted sequence were: TR = 6500 ms, TE =



Fig. 1. FA map of the population-specific tensor template used in the spatial normalization.

99 ms, 90° flip angle, number of averages = 6, matrix size =  $128 \times 128$ , slice thickness = 3.0 mm, spacing between slices = 3.0 mm, 40 axial slices with in-plane resolution of  $1.72 \times 1.72$  mm, resulting in voxel dimensions equal to  $1.72 \times 1.72 \times 3.0$  mm<sup>3</sup>.

## B. Image Preprocessing

The diffusion-weighted images were first corrected for motion and eddy-current artifacts according to the method reported in [64], prior to extracting brain parenchyma with the Brain Extraction Tool [65]. The diffusion tensor images were then reconstructed from the diffusion-weighted images using the standard linear regression approach [1]. Finally, the resulting tensor volumes were resampled to a voxel space of  $128 \times 128 \times 64$  with voxel dimensions equal to  $1.72 \times 1.72 \times$  $2.5 \text{ mm}^3$ . The resampled volume, with axial dimension equal to a power of 2, is better suited for registration algorithms that require the construction of standard multi-resolution image pyramids.

## C. Population-Specific Tensor and FA Templates

A population-specific tensor template was constructed from all 16 subjects using an iterative strategy similar to the one described in [66]. An initial template was computed as an average of the original subject diffusion tensor images. The template was then iteratively refined by repeating the following procedure: register the subjects to the template, then compute a refined template for the next iteration as an average of the normalized images. This procedure was repeated until the change between templates from consecutive iterations became sufficiently small. During each iteration, the diffusion tensor images were registered to the respective template estimate using the tensor registration algorithm described later in the section. The FA template was taken as the FA map derived from the tensor template and it is illustrated in Fig. 1.

## D. Spatial Normalization

The three spatial normalization methods we chose to evaluate are the SPM2 normalization via FA images representing low-dimensional approaches, the diffeomorphic normalization via FA images representing high-dimensional approaches using tensor-derived indices, and finally the tensor normalization via diffusion tensor images themselves representing highdimensional approaches using full tensor features.

1) Initial Alignment: The diffusion tensor images were first affinely aligned to the template. The tensor images after the affine alignment were provided as the input to the tensor registration algorithm while their respective FA maps were given as the input to the SPM2 and diffeomorphic FA image registration algorithms.

2) SPM2 Normalization: We chose the default parameters for the SPM2's non-linear registration algorithm which minimizes the residual sum of squared intensity differences with a 12-parameter affine transformation and a non-linear transformation comprising a linear combination of  $7 \times 8 \times 7$ smooth spatial basis functions [67].

*3) Diffeomorphic Normalization:* We chose to optimize a cross-correlation metric under the constraints of a diffeomorphic transformation model in multi-resolution and symmetric fashion [68].

4) Tensor Normalization: The diffusion tensor image registration algorithm used here is an extension of the deformable diffusion tensor image registration method recently proposed in [53]. The algorithm leverages full tensor-based similarity metrics while optimizing tensor orientation explicitly. It approximates smooth transformations using a dense piecewise affine parametrization which is sufficient when the required deformations are not large. The current extension handles larger deformations by iteratively composing smaller incremental deformations estimated using the algorithm in [53]. We used the tensor metric that measures the  $L^2$  distance between the anisotropic part of the apparent diffusion profiles associated with the diffusion tensors as described in [53]. Under this metric, the distance between two diffusion tensors  $D_1$  and  $D_2$  is equal to

$$\sqrt{rac{8\pi}{15}(\|\mathbf{D}_1-\mathbf{D}_2\|_C^2-rac{1}{3} ext{Tr}^2(\mathbf{D}_1-\mathbf{D}_2))}$$
 ,

where  $\|\mathbf{D}_1 - \mathbf{D}_2\|_C$  is the Euclidean distance between the two tensors and equal to  $\sqrt{\text{Tr}((\mathbf{D}_1 - \mathbf{D}_2)^2)}$ .

## E. Statistical Non-Parametric Mapping of WM Differences Between the ALS and Control Groups

The statistical non-parametric maps of WM differences were derived using the SnPM toolbox (freely available at http://www.sph.umich.edu/ni-stat/SnPM/) [69]. The SnPM toolbox uses the general linear model [70] to construct t-statistics images that are subsequently assessed for statistical significance using permutation testing, a standard non-parametric procedure for multiple comparison correction. Specifically for the present study, the appropriate t-statistics images were produced by statistically comparing the voxelwise data values of the ALS and control populations using two-sample t tests. We carried out the voxelwise t tests on the brain volumes that were spatially normalized to the template following the procedures outlined in Sec. III-D. Hence the

resulting t-statistics images were in the same space as the template brain volume. Furthermore, we restricted the t tests to the WM region, which we defined as the set of voxels with FA values greater or equal to 0.2 in the FA template according to [30]. The total number of such voxels is 48,768. Next, the t-statistics maps were assessed for statistical significance using suprathreshold cluster tests. Suprathreshold cluster tests threshold a statistic image at a predetermined primary threshold and assess the size of clusters, connected suprathreshold regions, for significance. Here we chose the primary threshold to be 3.72 corresponding to an uncorrected p-value of 0.001. The clusters with Family-Wise Error (FWE) corrected p-values less than 0.1 are reported. Among them, the ones with FWE corrected p-values less than 0.05 are considered to be highly significant. The FWE corrected pvalues were determined using permutation tests. Given the number of possible permutations for the current data set is only 12,870, the full permutation test is feasible and was used. This procedure of deriving statistical non-parametric maps was applied to identify both local volumetric and FA differences, and was repeated for each of the three spatial normalization methods. The respective statistical maps were subsequently compared to form a basis for comparison.

Local volumetric differences were derived following the standard approach of tensor-based morphometry [71], [72]. For each subject, a Jacobian determinant field was computed from the spatial transformation obtained from normalizing its brain volume to the template at the spatial normalization stage. Statistical testing was then performed over the natural logarithm of the relevant Jacobian determinant fields.

The normalized FA maps for statistical non-parametric mapping of FA differences were derived by warping the input FA maps using the spatial transformations estimated from spatial normalization. By generating the normalized FA maps in an identical fashion for all the normalization methods, we can attribute any differences in the normalized FA maps solely to the differences in the normalization methods themselves.

## IV. EVALUATION

The premise of our evaluation strategy is that the accuracy of spatial normalization directly affects the success in removing shape confounds which in turn influences the ability to detect WM differences. Specifically, each method was first assessed for its spatial normalization accuracy, which was measured in terms of two voxelwise statistics indicative of the extent of WM alignment. The ability of each method to detect local volumetric differences was analyzed next, which served to provide a direct measure of the extent to which shape confounds are removed, as well as to be an integral part of understanding group differences of WM, relative size differences of WM tracts in this case. Finally, each method's performance in detecting FA differences was determined. To evaluate the influence of shape confounds on the FA findings, we analyzed specifically the normalization quality of the voxels belonging to the clusters with significant FA differences using the two voxelwise statistics derived at the outset. In the following, each component of our evaluation is described in order.

## A. Evaluating the Overall Spatial Normalization Performance

Since diffusion anisotropy and the dominant direction of diffusion are two features that account for most of the variations in WM structures [2], misalignment that renders different WM structures being mapped to one another should yield large voxelwise variations in either one or both of the features. Therefore, we chose normalized FA standard deviation  $\bar{\sigma}_{FA}$ and *dyadic coherence*  $\kappa$ , as the two voxelwise statistics to gauge normalization quality. Given a set of diffusion tensors,  $\bar{\sigma}_{FA}$ , which is defined as the ratio of the standard deviation and mean of the FA values of these diffusion tensors, measures the variability in diffusion anisotropy, while  $\kappa$  [44], which takes values that range from 0 to 1 (0 for randomly oriented directions and 1 for identically oriented directions), captures the variability in the dominant direction of diffusion. In the current context, given a set of 16 subject images normalized to the template using one of the three normalization methods, the pair of descriptive statistics were calculated at each voxel within the WM region of the template. To compare the resulting  $\bar{\sigma}_{FA}$  and  $\kappa$  statistical maps across spatial normalization methods, we examined their respective empirical cumulative distribution functions (CDF). The method producing better spatial alignment should result in more reduction in  $\bar{\sigma}_{FA}$  and larger increase in  $\kappa$ , which in turn will be reflected as its  $\bar{\sigma}_{FA}$  and  $\kappa$  CDFs being more to the left and to the right, respectively.

## B. Evaluating Statistical Non-Parametric Mapping of Local Volumetric Differences

For each spatial normalization method, we first examined the number of significant suprathreshold clusters and the statistical significance of these clusters. The method that identifies more local volumetric differences and thereby removes more shape confounds (to FA analysis) should find more clusters with higher significance.

To evaluate specifically the normalization quality of the regions occupied by the significant suprathreshold clusters, two-dimensional scatter plots were created to visualize the distributions of both  $\bar{\sigma}_{FA}$  and  $\kappa$  of these voxels. The scatter plots show the distribution of  $\kappa$  along the horizontal axis, the distribution of  $\bar{\sigma}_{FA}$  along the vertical axis. We expect the method with better spatial normalization quality should have its voxels clustered more to the bottom right corner of the plot, i.e., with both lower  $\bar{\sigma}_{FA}$  and higher  $\kappa$ . For each method, we constructed a scatter plot of all the voxels within any of its clusters, and for each cluster, a scatter plot of all of its voxels.

## C. Evaluating Statistical Non-Parametric Mapping of FA Differences

Here we repeated the same analyses as above for FA differences. For each spatial normalization method, the number of significant suprathreshold clusters and their statistical significance are examined. The method that identifies more FA differences should find more clusters with higher significance. Given that FA changes have been consistently reported in the motor pathway [15]–[17], [28], particular attention was paid



Fig. 2. Comparison of the overall performance of the three spatial normalization methods in terms of the empirical cumulative distribution functions (CDF) of both normalized FA standard deviation  $\bar{\sigma}_{FA}$  and dyadic coherence  $\kappa$ computed for the voxels within the white matter region using their respective normalized images. Given a distribution of some variable, the CDF of the variable associated with the distribution maps a value of the variable to the percentage of the distribution with values less than or equal to that. The CDFs computed from the initial affine-aligned images are also shown as the baselines.

to compare each method's ability to detect such changes. The scatter plots designed to analyze the clusters with significant local volumetric differences in the previous section were also used here to study the clusters with significant FA differences.

## V. RESULTS

## A. Comparison of the Overall Spatial Normalization Performance

The CDFs of  $\bar{\sigma}_{FA}$  and  $\kappa$  for all three methods along with the ones computed from the initial affine-aligned images are shown in Fig. 2. Compared to the initial affine registration, the respective  $\bar{\sigma}_{FA}$  CDFs of all three methods shifted to the left demonstrating improved spatial alignment. The diffeomorphic and tensor methods showed much more significant reduction in  $\bar{\sigma}_{FA}$  than the SPM2 method with the diffeomorphic displaying the most reduction. In the case of  $\kappa$ , the CDFs moved towards the right of that for the initial affine registration reflecting improved alignment of the dominant directions of diffusion. The diffeomorphic and tensor methods again yield better performance than the SPM2 method but the tensor method produced more evident improvement in  $\kappa$  than the diffeomorphic method.

## B. Comparison of Statistical Non-Parametric Mapping of Local Volumetric Differences

The suprathreshold cluster tests were applied to identify regions of both increased and decreased volume in the ALS population. Only clusters with decreased volume were discovered. We found 0, 2 and 2 such clusters for the SPM2, diffeomorphic and tensor methods, respectively. The selected representative slices showing all the clusters are presented in Fig. 3 by overlaying the cluster voxels over the respective mean FA images. Further details of each cluster are listed in Table I. In these tables we have labeled each cluster and reported its center of mass (in voxel unit) and volume (in number of voxels), its FWE-corrected p-value, its anatomical location, and a reference to Fig. 3, indicating the axial slice level(s) on which it appears in the figure. The anatomical locations were determined with reference to the fiber tractbased atlas of human white matter anatomy [40].

Close inspection shows that all the identified clusters correspond to extramotor areas and that although the two highdimensional methods found the same number of clusters, the clusters that they identified correspond to different anatomies. The clusters found with the tensor method are slightly higher in statistical significance and larger in size.

The scatter plots of all the voxels within any of the clusters for the diffeomorphic and tensor methods are shown in Fig. 4. It shows that, compared to the tensor method, the significant clusters found with the diffeomorphic method are with lower  $\kappa$  (p-value =  $8.7 \times 10^{-9}$ ) and slightly lower  $\bar{\sigma}_{FA}$  (p-value =  $8.4 \times 10^{-3}$ ). The voxels with significantly lower  $\kappa$  belongs to the cluster LD2 (compared to the voxels of the other three clusters, p-value =  $8.5 \times 10^{-21}$ ), indicating poor alignment of principal directions of diffusion at those voxels.

## C. Comparison of Statistical Non-Parametric Mapping of FA Differences

The suprathreshold cluster tests were applied to identify regions of both increased and decreased FA in the ALS population. Only clusters with decreased FA were discovered. We found 3, 3 and 6 such clusters for the SPM2, diffeomorphic and tensor methods, respectively. The selected representative slices showing all the clusters are presented in Fig. 5. Details of each cluster are tabulated and listed in Table II.

Close inspection of the spatial locations of these clusters show that while the findings of the two high-dimensional methods are very consistent with one another, they differ significantly from that of the SPM2 method. For the clusters located in the motor pathway, the ones found with the diffeomorphic method agree well with the ones found with the tensor method: FD1 and FT1 corresponding to the left cerebral peduncle (CP) ; FD2 and FT4 corresponding to the posterior portion of the left superior corona radiata (SCR); FD3 and FT5/FT6 corresponding to the posterior portion of the right SCR. Only a single cluster along the motor pathway was found with the SPM2 method: FS2 corresponding to the posterior portion of the left SCR. While the diffeomorphic method identified no extramotor clusters, both the SPM2 and tensor methods found two such clusters each. However, these



Fig. 4. The scatter plots of the voxels within all the clusters with significant local volumetric differences found with the diffeomorphic (top) and tensor (bottom) methods. The scatter plot of a set of voxels displays the distributions of normalized FA standard deviation  $\bar{\sigma}_{FA}$  and dyadic coherence  $\kappa$  at those voxels along the vertical and horizontal axes, respectively. For each set of voxels, the mean and standard deviation of their  $\bar{\sigma}_{FA}$  and  $\kappa$  are reported in the figure legend in that order.

two sets of extramotor clusters do not correspond to the same anatomies.

Fig. 6 shows, for each method, the scatter plot of all the voxels belonging to any of its clusters, from which we can make the following observations: (1) compared to the scatter plots of the two high-dimensional methods, the scatter plot of the SPM2 method is both more sparse and scattered. The sparsity reflects the fact that the number of voxels in the clusters found with the SPM2 method is much smaller than those of the two high-dimensional methods: 100, 212 and 210 for the SPM2, diffeomorphic and tensor methods, respectively. The more scattered pattern indicates that a large number of the voxels in the clusters found with the SPM2 method were poorly aligned. The voxels of the clusters found with the SPM2 method have larger variances in both  $\bar{\sigma}_{FA}$  and  $\kappa$  than those of the clusters found with the high-dimensional methods. Furthermore, these voxels have much higher  $\bar{\sigma}_{FA}$ (p-value =  $5.2 \times 10^{-39}$ ) than those of the clusters found with the high-dimensional methods. Finally, they have much lower  $\kappa$  (p-value =  $6.2 \times 10^{-16}$ ) compared to those of the clusters found with the tensor method, but only slightly so (pvalue = 0.27) when compared to those of the clusters found with the diffeomorphic method. (2) Comparing the two highdimensional methods, the scatter plots show that while the clusters found with the diffeomorphic method had slightly lower  $\bar{\sigma}_{FA}$  (p-value =  $4.7 \times 10^{-5}$ ) than those found with the tensor method, the clusters found with the tensor method



Fig. 3. Clusters with significant local volumetric differences found with each of the three methods superimposed with the respective FA means of the normalized control and ALS subjects. Note that the images were rendered following the radiological convention: the displayed left is the physical right.

## TABLE I

DETAILS OF THE CLUSTERS WITH SIGNIFICANT LOCAL VOLUMETRIC DIFFERENCES FOUND WITH THE DIFFEOMORPHIC AND TENSOR METHODS. FOR EACH CLUSTER, WE LISTED ITS LABEL, ITS CENTER OF MASS (IN VOXEL UNIT) AND VOLUME (IN NUMBER OF VOXELS), ITS FWE-CORRECTED P-VALUE, ITS ANATOMICAL LOCATION, AND A REFERENCE TO FIG. 3, INDICATING THE AXIAL SLICE LEVEL(S) ON WHICH IT APPEARS IN THE FIGURE.

DIFFEOMORPHIC METHOD								
Label	Center of Mass $[x, y, z]$	Volume	$P_{FWE-corr}$	Anatomical Location	Axial Slice(s) Shown			
LD1	[62, 78, 35]	37	0.0879	genu of corpus callosum	35			
LD2	[68, 68, 40]	48	0.0515	left cingulum	40			
TENSOR METHOD								
Label	Center of Mass $[x, y, z]$	Volume	$P_{FWE-corr}$	Anatomical Location	Axial Slice(s) Shown			
LT1	[72, 89, 29]	44	0.0784	left forcep minor / left anterior corona radiata	28, 30			
ITA	[F9 96 90]	55	0.0456	right forcen minor / right anterior corona radiata	28 30			

had significantly higher  $\kappa$  (p-value  $= 1.6 \times 10^{-26})$  than those found with the diffeomorphic method.

Fig. 7 compares the scatter plots of the voxels within the clusters that are along the motor pathway and found with one of the three methods. The left panel compares the clusters in the left CP: FD1 and FT1 found with the diffeomorphic and tensor methods, respectively. FD1 is slightly larger in size but less in significance than FT1. The scatter plot shows that the voxels in FD1 have slightly lower  $\bar{\sigma}_{FA}$  (p-value =  $1.4 \times 10^{-3}$ ) but much lower  $\kappa$  (p-value =  $1.2 \times 10^{-11}$ ) than those of FT1. The middle panel compares the clusters in the posterior portion of the left SCR: FS2, FD2 and FT4 found with the SPM2,

diffeomorphic and tensor methods, respectively. FS2 is the smallest of the three and the least significant. In contrast, FD2 and FT4 are both larger and more significant. Compared to FT4, FD2 is much larger and includes voxels located more superiorly. However, FT4 is the cluster of the highest statistical significance. The scatter plot shows that the voxels in FT4 have significantly lower  $\kappa$  (p-value =  $4.2 \times 10^{-34}$ ) compared to the voxels in both FS2 and FD2. The right panel compares the clusters in the posterior portion of the right SCR: FD3 found with the diffeomorphic method, FT5 and FT6 found with the tensor method. Anatomically, we consider FT5 and FT6 jointly since the two clusters are only a few voxels apart



Fig. 5. Clusters with significant FA differences found with each of the three methods superimposed with the respective means of the normalized control and ALS subjects. Note that the images were rendered following the radiological convention: the displayed left is the physical right.

from each other. They are located directly contralateral to the FT4 and the inferior portion of the FD2, while FD3 is located directly contralateral to the superior portion of the FD2. The voxels in FT5 has a much larger variance in  $\kappa$  than those of FT6 and FD3. This is accounted for by 14 of the 27 voxels in FT5 that have  $\kappa$  larger than 0.85. The mean and the standard deviation of  $\kappa$  of this subset of FT5 are  $0.66 \pm 0.11$ , while those of the other voxels in FT5 are  $0.90 \pm 0.06$ . It turns out that all these voxels are located on the slice levels 40 and 41. Anatomically, they are at the interface of the right SCR and the corpus callosum, two fiber bundles traversing in distinctly

different directions, thus the principal directions of diffusion for those voxels are poorly defined.

Fig. 8 compares the scatter plots of the voxels within the extramotor clusters found with either the SPM2 or the tensor methods. Observe that FS1 and FS3, the two clusters found with the SPM2 method, have much larger  $\bar{\sigma}_{FA}$  (pvalue =  $1.4 \times 10^{-67}$ ) than any other clusters found with any of the three methods, indicating the poor alignment. In fact, these two clusters account for the overall larger  $\bar{\sigma}_{FA}$  of the clusters found with the SPM2 method than the ones found with the two high-dimensional methods as discussed above.

#### TABLE II

DETAILS OF THE CLUSTERS WITH SIGNIFICANT FA DIFFERENCES FOUND WITH THE SPM2, DIFFEOMORPHIC AND TENSOR METHODS. FOR EACH CLUSTER, WE LISTED ITS LABEL, ITS CENTER OF MASS (IN VOXEL UNIT) AND VOLUME (IN NUMBER OF VOXELS), ITS FWE-CORRECTED P-VALUE, ITS ANATOMICAL LOCATION, AND A REFERENCE TO FIG. 5, INDICATING THE AXIAL SLICE LEVEL(S) ON WHICH IT APPEARS IN THE FIGURE.

SPM2 METHOD								
Label	Center of Mass [x, y, z]	Volume	$P_{FWE-corr}$	Anatomical Location	Axial Slice(s) Shown			
FS1	[57, 80, 37]	28	0.0495	interface between the right cingulum and corpus	37			
				callosum				
FS2	[76, 52, 40]	47	0.0165	posterior portion of the left superior corona radiata	40-42			
FS3	[57, 62, 40]	25	0.0642	right cingulum	40, 41			
DIFFEOMORPHIC METHOD								
Label	Center of Mass [x, y, z]	Volume	P <sub>FWE-corr</sub>	Anatomical Location	Axial Slice(s) Shown			
FD1	[73, 63, 24]	29	0.0636	left cerebral peduncle	24			
FD2	[74, 52, 42]	135	0.0046	posterior portion of the left superior corona radiata	40-42, 46			
FD3	[54, 47, 46]	48	0.0269	posterior portion of the right superior corona radiata	46			
TENSOR METHOD								
Label	Center of Mass [x, y, z]	Volume	P <sub>FWE-corr</sub>	Anatomical Location	Axial Slice(s) Shown			
FT1	[73, 64, 23]	26	0.0371	left cerebral peduncle	24			
FT2	[64, 62, 37]	18	0.0785	midbody of corpus callosum	37			
FT3	[51, 82, 39]	43	0.0121	right forcep minor / anterior portion of the right	37, 40			
				superior corona radiata				
FT4	[76, 52, 42]	72	0.0037	posterior portion of the left superior corona radiata	40-42			
FT5	[52, 63, 42]	27	0.0345	posterior portion of the right superior corona radiata	40-42			
FT6	[52, 54, 42]	24	0.0435	posterior portion of the right superior corona radiata	41, 42			



Fig. 6. The scatter plots of the voxels within all the clusters with significant FA differences found with the SPM2 (left), diffeomorphic (middle), and tensor (right) methods. The scatter plot of a set of voxels displays the distributions of normalized FA standard deviation  $\bar{\sigma}_{FA}$  and dyadic coherence  $\kappa$  at those voxels along the vertical and horizontal axes, respectively. For each set of voxels, the mean and standard deviation of their  $\bar{\sigma}_{FA}$  and  $\kappa$  are reported in the figure legend in that order.



Fig. 7. The scatter plots of the voxels within the clusters with significant FA differences that are along the motor pathway and found with one of the three methods: the clusters in the left cerebral peduncle (left), the clusters in the posterior portion of the left superior coronal radiata (middle), and the clusters in the posterior portion of the right superior coronal radiata (right). The scatter plot of a set of voxels displays the distributions of normalized FA standard deviation  $\bar{\sigma}_{FA}$  and dyadic coherence  $\kappa$  at those voxels along the vertical and horizontal axes, respectively. For each set of voxels, the mean and standard deviation of their  $\bar{\sigma}_{FA}$  and  $\kappa$  are reported in the figure legend in that order.

Compared to the two high-dimensional methods, the voxels of FS2 have only slightly higher  $\bar{\sigma}_{FA}$  (p-value =  $4.9 \times 10^{-4}$ ), while the voxels of FS1 and FS3 have much higher  $\bar{\sigma}_{FA}$  (p-

value =  $2.2 \times 10^{-65}$ ). This can also be appreciated visually by inspecting the Figs. 3 and 5. In clear contrast to the visible similarity between the mean FA images of the ALS and control populations for the two high-dimensional methods, there is visible difference between those for the SPM2 method for the slices on which these clusters appear. Among all the clusters found with the tensor method, FT3 is similar to FT5 in that it also has relatively lower  $\kappa$  than the others. The anatomical location of FT3 also corresponds to the interface of two fiber bundles, the corpus callosum and the anterior portion of the right SCR, traversing in different directions. Thus the principal directions of diffusion are poorly defined for many voxels of FT3.



Fig. 8. The scatter plots of the voxels within the extramotor clusters with significant FA differences found with either the SPM2 or the tensor methods. The scatter plot of a set of voxels displays the distributions of normalized FA standard deviation  $\bar{\sigma}_{FA}$  and dyadic coherence  $\kappa$  at those voxels along the vertical and horizontal axes, respectively. For each set of voxels, the mean and standard deviation of their  $\bar{\sigma}_{FA}$  and  $\kappa$  are reported in the figure legend in that order.

## VI. DISCUSSION

In this paper, we evaluated and compared the efficacy of three different algorithms representing the spectrum of spatial normalization approaches for diffusion tensor images in the context of a typical voxel-based investigation into WM differences between populations. We used an ongoing ALS study as an example and measured the efficacy of each method by its ability to identify statistically significant differences between the ALS and control populations. This assessment was then correlated with the quality of spatial normalization. In doing so, we hoped to understand the impact of spatial normalization on the validity, specificity and sensitivity of the subsequent statistical inference of WM differences.

Our assessment of the quality of spatial normalization found that the SPM2 method produced lower spatial normalization quality than the other two methods. This is expected, given that the method utilizes a low-dimensional parametrization of spatial transformation that limits its ability to represent local deformation. Also as expected, the diffeomorphic method employing a state-of-the-art scalar image registration algorithm reduced the normalized FA standard deviation the most. The tensor method performed only slightly worse than the diffeomorphic method even though it does not use FA as a feature to drive the registration directly and does not employ the most general representation of spatial transformation. On the other hand, the tensor method demonstrated significantly larger increase of  $\kappa$ , i.e., more aligned dominant directions of diffusion, than the diffeomorphic method. This suggests that the tensor method, by leveraging tensor features to guide registration, better aligns the fiber tracts underlying WM tissues.

We found that the spatial normalization quality of these methods correlated well with their ability to capture significant local volumetric differences, i.e., shape differences. The SPM2 method that had the worst spatial normalization performance identified no significant shape differences, while the other two methods that produced similarly better normalization quality both extracted two clusters of significant shape differences. Compared to the clusters found with the diffeomorphic method, the ones found with the tensor method were slightly larger in size and higher in significance. But the two highdimensional methods were essentially comparable in their performances.

We also found that a method's ability to capture significant FA differences is similarly influenced by its spatial normalization quality. Overall, the total number of voxels in the clusters found with the SPM2 method was much smaller than those of the clusters found with the two high-dimensional methods. Excluding FS1 and FS3, the two clusters with much higher  $\bar{\sigma}_{FA}$  than all other clusters, the total number of voxels found with the SPM2 method became even smaller. In contrast, the total numbers of voxels in the clusters found with the two high-dimensional methods were essentially identical. In addtion, anatomically, there was evident similarity between the motor pathway findings of the diffeomorphic and tensor methods whereas, these clearly disagreed with those of the SPM2 method. The diffeomorphic and tensor methods had all their clusters in the motor pathway similarly located spatially. In contrast, FS2 was the only such cluster found with the SPM2 method. This pattern clearly mirrors that of the observed spatial normalization quality of all the methods.

The clear discrepancy between the SPM2 method's findings and those of the other two methods made evident the impact of the spatial normalization quality. The inadequate normalization performance of the SPM2 method translated into its inability to identify the shape differences uncovered by the two high-dimensional normalization methods. Consequently, these uncovered shape differences remained and became confounds to the findings of FA differences. Structural misalignment rendered the FA values of different structures as opposed to those of the same structures from different subjects to be compared voxelwise. Our study demonstrates that this can have a number of adverse effects on the findings. Firstly, it introduced false-positive clusters. The significant FA differences reported by these clusters did not reflect any differences in the diffusion properties of some common anatomy. Instead, they indicated that there existed some systematic shape differences. FS1 and FS3, the two clusters that had much larger  $\bar{\sigma}_{FA}$ than all other clusters, were evidently such false-positives. Secondly, it introduced false-negative clusters. In this case, the confounding effect of misalignment rendered the underlying FA differences less significant. This is most likely why the SPM2 method failed to capture the other clusters in the motor pathway that were found with the two high-dimensional methods. Finally, it reduced the sensitivity of the analysis to

the true-positive clusters. FS2, the sole motor pathway finding of the SPM2 method, is a good example of this. It is both much smaller and much less significant compared to FD2 and FT4, the two clusters at the same location that were found with the two high-dimensional methods.

The consistency in the FA findings of the diffeomorphic and tensor methods provides converging evidence for the credibility of their findings and thus for an important role for high-dimensional registration algorithms in studies of WM differences. The low-dimensional normalization techniques, exemplified by the SPM2 method, have proven successful in voxel based morphometry that extracts underlying shape or morphological differences via the analysis of residual differences. However, there are clear issues in applying them for recovering "true" FA differences as highlighted above. We showed that the high-dimensional normalization methods, such as the two examined here, were able to remove shape confounds by optimally aligning common structures. Consequently they were better able to present a more complete description of WM difference by simultaneously revealing both shape differences and "true" FA differences. The former identifies notable size changes in WM tracts while the latter highlights significant changes in the underlying tissue architecture.

The few findings that differed in the diffeomorphic and tensor methods appeared to be a reflection of the difference in each method's strength. On one hand, the diffeomorphic method has the advantage of employing the most general class of smooth transformations. Therefore, it is able to model larger shape differences than the piecewise affine transformations used with the tensor method. This is likely the explanation of the tensor method's failure to identify the cluster LD1 (corresponding to significant shape differences at the midsagittal portion of the corpus callosum). Similarly, the diffeomorphic method was able to identify more voxels with FA differences in the more superior part of the motor pathway closer to the cortex, as demonstrated by the clusters FD2 and FD3. In all these regions, there is little variation in the orientation of the underlying WM tracts, thus the tensor method's sensitivity to orientation does not gain an advantage while its more constrained transformation model perhaps limited its ability to remove shape confounds. On the other hand, the tensor method is sensitive to fiber orientations, which gives it the advantage of being able to disambiguate regions with similar FA values but different orientation patterns. For instance, the cluster FT2 found with the tensor method (corresponding to FA differences at the midbody of corpus callosum) was not identified by the diffeomorphic method. Instead, the diffeomorphic method identified the cluster LD2 corresponding to shape differences in the vicinity. Observe that the cluster LD2 is located across the interface of the right cingulum and the right portion of corpus callosum, two tracts traversing in two orthogonal directions but appearing similar on the FA maps. Combining with the fact that the cluster has a significant number of voxels with low  $\kappa$ , this discrepancy is most likely explained by the diffeomorphic method's misalignment. Similarly, this might be true for LT1, LT2 and FT3, the other extramotor clusters found with the tensor methods that the diffeomorphic methods did not identify. The better alignment of tract orientation might also contribute to the higher statistical significance of the clusters FT1 and FT4 found with the tensor method (corresponding to FA differences at the left CP and the posterior portion of the left SCR) than the clusters FD1 and FD2 found with the diffeomorphic method. In light of these observations, we believe that methods, such as [54], that employ both the diffeomorphic formulation of transformation and full tensor feature driven registration, have the promise to achieve more optimal normalization and hence better discovery of WM differences.

Since the two populations used in this study were not age matched, it is important to determine if our findings are confounded by the age differences before they can be compared to previously published ALS findings. Given that the age differences between the two populations are very significant (p-value = 0.0057), we can not use age as an covariate to correct for age differences. However, in a regionof-interest based analysis, it has been shown that no significant FA changes could be identified between an middle aged (mean age 52) healthy population and an older (mean age 71) healthy population [10]. This leads us to believe that our findings of FA reductions in the motor pathway are due to the disease effect and that they are consistent with previous findings. The extramotor cluster FT3 found with the tensor methods is consistent with the known findings of age-related frontal FA reduction [10]. The extramotor clusters LT1 and LT2 found with the tensor methods might represent the frontal WM size reduction associated with the FA reduction.

There are a number of issues that the current paper does not address but will be of interest for future research. Firstly, we did not examine the effect of image smoothing on statistical findings of different normalization methods. In the case of low-dimensional normalization, Jones et al have demonstrated that varying the amount of image smoothing can produce very different findings [73]. Examining the smoothing effect on high-dimensional normalization might help us to understand if misalignment contributes to the complex nature of the smoothing effect on low-dimensional normalization. Secondly, the effect of the choice of the template was not examined. The current template was generated using the tensor-based registration with the hope that the orientation features that don't affect the normalization methods using FA but are important to the normalization method using full tensor features were better maintained in the resulting template. But it will be valuable to study how different the results of this study will be if the template is generated using the diffeomorphic normalization instead. Additionally, the current populationspecific template was not optimally constructed to be most representative of the average shape of the population and can be improved by using algorithms, such as [68], that produce shape averages. Finally, we did not explore the possiblity of deriving the spatially normalized FA maps from the spatially normalized tensor images. Since the FA is a nonlinear function of the tensor, interpolating FA will produce different numerical results from first interpolating the corresponding tensors then computing the FA. It will be interesting to determine the optimal approach to produce spatial normalized FA maps in

the future. The current evaluation framework can be used to examine this question.

## VII. CONCLUSION

We presented an evaluation of three spatial normalization methods for diffusion tensor images representative of the current range of available approaches that include lowdimensional normalization using the FA, high-dimensional normalization using the FA and high-dimensional normalization using the full tensor features. The evaluation was done in the context of a typical whole brain voxel-based population study of WM differences, using an ongoing ALS study as the example. Our findings suggest that inadequate normalization with low-dimensional approaches can result in insufficient removal of shape differences which in turn can confound FA differences in a complex manner, and that utilizing highdimensional normalization can both significantly remedy the confounding effect of shape differences to FA differences and provide a more complete description of WM differences in terms of both size and tissue architecture differences. We also found that high-dimensional approaches, by leveraging full tensor features instead of tensor-drived indices, can further improve the alignment of WM tracts.

#### ACKNOWLEDGMENT

The authors gratefully acknowledge support of this work by the NIH via grants EB006266, NS045839, HD046159, HD042974 and MH068066. This work was additionally supported by an award from the Elliott Fund at Penn Neurological Institute to L B Elman.

#### REFERENCES

- P. J. Basser, J. Mattiello, and D. L. Bihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *J. Magn. Reson.*, vol. 103, pp. 247–254, 1994.
- [2] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. D. Chiro, "Diffusion tensor MR imaging of the human brain," *Radiology*, vol. 201, pp. 637–648, 1996.
- [3] O. Ciccarelli, A. T. Toosy, G. J. Parker, C. A. Wheeler-Kingshott, G. J. Barker, D. H. Miller, and A. J. Thompson, "Diffusion tractography based group mapping of major white-matter pathways in the human brain," *NeuroImage*, vol. 19, no. 4, pp. 1545–1555, Aug 2003.
- [4] T. R. Barrick, I. N. Lawes, C. E. Mackay, and C. A. Clark, "White matter pathway asymmetry underlies functional laterlization," *Cereb Cortex*, vol. 17, no. 3, pp. 591–598, Mar 2006.
- [5] C. Büchel, T. Raedler, M. Sommer, M. Sach, C. Weiller, and M. A. Koch, "White matter asymmetry in the human brain: a diffusion tensor MRI study," *Cereb Cortex*, vol. 14, no. 9, pp. 945–951, Apr 2004.
- [6] Y. Cao, S. Whalen, J. Huang, K. L. Berger, and M. C. DeLano, "Asymmetry of subinsular anisotropy by in vivo diffusion tensor imaging," *Human Brain Mapping*, vol. 20, pp. 82–90, 2003.
- [7] S. Zhang, M. E. Bastin, D. H. Laidlaw, S. Sinha, P. A. Armitage, and T. S. Deiboeck, "Visualization and analysis of white matter structural asymmetry in diffusion tensor MRI," *Magn. Reson. Med.*, vol. 51, no. 1, pp. 140–147, Jan 2004.
- [8] H. Huang, J. Zhang, S. Wakana, W. Zhang, T. Ren, L. J. Richards, P. Yarowsky, P. Donohue, E. Graham, P. C. van Zijl, and S. Mori, "White and gray matter development in human fetal, newborn and pediatric brains," *NeuroImage*, vol. 33, no. 1, pp. 27–38, Oct 2006.
- [9] D. J. Mabbott, M. Noseworthy, E. Bouffet, S. Laughlin, and C. Rockel, "White matter growth as a mechanism of cognitive development in children," *NeuroImage*, vol. 33, no. 3, pp. 936–946, Nov 2006.

- [10] D. H. Salat, D. S. Tuch, D. N. Greve, A. J. W. van der Kouwe, N. D. hevelone, A. K. Zaleta, B. R. Rosen, B. Fischl, S. Corkin, H. D. Rosas, and A. M. Dale, "Age-related alterations in white matter microstructure measured by diffusion tensor imaging," *Neurobiology of aging*, vol. 26, pp. 1215–1227, Aug 2005.
- [11] D. H. Salat, D. S. Tuch, N. D. hevelone, B. Fischl, S. Corkin, H. D. Rosas, and A. M. Dale, "Age-related changes in prefrontal white matter measured by diffusion tensor imaging," *Annals of the New York Academy* of Sciences, vol. 1064, pp. 37–49, Dec 2005.
- [12] M. A. Horsfield and D. K. Jones, "Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - a review," *NMR In Biomedicine*, vol. 15, no. 7-8, pp. 570–577, Nov-Dec 2002.
- [13] M. A. Horsfield, H. B. Larsson, D. K. Jones, and A. Gass, "Diffusion magnetic resonance imaging in multiple sclerosis," *J Neurol Neurosurg Psychiatry*, vol. 64, no. Suppl 1, pp. S80–S84, 1998.
- [14] M. Rovaris, A. Gass, R. Bammer, S. J. Hickman, O. Ciccarelli, D. H. Miller, and M. Filippi, "Diffusion MRI in multiple sclerosis," *Neurology*, vol. 65, pp. 1526–1532, 2005.
- [15] C. M. Ellis, A. Simmons, D. K. Jones, J. Bland, J. M. Dawson, and M. A. Horsfield, "Diffusion tensor MRI assesses corticospinal tract damage in ALS," *Neurology*, vol. 53, pp. 1051–1058, 1999.
- [16] M. Sach, G. Winkler, V. Glauche, J. Liepert, B. Heimbach, M. A. Koch, C. Büchel, and C. Weille, "Diffusion tensor MRI of early upper motor neuron involvement in amyotrophic lateral sclerosis," *Brain*, vol. 127, no. 2, pp. 340–350, 2004.
- [17] S. Wang, H. Poptani, M. Bilello, X. Wu, J. H. Woo, L. B. Elman, L. F. McCluskey, J. Krejza, and E. R. Melhem, "Diffusion tensor imaging in amyotrophic lateral sclerosis: volumetric analysis of the cortiospinal tract," *Am J Neuroradiol*, vol. 27, no. 6, pp. 1234–1238, Jun 2006.
- [18] J. M. Provenzale, M. Escolar, and J. Kurtzberg, "Quantitative analysis of diffusion tensor imaging data in serial assessment of Krabbe disease," *Ann N Y Acad Sci*, vol. 1064, pp. 220–229, Dec 2005.
- [19] M. Kubicki, C.-F. Westin, S. E. Maier, H. Mamata, M. Frumin, H. Ersner-Hershfield, R. Kikinis, F. A. Jolesz, R. McCarley, and M. E. Shenton, "Diffusion tensor imaging and its application to neuropsychiatric disorders," *Harvard Rev Psychiatry*, vol. 10, pp. 324–336, 2002.
- [20] M. S. Buchsbaum, J. Friedman, B. R. Buchsbaum, K.-W. Chu, E. A. Hazlett, R. Newmark, J. S. Schneiderman, Y. Torosjan, C. Tang, P. R. Hot, D. Stewart, K. L. Davis, and J. Gorman, "Diffusion tensor imaging in schizophrenia," *Biological psychiatry*, vol. 60, no. 11, pp. 1181–1187, Dec 2006.
- [21] M. Kubicki, C.-F. Westin, R. W. McCarley, and M. E. Shenton, "The application of DTI to investigate white matter abnormalities in schizophrenia," *Ann N Y Acad Sci*, vol. 1064, pp. 134–148, Dec 2005.
- [22] N. Molko, A. Cachia, D. Rivere, J. F. Mangin, M. Bruandet, D. LeBihan, L. Cohen, and S. Dehaene, "Brain anatomy in Turner syndrome: evidence for impaired social and spatial-numerical networks," *Cereb Cortex*, vol. 14, no. 8, pp. 840–850, Aug 2004.
- [23] T. J. Simon, L. Ding, J. P. Bish, D. M. McDonald-McGinn, E. H. Zackai, and J. Gee, "Volumetric, connective and morphologic changes in the brains of children with chromosome 22q11.2 deletion syndrome: an integrative study," *NeuroImage*, vol. 25, no. 1, pp. 169–180, Mar 2005.
- [24] C. H. Sotak, "The role of diffusion tensor imaging in the evaluation of ischemic brain injury - a review," *NMR In Biomedicine*, vol. 15, no. 7-8, pp. 561–569, Nov-Dec 2002.
- [25] D. J. Tisserand, G. Stanisz, N. Lobaugh, E. Gibson, T. Li, S. E. Black, and B. Levine, "Diffusion tensor imaging for the evaluation of white matter pathology in traumatic brain injury," *Brain Cogn*, vol. 60, no. 2, pp. 216–217, Mar 2006.
- [26] O. Ciccarelli, A. T. Toosy, S. J. Hickman, G. J. M. Parker, C. A. M. Wheeler-Kingshott, D. H. Miller, and A. J. Thompson, "Optic radiation changes after optic neuritis detected by tractography-based group mapping," *Human Brain Mapping*, vol. 25, no. 3, pp. 308–316, Apr 2005.
- [27] S. H. Eriksson, F. J. Rugg-Gunn, M. R. Symms, G. J. Barker, and J. S. Duncan, "Diffusion tensor imaging in patients with epilepsy and malformations of cortical development," *Brain*, vol. 124, pp. 617–626, Mar 2001.
- [28] O. Abe, H. Yamada, Y. Masutani, S. Aoki, A. Kunimatsu, H. Yamasue, R. Fukuda, K. Kasai, N. Hayashi, T. Masumoto, H. Mori, T. Soma, and K. Ohtomo, "Amyotrophic lateral sclerosis: diffusion tensor tractography and voxel-based analysis," *NMR In Biomedicine*, vol. 17, pp. 411–416, 2004.
- [29] K. J. Friston, J. Ashburner, C. D. Frith, J.-B. Poline, J. D. Heather, and R. S. J. Frackowiak, "Spatial registration and normalization of images," *Human Brain Mapping*, vol. 2, pp. 165–189, 1995.
- [30] S. M. Smith, M. Jenkinson, H. Johansen-Berg, D. Rueckert, T. E. Nichols, C. E. mackay, K. E. Watkins, O. Ciccarelli, M. Z. Cader, P. M.

- [31] H.-J. Park, C.-F. Westin, M. Kubicki, S. E. Maier, M. Niznikiewicz, A. Baer, M. Frumin, R. Kikinis, F. A. Jolesz, R. W. McCarley, and M. E. Shenton, "White matter hemisphere asymmetries in healthy subjects and in schizophrenia: a diffusion tensor MRI study," *NeuroImage*, vol. 23, pp. 213–223, 2004.
- [32] C. E. L. Stark and Y. Okado, "Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition," *J Neurosci*, vol. 23, no. 17, pp. 6748–6753, 2003.
- [33] M. I. Miller, M. F. Beg, C. Ceritoglu, and C. Stark, "Increasing the power of functional maps of the medial temporal lobe by using large deformation diffeomorphic metric mapping," *Proc Natl Acad Sci U S A*, vol. 102, no. 27, pp. 9685–9690, 2005.
- [34] P. A. Yushkevich, J. A. Detre, D. Mechanic-Hamilton, M. A. Fernández-Seara, K. Z. Tang, A. Hoang, M. Korczykowski, H. Zhang, and J. C. Gee, "Hippocampus-specific fMRI group activation using the continuous medial representation," *NeuroImage*, vol. 35, no. 4, pp. 1516–1530, 2007.
- [35] M. Moseley, Y. Cohen, M. Kucharaczyj, J. Mintorovitch, H. Asgari, M. Wendland, J. Tsuruda, and D. Norman, "Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system," *Radiology*, vol. 176, pp. 439–445, 1990.
- [36] C. Pierpaoli and P. J. Basser, "Toward a quantitative assessment of diffusion anisotropy," *Magn. Reson. Med.*, vol. 36, pp. 893–906, 1996.
- [37] S. Mori and P. C. M. van Zijl, "Fiber tracking: principles and strategies – a technical review," *NMR In Biomedicine*, vol. 15, pp. 468–480, 2002.
- [38] B. Stieltjes, W. E. Kaufmann, P. C. M. van Zijl, G. D. Pearlson, and S. Mori, "Diffusion tensor imaging and axonal tracking in the human brainstem," *NeuroImage*, vol. 14, pp. 723–735, 2001.
- [39] S. Mori, W. E. Kaufmann, C. Davatzikos, B. Stieltjes, L. Amodei, K. Fredericksen, G. D. Pearlson, M. Solaiiyappan, G. V. Raymond, H. W. Moser, and P. C. van Zijl, "Imaging cortical association tracts in the human brain using diffusion-tensor-based axonal tracking," *Magn. Reson. Med.*, vol. 47, pp. 215–223, 2002.
- [40] S. Wakana, H. Jiang, L. M. Nagae-Poetscher, P. C. van Zijl, and S. Mori, "Fiber tract-based atlas of human white matter anatomy," *Radiology*, vol. 230, no. 1, pp. 77–87, 2004.
- [41] J. C. Gee and D. C. Alexander, "Diffusion-tensor image registration," in *Visualization and Image Processing of Tensor Fields*, J. Welckert and H. Hagen, Eds. Berlin:Springer, 2005.
- [42] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee, "Spatial transformations of diffusion tensor magnetic resonance images," *IEEE Trans. Med. Imaging*, vol. 20, no. 11, pp. 1131–1139, Nov 2001.
- [43] A. Guimond, C. R. G. Guttmann, S. K. Warfield, and C.-F. Westin, "Deformable registration of DT-MRI data based on transformation invariant tensor characterisitics," in *IEEE International Symposium on Biomedical Imaging: Macro to Nano*, July 2002, pp. 761–764.
- [44] D. K. Jones, L. D. Griffin, D. C. Alexander, M. Catani, M. A. Horsfield, R. Howard, and S. C. R. Williams, "Spatial normalization and averaging of diffusion tensor MRI data sets," *NeuroImage*, vol. 17, pp. 592–617, 2002.
- [45] D. C. Alexander and J. C. Gee, "Elastic matching of diffusion tensor images," *Comput. Vis. Image Understand.*, vol. 77, no. 2, pp. 233–250, 2000.
- [46] J. Ruiz-Alzola, C.-F. Westin, S. K. Warfield, C. Alberola, S. Maier, and R. Kikinis, "Nonrigid registration of 3d tensor medical data," *Medical Image Analysis*, vol. 6, pp. 143–161, 2002.
- [47] H.-J. Park, M. Kubicki, M. E. Shenton, A. Guimond, R. W. McCarley, S. E. Maier, R. Kikinis, F. A. Jolesz, and C.-F. Westin, "Spatial normalization of diffusion tensor MRI using multiple channels," *NeuroImage*, vol. 20, pp. 1995–2009, 2003.
- [48] K. M. Curran and D. C. Alexander, "Diffusion tensor orientation matching for image registration," in *SPIE Medical Imaging*, vol. 5032, May 2003, pp. 149–156.
- [49] —, "Orientation coherence optimisation in tensor image registration," in *Proc. MIUA*, 2004.
- [50] H. Zhang, P. A. Yushkevich, and J. C. Gee, "Registration of diffusion tensor images," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, Jun 2004, pp. 842–847.
- [51] ——, "Towards diffusion profile image registration," in *IEEE International Symposium on Biomedical Imaging: Macro to Nano*, vol. 1, Apr 2004, pp. 324–327.
- [52] Y. Cao, M. Miller, R. L. Winslow, and L. Younes, "Large deformation diffeomorphic metric mapping of vector fields," *IEEE Trans. Med. Imaging*, vol. 24, no. 9, pp. 1216–1230, 2005.

- [53] H. Zhang, P. A. Yushkevich, D. C. Alexander, and J. C. Gee, "Deformable registration of diffusion tensor MR images with explicit orientation optimization," *Medical Image Analysis – Special Issue: The Eighth International Conference on Medical Imaging and Computer Assisted intervention – MICCAI 2005*, vol. 10, no. 5, pp. 764–785, Oct 2006.
- [54] Y. Cao, M. Miller, S. Mori, R. L. Winslow, and L. Younes, "Diffeomorphic matching of diffusion tensor images," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop*, 2006, p. 67.
- [55] J. I. Berman, P. Mukherjee, S. C. Partridge, S. P. Miller, D. M. Ferriero, A. J. Barkovich, D. B. Vigneron, and R. G. Henry, "Quantitative diffusion tensor MRI fiber tractography of sensorimotor white matter development in premature infants," *NeuroImage*, vol. 27, no. 4, Oct 2005.
- [56] O. A. Glenn, R. G. Henry, J. I. Berman, P. C. Chang, S. P. Miller, D. B. Vigneron, and A. J. Barkovich, "DTI-based three-dimensional tractography detects differences in the pyramidal tracts of infants and children with congenital hemiparesis," *J. Magn. Reson. Imaging*, vol. 18, pp. 641–648, 2003.
- [57] D. K. Jones, M. Catani, C. Pierpaoli, S. J. C. Reeves, S. S. Shergill, M. O. Sullivan, P. Golesworthy, P. McGuire, M. A. Horsfield, A. Simmons, S. C. R. Williams, and R. J. Howard, "Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia," *Human Brain Mapping*, vol. 27, pp. 230–238, 2006.
- [58] R. A. Kanaan, S. S. Shergill, G. J. Barker, M. Catani, V. W. Ng, R. Howard, P. K. McGuire, and D. K. Jones, "Tract-specific anisotropy measurements in diffusion tensor imaging," *Psychiatry Research: Neuroimaging*, vol. 146, no. 1, pp. 73–82, Jan 2006.
- [59] E. Pagani, M. Filippi, M. A. Rocca, and M. A. Horsfield, "A method for obtaining tract-specific diffusion tensor MRI measurements in the presence of disease: application to patients with clinically isolated syndromes suggestive of multiple sclerosis," *NeuroImage*, vol. 26, no. 1, pp. 258–265, May 2005.
- [60] I. Corouge, P. T. Fletcher, S. Joshi, S. Gouttard, and G. Gerig, "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," *Medical Image Analysis*, vol. 10, no. 5, pp. 786–798, Oct 2006.
- [61] D. K. Jones, A. R. Travis, G. Eden, C. Pierpaoli, and P. J. Basser, "PASTA: pointwise assessment of streamline tractography attributes," *Magn. Reson. Med.*, vol. 53, pp. 1462–1467, 2005.
- [62] F. Lin, C. Yu, T. Jiang, K. Li, X. Li, W. Qin, H. Sun, and P. Chan, "Quantitative analysis along the pyramidal tract by length-normalized parametrization based on diffusion tensor tractography: Application to patients with relapsing neuromyelitis optica," *NeuroImage*, vol. 33, no. 1, pp. 154–160, Oct 2006.
- [63] M. Kinoshita, K. Yamada, N. Hashimoto, A. Kato, S. Izumoto, T. Baba, M. Maruno, T. Nishimura, and T. Yoshimine, "Fiber-tracking does not accurately estiamte size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation," *NeuroImage*, vol. 25, no. 2, pp. 424–429, Apr 2005.
- [64] J. F. Mangin, C. Poupon, C. A. Clark, D. L. Bihan, and I. Bloch, "Distortion correction and robust tensor estimation for MR diffusiong imaging," *Medical Image Analysis*, vol. 8, pp. 191–198, 2002.
- [65] S. M. Smith, "Fast robust automated brain extraction," Human Brain Mapping, vol. 17, no. 3, pp. 143–155, Sep 2002.
- [66] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, no. S, pp. 151–160, 2004.
- [67] J. Ashburner and K. J. Friston, "Voxel-based morphometry: the methods," *NeuroImage*, vol. 11, pp. 805–821, 2000.
- [68] B. B. Avants and J. C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interpolation," *NeuroImage*, vol. 23, no. Suppl 1, pp. 139–150, 2004.
- [69] T. Nichols and A. P. Holmes, "Nonparametric analysis of PET functional neuroimaging experiments: a primer," *Human Brain Mapping*, vol. 15, pp. 1–25, 2001.
- [70] K. J. Friston and A. P. Holmes, "Statistical parametric maps in functional imaging: a general linear approach," *Human Brain Mapping*, vol. 2, pp. 189–210, 1995.
- [71] P. A. Freeborough and N. C. Fox, "Modeling brain deformations in Alzheimer disease by fluid registration of serial MR images," J. Comput. Assisted. Tomogr., vol. 22, pp. 838–843, 1998.
- [72] J. C. Gee and R. K. Bajcsy, "Elastic matching: continuum mechanical and probabilistic analysis," in *Brain warping*. Academic Press, San Diego, 1999, ch. 11, pp. 183–198.

[73] D. K. Jones, M. R. Symms, M. Cercignani, and R. J. Howard, "The effect of filter size on VBM analyses of DT-MRI data," *NeuroImage*, vol. 26, no. 2, pp. 546–554, Jun 2005.