

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/118318/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Ravikumar, Nishant, Gooya, Ali, Beltrachini, Leandro , Frangi, Alejandro F. and Taylor, Zeike A. 2019. Generalised coherent point drift for group-wise multi-dimensional analysis of diffusion brain MRI data. *Medical Image Analysis* 53 , pp. 47-63. 10.1016/j.media.2019.01.001

Publishers page: <https://doi.org/10.1016/j.media.2019.01.001>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Generalised coherent point drift for group-wise multi-dimensional analysis of diffusion brain MRI data

Nishant Ravikumar<sup>a,b,\*</sup>, Ali Gooya<sup>a,c</sup>, Leandro Beltrachini<sup>a,c</sup>, Alejandro F. Frangi<sup>a,c</sup>, Zeike A. Taylor<sup>a,b</sup>

<sup>a</sup>*CISTIB Centre for Computational Imaging and Simulation Technologies in Biomedicine, INSIGNEO Institute for in silico Medicine*

<sup>b</sup>*Department of Mechanical Engineering, The University of Sheffield*

<sup>c</sup>*Department of Electronic and Electrical Engineering, The University of Sheffield*

---

## Abstract

A probabilistic framework for registering generalised point sets comprising multiple voxel-wise data features such as positions, orientations and scalar-valued quantities, is proposed. It is employed for the analysis of magnetic resonance diffusion tensor image (DTI)-derived quantities, such as fractional anisotropy (FA) and fibre orientation, across multiple subjects. A hybrid Student's t-Watson-Gaussian mixture model-based non-rigid registration framework is formulated for the joint registration and clustering of voxel-wise DTI-derived data, acquired from multiple subjects. The proposed approach jointly estimates the non-rigid transformations necessary to register an unbiased mean template (represented as a 7-dimensional hybrid point set comprising spatial positions, fibre orientations and FA values) to white matter regions of interest (ROIs), and approximates the joint distribution of voxel spatial positions, their associated principal diffusion axes, and FA. Specific white matter ROIs, namely, the corpus callosum and cingulum, are analysed across healthy control (HC) subjects (K=20 samples) and patients diagnosed with mild cognitive impairment (MCI) (K=20 samples) or Alzheimer's disease (AD) (K=20 samples) using the proposed framework, facilitating inter-group comparisons of FA and fibre orientations. Group-wise analyses of the

---

\*Corresponding author

*Email addresses:* [mta08nr@gmail.com](mailto:mta08nr@gmail.com) (Nishant Ravikumar), [a.gooya@sheffield.ac.uk](mailto:a.gooya@sheffield.ac.uk) (Ali Gooya), [beltrachinil@cardiff.ac.uk](mailto:beltrachinil@cardiff.ac.uk) (Leandro Beltrachini), [A.Frangi@leeds.ac.uk](mailto:A.Frangi@leeds.ac.uk) (Alejandro F. Frangi), [Z.Taylor@leeds.ac.uk](mailto:Z.Taylor@leeds.ac.uk) (Zeike A. Taylor)

latter is not afforded by conventional approaches such as tract-based spatial statistics (TBSS) and voxel-based morphometry (VBM).

---

## 1. Introduction

Group-wise registration of multi-dimensional unstructured point sets comprising different types of data such as directional/axial and scalar-valued quantities is useful for a variety of medical imaging and computer vision applications. This study proposes a probabilistic approach for group-wise registration of generalised point sets comprising positions, associated axial orientations and scalar-valued measures. This is achieved through formulation of a hybrid mixture model (HdMM), combining suitable probability distributions to model disparate data features within a cohesive framework. As an exemplar application, the proposed framework is employed for the joint registration and clustering of magnetic resonance (MR) diffusion tensor image (DTI)-derived data, acquired from multiple subjects. The generality of the proposed framework however, makes it suitable for registering other types of hybrid point sets comprised of feature vectors containing principal curvatures, surface normals, integral descriptors, etc. High-dimensional feature vectors are in general more descriptive (than spatial positions alone, for example) and discriminative when establishing correspondences, due to the low probability of matching all features for non-corresponding points.

MR-DTI has found widespread use for studying structural changes within brain white matter (WM), and the potential of such changes as biomarkers for dementia and other neurodegenerative diseases. DT fields are estimated from diffusion weighted images (DWIs), which encode diffusion of water molecules along different gradient directions. MR-DTIs use a diffusion tensor model (Basser et al., 1994) that, under some assumptions, can be related to local tissue microstructure. They aid in voxel-wise quantification of diffusion characteristics, which may be expressed in terms of principal eigenvectors and eigenvalues of the estimated diffusion tensors. Tissue microstructure affects local diffusion properties. For example, water diffuses preferentially parallel to the major axis of a fibre bundle, as opposed to perpendicular to it and, consequently, gives rise to the sense of tissue anisotropy commonly observed in major WM tracts. Fractional anisotropy (FA), a measure frequently employed to describe tissue anisotropy (Pierpaoli and Basser, 1996), represents the degree of directional dependence in diffusion at a specific voxel. The primary eigenvector of a diffusion tensor represents the preferred direction for

35 the diffusion of water at any given voxel, and is often interpreted as reflecting  
36 the local fibre orientation within tissue.

37 Region of interest (ROI)-based analyses have been used to assess changes  
38 in local (Salat et al., 2005) and global (Cercignani et al., 2001) tissue diffusion  
39 properties. A limitation of such approaches is the need to accurately delin-  
40 eate ROIs across multiple patients’/subjects’ images. Consequently, they  
41 are affected by low reproducibility, leading to discrepancies across studies.  
42 Tract-based spatial statistics (TBSS) (Smith et al., 2006) and voxel-based  
43 morphometric (VBM) approaches (Ashburner and Friston, 2000) are suit-  
44 able alternatives that are fully automatic and enable analysis of localised  
45 changes to FA and other diffusion measures, across the entire WM volume.  
46 The quality of non-rigid registration used in VBM significantly influences  
47 the subsequent voxel-wise analysis. To overcome this issue, (Smith et al.,  
48 2006) proposed the widely used TBSS approach, which ensures that registra-  
49 tion quality has less influence on subsequent statistical analysis of FA (and  
50 other diffusion-derived quantities). TBSS constructs an alignment invariant  
51 mean FA skeleton following registration of subjects’ FA images to a template.  
52 Neighbouring voxels located perpendicular to the skeleton are identified for  
53 each subject, and the highest FA values are assigned to each skeleton voxel.  
54 The resulting projections to the skeleton enable statistical analysis across  
55 multiple subjects.

56 Alternative probabilistic techniques that jointly register and cluster WM  
57 fibre trajectories (obtained from diffusion tractography), and which enable  
58 quantitative analysis of diffusion measures over fibre pathways (rather than  
59 voxel-wise quantification), have also been proposed. For example, registra-  
60 tion of curves and fibre bundles using diffeomorphisms and currents, and  
61 a statistical framework to assess variability in geometry and fibre density  
62 across a population, was proposed in (Durrleman et al., 2009), (Durrleman  
63 et al., 2011). Maddah et al. (2008) employ a Gamma mixture modelling  
64 framework to register fibre trajectories by establishing probabilistic corre-  
65 spondences, and jointly cluster them into representative fibre bundles. The  
66 authors also note therein, through use of a suitable fibre tract atlas as a prior  
67 during the clustering procedure, correspondences may be estimated across fi-  
68 bre trajectories obtained from multiple subjects, thereby enabling statistical  
69 analysis of FA and other diffusion quantities across populations. Similarly,  
70 (Mayer et al., 2011) proposed a supervised approach for joint registration  
71 and segmentation WM tracts, wherein, the iterative closest fiber algorithm  
72 (Mayer and Greenspan, 2008) was used to register fibre sets between a manu-

73 ally annotated tractography atlas and a subject’s reconstructed set of fibres.  
74 The resulting segmentation was subsequently refined using a probabilistic  
75 boosting tree-based classifier. In (Zvitia et al., 2010), the authors propose a  
76 combined adaptive mean shift and Gaussian mixture model (GMM) formu-  
77 lation to jointly cluster fibre trajectories into compact fibre sets, and subse-  
78 quently register fibre sets obtained from multiple subjects. The registration  
79 of two clustered fibre sets is formulated as a problem of aligning two distinct  
80 GMMs, analogous to point set registration using GMMs (Jian and Vemuri,  
81 2005). Similar approaches to clustering fibre trajectories across a population,  
82 using spectral embedding, have also been proposed (O’Donnell and Westin,  
83 2007), facilitating the estimation of WM atlases and enabling automatic seg-  
84 mentation of major WM tracts. An unbiased, group-wise, whole-brain trac-  
85 tography registration approach was proposed by (O’Donnell et al., 2012).  
86 Kernel density estimation was used to approximate the probability distribu-  
87 tion of fibre trajectories within each brain and the overall distribution of the  
88 atlas, was modelled as a mixture of the former. Alignment of WM tracts was  
89 achieved by minimizing an entropic measure defined on the atlas distribution.  
90 In a follow up study ODonnell et al. (2017), this group-wise registration ap-  
91 proach was combined with their previous work on spectral clustering of fibre  
92 trajectories, to formulate an end-to-end automated framework for automated  
93 WM tract identification, thereby enabling statistical analyses of DTI-derived  
94 quantities. Garyfallidis et al. (2015) proposed a linear registration framework  
95 to align WM bundles directly in the space of streamlines. They also demon-  
96 strated the viability of their approach to construct bundle specific atlases.  
97 In a recent study (Benou et al., 2018), novel descriptors called Fiber-Flux  
98 Diffusion Density (FFDD), which jointly describe fibre bundle geometry and  
99 diffusivity measures were proposed, to facilitate localized quantification of  
100 WM fibre bundles. Additionally, a FFDD dissimilarity measure was formu-  
101 lated and a novel registration framework (based on the fast marching method)  
102 for WM tract-profiles was proposed, enabling inter-subject comparisons and  
103 group-wise statistical analysis. Such techniques are however, dependent on  
104 the tractography algorithm employed to estimate fibre trajectories, introduc-  
105 ing an additional potential source of error, and typically require some degree  
106 of user intervention (to define seeds for streamline generation for example).

107 Applications of the various methods described above have included, for  
108 example, identification of relationships between mild cognitive impairment  
109 (MCI) and Alzheimer’s disease (AD), and localised changes to WM diffusion  
110 characteristics. For example, in (Zhang et al., 2007), ROI-based analysis

111 was used to identify significant reduction in FA in the cingulum for patients  
112 diagnosed with MCI and AD, relative to healthy controls (HC). In (Medina  
113 et al., 2006), VBM was used to identify significant reduction in FA in poste-  
114 rior regions of the brain, for MCI and AD patient groups, using VBM. While  
115 (Liu et al., 2011) used the TBSS-approach and found reduced FA in the  
116 cingulum, corpus callosal and inferior/superior longitudinal fasciculus tracts,  
117 among others.

118 This study proposes a probabilistic approach to enable statistical anal-  
119 ysis of diffusion-derived measures, as an alternative to existing VBM- and  
120 TBSS-based approaches. The latter are based on non-rigid registration of  
121 subjects' FA images to a standard space to perform such analysis. Instead,  
122 our approach uses group-wise non-rigid point set registration based on a  
123 novel mixture modelling framework, which approximates the joint probabil-  
124 ity density of: (1) spatial positions (of voxel centroids within a region/tract  
125 of interest), (2) primary diffusion axes (henceforth referred to as fibre orien-  
126 tations for brevity), and (3) fractional anisotropy, estimated at the voxels of  
127 interest. The proposed framework is flexible and can be used to model other  
128 diffusion-derived data such as mean/radial diffusivity, relative anisotropy,  
129 tensor-eigenvalues, etc. — a functionality also afforded by TBSS. However,  
130 the proposed approach also enables analysis of the variation in fibre orienta-  
131 tions, across multiple subjects, which is not possible with conventional TBSS  
132 and VBM approaches.

133 Statistical analysis of fibre orientations across multiple subjects and com-  
134 parisons between patient groups was pursued in a previous study (Schwartz-  
135 man et al., 2005). Here, the authors followed a VBM-style approach where  
136 DTIs from multiple subjects were spatially normalized to a reference template  
137 using a spline-based tensor interpolation approach together with a tensor re-  
138 orientation mechanism designed to preserve the principal diffusion direction.  
139 Subsequently, Watson distributions were fitted by maximum likelihood esti-  
140 mation to the fibre orientations observed across a group, at each voxel,  
141 independently. This provides a measure of the mean orientation and disper-  
142 sion, observed across the group of subjects. A drawback of such an approach  
143 however, is the need to choose a single, appropriate template, for spatial  
144 normalization, which is particularly difficult for images exhibiting varying  
145 degrees of pathology-induced morphological changes. All subsequent reg-  
146 istrations performed and correspondences estimated are biased towards the  
147 chosen template. VBM-based approaches in general, are dependent on the ac-  
148 curacy of non-rigid registration and the exact estimation of correspondences,

149 to ensure validity in the subsequent voxel-wise statistical analyses. TBSS  
150 and our proposed approach are less restrictive in this regard. Registration of  
151 WM regions defined by hybrid point sets (comprising voxel positions, asso-  
152 ciated fibre orientations and FA values) across subjects, is achieved using a  
153 group-wise rigid, and subsequent non-rigid point set registration procedure,  
154 based on a HdMM. In the proposed approach, correspondence probabilities  
155 are estimated by approximating the joint probability density of position,  
156 fibre orientation and FA, which are iteratively revised as the registration  
157 progresses. Consequently, three distinct sources of information are leveraged  
158 to guide the registration of an unbiased, study-specific atlas (iteratively re-  
159 visited as the registration progresses), onto each subject’s WM tract/ROI. The  
160 evolving soft correspondences provide model-based estimates for the mean fi-  
161 bre orientation and FA value (for a given population) at each component in  
162 the mixture model and help mitigate any misalignment incurred during reg-  
163 istration.

### 164 *1.1. Motivation and Contributions*

165 The primary motivation for this study is to enable quantitative compar-  
166 isons of both voxel-wise scalar-valued (such as FA) and vector-valued (such as  
167 position and orientation) DTI data, across multiple subjects. Although the  
168 proposed framework is used to analyse voxel-wise diffusion-derived quantities  
169 in this study, the method itself is not intrinsically dependent on voxel-wise  
170 (or structured grid-wise) data, i.e. the framework could be used to register  
171 and analyse unstructured data as well. The proposed hybrid mixture model  
172 approximates the joint probability density function (*PDF*) of spatial posi-  
173 tions, associated fibre orientations and FA values, using Student’s *t*, Watson  
174 and Gaussian distributions, respectively. The proposed approach models the  
175 *PDF* of fibre orientations, rather than the directions of the observed primary  
176 diffusion eigenvectors, which tend to be random (as diffusion tensors are an-  
177 tipodally symmetric). To the best of our knowledge, this is the first study to  
178 formulate such a hybrid mixture model-based registration framework, which  
179 employs Watson distributions to model fibre orientations.

## 180 **2. Methods**

### 181 *2.1. Pre-processing*

182 MR-DWIs were acquired for 60 subjects (20 HC, 20 MCI, 20 AD), as  
183 part of prospective cohort of the VPH-DARE@IT project ([vph-dare.eu](http://vph-dare.eu)).

184 All images used in this study were acquired using identical protocols: 2  
 185 diffusion-weighted b-values (0, 800), with diffusivity gradients applied along  
 186 32 directions; image size of  $(240 \times 240 \times 120)$  slices, 2.5mm thick in the right-  
 187 left, anterior-posterior and inferior-superior directions, respectively. DTIs  
 188 were estimated from these for each subject using TORTOISE v 2.5.0 (Pier-  
 189 paoli et al., 2010), which employs state-of-the-art algorithms for motion and  
 190 eddy current correction, correcting B0 susceptibility induced EPI distortions  
 191 and B-matrix re-orientation artefacts. Tensor-fitting was then achieved us-  
 192 ing iRESTORE (Chang et al., 2012), based on non-linear iterative least-  
 193 squares. TORTOISE registers each subject’s DWIs to their corresponding  
 194 T2-weighted structural MRI during the aforementioned pre-processing steps.  
 195 As the latter were acquired at resolutions of  $(1.5 \times 1.5 \times 1.5mm)$ , all estimated  
 196 DTIs (and correspondingly, DTI-derived images) were up-sampled relative to  
 197 their raw DWIs. Finally, tensor-derived measures such as the eigenvector and  
 198 fractional anisotropy images were also estimated using TORTOISE.

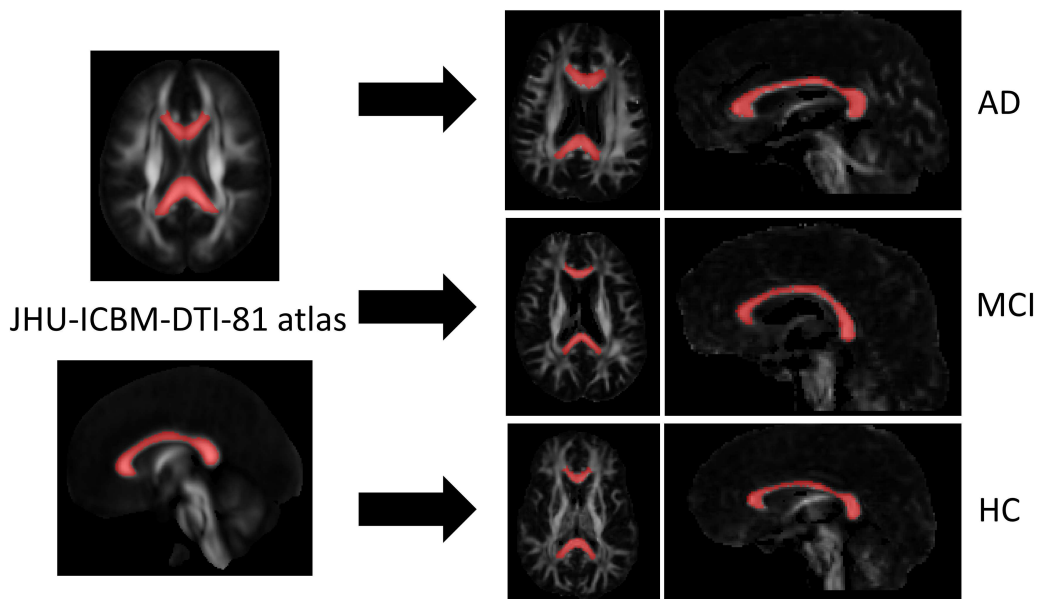


Figure 1: Nifty-Reg used to propagate labels for WM regions of interest from JHU-ICBM-DTI-81 atlas to each subject in AD, MCI and HC groups. Images depict propagation of the corpus callosum label from the atlas to subjects in AD, MCI and control groups.

199 The proposed framework is flexible and can consider the entire WM vol-



200 ume as the region of interest, eliminating the need for pre-processing steps  
201 in the form of *a priori* definition of the ROIs (using atlas-based label prop-  
202 agation for example). However, such an automated approach to analysing  
203 the entire WM volume across multiple subjects carries significant computa-  
204 tional burden. Consequently, for the purpose of this study, we restrict our  
205 attention to two WM regions, namely, the cingulum and corpus callosum.  
206 An atlas-based label propagation approach is used to segment the WM ROIs  
207 from all subjects' FA images. The fractional anisotropy image of the JHU-  
208 ICBM-DTI-81 atlas <sup>1</sup> (Mori et al., 2008) - (Hua et al., 2008) is non-rigidly  
209 registered to each subject's FA image (following an initial affine alignment),  
210 using Nifty-Reg v 1.3.9 (Ourselin et al., 2001), (Modat et al., 2010), a de-  
211 formable image registration algorithm based on cubic B-splines. Following  
212 FA image registration, the segmented labels for the cingulum and corpus cal-  
213 losum defined on the atlas (available along with the FA atlas), are resampled  
214 to the space of each subject's FA image. In this way, labels delineating the  
215 cingulum and corpus callosum in the atlas image, are propagated to each  
216 subject's image, segmenting the ROIs (as illustrated in Fig. 1).

---

<sup>1</sup>Available at: <http://www.loni.usc.edu/ICBM/Downloads/Downloads'DTI-81.shtml>

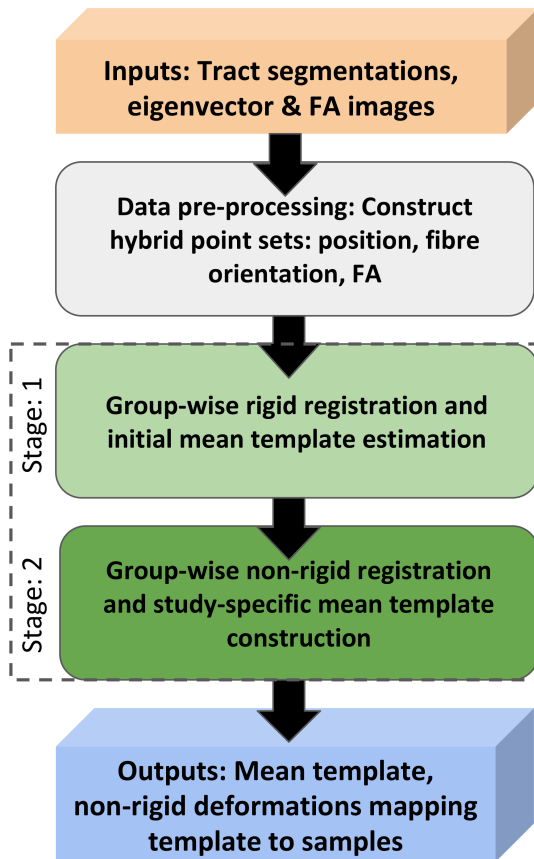


Figure 2: Summary of steps involved in the proposed framework to jointly register and cluster hybrid point sets comprising spatial positions, fibre orientations and FA values, for a WM tract/ROI. Dashed box outlines the two stages of the proposed algorithm.

218 The steps involved in the proposed approach are summarised by Fig. 2.  
 219 For a group of  $k = 1 \dots K$  subjects to be analysed (e.g. comprising control,  
 220 MCI and AD sub-groups), their tract segmentations, eigenvector and FA im-  
 221 ages were used to construct hybrid point sets  $\mathbf{D}_k$ , where each data point is a  
 222 7-dimensional vector denoted as  $\mathbf{d}_{ki} = [\mathbf{x}_{ki}, \mathbf{n}_{ki}, f_{ki}]$ . Here  $\mathbf{x}_{ki}$  represents the  
 223 spatial co-ordinate,  $\mathbf{n}_{ki}$  represents the primary diffusion eigenvector and  $f_{ki}$   
 224 denotes the FA value for the  $i^{\text{th}}$  voxel, in the  $k^{\text{th}}$  subject's image.  $\mathbf{x}_{ki}$  are  
 225 consequently, densely distributed points within the volumes/ROIs. The re-  
 226 sulting hybrid point sets were, subsequently, jointly registered and clustered

227 by fitting an  $M$ -component hybrid mixture model (comprising Student's t,  
 228 Watson and Gaussian distributions) to the data. This was achieved over  
 229 two stages (as depicted in Fig. 2): (1) Group-wise rigid registration of the  
 230 hybrid point sets  $\mathbf{D}_k$  and mean template  $\mathbf{M}$  construction; and (2) Group-  
 231 wise non-rigid registration, wherein the mean template estimated in stage  
 232 1 was non-rigidly registered to each sample from all patient groups simul-  
 233 taneously. The similarity transformation and the non-rigid transformation,  
 234 corresponding to stage 1 and 2 of the algorithm respectively, are both repre-  
 235 sented by  $\mathbf{T}_k$  throughout this study. For the former,  $\mathbf{T}_k = [s_k, \mathbf{R}_k, \mathbf{t}_k]$ . Here,  
 236  $s_k, \mathbf{R}_k, \mathbf{t}_k$  represent the scaling, rotation and translation (for the  $k^{\text{th}}$  sam-  
 237 ple), respectively, estimated in stage 1. These are used to align the hybrid  
 238 point sets to the estimated mean template and initialise the subsequent non-  
 239 rigid registration step (stage 2) by correcting global pose differences across  
 240 the data set. Stage 2 of the algorithm estimates non-rigid transformations  
 241  $\mathbf{T}_k$ , defined by a linear combination of radial basis functions (with a Gaus-  
 242 sian kernel). Together with a Gaussian kernel, the basis function weights  
 243  $\mathbf{W}_k$  estimated define point-wise displacements that map the mean template  
 244 to each sample within a subject group. In both stages of the algorithm,  
 245 estimation of the desired registration parameters was accompanied by the  
 246 joint clustering of positions, orientations and FA values. The parameters to  
 247 be estimated for each of the  $j = 1 \dots M$  components of the hybrid mixture  
 248 model include:  $\{\mathbf{m}_j^p, \sigma_p^2, \nu_j\} = \Theta_p$ , which represent mean spatial positions,  
 249 their variance and the degrees of freedom, respectively, for the Student's t-  
 250 distributions;  $\{\mathbf{m}_j^d, \kappa_j\} = \Theta_n$ , which represent the mean fibre orientations  
 251 and concentration around the means, respectively, for the Watson distribu-  
 252 tions;  $\{m_j^f, \sigma_f^2\} = \Theta_f$ , which denote the mean FA values and FA variance,  
 253 respectively, for the Gaussian distributions; and  $\pi_j$  which denote the mix-  
 254 ture coefficients. Following non-rigid registration, the study-specific mean  
 255 template estimated (for each WM ROI)  $\mathbf{M}$  thus comprises positions,  $\mathbf{m}_j^p$ ,  
 256 orientations  $\mathbf{m}_j^d$  and FA values  $m_j^f$ .

### 257 2.3. Joint Probabilistic Model of Position, Orientation and Anisotropy

258 The problem of joint registration and clustering of hybrid point sets is  
 259 formulated as one of maximum likelihood parameter estimation, using a hy-  
 260 brid mixture model that approximates the joint *PDF* of spatial positions (of  
 261 voxel centroids), fibre orientations, and fractional anisotropy. By assuming  
 262 voxel positions, fibre orientations, and FA values to be independent and iden-  
 263 tically distributed (i.i.d), for each subject and across multiple subjects, the

264 joint *PDF* can be approximated as a product of the individual conditional  
265 densities (Bishop, 2006) for position, orientation and FA. Consequently, by  
266 considering all data points  $\mathbf{d}_{ki} \in \mathbf{D}_k$ , from all  $K$  subjects, to be i.i.d. the con-  
267 ditional probability of an observation being sampled from an  $M$ -component  
268 HdMM is given by equation 1a. The set of all transformations (similarity  
269 or non-rigid) is represented by  $\mathbf{T}_k \in \mathbb{T}$ ;  $\Theta_p$  represents the set of model pa-  
270 rameters associated with the Student's t-distributions  $\mathcal{S}$ , used to model the  
271 distribution of voxel spatial positions;  $\Theta_n$  represents the parameters of the  
272 Watson distributions  $\mathcal{W}$  (modelling fibre orientations);  $\Theta_f$  denotes the set  
273 of parameters of the Gaussian distributions  $\mathcal{N}$  (modelling FA); and  $\pi_j \in \Pi$   
274 represents the set of mixture coefficients, of the HdMM. Here and through-  
275 out, subscript  $j = 1 \dots M$  denotes mixture components and the choice of  
276 distributions indicated earlier will be justified later in this Section. Using  
277 equation (1a) the log-likelihood function is formulated as shown in equation  
278 (1b), which defines the cost function to be optimised with respect to the mix-  
279 ture model and transformation parameters  $\{\Theta_p, \Theta_n, \Theta_f, \Pi, \mathbb{T}\} \in \Psi$ , to jointly  
280 register and cluster the hybrid point set data  $\mathbf{D}_k \in \mathbb{D}$ .

$$p(\mathbf{d}_{ki} | \Theta_p, \Theta_n, \Theta_f, \mathbf{T}_k) = \sum_{j=1}^M \pi_j \mathcal{S}(\mathbf{x}_{ki} | \Theta_p, \mathbf{T}_k) \mathcal{W}(\mathbf{n}_{ki} | \Theta_n, \mathbf{T}_k) \mathcal{N}(f_{ki} | \Theta_f, \mathbf{T}_k) \quad (1a)$$

$$\ln p(\mathbb{D} | \Psi) = \sum_{k=1}^K \sum_{i=1}^{N_k} \ln p(\mathbf{d}_{ki} | \Theta_p, \Theta_n, \Theta_f, \mathbf{T}_k) \quad (1b)$$

$$P_{kij}^t = \frac{\pi_j p(\mathbf{d}_{ki} | \Theta_p^t, \Theta_n^t, \Theta_f^t, \mathbf{T}_k)}{\sum_{l=1}^M \pi_l p(\mathbf{d}_{ki} | \Theta_p^t, \Theta_n^t, \Theta_f^t, \mathbf{T}_k)} \quad (1c)$$

$$Q(\Psi^{t+1} | \Psi^t) = \sum_{k=1}^K \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^t \left[ \ln \pi_j + Q(\Theta_{p_j}^{t+1}, \mathbf{T}_k^{t+1} | \Theta_{p_j}^t, \mathbf{T}_k^t) \right. \\ \left. + Q(\Theta_{n_j}^{t+1}, \mathbf{T}_k^{t+1} | \Theta_{n_j}^t, \mathbf{T}_k^t) + Q(\Theta_{f_j}^{t+1}, \mathbf{T}_k^{t+1} | \Theta_{f_j}^t, \mathbf{T}_k^t) \right] \quad (1d)$$

---

**Algorithm 1** Hybrid Mixture Model: HdMM

---

Inputs: Group of hybrid point sets  $\mathcal{D}_{k=1..K}$ , number of mixture components  $M$ , max.iterations

Outputs: Set of HdMM parameters  $\{\Theta_p, \Theta_n, \Theta_f\} \in \Psi$ , soft correspondences

```
1: INITIALIZATION
2: Initialize  $\mathbf{M}, \sigma_p^2, \sigma_f^2$  using K-means clustering.
3: All  $\pi_j = 1/M$  and  $\nu_j = 3.0, \kappa_j = 1.0$ 
4: procedure STAGE 1 EM:
5: GROUP-WISE RIGID REGISTRATION( $\mathbf{D}_k, \Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{T}_k$ )  $\triangleright$  EM
   initialized
6:   while Iteration < max.iterations do
7:     Compute  $P_{kij}$   $\triangleright$  E-step
8:     Update  $\mathbf{R}_k, s_k, \mathbf{t}_k$   $\triangleright$  M-step
9:     Update  $\Theta_p, \Theta_n, \Theta_f$   $\triangleright$  M-step
10:  end while
11:  return  $\Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{T}_k$ 
12: end procedure
13: Estimated mean template  $\mathbf{M}$ , mixture coefficients  $\Pi$  and similarity transformations  $\{\mathbf{T}_k\}_{k=1..K}$  initialise Stage 2.
14: procedure STAGE 2 EM:
15: GROUP-WISE NON-RIGID REGISTRATION( $\mathbf{D}_k, \Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{W}_k$ )  $\triangleright$  EM
   non-rigid initialized
16:  while Iteration < max.iterations do
17:    Compute  $P_{kij}$   $\triangleright$  E-step
18:    Update  $\mathbf{W}_k$   $\triangleright$  M-step
19:    Update  $\mathbf{M}, \sigma_p^2, \nu_j, \Theta_n, \Theta_f$   $\triangleright$  M-step
20:    Update spatial positions of each  $\mathbf{D}_k$ 
21:  end while
22:  return  $\Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{W}_k$ 
23: end procedure
```

---

281 A tractable approach to maximising equation 1b is achieved using the  
282 expectation-maximisation (EM) framework (Dempster et al., 1977), which  
283 iteratively alternates between: the expectation (E)-step, which evaluates the  
284 mixture component membership probabilities as shown in equation 1c (i.e.  
285 posterior probabilities  $P_{kij}^t$ , that define soft correspondences and are expec-

286 tations of the latent variables in the model) for the observed data, given  
 287 an estimate of the model parameters  $\Psi^t$ , at the  $t^{\text{th}}$  EM-iteration; and the  
 288 maximisation (M)-step, which uses the computed posterior probabilities  $P_{kij}^t$   
 289 to maximise the conditional expectation of the complete-data-log-likelihood  
 290 function  $Q$  (refer to equation 1d), with respect to each model parameter,  
 291 resulting in revised estimates  $\Psi^{t+1}$ . As shown in equation 1d,  $Q$  for the hy-  
 292 brid mixture model can be expressed as a sum of contributions from each  
 293 distribution and corresponding data feature (i.e. position, orientation and  
 294 FA), denoted,  $Q(\Theta_p^{t+1}|\Theta_p^t)$ ,  $Q(\Theta_n^{t+1}|\Theta_n^t)$ ,  $Q(\Theta_f^{t+1}|\Theta_f^t)$ , respectively. The com-  
 295 plete algorithm for the proposed hybrid mixture model, to jointly register  
 296 and cluster a group  $\mathbb{D}$  of hybrid point sets, is summarized in Algorithm 1.  
 297 Subsequent sections discuss each probability distribution and estimation of  
 298 their associated parameters, within the proposed framework, in more detail.

#### 299 2.4. Mixture Model for Primary Diffusion Axes

300 In addition to modelling the spatial distribution of voxels defining ROIs,  
 301 the proposed approach also deals with axial data distributed over the  $S^2$   
 302 sphere, i.e. fibre orientations defined by primary diffusion eigenvectors.  
 303 GMMs and TMMs, comprising Gaussian and Student’s t-distributions, re-  
 304 spectively, are inappropriate for clustering such data and consequently, a  
 305 mixture of Watson distributions, also defined over the spherical domain, is  
 306 employed in this study. While Von-Mises-Fisher distributions are frequently  
 307 used for clustering directional data, they are unsuitable for axial data, as  
 308 they lack of antipodal symmetry. Watson distributions on the other hand,  
 309 are naturally suited to model diffusion data as they are antipodally symmet-  
 310 ric (i.e. the probability density is the same along an axis in either direction)  
 311 and as the aim here is to model the *PDF* of diffusion axes at correspond-  
 312 ing spatial locations, rather than any specific direction along the axes (Jupp  
 313 and Mardia, 1989). They are fully defined by two parameters, namely, the  
 314 mean/principal axis ( $\pm \mathbf{m}^d$ , about which the distribution is rotationally sym-  
 315 metric) and a scalar concentration parameter  $\kappa$ . The latter describes the  
 316 degree of concentration about the mean axis of the distribution, with high  
 317 values indicating high concentration. The *PDF* of a Watson distribution  
 318 with mean direction  $\mathbf{m}^d$  and concentration  $\kappa$  is expressed as equation 2a,  
 319 for antipodally symmetric 3D unit vectors  $\pm \mathbf{n}$ . Here,  $M(\cdot)$  represents the  
 320 Kummer function. Watsons are in general more flexible than Fisher distri-  
 321 butions as there is no positivity constraint on  $\kappa$  and they can be used to  
 322 model both directional and axial data. (Bijral et al., 2007) proposed an ef-

323 efficient EM-based clustering framework for axially-distributed data, using a  
 324 WMM, employed in this study to cluster fibre orientations.

$$p(\pm \mathbf{n} | \mathbf{m}^d, \kappa) = M\left(\frac{1}{2}, \frac{D}{2}, \kappa\right)^{-1} \exp^{\kappa(\mathbf{m}^{dT} \mathbf{n})^2} \quad (2a)$$

$$p(\mathbb{N} | \Theta_n) = \sum_{k=1}^K \sum_{i=1}^{N_k} \ln \sum_{j=1}^M \pi_j p(\pm \mathbf{n}_{ki} | \mathbf{m}_j^d, \kappa_j) \quad (2b)$$

325 The joint likelihood of the diffusion eigenvectors  $\pm \mathbf{n}_{ki} \in \mathbf{N}_k$  observed  
 326 across all  $N_k$  points in all  $K$  hybrid point sets, given Watson distributions  
 327 with mean directions and concentrations  $\{\mathbf{m}_j^d, \kappa_j\}_{j=1 \dots M} \in \Theta_n$ , is evaluated as  
 328 shown in equation 2b. Here,  $\mathbf{N}_k \in \mathbb{N}$  denotes the set of all observed diffusion  
 329 vectors across the entire population. It is important to note at this point  
 330 that, as the clustering of fibre orientations is initially performed jointly with  
 331 rigid registration of the hybrid point sets  $\mathbf{D}_k$ , the estimated rotations  $\mathbf{R}_k^{(t)}$  at  
 332 the  $t^{\text{th}}$  EM-iteration, are applied to the current estimate of the mean fibre  
 333 orientations  $\mathbf{m}_j^{d^{(t)}}$ , prior to the evaluation of the posterior probabilities  $P_{kij}$ ,  
 334 and concentrations  $\kappa_j$ , in the E- and M-steps, respectively. Additionally, for  
 335 the estimation of  $\mathbf{m}_j^d$  the inverse of the estimated rotations  $\mathbf{R}_k^T$  were applied  
 336 to their corresponding sample's diffusion eigenvectors  $\mathbf{n}_{ki}$ , to align the  $k^{\text{th}}$   
 337 sample to the current estimate of the mean template (refer to equation 3c).

$$Q(\Theta_n^{t+1} | \Theta_n^t) = \sum_{k=1}^K \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^{(t)} \ln p(\pm \mathbf{n}_{ki} | \mathbf{R}_k^{(t)} \mathbf{m}_j^{d^{(t)}}, \kappa_j^{(t)}) \quad (3a)$$

$$Q(\Theta_n^{t+1} | \Theta_n^t) = \sum_{k=1}^K \sum_{i=1}^{N_k} \sum_{j=1}^M [P_{kij}^{(t)} \ln p(\pm \mathbf{n}_{ki} | \mathbf{R}_k^{(t)} \mathbf{m}_j^d, \kappa_j) + \lambda_j (1 - (\mathbf{R}_k^{(t)} \mathbf{m}_j^d)^T \mathbf{R}_k^{(t)} \mathbf{m}_j^d)] \quad (3b)$$

$$\mathbf{m}_j^{d^{(t)}} - \frac{\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij}^{(t)} ((\mathbf{R}_k^{T(t+1)} \mathbf{n}_{ki})^T \mathbf{m}_j^{d^{(t)}}) \mathbf{R}_k^{T(t+1)} \mathbf{n}_{ki}}{\|\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij}^{(t)} ((\mathbf{R}_k^{T(t+1)} \mathbf{n}_{ki})^T \mathbf{m}_j^{d^{(t)}}) \mathbf{R}_k^{T(t+1)} \mathbf{n}_{ki}\|} = 0 \quad (3c)$$

$$\left[ \frac{M'(\kappa_j)}{M(\kappa_j)} \right]^{(t+1)} = \frac{\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij}^{(t)} (\mathbf{n}_{ki}^T \mathbf{m}_j^{d(t+1)})^2}{\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij}^{(t)}} \quad (3d)$$

$$\kappa_j^{(t+1)} \approx \frac{1}{2} \left[ \frac{1 - \left[ \frac{M'(\kappa_j)}{M(\kappa_j)} \right]^{(t+1)} D}{\left[ \left( \frac{M'(\kappa_j)}{M(\kappa_j)} \right)^2 \right]^{(t+1)} - \left[ \frac{M'(\kappa_j)}{M(\kappa_j)} \right]^{(t+1)}} \right] \quad (3e)$$

338 Maximum likelihood estimates for the associated parameters are evalu-  
 339 ated at each M-step of the algorithm by maximising the expectation of the  
 340 complete data likelihood (equation 3a), with respect to  $\mathbf{m}_j^d$  and  $\kappa_j$ , subject  
 341 to the constraint  $\mathbf{m}_j^{dT} \mathbf{m}_j^d = 1$  (Bijral et al., 2007). This is achieved by max-  
 342 imising the Lagrangian form of  $Q$  shown in equation 3b. Mean directions  $\mathbf{m}_j^d$   
 343 are estimated numerically, using fixed-point iteration, to solve the non-linear  
 344 equation (shown in equation 3c) obtained from differentiating  $Q$  (3b) with  
 345 respect to  $\mathbf{m}_j^d$ .  $\kappa_j$  on the other hand is approximated (refer to equation 3e)  
 346 using the continued fraction representation for the ratio of, the derivative  
 347 of the Kummer function and the function itself, i.e.  $\frac{M'(\kappa_j)}{M(\kappa_j)}$  (equation 3d).  
 348 In a recent study (Sra and Karp, 2013) derived two-sided bounds for ap-  
 349 proximating  $\kappa$ , particularly useful when dealing with high dimensional data.  
 350 However, for 3D data (as in this study) the approximation presented in equa-  
 351 tion 3e is sufficient (as noted by (Bijral et al., 2007), (Sra and Karp, 2013)).  
 352 Better approximations for  $\kappa_j$  may be obtained using numerical techniques  
 353 such as Newton’s method, however, at the expense of significant increase in  
 354 computational burden.

### 355 2.5. Mixture Model for Fractional Anisotropy

356 The distribution of voxel-wise FA in WM ROIs across a population, is  
 357 modelled using a univariate GMM. GMM was chosen as the resulting model-  
 358 predicted FA values at the estimated spatial correspondences, across subjects,  
 359 is guaranteed to be normally distributed — a useful property for subsequent  
 360 statistical analyses, as noted in (Smith et al., 2006), where the authors also  
 361 show that FA values at corresponding spatial positions across populations are  
 362 indeed approximately normally-distributed. Additionally, GMMs are com-  
 363 putationally efficient, as analytical solutions exist for revising estimates of  
 364 the associated model parameters (mean  $m_j^f$  and variance  $\sigma_j^2$  of FA), at each  
 365 EM-iteration. Assuming the observed FA values  $f_{ki}$  at voxels in ROIs, across



366 a group of subjects  $\mathbf{F}_k \in \mathbb{F}$  are i.i.d, the joint log-likelihood  $\log p(\mathbb{F}|\Theta_f)$ , is  
 367 expressed as equations 4a, 4b. Consequently, the conditional expectation of  
 368 the complete data log likelihood  $Q$ , maximised with respect to the model pa-  
 369 rameters associated with the Gaussian distributions in the mixture, is given  
 370 by equation 4c (only terms dependent on  $m_j^f$  and  $\sigma_f^2$  are retained in  $Q$ ). As  
 371 GMM-based clustering of FA values is performed jointly with the registra-  
 372 tion of WM ROIs, and clustering of voxel positions and the associated fibre  
 373 orientations, the influence of a Gaussian component in the mixture model  
 374 is automatically limited to its local neighbourhood. This helps ensure that  
 375 only voxels in close proximity to each other contribute significantly to the  
 376 estimation of mean FA values at each mixture component. Estimates for  
 377 the GMM parameters  $m_j^f$  and  $\sigma_f^2$  in the M-step of the algorithm are derived  
 378 analytically, as shown in (Bishop, 2006).

$$p(\mathbf{F}_k|m_j^f, \sigma_f^2) = \prod_{i=1}^{N_k} \sum_{j=1}^M \pi_j \mathcal{N}(f_{ki}|m_j^f, \sigma_f^2) \quad (4a)$$

$$\ln p(\mathbb{F}|\Theta_f) = \sum_{k=1}^K \ln p(\mathbf{F}_k|\Theta_f) \quad (4b)$$

$$Q(\Theta_f^{t+1}|\Theta_f^t) = -\frac{1}{2} \sum_{k=1}^K \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^t \left[ \frac{(f_{ki} - m_j^f)^2}{\sigma_f^2} \right] \quad (4c)$$

### 379 2.6. Rigid Alignment and Template Construction

380 Previously, we proposed a group-wise rigid point set registration frame-  
 381 work based on Student’s t-mixture model (Ravikumar et al., 2016), (Raviku-  
 382 mar et al., 2018), which exploits the inherent robustness of Student’s t-  
 383 distribution for robust registration of shapes in the presence of missing data  
 384 and significant proportions of outliers. Additionally, in a more recent study  
 385 (Ravikumar et al., 2017) we proposed a variant of the hybrid mixture model-  
 386 based registration framework formulated in this study. In (Ravikumar et al.,  
 387 2017) Von-Mises-Fisher distributions were used in place of the Watson distri-  
 388 butions used in this study, to model directional data such as surface normal  
 389 vectors, for rigid and non-rigid shape registration. A Watson distribution-  
 390 based variant of (Ravikumar et al., 2017) is employed in the present study  
 391 as an initial step, to rigidly align WM ROIs (hybrid point sets representing  
 392 voxel centroid positions, fibre orientations and FA values), segmented from

393 all subjects’ images, whilst simultaneously estimating a mean model. The  
 394 latter subsequently serves as an unbiased, study-specific template for non-  
 395 rigid registration. Rigid group-wise registration is preferred to a pair-wise  
 396 approach as it enables estimation of a mean template and the desired sim-  
 397 ilarity transformations in an unbiased manner. Rigid alignment also helps  
 398 initialise the subsequent non-rigid registration by recovering global differ-  
 399 ences in pose between sample shapes, and establishes soft correspondences  
 400 across subjects.

401 Group-wise point set registration using mixture models assumes that the  
 402 point sets to be aligned are transformed observations of a central mixture  
 403 model (which we refer to as the mean template) (Gooya et al., 2015). Con-  
 404 sequently, the optimal transformations that align the template to the group  
 405 of shapes are those that maximise the likelihood of the data (or equivalently,  
 406 minimise the negative log-likelihood function). The desired similarity trans-  
 407 formations are thus iteratively refined along with the template itself at each  
 408 M-step of the algorithm. The main differences between EM-based estima-  
 409 tion of parameters for TMMs and GMMs are: (1) TMMs have two associated  
 410 latent variables (as opposed to just one with GMMs, which represent the mix-  
 411 ture component membership of the data), whose expectations are evaluated  
 412 in the E-step and used to compute a set of corrected posterior probabilities  
 413  $P_{kij}^*$ , estimated identically to (Ravikumar et al., 2016), (Ravikumar et al.,  
 414 2018) (refer to the Appendix); and (2) Student’s t-distributions are defined  
 415 by three parameters (as opposed to two for Gaussians). The additional pa-  
 416 rameter is referred to as the degrees of freedom/shape parameter  $\nu$ , which is  
 417 responsible for controlling the heaviness of the tails of the distribution (and  
 418 consequently, the degree of robustness to outliers). The behaviour of the  
 419 t-distribution tends towards that of a Gaussian as  $\nu \rightarrow \infty$ .

$$\log p(\mathbb{X}|\Theta_p, \mathbb{T}) = \sum_{k=1}^K \sum_{i=1}^{N_k} \log \sum_{j=1}^M \pi_j \mathcal{S}(\mathbf{x}_{ki}|\mathbf{T}_k(\mathbf{m}_j^p), \sigma_p^2, \nu_j) \quad (5a)$$

$$Q(\Theta_p^{t+1}, \mathbb{T}^{t+1}|\Theta_p^t, \mathbb{T}^t) \propto -\frac{1}{2\sigma_p^2} \sum_{k=1}^K \sum_i^{N_k} \sum_{j=1}^M P_{kij}^{*t} \|\mathbf{x}_{ki} - s_k \mathbf{R}_k \mathbf{m}_j^p - \mathbf{b}_k\|^2 \quad (5b)$$

420 The joint *PDF* of voxel positions  $\mathbf{x}_{ki} \in \mathbf{X}_k$ , across all  $K$  subjects in  
 421 a group (denoted,  $\mathbf{X}_k \in \mathbb{X}$ ), is given by equation 5a (assuming they are  
 422 i.i.d transformed observations of a TMM). In equation 5a,  $\mathbf{T}_k$  represents the

423 similarity transformation (comprising rotation  $\mathbf{R}_k$ , scaling  $s_k$  and translation  
 424  $\mathbf{b}_k$ ), to align the positions  $\mathbf{m}_j^p$  defining the mean template, to the  $k^{\text{th}}$  sample  
 425 in the group. In our recent work (Ravikumar et al., 2016), (Ravikumar et al.,  
 426 2018), we showed that the form of  $Q$  to be maximised, to estimate the desired  
 427 similarity transformations  $\mathbf{T}_k \in \mathbb{T}$  and mixture component parameters  $\Theta_p$ ,  
 428 is given by equation 5b. Closed form expressions are derived for the M-  
 429 step update equations of all TMM and transformation parameters, which  
 430 are presented in the Appendix. Fibre orientations and FA are invariant to  
 431 translation  $\mathbf{b}_k$  and scaling  $s_k$ , consequently, these transformation parameters  
 432 are estimated identically as in (Ravikumar et al., 2016), (Ravikumar et al.,  
 433 2018). Although the former are rotationally dependent, the contribution of  
 434 fibre orientations to the estimation of  $\mathbf{R}_k$  is ignored as the direction of the  
 435 observed diffusion eigenvectors tend to be random. Consequently, rotations  
 436  $\mathbf{R}_k$  are derived based on the spatial positions of hybrid point sets alone,  
 437 by optimising the form of  $Q$  shown in equation 5b, similar to (Ravikumar  
 438 et al., 2016), (Ravikumar et al., 2018). However, following estimation of  
 439 the desired rotations  $\mathbf{R}_k$  at each EM-iteration, the current estimate of the  
 440 mean template is transformed by rotating both spatial positions  $\mathbf{m}_j^p$  and  
 441 their associated fibre orientations  $\mathbf{m}_j^d$ , to align it with the  $k^{\text{th}}$  sample in the  
 442 group. Additionally, it is important to note that, while the fibre orientations  
 443 and FA values are ignored in the derivation of the desired transformation  
 444 parameters, they are intrinsic to the estimation of the posterior probabilities  
 445  $P_{kij}$  at each E-step of the algorithm. Consequently, they drive the estimation  
 446 of soft correspondences, which in turn affect the transformations evaluated  
 447 at each M-step of the algorithm.

### 448 2.7. Non-rigid Point Set Registration

449 Coherent point drift (CPD) (Myronenko and Song, 2010) is a well known  
 450 pair-wise, non-rigid point set registration technique based on motion coher-  
 451 ence theory. The spatial transformation between two point sets is considered  
 452 to be an initial position (of the moving point set) plus some unknown dis-  
 453 placement (or velocity) function mapping it to the target point set. This un-  
 454 known transformation is regularized using Tikhonov regularization, to ensure  
 455 estimation of a smooth displacement function, and is expressed in the Repro-  
 456 ducing Kernel Hilbert Space (RKHS). Using variational calculus, Myronenko  
 457 and Song (2010) showed that the optimal displacement function under such  
 458 smoothness constraints, can be expressed as a linear combination of kernel  
 459 functions (i.e. Gaussian radial basis functions). Similarly, our approach also

460 employs Gaussian radial basis functions to parametrize the non-linear trans-  
 461 formation, and the associated basis function weights are estimated by max-  
 462 imising the likelihood function using EM (similar to estimation of rotation,  
 463 translation and scaling, in the rigid registration approach discussed in the  
 464 previous section). CPD models the target point set as a transformed obser-  
 465 vation of the source point set (i.e. the point set to be registered). The latter  
 466 is consequently considered to represent the centroids of a Gaussian mixture  
 467 model, which is fit to the former using EM, and the transformation necessary  
 468 to register the source to the target set is estimated as parameters of the mix-  
 469 ture model. In addition to the Gaussian components in the mixture model,  
 470 CPD incorporates a uniform distribution component to model noise/outliers  
 471 present in the data. This confers added robustness to the registration pro-  
 472 cess. However, a user-defined parameter is used to balance the weight of the  
 473 uniform distribution component relative to its Gaussian counterparts, which  
 474 needs to be tuned for different applications and data sets, for optimal regis-  
 475 tration. To ameliorate the need for parameter tuning, we employ Student’s  
 476 t-distributions in place of the Gaussian and uniform distributions used in  
 477 CPD and re-formulate the approach in a group-wise non-rigid registration  
 478 framework. As stated previously, the robust nature of t-distributions makes  
 479 them well suited to registration applications requiring automatic robustness  
 480 to outliers. A similar approach for pair-wise registration of 2D/3D point sets  
 481 was proposed previously, by (Zhou et al., 2014).

482 The mean tract template estimated during the initial group-wise rigid  
 483 registration step (discussed in section 2.6), is non-rigidly registered to each  
 484 patient group (AD, MCI and HC) independently. The desired non-rigid  
 485 transformations are defined with respect to the template  $\mathcal{M}$  as:  $\mathbf{M} + v^k(\mathbf{M})$   
 486 (considering spatial positions  $\mathbf{m}_j^p$  alone), where  $v$  is a displacement func-  
 487 tion mapping the template to the  $k^{\text{th}}$  sample in the group. In (Myronenko  
 488 and Song, 2010) the authors show that the desired displacement field is con-  
 489 strained to be smooth by employing Tikhonov regularization (or regularizing  
 490 the norm of  $v$ , expressed in RKHS). This forces points in close proximity, to  
 491 move together. Regularization of this nature is akin to employing a prior on  
 492 the displacement field of the form  $p(v) = \exp^{-\frac{\lambda}{2}\phi(v)}$ , where  $\phi(v)$  represents the  
 493 regularization term and  $\lambda$  controls the trade-off between registration accuracy  
 494 and smoothness of the deformation field. The prior on the displacement field  
 495 is incorporated into the TMM, resulting in a log-likelihood function expressed  
 496 as equation 6a. As stated previously, (Myronenko and Song, 2010) show that

497 the function  $v$ , which maximises the data likelihood, can be expressed as a  
498 linear combination of radial basis functions (refer to equation 6b). Conse-  
499 quently, to register the study-specific mean template to each sample from  
500 all patient groups simultaneously, the objective function to be maximised  
501 with respect to the basis function weights  $w_{kj} \in \mathbf{W}_k$ , is expressed as shown  
502 in equation 6c, where  $\mathbf{G}$  represents the Gaussian kernel/Gram matrix. The  
503 basis function weights required to register the study-specific mean template  
504 to each sample are estimated as shown in 6d, by computing the derivative of  
505  $Q$  with respect to the weights, similarly to (Myronenko and Song, 2010). In  
506 equation 6d  $\mathbf{P}_k^s = \sum_{i=1}^{N_k} P_{kij}^{*t}$ ,  $\mathbf{P}_k^T$  is the transpose of the posterior probability  
507 matrix for the  $k^{\text{th}}$  sample,  $\text{diag}$  is a diagonal matrix, and  $\mathbf{I}$  is the identity  
508 matrix. Subsequently, the mean template is deformed to match each  $k^{\text{th}}$   
509 sample (in the entire population) as described by equation 6e.

$$\log p(\mathbb{X}|\Theta_p) = \sum_{k=1}^K \sum_{i=1}^{N_k} \log \sum_{j=1}^M \pi_j \mathcal{S}(\mathbf{x}_{ki}|v^k(\mathbf{m}_j^p), \sigma^2, \nu_j) + \frac{\lambda}{2} \phi(v^k) \quad (6a)$$

$$v^k(\mathbf{q}) = \sum_{j=1}^M w_{kj} G(\mathbf{q} - \mathbf{m}_j^p) \quad (6b)$$

$$\begin{aligned} Q(\Theta_p^{t+1}, \mathbf{W}_k^{t+1}|\Theta_p^t, \mathbf{W}_k^t) = \\ -\frac{1}{2\sigma_p^2} \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij}^{*t} \|\mathbf{x}_{ki} - (\mathbf{m}_j^p + v^k(\mathbf{m}_j^p))\|^2 + \frac{\lambda}{2} \mathbf{W}_k^T \mathbf{G} \mathbf{W}_k \end{aligned} \quad (6c)$$

$$\mathbf{W}_k^{(t+1)} = [\text{diag}(\mathbf{P}_k^{s^t}) \mathbf{G} + \lambda \sigma_p^2 \mathbf{I}]^{-1} [\mathbf{P}_k^{T^t} \mathbf{X}_k - \text{diag}(\mathbf{P}_k^{s^t}) \mathbf{M}^t] \quad (6d)$$

$$\mathbf{M}_k^{(t+1)} = \mathbf{T}_k^t(\mathbf{M}_k^t, \mathbf{W}_k^t) = \mathbf{M}_k^t + \mathbf{G} \mathbf{W}_k^t \quad (6e)$$

510 Following convergence of the non-rigid registration step, a study-specific  
511 mean template comprising, mean spatial positions, mean fibre orientations  
512 and mean FA values (representative of the entire population of AD, MCI and  
513 HC subjects), is estimated. Additionally, point-wise displacements mapping  
514 this mean template to each sample in the entire population (as described

515 by equation 6e), is also obtained, thereby establishing the spatial correspon-  
516 dences used for any subsequent inter-group statistical comparisons. These  
517 correspondences play a similar role to the mean FA skeleton estimated in  
518 TBSS. In addition to these spatial correspondences, we also compute “model-  
519 predicted” values for FA and fibre orientation, at each correspondence, for all  
520 subjects. These model-predicted values are probabilistic weighted averages  
521 of the FA values and fibre orientations associated with the voxels in the orig-  
522 inal DTI-derived FA and eigenvector images (i.e. the original hybrid point  
523 sets). The weighted averages are assigned to each spatial correspondence  
524 point and are analogous to the ‘soft/probabilistic spatial correspondences’  
525 estimated in previous studies, such as (Hufnagel et al., 2008), (Gooya et al.,  
526 2015) for example. Here, the weights are defined by the posterior probabil-  
527 ities estimated for each voxel, of each subject’s original FA and eigenvector  
528 images ( $P_{kij}$ ), following non-rigid registration. Equations describing the es-  
529 timation of model-predicted FA values and fibre orientations are included in  
530 the Appendix (refer to equations 19a - 19b). Although point set registration  
531 techniques are typically employed to register 3D point sets (comprising only  
532 spatial positions) representing the surface/boundary of an object, this study  
533 incorporates additional image-based features (such as fibre orientations and  
534 FA values), that enable registration of dense point sets, defined by voxel  
535 centroids located at the boundary of, and within a region of interest.

### 536 3. Results and Discussion

#### 537 3.1. Rigid Registration Accuracy

538 Rigid registration accuracy of the proposed framework and the robust-  
539 ness of Student’s t-distributions to outliers is assessed using synthetic data  
540 comprising point sets containing positions, associated fibre orientations and  
541 FA values. The synthetic data set was generated by rigidly transforming  
542 a corpus callosum hybrid point set by varying amounts. Four distinct syn-  
543 thetic samples (Samples 1-4) were generated in this manner from the original  
544 ground truth point set (referred to as Sample 0), as illustrated by Fig. 3.

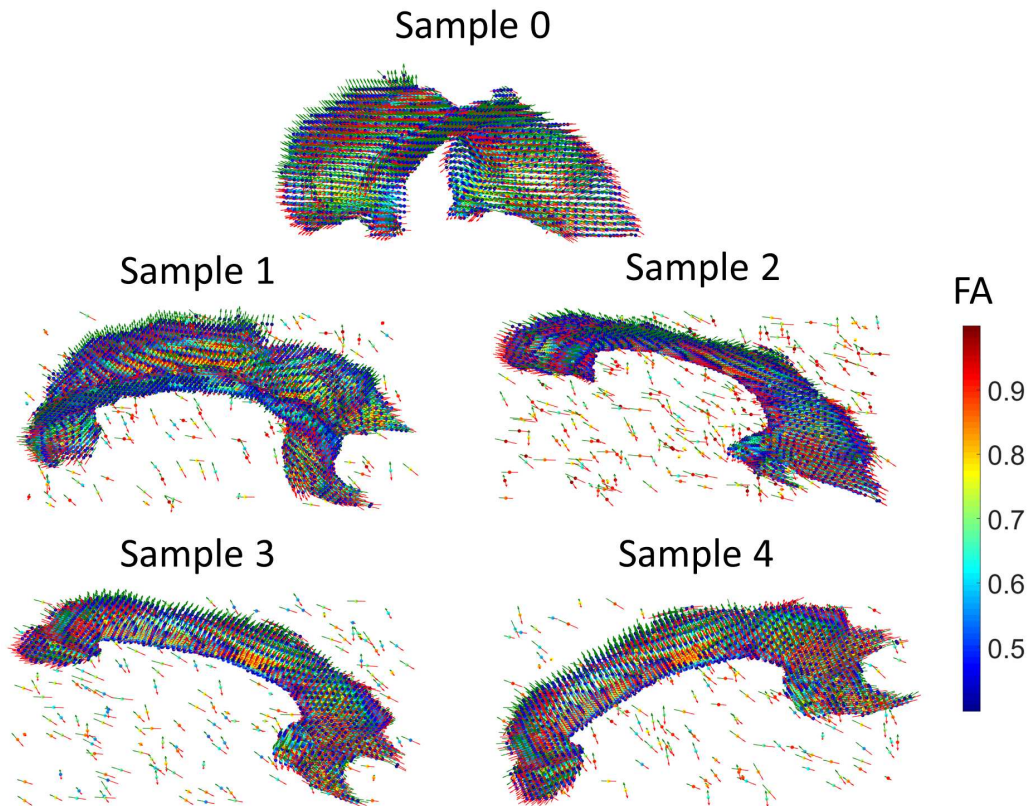


Figure 3: Synthetic corpus callosum data set comprising: Sample 0, the ground truth hybrid point set; and Samples 1-4, which are rotated and modified versions of Sample 0.

545 The rigidly transformed point sets were also modified by the addition of  
 546 varying proportions of random outliers (comprising positions, orientations  
 547 and FA values). Fibre orientations associated with the outliers were gener-  
 548 ated from normalized 3D points. While their FA values were uniformly  
 549 sampled within the range  $[0.2, 0.8]$ . The FA values associated with the voxels  
 550 of each modified hybrid point set were also varied by  $\pm 0.1$ , relative to the  
 551 ground truth point set. This was necessary in order to emulate real data  
 552 as FA values typically vary at corresponding anatomical locations, between  
 553 subjects. This process was repeated 10 times, to generate 10 unique syn-  
 554 thetic data sets (each comprising one ground truth and 4 modified, unique  
 555 samples), which were subsequently rigidly aligned using the proposed Wat-  
 556 son distribution-based HdMM algorithm (i.e. 10 distinct registration exper-

557 iments). Random rotations and proportions of outliers were generated for  
 558 each experiment, within the range of  $[-30^\circ, 30^\circ]$  and  $[2\%, 5\%]$ , respectively  
 559 (as illustrated in Fig. 3). Table 1 summarises the mean ground truth eu-  
 560 clidean distances between Samples 1-4 and Sample 0 across all 10 experiments  
 561 (prior to registration), and the axes about which rotations were applied to  
 562 generate each sample in each experiment. The average rigid registration er-  
 563 rors following alignment of the synthetic data sets (with  $M = 2000$  mixture  
 564 components) using the proposed framework are also reported in Table 1.

565 Rigid registration accuracy was evaluated by: (a) computing the in-  
 566 trinsic distance between the estimated and ground truth rotations (Huynh,  
 567 2009), for easy interpretation of the rotation errors ( $\theta_{err}$ ), in degrees (refer to  
 568 equation 7); and (b) computing the mean Euclidean distance (ED) between  
 569 (transformed) Samples 1-4 and Sample 0 (averaged across all points). Ta-  
 570 ble 1 summarises average rotation and Euclidean distance errors (computed  
 571 across all 10 experiments). Point-wise Euclidean distances are first evaluated  
 572 between each modified sample (Samples 1-4) and Sample 0, following rigid  
 573 registration, and subsequently averaged across all points. The resulting mean  
 574 Euclidean distance is then averaged once again across all 10 experiments and  
 575 is reported in Table 1.

$$\theta_{err} = \arccos \left[ \frac{\text{tr}((\mathbf{R}_k^g (\mathbf{R}_k \mathbf{R}_1^T)^T) - 1)}{2} \right] \quad (7)$$

Table 1: Summary of rigid registration errors across 10 experiments using synthetic corpus callosum data sets.

Sample #	Rotated Around	Ground Truth Euc. Dist. (mm.)	Rot. Err. (degrees)	Euc. Dist. (mm.)
1	x,y	43.57 ± 19.85	0.06 ± 0.03	0.34 ± 0.15
2	y,z	42.85 ± 13.12	0.05 ± 0.03	0.30 ± 0.16
3	z,x	42.77 ± 8.74	0.04 ± 0.03	0.23 ± 0.13
4	x,y,z	35.52 ± 17.19	0.04 ± 0.03	0.25 ± 0.17



576 The average Euclidean distance errors reported in Table 1 indicate that  
577 the proposed Watson-based HdMM framework achieved very low errors (de-  
578 spite the presence of random outliers) as all values are substantially lower  
579 than the voxel size of the original eigenvector and FA image (refer to section  
580 2.1), from which the ground truth corpus callosum hybrid point set (sample  
581 0) was generated. Robustness to outliers may be attributed to the con-  
582 stituent t-distributions in the HdMM, modelling spatial positions. Similarly  
583 the proposed approach was also able to accurately recover the applied ground  
584 truth rotations, resulting in very low rotation errors for all samples (as shown  
585 in Table 1), relative to the magnitude of the rotations applied to generate  
586 the synthetic data set. The proposed approach therefore, is considered to  
587 successfully approximate the joint density of position, fibre orientation and  
588 FA, for the synthetic corpus callosum data set, and accurately recover the  
589 applied rigid transformations.

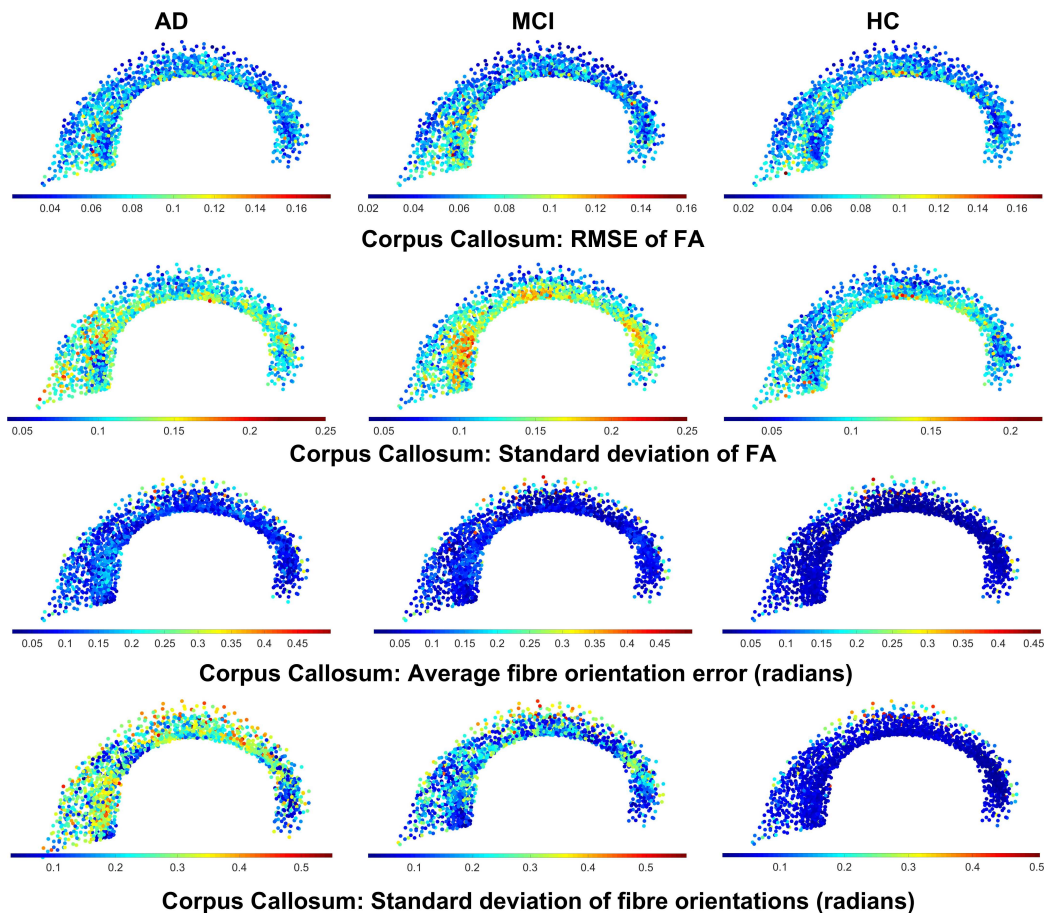


Figure 4: Model quality evaluated for the corpus callosum, independently for AD, MCI and HC groups, using  $M = 2000$  mixture components. Rows one and two: RMSE of FA and standard deviations of the same computed across subjects; Rows three and four: Angular errors for fibre orientations (in radians) and standard deviations of the same computed across subjects.

591 The ability of the HdMM to model DTI-derived quantities was assessed  
 592 using clinical data, acquired from the VPH-DARE@IT prospective cohort,  
 593 described in section 2.1. Specifically, model quality was quantified by eval-  
 594 uating the similarity between the estimated correspondences (resulting from  
 595 non-rigidly registering the the unbiased study-specific mean template to each

596 sample from all patient groups) and the nearest neighbour voxels in the corre-  
597 sponding subject’s original FA and eigenvector images. FA accuracy is quan-  
598 tified as the root-mean-squared error (RMSE), evaluated between the model-  
599 predicted and original voxel-wise FA values, across all correspondences, for  
600 each subject. The group-wise average error (for each subject group) of FA  
601 was subsequently computed. The minimum arc length (measured in radians)  
602 between two unit vectors is used to measure the accuracy of local fibre orien-  
603 tation in a similar manner. As discussed in section 2.4, the proposed frame-  
604 work models axial data rather than directional data. When computing fibre  
605 orientation errors, corresponding unit vectors between the model-predicted  
606 and original voxel-wise eigenvectors are first identified. This is achieved by  
607 evaluating their scalar product and ensuring it is positive — i.e. if the dot  
608 product is negative, the antipodal counterpart of the model-predicted vector  
609 is used instead. The resulting measure thus quantifies the angular error in  
610 fibre orientation between the model-predicted and original voxel-wise data  
611 (in the eigenvector image), for each subject. These measures represent reg-  
612 istration residuals which describe the quality of correspondences established  
613 by the proposed HdMM (i.e. how well the HdMM can model the observed  
614 DTI-derived data), and only indirectly reflect registration ‘accuracy’. To pro-  
615 vide a more general view of registration accuracy, the mean-squared distance  
616 (MSD, formulated as shown in the Appendix), quantifying spatial position  
617 errors was also evaluated between the registered study-specific mean template  
618 and the original hybrid point sets from all patient groups (Note: MSD values  
619 were evaluated between dense volumetric point sets). It is important to note  
620 that the model-predicted values for FA and fibre orientation assigned to the  
621 spatial correspondences established using the proposed approach, are proba-  
622 bilistic in nature (as discussed in section 2.7). Consequently, they reflect the  
623 DTI-derived quantities of voxels located in the local spatial neighbourhood  
624 of the correspondences.

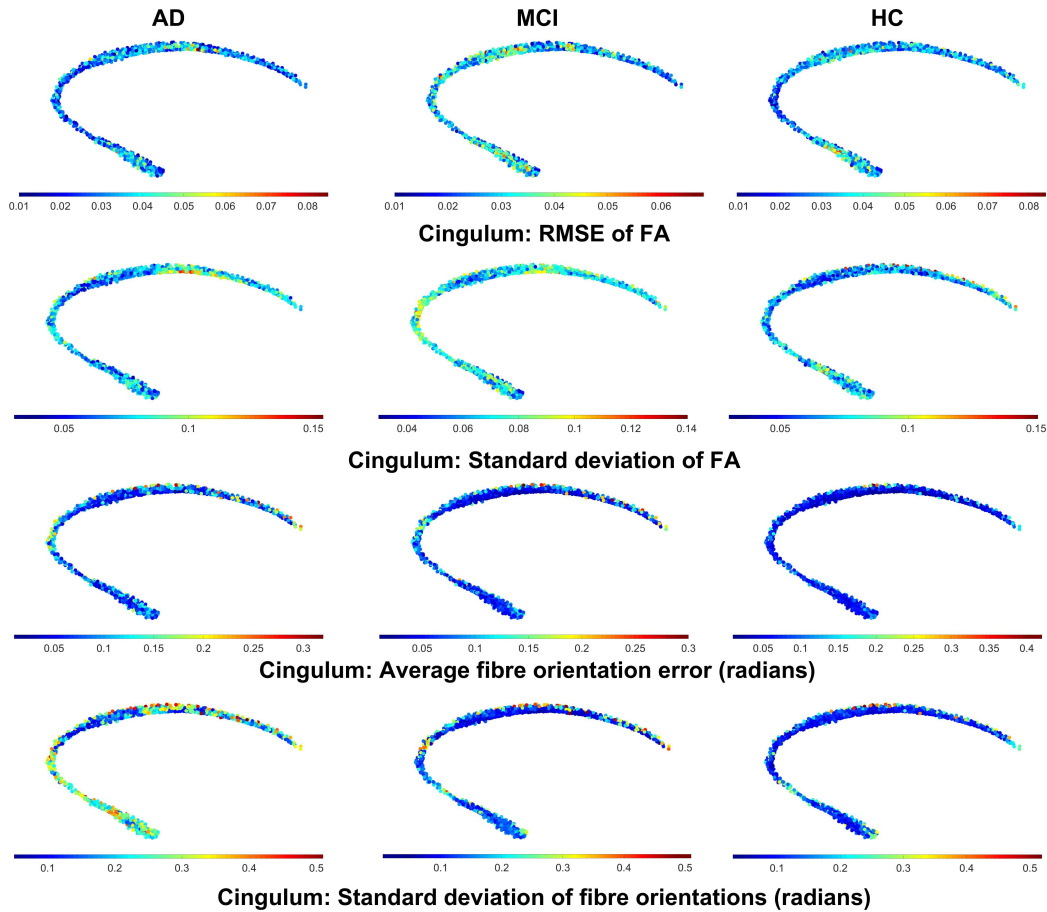


Figure 5: Model quality evaluated for the cingulum, independently for AD, MCI and HC groups, using  $M = 1500$  mixture components. Rows one and two: RMSE of FA and standard deviations of the same computed across subjects; Rows three and four: Angular errors for fibre orientations (in radians) and standard deviations of the same computed across subjects.

625 Results summarizing the ability of the proposed framework to model DTI-  
 626 derived quantities across all 60 subjects are presented in Fig. 4 - Fig. 7 and  
 627 Tables 2 - 7. Fig. 4 and Fig. 5 help visualise the spatial distribution of mean  
 628 registration errors and the standard deviations of FA values and fibre ori-  
 629 entations observed across subjects within each patient group, for the corpus  
 630 callosum and cingulum, respectively. We would like to highlight that while  
 631 samples from all patient groups were registered simultaneously, the registra-

632 tion errors presented in Fig. 4 - Fig. 7 and Tables 2 - 7 alone were evaluated  
633 for each patient group separately. This was done in order to identify any  
634 group-specific trends that exist in the registration accuracy afforded by the  
635 proposed approach. In Fig. 4 and Fig. 5 the RMSE values of FA were  
636 computed by averaging across subjects in each group, at each corresponding  
637 position. Similarly, the standard deviations were also evaluated point-wise  
638 across subjects for each group. The depicted mean angular errors were av-  
639 eraged across subjects, quantifying the fibre orientation accuracy at each  
640 corresponding position, and point-wise estimates for the standard deviations  
641 in fibre orientation were also evaluated. The presented standard deviations  
642 in Fig. 4 and Fig. 5 aid in interpretation of the error measures evaluated,  
643 and provide a frame of reference, for both WM regions. The spatial dis-  
644 tribution of the variation in FA across subjects within each patient group,  
645 was evaluated as follows: (a) the nearest neighbour voxel in the original hy-  
646 brid point sets were first identified based on the spatial positions estimated  
647 by non-rigid registration of the study-specific mean template, to each cor-  
648 responding sample ; (b) the FA values associated with the voxels identified  
649 for each subject were in turn used to compute the standard deviation across  
650 subjects, within each patient group; and (c) these values were subsequently  
651 mapped on to the study-specific mean template estimated for the corpus cal-  
652 losum and cingulum, for easy comparison with the registration errors plotted  
653 in a similar manner, as shown in Fig. 4. Similarly, the standard deviations  
654 in fibre orientations about the mean, were also evaluated across subjects,  
655 within each patient group, for both WM regions. Here, the difference be-  
656 tween the mean fibre orientation estimated at each correspondence point in  
657 the study-specific mean template, and the nearest neighbour voxels identified  
658 (refer to (a) above) in the original hybrid point sets, was evaluated as the  
659 minimum arc length (in radians) between each other. This in turn was em-  
660 ployed to compute the standard deviation in fibre orientations and visualize  
661 their spatial distribution across both WM regions.

662 Based on these results, the proposed HdMM is considered to establish  
663 valid correspondences across patients, as the estimated fibre orientation and  
664 FA errors are low across the majority of correspondences. Fibre orientation  
665 errors were consistently  $< 0.2$  radians across most correspondences for both  
666 WM ROIs (refer to first and third row in Fig. 4). FA errors meanwhile, were  
667  $< 0.1$  for the corpus callosum and cingulum (refer to second and fourth row in  
668 Fig. 4), across all patient groups. For the former WM region, FA errors below  
669  $0.1$  were produced for  $> 92\%$  of all established correspondences. While for

670 the latter, all correspondences, had errors below 0.1. Fibre orientation errors  
671 were  $< 0.2$  across  $> 94\%$  of correspondences estimated for both WM ROIs,  
672 in all patient groups. Errors of this magnitude are considered reasonable as  
673 the model-predicted FA values and fibre orientations evaluated at correspon-  
674 dences are based on the soft-assignment approach (refer to section 2.7), using  
675 the estimated posterior probabilities. Consequently, they reflect weighted av-  
676 erages of FA and fibre orientations of neighbouring voxels. FA variations of  
677  $\approx 0.1$  may occur due to partial volume effects at WM-GM and WM-CSF  
678 interfaces (Smith et al., 2006), particularly when WM tracts/ROIs are very  
679 thin compared to the voxel size (often the case following dementia-related  
680 atrophy of brain tissue), potentially further contributing to the observed er-  
681 rors. Additionally, significant variations in DTI-data in a select few cases  
682 within individual patient groups may be another source of the high average  
683 errors evaluated, in a small proportion of correspondences. These results are  
684 further supported by the standard deviations of FA and fibre orientations  
685 depicted in Fig. Fig. 4 and Fig. 5, which highlight the high degree of vari-  
686 ation in FA and fibre orientations (across subjects), respectively, relative to  
687 the corresponding errors evaluated across both WM regions.

688 These results are further verified by the histograms of errors in fibre  
689 orientation and FA presented in Fig. 6 and 7, respectively, summarising the  
690 correspondence-wise errors evaluated for each subject in the population. In  
691 this case, fibre orientation errors were computed as in preceding experiments,  
692 while FA errors were evaluated as the root-squared-error (RSE) between  
693 the model-predicted values and the closest voxels in the corresponding FA  
694 images. In general, high errors occur at only a few correspondences, across  
695 both the cingulum and corpus callosum. Registration errors for the AD and  
696 MCI groups were higher than for the HC group for both ROIs. This is  
697 attributed to the presence of varying degrees of pathology-induced changes  
698 in a few subjects in these groups, verified by Figs. 6 and 7, and by computing  
699 region-wise mean and standard deviations of FA and fibre orientation errors,  
700 presented in Tables 3 - 7.

701 Tables 2 - 7 report the average spatial position, fibre orientation and FA  
702 errors evaluated across correspondences and subjects. Statistically signifi-  
703 cant reduction in mean spatial position errors across experiments conducted  
704 using differing model complexities (i.e. different number of mixture compo-  
705 nents) are highlighted in bold in Tables 2 and 5, considering a significance  
706 level of 5%. In Tables 4 and 7 the reported mean FA errors were estimated  
707 by first computing the RMSE, this time averaging across correspondences,

708 and subsequently computing the mean RMSE across subjects. Tables 3 and  
709 6 summarise the mean angular error values, first averaged across correspon-  
710 dences and subsequently across subjects. These alternate error measures are  
711 presented to assess model quality of the HdMM across regions, and comple-  
712 ment the correspondence-wise errors presented in Fig. 4 - 5. From Tables 2  
713 - 7, the number of mixture components required to adequately characterise  
714 the entire population was identified as  $M = 1500$  and  $M = 2000$  for the cin-  
715 gulum and corpus callosum, respectively. The fibre orientation and FA errors  
716 depicted in Fig. 4 - 5 were evaluated using these values. All subsequent inter-  
717 group statistical analyses conducted employed these model complexities for  
718 the respective WM regions.

Table 2: Model quality of HdMM for the cingulum, assessed in terms of the mean spatial position error evaluated across correspondences and subjects, using the MSD metric, for each patient group, and for varying model complexities. Bold values indicate statistically significant reduction in errors.

# Mixture Components	Spatial Position Error: MSD (mm.)		
	AD	MCI	HC
500	0.86 ± 0.11	0.84 ± 0.09	0.82 ± 0.09
1000	0.73 ± 0.10	0.72 ± 0.08	0.71 ± 0.08
1500	<b>0.67 ± 0.09</b>	<b>0.66 ± 0.07</b>	<b>0.64 ± 0.07</b>
2000	<b>0.65 ± 0.09</b>	<b>0.63 ± 0.07</b>	<b>0.62 ± 0.07</b>

Table 3: Model quality of HdMM for the cingulum, assessed as the mean fibre orientation error evaluated across correspondences and subjects, for each patient group, and for varying model complexities.

# Mixture Components	Mean Fibre Orientation Error (radians)		
	AD	MCI	HC
300	0.11 ± 0.10	0.08 ± 0.02	0.07 ± 0.02
600	0.09 ± 0.08	0.07 ± 0.02	0.06 ± 0.01
1200	0.09 ± 0.08	0.06 ± 0.01	0.06 ± 0.01
1500	0.08 ± 0.08	0.06 ± 0.01	0.05 ± 0.01

Table 4: Model quality of HdMM for the cingulum, assessed as the average RMSE of FA evaluated over correspondences and averaged across subjects, for each patient group, and for varying model complexities.

# Mixture Components	Mean RMSE of FA		
	AD	MCI	HC
300	$0.06 \pm 0.01$	$0.06 \pm 0.01$	$0.06 \pm 0.01$
600	$0.06 \pm 0.01$	$0.06 \pm 0.01$	$0.06 \pm 0.01$
1200	$0.05 \pm 0.01$	$0.05 \pm 0.01$	$0.05 \pm 0.01$
1500	$0.05 \pm 0.01$	$0.05 \pm 0.01$	$0.05 \pm 0.01$

Table 5: Model quality of HdMM for the corpus callosum, assessed in terms of the mean spatial position error evaluated across correspondences and subjects, using the MSD metric, for each patient group, and for varying model complexities. Bold values indicate statistically significant reduction in errors.

# Mixture Components	Spatial Position Error: MSD (mm.)		
	AD	MCI	HC
500	$1.15 \pm 0.17$	$1.14 \pm 0.10$	$1.09 \pm 0.12$
1000	$0.99 \pm 0.15$	$0.98 \pm 0.09$	$0.94 \pm 0.10$
1500	<b><math>0.91 \pm 0.13</math></b>	$0.90 \pm 0.08$	<b><math>0.85 \pm 0.09</math></b>
2000	<b><math>0.86 \pm 0.12</math></b>	<b><math>0.85 \pm 0.07</math></b>	<b><math>0.81 \pm 0.08</math></b>

Table 6: Model quality of HdMM for the corpus callosum, assessed as the mean fibre orientation error evaluated across correspondences and subjects, for each patient group, and for varying model complexities.

# Mixture Components	Mean Fibre Orientation Error (radians)		
	AD	MCI	HC
500	$0.13 \pm 0.19$	$0.10 \pm 0.14$	$0.06 \pm 0.01$
1000	$0.13 \pm 0.19$	$0.13 \pm 0.16$	$0.05 \pm 0.01$
1500	$0.12 \pm 0.19$	$0.09 \pm 0.13$	$0.05 \pm 0.01$
2000	$0.12 \pm 0.18$	$0.09 \pm 0.13$	$0.05 \pm 0.01$



Table 7: Model quality of HdMM for the corpus callosum, assessed as the average RMSE of FA evaluated over correspondences and averaged across subjects, for each patient group, and for varying model complexities.

# Mixture Components	Mean RMSE of FA		
	AD	MCI	HC
500	$0.11 \pm 0.03$	$0.11 \pm 0.02$	$0.10 \pm 0.01$
1000	$0.10 \pm 0.03$	$0.10 \pm 0.02$	$0.09 \pm 0.01$
1500	$0.09 \pm 0.03$	$0.09 \pm 0.03$	$0.08 \pm 0.004$
2000	$0.09 \pm 0.03$	$0.08 \pm 0.03$	$0.07 \pm 0.01$

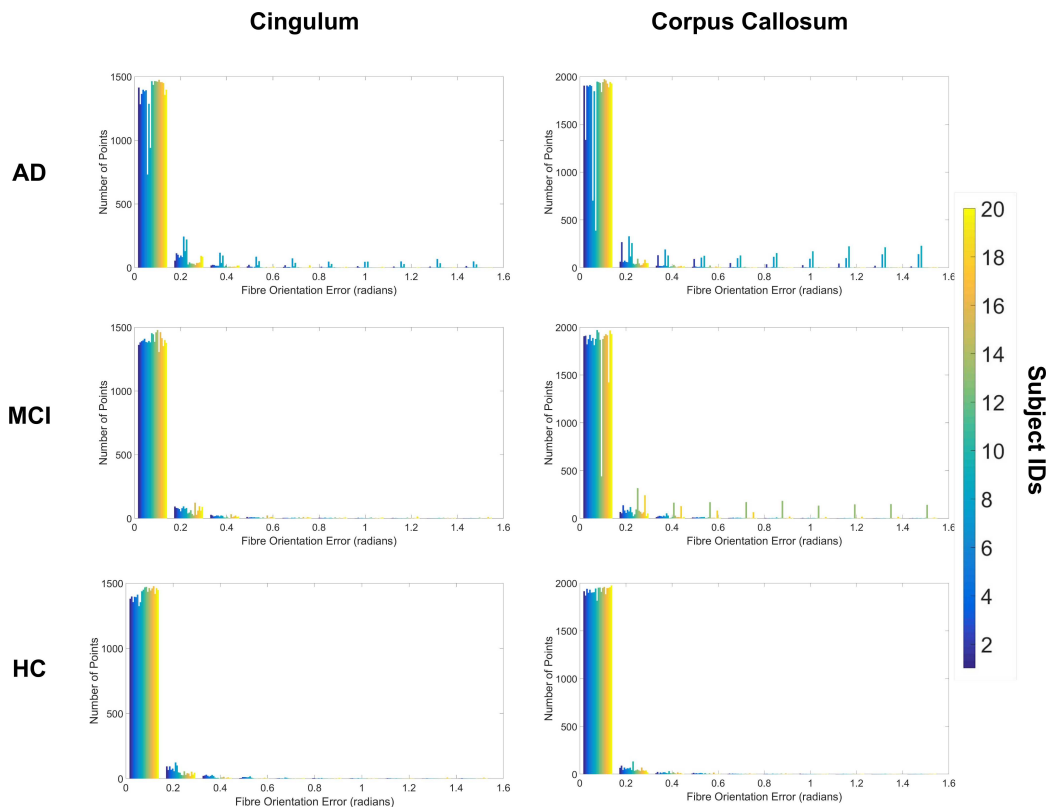


Figure 6: Histograms of fibre orientation errors for each subject in AD, MCI and HC groups, evaluated between established correspondences and ground truth voxels.

719 Results in Fig. 6 and 7 indicate that the proposed framework achieves  
 720 low fibre orientation and FA errors at each estimated correspondence, for all

721 subjects in the HC group (for both WM ROIs). The estimated correspon-  
722 dences were less accurate for two cases in the AD group (for both cingulum  
723 and corpus callosum) and for one case in the MCI group (only corpus cal-  
724 losum), which is attributed to significant variation in fibre orientations and  
725 FA values in these cases and ROIs, relative to the remaining samples in their  
726 corresponding patient groups. As discussed previously, this may be a re-  
727 sult of varying degrees of pathology-induced changes in these cases relative  
728 to the rest of their group. Consequently, the accuracy of the HdMM when  
729 fitting to these few cases, is reduced. The proposed framework, however,  
730 established accurate correspondences for the remaining samples in the AD  
731 and MCI groups across both WM ROIs. The high deviations from the mean  
732 fibre orientation errors in the corpus callosum for these groups (Table 6)  
733 are thus attributed to the outlier subjects identified from the corresponding  
734 histograms (Fig. 6). Similarly, for the cingulum, the high standard devia-  
735 tions observed for the AD group are attributed to the two subjects mentioned  
736 above. However, no apparent outliers were identified in the MCI group based  
737 on the registration errors and, by extension, the mean FA and fibre orien-  
738 tation errors reported in Tables 4 and 3, are low and consistent with their  
739 corresponding histogram plots (Fig. 7 and Fig. 6).

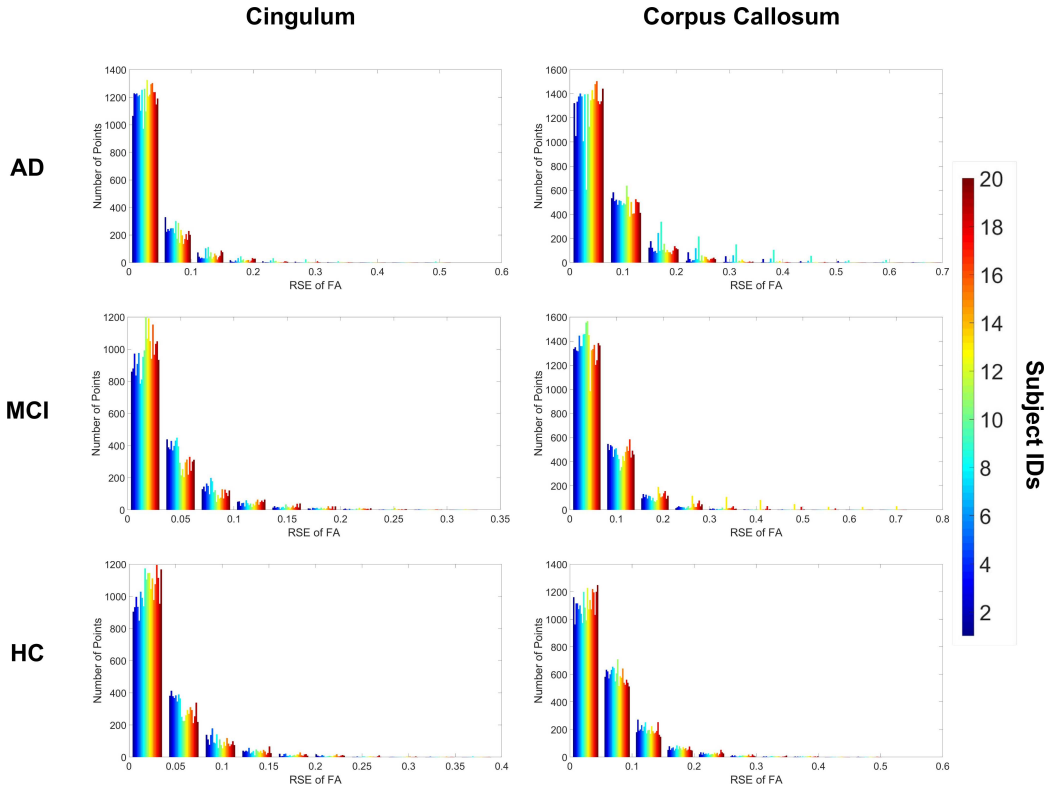


Figure 7: Histograms of root-squared-error (RSE) of FA for each subject in AD, MCI and HC groups, evaluated between established correspondences and ground truth voxels.

740 The foregoing results suggest the proposed framework established valid  
 741 correspondences for both WM ROIs across all subjects in the HC group and  
 742 for the majority of cases in the AD and MCI groups. This is indicative of the  
 743 ability of the proposed HdMM to approximate the joint  $PDF$  of positions,  
 744 fibre orientations and FA values across multiple subjects.

### 745 3.3. Group Comparisons

746 The ability of the proposed framework to identify significant differences  
 747 between patient groups was assessed by comparing each pair of patient groups  
 748 in terms of the variation in FA. These results were compared with those ob-  
 749 tained from the widely used TBSS approach. Un-paired two-sample t-tests,  
 750 assuming equal variances, were performed to compare FA values at corre-  
 751 sponding spatial positions between patient groups. The procedure proposed

752 in (Benjamini and Yekutieli, 2001) was used to correct for multiple compar-  
 753 isons by controlling the false discovery rate (FDR) for the set of hypothesis  
 754 tests. The desired FDR was fixed at 1% for all experiments. However, no  
 755 statistically significant reduction in FA was identified between any of the  
 756 groups, using the proposed approach, TBSS and VBM.

757 Interquartile ranges (IQRs) for the mean FA values estimated using each  
 758 approach were also evaluated to provide a quantitative means of comparing  
 759 the range of estimated FA values for both WM ROIs. This measure is adopted  
 760 as it provides a robust means of assessing dispersion in data. IQRs are  
 761 summarised in Table 8 for both WM ROIs, from which we infer that all  
 762 three methods do indeed show similarities in the range of estimated mean  
 763 FA values, for the corpus callosum. Conversely, for the cingulum, while  
 764 VBM and the proposed approach show similar IQRs, the ranges estimated  
 765 for TBSS are lower. This is because TBSS models the central skeleton of  
 766 the ROI, and there is substantial variation in FA between the center and  
 767 peripheral regions of cingulum region. Consequently, the variation in mean  
 768 FA values in the skeleton voxels is lower in comparison to the entire ROI (as  
 769 modelled by VBM and HdMM).

Table 8: Interquartile ranges for mean FA values estimated using each approach for both WM ROIs.

Method	Corpus Callosum: IQR of mean FA			Cingulum: IQR of mean FA		
	AD	MCI	HC	AD	MCI	HC
HdMM	0.24	0.24	0.24	0.17	0.16	0.16
TBSS	0.20	0.21	0.21	0.08	0.08	0.09
VBM	0.21	0.21	0.21	0.14	0.13	0.14

770 As discussed previously, the primary advantage of the proposed HdMM  
 771 framework is its ability to model fibre orientations and facilitate their compar-  
 772 ison across multiple subjects, which is not offered by conventional approaches  
 773 such as TBSS and VBM. Furthermore, the proposed method does not require  
 774 extraction of fibre trajectories using tractography in order to model fibre ori-  
 775 entations as it operates directly on the raw DTI-derived eigenvectors, unlike  
 776 state-of-the-art approaches such as those proposed in (Garyfallidis et al.,  
 777 2015) and (ODonnell et al., 2017). Inter-group statistical comparisons of  
 778 the angular deviation in fibre orientations, relative to study-specific mean

779 template, were also conducted. Here, the angular deviation of the model-  
780 predicted fibre orientations at each spatial correspondence was first evalu-  
781 ated relative to the corresponding mean fibre orientation (for patients from  
782 all groups), as the minimum arc length between unit vectors. Subsequently,  
783 these deviations were compared between each pair of patient groups, while  
784 correcting for multiple comparisons using FDR. However, as with the FA  
785 analyses, no statistically significant differences were identified.

786 The proposed HdMM for the joint registration and clustering of data com-  
787 prising positions, orientations and scalar-valued features (such as FA) shows  
788 promise for statistical analysis of diffusion derived measures across multiple  
789 subjects and patient populations. Although the inter-group statistical com-  
790 parisons conducted to analyse the variation in FA and fibre orientations re-  
791 vealed no significant differences between patient groups, our results matched  
792 those obtained using TBSS and VBM, in the case of the former. This may  
793 be due to the underlying nature of the data as the samples used throughout  
794 this study were part of the prospective cohort of the VPH-DARE@IT project.  
795 Consequently, it is possible that no significant differences in FA and fibre ori-  
796 entation exist in the WM ROIs considered, between the subjects assigned to  
797 the AD, MCI and HC groups. However, we believe the proposed approach  
798 still holds merit due to the flexibility it affords, as: (a) it enables analysis of  
799 various scalar-valued diffusion measures (although just FA was considered in  
800 this study), similar to existing approaches such as TBSS and VBM; and (b)  
801 also permits analysis of local fibre orientation, defined by primary diffusion  
802 axes, a capability not afforded by existing techniques. Although approaches  
803 based on clustering of fibre trajectories enable such analyses, they require  
804 diffusion-tractography derived fibres to do so. The present work ameliorates  
805 this need and acts directly on the raw eigenvector images. Additionally, our  
806 approach is not restricted to a specific anatomical region or analysing voxel-  
807 wise (or structured grid) data and may be employed to jointly register and  
808 cluster unstructured data as well.

809 A current limitation of the proposed approach is it only enables anal-  
810 ysis of DTI data generated using a single tensor model. However, the pro-  
811 posed HdMM framework could be imbued with greater flexibility by replacing  
812 the Watson distributions with the Kent or the general 8-parameter Fisher-  
813 Bingham distribution, to model multi-fibre (or crossing fibre) regions by fit-  
814 ting to orientation distribution functions obtained from high angular diffusion  
815 images. Extensions to the Von-Mises-Fisher mixture model for example, have  
816 been proposed previously to accommodate antipodal symmetry and model

817 diffusion ODFs (McGraw et al., 2006).

818 The sensitivity and discriminative capacity of the proposed framework  
819 in comparison to existing approaches requires further investigation and val-  
820 idation, which will be the subject of future work. Natural extensions to the  
821 proposed framework include whole WM volume analysis across multiple sub-  
822 jects, WM parcellation, and automatic region-of-interest analysis, to name  
823 a few. As discussed previously, the proposed approach can be employed to  
824 analyse the entire WM volume across subjects, i.e. *a priori* definition of ROIs  
825 is not required, though the computational burden at present is substantial.  
826 Such an approach naturally leads to the unsupervised parcellation of WM  
827 into distinct clusters defined by the centroids of the HdMM, across multiple  
828 subjects. This in turn provides a mechanism for automatic ROI-type analy-  
829 ses, as the generated clusters for each subject will correspond to similar WM  
830 regions in terms of spatial position, fibre orientation and FA (or some other  
831 scalar measure of interest). Furthermore, by employing a suitable prior/atlas  
832 containing pre-defined labels for WM tracts of interest, the presented frame-  
833 work could be employed for automatic tractography segmentation (similar  
834 to (O’Donnell and Westin, 2007)). The proposed approach can also be em-  
835 ployed to track and identify localised changes in WM over time for a single  
836 subject, resulting from the progression of neuro-degenerative disorders such  
837 as dementia, for example. Although WM changes in the brain were consid-  
838 ered in this study, the generic nature of the proposed framework permits its  
839 application to other organs exhibiting tissue anisotropy, such as cardiac dif-  
840 fusion data, and modelling bone micro-architecture. Additionally, it can be  
841 employed for a variety of other applications, such as vessel centerlines-based  
842 image registration, as demonstrated by our recent study (Bayer et al., 2018).

#### 843 4. Conclusions

844 In this study, a Watson-distribution based hybrid mixture model was pre-  
845 sented for jointly registering and clustering DTI-derived data from multiple  
846 subjects and patient populations. This approach was shown to model the  
847 observed fibre orientations and FA values accurately for all subjects within  
848 the HC group, for both of the studied WM ROIs, namely, the cingulum and  
849 corpus callosum. Registration to subjects in AD and MCI groups was suc-  
850 cessful for the majority of cases, with two in the former and one in the latter  
851 resulting in high registration errors, due to significant pathology induced  
852 changes in these cases. Group comparisons of FA values in the WM ROIs

853 using the proposed approach showed no statistically significant reductions in  
854 FA between the AD, MCI and HC groups, as with TBSS and VBM. Similarly,  
855 no significant variations in fibre orientation were identified between patient  
856 groups. However, the proposed method has potential for use in a variety of  
857 applications involving statistical analysis of diffusion data. Its generic and  
858 flexible nature make it well suited to a variety of other computer vision and  
859 medical image analysis tasks, such as: point set registration with the inte-  
860 gration of surface normals, vessel-based image registration, joint registration  
861 and clustering of geometries with associated velocity fields (estimated from  
862 computational fluid dynamic simulations for example) and texture mapping,  
863 to name a few. The fidelity and extensibility of the proposed framework is  
864 thus compelling as a general tool for multi-dimensional medical image anal-  
865 ysis.

## 866 Acknowledgments

867 This study was funded by the European Unions Seventh Framework Pro-  
868 gramme (FP7/2007 2013) as part of the project VPH-DARE@IT (grant  
869 agreement no. 601055), and by the Engineering and Physical Sciences Re-  
870 search Council through the OCEAN project (EP/M006328/1). The authors  
871 would like to thank the partners of the VPH-DARE@IT consortium for the  
872 acquisition of the brain DW-MRI data used in this study and for providing  
873 access to the same.

## 874 Appendix

875 M-step update equations for the Student’s t-distribution parameters in  
876 the HdMM and rigid registration parameters at the  $(t + 1)^{\text{th}}$  EM-iteration,  
877 discussed in section 2.6, are derived by maximizing the complete data log-  
878 likelihood  $Q(\Theta_p^{t+1}, \mathbb{T}^{t+1} | \Theta_p^t, \mathbb{T}^t)$  with respect to each parameter as follows:

- Estimation of TMM centroids  $\boldsymbol{\mu}_j$  at the  $(t + 1)^{\text{th}}$  EM-iteration:

$$Q(\Theta_p^{t+1}, \mathbb{T}^{t+1} | \Theta_p^t, \mathbb{T}^t) = -\frac{1}{2} \sum_{k,i,j} P_{kij}^{\star t} \Delta_{kij} + O.T. \quad (8a)$$

$$\Delta_{kij} = \frac{(\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)^T (\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)}{\sigma^2} \quad (8b)$$

$O.T.$  summarizes terms in  $Q$  independent of  $\boldsymbol{\mu}_j$ .

$$\langle \partial Q, \partial \boldsymbol{\mu}_j \rangle = \left[ -\frac{1}{2} \sum_{k,i} P_{kij}^* \Delta_{kij}^{\boldsymbol{\mu}_j + \partial \boldsymbol{\mu}_j} \right] - \left[ -\frac{1}{2} \sum_{k,i} P_{kij}^* \Delta_{kij}^{\boldsymbol{\mu}_j} \right] \quad (9a)$$

$$\langle \partial Q, \partial \boldsymbol{\mu}_j \rangle = \sum_{k,i} P_{kij}^* [(\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)^T s_k \mathbf{R}_k] \partial \boldsymbol{\mu}_j \quad (9b)$$

$$\langle \partial Q, \partial \boldsymbol{\mu}_j \rangle = 0 \implies \sum_{k,i} P_{kij}^* [(\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)^T s_k \mathbf{R}_k] = 0 \quad (9c)$$

$$\sum_{k,i} P_{kij}^* s_k \mathbf{R}_k^T (\mathbf{x}_{ki} - \mathbf{t}_k) = \sum_{k,i} P_{kij}^* s_k \mathbf{R}_k^T \mathbf{R}_k s_k \boldsymbol{\mu}_j \quad (9d)$$

$$\boldsymbol{\mu}_j = \frac{\sum_{k,i} P_{kij}^* s_k^{-1} \mathbf{R}_k^T (\mathbf{x}_{ki} - \mathbf{t}_k)}{\sum_{k,i} P_{kij}^*} \quad (9e)$$

- Estimation of model variance  $\sigma^2$ :

$$\frac{\partial Q}{\partial \sigma^2} = \frac{\partial \sum_{k,i,j} \left[ -\frac{P_{kij}}{2} [\log(\sigma^6)] - \frac{P_{kij}^*}{2} [\Delta_{kij}] \right]}{\partial \sigma^2} = 0 \quad (10a)$$

$$\implies \sum_{k,i,j} -P_{kij} \frac{3}{\sigma} + P_{kij}^* \frac{(\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)^T (\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)}{\sigma^3} = 0 \quad (10b)$$

$$\sigma^2 = \frac{\sum_{k,i,j} P_{kij}^* (\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)^T (\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)}{3 \sum_{kij} P_{kij}} \quad (10c)$$

- Estimation of translation  $\mathbf{t}_k$ :

$$\langle \partial Q, \partial \mathbf{t}_k \rangle = \left[ -\frac{1}{2} \sum_{i,j} P_{kij}^* \Delta_{kij}^{\mathbf{t}_k + \partial \mathbf{t}_k} \right] - \left[ -\frac{1}{2} \sum_{i,j} P_{kij}^* \Delta_{kij}^{\mathbf{t}_k} \right] \quad (11a)$$

$$\langle \partial Q, \partial \mathbf{t}_k \rangle = \sum_{i,j} P_{kij}^* [(\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j - \mathbf{t}_k)^T] \partial \mathbf{t}_k \quad (11b)$$

$$\langle \partial Q, \partial \mathbf{t}_k \rangle = 0 \implies \sum_{i,j} P_{kij}^* (\mathbf{x}_{ki} - s_k \mathbf{R}_k \boldsymbol{\mu}_j)^T = \sum_{i,j} P_{kij}^* \mathbf{t}_k^T \quad (11c)$$



$$\mathbf{t}_k = \frac{\sum_{i,j} P_{kij}^* \mathbf{x}_{ki}}{\sum_{i,j} P_{kij}^*} - s_k \mathbf{R}_k \frac{\sum_{i,j} P_{kij}^* \boldsymbol{\mu}_j}{\sum_{i,j} P_{kij}^*} \quad (11d)$$

Setting the first term as  $\mathbf{d}_k$  and the second term as  $\mathbf{m}_k$  we get:

$$\mathbf{t}_k = \mathbf{d}_k - s_k \mathbf{R}_k \mathbf{m}_k \quad (11e)$$

- Estimation of strictly orthogonal rotation  $\mathbf{R}_k$ : Using the lemma outlined in (Myronenko and Song, 2010), the optimal rotation matrix maximises  $\text{tr}(\mathbf{C}_k^T \mathbf{R}_k)$  where  $\mathbf{C}_k$  represents a real covariance matrix (refer to equation 12d).

$$\tilde{\mathbf{x}}_{ki} = \mathbf{x}_{ki} - \mathbf{d}_k, \tilde{\mathbf{m}}_{kj} = \boldsymbol{\mu}_j - \mathbf{m}_k \quad (12a)$$

Using equations (11e) and (12a) we get:

$$Q(\Theta_p^{t+1}, \mathbb{T}^{t+1} | \Theta_p^t, \mathbb{T}^t) \propto \sum_{i,j} P_{kij}^{*t} (\tilde{\mathbf{x}}_{ki}^T \mathbf{R}_k \tilde{\mathbf{m}}_{kj}) \quad (12b)$$

$$Q(\Theta_p^{t+1}, \mathbb{T}^{t+1} | \Theta_p^t, \mathbb{T}^t) \propto \sum_{i,j} P_{kij}^{*t} \text{tr}[\tilde{\mathbf{m}}_{kj} \tilde{\mathbf{x}}_{ki}^T \mathbf{R}_k] \quad (12c)$$

As equation (12c) must be maximised with respect to  $\mathbf{R}_k$ ,

$$\mathbf{C}_k = \sum_{i,j} P_{kij}^* \tilde{\mathbf{x}}_{ki} \tilde{\mathbf{m}}_{kj}^T \quad (12d)$$

879  $\mathbf{R}_k = \mathbf{U}\mathbf{S}\mathbf{V}^T$ , where  $\mathbf{U}, \mathbf{V}$  are unitary matrices computed by singular  
 880 value decomposition of  $\mathbf{C}_k$  and  $\mathbf{S} = \text{diag}(1, 1, \det(\mathbf{U}\mathbf{V}^T))$  is a diagonal  
 881 matrix that prevents reflections.

- Estimation of scaling  $s_k$ :

$$\frac{\partial Q}{\partial s_k} = -\frac{1}{2} \frac{\partial \sum_{i,j} P_{kij}^* \Delta_{kij}}{\partial s_k} = 0 \quad (13a)$$

$$\sum_{i,j} P_{kij}^* \frac{(\tilde{\mathbf{x}}_{ki} - s_k \mathbf{R}_k \tilde{\mathbf{m}}_{kj})^T (\mathbf{R}_k \tilde{\mathbf{m}}_{kj})}{\sigma^2} = 0 \quad (13b)$$

$$\sum_{i,j} P_{kij}^* [(\tilde{\mathbf{x}}_{ki})^T (\mathbf{R}_k \tilde{\mathbf{m}}_{kj})] = s_k \sum_{i,j} P_{kij}^* [\tilde{\mathbf{m}}_{kj}^T \mathbf{R}_k^T \mathbf{R}_k \tilde{\mathbf{m}}_{kj}] \quad (13c)$$

$$s_k = \frac{\text{tr}[\tilde{\mathbf{m}}_{kj} \tilde{\mathbf{x}}_{ki}^T] \mathbf{R}_k}{\text{tr}[\tilde{\mathbf{m}}_{kj} \tilde{\mathbf{m}}_{kj}^T]} = \frac{\text{tr}[\mathbf{C}_k^T \mathbf{R}_k]}{\text{tr}[\tilde{\mathbf{m}}_{kj} \tilde{\mathbf{m}}_{kj}^T]} \quad (13d)$$

- Estimation of degrees of freedom  $\nu_j$ :

$$Q(\Theta_p^{t+1}, \mathbb{T}^{t+1} | \Theta_p^t, \mathbb{T}^t) = \sum_{k,i,j} P_{kij}^t [-\log \Gamma(\frac{\nu_j}{2}) + \frac{1}{2} \nu_j \log(\frac{\nu_j}{2}) + \frac{\nu_j}{2} [\log(U_{kij}^t) - U_{kij}^t + \Psi(\frac{\nu_j + D}{2}) - \log(\frac{\nu_j^t + D}{2})]] + O.T. \quad (14a)$$

$O.T.$  summarizes terms in  $Q$  independent of  $\nu_j$ .

$$\frac{\partial Q}{\partial \nu_j} = -\Psi(\frac{\nu_j}{2}) + \log(\frac{\nu_j}{2}) + 1 + \frac{1}{\sum_{k,i} P_{kij}^t} \sum_{k,i} P_{kij}^t (\log(U_{kij}^t) - U_{kij}^t) + \Psi(\frac{\nu_j^t + D}{2}) - \log(\frac{\nu_j^t + D}{2}) = 0 \quad (14b)$$

882 Equation (14b) is solved using Newton's method to estimate the degrees  
883 of freedom  $\nu_j$ .

- 884 • Derivations for the M-step updates (refer to equations 3c - 3e) of the  
885 mean fibre orientation  $\mathbf{m}_j^d$  and fibre concentration  $\kappa_j$  parameters asso-  
886 ciated with Watson distributions in the HdMM, presented in section  
887 2.4, are derived by maximizing the complete data log-likelihood  $Q$  (re-  
888 fer to equation 15a), with respect to each model parameter as follows:  
889 (Here  $M(\kappa_j)$  denotes the Kummer function).

$$Q(\Theta_n^{t+1} | \Theta_n^t) = \sum_{k=1}^K \sum_{i=1}^{N_k} \sum_{j=1}^M P_{kij} \log p(\pm \mathbf{n}_{ki} | \mathbf{m}_j^d, \kappa_j) + \lambda_j (1 - \mathbf{m}_j^{dT} \mathbf{m}_j^d) \quad (15a)$$

$$\langle \partial Q, \partial \mathbf{m}_j^d \rangle = 0 \implies \lambda_j \mathbf{m}_j^d = \kappa_j \sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} (\mathbf{n}_{ki}^T \mathbf{m}_j^d) \mathbf{n}_{ki} \quad (15b)$$

$$\langle \partial Q, \partial \kappa_j \rangle = 0 \implies \frac{M'(\kappa_j)}{M(\kappa_j)} \sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} = \sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} (\mathbf{n}_{ki}^T \mathbf{m}_j^d)^2 \quad (15c)$$

$$\mathbf{m}_j^{dT} \mathbf{m}_j^d = 1 \implies \lambda_j = \kappa_j \left\| \sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} (\mathbf{n}_{ki}^T \mathbf{m}_j^d) \mathbf{n}_{ki} \right\| \quad (15d)$$

Substituting equation (15d) in (15b) results in a non-linear equation (16), which is solved numerically by fixed-point iteration.

$$\mathbf{m}_j^d = \frac{\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} (\mathbf{n}_{ki}^T \mathbf{m}_j^d) \mathbf{n}_{ki}}{\left\| \sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} (\mathbf{n}_{ki}^T \mathbf{m}_j^d) \mathbf{n}_{ki} \right\|} \quad (16)$$

Based on equation (15c), the ratio of the derivative of the Kummer function to the function itself, is expressed as shown in equation (17a). This ratio may be expressed as a continued fraction, as shown in equation (17b). Consequently, using equations (17a) and (17b), the concentration parameters  $\kappa_j$  can be approximated as shown in equation (17d), by solving the linear equation (17c) (similarly to (Bijral et al., 2007)).

$$\frac{M'(\kappa_j)}{M(\kappa_j)} = \frac{\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij} (\mathbf{n}_{ki}^T \mathbf{m}_j^d)^2}{\sum_{k=1}^K \sum_{i=1}^{N_k} P_{kij}} \quad (17a)$$

$$\frac{\kappa_j M'(\kappa_j)}{M(\kappa_j)} = \frac{\kappa_j/2}{(D/2) - \kappa_j + \frac{(3/2)\kappa_j}{(\frac{D}{2}+1) - \kappa_j + \dots}} \quad (17b)$$

$$\frac{\kappa_j M'(\kappa_j)}{M(\kappa_j)} \approx \frac{\kappa_j/2}{(D/2) - \kappa_j + \frac{\kappa_j M'(\kappa_j)}{M(\kappa_j)}} \quad (17c)$$

$$\kappa_j \approx \frac{1}{2} \left[ \frac{1 - \frac{M'(\kappa_j)}{M(\kappa_j)} D}{\left(\frac{M'(\kappa_j)}{M(\kappa_j)}\right)^2 - \frac{M'(\kappa_j)}{M(\kappa_j)}} \right] \quad (17d)$$

- The mean-squared distance (MSD) metric (refer to equation (18)) is used to assess registration errors in terms of spatial position. MSD values were evaluated between the correspondences established following registration of the (study-specific) mean template, and the corresponding original hybrid point sets (i.e. between the estimated correspondences and the voxel centroids defining the WM ROIs). In equation

(18)  $\mathbf{d}_{\min}(A, B)$  denotes the minimum Euclidean distance between each point in sample A and sample B.

$$MSD = \text{mean}(\text{mean}(\mathbf{d}_{\min}(A, B)), \text{mean}(\mathbf{d}_{\min}(B, A))) \quad (18)$$

- The “model-predicted” values for FA ( $\hat{f}_{kj}$ ) and fibre orientation ( $\hat{\mathbf{n}}_{kj}$ ) estimated at each established spatial correspondence, for each patient, are weighted averages of the neighbouring voxels in their original DTI-derived images (original hybrid point sets), where the weights are defined by the estimated posterior probabilities following non-rigid registration of the study-specific mean template to each sample. These values were estimated for FA and fibre orientation as described by equations 19a and 19b, respectively.

$$\hat{f}_{kj} = \sum_{i=1}^{N_k} \frac{P_{kij} f_{ki}}{\sum_l P_{klj}} \quad (19a)$$

$$\hat{\mathbf{n}}_{kj} = \sum_{i=1}^{N_k} \frac{P_{kij} \mathbf{n}_{ki}}{\sum_l P_{klj}} \quad (19b)$$

890

## 891 References

- 892 Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry — the meth-  
893 ods. *Neuroimage* 11, 805–821.
- 894 Basser, P.J., Mattiello, J., LeBihan, D., 1994. Estimation of the effective self-  
895 diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance*,  
896 Series B 103, 247–254.
- 897 Bayer, S., Ravikumar, N., Strumia, M., Tong, X., Gao, Y., Ostermeier, M.,  
898 Fahrig, R., Maier, A., 2018. Intraoperative brain shift compensation using  
899 a hybrid mixture model, in: *International Conference on Medical Image*  
900 *Computing and Computer-Assisted Intervention*, Springer. pp. 116–124.
- 901 Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in  
902 multiple testing under dependency. *Annals of Statistics* , 1165–1188.

- 903 Benou, I., Veksler, R., Friedman, A., Raviv, T.R., 2018. Fiber-flux diffusion  
904 density for white matter tracts analysis: Application to mild anomalies  
905 localization in contact sports players, in: Computational Diffusion MRI,  
906 Springer. pp. 191–204.
- 907 Bijral, A.S., Breitenbach, M., Grudic, G.Z., 2007. Mixture of watson distribu-  
908 tions: A generative model for hyperspherical embeddings., in: AISTATS,  
909 pp. 35–42.
- 910 Bishop, C.M., 2006. Pattern Recognition and Machine Learning (Information  
911 Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ,  
912 USA.
- 913 Cercignani, M., Inglese, M., Pagani, E., Comi, G., Filippi, M., 2001. Mean  
914 diffusivity and fractional anisotropy histograms of patients with multiple  
915 sclerosis. *American Journal of Neuroradiology* 22, 952–958.
- 916 Chang, L.C., Walker, L., Pierpaoli, C., 2012. Informed restore: a method for  
917 robust estimation of diffusion tensor from low redundancy datasets in the  
918 presence of physiological noise artifacts. *Magnetic Resonance in Medicine*  
919 68, 1654–1663.
- 920 Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from  
921 incomplete data via the em algorithm. *Journal of the Royal Statistical*  
922 *Society. Series B (methodological)* , 1–38.
- 923 Durrleman, S., Fillard, P., Pennec, X., Trouvé, A., Ayache, N., 2011. Reg-  
924 istration, atlas estimation and variability analysis of white matter fiber  
925 bundles modeled as currents. *NeuroImage* 55, 1073–1090.
- 926 Durrleman, S., Pennec, X., Trouvé, A., Ayache, N., 2009. Statistical models  
927 of sets of curves and surfaces based on currents. *Medical Image Analysis*  
928 13, 793–808.
- 929 Garyfallidis, E., Ocegueda, O., Wassermann, D., Descoteaux, M., 2015. Ro-  
930 bust and efficient linear registration of white-matter fascicles in the space  
931 of streamlines. *NeuroImage* 117, 124–140.
- 932 Gooya, A., Davatzikos, C., Frangi, A.F., 2015. A bayesian approach to sparse  
933 model selection in statistical shape models. *SIAM Journal on Imaging*  
934 *Sciences* 8, 858–887.

- 935 Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi,  
936 P.A., Pekar, J.J., van Zijl, P.C., Mori, S., 2008. Tract probability maps  
937 in stereotaxic spaces: analyses of white matter anatomy and tract-specific  
938 quantification. *Neuroimage* 39, 336–347.
- 939 Hufnagel, H., Pennec, X., Ehrhardt, J., Ayache, N., Handels, H., 2008. Gen-  
940 eration of a statistical shape model with probabilistic point correspon-  
941 dences and the expectation maximization-iterative closest point algorithm.  
942 *International journal of computer assisted radiology and surgery* 2, 265–  
943 273.
- 944 Huynh, D.Q., 2009. Metrics for 3d rotations: Comparison and analysis.  
945 *Journal of Mathematical Imaging and Vision* 35, 155–164.
- 946 Jian, B., Vemuri, B.C., 2005. A robust algorithm for point set registration  
947 using mixture of gaussians, in: *Tenth IEEE International Conference on*  
948 *Computer Vision (ICCV’05) Volume 1, IEEE*. pp. 1246–1251.
- 949 Jupp, P., Mardia, K., 1989. A unified view of the theory of directional statis-  
950 tics, 1975-1988. *International Statistical Review/Revue Internationale de*  
951 *Statistique* , 261–294.
- 952 Liu, Y., Spulber, G., Lehtimäki, K.K., Könönen, M., Hallikainen, I., Gröhn,  
953 H., Kivipelto, M., Hallikainen, M., Vanninen, R., Soininen, H., 2011. Diffu-  
954 sion tensor imaging and tract-based spatial statistics in alzheimer’s disease  
955 and mild cognitive impairment. *Neurobiology Of Aging* 32, 1558–1571.
- 956 Maddah, M., Grimson, W.E.L., Warfield, S.K., Wells, W.M., 2008. A unified  
957 framework for clustering and quantitative analysis of white matter fiber  
958 tracts. *Medical Image Analysis* 12, 191–202.
- 959 Mayer, A., Greenspan, H., 2008. Bundles of interest based registration  
960 of white matter tractographies, in: *Biomedical Imaging: From Nano to*  
961 *Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, IEEE*.  
962 pp. 919–922.
- 963 Mayer, A., Zimmerman-Moreno, G., Shadmi, R., Batkoff, A., Greenspan,  
964 H., 2011. A supervised framework for the registration and segmentation  
965 of white matter fiber tracts. *IEEE transactions on medical imaging* 30,  
966 131–145.

- 967 McGraw, T., Vemuri, B.C., Yeziarski, B., Mareci, T., 2006. Von mises-fisher  
968 mixture model of the diffusion odf, in: 3rd IEEE International Symposium  
969 on Biomedical Imaging: Nano to Macro, 2006., IEEE. pp. 65–68.
- 970 Medina, D., Urresta, F., Gabrieli, J.D., Moseley, M., Fleischman, D., Ben-  
971 nett, D.A., Leurgans, S., Turner, D.A., Stebbins, G.T., et al., 2006. White  
972 matter changes in mild cognitive impairment and ad: a diffusion tensor  
973 imaging study. *Neurobiology Of Aging* 27, 663–672.
- 974 Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes,  
975 D.J., Fox, N.C., Ourselin, S., 2010. Fast free-form deformation using graph-  
976 ics processing units. *Computer Methods And Programs In Biomedicine* 98,  
977 278–284.
- 978 Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria,  
979 A.V., Mahmood, A., Woods, R., et al., 2008. Stereotaxic white matter  
980 atlas based on diffusion tensor imaging in an icbm template. *Neuroimage*  
981 40, 570–582.
- 982 Myronenko, A., Song, X., 2010. Point set registration: Coherent point  
983 drift. *IEEE Transactions On Pattern Analysis And Machine Intelligence*  
984 32, 2262–2275.
- 985 O'Donnell, L.J., Suter, Y., Rigolo, L., Kahali, P., Zhang, F., Norton, I., Albi,  
986 A., Olubiyi, O., Meola, A., Essayed, W.I., et al., 2017. Automated white  
987 matter fiber tract identification in patients with brain tumors. *NeuroImage:*  
988 *Clinical* 13, 138–153.
- 989 O'Donnell, L.J., Wells, W.M., Golby, A.J., Westin, C.F., 2012. Unbiased  
990 groupwise registration of white matter tractography, in: *International Con-*  
991 *ference on Medical Image Computing and Computer-assisted Intervention,*  
992 *Springer.* pp. 123–130.
- 993 O'Donnell, L.J., Westin, C.F., 2007. Automatic tractography segmentation  
994 using a high-dimensional white matter atlas. *IEEE Transactions On Med-*  
995 *ical Imaging* 26, 1562–1575.
- 996 Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N., 2001. Recon-  
997 structing a 3d structure from serial histological sections. *Image and Vision*  
998 *Computing* 19, 25–31.

- 999 Pierpaoli, C., Basser, P.J., 1996. Toward a quantitative assessment of diffu-  
1000 sion anisotropy. *Magnetic Resonance in Medicine* 36, 893–906.
- 1001 Pierpaoli, C., Walker, L., Irfanoglu, M., Barnett, A., Basser, P., Chang,  
1002 L., Koay, C., Pajevic, S., Rohde, G., Sarlls, J., et al., 2010. Tortoise:  
1003 an integrated software package for processing of diffusion mri data, in:  
1004 ISMRM 18th annual meeting, p. 1597.
- 1005 Ravikumar, N., Gooya, A., Çimen, S., Frangi, A.F., Taylor, Z.A., 2016. A  
1006 multi-resolution t-mixture model approach to robust group-wise alignment  
1007 of shapes, in: *International Conference on Medical Image Computing and*  
1008 *Computer-Assisted Intervention*, Springer. pp. 142–149.
- 1009 Ravikumar, N., Gooya, A., Çimen, S., Frangi, A.F., Taylor, Z.A., 2018.  
1010 Group-wise similarity registration of point sets using students t-mixture  
1011 model for statistical shape models. *Medical Image Analysis* 44, 156–176.
- 1012 Ravikumar, N., Gooya, A., Frangi, A.F., Taylor, Z.A., 2017. Generalised co-  
1013 herent point drift for group-wise registration of multi-dimensional point  
1014 sets, in: *International Conference on Medical Image Computing and*  
1015 *Computer-Assisted Intervention*, Springer. pp. 309–316.
- 1016 Salat, D., Tuch, D., Greve, D., Van Der Kouwe, A., Hevelone, N., Zaleta,  
1017 A., Rosen, B., Fischl, B., Corkin, S., Rosas, H.D., et al., 2005. Age-related  
1018 alterations in white matter microstructure measured by diffusion tensor  
1019 imaging. *Neurobiology Of Aging* 26, 1215–1227.
- 1020 Schwartzman, A., Dougherty, R.F., Taylor, J.E., 2005. Cross-subject compar-  
1021 ison of principal diffusion direction maps. *Magnetic Resonance in Medicine*  
1022 53, 1423–1431.
- 1023 Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E.,  
1024 Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M.,  
1025 et al., 2006. Tract-based spatial statistics: voxelwise analysis of multi-  
1026 subject diffusion data. *Neuroimage* 31, 1487–1505.
- 1027 Sra, S., Karp, D., 2013. The multivariate watson distribution: Maximum-  
1028 likelihood estimation and other aspects. *Journal of Multivariate Analysis*  
1029 114, 256–269.



- 1030 Zhang, Y., Schuff, N., Jahng, G.H., Bayne, W., Mori, S., Schad, L., Mueller,  
1031 S., Du, A.T., Kramer, J., Yaffe, K., et al., 2007. Diffusion tensor imaging  
1032 of cingulum fibers in mild cognitive impairment and alzheimer disease.  
1033 Neurology 68, 13–19.
- 1034 Zhou, Z., Zheng, J., Dai, Y., Zhou, Z., Chen, S., 2014. Robust non-rigid  
1035 point set registration using student’s-t mixture model. PloS one 9, e91381.
- 1036 Zvitia, O., Mayer, A., Shadmi, R., Miron, S., Greenspan, H.K., 2010. Co-  
1037 registration of white matter tractographies by adaptive-mean-shift and  
1038 gaussian mixture modeling. IEEE Transactions On Medical Imaging 29,  
1039 132–145.