Reducing variability in along-tract analysis with diffusion profile realignment*

Samuel St-Jean[a,*], Maxime Chamberland[b], Max A. Viergever[a], Alexander Leemans[a]

[a] *Image Sciences Institute, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands*
[b] *Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Maindy Road, Cardiff, CF24 4HQ, United Kingdom*

Abstract

Diffusion weighted magnetic resonance imaging (dMRI) provides a non invasive virtual reconstruction of the brain's white matter structures through tractography. Analyzing dMRI measures along the trajectory of white matter bundles can provide a more specific investigation than considering a region of interest or tract-averaged measurements. However, performing group analyses with this along-tract strategy requires correspondence between points of tract pathways across subjects. This is usually achieved by creating a new common space where the representative streamlines from every subject are resampled to the same number of points. If the underlying anatomy of some subjects was altered due to, e.g., disease or developmental changes, such information might be lost by resampling to a fixed number of points. In this work, we propose to address the issue of possible misalignment, which might be present even after resampling, by realigning the representative streamline of each subject in this 1D space with a new method, coined diffusion profile realignment (DPR). Experiments on synthetic datasets show that DPR reduces the coefficient of variation for the mean diffusivity, fractional anisotropy and apparent fiber density when compared to the unaligned case. Using 100 in vivo datasets from the human connectome project, we simulated changes in mean diffusivity, fractional anisotropy and apparent fiber density. Pairwise Student's t-tests between these altered subjects and the original subjects indicate that regional changes are identified after realignment with the DPR algorithm, while preserving differences previously detected in the unaligned case. This new correction strategy contributes to revealing effects of interest which might be hidden by misalignment and has the potential to improve the specificity in longitudinal population studies beyond the traditional region of interest based analysis and along-tract analysis workflows.

Keywords: Diffusion profile realignment, Along-tract analysis, Tractometry, Tractography, Diffusion MRI, White matter

1. Introduction

Diffusion weighted magnetic resonance imaging (dMRI) is a noninvasive technique which can be used to study micro-structure in living tissues based on the displacement of water molecules. Since neurological diseases (e.g., multiple sclerosis (MS) (Cercignani and Gandini Wheeler-Kingshott, 2018), amyotrophic lateral sclerosis (ALS) (Haakma et al., 2017)) involve many processes that affect the density and properties of the underlying tissue, the corresponding changes are reflected on scalar values extracted from dMRI (Bodini and Ciccarelli, 2009). However, it remains challenging to accurately pinpoint the underlying cause as many of these changes (e.g., axonal damage, demyelination) may be reflected similarly by changes in measurements from dMRI (Beaulieu, 2002). Such changes could even be due to acquisition artifacts or from the use of a different processing method during data analysis (Jones and Cercignani, 2010), making dMRI sensitive,

*The same document is also available over at https://arxiv.org/abs/1902.01399

*Corresponding author

Email addresses: samuel@isi.uu.nl (Samuel St-Jean), chamberlandm@cardiff.ac.uk (Maxime Chamberland), max@isi.uu.nl (Max A. Viergever), A.Leemans@umcutrecht.nl (Alexander Leemans)

but not necessarily specific, to the various mechanisms involved in those changes (O'Donnell and Pasternak, 2015). Accurate characterization of the underlying processes affecting scalar metrics computed from dMRI still remains an open question.

A successful application of dMRI is to reconstruct the structure of the underlying tissues, a process known as tractography (see, e.g., Jeurissen et al. (2017); Mori and Van Zijl (2002) for a review). Tractography enables a virtual reconstruction of the white matter bundles and pathways of the brain, which is central to preoperative neurosurgical planning (Nimsky et al., 2016) and at the heart of connectomics (Hagmann et al., 2007; Sporns et al., 2005). Over the last years, various analysis strategies have arisen to study scalar values computed from dMRI models. Two popular schools of techniques consist in using anatomical regions of interests (ROIs), either by manual or automatic delineation (Froeling et al., 2016; Smith et al., 2006), or using spatial information additionally brought by tractography to analyze scalar metrics along reconstructed bundles (Colby et al., 2012; Corouge et al., 2006; Cousineau et al., 2017; Jones et al., 2005; Yeatman et al., 2012). Both approaches involve various user defined settings and have their respective criticisms and drawbacks; ROI based analysis requires accurate groupwise registration (Bach et al., 2014), whereas tractography based analysis needs to deal with false positives streamlines which can also look anatomically plausible (Maier-Hein et al., 2017). One key point shared between these methods is that they both require some form of correspondence between the studied structure of interest for each subjects, either by spatial registration to align the delineated ROIs (Froeling et al., 2016; Smith et al., 2006) or along the streamlines by resampling to a common number of points (Colby et al., 2012; Yeatman et al., 2012). Tractography based approaches can analyze the voxels traversed by a specific white matter bundle in a data driven way and reveal subtle local changes inside a bundle, while ROI based analysis discards the 3D spatial information but reveal widespread changes in the bundle (O'Hanlon et al., 2015). For tractography based analysis, metrics are either averaged by using all points forming a common bundle (Wakana et al., 2007) or collapsed as a representative pathway of the bundle (Colby et al., 2012; Cousineau et al., 2017; Yeatman et al., 2012) to study changes in scalar values along its length. Once this per subject representative streamline has been defined, it is used to index scalar values along the length of this pathway (O'Hanlon et al., 2015; Szczepankiewicz et al., 2013). Recent applications include studying changes in diffusion metrics due to Alzheimer's disease (AD) (Jin et al., 2017), which helped to uncover changes in mean diffusivity (MD) along the fornix for example. Studies in ALS patients also identified a diminution in fractional anisotropy (FA) along the corticospinal tract depending on the origin of the disease (Blain et al., 2011). Information from other MRI weighting such as myelin water fraction maps derived from T2 relaxometry have also been included to study changes due to MS (Dayan et al., 2016). As each subject respective morphology is different (i.e., reconstructed bundles from different subjects vary in shape and size) just as in ROIs based analysis, one needs to ensure correspondence between each segment of the studied bundle for all subjects. This correspondence is usually achieved by creating a new common space where all of the subjects representative streamlines are resampled to a common number of points. As noted by Colby et al. (2012), resampling to the same number of points makes the implicit assumption that the end points (and every point in between) are in correspondence across each subject. Yeatman et al. (2012) also mention that "it is important to recognize that the distal portions of the tract may not be in register across subjects", even though the resampling step creates a new 1D space for point-by-point comparison. O'Donnell et al. (2009) previously noticed the potential issue introduced by misalignment between subjects mentioning that "improved cross-subject alignment is of interest [...] as the high-frequency variations seen in individual subjects [...] are smoothed in the group average". While many methods for registering dMRI volumes or streamlines were developed (see, e.g., O'Donnell et al. (2017) for a review), they do not directly address the issue of possible residual misalignment between the end points *after* extracting the representative streamlines of each subject. To ensure an adequate comparison between subjects, one must make sure that each streamline corresponds to the same underlying anatomical location.

In the present work, which extends our preliminary work presented at the ISMRM (St-Jean et al., 2016), we focus on the issue of possible misalignment between the final representative streamlines before conducting statistical analysis. To prevent this issue, we propose to realign the representative streamline of each subject while ensuring that the distance between each point is preserved, by resampling to a larger number of points than initially present. This strategy preserves the original 1D resolution of each subject,

allowing a groupwise realignment based on maximizing the overall similarity by using the Fourier transform. After this realignment, points from individual streamlines which are identified as outliers can be discarded as they would not overlap with the rest of the subjects. The representative, and now realigned, streamlines can be resampled to a lower number of points such as approximately one unit voxel size to facilitate group comparison and statistics (Colby et al., 2012). Fig. 1 shows an example of a typical workflow to analyze dMRI datasets and how the proposed diffusion profile realignment (DPR) methodology can be used in preexisting pipelines.

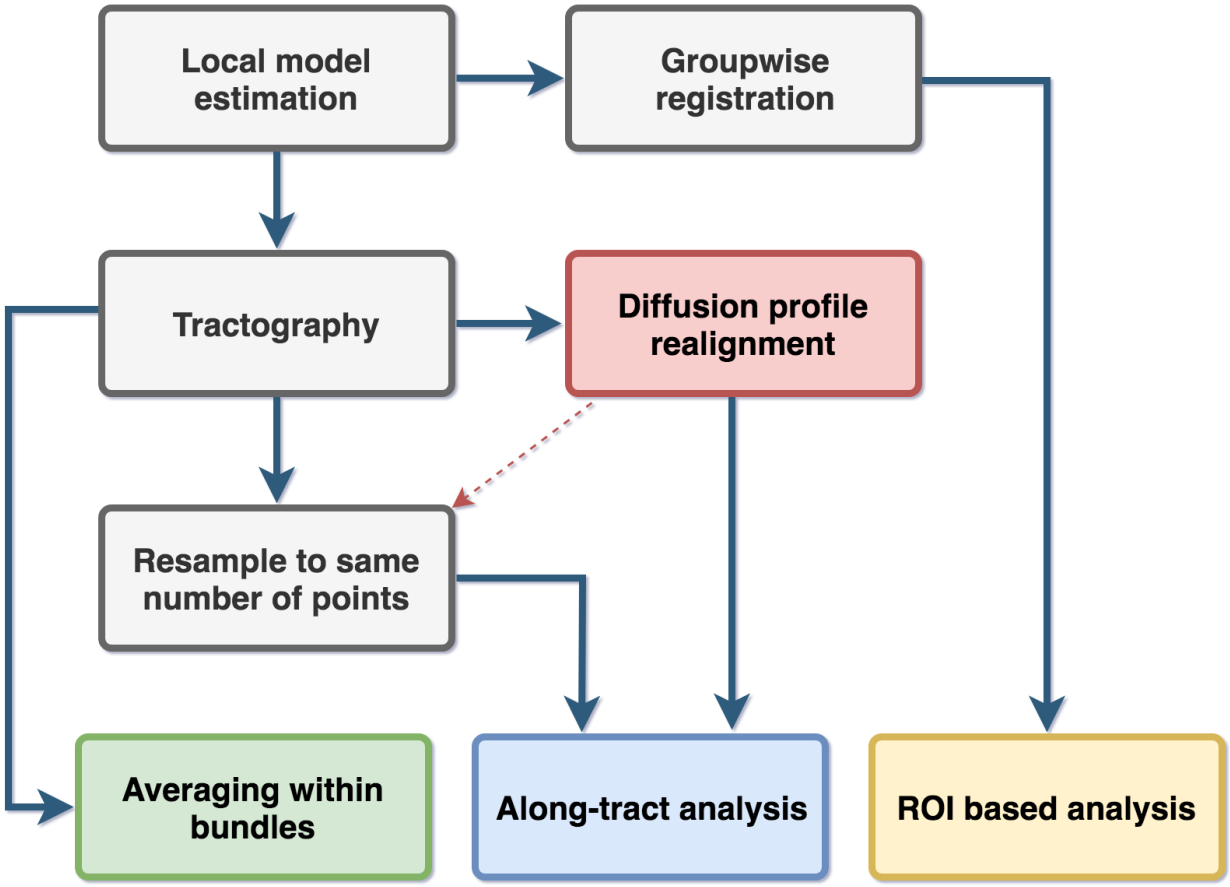

Figure 1: Flowchart of current approaches and the proposed methodology. The diffusion profile realignment inserts itself in existing along-tract analysis workflows (red box) by combining a different resampling strategy with a realignment step. It is also possible to resample each streamlines to a smaller number of points (red arrow) if desired.

2. Theory

Each subject's representative streamline is a 3D object, but the scalar metrics extracted along the tract can be viewed as a discrete 1D signal that may be non-stationary. In this work, we consider the 1D scalar metric profile to be a discrete signal equally sampled at each step of the tractography which has a value of 0 outside the region delineated by the bundle it represents. We now present a realignment technique for 1D signals based on maximizing the cross-correlation function (CCF).

We can define the CCF using the fast Fourier transform (FFT) (Cooley and Tukey, 1965) as

$$\mathrm{CCF}(x, y) = \mathcal{F}^{-1}(\mathcal{F}(x) \odot \mathcal{F}(y)^*), \tag{1}$$

where $\mathcal{F}(\mathrm{x})$ and $\mathcal{F}^{-1}(\mathrm{x})$ is the Fourier transform of x and its inverse, * is the complex conjugation and $\odot$ the pointwise Hadamard product. The required shift to realign the vectors is given at the maximum coordinate of the CCF. The CCF measures the similarity between two vectors x and y assuming that the data is 1) stationary, 2) equally spaced between all points and 3) normally distributed (Denman, 1975; Platt and Denman, 1975). Stationarity can be achieved by fitting and subtracting a low degree polynomial from each vector before computing the cross-correlation, see, e.g., Box et al. (2008); Stoica and Moses (2005) and references therein for more details. Equal spacing between each points can be obtained by resampling the data. The normality assumption seems less of an issue for large samples in practice (Platt and Denman, 1975). If the two vectors x and y have a different amplitude, the cross-correlation can be normalized by subtracting the mean and dividing by the standard deviation of each vector beforehand (Lewis, 1995). The shift computed at the maximum of the CCF is an integer displacement which can be refined by finding the maximum of the parabola around this point. Fig. 2 shows an example of the cross-correlation for both the stationary and non stationary case on two vectors. The first vector was randomly sampled from the standard normal distribution $\mathcal{N}(0, 1)$. The second vector was generated from the first vector by changing the offset and amplitude and then zero padding it at both end to create an artificial displacement.

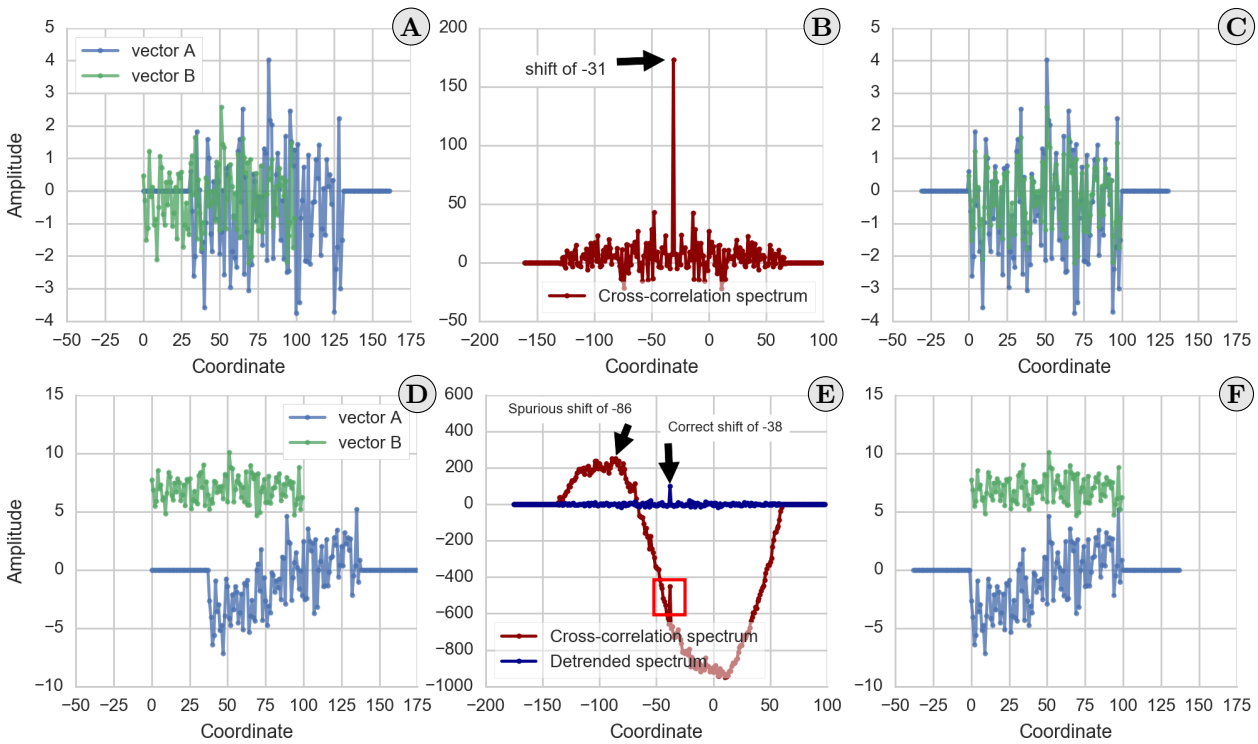

Figure 2: A synthetic example of the CCF between two randomly generated vectors. The top graphs showcase how the CCF spectrum can be used to find the displacement required to realign two different vectors by finding its maximum. **A)** Two vectors which are displaced with respect to each other, where vector B has a different amplitude from vector A. **B)** The cross-correlation spectrum, where the peak indicates the required shift to maximize the overlap between both vectors. **C)** The vectors after realignment, which is the exact displacement that had been applied. On the bottom graphs, removing the linear trend and normalizing the vectors satisfies the assumption of stationarity required by Eq. (1) and allows recovering the correct shift. **D)** Two unaligned vectors of different amplitude where vector B is also non stationary. **E)** The cross-correlation spectrum with detrending and normalization (in blue) and without these steps (in red). The detrended version recovers the correct shift, while the original CCF exhibits a variation in amplitude which hides the correct peak as a local (red box), but not global, maxima due to non stationarity. **F)** The vectors after realignment with the shift as computed by the detrended CCF. Both vectors are now realigned after shifting vector B with the shift computed in **E)**.

3. Materials and methods

To evaluate the proposed realignment procedure, we 1) generated synthetic datasets comprised of crossing bundles and 2) compared realignment on in vivo datasets with an altered version of their diffusion metrics. We now detail the various steps needed to perform an along-tract analysis and how the proposed realignment algorithm can be applied before performing a statistical analysis between subjects.

3.1. Resampling strategies for comparison between subjects

Various resampling strategies have been discussed in previous along-tract frameworks, with a common idea advocating resampling all representative streamlines to the same number of points. In Cousineau et al. (2017), the authors used a fixed number of points by resampling all of the studied bundles to 20 points while Yeatman et al. (2012) instead used 100 points. Colby et al. (2012) opted for resampling each bundle based on their average group length, ensuring that approximately one point per voxel was present. In this representation, each point of the streamlines is considered to correspond to the same anatomical location across subjects and is therefore blind to the intrinsic variance in shape or length between subjects. As each representative streamline most likely had a different length initially, the distance in millimeters between each sampled coordinate will be different for each subject. If the underlying anatomy of some subjects was altered due to, e.g., disease or developmental changes, such information might be lost by resampling to a fixed number of points as a first step. This can be prevented by ensuring that the new sampling resolution is at least equal or larger than the initial resolution used during tractography.

As a bundle is comprised of many individual streamlines, they are usually collapsed to a single representative pathway to facilitate subsequent analysis. This representative streamline is therefore an aggregation of many streamlines of various length and can be obtained either by averaging (Colby et al., 2012; Yeatman et al., 2012) or by finding representative clusters (Cousineau et al., 2017). Other assignment strategies towards a single representative pathway have been discussed in (Corouge et al., 2006; O'Donnell et al., 2009). To ensure correspondence during this aggregation step, individual streamlines are usually resampled to a common number of points for all subjects. While this resampling is needed to obtain the representative streamline, it may also reduces the sampling resolution from the original streamlines given by the step size used for tractography if not enough points are kept. The representative streamline of each subject may also have a different orientation altogether and therefore might need to be flipped, ensuring that they share a common coordinate system (Colby et al., 2012).

In the present work, we instead advocate a novel two-step resampling strategy which builds upon the classical resampling strategy. After extracting the representative streamlines $(S_1, \ldots, S_n)$ for $i = 1, \ldots, n$ of each subject, each representative streamline S_i is defined by its number of points N_i and the distance between its points δ_i. All streamlines are first resampled to $M_i = N_i \times \delta_i / \delta_{\min}$ points, ensuring an equal distance between each point $\delta_{\min} = \min(\delta_1, \ldots, \delta_n)$. In the end, the streamlines still have a different number of points $M_i \geq \min(N_1, \ldots, N_n)$ and points at the same coordinates across subjects do not implicitly assume to represent the same anatomical location. However, the distance δ_i between each point M_i is now constant across subjects. While this idea may seem counterintuitive, the motivation behind this choice is due to Eq. (1), which relies on the FFT, and as such, needs equally sampled vectors and benefits from a high sampling resolution.

After the displacement has been applied , one can use the classical resampling strategies presented by other authors, therefore making our approach fully compatible with already existing analysis techniques. We opted to use the methodology of Colby et al. (2012) since more than one point per unit voxel size would not carry additional information from the original data. This also alleviates further complications arising from multiple comparisons (Benjamini and Hochberg, 1995) for the subsequent statistical analysis one seeks to apply afterwards. Fig. 3 illustrates schematically the classical resampling versus our novel two-step resampling strategy.

3.2. Proposed algorithm for diffusion profile realignment

DPR works in three steps once the 1D profiles have been resampled to an equal spacing as presented in Section 3.1. We also ensure stationarity of the data by fitting and subtracting a polynomial of degree one

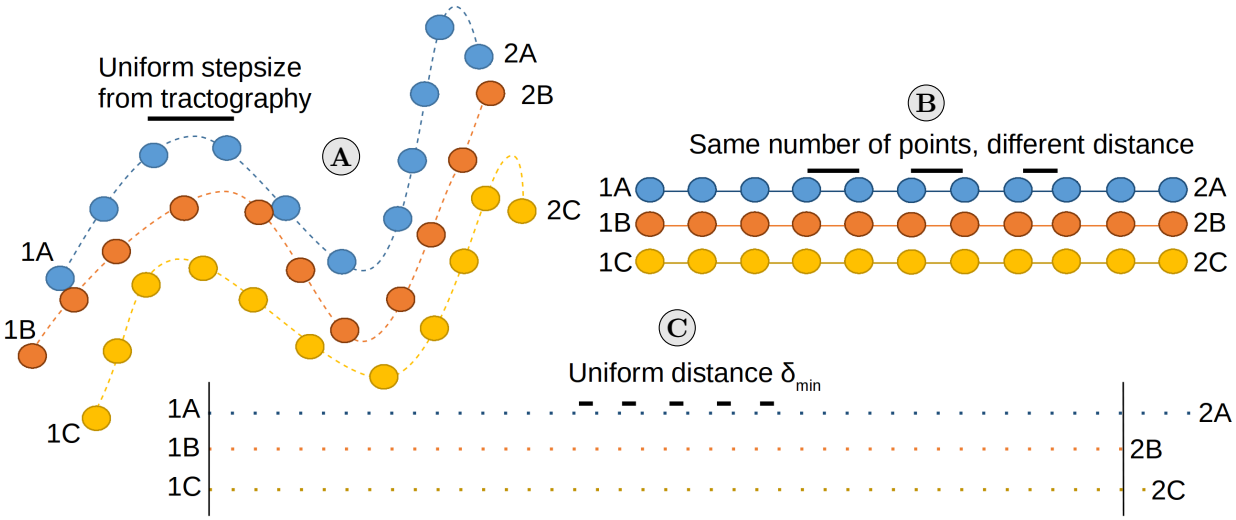

Figure 3: An example of the classical and proposed resampling strategies on three representative streamlines. In **A)**, three representative streamlines which have different shapes and lengths with their start (1A, 1B and 1C) and end points (2A, 2B and 2C) at different spatial locations. In **B)**, the classical strategy of resampling to the same number of points (circles) introduces a common space to easily compare them. However, the end points of the underlying anatomies are artificially aligned when compared to their original representation and each point is at a different distance (black lines). In **C)**, the proposed resampling strategy ensures that the distance $\delta_{\min}$ (black lines) between every point is constant. Even though each streamline length is different as indicated by the location of the end points, they can now be realigned to identify the common anatomical positions between all subjects.

(i.e., a straight line) to each subject. It is important to mention here that this step is only to satisfy the stationarity assumption of Eq. (1) and does not modify the extracted diffusion profiles afterwards.

Firstly, a matrix of displacement is computed between every pairs of subjects and subsequently refined with parabola fitting as previously defined in Section 2. A maximum possible displacement in mm is then chosen. From the displacement matrix, the subject realigning the largest number of streamlines inside this maximum displacement is chosen automatically as the template subject. As Eq. (1) is symmetric, realigning subject A to B or subject B to A will have the same outcome in practice.

Secondly, all outliers with a displacement larger than the chosen threshold from the first step are realigned with the help of a new per-streamline template. For each outlier, a new template is selected amongst the remaining non-outlier subjects which minimizes the total displacement between the original template from the first step and the current outlier. If the new minimum displacement is inside the chosen threshold, the subject which was previously an outlier is now registered through this new template. If no new template providing realignment inside the threshold can be found, then this subject is declared as an outlier and is not realigned at all. Fig. 4 shows the spectra of a normal subject and of an outlier for spectra computed with Eq. (1) from the HCP datasets. Even if the optimum displacement lies outside the chosen threshold, the outlier can still be realigned by finding a new template subject.

Finally, after realigning all the admissible streamlines to the template, there will be a different number of overlapping subjects for each coordinate. Just as ROIs were previously used to truncate the bundles' end points (recall Fig. 6), the resulting aligned streamlines should be truncated once again to reduce their uncertainty since not all coordinates have the same number of overlapping streamlines anymore. A pseudocode version of the proposed algorithm is outlined in Appendix A. Our reference implementation is freely available as a standalone[1] ([Code] St-Jean, 2019), and will also be included in *ExploreDTI* (Leemans et al., 2009). We also make available the synthetic datasets and metrics extracted along the representative streamlines of

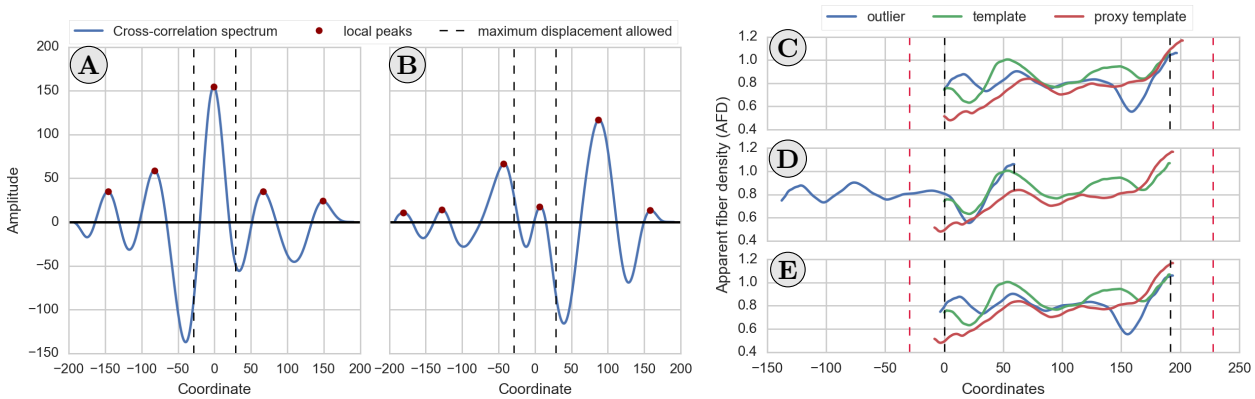

Figure 4: An example of a cross-correlation spectra (left) and finding a new template to realign outliers (right) using the HCP datasets. On the **left**, a threshold of 15% of the total streamline length is selected as the maximum allowed displacement (dashed vertical lines). **A)** A streamline with the global maximum of the CCF inside the chosen threshold. The maximum indicates the shift needed to realign it to the template. **B)** A streamline with a local maximum, but not the global maximum, of the CCF inside the chosen threshold. In this case, the two streamlines would not be realigned together as only small shifts should be needed for realignment. On the **right**, an example of realigning an outlier subject (in blue) to the original template (in green) via the closest matching new template (in red) using the AFD metric. The black dashed bars indicate the region where all three streamlines fully overlap and the red dashed bars shows the maximum allowed displacement of 15%. **C)** The three streamlines before realignment. **D)** Realigning the blue streamline with the template (in green) as given by the maximum of the CCF results in an outlier as in case **B)**. **E)** To circumvent the issue, a new template (in red) is found amongst the non-outlier subjects which minimizes the total displacement with the original template. The blue streamline is therefore not an outlier anymore as it now lies inside the displacement threshold as in case **A)**.

the HCP datasets which are used in this manuscript ([Dataset] St-Jean et al., 2018).

3.3. Datasets and acquisition parameters

Synthetic datasets. A synthetic phantom consisting of 3 straight bundles crossing in the center at 60 degrees with a voxel size of 2 mm was created with phantomas (Caruyer et al., 2014). Each bundle has some partial voluming present on the outer edge to mimic the white matter / gray matter interface. We simulated 64 diffusion weighted images (DWIs) using gradient directions uniformly distributed on a half sphere and one b = 0 s/mm^2 image with a signal-to-noise ratio (SNR) of 10, 20 and 30 with uniformly distributed Rician noise and a noiseless reference volume. Two distinct diffusion weightings of b = 1000 s/mm^2 and b = 3000 s/mm^2 were used, producing a total of 8 different synthetic datasets. The SNR was defined as SNR = S_0/σ, where S_0 is the non-diffusion weighted signal and σ is the Gaussian noise standard deviation.

HCP datasets. 100 subjects (50 males, 50 females) from the in vivo Human Connectome Project (HCP) database (Van Essen et al., 2012) aged between 26 and 30 years old were selected. All 18 b = 0 s/mm^2 volumes were kept along with the 90 volumes at b = 3000 s/mm^2 in order to maximize the angular resolution (Tournier et al., 2013). The acquisition parameters were a voxel size of 1.25 mm isotropic, a gradient strength of 100 mT/m, a multiband acceleration factor of 3 and TR / TE = 5520 ms / 89.5 ms. We used the minimally preprocessed datasets which are already corrected for subject motion, EPI distortions and eddy currents induced distortions (Van Essen et al., 2012).

3.4. Local model reconstruction and fiber tractography

We used the constrained spherical deconvolution (CSD) algorithm (Tournier et al., 2007) with a recursive calibration of the response function (Tax et al., 2014) and spherical harmonics of order 8 to estimate the fiber

[1] https://github.com/samuelstjean/dpr

orientation distribution functions (fODFs). We also computed the diffusion tensors using the REKINDLE approach (Tax et al., 2015) to exclude potential outliers from the data. We subsequently computed the apparent fiber density (AFD) maps (Dell'Acqua et al., 2013; Raffelt et al., 2012) from the fODFs and the FA and MD maps from the diffusion tensors (Basser and Pierpaoli, 1996) in all experiments. Whole-brain deterministic tractography was performed using the fODFs with *ExploreDTI* (Leemans et al., 2009) with a step size of 0.5 mm, a fODFs threshold of 0.1 and an FA threshold of 0.2 for all datasets. The angle threshold, seeding grid resolution and streamlines length threshold used during tractography were different for the synthetic and HCP datasets as detailed below.

Tractography parameters for the synthetic datasets. Tractography was performed with an angle threshold of 30 degrees and a seeding grid resolution of 0.5 mm on each axis to ensure a dense coverage of each bundle. Only the streamlines with a length of at least 10 mm and up to 150 mm were kept to prevent the presence of spurious streamlines. Two ROIs were manually drawn on one bundle to select only straight streamlines belonging to this bundle as shown in Fig. 5. The streamlines were kept at their full extent, including some small variations near the end points due to partial voluming, which ensures that the intersection of the three bundles is approximately at the center. To mimic similar representative streamlines extracted from various subjects, 150 streamlines were randomly selected and cut randomly from 1% up to 10% of their total length at both end points. Two sets of representative streamlines were created using classical resampling to the same number of points and our novel two-step resampling strategy, which is detailed in Section 3.1. In the first case, all streamlines were resampled to 50 points, which is approximately one unit point size per voxel. As each synthetic representative streamline had a different length after truncation, resampling to the same number of points allows a direct comparison between each coordinate, even if they do not match the same "anatomical" location by design of the experiment. No resampling was needed to simulate our proposed resampling strategy as the distance between each point is already equal for this particular synthetic example.

Tractography parameters for the HCP datasets. Whole-brain tractography was performed with an angle threshold of 45 degrees and a seeding grid resolution of 2 mm on each axis. Only the streamlines with a length of at least 10 mm and up to 300 mm were kept to limit the presence of spurious streamlines. ROIs were manually drawn to segment the left and right arcuate fasciculus (AF) and the left and right corticospinal tract (CST) on an exemplar subject (Wakana et al., 2007) as shown in Fig. 5. This exemplar subject FA map was used as a template and subsequently non linearly registered to each other subject respective FA map using Elastix (Klein et al., 2010). The obtained transformation was then applied on each ROIs drawn on the exemplar subject defining the four bundles, therefore warping the original ROIs unto each subject's respective diffusion space as in Lebel et al. (2008). Only the segments between the ROIs were kept to only retain the straight sections and to remove spurious end points e.g., before the fanning in the CST. An alternative approach could be to extract the bundles automatically using a parcellation of the white matter obtained from each subject's T1-weighted MR image (Cousineau et al., 2017; Wassermann et al., 2016).

Extracting representative streamlines for the HCP datasets. To extract the representative streamline of each subject, all streamlines forming a given bundle were linearly resampled to the same number of points, chosen as the number of points of the top 5% longest streamlines to reduce the effect of possible outliers. This choice is robust to possible outliers which might be longer (or much smaller) than the rest of the streamlines due to spurious results from tractography while also keeping a high sampling resolution, a desirable property for Eq. (1). In the present work, the mean streamline per bundle was extracted and finally resampled in two different ways: 1) using a fixed number of points for all subjects and 2) ensuring an equal distance between each point. For the classical resampling strategy, we resampled all subjects to 70 points for the arcuate fasciculi and 105 points for the corticospinal tracts. The second resampling strategy ensured that the distance $\delta_{\min}$ (in mm) between each point is the same for all subjects. This also means that the representative streamlines of each subject do *not* have the same number of points and can not be compared directly at this stage when using this resampling strategy. Fig. 6 shows an example of selecting a representative segment between two ROIs as would be done for the uncinate fasciculus.

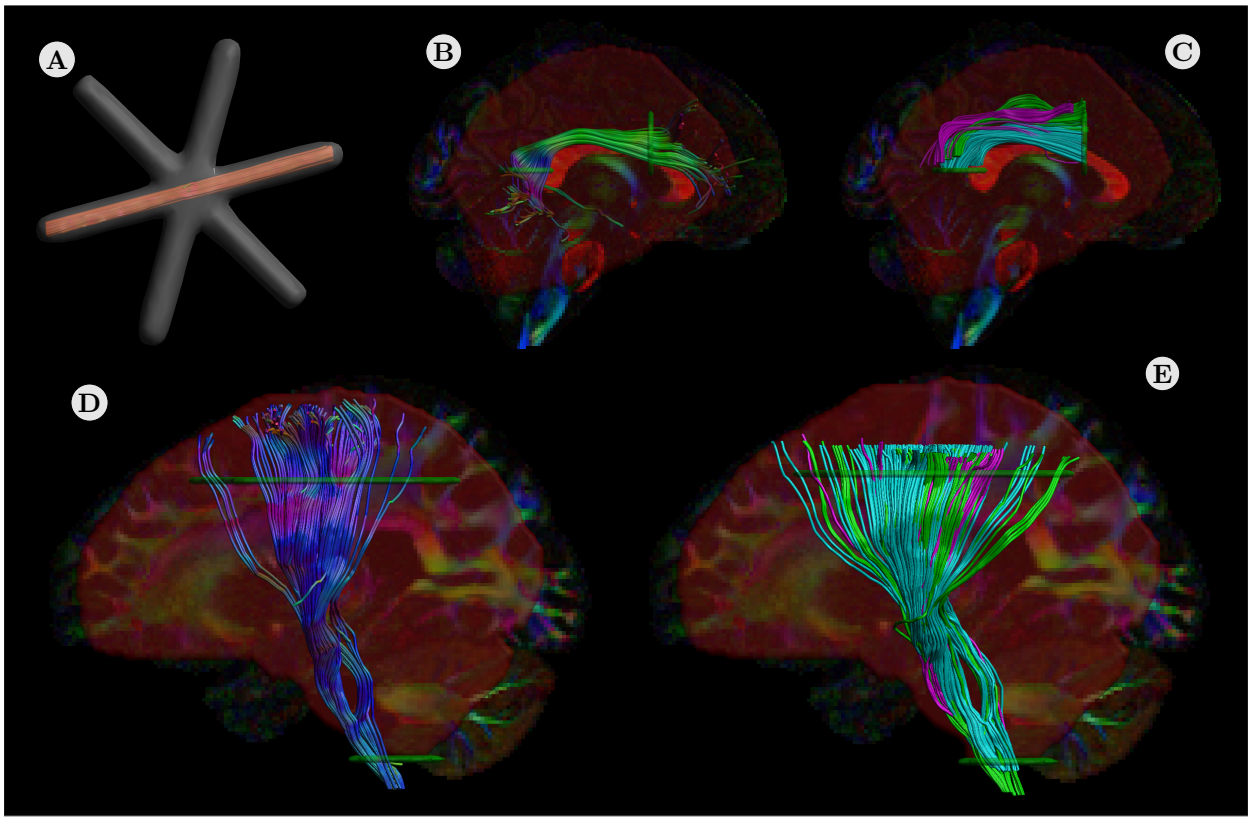

Figure 5: The synthetic bundles dataset and the locations of the ROIs used to segment some of the in vivo bundles on the exemplar subject with their automatically extracted counterpart for three subjects. In the top row, **A)** streamlines in a straight bundle of the synthetic datasets. Note that the streamlines are not truncated at the end points, but rather cover the full length of the red bundle so that they cross exactly at the center. The two inclusions (in green) and one exclusion (in red) ROIs segmenting **B)** the right AF on the exemplar subject and **C)** three automatically extracted right AF drawn in the exemplar subject native space (shown in green, cyan and magenta). On the bottom row, **D)** the left CST on the exemplar subject and **E)** three automatically extracted left CST bundles (shown in green, cyan and magenta) drawn in the exemplar subject native space. Note that each subject's bundle would correspond roughly to the same anatomical location in its own native space.

3.5. Extracting diffusion metrics for along-tract analysis

Once every representative streamline has been obtained, it can be used to collect diffusion derived metrics along the 3D pathway indexing a volume of interest. We collected the values of MD, FA and AFD for each subject along the streamline trajectory as in Colby et al. (2012). The resulting 1D segment is a vector of values varying along the length of the representative streamline. This single representative pathway can now be realigned in a pointwise fashion to ensure correspondence between subjects before moving on to statistical analysis.

3.6. Applying the diffusion profile realignment on representative streamlines

Realignment of uniformly resampled and variable length streamlines. To evaluate the reduction in variability brought by our proposed DPR algorithm, we estimated the coefficient of variation (CV) at each coordinate along the streamlines before and after realignment using both resampling strategies. The CV, defined as $CV = \sigma/\mu$ with σ the standard deviation and μ the mean of each metric, is a unitless standardized measure of dispersion where a lower CV indicates a lower standard deviation around the mean value.

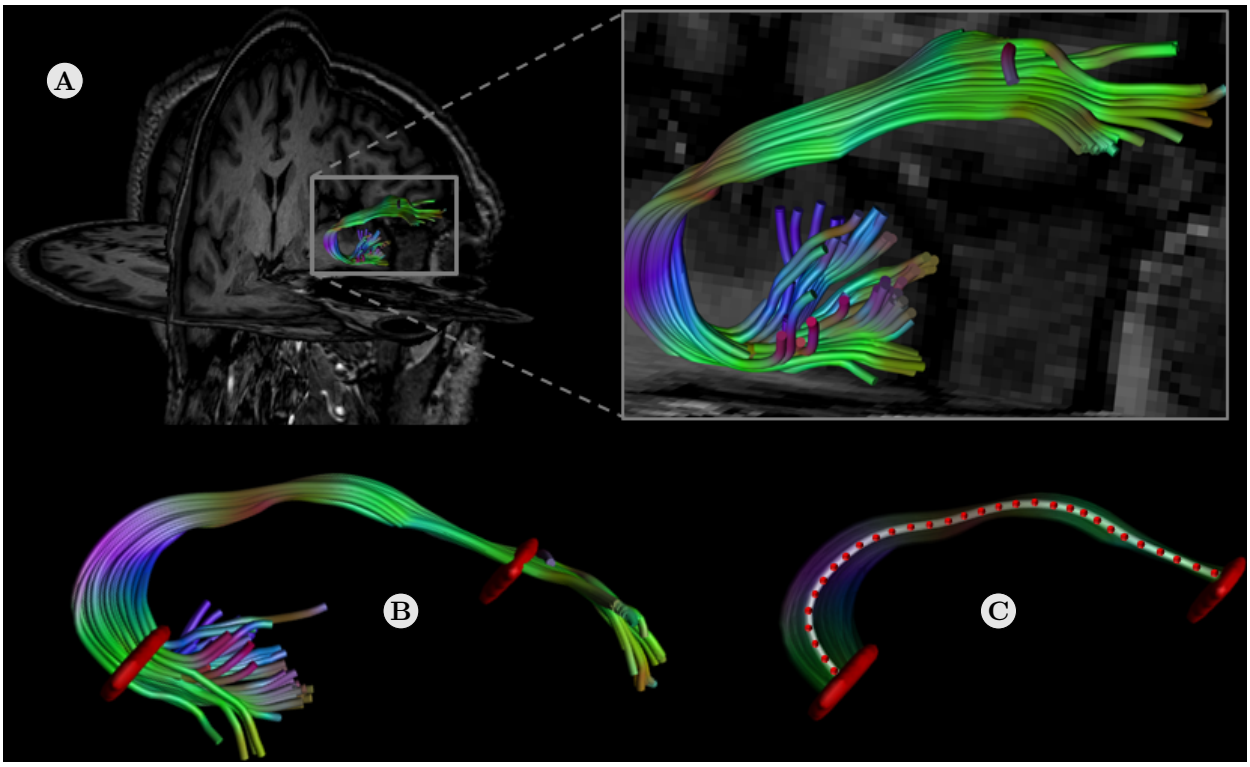

Figure 6: An example of along-tract analysis. **A)** The uncinate fasciculus is first segmented from a whole-brain tractography on an exemplar subject. **B)** The two ROIs (shown in red) that were defined to segment the uncinate fasciculus. Warping these ROIs to each subject provides an automatic dissection of the bundle. **C)** Only the portion of the mean streamline (shown in white) between the two ROIs is discretized (shown by the red dots), which allows mapping scalar metrics along the bundle itself.

For all experiments, we used a maximum displacement threshold of 15% to find the subject serving as a template during realignment. We computed the CV before and after realignment of the representative streamlines using both resampling strategies. To compare the variability due to truncation of the end points, only the segments where 1% (at least one streamline present), 50%, 75% and 100% (all streamlines are fully overlapping) of the realigned streamlines were kept for computing the CV. In the synthetic datasets experiments, we weighted the CV by the number of points at each coordinate to account for the different number of points of the unaligned bundle. For experiments with the HCP datasets, we instead did a final resampling to the same number of points (if appropriate) after the realignment as previously used for the classical resampling strategy in order to ensure a fair comparison between both approaches.

Simulating abnormal values of diffusion metrics in HCP subjects. An example application of the along-tract analysis framework could be to study neurological changes in a given population. These changes would presumably affect some specific white matter bundles and their underlying scalar values extracted from dMRI. Both the location and magnitude of these changes could reveal an effect of interest that might be hidden at first due to potential misalignment between subjects. To simulate a change in scalar metrics, 50 subjects were chosen randomly and had their representative streamlines profile modified while the other 50 subjects were left untouched. These 50 modified subjects are now classified as the "altered" subjects and the other untouched 50 subjects as the "controls" subjects in the subsequent experiments. For each altered subject, a location covering two times the affected length on both sides was chosen at random starting from the middle and the metrics were modified at this location. Two separate set of experiments were performed where the changes in metric was at first +10% and then −10% of its original value over 15% of the length.

An additional set of experiments simulating highly focused damage of ±25% and ±50% of the metrics over 5% or 1% of the bundle length was performed. For the three cases, the randomly chosen location was at a position from 20% to 80%, 40% to 60% and 48% to 52% of the bundle length. This process is repeated for each metric and each bundle, creating a different set of randomly modified subjects every time. The representative streamlines were finally realigned separately per group. As the control and altered subjects likely have different 1D profiles, realigning them separately makes it possible to select the best template for each group by itself. This strategy implicitly assumes that the neurological changes induce a similar increase or decrease in the diffusion metrics of each subject and that after realignment, each anatomical location is in correspondence between both groups. Correspondence between groups is also implied in classical along-tract analysis when resampling to the same number of points for comparison. Limiting the maximum displacement allowed also ensures that information carried by the diffusion metrics stays locally around the same position. The correspondence after separate realignment is assumed by resampling to the same number of points as the final step before analysis. In a clinical study setting, this could reflect neurological changes as induced by, e.g., a neurodegenerative disease or aging. The idea is to induce some changes in the extracted scalar values *only*, without modifying the underlying raw data or performing tractography and representative streamlines extraction once again. This choice of working in the extracted metric space only is to assess the changes on the metrics and realignment, in opposition to changes affecting the raw data itself. As the tractography process and extracted streamlines would most likely be slightly different due to the inherent challenges in reproducing tractography (Maier-Hein et al., 2017), the subsequent interpretation of the results could be confounded if tractography would be done anew.

Statistical tests between HCP subjects. We conducted a Student's t-test for related pairs of samples between the controls and altered HCP subjects with a correction for the false discovery rate (FDR) of $\alpha = 0.05$ (Benjamini and Hochberg, 1995) for one metric on each bundle. The t-test was realized on the datasets before and after realignment of the representative streamlines metrics. However, the FDR correction only ensures an upper bound on the occurrence of false effects and do not indicate their location nor how many are present.

4. Results

4.1. Simulations with the synthetic datasets

We now present numerical simulations involving the synthetic datasets presented in Section 3.3, comparing the two resampling strategies from Section 3.1 before applying the DPR algorithm. Fig. 7 shows the reduced CV for the realignment of the AFD metric on the SNR 20 dataset at b = 3000 s/mm^2 when compared to their non realigned counterpart. After realignment, the standard deviation at each coordinate is now generally lower, especially in the center portion where the three bundles are crossing. In the case of resampling to an equal distance $\delta_{\min}$, a few streamlines are overlapping at the end points, which might reduce statistical power for these regions during subsequent analyses. As previously mentioned in Section 3.2, portions where only a few streamlines are overlapping should be truncated accordingly to prevent these degenerate cases. Fig. 8 shows summary boxplots of the CV in addition to the mean CV across all coordinates for the synthetic datasets for the MD, FA and AFD. In all cases, realignment provides a lower CV than the non realigned synthetic streamlines.

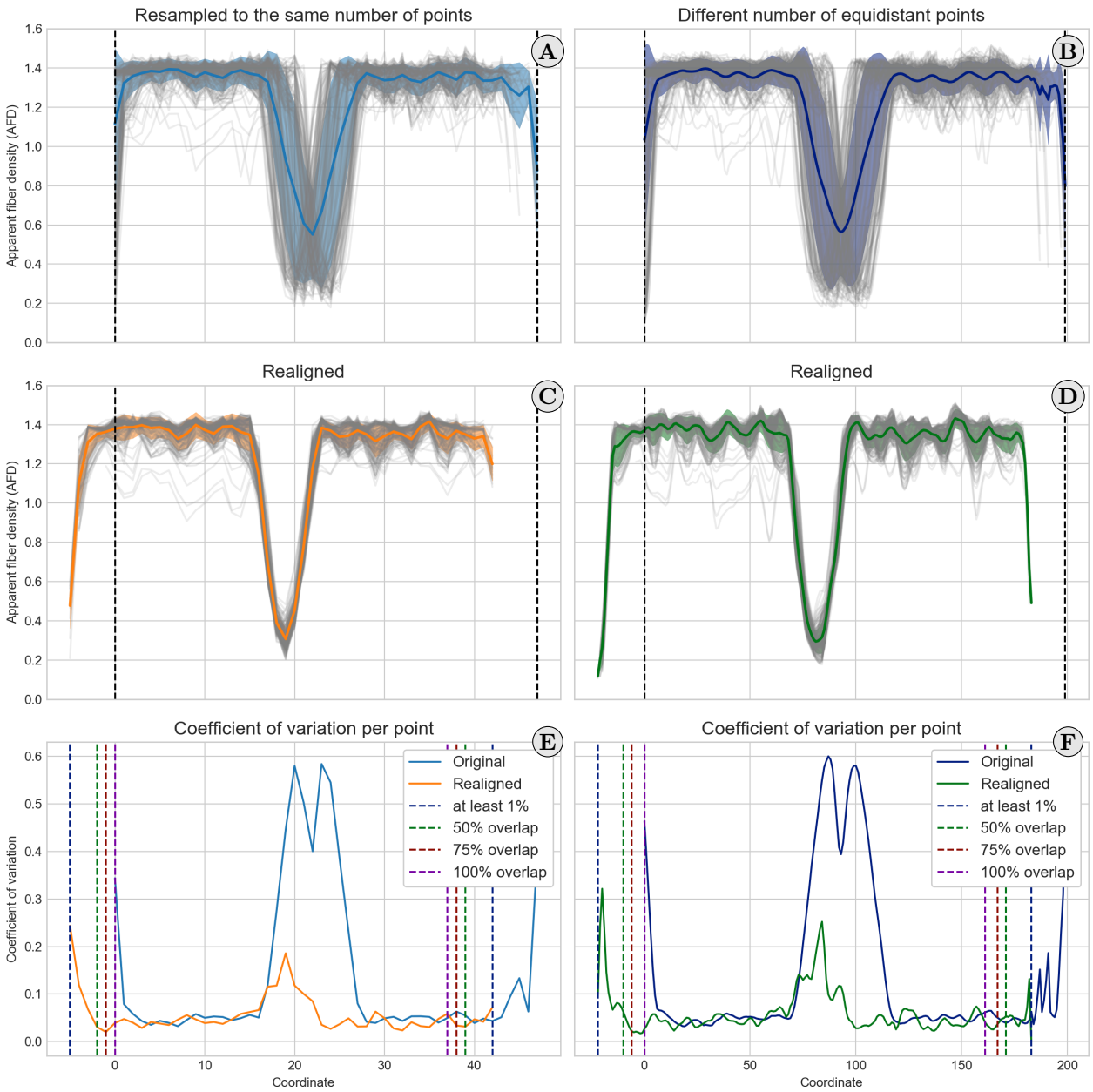

Figure 7: Realignment of representative streamlines resampled to 50 points (left column) and with an equal distance $\delta_{\min}$ (right column) for the AFD case at SNR 20 and b = 3000 s/mm^2. Each individual streamline is plotted in light gray, with the mean value in color and the standard deviation as the shaded area. The black vertical bars indicate the location of the original, non realigned streamlines. The colored vertical bars indicate the number of overlapping streamlines, ranging from at least 1 (all subjects, purple lines) to all of them (100%, red lines). Panels **A)** and **B)** show the streamlines before realignment. Note how individual streamlines are rather dispersed around the mean. Panels **C)** and **D)** show the streamlines after realignment, with the mean value being closer to all of the subjects and a smaller standard deviation than in panels **A)** and **B)**. However, due to the realignment, the end points have less subjects contributing to the mean value and should be truncated according to the number of overlapping subjects. Panels **E)** and **F)** show the coefficient of variation (CV, where lower is better) for each point, which is in general lower or equal than the non realigned version in both cases. Note how the largest reduction in CV is in the crossing region, where the standard deviation is approximately three times smaller in the realigned case than for the unaligned case.

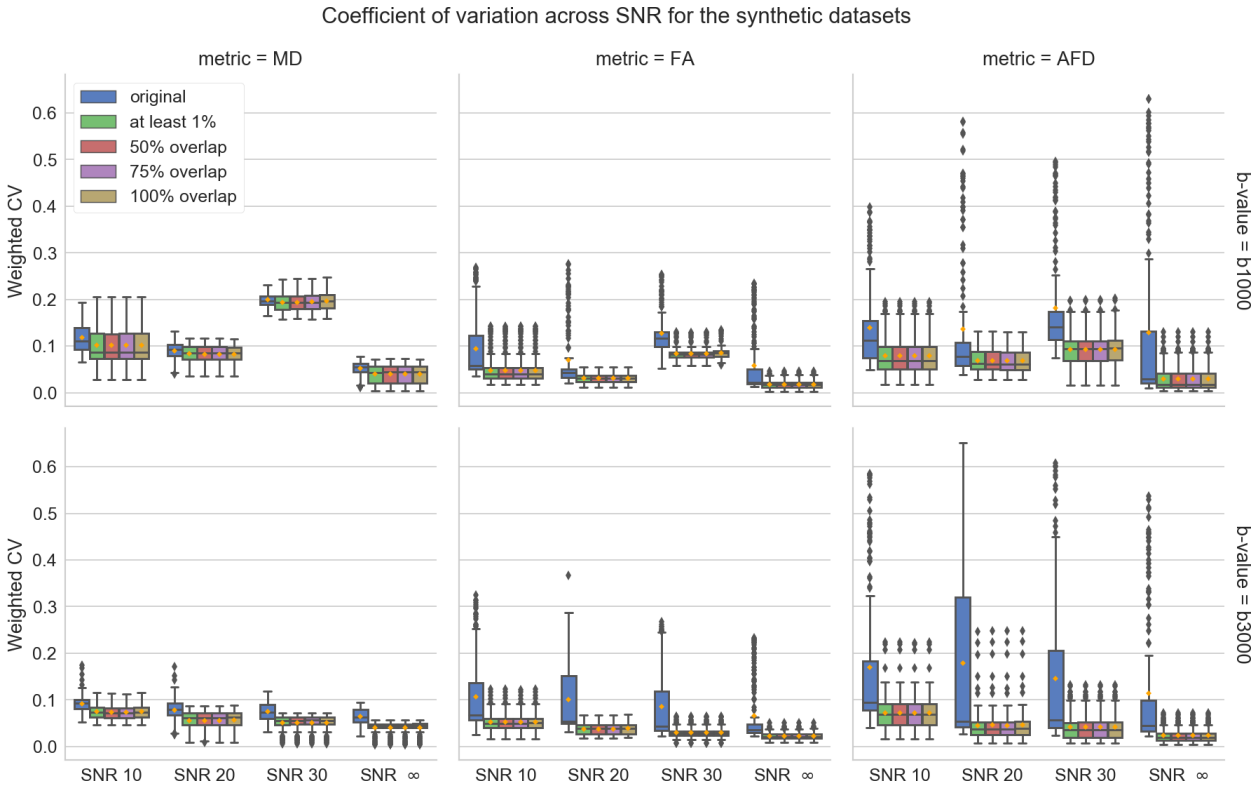

Figure 8: Boxplots of the CV for each point weighted by the number of overlapping subjects, for the MD (left), FA (center) and AFD (right) metrics and their respective mean value (in orange). The top row shows results for b = 1000 s/mm^2 on the synthetics datasets at SNR 10, 20, 30 and in the noiseless case while the bottom row shows results for b = 3000 s/mm^2. In all cases, the realigned metrics (for any truncation percentage) have a lower or equal CV on average than the non realigned metrics (in blue). The FA and AFD metrics have a CV in the realigned case which is on average approximately two times smaller than the non realigned case across all SNRs and both b-values. This gain is smaller for the MD, which might be due to the relative homogeneity of the MD values.

4.2. Realignment of the in vivo HCP datasets

Realignment of the arcuate fasciculi and corticospinal tracts. To quantify the improvements brought by the DPR algorithm for the in vivo datasets, we realigned the representative streamlines extracted from the 100 HCP datasets. Fig. 9 shows the final outcome with the two previously discussed pipelines for producing along-tract averaged profiles: resampling to the same number of points as is conventionally done and after realignment with the DPR algorithm. For the realigned case, we kept only the segments where at least 75% of the subjects are overlapping and finally resampled all subjects to the same number of points. This last resampling step could be considered optional and is used to allow an easier visual comparison between the unaligned and realigned group profiles. While the overall shape of each profile is similar between the unaligned and realigned version, the end points and location of salient features are slightly different due to the realignment and the truncation threshold we used. As the maximum displacement threshold dictates which subject is used as a template for the realignment, average group profiles using a maximum displacement threshold of 5, 10 and 20% are shown in the supplementary materials Appendix B.1. To assess the effect of truncation on variance near the end points, we computed the CV for each metric at various truncation thresholds and for the unaligned metrics. Fig. 10 shows the CV for the HCP datasets when the bundles are first resampled to the same number of points and after realignment (in brown). In all cases, the CV is approximately equal or lower after realignment with the DPR algorithm than when the representative streamlines are unaligned and resampled to the same number of points. We also show the CV in the unaligned case where all streamlines have an equal distance $\delta_{\min}$ between points and for four truncation thresholds after applying the DPR algorithm (no truncation, 50%, 75% and 100% of overlap). In this particular case, the resampled and realigned bundles (light brown) and the realigned bundles with no truncation (green) are mostly equivalent as they are resampled to the same final number of points after realignment for comparison purposes. The main tendency shows a lower mean CV after realignment when compared to the non realigned cases. The CV values are also generally lower with increasing truncation thresholds as the number of overlapping points per coordinates is also increasing, contributing to a lower standard deviation of each metric.

Robustness of the shapes of averaged profiles towards different metrics. When performing an along-tract analysis, tractography plays a key role as a spatial indexation method for *extracting* the 1D metric profiles along the streamline. Given a particular subject representative streamline, the various scalar metrics that can be extracted each have their own distinct 1D profile along the streamline. In order to assess the robustness of our proposed DPR algorithm, we investigated whether for a given metric and template the resulting average group profile would be similar using the displacement computed from the other metrics. As the displacement depends on the spectrum of each 1D profile, each metric may use a different template and apply a different displacement for each subject. This may ultimately lead to a different group average profile due to our algorithm automatically choosing the template amongst the subjects. However, the *relative* displacement due to a change of template (and hence the resulting group average 1D profile) may be unaffected by this choice, leading to a similar group average profile. Fig. 11 shows the resulting average group profiles for each metric when using the original realignment and the realignment that would be applied from the two other remaining metrics with a maximum displacement threshold of 15%. As the AF is slowly varying in terms of diffusion metrics along its extracted path, the realignment of the MD metric is similar even when using the displacement computed from the FA or AFD metric. On the other hand, applying the realignment given by the MD to the FA and AFD profiles leads to different optimal realignments and a change in their overall profile. For the CST, as the representative streamline crosses other anatomical bundles along its path, the 1D profiles have more variation along coordinates than in the AF case. This is mostly notable in the MD metric profile which is now similarly realigned when using either the FA or AFD. Due to these anatomical "landmarks", the displacement given by the MD also yields similar profiles when applied to the FA and AFD metrics. Results for maximum displacement thresholds of 5, 10 and 20% produced similar trends which are shown in the supplementary materials Appendix B.2.

Realignment with simulated diffusion abnormalities in HCP datasets. We first focus on the new strategy of resampling the representative streamlines, while ensuring that the distance between each point $\delta_{\min}$ is the

same. As one can always resample to a common number of points after realignment, this prevents a reduced sampling resolution when using Eq. (1). Automatically selecting a template from the subjects themselves allows the DPR algorithm to be as flexible as possible. The changes in scalar metrics (e.g., introduced by local alteration of tissue microstructure following disease) might not be obviously identified on the group average for the unaligned streamlines case, but the variations in shape of the realigned group average may be uncovered by selecting a new template. Fig. 12 shows four examples of the unaligned and realigned profiles of the scalar metrics for the datasets with and without simulated diffusion abnormalities for each white matter fiber bundle. Note how the original and altered unaligned streamlines have a similar profile for both metrics at first, but the realigned altered streamlines now have a different profile which was uncovered by realignment with the DPR algorithm (see the red boxes in Fig. 12). This is especially prevalent in the case of the MD metric where the unaligned profiles are similar for the control and altered subject data, while realignment uncovers the higher MD values that were originally simulated.

Statistical hypothesis testing. We now look at uncovering groupwise differences between the control and altered HCP subjects over the affected regions. Fig. 13 shows the results of the paired t-test for the HCP datasets before and after realignment for the **A)** AF left with the MD metric, **B)** AF right with the FA metric, **C)** CST left with the FA metric and **D)** CST right with the AFD metric, as previously shown in Fig. 12. All of the regions uncovered before using realignment are also identified as statistically significant at the level of p-value < 0.05 after realignment. This indicates that findings for the unaligned case are preserved when using our proposed algorithm, with the addition of new affected regions which might have been averaged out due to misalignment in the first place. For example, the left AF and left CST showcase an affected portion which is statistically significant only after realignment. However, using a lower statistical threshold or a higher level α for the FDR might reveal more affected regions at the cost of introducing potential false positives. Fig. 14 shows a second set of experiments on the four bundles realized with large alterations of the metrics which are spatially focused e.g., in the presence of tumors. Specifically, alterations in the metrics of 25% or 50% were induced over 1% or 5% of each bundle length and each group subsequently realigned with DPR. Paired t-test before and after realignment are conducted between the two groups at each location as in Fig. 13. Almost all affected regions are identified before and after realignment when the affected length is of 5%. For the CST left, the affected region is only identified after realignment when the alteration is of 25%. When only 1% of the bundle length is affected, no changes are identified before realignment, but are partly uncovered after realignment with the DPR algorithm in 6 of the 8 tested cases. Results obtained with maximum displacement thresholds of 5%, 10% and 100% are shown in the supplementary materials Appendix B.3.

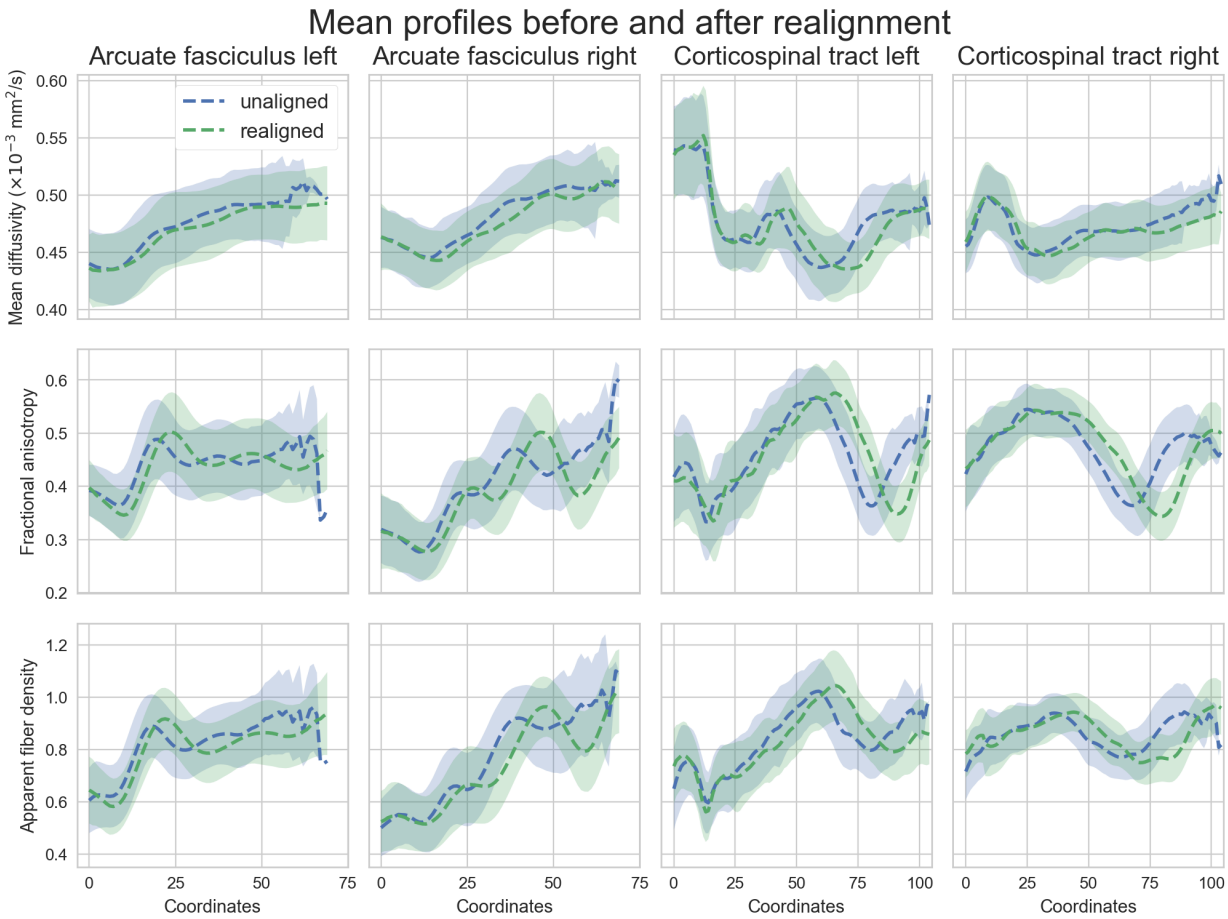

Figure 9: Along-tract averaged profiles (and standard deviation as the shaded area) of the unaligned (blue) and realigned (green) HCP subjects truncated to 75% of overlap with a final resampling to the same number of points. Each row shows the profile for one diffusion metric (MD, FA and AFD) while each column shows one of the studied bundles (AF left/right from anterior (coordinate 0) to posterior and CST left/right from inferior (coordinate 0) to superior). After realignment and truncation, the profiles are slightly different from their unaligned version at the end points while the center profile is similar. This is likely due to the misalignment mostly affecting the initial end points which are defined by the original truncation from the ROIs.

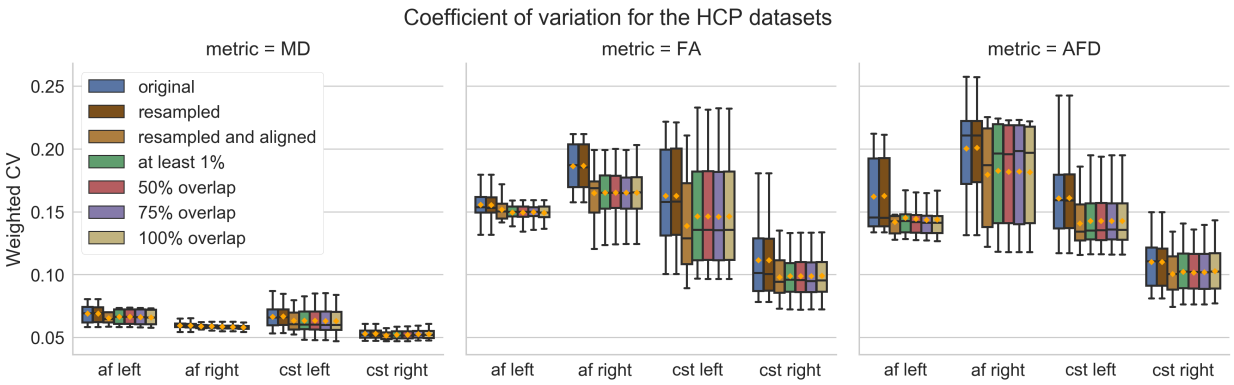

Figure 10: Boxplots of the CV for each point weighted by the number of overlapping subjects, for the MD (left), FA (center) and AFD (right) metrics and their respective mean value (in orange) for the four studied bundles. Similar to the synthetic datasets experiments, the in vivo datasets have a lower CV after realignment (green, red, purple and yellow boxplots) than when they are unaligned (brown boxplots). Even if the representative streamlines are truncated to the shortest number of points (yellow boxplot) or are resampled to the same length (light brown boxplot), the CV is smaller in the realigned cases than in the unaligned cases (brown and blue boxplots respectively). The gain in CV is once again smaller for the MD but larger for the FA and AFD in favor of the realigned cases, which is in line with the synthetic experiments.

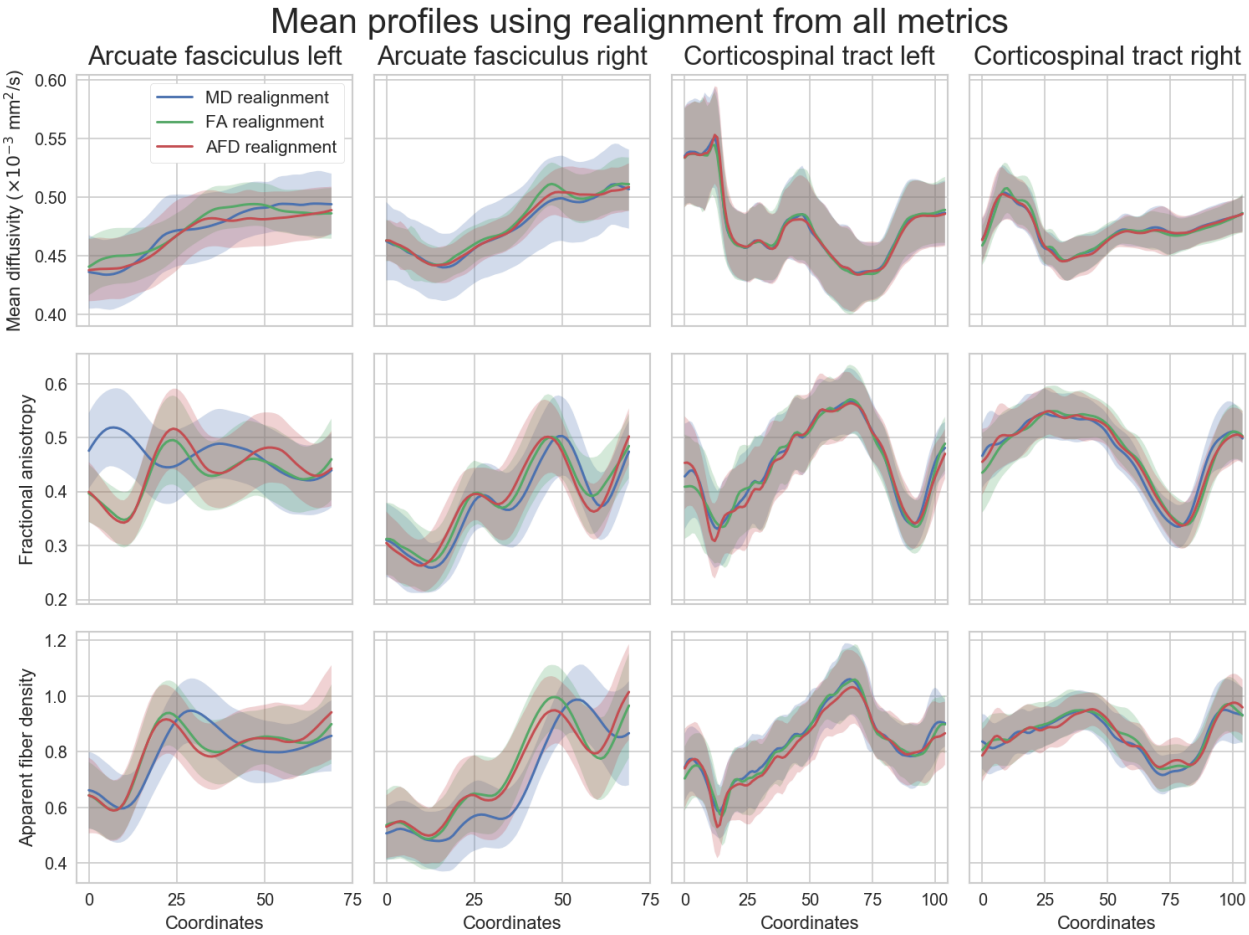

Figure 11: Along-tract averaged profiles (and standard deviation as the shaded area) of the white matter fiber bundles (columns) from the HCP datasets after realignment for each studied metric (rows). The metrics were truncated to 75% of overlap after realignment with a final resampling to the same number of points. On each row, the along-tract profile after realignment is shown for a given metric (MD on the first row, FA on the second row and AFD on the third row) using the displacement computed by the MD (blue), FA (green) and AFD (red). The AF are displayed from anterior (coordinate 0) to posterior and the CST from inferior (coordinate 0) to superior.

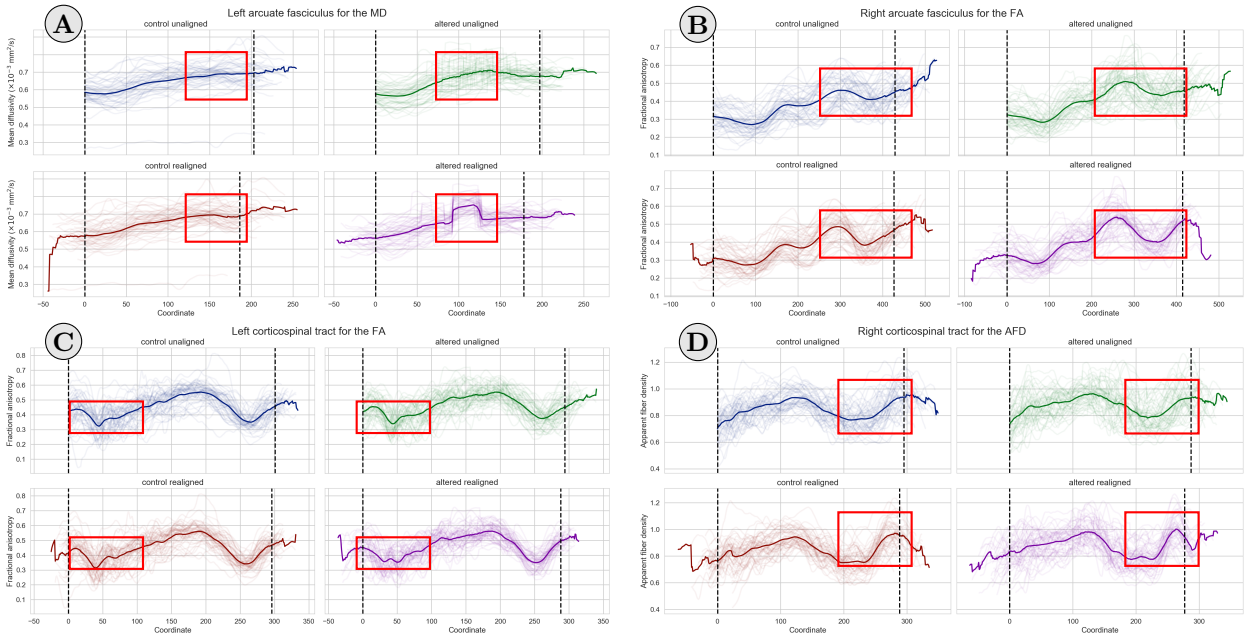

Figure 12: Comparisons between the unaligned and realigned profiles for the HCP datasets without (control column) and with (altered column) simulated diffusion alterations in the white matter fiber bundles. A different bundle for a specific metric is shown in each subfigure: **A)** AF left for the MD, **B)** AF right for the FA, **C)** CST left for the FA and **D)** CST right for the AFD. The AF are displayed from anterior (coordinate 0) to posterior and the CST from inferior (coordinate 0) to superior. Each subject representative streamline is rendered transparently and the group average representative streamline is represented by the solid line. The black bars indicate where at least 75% of the subjects are overlapping. Some key visual differences (red boxes) are hidden by misalignment between the control and altered subject data when they are unaligned, while realignment helps to uncover those hidden degeneracies. Note that the red boxes in the subgraphs have the same size and are aligned for easier visual comparison. The most striking example is in **A)** where the change in MD is easier to see after realignment as the control subjects are keeping their original shape while the altered datasets exhibit a drop in their scalar value around the same region. The unaligned group average streamline however makes this difference harder to uncover.

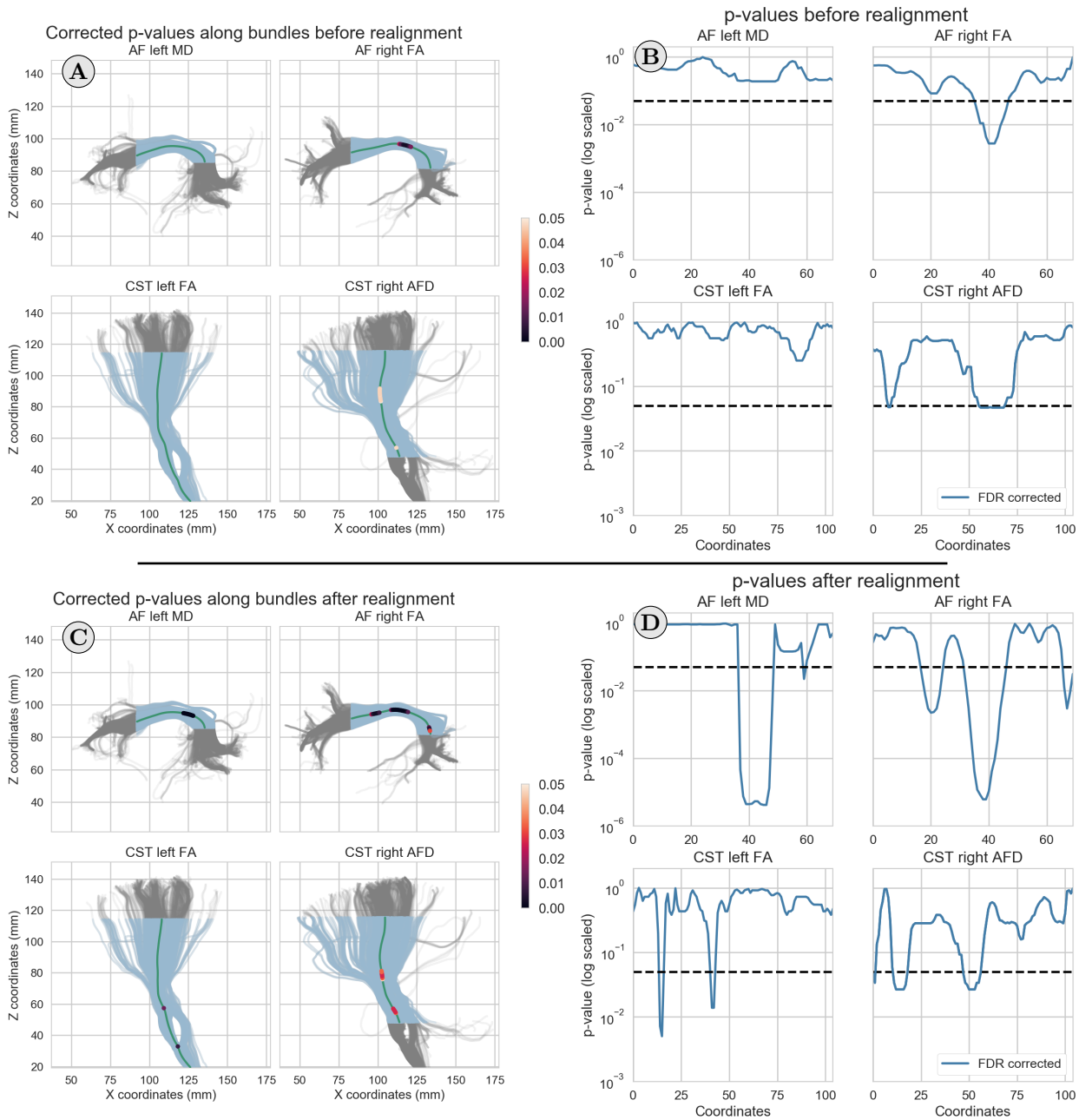

Figure 13: Paired t-test corrected for false discovery rate (FDR) at $\alpha = 0.05$ overlaid on the exemplar subject bundle for the same cases as in Fig. 12. On the left, fiber trajectories of the exemplar subject (in gray) and truncated portions of these pathways between the ROIs (in blue) expressed in world coordinates **A**) before realignment and **C**) after realignment with the DPR algorithm. The p-values at locations deemed statistically significant in the present work ($p < 0.05$) are overlaid on the average streamline (in green). On the right, the p-values on a log scale after FDR correction along the average streamlines **B**) before realignment and **D**) after realignment with the DPR algorithm, but expressed as along-tract 1D point coordinates. The horizontal black bar is located at p-value = 0.05. In the realigned data case, the p-values are lower in the significant regions (corticospinal tract right) or even show affected regions which are not detected when the data is unaligned (arcuate fasciculi and corticospinal tract left). The most prominent case is for the left arcuate fasciculus, where the affected portion is not identified in the unaligned case (for our chosen significance threshold of 0.05), but has a corrected p-value of approximately 10^{-5} after realignment.

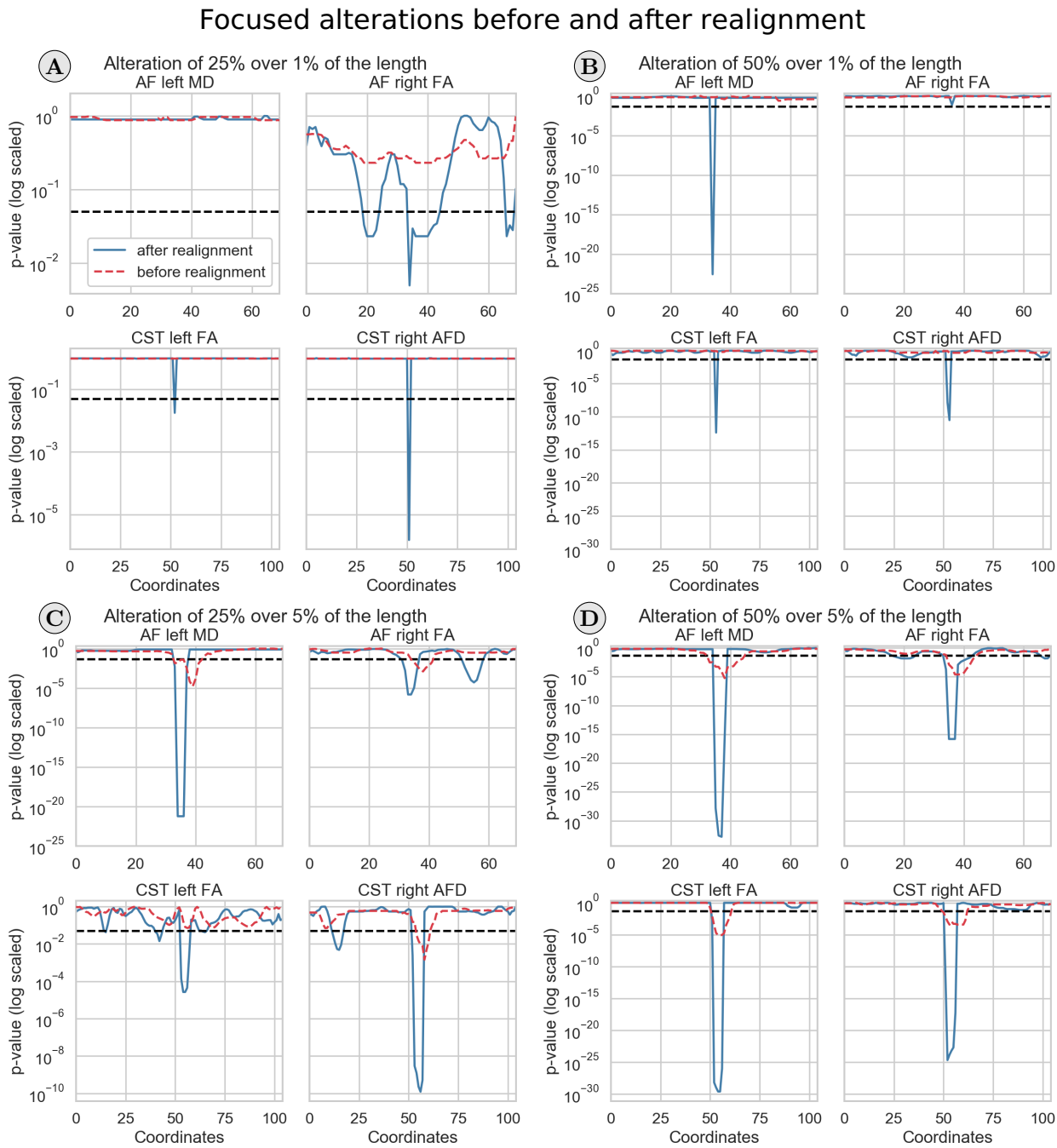

Figure 14: Paired t-test (FDR corrected at $\alpha = 0.05$) with focused alterations of the metrics for each bundle of **A)** 25% over 1% of the length, **B)** 50% over 1% of the length, **C)** 25% over 5% of the length and **D)** 50% over 5% of the length. The AF left/right are represented from anterior (coordinate 0) to posterior and the CST left/right from inferior (coordinate 0) to superior. The p-values are on a log scale along the average streamline before realignment (dashed red lines) and after realignment (solid blue lines) with the DPR algorithm. The horizontal dashed black lines indicate p-value = 0.05. When alterations cover 1% of the length, the affected profiles are identified only after realignment. At 5% of the length, the uncovered regions after realignment are concentrated around smaller sections than their counterpart before realignment.

5. Discussion

5.1. Reducing variability along bundles

Using simulations, we have shown how residual misalignment may hide the expected average profile of an along-tract analysis. Fig. 7 shows this effect directly as the group mean profile from a set of streamlines only roughly corresponds to their individual, but in truth identical, shape as their spatial location differs due to small differences in their length. Realignment not only restores the expected group profile, but also reduces the pointwise variability of the metrics as the unequal streamlines are now aligned as reflected by the lower overall CV. Each individual subject is therefore participating to the group average instead of being spread out and biasing the estimated mean scalar value of the overall bundle in the crossing region. This is also true if the streamlines are first resampled to the same number of points. In this case the variance at the end points is larger, possibly due to a loss in spectral resolution caused by resampling to a lower number of points than originally present. Resampling early in the along-tract analysis pipeline may not only inadvertently hide information for the realignment, but also hamper statistical testing by reducing the spatial specificity of the data (O'Donnell et al., 2009).

For the realignment of the in vivo HCP datasets, Fig. 9 shows that realignment alters the group profile at the end points while preserving the overall shape and the central portion of the bundle. This leads to a reduction of the CV, likely due to the reduction in variance at the end points while the overall mean profile is preserved as shown in Fig. 10. As the realigned end points will also have less data from different subjects present at each coordinate, subsequent truncation further reduces the CV once again. The change of shape after realignment is possibly due to the difference in length between each subject and the subsequent mapping to their 1D metric profile. This 1D space hides the spatial 3D coordinates misalignment which may be present between subjects. However, this misalignment can still be mitigated afterwards. Even if the representatives streamlines are shifted as a whole with the realignment, preservation of the overall shape and center portion might indicate that only the end points were dissimilar. The lower end point variance effect is also present when using the classical resampling strategy and subsequently realigning the representative streamlines. The misalignment at the end points between subjects is due in part to the truncation effect of the ROIs and to the nature of tractography itself and its many user defined settings (Chamberland et al., 2014). The use of termination criteria (e.g., FA threshold, white matter mask, maximum curvature) or seeding strategy (e.g., white matter versus cortex seeding) (Girard et al., 2014) may prematurely terminate tractography in the middle of a white matter bundle, contributing in producing shorter streamlines which end before fully reaching the gray matter (Maier-Hein et al., 2017). New algorithms and seeding strategies are developed to enhance tractography end points near the cortex (St-Onge et al., 2018) and could help to reduce this truncation effect.

5.2. Effect of exchanging metrics for realignment

We have shown in Fig. 11 the effect of applying the realignment computed from different metrics on the mean group tract profile. From these results, we can observe the different displacement values obtained from the dMRI metrics, even though the representative streamlines arise from the same anatomical location. This is due to the fact that our framework is fully driven by the 1D profiles of the studied metric which all have different shapes and features, leading to slightly different realignment outcome depending on the bundle and the metric that is used. As the FA and AFD profiles are similar in the four studied bundles, exchanging their value still leads to the same overall profile in most cases. For the MD, results showed that the CST is also stable. This is most likely due to the complex 1D profile of the CST for the three metrics, as it defines unique landmarks that are picked by our algorithm for accurate realignment. Regarding the AF, exchanging the displacement from the FA or AFD yields similar profiles, an observation which does not hold for the MD metric. As the MD metric for the AF has a rather flat profile, the algorithm might pick up a spurious displacement due to the lack of well defined features to exploit. Avants et al. (2011) also reached a similar conclusion in the context of 3D volume registration when using different metrics such as the mean square difference, cross correlation or mutual information; using different metrics, type of registration or registering subject A onto subject B (and vice versa) leads to slightly different outcomes. We have fixed the maximally allowed displacement to 15% of the length of the bundle, but similar conclusions also applied for 10% and

20% of maximum displacement as shown in Appendix B.2. When the maximum displacement is only 5%, the AF show similar mean profiles for the three metrics, whereas the opposite is seen in the CST. This indicates that the maximally allowed displacement should be chosen per bundle and is data dependent. Short, straight and simpler bundles, such as the AF, might only need small realignment, whereas more complex structures with fanning, intersecting bundles and possibly large anatomical variations between subjects, such as the CST, likely benefit from larger maximum displacement thresholds to find their full overlap between subjects.

5.3. Identifying brain regions affected by abnormalities along-tract

One of the end goal of along-tract analysis is to uncover alteration of the white matter due to, e.g., disease at their specific locations. This is at the cost of trading the sensitivity of ROI averaging based analysis for additional specificity along the bundle, which also depends on the discretization of the points forming the streamlines (O'Donnell et al., 2009). Using simulated changes in scalar metrics from the HCP subjects, we have shown in Fig. 12 how misalignment can artificially reduce the specificity of along-tract analysis. As the affected portion of the bundle is usually unknown *a priori*, morphological differences between subjects might map the affected area to different points in their 1D profile during the representative streamlines extraction. The unaligned metrics might exhibit similar mean profiles between the control and altered subjects in this case, as the affected portions would be originating from an adjacent anatomical location in each subject's original space, but would not be aligned in the 1D space. The mean representative streamline at the group level could therefore average out each subject's individual difference due to residual misalignment, hiding the effect of interest in the process. As we have previously mentioned in Section 1, this effect of averaging out important information has also been theorized by O'Donnell et al. (2009). However, the same effects can also become easier to detect after realignment since the control subjects mean profile will potentially be different from the altered subjects mean profile. This is thanks to the particular features of their 1D profile now being realigned instead of averaged out. In a similar way, if changes in the diffusion metrics are potentially present across the whole length of a white matter bundle, the maximum displacement threshold should be increased. This may reduce the number of subjects identified as outliers by using a smaller maximum displacement, which would not have been realigned in the first place. The tradeoff in allowing a larger maximum displacement is a potential reduction in statistical power or false discoveries as less subjects may be present at each along-tract location for statistical testing.

In our simulations, changes on the left AF and left CST are identifiable only after realignment whereas the original control and altered average profiles are mostly similar since each individual contribution is lost in the unaligned group averaging. After realignment, the altered region can be identified as each realigned subject now contributes to the group average at the same location. This effect is similar to what we observed in our simulations in Section 4.1, where the CV is lower in the crossing-bundles region after realignment and how the mean group profile is also lower after realignment. It is also noticeable on the right AF bundle with the FA metric or on the CST bundles, but to a lesser extent, as the overall morphology of the CST bundles stays relatively similar even after altering the scalar metrics. Interestingly, the altered group profile seems to be subject to larger morphometric changes after realignment than the control group counterpart. This might indicate that sharp profile changes in each subject's shape due to disease are automatically picked up by our algorithm, providing realignment based on this change.

We also conducted paired Student's t-tests to statistically identify the altered regions on the same bundles and metrics as shown in Fig. 13. While we used an FDR correction of $\alpha = 0.05$, different results could be obtained by choosing a different value of α. However, the main conclusion should still be valid; statistical testing performed on the realigned datasets uncovered affected regions which were not identified in the unaligned case as shown from the global p-values plot. This difference could be partly due to the residual misalignment between subjects inadvertently canceling out the effect of interest as coordinates are not overlapping. In this study, we considered statistical testing at the spatial resolution in the order of magnitude of one voxel size (1.25 mm in our case), but studying larger bundle segments could be used as a compromise between averaging data over the whole bundle in order to uncover effects of interest at the expense of spatial specificity (O'Donnell et al., 2009).

5.4. Mapping to 1D space versus registration methods

In the present work, we concentrated on reducing the effect of residual misalignment between representative streamlines. As tractography is a mandatory step before using our approach, registration methods for raw dMRI datasets would likely not reduce the misalignment resulting from streamlines extraction. Some registration methods specifically work directly on the streamlines or bundles space (e.g., Leemans et al. (2006)), but the same transformation should be applied to the underlying 3D volume containing the metric of interest. This is because we work on metrics extracted from representative streamlines, and not directly in the streamlines space itself, see e.g., Glozman et al. (2018); O'Donnell et al. (2017) and references therein for a review of registration methods in dMRI.

O'Donnell et al. (2009) state that "because within a bundle fibers have varying lengths and their point correspondence is not known a priori, it is not possible to directly average fiber coordinates to calculate a mean fiber"; care must be taken during the representative streamline extraction step that is at the core of the along-tract analysis framework. As such, the required step dictating this possible misalignment is the mapping strategy used to extract the representative streamline and how its end points are defined. Various schemes have been proposed such as assignment to perpendicular planes (Corouge et al., 2006), variants reducing the effect of outliers by additionally considering the spatial distance between streamlines (O'Donnell et al., 2009), extracting representative core streamlines with splines (Chamberland et al., 2018) or resampling to a common number of points (Colby et al., 2012). All these choices inevitably lead to differences and a mismatch across subjects after metric extraction, even if the original underlying anatomy would be perfectly aligned as we have shown in our synthetic experiments in Fig. 7. Assignment and truncation strategies between the common points of bundles have been explored in Colby et al. (2012) with the authors noting that all compared methods are generally successful in extracting a meaningful (but slightly different) representation as they use different strategies and parameters. Close similarities in the extracted metrics using the representative streamline could explain why 1D misalignment, while still present, had not been thoroughly investigated previously. Reliably extracting the information from fanning regions (e.g., CST towards the motor cortex) or from a splitting configuration (e.g., anterior pillars of the fornix) in a single representative streamline still remains an open problem (Chamberland et al., 2018).

5.5. Assumptions of the DPR algorithm and limitations of this study

In the present work, we exchanged the classical assumption of 1D *spatial correspondence* between points for the assumption of an equal 1D *spatial distance* between points. This latter requirement is usually fulfilled with the use of a fixed step size during tractography, but might be void by the representative streamline extraction. Without loss of generality, we chose to resample each subjects' representative streamline a second time to ensure an equal distance $\delta_{\min}$ between each point. We advocate resampling to a larger number of points than initially present to reduce possible complications due to aliasing or using windowing functions for filtering (Stoica and Moses, 2005). While this theoretically increases the computational complexity of the DPR algorithm, it also preserves the full spectra when applying Eq. (1). This is not a problem in practice owing to the existence of efficient FFT implementations; our algorithm can realign the 100 HCP subjects in less than 3 seconds on a standard desktop with a 3.5 GHz Intel processor. The resulting realigned metrics can then be resampled back to approximately one point per unit voxel size to minimize the effect of multiple comparisons during statistical testing. With the development of new methods that go beyond fixed step size tractography, such as the use of compressed streamlines (Rheault et al., 2017), it might be beneficial to avoid this resampling step for computational reasons after sampling metrics along non regularly spaced streamlines. Another approach to remove the need of resampling could be to use an FFT implementation dealing with non equal sampling of the data (Dutt and Rokhlin, 1993; Scargle, 1989), but such implementations may not be as widely (and easily) available as the classical equispaced version of the FFT algorithm.

Due to the difficulty in reproducing tractography (Maier-Hein et al., 2017), our simulations on the in vivo datasets were designed around altered versions of already extracted scalar values. One would however expect true neurodegenerative changes to additionally influence the steps prior to tractography such as the main orientations extracted from tensors or fODF. The results we obtained should translate as long as a representative streamline for each bundle of interest can be reliably delineated for all subjects. Similar

recommendations apply if the white matter bundle of interest is largely affected by disease or altered when compared to the expected overall shape from a healthy subject. Specific care should also be taken during the prior step of extracting the representative streamlines in these cases to ensure that relevant portions of the bundles of interest are present in all subjects (Parker et al., 2016).

Although not considered in the present work, any quantitative diffusion metric such as the diffusion kurtosis metrics (e.g., mean kurtosis (MK)) (Jensen and Helpern, 2010), the axon diameter (Assaf et al., 2008), or metrics provided by NODDI (Zhang et al., 2012) could be studied using our proposed framework. In cases of physical alterations of the white matter (e.g., tumors, lesions), the diffusion metrics themselves may not provide accurate landmarks for realignment due to differences in tractography when extracting the representative streamline of each subject. The use of shape descriptors, such as torsion or curvature of the bundles themselves (Leemans et al., 2006), could also be employed with DPR instead of diffusion metrics as done in the present work. These descriptors may also be useful in cases where using a large maximum displacement threshold may yield false positives detections if the effects are small, see the supplementary materials Appendix B.3 for examples. In a similar fashion, any other volume (e.g., T1 or T2 relaxometry values (Deoni et al., 2008)) providing anatomical information of interest can be used once co-registered to each subject's native diffusion space.

We did not investigate realignment of lateralized bundles (e.g., realignment of the left and right AF together instead of separately) which can be useful for studying intra-hemispheric differences between subjects (Catani et al., 2007). Variations between left and right anatomical locations also implicitly assumes that each coordinate in the 1D space is matched against its inter hemispheric counterpart. To facilitate this mapping between hemispheres, O'Donnell et al. (2009) proposed to mirror all streamlines from one hemisphere to the other, allowing a direct correspondence between the subsequently extracted representative streamlines as they would be effectively identical. However, the 3D volume used to extract the scalar metrics of interest would possibly be different in each hemisphere. In this context, the realignment could be done separately for each side, providing different profiles reflecting lateralization.

6. Conclusion

In this paper, we developed a new correction strategy, the diffusion profile realignment (DPR), which is designed to address residual misalignments between subjects in along-tract analysis. Through simulations on synthetic and in vivo datasets, we have shown how realignment based on our novel approach can reduce variability at the group level between subjects. Furthermore, realignment of the in vivo datasets provided new insights and improved sensitivity about the location of the induced changes, which could not be completely identified at first when misalignment was present. The DPR algorithm can be integrated in preexisting along-tract analysis pipelines as it comes just before conducting statistical analysis. It can be used to reveal effects of interest which may be hidden by misalignment and has the potential to improve the specificity in longitudinal population studies beyond the traditional ROI based analysis and along-tract analysis workflows.

Appendix A. The diffusion profile realignment algorithm

This appendix outlines the diffusion profile realignment (DPR) algorithm. Our implementation is also freely available at `https://github.com/samuelstjean/dpr` ([Code] St-Jean, 2019) and will be a part of *ExploreDTI* (Leemans et al., 2009). The synthetic datasets and metrics extracted along the representative streamlines of the HCP datasets used in this manuscript are also available ([Dataset] St-Jean et al., 2018).

To complement Eq. (1), the shift needed to maximize the overlap between the vector x and y is the maximum of the CCF, given by

$$\text{shift}(x, y) = \arg\max(\text{CCF}(x, y)). \tag{A.1}$$

In practice, x and y are discrete and must be both zero-padded sufficiently, that is, zeros are appended to each vector and make them artificially longer to prevent border effects when computing the linear cross-correlation (Stoica and Moses, 2005).

Algorithm 1: The proposed diffusion profile realignment (DPR) algorithm.

Data: Metrics extracted from streamlines discretized (with an equal distance $\delta_{\min}$ and stationary metrics), displacement threshold $\mathbf{t}$, percentage of overlap $\mathbf{p}\%$

Result: Realigned metrics

```
Step 1 : Finding a common template;
foreach streamline do
    Compute the displacement d with each other streamline using Eq. (A.1);
end

Define the template as the subject which realigns the most streamlines below the threshold t;

foreach streamline do
    if |d| ≤ t then
        Realign the streamline unto the candidate template by its displacement d;
    else
        Do not touch the streamline and flag it as an outlier;
    end
end

Step 2 : Realigning outliers;
foreach outlier do
    Compute the new displacement nd between the template, the outlier and each other non outlier;
    if min(|d + nd|) < t then
        Realign the streamline unto the template using the new displacement d + nd (see Fig. 4);
        Add the streamline to the pool of non outliers candidates such that it can now be used;
    else
        Do not touch the streamline and flag it as an outlier;
    end
end

Step 3 : Truncating to overlapping coordinates;
Truncate the realigned metrics to have at least p% of overlapping streamlines;
If outliers are still present from Step 2, (optionally) exclude them from further analysis as they can
not be realigned inside the chosen displacement threshold t;
```

Acknowledgments

We would like to thank Victor Bodiut for useful comments and discussions. Samuel St-Jean was supported by the Fonds de recherche du Québec - Nature et technologies (FRQNT), (Dossier 192865). Maxime Chamberland was supported by the Postdoctoral Fellowships Program from the Natural Sciences and Engineering Research Council of Canada (NSERC) (PDF-502385-2017) This research is supported by VIDI Grant 639.072.411 from the Netherlands Organization for Scientific Research (NWO). The funding agencies were not involved in the design, data collection nor interpretation of this study. Declarations of interest: none. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J., 2008. AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. Magnetic resonance in medicine 59, 1347–54. doi:10.1002/mrm.21577.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage 54, 2033–44. doi:10.1016/j.neuroimage.2010.09.025.

Bach, M., Laun, F.B., Leemans, A., Tax, C.M., Biessels, G.J., Stieltjes, B., Maier-Hein, K.H., 2014. Methodological considerations on tract-based spatial statistics (TBSS). NeuroImage 100, 358–369. doi:10.1016/j.neuroimage.2014.06.021.

Basser, P.J., Pierpaoli, C., 1996. Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI. Journal of Magnetic Resonance, Series B 111, 209–219. doi:10.1006/jmrb.1996.0086.

Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system - a technical review. NMR in Biomedicine 15, 435–455. doi:10.1002/nbm.782.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57, 289–300. doi:10.2307/2346101.

Blain, C.R.V., Brunton, S., Williams, V.C., Leemans, A., Turner, M.R., Andersen, P.M., Catani, M., Stanton, B.R., Ganesalingham, J., Jones, D.K., Williams, S.C.R., Leigh, P.N., Simmons, A., 2011. Differential corticospinal tract degeneration in homozygous 'D90A' SOD-1 ALS and sporadic ALS. Journal of neurology, neurosurgery, and psychiatry 82, 843–9. doi:10.1136/jnnp.2010.236018.

Bodini, B., Ciccarelli, O., 2009. Diffusion MRI in Neurological Disorders, in: Diffusion MRI. Elsevier, pp. 175–203. doi:10.1016/B978-0-12-374709-9.00009-2.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 2008. Time Series Analysis. John Wiley & Sons, Inc., Hoboken, NJ. doi:10.1002/9781118619193.

Caruyer, E., Daducci, A., Descoteaux, M., Houde, J.C., Thiran, J.P., Verma, R., 2014. Phantomas: a flexible software library to simulate diffusion MR phantoms, in: International Symposium on Magnetic Resonance in Medicine (ISMRM'14).

Catani, M., Allin, M.P.G., Husain, M., Pugliese, L., Mesulam, M.M., Murray, R.M., Jones, D.K., 2007. Symmetries in human brain language pathways correlate with verbal recall. Proceedings of the National Academy of Sciences 104, 17163–17168. doi:10.1073/pnas.0702116104.

Cercignani, M., Gandini Wheeler-Kingshott, C., 2018. From micro- to macro-structures in multiple sclerosis: what is the added value of diffusion imaging. NMR in Biomedicine , e3888doi:10.1002/nbm.3888.

Chamberland, M., St-jean, S., Tax, C.M.W., Derek, K., 2018. Obtaining representative core streamlines for white matter tractometry of the human brain, in: Bonet-Carne, E., Grussu, F., Ning, L., Sepehrband, F., Tax, C.M.W. (Eds.), Computational Diffusion MRI. Springer International Publishing, Granada, pp. 3–19. doi:10.1007/978-3-030-05831-9_28.

Chamberland, M., Whittingstall, K., Fortin, D., Mathieu, D., Descoteaux, M., 2014. Real-time multi-peak tractography for instantaneous connectivity display. Frontiers in neuroinformatics 8, 59. doi:10.3389/fninf.2014.00059.

Colby, J.B., Soderberg, L., Lebel, C., Dinov, I.D., Thompson, P.M., Sowell, E.R., 2012. Along-tract statistics allow for enhanced tractography analysis. NeuroImage 59, 3227–3242. doi:10.1016/j.neuroimage.2011.11.004.

Cooley, J.W., Tukey, J.W., 1965. An algorithm for the machine calculation of complex Fourier series. Mathematics of Computation 19, 297–297. doi:10.1090/S0025-5718-1965-0178586-1.

Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G., 2006. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. Medical Image Analysis 10, 786–798. doi:10.1016/j.media.2006.07.003.

Cousineau, M., Jodoin, P.M., Morency, F.C., Rozanski, V., Grand'Maison, M., Bedell, B.J., Descoteaux, M., 2017. A test-retest study on Parkinson's PPMI dataset yields statistically significant white matter fascicles. NeuroImage. Clinical 16, 222–233. doi:10.1016/j.nicl.2017.07.020.

Dayan, M., Monohan, E., Pandya, S., Kuceyeski, A., Nguyen, T.D., Raj, A., Gauthier, S.A., 2016. Profilometry: A new statistical framework for the characterization of white matter pathways, with application to multiple sclerosis. Human Brain Mapping 37, 989–1004. doi:10.1002/hbm.23082.

Dell'Acqua, F., Simmons, A., Williams, S.C., Catani, M., 2013. Can spherical deconvolution provide more information than fiber orientations? Hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. Human Brain Mapping 34, 2464–2483. doi:10.1002/hbm.22080.

Denman, K.L., 1975. Spectral Analysis: A summary of the theory and techniques. Fish. Mar. Servo Res. Dev. Tech. Rep. 539 , 1–32.

Deoni, S.C.L., Rutt, B.K., Arun, T., Pierpaoli, C., Jones, D.K., 2008. Gleaning multicomponent T1 and T2 information from steady-state imaging data. Magnetic Resonance in Medicine 60, 1372–1387. doi:10.1002/mrm.21704.

Dutt, A., Rokhlin, V., 1993. Fast Fourier Transforms for Nonequispaced Data. SIAM Journal on Scientific Computing 14, 1368–1393. doi:10.1137/0914081.

Froeling, M., Pullens, P., Leemans, A., 2016. DTI Analysis Methods: Region of Interest Analysis, in: Van Hecke, W., Emsell, L., Sunaert, S. (Eds.), Diffusion Tensor Imaging. Springer New York, New York, NY. October 2017, pp. 175–182. doi:10.1007/978-1-4939-3118-7_9.

Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M., 2014. Towards quantitative connectivity analysis: reducing tractography biases. NeuroImage 98, 266–78. doi:10.1016/j.neuroimage.2014.04.074.

Glozman, T., Bruckert, L., Pestilli, F., Yecies, D.W., Guibas, L.J., Yeom, K.W., 2018. Framework for shape analysis of white matter fiber bundles. NeuroImage 167, 466–477. doi:10.1016/j.neuroimage.2017.11.052.

Haakma, W., Jongbloed, B.A., Froeling, M., Goedee, H.S., Bos, C., Leemans, A., van den Berg, L.H., Hendrikse, J., van der Pol, W.L., 2017. MRI shows thickening and altered diffusion in the median and ulnar nerves in multifocal motor neuropathy. European Radiology 27, 2216–2224. doi:10.1007/s00330-016-4575-0.

Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V.J., Meuli, R., Thiran, J.P., 2007. Mapping Human Whole-Brain Structural Networks with Diffusion MRI. PLoS ONE 2, e597. doi:10.1371/journal.pone.0000597.

Jensen, J.H., Helpern, J.A., 2010. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. NMR in biomedicine 23, 698–710. doi:10.1002/nbm.1518.

Jeurissen, B., Descoteaux, M., Mori, S., Leemans, A., 2017. Diffusion MRI fiber tractography of the brain. NMR in Biomedicine , e3785doi:10.1002/nbm.3785.

Jin, Y., Huang, C., Daianu, M., Zhan, L., Dennis, E.L., Reid, R.I., Jack, C.R., Zhu, H., Thompson, P.M., Alzheimer's Disease Neuroimaging Initiative, 2017. 3D tract-specific local and global analysis of white matter integrity in Alzheimer's disease. Human brain mapping 38, 1191–1207. doi:10.1002/hbm.23448.

Jones, D.K., Cercignani, M., 2010. Twenty-five pitfalls in the analysis of diffusion MRI data. NMR in Biomedicine 23, 803–820. doi:10.1002/nbm.1543.

Jones, D.K., Travis, A.R., Eden, G., Pierpaoli, C., Basser, P.J., 2005. PASTA: Pointwise assessment of streamline tractography attributes. Magnetic Resonance in Medicine 53, 1462–1467. doi:10.1002/mrm.20484.

Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J., 2010. elastix: A Toolbox for Intensity-Based Medical Image Registration. IEEE Transactions on Medical Imaging 29, 196–205. doi:10.1109/TMI.2009.2035616.

Lebel, C., Walker, L., Leemans, A., Phillips, L., Beaulieu, C., 2008. Microstructural maturation of the human brain from childhood to adulthood. NeuroImage 40, 1044–1055. doi:10.1016/j.neuroimage.2007.12.053.

Leemans, A., Jeurissen, B., Sijbers, J., Jones, D., 2009. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine 17, 3537.

Leemans, a., Sijbers, J., De Backer, S., Vandervliet, E., Parizel, P., 2006. Multiscale white matter fiber tract coregistration: A new feature-based approach to align diffusion tensor data. Magnetic Resonance in Medicine 55, 1414–1423. doi:10.1002/mrm.20898.

Lewis, J.P., 1995. Fast normalized cross-correlation. Vision interface 10, 120–123.

Maier-Hein, K.H., Neher, P.F., Houde, J.C., Côté, M.A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.C., Lin, Y.C., Ji, Q., Reddick, W.E., Glass, J.O., Chen, D.Q., Feng, Y., Gao, C., Wu, Y., Ma, J., Renjie, H., Li, Q., Westin, C.F., Deslauriers-Gauthier, S., González, J.O.O., Paquette, M., St-Jean, S., Girard, G., Rheault, F., Sidhu, J., Tax, C.M.W., Guo, F., Mesri, H.Y., Dávid, S., Froeling, M., Heemskerk, A.M., Leemans, A., Boré, A., Pinsard, B., Bedetti, C., Desrosiers, M., Brambati, S., Doyon, J., Sarica, A., Vasta, R., Cerasa, A., Quattrone, A., Yeatman, J., Khan, A.R., Hodges, W., Alexander, S., Romascano, D., Barakovic, M., Auría, A., Esteban, O., Lemkaddem, A., Thiran, J.P., Cetingul, H.E., Odry, B.L., Mailhe, B., Nadar, M.S., Pizzagalli, F., Prasad, G., Villalon-Reina, J.E., Galvis, J., Thompson, P.M., Requejo, F.D.S., Laguna, P.L., Lacerda, L.M., Barrett, R., Dell'Acqua, F., Catani, M., Petit, L., Caruyer, E., Daducci, A., Dyrby, T.B., Holland-Letz, T., Hilgetag, C.C., Stieltjes, B., Descoteaux, M., 2017. The challenge of mapping the human connectome based on diffusion tractography. Nature Communications 8, 1349. doi:10.1038/s41467-017-01285-x.

Mori, S., Van Zijl, P.C.M., 2002. Fiber tracking: Principles and strategies - A technical review. NMR in Biomedicine 15, 468–480. doi:10.1002/nbm.781.

Nimsky, C., Bauer, M., Carl, B., 2016. Merits and Limits of Tractography Techniques for the Uninitiated, in: Advances and technical standards in neurosurgery. 43, pp. 37–60. doi:10.1007/978-3-319-21359-0_2.

O'Donnell, L.J., Daducci, A., Wassermann, D., Lenglet, C., 2017. Advances in computational and statistical diffusion MRI. NMR in Biomedicine , e3805doi:10.1002/nbm.3805.

O'Donnell, L.J., Pasternak, O., 2015. Does diffusion MRI tell us anything about the white matter? An overview of methods and pitfalls. Schizophrenia Research 161, 133–141. doi:10.1016/j.schres.2014.09.007.

O'Donnell, L.J., Westin, C.F., Golby, A.J., 2009. Tract-based morphometry for white matter group analysis. NeuroImage 45, 832–844. doi:10.1016/j.neuroimage.2008.12.023.

O'Hanlon, E., Leemans, A., Kelleher, I., Clarke, M.C., Roddy, S., Coughlan, H., Harley, M., Amico, F., Hoscheit, M.J., Tiedt, L., Tabish, J., McGettigan, A., Frodl, T., Cannon, M., 2015. White matter differences among adolescents reporting psychotic experiences: A population-based diffusion magnetic resonance imaging study. JAMA Psychiatry 72, 668–677. doi:10.1001/jamapsychiatry.2015.0137.

Parker, G.D., LLoyd, D., Jones, D.K., 2016. The best of both worlds: Combining the strengths of TBSS and tract-specific mea-

surements for group-wise comparison of white matter microstructure, in: International Symposium on Magnetic Resonance in Medicine (ISMRM'16).

Platt, T., Denman, K.L., 1975. Spectral Analysis in Ecology. Annual Review of Ecology and Systematics 6, 189–210. doi:10.1146/annurev.es.06.110175.001201.

Raffelt, D., Tournier, J.D., Rose, S., Ridgway, G.R., Henderson, R., Crozier, S., Salvado, O., Connelly, A., 2012. Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. NeuroImage 59, 3976–3994. doi:10.1016/j.neuroimage.2011.10.045.

Rheault, F., Houde, J.C., Descoteaux, M., 2017. Visualization, Interaction and Tractometry: Dealing with Millions of Streamlines from Diffusion MRI Tractography. Frontiers in Neuroinformatics 11, 42. doi:10.3389/fninf.2017.00042.

Scargle, J.D., 1989. Studies in astronomical time series analysis. III - Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data. The Astrophysical Journal 343, 874. doi:10.1086/167757.

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., Behrens, T.E., 2006. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. NeuroImage 31, 1487–1505. doi:10.1016/j.neuroimage.2006.02.024.

Sporns, O., Tononi, G., Kötter, R., 2005. The human connectome: A structural description of the human brain. PLoS Computational Biology 1, 0245–0251. doi:10.1371/journal.pcbi.0010042.

St-Jean, S., 2019. samuelstjean/dpr: [v0.1.1b] - 2019-03-13. doi:10.5281/zenodo.2592014.

St-Jean, S., Chamberland, M., Viergever, M.A., Leemans, A., 2018. Datasets for 'Reducing variability in along-tract analysis with diffusion profile realignment'. Zenodo doi:10.5281/zenodo.2483169.

St-Jean, S., Viergever, M.A., Biessels, G.J., Leemans, A., 2016. Correcting spatial misalignment between fiber bundles segments for along-tract group analysis, in: International Symposium on Magnetic Resonance in Medicine (ISMRM'16).

St-Onge, E., Daducci, A., Girard, G., Descoteaux, M., 2018. Surface-enhanced tractography (SET). NeuroImage 169, 524–539. doi:10.1016/j.neuroimage.2017.12.036.

Stoica, P., Moses, R., 2005. Spectral Analysis of Signals. Prentice Hall.

Szczepankiewicz, F., Lätt, J., Wirestam, R., Leemans, A., Sundgren, P., van Westen, D., Ståhlberg, F., Nilsson, M., 2013. Variability in diffusion kurtosis imaging: Impact on study design, statistical power and interpretation. NeuroImage 76, 145–154. doi:10.1016/j.neuroimage.2013.02.078.

Tax, C.M., Jeurissen, B., Vos, S.B., Viergever, M.A., Leemans, A., 2014. Recursive calibration of the fiber response function for spherical deconvolution of diffusion MRI data. NeuroImage 86, 67–80.

Tax, C.M.W., Otte, W.M., Viergever, M.A., Dijkhuizen, R.M., Leemans, A., 2015. REKINDLE: Robust Extraction of Kurtosis INDices with Linear Estimation. Magnetic Resonance in Medicine 73, 794–808. doi:10.1002/mrm.25165.

Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. NeuroImage 35, 1459–72.

Tournier, J.D., Calamante, F., Connelly, A., 2013. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. NMR in biomedicine 26, 1775–86. doi:10.1002/nbm.3017.

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., 2012. The Human Connectome Project: A data acquisition perspective. NeuroImage 62, 2222–2231. doi:10.1016/j.neuroimage.2012.02.018.

Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. NeuroImage 36, 630–644. doi:10.1016/j.neuroimage.2007.02.049.

Wassermann, D., Makris, N., Rathi, Y., Shenton, M., Kikinis, R., Kubicki, M., Westin, C.F., 2016. The white matter query language: a novel approach for describing human white matter anatomy. Brain Structure and Function 221, 4705–4721. doi:10.1007/s00429-015-1179-4.

Yeatman, J.D., Dougherty, R.F., Myall, N.J., Wandell, B.A., Feldman, H.M., 2012. Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. PLoS ONE 7, e49790. doi:10.1371/journal.pone.0049790.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. NeuroImage 61, 1000–16. doi:10.1016/j.neuroimage.2012.03.072.

Appendix B. Supplementary materials

Appendix B.1. Realignment of the HCP datasets

Appendix B.1 presents counterpart results to Fig. 9, comparing along-tract averaged profiles before and after realignment, but instead using a maximally allowed displacement of 5%, 10% or 20%. Coordinates for the AF are from anterior (coordinate 0) to posterior and the CST are drawn from inferior (coordinate 0) to superior. In general, the overall mean profile is similar for every value of the maximally allowed displacement that were tested.

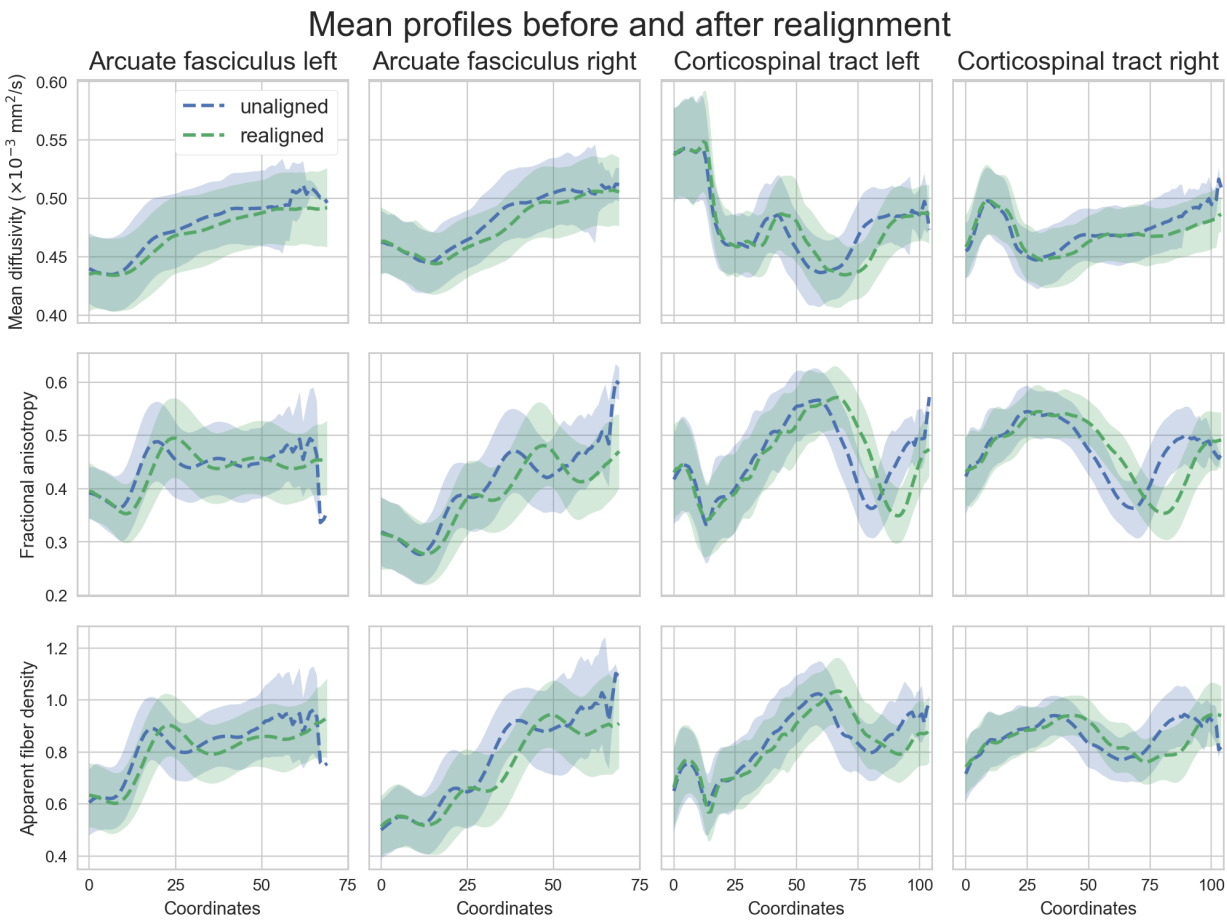

Figure B.15: Along-tract averaged profiles (and standard deviation as the shaded area) of the unaligned (blue) and realigned (green) HCP subjects truncated to 75% of overlap with a final resampling to the same number of points. These results are obtained by using a maximally allowed displacement of 5%.

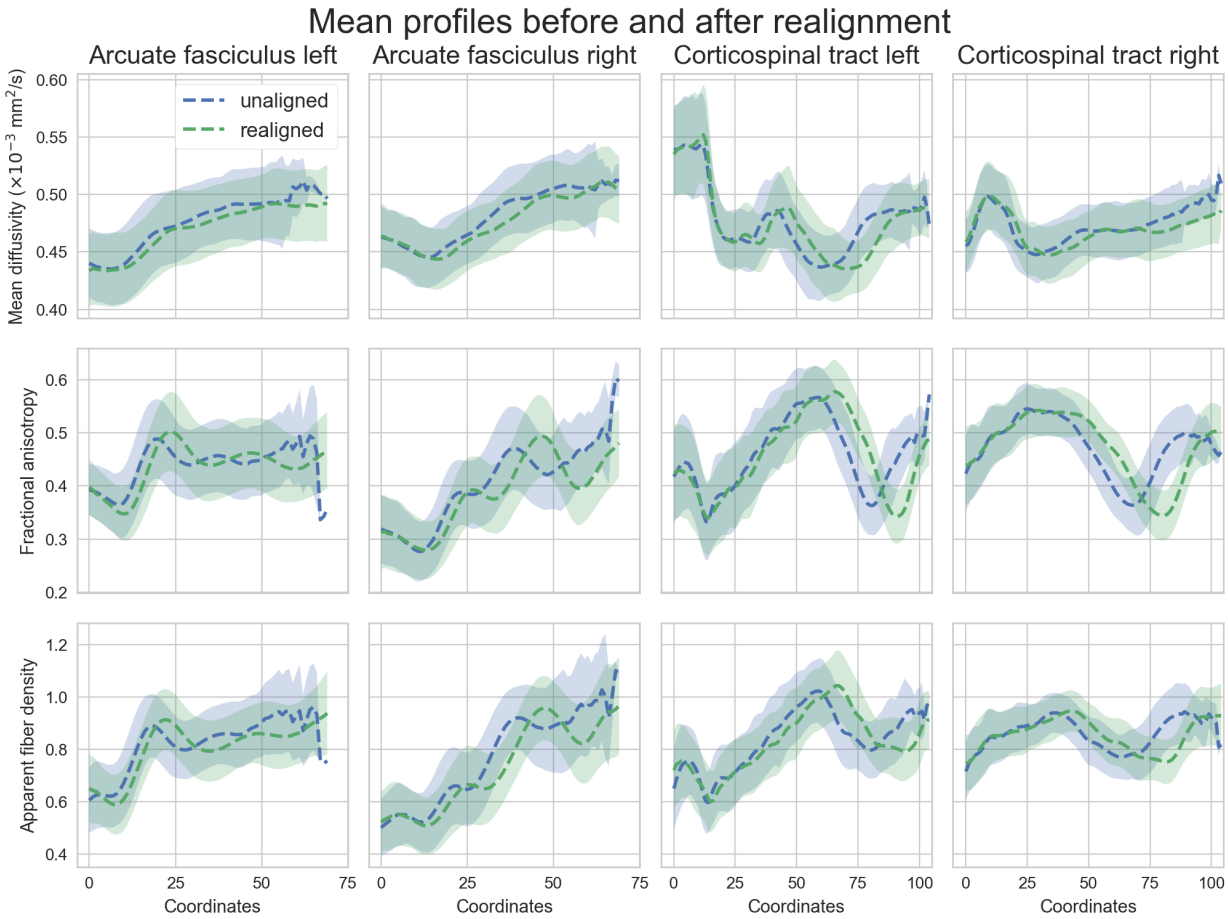

Figure B.16: Along-tract averaged profiles (and standard deviation as the shaded area) of the unaligned (blue) and realigned (green) HCP subjects truncated to 75% of overlap with a final resampling to the same number of points. These results are obtained by using a maximally allowed displacement of 10%.

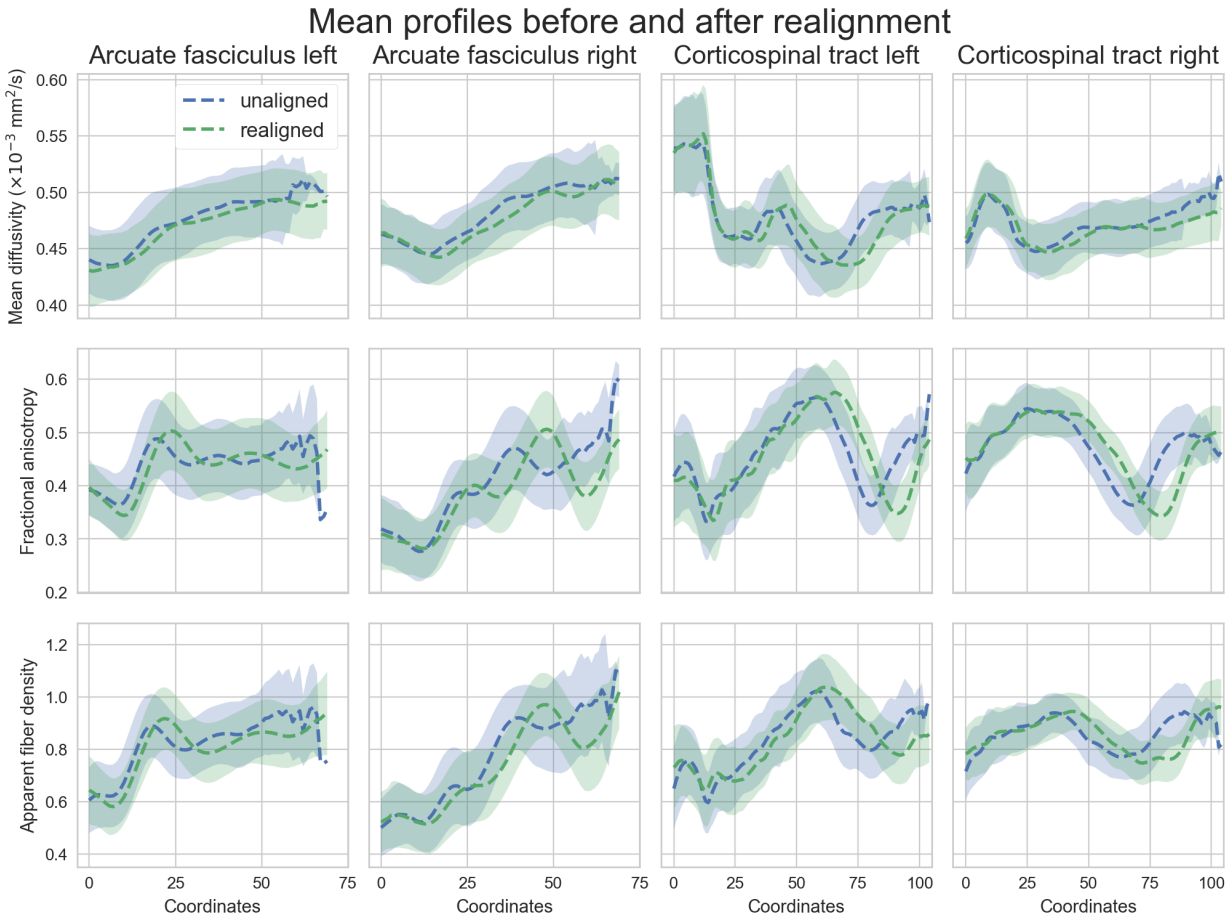

Figure B.17: Along-tract averaged profiles (and standard deviation as the shaded area) of the unaligned (blue) and realigned (green) HCP subjects truncated to 75% of overlap with a final resampling to the same number of points. These results are obtained by using a maximally allowed displacement of 20%.

Appendix B.2. Displacement of the HCP datasets

Appendix B.2 presents counterpart results to Fig. 11 using realignment from other metrics, but instead using a maximally allowed displacement of 5%, 10% or 20%. Coordinates for the AF are from anterior (coordinate 0) to posterior and the CST are drawn from inferior (coordinate 0) to superior. Most of the trends observed previously when the displacement was of 15% are still valid.

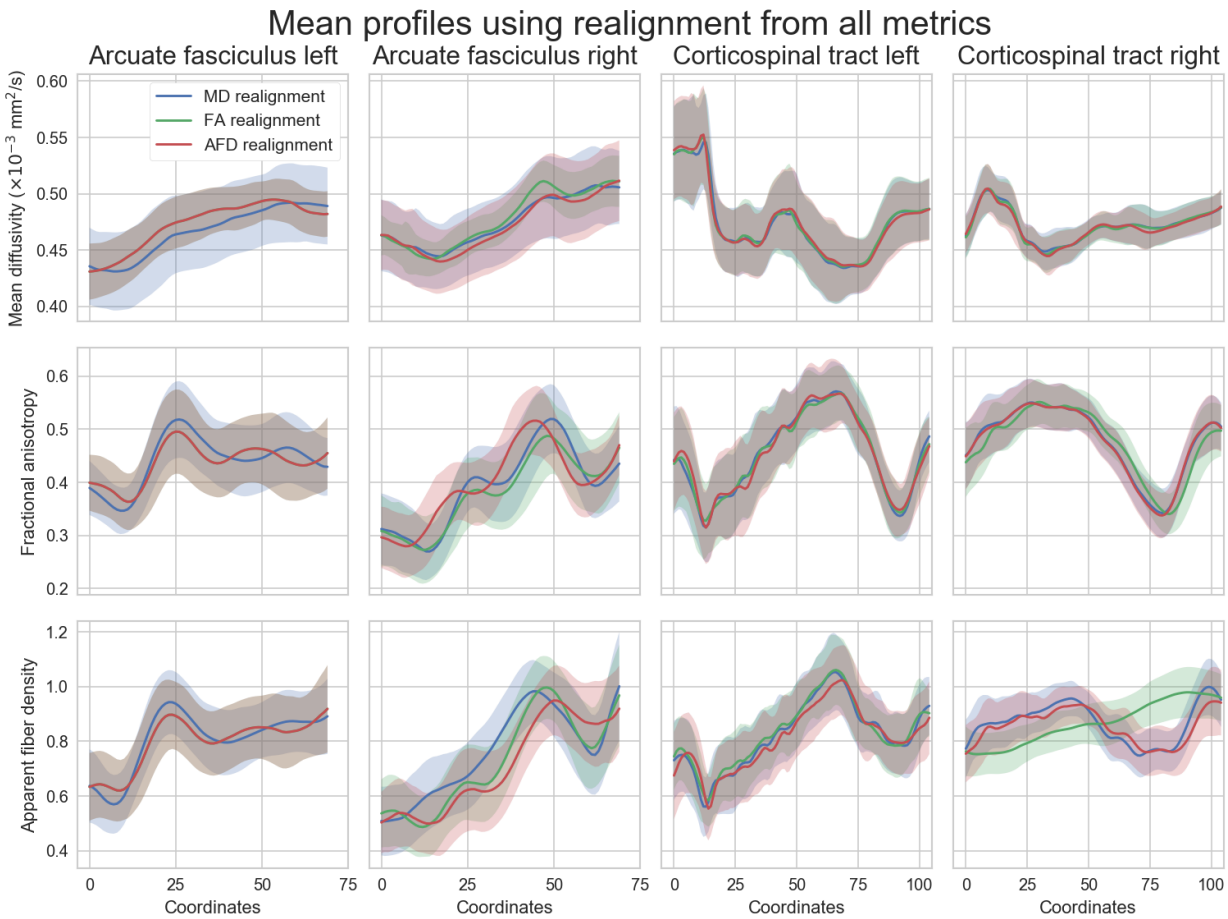

Figure B.18: Along-tract averaged profiles (and standard deviation as the shaded area) of the white matter fiber bundles (columns) from the HCP datasets after realignment for each studied metric (rows). These results are obtained by using a maximally allowed displacement of 5%.

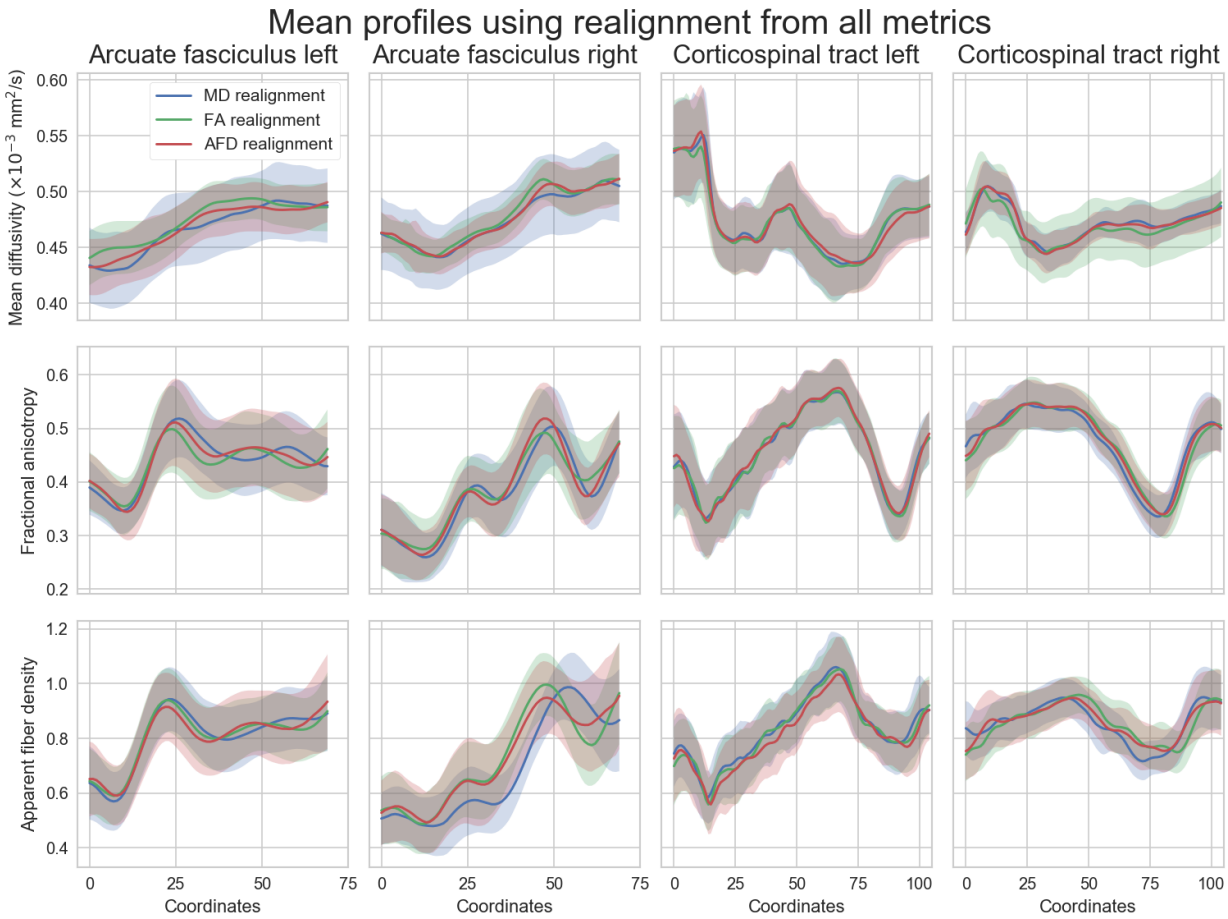

Figure B.19: Along-tract averaged profiles (and standard deviation as the shaded area) of the white matter fiber bundles (columns) from the HCP datasets after realignment for each studied metric (rows). These results are obtained by using a maximally allowed displacement of 10%.

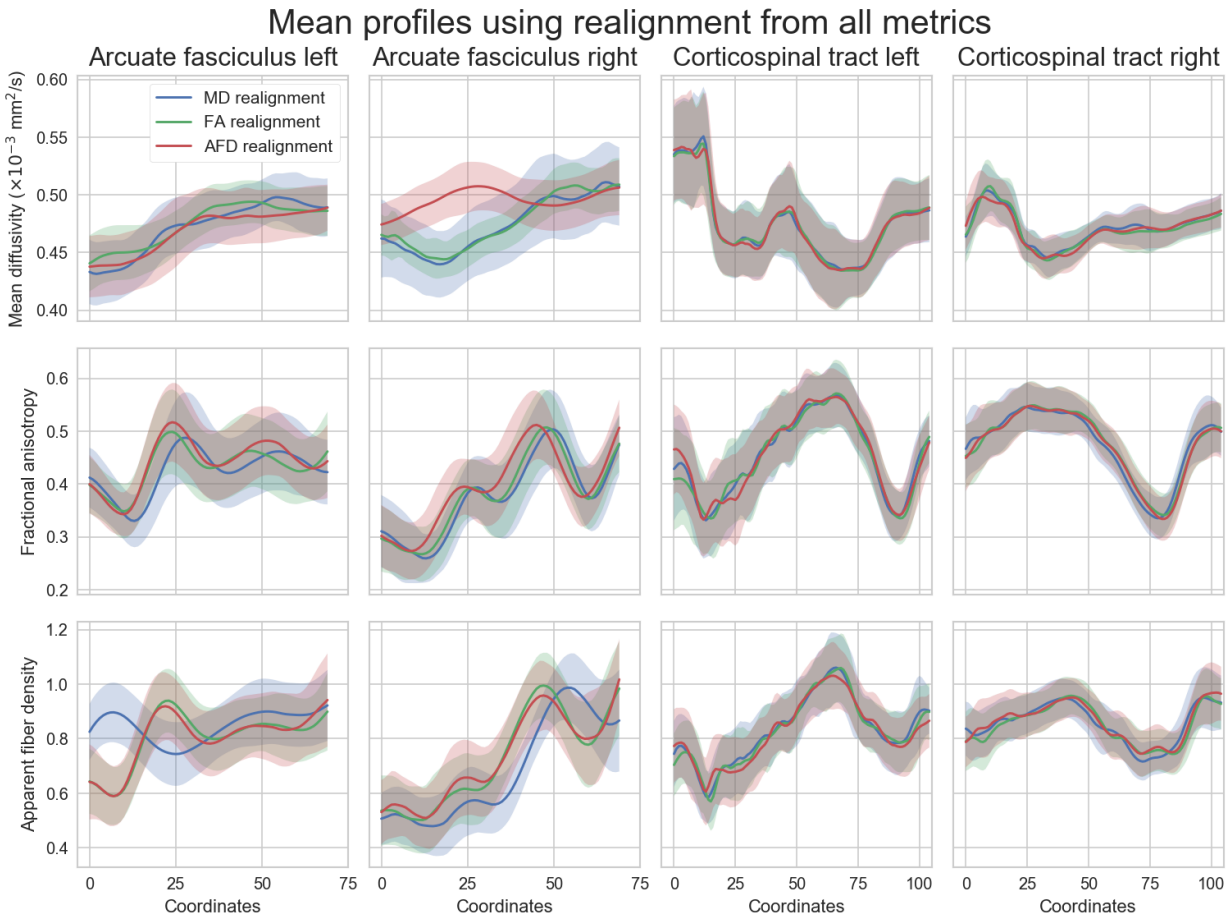

Figure B.20: Along-tract averaged profiles (and standard deviation as the shaded area) of the white matter fiber bundles (columns) from the HCP datasets after realignment for each studied metric (rows). These results are obtained by using a maximally allowed displacement of 20%.

Appendix B.3. Localized alterations of the HCP datasets

Appendix B.3 presents counterpart results to Fig. 14, but instead using a maximally allowed displacement of 5%, 10% or 100% (no limit). Paired t-test (FDR corrected at $\alpha = 0.05$) with focused alterations of the metrics for each bundle of **A)** 25% over 1% of the length, **B)** 50% over 1% of the length, **C)** 25% over 5% of the length and **D)** 50% over 5% of the length are shown. The AF left/right are represented from anterior (coordinate 0) to posterior and the CST left/right from inferior (coordinate 0) to superior. The p-values are on a log scale along the average streamline before realignment (dashed red lines) and after realignment (solid blue lines) with the DPR algorithm. The horizontal dashed black lines indicate p-value = 0.05. In general, alterations of 5% of the bundle can be detected, whereas small changes of 1% here

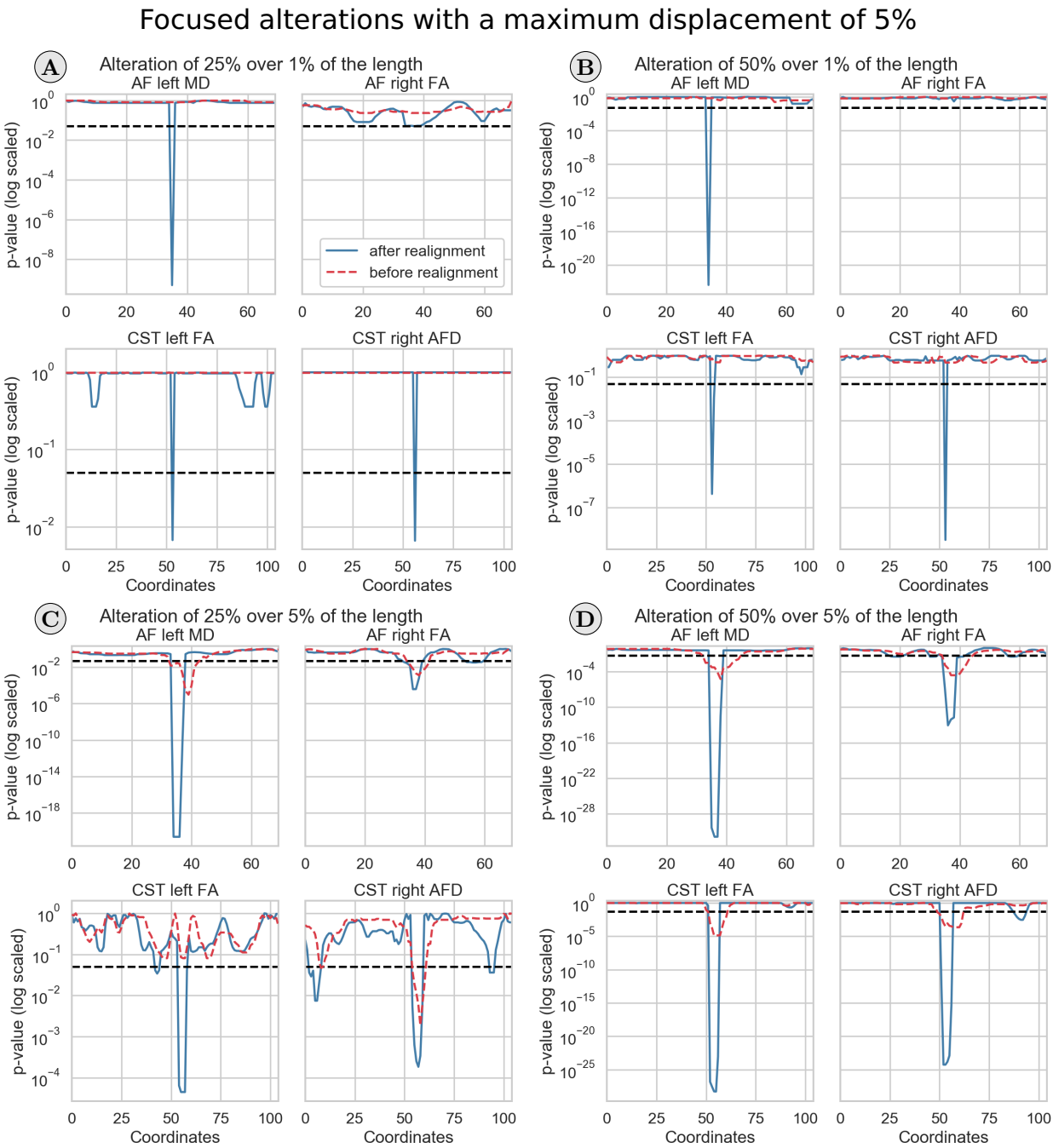

Figure B.21: Paired t-test before and after realignment for the four bundles. These results are obtained by using a maximally allowed displacement of 5%.

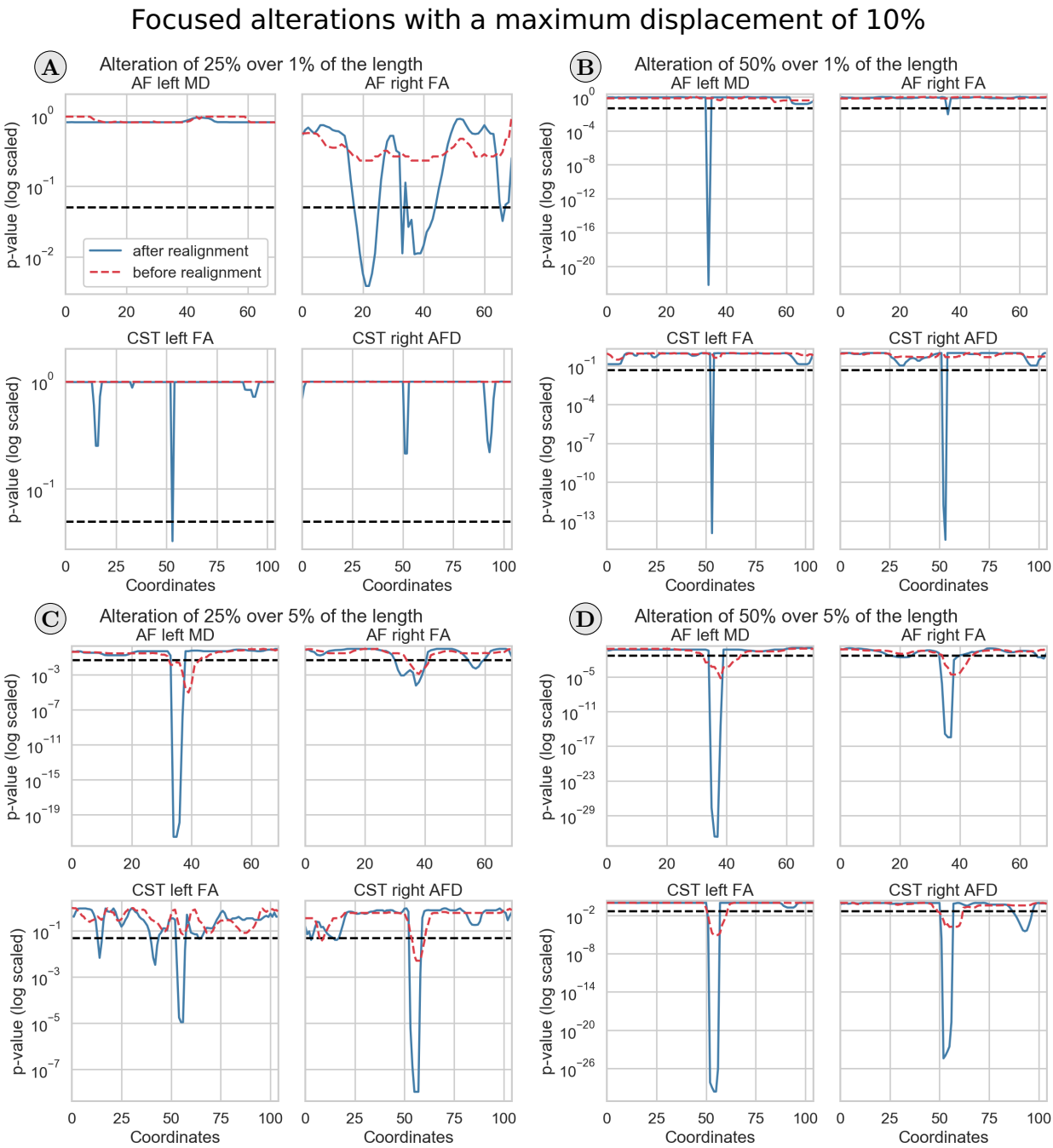

Figure B.22: Paired t-test before and after realignment for the four bundles. These results are obtained by using a maximally allowed displacement of 10%.

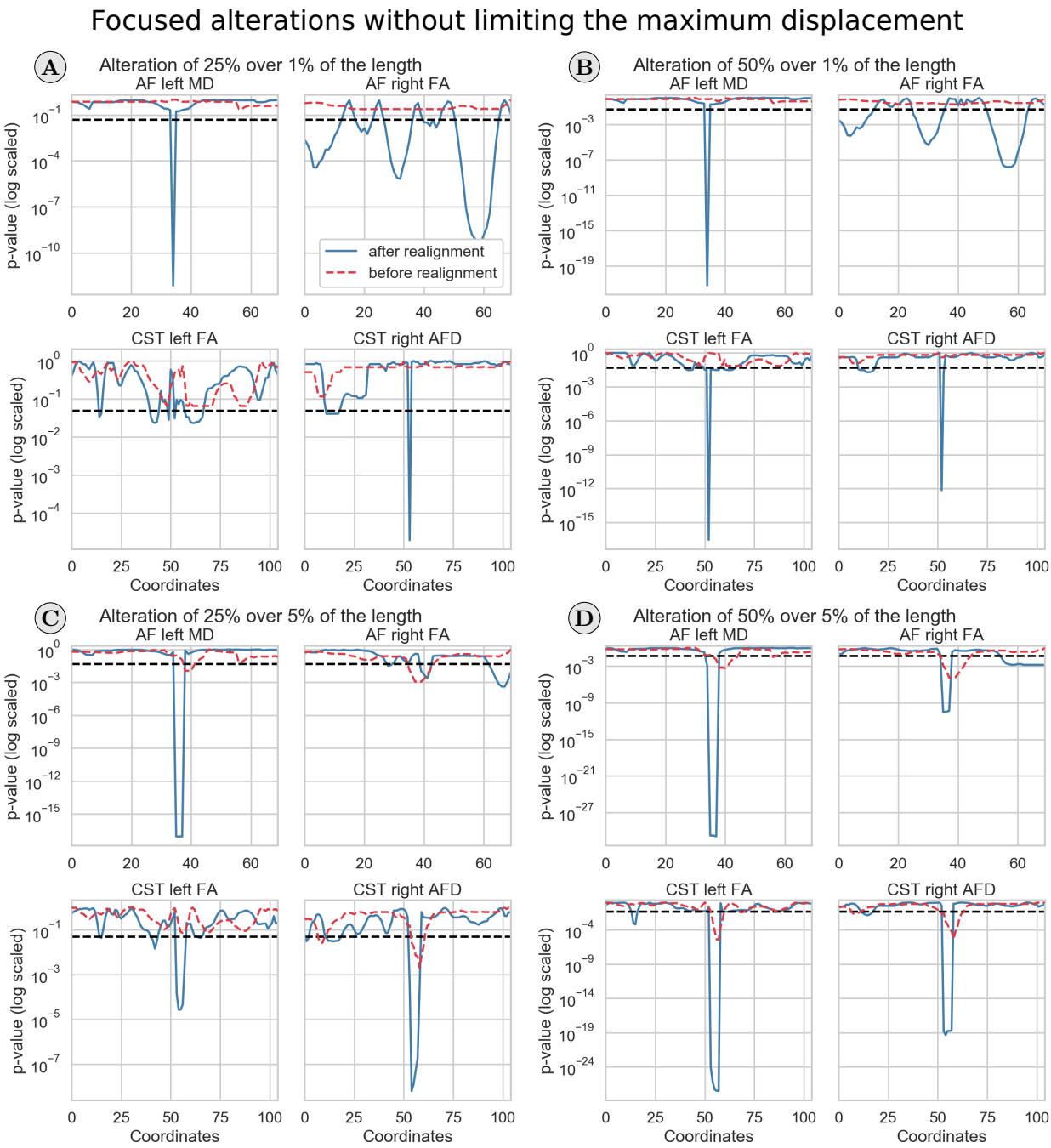

Figure B.23: Paired t-test before and after realignment for the four bundles. These results are obtained without limiting the allowed maximum displacement. This leads to false effects for the AF right bundles, presumably because structural differences, rather than local alterations, are driving the realignment process.