Article:

Ravikumar, N, Gooya, A, Beltrachini, L et al. (2019) Generalised coherent point drift for group-wise multi-dimensional analysis of diffusion brain MRI data. Medical Image Analysis, 53. pp. 47-63. ISSN: 1361-8415

https://doi.org/10.1016/j.media.2019.01.001

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

Nishant Ravikumar[a,*], Ali Gooya[d,f,g], Leandro Beltrachini[a], Alejandro F. Frangi[b,c,d,f,g], Zeike A. Taylor[c,e,f,g]

[a]*CISTIB Centre for Computational Imaging and Simulation Technologies in Biomedicine, INSIGNEO Institute for in silico Medicine, University of Sheffield, Sheffield, UK*
[b]*LICAMM Leeds Institute of Cardiovascular and Metabolic Medicine, School of Medicine, University of Leeds, Leeds, UK*
[c]*iMBE Institute of Medical and Biological Engineering, University of Leeds, Leeds, UK*
[d]*School of Computing, University of Leeds, Leeds, UK*
[e]*School of Mechanical Engineering, University of Leeds, Leeds, UK*
[f]*CISTIB Center for Computational Imaging and Simulation Technologies in Biomedicine, University of Leeds, Leeds, UK*
[g]*Previously at CISTIB Centre for Computational Imaging and Simulation Technologies in Biomedicine, University of Sheffield, Sheffield, UK*

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

---

[*] Corresponding author

Email addresses: `mta08nr@gmail.com` (Nishant Ravikumar), `A.Gooya@leeds.ac.uk` (Ali Gooya), `beltrachinil@cardiff.ac.uk` (Leandro Beltrachini), `A.Frangi@leeds.ac.uk` (Alejandro F. Frangi), `Z.Taylor@leeds.ac.uk` (Zeike A. Taylor)

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

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

Alternative probabilistic techniques that jointly register and cluster WM fibre trajectories (obtained from diffusion tractography), and which enable quantitative analysis of diffusion measures over fibre pathways (rather than voxel-wise quantification), have also been proposed. For example, registration of curves and fibre bundles using diffeomorphisms and currents, and a statistical framework to assess variability in geometry and fibre density across a population, was proposed in (Durrleman et al., 2009), (Durrleman et al., 2011). Maddah et al. (2008) employ a Gamma mixture modelling framework to register fibre trajectories by establishing probabilistic correspondences, and jointly cluster them into representative fibre bundles. The authors also note therein, through use of a suitable fibre tract atlas as a prior during the clustering procedure, correspondences may be estimated across fibre trajectories obtained from multiple subjects, thereby enabling statistical analysis of FA and other diffusion quantities across populations. Similarly, (Mayer et al., 2011) proposed a supervised approach for joint registration and segmentation WM tracts, wherein, the iterative closest fiber algorithm (Mayer and Greenspan, 2008) was used to register fibre sets between a manually annotated tractography atlas and a subject's reconstructed set of fibres. The resulting segmentation was subsequently refined using a probabilistic boosting tree-based classifier. In (Zvitia et al., 2010), the authors propose a combined adaptive mean shift and Gaussian mixture model (GMM) formulation to jointly cluster fibre trajectories into compact fibre sets, and subsequently register fibre sets obtained from multiple subjects. The registration of two clustered fibre sets is formulated as a problem of aligning two distinct GMMs, analogous to point set registration using GMMs (Jian and Vemuri, 2005). Similar approaches to clustering fibre trajectories across a population, using spectral embedding, have also been proposed (O'Donnell and Westin, 2007), facilitating the estimation of WM atlases and enabling automatic segmentation of major WM tracts. An unbiased, group-wise, whole-brain tractography registration approach was proposed by (O'Donnell et al., 2012). Kernel density estimation was used to approximate the probability distribution of fibre trajectories within each brain and the overall distribution of the atlas, was modelled as a mixture of the former. Alignment of WM tracts was achieved by minimizing an entropic measure defined on the atlas distribution. In a follow up study ODonnell et al. (2017), this group-wise registration approach was combined with their previous work on spectral clustering of fibre trajectories, to formulate an end-to-end automated framework for automated WM tract identification, thereby enabling statistical analyses of DTI-derived quantities. Garyfallidis et al. (2015) proposed a linear registration framework to align WM bundles directly in the space of streamlines. They also demonstrated the viability of their approach to construct bundle specific atlases. In a recent study (Benou et al., 2018), novel descriptors called Fiber-Flux Diffusion Density (FFDD), which jointly describe fibre bundle geometry and diffusivity measures were proposed, to facilitate localized quantification of WM fibre bundles. Additionally, a FFDD dissimilarity measure was formulated and a novel registration framework (based on the fast marching method) for WM tract-profiles was proposed, enabling inter-subject comparisons and group-wise statistical analysis. Such techniques are however, dependent on the tractography algorithm employed to estimate fibre trajectories, introducing an additional potential source of error, and typically require some degree of user intervention (to define seeds for streamline generation for example).

Applications of the various methods described above have included, for example, identification of relationships between mild cognitive impairment (MCI) and Alzheimer's disease (AD), and localised changes to WM diffusion characteristics. For example, in (Zhang et al., 2007), ROI-based analysis was used to identify significant reduction in FA in the cingulum for patients diagnosed with MCI and AD, relative to healthy controls (HC). In (Medina et al., 2006), VBM was used to identify significant reduction in FA in posterior regions of the brain, for MCI and AD patient groups, using VBM. While (Liu et al., 2011) used the TBSS-approach and found reduced FA in the cingulum, corpus callosal and inferior/superior longitudinal fasiculus tracts, among others.

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

Statistical analysis of fibre orientations across multiple subjects and comparisons between patient groups was pursued in a previous study (Schwartzman et al., 2005). Here, the authors followed a VBM-style approach where DTIs from multiple subjects were spatially normalized to a reference template using a spline-based tensor interpolation approach together with a tensor re-orientation mechanism designed to preserve the principal diffusion direction. Subsequently, Watson distributions were fitted by maximum likelihood estimation to the fibre orientations observed across a group, at each voxel, independently. This provides a measure of the mean orientation and dispersion, observed across the group of subjects. A drawback of such an approach however, is the need to choose a single, appropriate template, for spatial normalization, which is particularly difficult for images exhibiting varying degrees of pathology-induced morphological changes. All subsequent registrations performed and correspondences estimated are biased towards the chosen template. VBM-based approaches in general, are dependent on the accuracy of non-rigid registration and the exact estimation of correspondences, to ensure validity in the subsequent voxel-wise statistical analyses. TBSS and our proposed approach are less restrictive in this regard. Registration of WM regions defined by hybrid point sets (comprising voxel positions, associated fibre orientations and FA values) across subjects, is achieved using a group-wise rigid, and subsequent non-rigid point set registration procedure, based on a HdMM. In the proposed approach, correspondence probabilities are estimated by approximating the joint probability density of position, fibre orientation and FA, which are iteratively revised as the registration progresses. Consequently, three distinct sources of information are leveraged to guide the registration of an unbiased, study-specific atlas (iteratively revised as the registration progresses), onto each subject's WM tract/ROI. The evolving soft correspondences provide model-based estimates for the mean fibre orientation and FA value (for a given population) at each component in the mixture model and help mitigate any misalignment incurred during registration.

1.1. Motivation and Contributions

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

2. Methods

2.1. Pre-processing

MR-DWIs were acquired for 60 subjects (20 HC, 20 MCI, 20 AD), as part of prospective cohort of the VPH-DARE@IT project (`vph-dare.eu`). All images used in this study were acquired using identical protocols: 2 diffusion-weighted b-values $(0, 800)$, with diffusivity gradients applied along 32 directions; image size of $(240 \times 240 \times 120)$ slices, 2.5mm thick in the right-left, anterior-posterior and inferior-superior directions, respectively. DTIs were estimated from these for each subject using TORTOISE v 2.5.0 (Pierpaoli et al., 2010), which employs state-of-the-art algorithms for motion and eddy current correction, correcting B0 susceptibility induced EPI distortions and B-matrix re-orientation artefacts. Tensor-fitting was then achieved using iRESTORE (Chang et al., 2012), based on non-linear iterative least-squares. TORTOISE registers each subject's DWIs to their corresponding T2-weighted structural MRI during the aforementioned pre-processing steps. As the latter were acquired at resolutions of $(1.5 \times 1.5 \times 1.5mm)$, all estimated DTIs (and correspondingly, DTI-derived images) were up-sampled relative to their raw DWIs. Finally, tensor-derived measures such as the eigenvector and fractional anisotropy images were also estimated using TORTOISE.

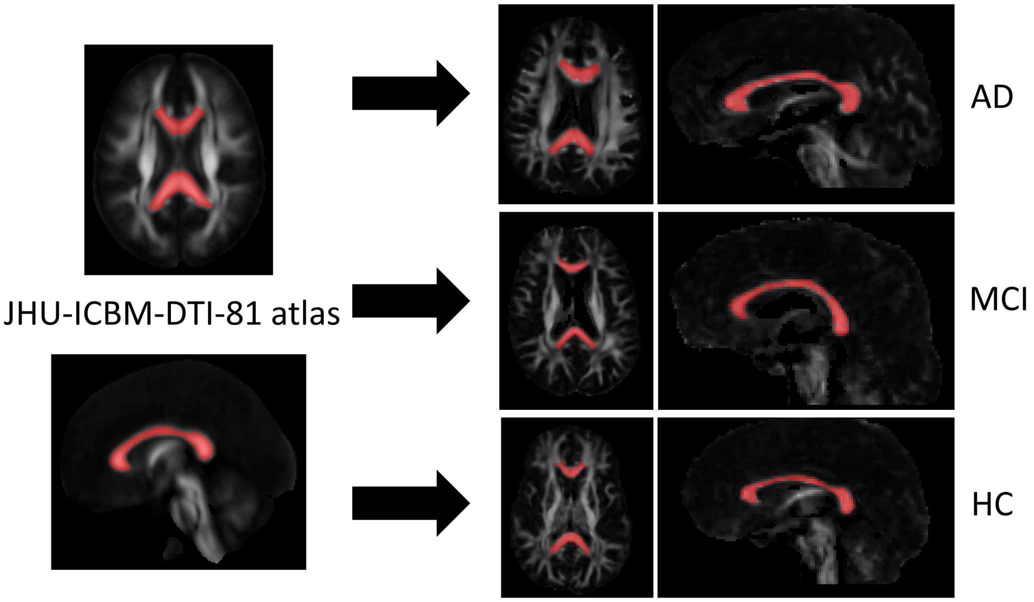

Figure 1: Nifty-Reg used to propagate labels for WM regions of interest from JHU-ICBM-DTI-81 atlas to each subject in AD, MCI and HC groups. Images depict propagation of the corpus callosum label from the atlas to subjects in AD, MCI and control groups.

The proposed framework is flexible and can consider the entire WM volume as the region of interest, eliminating the need for pre-processing steps in the form of *a priori* definition of the ROIs (using atlas-based label propagation for example). However, such an automated approach to analysing the entire WM volume across multiple subjects carries significant computational burden. Consequently, for the purpose of this study, we restrict our attention to two WM regions, namely, the cingulum and corpus callosum. An atlas-based label propagation approach is used to segment the WM ROIs from all subjects' FA images. The fractional anisotropy image of the JHU-ICBM-DTI-81 atlas [1] (Mori et al., 2008) - (Hua et al., 2008) is non-rigidly registered to each subject's FA image (following an initial affine alignment), using Nifty-Reg v 1.3.9 (Ourselin et al., 2001), (Modat et al., 2010), a deformable image registration algorithm based on cubic B-splines. Following FA image registration, the segmented labels for the cingulum and corpus cal-

---

[1] Available at: http://www.loni.usc.edu/ICBM/Downloads/Downloads˙DTI-81.shtml

losum defined on the atlas (available along with the FA atlas), are resampled to the space of each subject's FA image. In this way, labels delineating the cingulum and corpus callosum in the atlas image, are propagated to each subject's image, segmenting the ROIs (as illustrated in Fig. 1).

2.2. Algorithm Overview

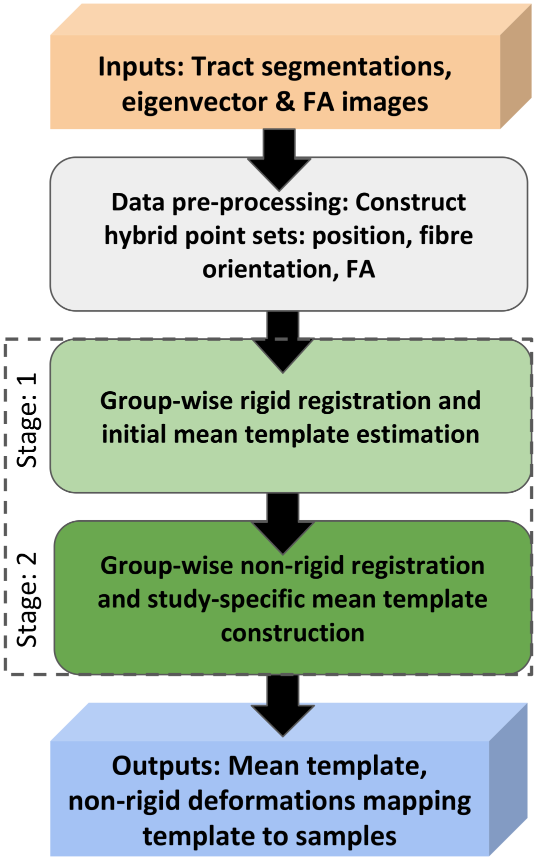

Figure 2: Summary of steps involved in the proposed framework to jointly register and cluster hybrid point sets comprising spatial positions, fibre orientations and FA values, for a WM tract/ROI. Dashed box outlines the two stages of the proposed algorithm.

The steps involved in the proposed approach are summarised by Fig. 2. For a group of $k = 1...K$ subjects to be analysed (e.g. comprising control, MCI and AD sub-groups), their tract segmentations, eigenvector and FA images were used to construct hybrid point sets $\mathbf{D}_k$, where each data point is a

7-dimensional vector denoted as $\mathbf{d}_{ki} = [\mathbf{x}_{ki}, \mathbf{n}_{ki}, f_{ki}]$. Here $\mathbf{x}_{ki}$ represents the spatial co-ordinate, $\mathbf{n}_{ki}$ represents the primary diffusion eigenvector and f_{ki} denotes the FA value for the i^{th} voxel, in the k^{th} subject's image. $\mathbf{x}_{ki}$ are consequently, densely distributed points within the volumes/ROIs. The resulting hybrid point sets were, subsequently, jointly registered and clustered by fitting an M-component hybrid mixture model (comprising Student's t, Watson and Gaussian distributions) to the data. This was achieved over two stages (as depicted in Fig. 2): (1) Group-wise rigid registration of the hybrid point sets $\mathbf{D}_k$ and mean template $\mathbf{M}$ construction; and (2) Group-wise non-rigid registration, wherein the mean template estimated in stage 1 was non-rigidly registered to each sample from all patient groups simultaneously. The similarity transformation and the non-rigid transformation, corresponding to stage 1 and 2 of the algorithm respectively, are both represented by $\mathbf{T}_k$ throughout this study. For the former, $\mathbf{T}_k = [s_k, \mathbf{R}_k, \mathbf{t}_k]$. Here, $s_k, \mathbf{R}_k, \mathbf{t}_k$ represent the scaling, rotation and translation (for the k^{th} sample), respectively, estimated in stage 1. These are used to align the hybrid point sets to the estimated mean template and initialise the subsequent non-rigid registration step (stage 2) by correcting global pose differences across the data set. Stage 2 of the algorithm estimates non-rigid transformations $\mathbf{T}_k$, defined by a linear combination of radial basis functions (with a Gaussian kernel). Together with a Gaussian kernel, the basis function weights $\mathbf{W}_k$ estimated define point-wise displacements that map the mean template to each sample within a subject group. In both stages of the algorithm, estimation of the desired registration parameters was accompanied by the joint clustering of positions, orientations and FA values. The parameters to be estimated for each of the $j = 1...M$ components of the hybrid mixture model include: $\{\mathbf{m}_j^p, \sigma_p^2, \nu_j\} = \Theta_p$, which represent mean spatial positions, their variance and the degrees of freedom, respectively, for the Student's t-distributions; $\{\mathbf{m}_j^d, \kappa_j\} = \Theta_n$, which represent the mean fibre orientations and concentration around the means, respectively, for the Watson distributions; $\{m_j^f, \sigma_f^2\} = \Theta_f$, which denote the mean FA values and FA variance, respectively, for the Gaussian distributions; and π_j which denote the mixture coefficients. Following non-rigid registration, the study-specific mean template estimated (for each WM ROI) $\mathbf{M}$ thus comprises positions, $\mathbf{m}_j^p$, orientations $\mathbf{m}_j^d$ and FA values m_j^f.

2.3. Joint Probabilistic Model of Position, Orientation and Anisotropy

The problem of joint registration and clustering of hybrid point sets is formulated as one of maximum likelihood parameter estimation, using a hybrid mixture model that approximates the joint *PDF* of spatial positions (of voxel centroids), fibre orientations, and fractional anisotropy. By assuming voxel positions, fibre orientations, and FA values to be independent and identically distributed (i.i.d), for each subject and across multiple subjects, the joint *PDF* can be approximated as a product of the individual conditional densities (Bishop, 2006) for position, orientation and FA. Consequently, by considering all data points $\mathbf{d}_{ki} \in \mathbf{D}_k$, from all K subjects, to be i.i.d. the conditional probability of an observation being sampled from an M-component HdMM is given by equation 1a. The set of all transformations (similarity or non-rigid) is represented by $\mathbf{T}_k \in \mathbb{T}$; Θ_p represents the set of model parameters associated with the Student's t-distributions $\mathcal{S}$, used to model the distribution of voxel spatial positions; Θ_n represents the parameters of the Watson distributions $\mathcal{W}$ (modelling fibre orientations); Θ_f denotes the set of parameters of the Gaussian distributions $\mathcal{N}$ (modelling FA); and $\pi_j \in \Pi$ represents the set of mixture coefficients, of the HdMM. Here and throughout, subscript $j = 1...M$ denotes mixture components and the choice of distributions indicated earlier will be justified later in this Section. Using equation (1a) the log-likelihood function is formulated as shown in equation (1b), which defines the cost function to be optimised with respect to the mixture model and transformation parameters $\{\Theta_p, \Theta_n, \Theta_f \Pi, \mathbb{T}\} \in \Psi$, to jointly register and cluster the hybrid point set data $\mathbf{D}_k \in \mathbb{D}$.

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

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

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

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

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
INITIALIZATION
Initialize $\mathbf{M}, \sigma_p^2, \sigma_f^2$ using K-means clustering.
All $\pi_j = 1/\text{M}$ and $\nu_j = 3.0$, $\kappa_j = 1.0$
procedure STAGE 1 EM:
GROUP-WISE RIGID REGISTRATION($\mathbf{D}_k, \Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{T}_k$)    ▷ EM initialized
    while Iteration < max.iterations do
        Compute $P_{kij}$                                   ▷ E-step
        Update $\mathbf{R}_k, s_k, \mathbf{t}_k$            ▷ M-step
        Update $\Theta_p, \Theta_n, \Theta_f$               ▷ M-step
    end while
    return $\Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{T}_k$
end procedure
Estimated mean template $\mathbf{M}$, mixture coefficients $\Pi$ and similarity transformations $\{\mathbf{T}_k\}_{k=1...K}$ initialise Stage 2.
procedure STAGE 2 EM:
GROUP-WISE NON-RIGID REGISTRATION($\mathbf{D}_k, \Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{W}_k$)    ▷ EM non-rigid initialized
    while Iteration < max.iterations do
        Compute $P_{kij}$                                   ▷ E-step
        Update $\mathbf{W}_k$                               ▷ M-step
        Update $\mathbf{M}, \sigma_p^2, \nu_j, \Theta_n, \Theta_f$    ▷ M-step
        Update spatial positions of each $\mathbf{D}_k$
    end while
    return $\Theta_p, \Theta_n, \Theta_f, \Pi, \mathbf{W}_k$
end procedure
```

A tractable approach to maximising equation 1b is achieved using the expectation-maximisation (EM) framework (Dempster et al., 1977), which iteratively alternates between: the expectation (E)-step, which evaluates the mixture component membership probabilities as shown in equation 1c (i.e. posterior probabilities P_{kij}^t, that define soft correspondences and are expectations of the latent variables in the model) for the observed data, given an estimate of the model parameters Ψ^t, at the t^{th} EM-iteration; and the maximisation (M)-step, which uses the computed posterior probabilities P^t_{kij} to maximise the conditional expectation of the complete-data-log-likelihood function Q (refer to equation 1d), with respect to each model parameter, resulting in revised estimates Ψ^{t+1}. As shown in equation 1d, Q for the hybrid mixture model can be expressed as a sum of contributions from each distribution and corresponding data feature (i.e. position, orientation and 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 complete algorithm for the proposed hybrid mixture model, to jointly register and cluster a group $\mathbb{D}$ of hybrid point sets, is summarized in Algorithm 1. Subsequent sections discuss each probability distribution and estimation of their associated parameters, within the proposed framework, in more detail.

2.4. Mixture Model for Primary Diffusion Axes

In addition to modelling the spatial distribution of voxels defining ROIs, the proposed approach also deals with axial data distributed over the S^2 sphere, i.e. fibre orientations defined by primary diffusion eigenvectors. GMMs and TMMs, comprising Gaussian and Student's t-distributions, respectively, are inappropriate for clustering such data and consequently, a mixture of Watson distributions, also defined over the spherical domain, is employed in this study. While Von-Mises-Fisher distributions are frequently used for clustering directional data, they are unsuitable for axial data, as they lack of antipodal symmetry. Watson distributions on the other hand, are naturally suited to model diffusion data as they are antipodally symmetric (i.e. the probability density is the same along an axis in either direction) and as the aim here is to model the *PDF* of diffusion axes at corresponding spatial locations, rather than any specific direction along the axes (Jupp and Mardia, 1989). They are fully defined by two parameters, namely, the mean/principal axis ($\pm\mathbf{m}^d$, about which the distribution is rotationally symmetric) and a scalar concentration parameter κ. The latter describes the degree of concentration about the mean axis of the distribution, with high values indicating high concentration. The *PDF* of a Watson distribution with mean direction $\mathbf{m}^d$ and concentration κ is expressed as equation 2a, for antipodally symmetric 3D unit vectors $\pm\mathbf{n}$. Here, $M(\cdot)$ represents the Kummer function. Watsons are in general more flexible than Fisher distributions as there is no positivity constraint on κ and they can be used to model both directional and axial data. (Bijral et al., 2007) proposed an efficient EM-based clustering framework for axially-distributed data, using a WMM, employed in this study to cluster fibre orientations.

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

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

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

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

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

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

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

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

Maximum likelihood estimates for the associated parameters are evaluated at each M-step of the algorithm by maximising the expectation of the complete data likelihood (equation 3a), with respect to $\mathbf{m}_j^d$ and κ_j, subject to the constraint ${\mathbf{m}_j^d}^T\mathbf{m}_j^d = 1$ (Bijral et al., 2007). This is achieved by maximising the Lagrangian form of Q shown in equation 3b. Mean directions $\mathbf{m}_j^d$ are estimated numerically, using fixed-point iteration, to solve the non-linear equation (shown in equation 3c) obtained from differentiating Q (3b) with respect to $\mathbf{m}_j^d$. κ_j on the other hand is approximated (refer to equation 3e) using the continued fraction representation for the ratio of, the derivative of the Kummer function and the function itself, i.e. $\frac{M'(\kappa_j)}{M(\kappa_j)}$ (equation 3d). In a recent study (Sra and Karp, 2013) derived two-sided bounds for approximating κ, particularly useful when dealing with high dimensional data. However, for 3D data (as in this study) the approximation presented in equation 3e is sufficient (as noted by (Bijral et al., 2007),(Sra and Karp, 2013)). Better approximations for κ_j may be obtained using numerical techniques such as Newton's method, however, at the expense of significant increase in computational burden.

2.5. Mixture Model for Fractional Anisotropy

The distribution of voxel-wise FA in WM ROIs across a population, is modelled using a univariate GMM. GMM was chosen as the resulting model-predicted FA values at the estimated spatial correspondences, across subjects, is guaranteed to be normally distributed — a useful property for subsequent statistical analyses, as noted in (Smith et al., 2006), where the authors also show that FA values at corresponding spatial positions across populations are indeed approximately normally-distributed. Additionally, GMMs are computationally efficient, as analytical solutions exist for revising estimates of the associated model parameters (mean m_j^f and variance σ_f^2 of FA), at each EM-iteration. Assuming the observed FA values f_{ki} at voxels in ROIs, across 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 expressed as equations 4a, 4b. Consequently, the conditional expectation of the complete data log likelihood Q, maximised with respect to the model parameters associated with the Gaussian distributions in the mixture, is given by equation 4c (only terms dependent on m_j^f and σ_f^2 are retained in Q). As GMM-based clustering of FA values is performed jointly with the registration of WM ROIs, and clustering of voxel positions and the associated fibre orientations, the influence of a Gaussian component in the mixture model is automatically limited to its local neighbourhood. This helps ensure that only voxels in close proximity to each other contribute significantly to the estimation of mean FA values at each mixture component. Estimates for the GMM parameters m_j^f and σ_f^2 in the M-step of the algorithm are derived analytically, as shown in (Bishop, 2006).

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

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

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

2.6. Rigid Alignment and Template Construction

Previously, we proposed a group-wise rigid point set registration framework based on Student's t-mixture model (Ravikumar et al., 2016), (Ravikumar et al., 2018), which exploits the inherent robustness of Student's t-distribution for robust registration of shapes in the presence of missing data and significant proportions of outliers. Additionally, in a more recent study (Ravikumar et al., 2017) we proposed a variant of the hybrid mixture model-based registration framework formulated in this study. In (Ravikumar et al., 2017) Von-Mises-Fisher distributions were used in place of the Watson distributions used in this study, to model directional data such as surface normal vectors, for rigid and non-rigid shape registration. A Watson distribution-based variant of (Ravikumar et al., 2017) is employed in the present study as an initial step, to rigidly align WM ROIs (hybrid point sets representing voxel centroid positions, fibre orientations and FA values), segmented from all subjects' images, whilst simultaneously estimating a mean model. The latter subsequently serves as an unbiased, study-specific template for non-rigid registration. Rigid group-wise registration is preferred to a pair-wise approach as it enables estimation of a mean template and the desired similarity transformations in an unbiased manner. Rigid alignment also helps initialise the subsequent non-rigid registration by recovering global differences in pose between sample shapes, and establishes soft correspondences across subjects.

Group-wise point set registration using mixture models assumes that the point sets to be aligned are transformed observations of a central mixture model (which we refer to as the mean template) (Gooya et al., 2015). Consequently, the optimal transformations that align the template to the group of shapes are those that maximise the likelihood of the data (or equivalently, minimise the negative log-likelihood function). The desired similarity transformations are thus iteratively refined along with the template itself at each M-step of the algorithm. The main differences between EM-based estimation of parameters for TMMs and GMMs are: (1) TMMs have two associated latent variables (as opposed to just one with GMMs, which represent the mixture component membership of the data), whose expectations are evaluated in the E-step and used to compute a set of corrected posterior probabilities $P^{\star}_{kij}$, estimated identically to (Ravikumar et al., 2016), (Ravikumar et al., 2018) (refer to the Appendix); and (2) Student's t-distributions are defined by three parameters (as opposed to two for Gaussians). The additional parameter is referred to as the degrees of freedom/shape parameter ν, which is responsible for controlling the heaviness of the tails of the distribution (and consequently, the degree of robustness to outliers). The behaviour of the t-distribution tends towards that of a Gaussian as $\nu \to \infty$.

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

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

The joint *PDF* of voxel positions $\mathbf{x}_{ki} \in \mathbf{X}_k$, across all K subjects in a group (denoted, $\mathbf{X}_k \in \mathbb{X}$), is given by equation 5a (assuming they are i.i.d transformed observations of a TMM). In equation 5a, $\mathbf{T}_k$ represents the similarity transformation (comprising rotation $\mathbf{R}_k$, scaling s_k and translation $\mathbf{b}_k$), to align the positions $\mathbf{m}_j^p$ defining the mean template, to the k^{th} sample in the group. In our recent work (Ravikumar et al., 2016), (Ravikumar et al., 2018), we showed that the form of Q to be maximised, to estimate the desired similarity transformations $\mathbf{T}_k \in \mathbb{T}$ and mixture component parameters Θ_p, is given by equation 5b. Closed form expressions are derived for the M-step update equations of all TMM and transformation parameters, which are presented in the Appendix. Fibre orientations and FA are invariant to translation $\mathbf{b}_k$ and scaling s_k, consequently, these transformation parameters are estimated identically as in (Ravikumar et al., 2016), (Ravikumar et al., 2018). Although the former are rotationally dependent, the contribution of fibre orientations to the estimation of $\mathbf{R}_k$ is ignored as the direction of the observed diffusion eigenvectors tend to be random. Consequently, rotations $\mathbf{R}_k$ are derived based on the spatial positions of hybrid point sets alone, by optimising the form of Q shown in equation 5b, similar to (Ravikumar et al., 2016), (Ravikumar et al., 2018). However, following estimation of the desired rotations $\mathbf{R}_k$ at each EM-iteration, the current estimate of the mean template is transformed by rotating both spatial positions $\mathbf{m}_j^p$ and their associated fibre orientations $\mathbf{m}_j^d$, to align it with the k^{th} sample in the group. Additionally, it is important to note that, while the fibre orientations and FA values are ignored in the derivation of the desired transformation parameters, they are intrinsic to the estimation of the posterior probabilities P_{kij} at each E-step of the algorithm. Consequently, they drive the estimation of soft correspondences, which in turn affect the transformations evaluated at each M-step of the algorithm.

2.7. Non-rigid Point Set Registration

Coherent point drift (CPD) (Myronenko and Song, 2010) is a well known pair-wise, non-rigid point set registration technique based on motion coherence theory. The spatial transformation between two point sets is considered to be an initial position (of the moving point set) plus some unknown displacement (or velocity) function mapping it to the target point set. This unknown transformation is regularized using Tikhonov regularization, to ensure estimation of a smooth displacement function, and is expressed in the Reproducing Kernel Hilbert Space (RKHS). Using variational calculus, Myronenko and Song (2010) showed that the optimal displacement function under such smoothness constraints, can be expressed as a linear combination of kernel functions (i.e. Gaussian radial basis functions). Similarly, our approach also employs Gaussian radial basis functions to parametrize the non-linear transformation, and the associated basis function weights are estimated by maximising the likelihood function using EM (similar to estimation of rotation, translation and scaling, in the rigid registration approach discussed in the previous section). CPD models the target point set as a transformed observation of the source point set (i.e. the point set to be registered). The latter is consequently considered to represent the centroids of a Gaussian mixture model, which is fit to the former using EM, and the transformation necessary to register the source to the target set is estimated as parameters of the mixture model. In addition to the Gaussian components in the mixture model, CPD incorporates a uniform distribution component to model noise/outliers present in the data. This confers added robustness to the registration process. However, a user-defined parameter is used to balance the weight of the uniform distribution component relative to its Gaussian counterparts, which needs to be tuned for different applications and data sets, for optimal registration. To ameliorate the need for parameter tuning, we employ Student's t-distributions in place of the Gaussian and uniform distributions used in CPD and re-formulate the approach in a group-wise non-rigid registration framework. As stated previously, the robust nature of t-distributions makes them well suited to registration applications requiring automatic robustness to outliers. A similar approach for pair-wise registration of 2D/3D point sets was proposed previously, by (Zhou et al., 2014).

The mean tract template estimated during the initial group-wise rigid registration step (discussed in section 2.6), is non-rigidly registered to each patient group (AD, MCI and HC) independently. The desired non-rigid transformations are defined with respect to the template $\mathcal{M}$ as: $\mathbf{M} + v^k(\mathbf{M})$ (considering spatial positions $\mathbf{m}_j^p$ alone), where v is a displacement function mapping the template to the k^{th} sample in the group. In (Myronenko and Song, 2010) the authors show that the desired displacement field is constrained to be smooth by employing Tikhonov regularization (or regularizing the norm of v, expressed in RKHS). This forces points in close proximity, to move together. Regularization of this nature is akin to employing a prior on the displacement field of the form $p(v) = \exp^{-\frac{\lambda}{2}\phi(v)}$, where $\phi(v)$ represents the regularization term and λ controls the trade-off between registration accuracy and smoothness of the deformation field. The prior on the displacement field is incorporated into the TMM, resulting in a log-likelihood function expressed as equation 6a. As stated previously, (Myronenko and Song, 2010) show that the function v, which maximises the data likelihood, can be expressed as a linear combination of radial basis functions (refer to equation 6b). Consequently, to register the study-specific mean template to each sample from all patient groups simultaneously, the objective function to be maximised with respect to the basis function weights $w_{kj} \in \mathbf{W}_k$, is expressed as shown in equation 6c, where $\mathbf{G}$ represents the Gaussian kernel/Gram matrix. The basis function weights required to register the study-specific mean template to each sample are estimated as shown in 6d, by computing the derivative of Q with respect to the weights, similarly to (Myronenko and Song, 2010). In equation 6d $\mathbf{P}_k^s = \sum\limits_{i=1}^{N_k} P_{kij}^{\star t}$, $\mathbf{P}_k^T$ is the transpose of the posterior probability matrix for the k^{th} sample, diag is a diagonal matrix, and $\mathbf{I}$ is the identity matrix. Subsequently, the mean template is deformed to match each k^{th} sample (in the entire population) as described by equation 6e.

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

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

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

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

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

Following convergence of the non-rigid registration step, a study-specific mean template comprising, mean spatial positions, mean fibre orientations and mean FA values (representative of the entire population of AD, MCI and HC subjects), is estimated. Additionally, point-wise displacements mapping this mean template to each sample in the entire population (as described by equation 6e), is also obtained, thereby establishing the spatial correspondences used for any subsequent inter-group statistical comparisons. These correspondences play a similar role to the mean FA skeleton estimated in TBSS. In addition to these spatial correspondences, we also compute"model-predicted" values for FA and fibre orientation, at each correspondence, for all subjects. These model-predicted values are probabilistic weighted averages of the FA values and fibre orientations associated with the voxels in the original DTI-derived FA and eigenvector images (i.e. the original hybrid point sets). The weighted averages are assigned to each spatial correspondence point and are analogous to the 'soft/probabilistic spatial correspondences' estimated in previous studies, such as (Hufnagel et al., 2008), (Gooya et al., 2015) for example. Here, the weights are defined by the posterior probabilities estimated for each voxel, of each subject's original FA and eigenvector images (P_{kij}), following non-rigid registration. Equations describing the estimation of model-predicted FA values and fibre orientations are included in the Appendix (refer to equations 19a - 19b). Although point set registration techniques are typically employed to register 3D point sets (comprising only spatial positions) representing the surface/boundary of an object, this study incorporates additional image-based features (such as fibre orientations and FA values), that enable registration of dense point sets, defined by voxel centroids located at the boundary of, and within a region of interest.

3. Results and Discussion

3.1. Rigid Registration Accuracy

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

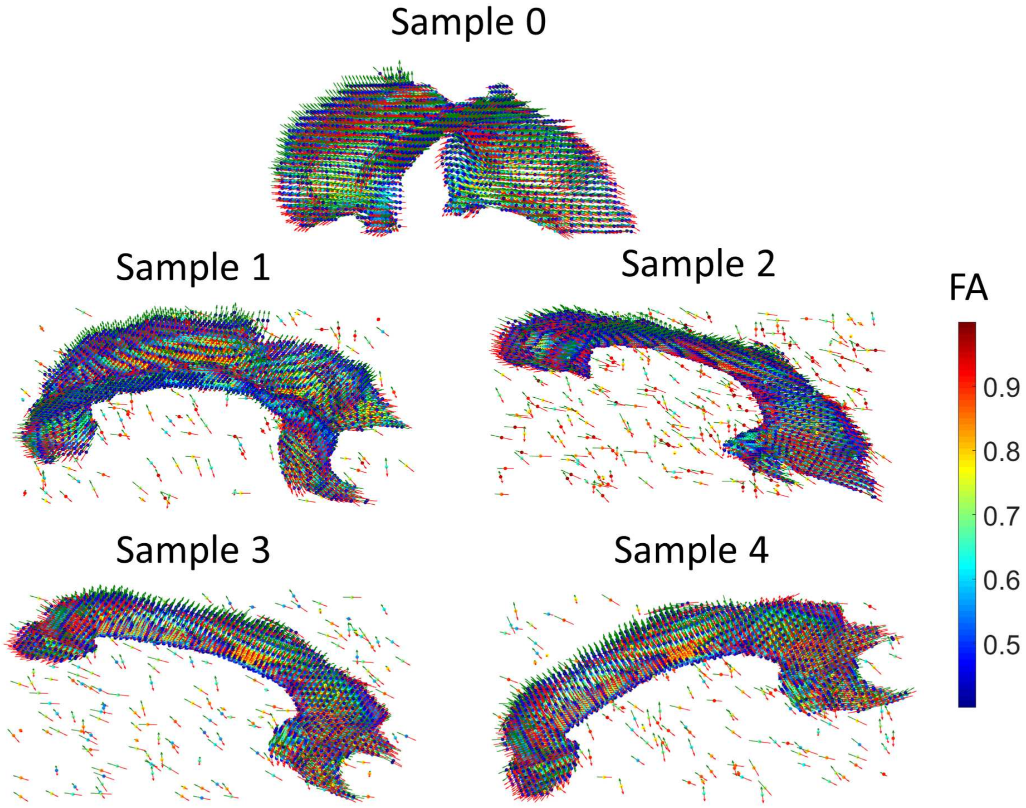

Figure 3: Synthetic corpus callosum data set comprising: Sample 0, the ground truth hybrid point set; and Samples 1-4, which are rotated and modified versions of Sample 0.

The rigidly transformed point sets were also modified by the addition of varying proportions of random outliers (comprising positions, orientations and FA values). Fibre orientations associated with the outliers were generated from normalized 3D points. While their FA values were uniformly sampled within the range $[0.2, 0.8]$. The FA values associated with the voxels of each modified hybrid point set were also varied by ± 0.1, relative to the ground truth point set. This was necessary in order to emulate real data as FA values typically vary at corresponding anatomical locations, between subjects. This process was repeated 10 times, to generate 10 unique synthetic data sets (each comprising one ground truth and 4 modified, unique samples), which were subsequently rigidly aligned using the proposed Watson distribution-based HdMM algorithm (i.e. 10 distinct registration experiments). Random rotations and proportions of outliers were generated for each experiment, within the range of $[-30°, 30°]$ and $[2\%, 5\%]$, respectively (as illustrated in Fig. 3). Table 1 summarises the mean ground truth euclidean distances between Samples 1-4 and Sample 0 across all 10 experiments (prior to registration), and the axes about which rotations were applied to generate each sample in each experiment. The average rigid registration errors following alignment of the synthetic data sets (with $M = 2000$ mixture components) using the proposed framework are also reported in Table 1.

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

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

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

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

3.2. Model Quality

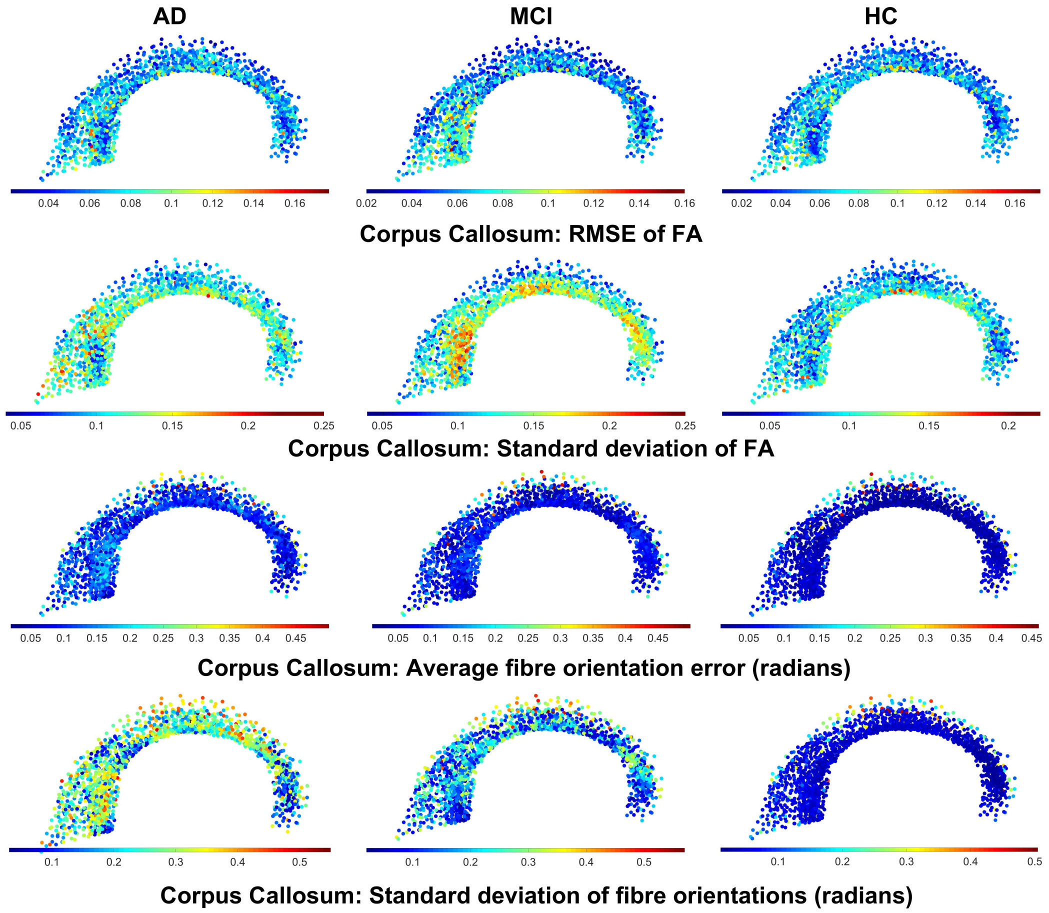

Figure 4: Model quality evaluated for the corpus callosum, independently for AD, MCI and HC groups, using $M = 2000$ mixture components. Rows one and two: RMSE of FA and standard deviations of the same computed across subjects; Rows three and four: Angular errors for fibre orientations (in radians) and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

The ability of the HdMM to model DTI-derived quantities was assessed using clinical data, acquired from the VPH-DARE@IT prospective cohort, described in section 2.1. Specifically, model quality was quantified by evaluating the similarity between the estimated correspondences (resulting from non-rigidly registering the the unbiased study-specific mean template to each sample from all patient groups) and the nearest neighbour voxels in the corresponding subject's original FA and eigenvector images. FA accuracy is quantified as the root-mean-squared error (RMSE), evaluated between the model-predicted and original voxel-wise FA values, across all correspondences, for each subject. The group-wise average error (for each subject group) of FA was subsequently computed. The minimum arc length (measured in radians) between two unit vectors is used to measure the accuracy of local fibre orientation in a similar manner. As discussed in section 2.4, the proposed framework models axial data rather than directional data. When computing fibre orientation errors, corresponding unit vectors between the model-predicted and original voxel-wise eigenvectors are first identified. This is achieved by evaluating their scalar product and ensuring it is positive — i.e. if the dot product is negative, the antipodal counterpart of the model-predicted vector is used instead. The resulting measure thus quantifies the angular error in fibre orientation between the model-predicted and original voxel-wise data (in the eigenvector image), for each subject. These measures represent registration residuals which describe the quality of correspondences established by the proposed HdMM (i.e. how well the HdMM can model the observed DTI-derived data), and only indirectly reflect registration 'accuracy'. To provide a more general view of registration accuracy, the mean-squared distance (MSD, formulated as shown in the Appendix), quantifying spatial position errors was also evaluated between the registered study-specific mean template and the original hybrid point sets from all patient groups (Note: MSD values were evaluated between dense volumetric point sets). It is important to note that the model-predicted values for FA and fibre orientation assigned to the spatial correspondences established using the proposed approach, are probabilistic in nature (as discussed in section 2.7). Consequently, they reflect the DTI-derived quantities of voxels located in the local spatial neighbourhood of the correspondences.

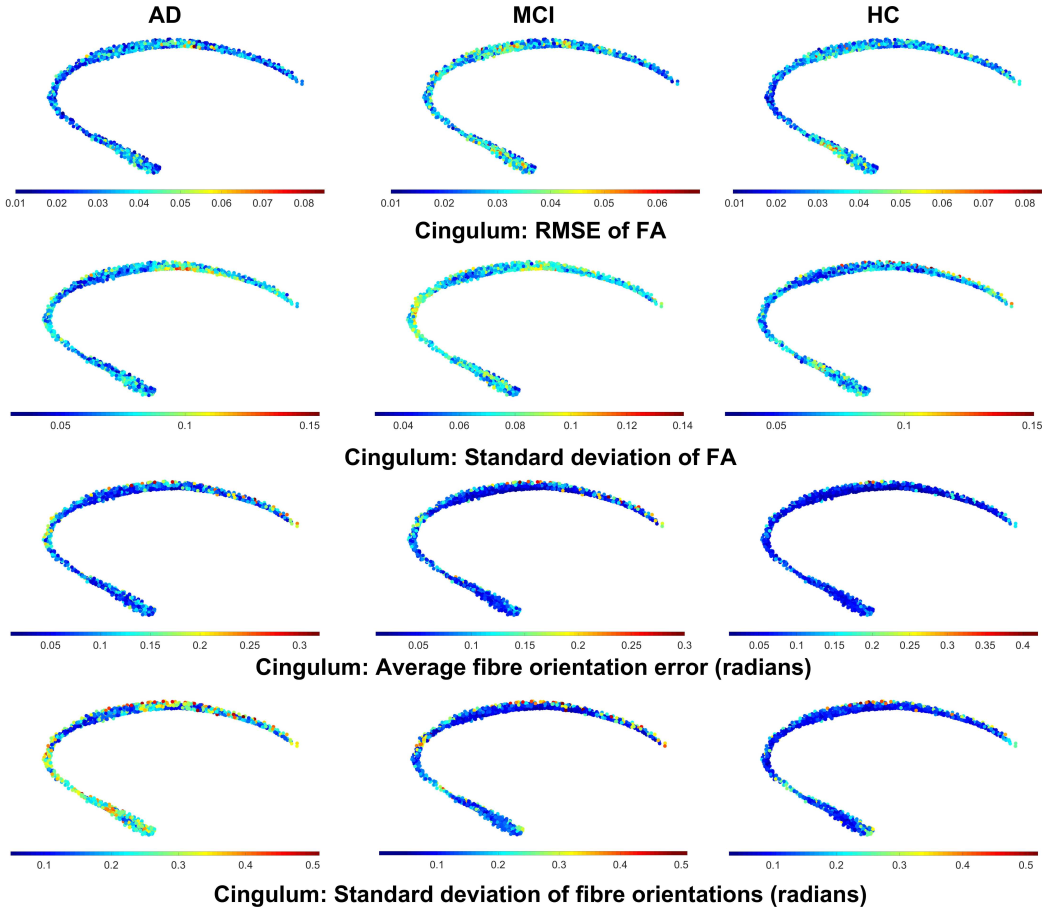

Figure 5: Model quality evaluated for the cingulum, independently for AD, MCI and HC groups, using $M = 1500$ mixture components. Rows one and two: RMSE of FA and standard deviations of the same computed across subjects; Rows three and four: Angular errors for fibre orientations (in radians) and standard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

Results summarizing the ability of the proposed framework to model DTI-derived quantities across all 60 subjects are presented in Fig. 4 - Fig. 7 and Tables 2 - 7. Fig. 4 and Fig. 5 help visualise the spatial distribution of mean registration errors and the standard deviations of FA values and fibre orientations observed across subjects within each patient group, for the corpus callosum and cingulum, respectively. We would like to highlight that while samples from all patient groups were registered simultaneously, the registration errors presented in Fig. 4 - Fig. 7 and Tables 2 - 7 alone were evaluated for each patient group separately. This was done in order to identify any group-specific trends that exist in the registration accuracy afforded by the proposed approach. In Fig. 4 and Fig. 5 the RMSE values of FA were computed by averaging across subjects in each group, at each corresponding position. Similarly, the standard deviations were also evaluated point-wise across subjects for each group. The depicted mean angular errors were averaged across subjects, quantifying the fibre orientation accuracy at each corresponding position, and point-wise estimates for the standard deviations in fibre orientation were also evaluated. The presented standard deviations in Fig. 4 and Fig. 5 aid in interpretation of the error measures evaluated, and provide a frame of reference, for both WM regions. The spatial distribution of the variation in FA across subjects within each patient group, was evaluated as follows: (a) the nearest neighbour voxel in the original hybrid point sets were first identified based on the spatial positions estimated by non-rigid registration of the study-specific mean template, to each corresponding sample ; (b) the FA values associated with the voxels identified for each subject were in turn used to compute the standard deviation across subjects, within each patient group; and (c) these values were subsequently mapped on to the study-specific mean template estimated for the corpus callosum and cingulum, for easy comparison with the registration errors plotted in a similar manner, as shown in Fig. 4. Similarly, the standard deviations in fibre orientations about the mean, were also evaluated across subjects, within each patient group, for both WM regions. Here, the difference between the mean fibre orientation estimated at each correspondence point in the study-specific mean template, and the nearest neighbour voxels identified (refer to (a) above) in the original hybrid point sets, was evaluated as the minimum arc length (in radians) between each other. This in turn was employed to compute the standard deviation in fibre orientations and visualize their spatial distribution across both WM regions.

Based on these results, the proposed HdMM is considered to establish valid correspondences across patients, as the estimated fibre orientation and FA errors are low across the majority of correspondences. Fibre orientation errors were consistently < 0.2 radians across most correspondences for both WM ROIs (refer to first and third row in Fig. 4). FA errors meanwhile, were < 0.1 for the corpus callosum and cingulum (refer to second and fourth row in Fig. 4), across all patient groups. For the former WM region, FA errors below 0.1 were produced for $> 92\%$ of all established correspondences. While for the latter, all correspondences, had errors below 0.1. Fibre orientation errors were < 0.2 across $> 94\%$ of correspondences estimated for both WM ROIs, in all patient groups. Errors of this magnitude are considered reasonable as the model-predicted FA values and fibre orientations evaluated at correspondences are based on the soft-assignment approach (refer to section 2.7), using the estimated posterior probabilities. Consequently, they reflect weighted averages of FA and fibre orientations of neighbouring voxels. FA variations of ≈ 0.1 may occur due to partial volume effects at WM-GM and WM-CSF interfaces (Smith et al., 2006), particularly when WM tracts/ROIs are very thin compared to the voxel size (often the case following dementia-related atrophy of brain tissue), potentially further contributing to the observed errors. Additionally, significant variations in DTI-data in a select few cases within individual patient groups may be another source of the high average errors evaluated, in a small proportion of correspondences. These results are further supported by the standard deviations of FA and fibre orientations depicted in Fig. Fig. 4 and Fig. 5, which highlight the high degree of variation in FA and fibre orientations (across subjects), respectively, relative to the corresponding errors evaluated across both WM regions.

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

Tables 2 - 7 report the average spatial position, fibre orientation and FA errors evaluated across correspondences and subjects. Statistically significant reduction in mean spatial position errors across experiments conducted using differing model complexities (i.e. different number of mixture components) are highlighted in bold in Tables 2 and 5, considering a significance level of 5%. In Tables 4 and 7 the reported mean FA errors were estimated by first computing the RMSE, this time averaging across correspondences, and subsequently computing the mean RMSE across subjects. Tables 3 and 6 summarise the mean angular error values, first averaged across correspondences and subsequently across subjects. These alternate error measures are presented to assess model quality of the HdMM across regions, and complement the correspondence-wise errors presented in Fig. 4 - 5. From Tables 2 - 7, the number of mixture components required to adequately characterise the entire population was identified as $M = 1500$ and $M = 2000$ for the cingulum and corpus callosum, respectively. The fibre orientation and FA errors depicted in Fig. 4 - 5 were evaluated using these values. All subsequent inter-group statistical analyses conducted employed these model complexities for 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	**0.67 ± 0.09**	**0.66 ± 0.07**	**0.64 ± 0.07**
2000	**0.65 ± 0.09**	**0.63 ± 0.07**	**0.62 ± 0.07**

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 ± 0.01	0.06 ± 0.01	0.06 ± 0.01
600	0.06 ± 0.01	0.06 ± 0.01	0.06 ± 0.01
1200	0.05 ± 0.01	0.05 ± 0.01	0.05 ± 0.01
1500	0.05 ± 0.01	0.05 ± 0.01	0.05 ± 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 ± 0.17	1.14 ± 0.10	1.09 ± 0.12
1000	0.99 ± 0.15	0.98 ± 0.09	0.94 ± 0.10
1500	**0.91 ± 0.13**	0.90 ± 0.08	**0.85 ± 0.09**
2000	**0.86 ± 0.12**	**0.85 ± 0.07**	**0.81 ± 0.08**

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 ± 0.19	0.10 ± 0.14	0.06 ± 0.01
1000	0.13 ± 0.19	0.13 ± 0.16	0.05 ± 0.01
1500	0.12 ± 0.19	0.09 ± 0.13	0.05 ± 0.01
2000	0.12 ± 0.18	0.09 ± 0.13	0.05 ± 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 ± 0.03	0.11 ± 0.02	0.10 ± 0.01
1000	0.10 ± 0.03	0.10 ± 0.02	0.09 ± 0.01
1500	0.09 ± 0.03	0.09 ± 0.03	0.08 ± 0.004
2000	0.09 ± 0.03	0.08 ± 0.03	0.07 ± 0.01

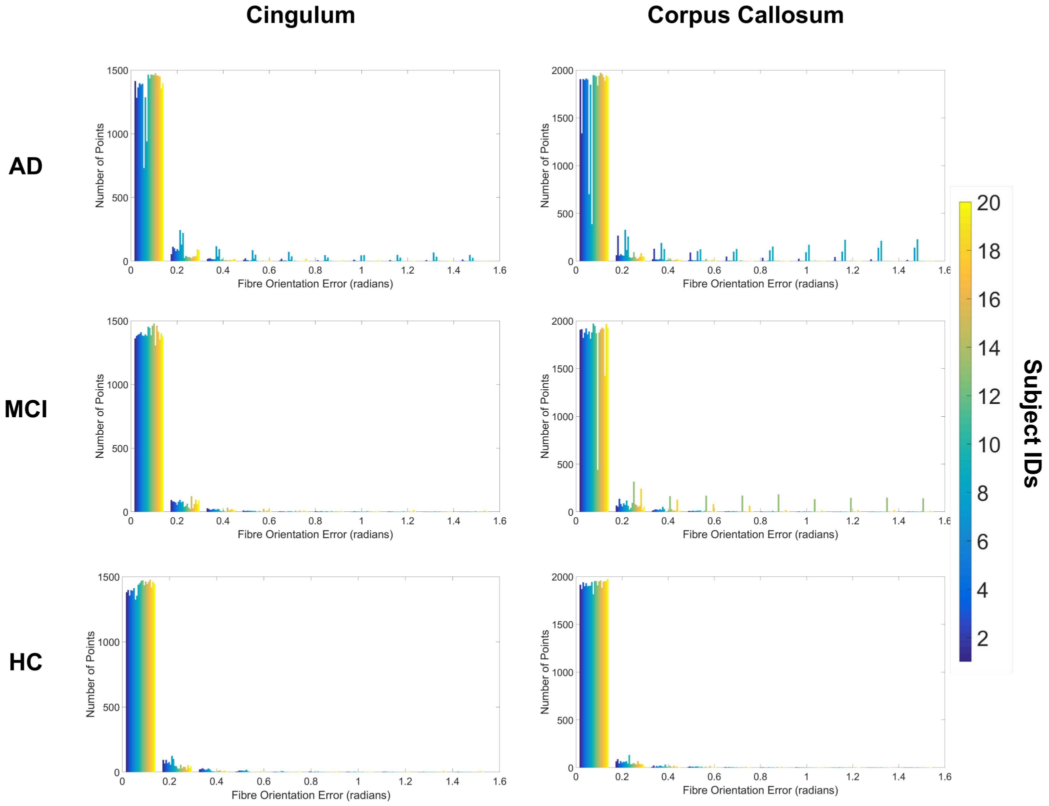

Figure 6: Histograms of fibre orientation errors for each subject in AD, MCI and HC groups, evaluated between established correspondences and ground truth voxels.

Results in Fig. 6 and 7 indicate that the proposed framework achieves low fibre orientation and FA errors at each estimated correspondence, for all subjects in the HC group (for both WM ROIs). The estimated correspondences were less accurate for two cases in the AD group (for both cingulum and corpus callosum) and for one case in the MCI group (only corpus callosum), which is attributed to significant variation in fibre orientations and FA values in these cases and ROIs, relative to the remaining samples in their corresponding patient groups. As discussed previously, this may be a result of varying degrees of pathology-induced changes in these cases relative to the rest of their group. Consequently, the accuracy of the HdMM when fitting to these few cases, is reduced. The proposed framework, however, established accurate correspondences for the remaining samples in the AD and MCI groups across both WM ROIs. The high deviations from the mean fibre orientation errors in the corpus callosum for these groups (Table 6) are thus attributed to the outlier subjects identified from the corresponding histograms (Fig. 6). Similarly, for the cingulum, the high standard deviations observed for the AD group are attributed to the two subjects mentioned above. However, no apparent outliers were identified in the MCI group based on the registration errors and, by extension, the mean FA and fibre orientation errors reported in Tables 4 and 3, are low and consistent with their corresponding histogram plots (Fig. 7 and Fig. 6).

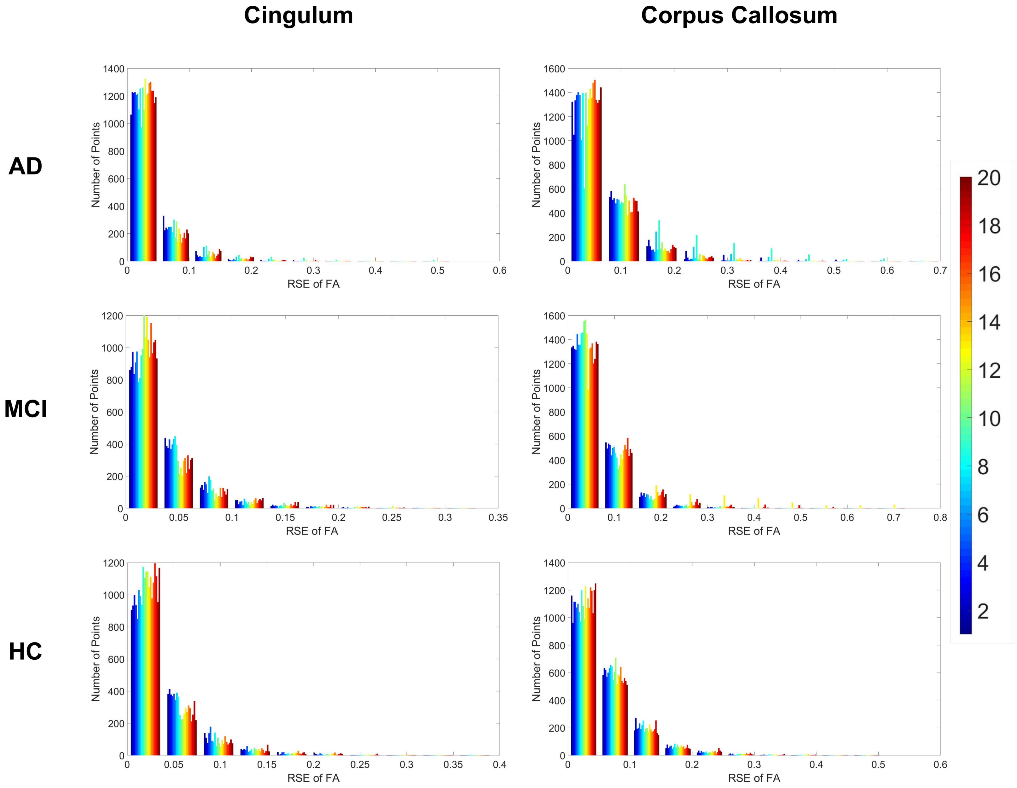

Figure 7: Histograms of root-squared-error (RSE) of FA for each subject in AD, MCI and HC groups, evaluated between established correspondences and ground truth voxels.

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

3.3. Group Comparisons

The ability of the proposed framework to identify significant differences between patient groups was assessed by comparing each pair of patient groups in terms of the variation in FA. These results were compared with those obtained from the widely used TBSS approach. Un-paired two-sample t-tests, assuming equal variances, were performed to compare FA values at corresponding spatial positions between patient groups. The procedure proposed in (Benjamini and Yekutieli, 2001) was used to correct for multiple comparisons by controlling the false discovery rate (FDR) for the set of hypothesis tests. The desired FDR was fixed at 1% for all experiments. However, no statistically significant reduction in FA was identified between any of the groups, using the proposed approach, TBSS and VBM.

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

As discussed previously, the primary advantage of the proposed HdMM framework is its ability to model fibre orientations and facilitate their comparison across multiple subjects, which is not offered by conventional approaches such as TBSS and VBM. Furthermore, the proposed method does not require extraction of fibre trajectories using tractography in order to model fibre orientations as it operates directly on the raw DTI-derived eigenvectors, unlike state-of-the-art approaches such as those proposed in (Garyfallidis et al., 2015) and (ODonnell et al., 2017). Inter-group statistical comparisons of the angular deviation in fibre orientations, relative to study-specific mean template, were also conducted. Here, the angular deviation of the model-predicted fibre orientations at each spatial correspondence was first evaluated relative to the corresponding mean fibre orientation (for patients from all groups), as the minimum arc length between unit vectors. Subsequently, these deviations were compared between each pair of patient groups, while correcting for multiple comparisons using FDR. However, as with the FA analyses, no statistically significant differences were identified.

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

A current limitation of the proposed approach is it only enables analysis of DTI data generated using a single tensor model. However, the proposed HdMM framework could be imbued with greater flexibility by replacing the Watson distributions with the Kent or the general 8-parameter Fisher-Bingham distribution, to model multi-fibre (or crossing fibre) regions by fitting to orientation distribution functions obtained from high angular diffusion images. Extensions to the Von-Mises-Fisher mixture model for example, have been proposed previously to accommodate antipodal symmetry and model diffusion ODFs (McGraw et al., 2006).

The sensitivity and discriminative capacity of the proposed framework in comparison to existing approaches requires further investigation and validation, which will be the subject of future work. Natural extensions to the proposed framework include whole WM volume analysis across multiple subjects, WM parcellation, and automatic region-of-interest analysis, to name a few. As discussed previously, the proposed approach can be employed to analyse the entire WM volume across subjects, i.e. *a priori* definition of ROIs is not required, though the computational burden at present is substantial. Such an approach naturally leads to the unsupervised parcellation of WM into distinct clusters defined by the centroids of the HdMM, across multiple subjects. This in turn provides a mechanism for automatic ROI-type analyses, as the generated clusters for each subject will correspond to similar WM regions in terms of spatial position, fibre orientation and FA (or some other scalar measure of interest). Furthermore, by employing a suitable prior/atlas containing pre-defined labels for WM tracts of interest, the presented framework could be employed for automatic tractography segmentation (similar to (O'Donnell and Westin, 2007)). The proposed approach can also be employed to track and identify localised changes in WM over time for a single subject, resulting from the progression of neuro-degenerative disorders such as dementia, for example. Although WM changes in the brain were considered in this study, the generic nature of the proposed framework permits its application to other organs exhibiting tissue anisotropy, such as cardiac diffusion data, and modelling bone micro-architecture. Additionally, it can be employed for a variety of other applications, such as vessel centerlines-based image registration, as demonstrated by our recent study (Bayer et al., 2018).

4. Conclusions

In this study, a Watson-distribution based hybrid mixture model was presented for jointly registering and clustering DTI-derived data from multiple subjects and patient populations. This approach was shown to model the observed fibre orientations and FA values accurately for all subjects within the HC group, for both of the studied WM ROIs, namely, the cingulum and corpus callosum. Registration to subjects in AD and MCI groups was successful for the majority of cases, with two in the former and one in the latter resulting in high registration errors, due to significant pathology induced changes in these cases. Group comparisons of FA values in the WM ROIs using the proposed approach showed no statistically significant reductions in FA between the AD, MCI and HC groups, as with TBSS and VBM. Similarly, no significant variations in fibre orientation were identified between patient groups. However, the proposed method has potential for use in a variety of applications involving statistical analysis of diffusion data. Its generic and flexible nature make it well suited to a variety of other computer vision and medical image analysis tasks, such as: point set registration with the integration of surface normals, vessel-based image registration, joint registration and clustering of geometries with associated velocity fields (estimated from computational fluid dynamic simulations for example) and texture mapping, to name a few. The fidelity and extensibility of the proposed framework is thus compelling as a general tool for multi-dimensional medical image analysis.

Acknowledgments

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

Appendix

M-step update equations for the Student's t-distribution parameters in the HdMM and rigid registration parameters at the $(t+1)^{\text{th}}$ EM-iteration, discussed in section 2.6, are derived by maximizing the complete data log-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:

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

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

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

$$< \partial Q, \partial\boldsymbol{\mu}_j >= [-\frac{1}{2}\sum_{k,i} P^{\star}_{kij}\Delta^{\boldsymbol{\mu}_j+\partial\boldsymbol{\mu}_j}_{kij}] - [-\frac{1}{2}\sum_{k,i} P^{\star}_{kij}\Delta^{\boldsymbol{\mu}_j}_{kij}] \tag{9a}$$

$$< \partial Q, \partial\boldsymbol{\mu}_j >= \sum_{k,i} P^{\star}_{kij}[(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T s_k\mathbf{R}_k]\partial\boldsymbol{\mu}_j \tag{9b}$$

$$< \partial Q, \partial\boldsymbol{\mu}_j >= 0 \implies \sum_{k,i} P^{\star}_{kij}[(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T s_k\mathbf{R}_k] = 0 \tag{9c}$$

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

$$\boldsymbol{\mu}_j = \frac{\sum\limits_{k,i} P^{\star}_{kij} s_k^{-1}\mathbf{R}^T(\mathbf{x}_{ki} - \mathbf{t}_k)}{\sum\limits_{k,i} P^{\star}_{kij}} \tag{9e}$$

- Estimation of model variance σ^2:

$$\frac{\partial Q}{\partial\sigma^2} = \frac{\partial\sum\limits_{k,i,j}[-\frac{P_{kij}}{2}[\log(\sigma^6)] - \frac{P^{\star}_{kij}}{2}[\Delta_{kij}]]}{\partial\sigma^2} = 0 \tag{10a}$$

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

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

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

$$< \partial Q, \partial\mathbf{t}_k >= [-\frac{1}{2}\sum_{i,j} P^{\star}_{kij}\Delta^{\mathbf{t}_k+\partial\mathbf{t}_k}_{kij}] - [-\frac{1}{2}\sum_{i,j} P^{\star}_{kij}\Delta^{\mathbf{t}_k}_{kij}] \tag{11a}$$

$$< \partial Q, \partial\mathbf{t}_k >= \sum_{i,j} P^{\star}_{kij}[(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j - \mathbf{t}_k)^T]\partial\mathbf{t}_k \tag{11b}$$

$$< \partial Q, \partial\mathbf{t}_k >= 0 \implies \sum_{i,j} P^{\star}_{kij}(\mathbf{x}_{ki} - s_k\mathbf{R}_k\boldsymbol{\mu}_j)^T = \sum_{i,j} P^{\star}_{kij}\mathbf{t}_k^T \tag{11c}$$

$$\mathbf{t}_k = \frac{\sum\limits_{i,j} P^{\star}_{kij}\mathbf{x}_{ki}}{\sum\limits_{i,j} P^{\star}_{kij}} - s_k\mathbf{R}_k\frac{\sum\limits_{i,j} P^{\star}_{kij}\boldsymbol{\mu}_j}{\sum\limits_{i,j} P^{\star}_{kij}} \tag{11d}$$

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

$$\mathbf{t}_k = \mathbf{d}_k - s_k\mathbf{R}_k\mathbf{m}_k \tag{11e}$$

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

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

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

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

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

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

  $$\mathbf{C}_k = \sum_{i,j} P^{\star}_{kij}\tilde{\mathbf{x}}_{ki}\tilde{\mathbf{m}}_{kj}^T \tag{12d}$$

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

- Estimation of scaling s_k:

  $$\frac{\partial Q}{\partial s_k} = -\frac{1}{2}\frac{\partial \sum\limits_{i,j} P^{\star}_{kij}\Delta_{kij}}{\partial s_k} = 0 \tag{13a}$$

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

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

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

- Estimation of degrees of freedom ν_j:

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

$O.T.$ summarizes terms in Q independent of ν_j.

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

Equation (14b) is solved using Newton's method to estimate the degrees of freedom ν_j.

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

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

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

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

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

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

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

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 κ_j can be approximated as shown in equation (17d), by solving the linear equation (17c) (similarly to (Bijral et al., 2007)).

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

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

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

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

- 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.

$$MSD = mean(mean(\mathbf{d_{min}}(A, B)), mean(\mathbf{d_{min}}(B, A))) \tag{18}$$

- 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.

$$\hat{f}_{kj} = \sum_{i=1}^{N_k} \frac{P_{kij} f_{ki}}{\sum\limits_{l} P_{klj}} \tag{19a}$$

$$\hat{\mathbf{n}}_{kj} = \sum_{i=1}^{N_k} \frac{P_{kij} \mathbf{n}_{ki}}{\sum\limits_{l} P_{klj}} \tag{19b}$$

References

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry — the methods. Neuroimage 11, 805–821.

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

Bayer, S., Ravikumar, N., Strumia, M., Tong, X., Gao, Y., Ostermeier, M., Fahrig, R., Maier, A., 2018. Intraoperative brain shift compensation using a hybrid mixture model, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 116–124.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Annals of Statistics , 1165–1188.

Benou, I., Veksler, R., Friedman, A., Raviv, T.R., 2018. Fiber-flux diffusion density for white matter tracts analysis: Application to mild anomalies localization in contact sports players, in: Computational Diffusion MRI, Springer. pp. 191–204.

Bijral, A.S., Breitenbach, M., Grudic, G.Z., 2007. Mixture of watson distributions: A generative model for hyperspherical embeddings., in: AISTATS, pp. 35–42.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Cercignani, M., Inglese, M., Pagani, E., Comi, G., Filippi, M., 2001. Mean diffusivity and fractional anisotropy histograms of patients with multiple sclerosis. American Journal of Neuroradiology 22, 952–958.

Chang, L.C., Walker, L., Pierpaoli, C., 2012. Informed restore: a method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. Magnetic Resonance in Medicine 68, 1654–1663.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society. Series B (methodological) , 1–38.

Durrleman, S., Fillard, P., Pennec, X., Trouvé, A., Ayache, N., 2011. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. NeuroImage 55, 1073–1090.

Durrleman, S., Pennec, X., Trouvé, A., Ayache, N., 2009. Statistical models of sets of curves and surfaces based on currents. Medical Image Analysis 13, 793–808.

Garyfallidis, E., Ocegueda, O., Wassermann, D., Descoteaux, M., 2015. Robust and efficient linear registration of white-matter fascicles in the space of streamlines. NeuroImage 117, 124–140.

Gooya, A., Davatzikos, C., Frangi, A.F., 2015. A bayesian approach to sparse model selection in statistical shape models. SIAM Journal on Imaging Sciences 8, 858–887.

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C., Mori, S., 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 39, 336–347.

Hufnagel, H., Pennec, X., Ehrhardt, J., Ayache, N., Handels, H., 2008. Generation of a statistical shape model with probabilistic point correspondences and the expectation maximization-iterative closest point algorithm. International journal of computer assisted radiology and surgery 2, 265–273.

Huynh, D.Q., 2009. Metrics for 3d rotations: Comparison and analysis. Journal of Mathematical Imaging and Vision 35, 155–164.

Jian, B., Vemuri, B.C., 2005. A robust algorithm for point set registration using mixture of gaussians, in: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, IEEE. pp. 1246–1251.

Jupp, P., Mardia, K., 1989. A unified view of the theory of directional statistics, 1975-1988. International Statistical Review/Revue Internationale de Statistique , 261–294.

Liu, Y., Spulber, G., Lehtimäki, K.K., Könönen, M., Hallikainen, I., Gröhn, H., Kivipelto, M., Hallikainen, M., Vanninen, R., Soininen, H., 2011. Diffusion tensor imaging and tract-based spatial statistics in alzheimer's disease and mild cognitive impairment. Neurobiology Of Aging 32, 1558–1571.

Maddah, M., Grimson, W.E.L., Warfield, S.K., Wells, W.M., 2008. A unified framework for clustering and quantitative analysis of white matter fiber tracts. Medical Image Analysis 12, 191–202.

Mayer, A., Greenspan, H., 2008. Bundles of interest based registration of white matter tractographies, in: Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, IEEE. pp. 919–922.

Mayer, A., Zimmerman-Moreno, G., Shadmi, R., Batikoff, A., Greenspan, H., 2011. A supervised framework for the registration and segmentation of white matter fiber tracts. IEEE transactions on medical imaging 30, 131–145.

McGraw, T., Vemuri, B.C., Yezierski, B., Mareci, T., 2006. Von mises-fisher mixture model of the diffusion odf, in: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006., IEEE. pp. 65–68.

Medina, D., Urresta, F., Gabrieli, J.D., Moseley, M., Fleischman, D., Bennett, D.A., Leurgans, S., Turner, D.A., Stebbins, G.T., et al., 2006. White matter changes in mild cognitive impairment and ad: a diffusion tensor imaging study. Neurobiology Of Aging 27, 663–672.

Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. Computer Methods And Programs In Biomedicine 98, 278–284.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A.V., Mahmood, A., Woods, R., et al., 2008. Stereotaxic white matter atlas based on diffusion tensor imaging in an icbm template. Neuroimage 40, 570–582.

Myronenko, A., Song, X., 2010. Point set registration: Coherent point drift. IEEE Transactions On Pattern Analysis And Machine Intelligence 32, 2262–2275.

ODonnell, L.J., Suter, Y., Rigolo, L., Kahali, P., Zhang, F., Norton, I., Albi, A., Olubiyi, O., Meola, A., Essayed, W.I., et al., 2017. Automated white matter fiber tract identification in patients with brain tumors. NeuroImage: Clinical 13, 138–153.

O'Donnell, L.J., Wells, W.M., Golby, A.J., Westin, C.F., 2012. Unbiased groupwise registration of white matter tractography, in: International Conference on Medical Image Computing and Computer-assisted Intervention, Springer. pp. 123–130.

O'Donnell, L.J., Westin, C.F., 2007. Automatic tractography segmentation using a high-dimensional white matter atlas. IEEE Transactions On Medical Imaging 26, 1562–1575.

Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N., 2001. Reconstructing a 3d structure from serial histological sections. Image and Vision Computing 19, 25–31.

Pierpaoli, C., Basser, P.J., 1996. Toward a quantitative assessment of diffusion anisotropy. Magnetic Resonance in Medicine 36, 893–906.

Pierpaoli, C., Walker, L., Irfanoglu, M., Barnett, A., Basser, P., Chang, L., Koay, C., Pajevic, S., Rohde, G., Sarlls, J., et al., 2010. Tortoise: an integrated software package for processing of diffusion mri data, in: ISMRM 18th annual meeting, p. 1597.

Ravikumar, N., Gooya, A., Çimen, S., Frangi, A.F., Taylor, Z.A., 2016. A multi-resolution t-mixture model approach to robust group-wise alignment of shapes, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 142–149.

Ravikumar, N., Gooya, A., Çimen, S., Frangi, A.F., Taylor, Z.A., 2018. Group-wise similarity registration of point sets using students t-mixture model for statistical shape models. Medical Image Analysis 44, 156–176.

Ravikumar, N., Gooya, A., Frangi, A.F., Taylor, Z.A., 2017. Generalised coherent point drift for group-wise registration of multi-dimensional point sets, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 309–316.

Salat, D., Tuch, D., Greve, D., Van Der Kouwe, A., Hevelone, N., Zaleta, A., Rosen, B., Fischl, B., Corkin, S., Rosas, H.D., et al., 2005. Age-related alterations in white matter microstructure measured by diffusion tensor imaging. Neurobiology Of Aging 26, 1215–1227.

Schwartzman, A., Dougherty, R.F., Taylor, J.E., 2005. Cross-subject comparison of principal diffusion direction maps. Magnetic Resonance in Medicine 53, 1423–1431.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., et al., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505.

Sra, S., Karp, D., 2013. The multivariate watson distribution: Maximum-likelihood estimation and other aspects. Journal of Multivariate Analysis 114, 256–269.

Zhang, Y., Schuff, N., Jahng, G.H., Bayne, W., Mori, S., Schad, L., Mueller, S., Du, A.T., Kramer, J., Yaffe, K., et al., 2007. Diffusion tensor imaging of cingulum fibers in mild cognitive impairment and alzheimer disease. Neurology 68, 13–19.

Zhou, Z., Zheng, J., Dai, Y., Zhou, Z., Chen, S., 2014. Robust non-rigid point set registration using student's-t mixture model. PloS one 9, e91381.

Zvitia, O., Mayer, A., Shadmi, R., Miron, S., Greenspan, H.K., 2010. Co-registration of white matter tractographies by adaptive-mean-shift and gaussian mixture modeling. IEEE Transactions On Medical Imaging 29, 132–145.