NeuroImage 200 (2019) 89–100

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Dimensionality reduction of diffusion MRI measures for improved tractometry of the human brain

Maxime Chamberland [a,*], Erika P. Raven [a], Sila Genc [a,c], Kate Duffy [a], Maxime Descoteaux [b], Greg D. Parker [a], Chantal M.W. Tax [a], Derek K. Jones [a,d]

[a] Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Cardiff, UK
[b] Sherbrooke Connectivity Imaging Lab, Department of Computer Science, University of Sherbrooke, Sherbrooke, Canada
[c] Developmental Imaging, Murdoch Children's Research Institute, Parkville, Australia
[d] Mary McKillop Institute for Health Research, Australian Catholic University, Victoria, Australia

ARTICLE INFO

Keywords:
Diffusion MRI
DTI
HARDI
Tractography
Tractometry
Dimensionality reduction

ABSTRACT

Various diffusion MRI (dMRI) measures have been proposed for characterising tissue microstructure over the last 15 years. Despite the growing number of experiments using different dMRI measures in assessments of white matter, there has been limited work on: 1) examining their covariance along specific pathways; and on 2) combining these different measures to study tissue microstructure. Indeed, it quickly becomes intractable for existing analysis pipelines to process multiple measurements at each voxel and at each vertex forming a streamline, highlighting the need for new ways to visualise or analyse such high-dimensional data. In a sample of 36 typically developing children aged 8–18 years, we profiled various commonly used dMRI measures across 22 brain pathways. Using a data-reduction approach, we identified two biologically-interpretable components that capture 80% of the variance in these dMRI measures. The first derived component captures properties related to hindrance and restriction in tissue microstructure, while the second component reflects characteristics related to tissue complexity and orientational dispersion. We then demonstrate that the components generated by this approach preserve the biological relevance of the original measurements by showing age-related effects across developmentally sensitive pathways. In summary, our findings demonstrate that dMRI analyses can benefit from dimensionality reduction techniques, to help disentangling the neurobiological underpinnings of white matter organisation.

1. Introduction

The human brain is composed of multiple white matter fibres connecting gray matter areas dedicated to processes such as memory, cognition, language, or consciousness. Diffusion MRI (dMRI) (Basser et al., 1994, 2000; Basser and Jones, 2002; LeBihan et al., 2001) has become the preferred tool to probe the brain's tissue microstructure non-invasively. Measures derived from diffusion tensor imaging (DTI) (Basser et al., 1994) can be obtained at each imaging voxel, including fractional anisotropy (FA) which reflects the degree of diffusion anisotropy (Pierpaoli and Basser, 1996), and mean diffusivity (MD), an indicator of the overall magnitude of diffusion. Based on local estimates of underlying trajectories at every voxel, dMRI is also capable of virtually reconstructing the structural architecture of the brain white matter pathways using tractography (Conturo et al., 1999; Mori and Van Zijl, 2002). The conventional approach to merge the quantitative nature of diffusion measures with the qualitative nature of tractography is to collapse voxel-based measures into a single scalar value per bundle (e.g., by averaging values over all vertices of a streamline; Jones et al. (2006); Kanaan et al. (2006); Jones et al. (2005a)). Individual differences in such summary diffusion-related measures can then be correlated, for example, with individual differences in cognition or behaviour. However, despite its well documented sensitivity, DTI has its limitations (Tournier et al., 2011; Jeurissen et al., 2013). For example, FA and MD lack specificity to the various physical properties of white matter, such as crossing fibres (Jeurissen et al., 2013), axon density and myelination (Beaulieu, 2002; Jones et al., 2013). Moreover, the average profile of those measures may vary along a given pathway depending on the underlying fibre architecture (Vos et al., 2012; Yeatman et al., 2012). Furthermore, only a subset of DTI measures are known to be orthogonal with each other (e.g.,

* Corresponding author.
E-mail address: ChamberlandM@cardiff.ac.uk (M. Chamberland).

https://doi.org/10.1016/j.neuroimage.2019.06.020
Received 8 March 2019; Received in revised form 5 June 2019; Accepted 7 June 2019
Available online 20 June 2019

FA, MD or tensor norm; Ennis and Kindlmann (2006); Kindlmann et al. (2007); De Santis et al. (2014)).

Recent advances in diffusion hardware, acquisition and modelling (Sotiropoulos et al., 2013; Jones et al., 2018; Assaf and Basser, 2005; Tournier et al., 2012; Jeurissen et al., 2014) have been introduced to overcome the limitations of DTI, giving access to previously inaccessible measures. High angular resolution diffusion imaging (HARDI; Tuch et al. (2002)) was originally developed to not only provide new anisotropy measures (Tournier et al., 2011) but also to solve the so-called *crossing fibre* problem, making tractography more robust (Descoteaux, 2015). Multi-shell acquisitions (Wedeen et al., 2005) have also facilitated new ways to link relevant tissue properties to the signal such as CHARMED (Assaf and Basser, 2005), AxCaliber (Assaf et al., 2008), ActiveAx (Alexander et al., 2010), multi-tensor models (Scherrer et al., 2016) and NODDI (Zhang et al., 2012) among others (for review, see Alexander et al. (2017)). In general, such models aim to extract parameters from intra- and extracellular compartments, and to estimate parameters such as axon diameter distributions and other high-order information.

Multi-shell acquisitions have also shown to improve the angular resolution of orientation distribution functions (ODFs) (Descoteaux et al., 2011; Jeurissen et al., 2014; Chamberland et al., 2018). In conjunction, new frameworks such as fixel-based analysis (Raffelt et al., 2012) have been proposed to map fibre-specific measures by looking at the apparent fibre density (AFD), a measure proportional to the underlying fibre density, as opposed to having voxel-specific scalar maps. The combination of frameworks such as along-tract profiling (Jones et al., 2005b; Corouge et al., 2006; Yeatman et al., 2012; De Santis et al., 2014; Colby et al., 2012; Cousineau et al., 2017) and tractometry (e.g., combining multiple measures (Bells et al., 2011)) allows for a comprehensive assessment of white matter microstructure. Both frameworks have the advantage of providing higher sensitivity to microstructural features of fibre pathways by mapping a set of MR-derived measures over white matter bundles. Recently, along-tract profiling has been successfully applied to study normal brain development (Geeraert et al., 2018) and to characterise areas of the brain with abnormal properties in various brain conditions (Dayan et al., 2016; Cousineau et al., 2017; Groeschel et al., 2014).

However, one problem arises with having access to multiple new measurements at each voxel and at each vertex forming a streamline: it quickly becomes intractable for existing analysis pipelines to process such high-dimensional data (a problem often referred to as the *curse of dimensionality*; Bellman (1961)), highlighting the need for new ways to visualise or analyse such data. Moreover, dMRI measures may share overlapping information which can cause redundancies (in the sense of correlation) in data analysis and ultimately decrease statistical power if strictly correcting for Type I errors (Penke et al., 2010; Metzler-Baddeley et al., 2017; Bourbon-Teles et al., 2017). A solution to this problem resides in dimensionality reduction, an established technique that has been successfully applied in the past by the neuroimaging community (for review, see Mwangi et al. (2014)). Despite the growing number of experiments using different microstructural measures in assessments of white matter, there has been limited work on combining these different measures and on examining their covariance along specific pathways.

In this work, we explore the covariance of commonly-derived dMRI measures (De Santis et al., 2014). We propose a data reduction framework that takes advantage of those redundancies and aims to provide a better insight into patterns of associations between DTI and HARDI measures. Specifically, we identified common components that explain the maximal variance in measures profiled along multiple fibre bundles. We demonstrate the utility of our framework by showing enhanced sensitivity to the detection of age-related differences in tissue microstructure across developmentally sensitive pathways compared with the individual dMRI measures. Finally, we provide recommendations for future studies with limited capabilities in terms of data acquisition and processing.

2. Methods

2.1. Participants

This study reports on a sample of typically developing children aged 8–18 years (mean = 12.2 ± 2.8) participating in the Cardiff University Brain Research Imaging Centre (CUBRIC, School of Psychology) Kids study. The study was performed with ethics approval from the internal ethics review board and informed consent was provided from the primary caregiver of children enrolled in the study. Exclusion criteria included previous history of a neurological condition or epilepsy.

2.2. Data acquisition

Data from thirty-six (n = 36, 13 male) children were acquired using a multi-shell HARDI protocol on a Siemens 3T Connectom system with maximum gradient amplitude = 300 mT/m. The acquisition protocol consisted of 14 b_0 images, 30 diffusion directions at b = 500, 1200 s/mm^2 and 60 diffusion directions at b = 2400, 4000, 6000 s/mm^2 with 2 × 2 × 2 mm^3 voxels (TE/TR: 59/3000 ms, δ/Δ: 7.0/23.3 ms).

2.3. Data pre-processing

Data quality assurance was performed on the raw diffusion volumes using slicewise outlier detection (SOLID; Sairanen et al. (2018)). Each dataset was then denoised in MRtrix (Veraart et al., 2016) and corrected for signal drift (Vos et al., 2017), subject motion (Andersson and Sotiropoulos, 2016), field distortion (Andersson et al., 2003), gradient non-linearities (Glasser et al., 2013; Suryanarayana et al., 2018) and Gibbs ringing artefacts (Kellner et al., 2016).

2.4. Local representation

Multi-shell multi-tissue constrained spherical deconvolution (MSMT-CSD; Jeurissen et al. (2014)) was applied to the pre-processed images to obtain voxel-wise estimates of fibre ODFs (fODFs; Tournier et al. (2004, 2007); Seunarine and Alexander (2009); Descoteaux et al. (2009)) with maximal spherical harmonics order l_{max} = 8. The fODFs were generated using a set of 3-tissue group-averaged response functions (Dhollander et al., 2016) followed by joint bias field and image intensity normalisation in MRtrix (Tournier et al., 2012), enabling the direct comparison of fODF amplitudes across subjects (Raffelt et al., 2012). Diffusion tensors were also generated using linearly weighted least squares estimation (for b < 1200 s/mm^2 data) providing the following quantitative scalar measures: FA, axial diffusivity (AD), radial diffusivity (RD), MD, geodesic anisotropy (GA; Fletcher et al. (2004)) and tensor mode representing the shape of the tensor (Kindlmann et al., 2007). In addition, HARDI measures were extracted from the fODFs of each subject. Those measures include fibre-specific AFD (Raffelt et al., 2012) for the bundles described in the next section, AFD_{tot} (spherical harmonics $l = 0$) and the Number of Fibre Orientations (NuFO) based on the number of local fODF peaks (Dell'Acqua et al., 2013). Finally, restricted signal fraction maps (FR, adapted from CHARMED to remove potential isotropic partial volume contamination; Assaf and Basser (2005)) were also computed using the fODFs peaks to initialise and regularise model-fitting. To summarise, ten dMRI measures related to tissue microstructure ($m = 10$) were generated for each subject.

2.5. Tractography and tractometry

Whole-brain streamline tractography was performed using FiberNavigator (Chamberland et al., 2014) using 8 seeds/voxel evenly distributed across the whole brain (approximating 1.8 M seeds), a minimum fODF amplitude of 0.1, a 1 mm step size (i.e. 0.5 × voxel size), a 45° maximum curvature angle and streamlines whose lengths were outside a range of 20 mm–300 mm were discarded. Twenty-two bundles of interest ($t = 22$) were then interactively dissected in the native space of each subject using a

combination of include and exclude regions of interests (ROIs). Anatomical definitions and ROIs used to delineate each pathway are listed in the Supplementary Materials. The virtual dissection plan included:

Commissural bundles: anterior commissure (AC), body of the corpus callossum (CC), forceps minor (Genu), forceps major (Splenium).
Association bundles (bilateral): arcuate fasciculus (AF), cingulum (Cg), inferior fronto-occipital fasciculus (iFOF), inferior longitudinal fasciculus (ILF), optic radiations (OR), superior longitudinal fasciculus (SLF), uncinate fasciculus (UF).
Projection bundles (bilateral): corticospinal tract (CST), frontal aslant tract (FAT).

At this stage, we examined the covariance of the averaged diffusion measures for all bundles using Pearson's correlation (r). Next, along-tract profiling was performed for each bundle using the Python toolbox developed by Cousineau et al. (2017), combined with DIPY functions (Garyfallidis et al., 2014). Bundles were first pruned to remove outliers as in (Cousineau et al., 2017; Garyfallidis et al., 2018). If necessary, the order in which the vertex-wise measures were stored was reversed, to ensure consistency across the subjects profiles (i.e., from left-to-right for commissural bundles, from inferior-to-superior for projection bundles and from posterior-to-anterior for association bundles). A representative core streamline was generated for each bundle (i.e., mean streamline of the pathway) and was subsequently resampled to $s = 20$ equidistant segments. Then, every vertex of every streamline forming the pathway was assigned to its closest segment along the core. The measure values of each vertex were then projected and averaged along each segment of the pathway, weighted by their geodesic distance from the core (Cousineau et al., 2017). An along-tract profile was finally generated for every combination of measure and pathway.

2.6. Dimensionality reduction

Each dataset comprised $m = 10$ dMRI-derived measures mapped along 440 white matter regions ($t = 22$ bundles $\times$ $s = 20$ segments). To explore the possible redundancy (in the context of data reduction) and complementarity of each measure, a principal component analysis (PCA) was performed on the concatenated set of profiles across subjects and bundles (Table 1) using the *tidy* data standard (Wickham et al., 2014). Performing PCA over tract segments rather than over all voxels alleviates the need to register each diffusion measure to a common space. PCA reduces data dimensionality by extracting principal components that reflect relevant features in the data (Jolliffe, 2002; Abdi and Williams, 2010). The benefit is that a significant proportion of the variance in the data can be explained by a reduced number of orthogonal components, compared to the total number of raw input variables. PCA was performed by singular value decomposition of the z-transformed tract profiles via the *prcomp* function in R (R Core Team, 2018). Here, the goal was to end up with the minimum number of components that summarise the maximum amount of information contained in the original set of diffusion measures. However, in order to avoid instability around the component loadings that comprise the principal components (Garg and Tai, 2013), measures showing significant covariance were discarded based on their correlation scores ($|r| > 0.8$) and the PCA was re-computed. Finally, the minimal number of principal components that accounted for the most variability was selected based on: 1) their interpretability (Metzler-Baddeley et al., 2017); and 2) the inspection of scree plots (Cattell, 1966) to select ranked components with an eigenvalue > 1.

Table 1
Data structure input for PCA. Individual subjects (n = 36), bundles ($t = 22$) and segments ($s = 20$) are concatenated to form observations while variables represent the measures ($m = 10$) derived from dMRI.

Subject	Bundle	Section	FA	AD	…	FR
S_1	$Bundle_1$	$Section_1$	FA_{111}	AD_{111}	…	FR_{111}
S_2	$Bundle_1$	$Section_1$	FA_{211}	AD_{211}	…	FR_{211}
⋮	⋮	⋮	⋮	⋮	…	⋮
S_1	$Bundle_1$	$Section_2$	FA_{112}	AD_{112}	…	FR_{112}
⋮	⋮	⋮	⋮	⋮	⋱	⋮
S_n	$Bundle_b$	$Section_s$	FA_{nbs}	AD_{nbs}	…	FR_{nbs}

2.7. Statistical analysis

PCA results were tested for sampling adequacy using a Kaiser-Meyer-Olkin (KMO; Dziuban and Shirkey (1974)) test followed by Bartlett's test of sphericity to test whether the covariance matrix is significantly different from identity. We then ran an exploratory linear regression analysis to see whether profiles extracted from the PCA can provide increased sensitivity in the detection of age-related differences in tissue microstructure (as opposed to using the full set of $m = 10$ measures). It is important to recall that PCA results are always orthogonal, and therefore are statistically independent of one another. To address the multiple comparisons problem, a strict Bonferroni correction was applied to all linear models whereby statistical significance was defined as: $p < 0.05/(m$ measures $\times$ t bundles $\times$ s segments) resulting in $p < 1.14e\text{-}5$ for the ten raw measures, and $p < 5.68e\text{-}5$ for the first two principal components. All statistical analyses were carried out using R v3.5.1 (R Core Team, 2018) and RStudio v1.1.456 (RStudio Team, 2016).

3. Results

3.1. Measures covariance and profiling

The entire set of bundles and measures was successfully reconstructed in all subjects. Fig. 1 shows the relationship between the various input measures averaged on different white matter pathways using a cross-correlation matrix representation. Matrices are re-organised using hierarchical clustering (Murtagh and Legendre, 2014), placing higher correlations closer to the diagonal in order to regroup measures that have similar correlations together. In general, the measures form two or three clusters across all bundles. Across the whole set of bundles (Fig. 1, middle), we observe a first cluster of positive correlations ($r > 0.5$) between AD, FA, GA, Mode, AFD, AFD_{tot} and FR measures. A second cluster of positive correlations is formed of MD, RD and NuFO. The group-averaged diffusion measures of each bundle are reported in Suppl. Table 1.

However, important details about the spatial heterogeneity of the various input measures of interest appear when profiled along pathways. For example, FA, AFD and FR values all get progressively smaller along the CST as they approach the cortex (Figs. 2 and 3). Furthermore, the high number of fibre crossings near the centrum semiovale is reflected by a high NuFO index and is also marked by a low FA (Fig. 2). As might be expected, HARDI-derived measures such as FR and AFD_{tot} seem to be less affected by the intra-voxel orientational heterogeneity of crossing regions than the tensor-based measures like FA, AD and RD. The correlation matrix in Fig. 3 also highlights the similarity between the various microstructural profiles, indicating potential overlap in the amount of information conveyed by the dMRI measures.

3.2. Principal component analysis

The loading vector plots in Fig. 4 show association patterns between the various input measures. The left panel shows PCA results performed on the entire set of measures. If two vectors subtend a small angle to each other, the two variables they represent are strongly correlated. When such vectors were found to be close, the one showing higher correlations with any other measures was removed (see Section 2.6). In line with the aforementioned results, shared covariance is observed between AD and tensor mode ($r = 0.8$), as well as between FA and GA ($r = 0.95$). After pruning up measures for multicollinearity (Fig. 4, right panel), PCA results show that 80% of the variability in the data is accounted by the first two principal components (KMO: 0.64, sphericity: $p < 2.2e\text{-}16$). As

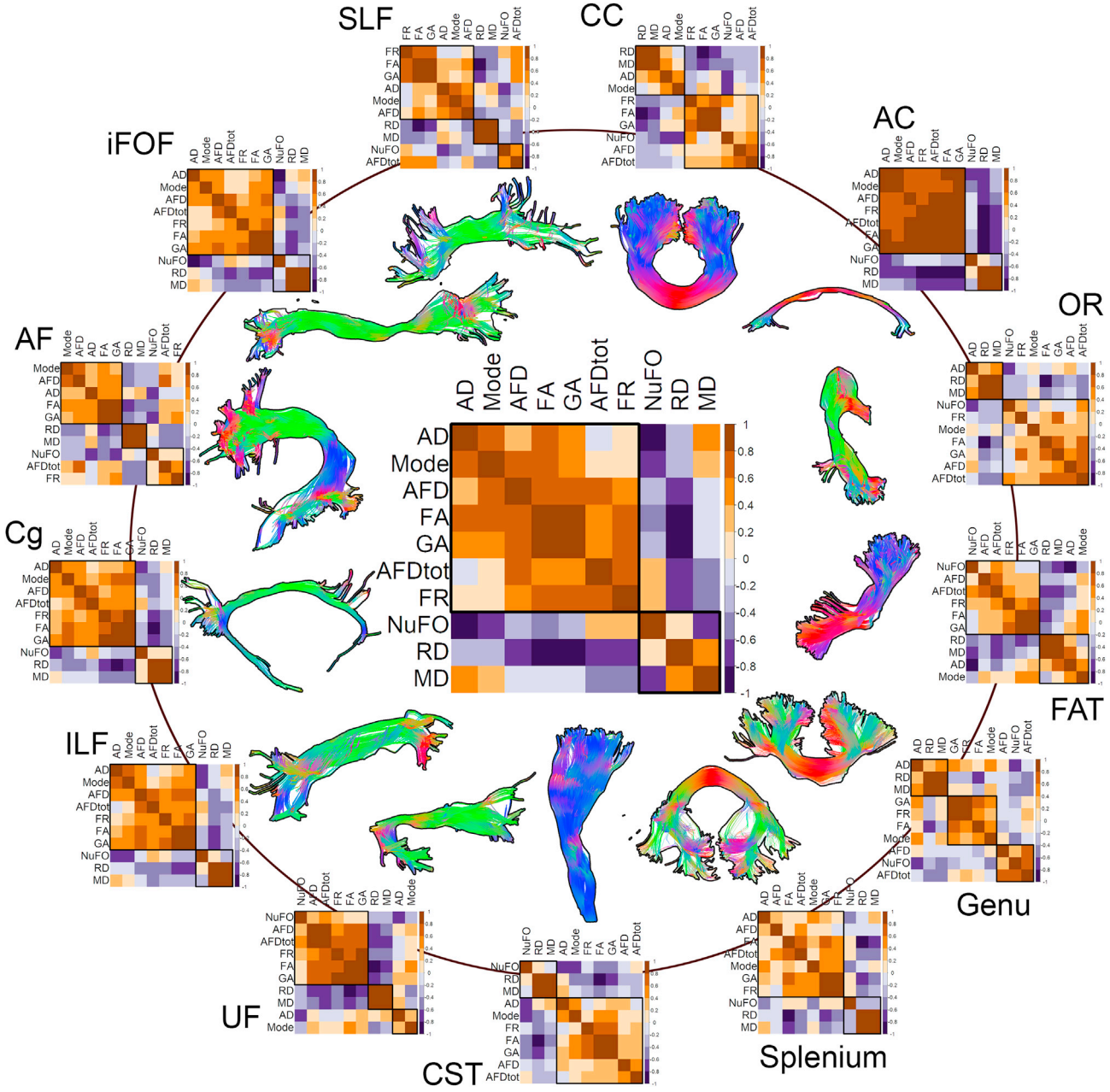

Fig. 1. Correlation matrices of the ten diffusion measures, group-averaged for each extracted bundles. The middle image represents the average of all white matter bundles. Matrices are re-organised using hierarchical clustering, grouping measures that have similar correlations together. A first cluster of positive correlations ($r > 0.5$) is observed between most of the bundles for measures like AD, FA, GA, AFD, AFD_{tot}, Mode and FR. A second set of positively correlated measures (NuFO, MD, RD) forms the second cluster. Note that for bilateral pathways, the left and right values were combined prior performing the correlation.

shown in Fig. 5, the PC that explains the largest proportion of the variance (PC1, 48%, $\lambda = 3.4$) is composed of hindrance-sensitive measures (Dell'Acqua et al., 2013) with AFD, FR and AFD_{tot} contributing positively (24%, 21% and 16%, respectively) and AD contributing negatively (25%). The second PC (PC2) represents 32% of the variance in the data ($\lambda = 2.2$) and is mostly driven by orientational dispersion and complexity-sensitive measurements, with its largest positive contribution from NuFO (34%), and negative contributions from AD (26%) and MD (25%) (Fig. 5, PC2).

3.3. Detecting differences in tissue microstructure

3.3.1. Bundle averages

To asses the relevance of the two principal components, we first report developmental changes in white matter tissue microstructure using bundle-averaged measures ($m = 10$, $s = 1$) and PCA components ($m = 2$, $s = 1$). Significance thresholds were Bonferroni-corrected to account for multiple comparisons ($p < 2.27e\text{-}4$ for the ten raw measures, and $p < 1.14e\text{-}3$ for the two principal components). Fig. 6 shows a significant increase in PC1 as a function of age for the left iFOF and CST, whereas no correlation with age was observed between individual hindrance-related measures. Significant positive correlations were found between PC1 (*restriction-related* component) and age in the following association pathways: AF (left: R^2: 0.34, $p = 1.06e\text{-}4$), iFOF (left: R^2: 0.31, $p = 2.51e\text{-}4$), FAT (left: R^2: 0.43, $p = 9.06e\text{-}6$, right: R^2: 0.43, $p = 7.77e\text{-}6$), UF (right: R^2: 0.26, $p = 9.76e\text{-}4$) and motor pathways: CST (left: R^2: 0.40, $p = 1.92e\text{-}5$, right: R^2: 0.40, $p = 2.0e\text{-}5$), CC (R^2: 0.29, $p = 4.51e\text{-}4$). One significant positive correlation between PC2 (*complexity-related*

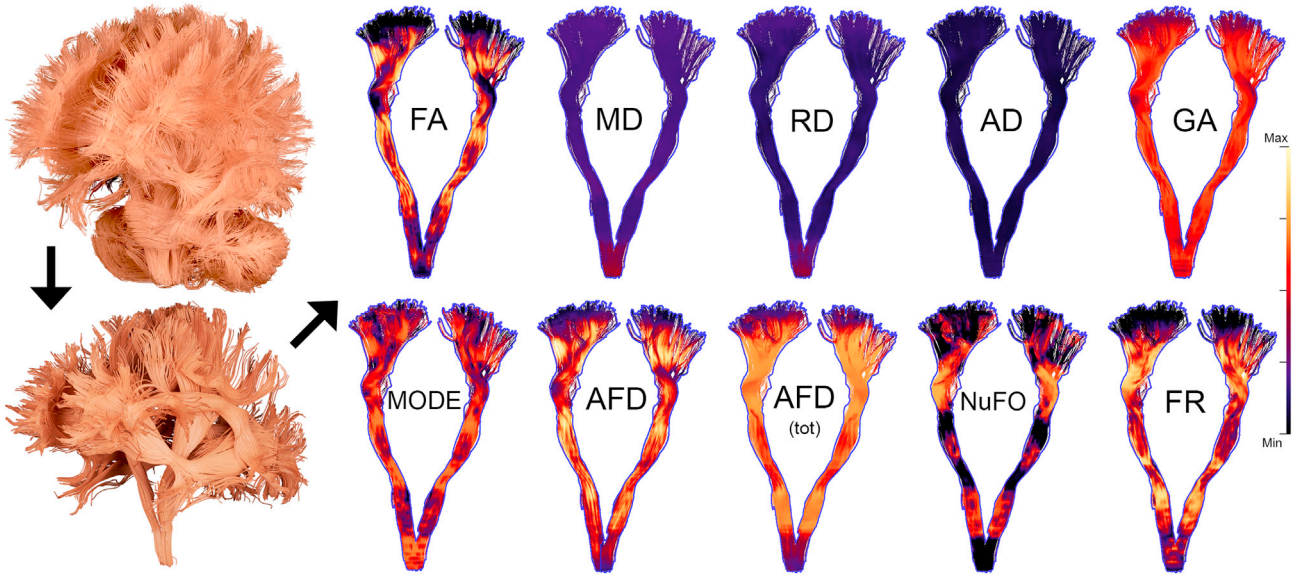

Fig. 2. Overview of the ten input measures overlaid on the CST of a representative subject. Whole-brain tractograms (top-left) were manually dissected into $t = 22$ bundles (bottom-left) and measures were subsequently mapped along each pathway, providing information about their spatial heterogeneity.

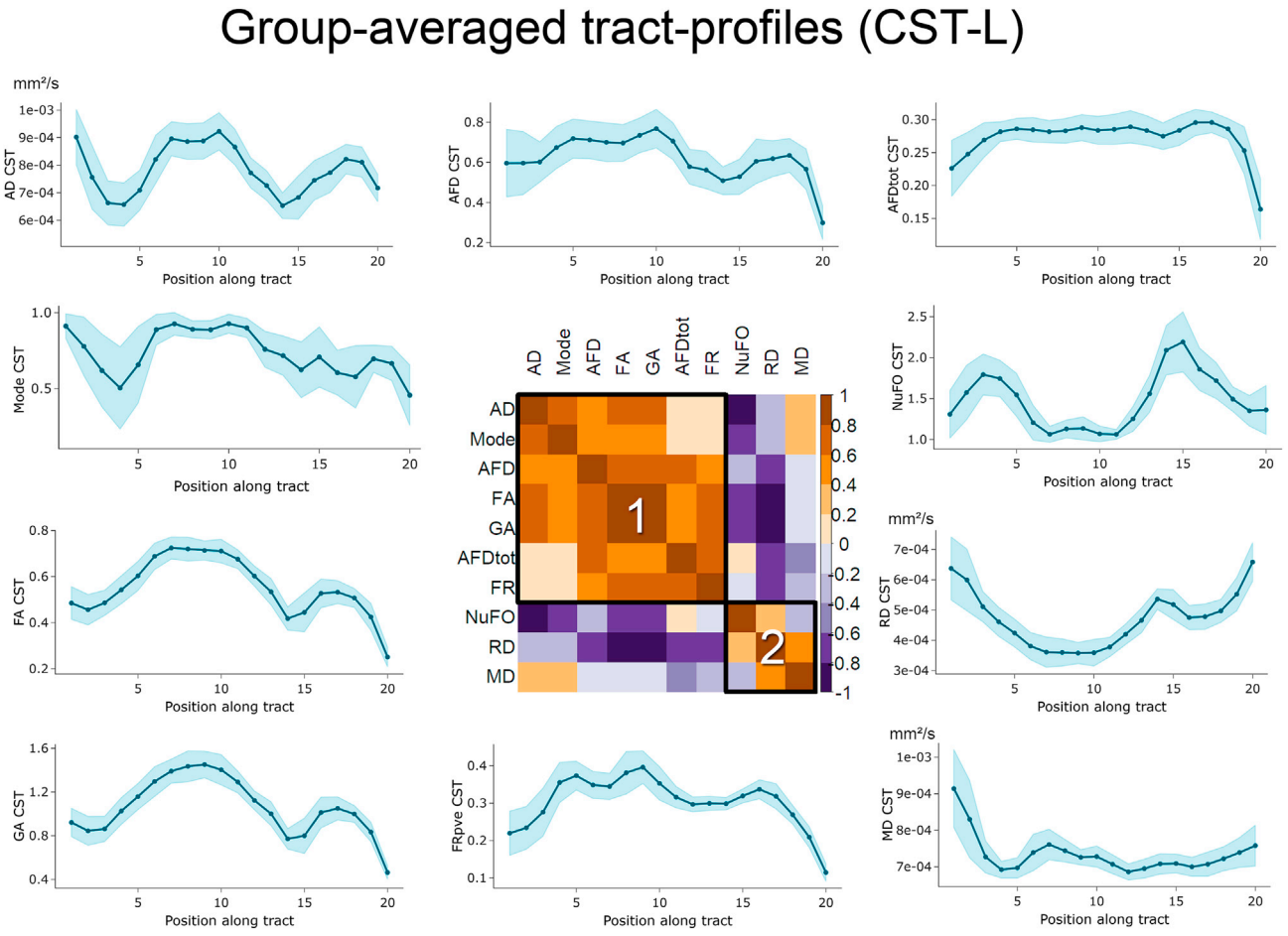

Fig. 3. Group-average profiling of the ten input diffusion measures along the left CST for $s = 20$ segments, spanning from the brainstem ($s = 1$) to the cortex ($s = 20$). Heterogeneity in the profiles along the tract highlights the need for a vertex-wise assessment of the measures. Similarity between profiles also shows shared covariance between the measures, indicated by the two clusters (1 and 2) on the correlation matrix (sorting: hierarchical clustering). Shaded tract-profile area: ±1 standard deviation.

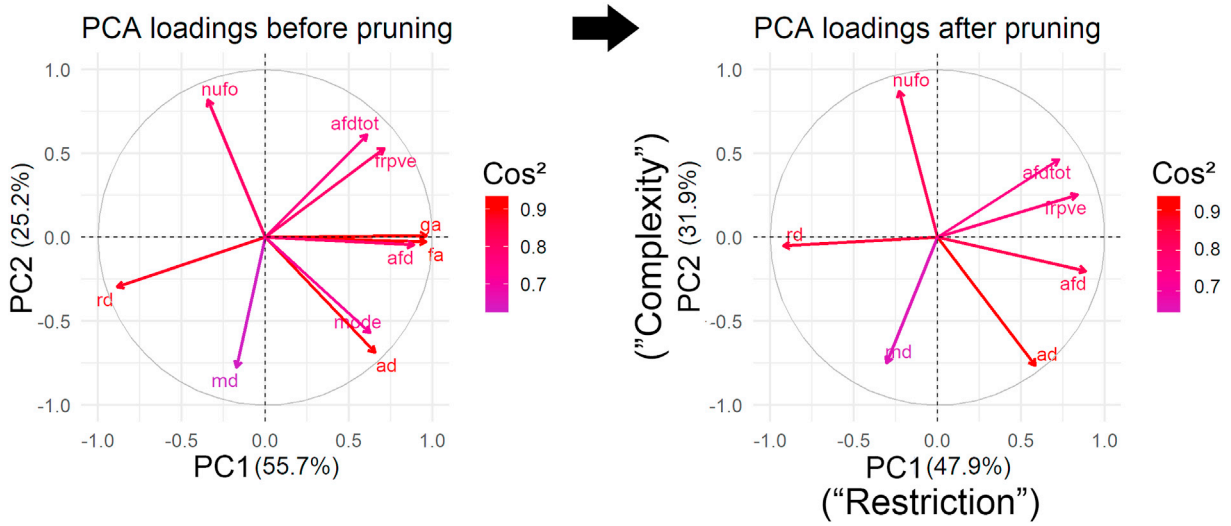

Fig. 4. PCA results before (left) and after (right) multicollinearity analysis. To improve stability around PC1, AFD was kept over GA and FA due to its fibre specificity properties. Tensor mode was also discarded based on its collinearity with AD. On the right PCA, one can observe separation between the various measures, generating a hindrance-related component (PC1) that loads on AFD, AFD_{tot}, RD and FR and a complexity-related component (PC2) that loads on NuFO, AD and MD. Here, the squared cosine notation (cos^2) shows the importance of a measure for a given PC. A high cos^2 indicates a good representation of the measure for a given principal component. Figure generated using the *FactoExtra* package (Kassambara and Mundt, 2017) in R.

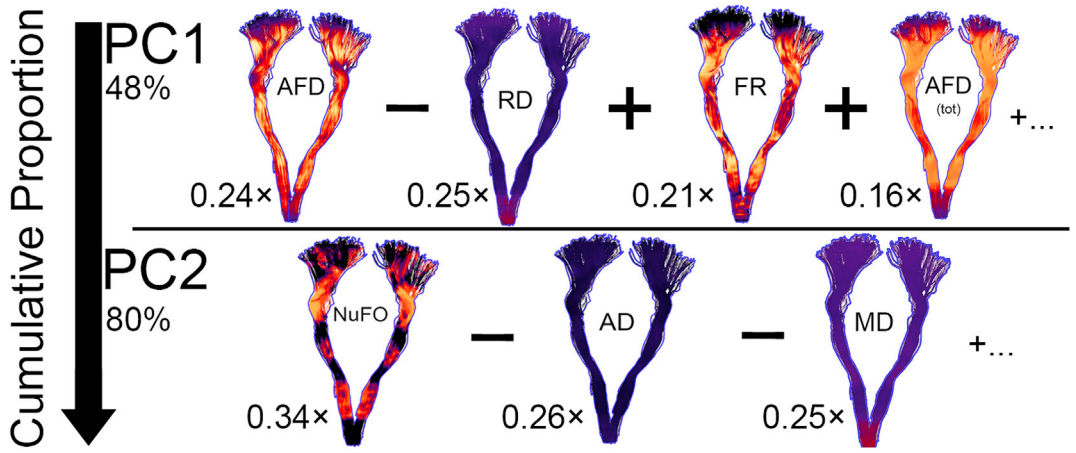

Fig. 5. Visual overview of PC1 and PC2 with the contribution of each measure to that component. The first PC captures most of the hindrance- and restriction-related measures (AFD, RD, FR, AFD_{tot}). The second PC mostly represents tissue complexity and orientational dispersion properties associated with NuFO, AD and MD. The first two components accounted for 80% of the variance in the diffusion measures.

component) and age was found in the SLF (right: R^2: 0.27, p = 6.82e-4).

No significant age relationships were found between any of the bundles and FA, GA, Mode, AFD, AFD_{tot}, FR and NuFO. Significant negative correlations with age were found in RD for the AF (left: R^2: 0.37, p = 5.40e-5, right: R^2: 0.34, p = 1.02e-4), Cg (left: R^2: 0.33, p = 1.45e-4, right: R^2: 0.47, p = 2.22e-6), CST (left: R^2: 0.35, p = 7.36e-5, right: R^2: 0.33, p = 1.28e-4), FAT (left: R^2: 0.37, p = 4.65e-5, right: R^2: 0.28, p = 4.67e-4). Significant negative correlations with age were found in MD for the FAT (left: R^2: 0.33, p = 1.22e-4, right: R^2: 0.31, p = 2.18e-4), AF (left: R^2: 0.39, p = 2.48e-5, right: R^2: 0.36, p = 6.28e-5), CST (left: R^2: 0.40, p = 2.06e-5, right: R^2: 0.42, p = 1.12e-5), SLF (right: R^2: 0.38, p = 3.21e-5) and Cg (right: R^2: 0.37, p = 4.06e-5). One significant negative correlation was found between AD and age in the SLF (right: R^2: 0.41, p = 1.80e-5).

3.3.2. *Along-tract profiling*

Here, we report on developmental changes in tissue microstructure seen with along-tract profiling (see Fig. 7). Table 2 reports the measures and tract segment mapped along different pathways where significant correlation with age was observed. Significant positive correlations were found between PC1 (*restriction-related* component) and age near the motor cortex area for the CST_{20} (right: R^2: 0.37, p = 4.91e-5) and CC_3 (R^2: 0.37, p = 4.62e-5). Significant age-related positive correlations with PC2 were observed for motor-related pathways: CC_2 (R^2: 0.37, p = 5.63e-5), CST_{20} (right: R^2: 0.49, p = 1.41e-6), CST_{19} (left: R^2: 0.43, p = 8.59e-6) and association pathways: FAT_{19} (right: R^2: 0.38, p = 4.08e-5), $iFOF_{19}$ (right: R^2: 0.39, p = 2.89e-5), ILF_1 (right: R^2: 0.35, p = 9.26e-5), $UF_{3,4}$ (right: R^2: 0.43 & 0.39, p = 7.71e-6 & 3.0e-5) and $SLF_{1,2}$ (R^2: 0.46 & 0.42, p = 3.41e-6 & 1.14e-5). Suppl. Fig. 2 shows that the PCA profiles preserve the spatial heterogeneity of the input diffusion measures along the CST.

No significant age relationships were found in any bundles for FA, GA, Mode, AD and AFD. Significant negative correlations were observed for RD in the FAT_8 (right: R^2: 0.42, p = 1.08e-5) and for MD in the AF_{16} (right: R^2: 0.44, p = 5.62e-6) and Cg_7 (left: R^2: 0.43, p = 6.5e-6). Significant positive correlations were also found for AFD_{tot} in the CST_{20} (left: R^2: 0.50, p = 9.02e-7, right: R^2: 0.53, p = 2.73e-7) and CST_1 (left: R^2: 0.46, p = 2.88e-6). For NuFO, significant age-related differences in tissue

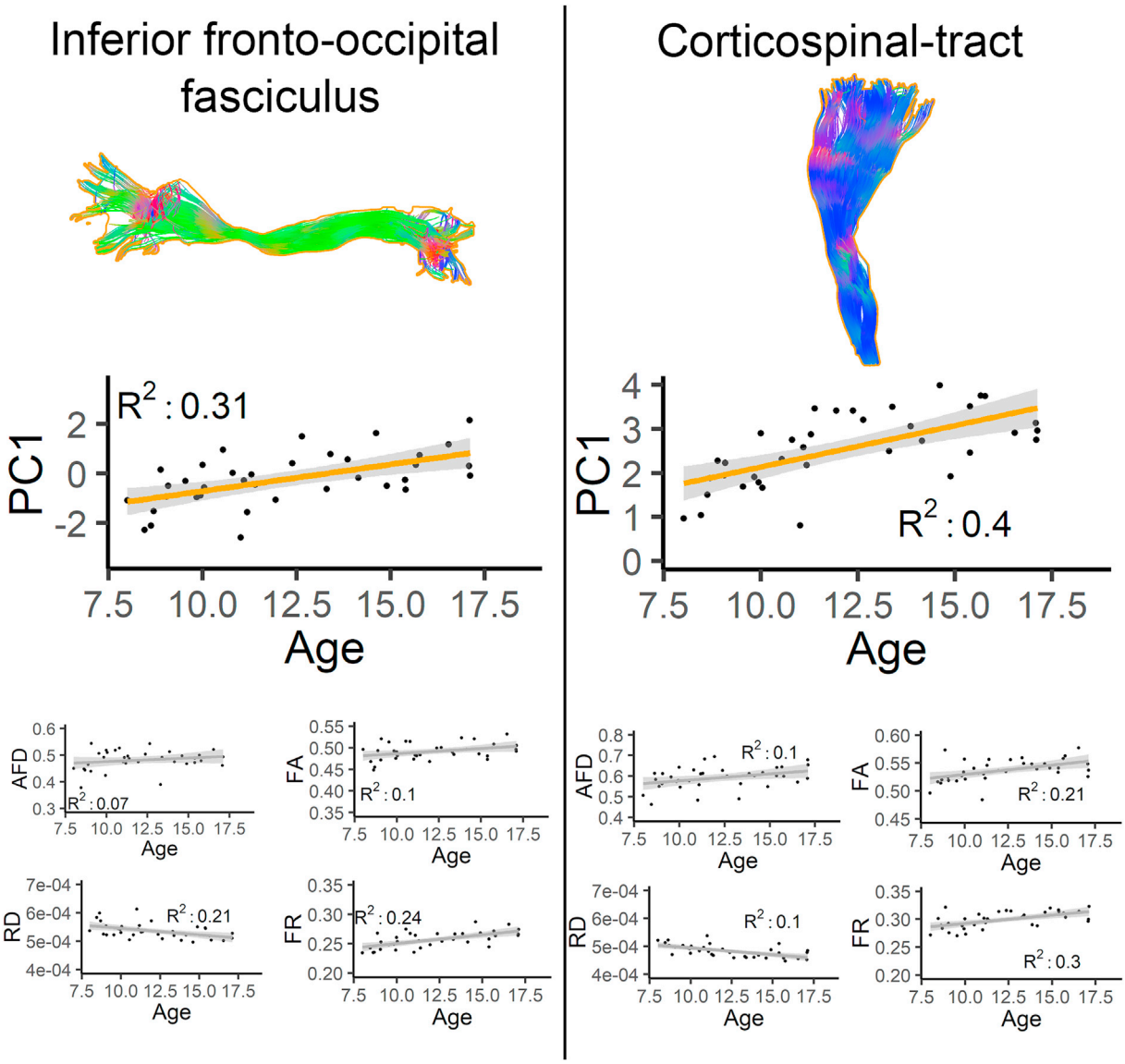

Fig. 6. Developmental changes in iFOF and CST bundles. PC1 show significant positive correlation with age, whereas no correlation was observed between the individual hindrance-sensitive measures.

complexity were observed in the CST_{20} (right: R^2: 0.54, p = 1.56e-7) and CC_3 (R^2: 0.43, p = 7.09e-6). Finally, one significant positive correlation in FR was found for the FAT_6 (right: R^2: 0.57, p = 4.9e-8).

4. Discussion

4.1. Extraction of interpretable components

The aim of this study was to systematically examine any potential covariance between various diffusion measures mapped along white matter fibre bundles extracted from a cohort of typically-developing children and adolescents. We first examined the covariance of the measures averaged over different bundles, which revealed two clusters of inter-dependent measures. The first cluster revealed that measures sensitive to restricted diffusion shared high correlations with each other. Similarly, measures which are known to be sensitive to local complexity or orientational-dispersion co-varied. When profiled along pathways, measures showed heterogeneity across the trajectory of diverse pathways, but their interdependency remained marked by a two-cluster formation. This provided motivation for the next step of our analysis, where we performed a PCA to collapse the inter-dependent measures into the principal modes of variance. This was done by profiling multiple brain fibre systems based on their dMRI features, and then deriving a set of principal components that best represent those individual measures. We then showed the sensitivity of these new components to the detection of differences in tissue microstructure of white matter pathways by exploring their relationship with the age of participants.

A common problem with PCA is that the interpretation of the resulting components can be challenging. Here, the principal components loaded onto variables that shared similarities in their sensitivity to different tissue properties, making the interpretation of the resulting components more meaningful. Measures accounting for the largest percentage of variance in the data (forming PC1) are those known to be most sensitive to hindrance or restriction in the signal, including RD, AFD and FR. In contrast, PC2 features measures that could reflect complexity or orientational-dispersion in the signal, such as NuFO, AD and MD.

The raw tract-averaged diffusion measures showed significant negative correlations in MD and RD with age across a range of developmentally sensitive tracts, which is in line with previous studies that also report a decrease in MD and RD with age, whereas FA shows slower increases with age in late childhood (Lebel et al., 2008). Additionally, we observed significant age relationships for PC1 and PC2 in

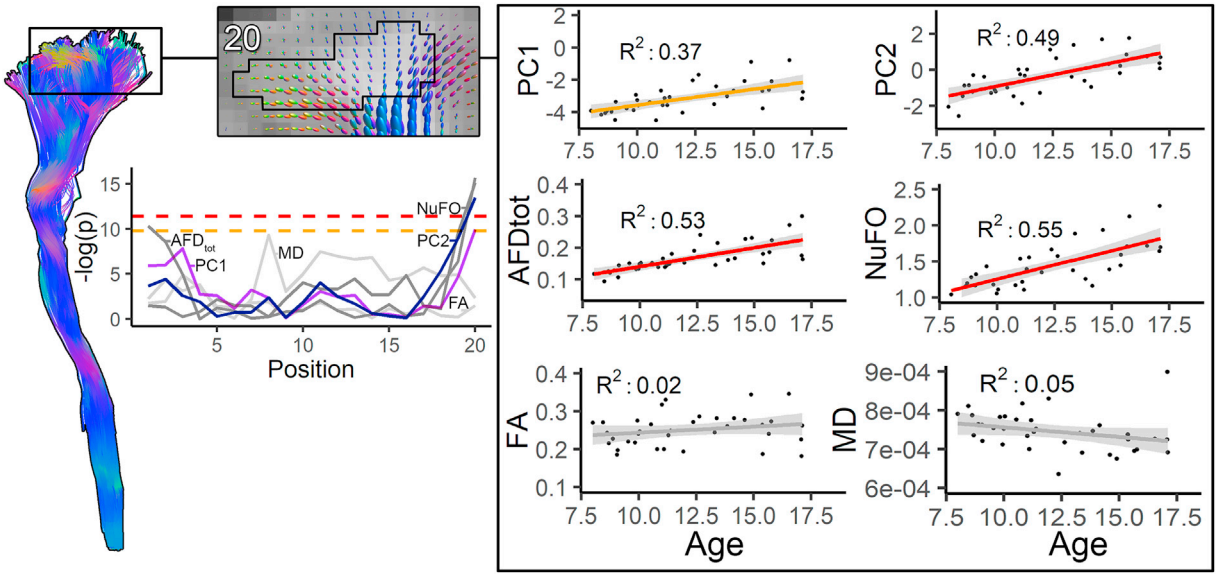

Fig. 7. Age relationships captured by PC1 and PC2 over the left CST. Highlighted section of an axial slice overlaid with fODFs reconstruction of a representative participant (top left) shows the contoured area (black line) where streamlines terminated to form segment 20. At the group level, significant positive correlations with age were found with PC1 and PC2 (top right). Significant positive correlations were also found for HARDI measures AFD_{tot} and NuFO (middle right). No significant correlations were observed for any of the DTI measures (bottom right). Profile plots indicate where significant differences in tissue microstructure were located along the CST (-log(p) scale).

Table 2
Segments of white matter bundles where significant correlation between diffusion measures and age was observed. Subscript ordering for along-tract positions: left ($s = 1$) to right ($s = 20$) for commissural bundles, inferior ($s = 1$) to superior ($s = 20$) for projections bundles and posterior ($s = 1$) to anterior ($s = 20$) for associations bundles. Positive and negative correlations are indicated by increasing (↗) and decreasing (↗) arrows, respectively. Significance thresholds for the measures and components were set as $p < 1.13e\text{-}5$ and $p < 5.68e\text{-}5$, respectively (adjusted $R^2 > 0.3$).

Individual diffusion measures					PCA	
RD	MD	AFD_{tot}	NuFO	FR	PC1	PC2
↘ r-FAT_8	↘ r-AF_{16}	↗ l-$CST_{1,20}$	↗ CC_1	↗ r-FAT_6	↗CC_3	↗CC_2
	↘ l-Cg_7	↗ r-CST_{20}	↗r-CST_{20}		↗r-CST_{20}	↗r-CST_{20}
						↗ l-CST_{19}
						↗ r-FAT_{19}
						↗ r-$iFOF_{19}$
						↗ r-ILF_1
						↗ r-$UF_{3,4}$
						↗ r-$SLF_{1,2}$

developmental-sensitive tracts involved in language (i.e., AF, SLF, FAT, iFOF, ILF) and motor functions (i.e., CST and CC), which is also in line with previous reports in the literature (Genc et al., 2017, 2018, 2018; Lebel et al., 2017; Geeraert et al., 2018). Examining the tract-averaged values for PC2, only one significant correlation with age was found for the right SLF. However, additional significant correlation between age and PC2 were observed when performing along-tract profiling. A potential explanation for this findings resides in the nature of PC2, with most of its contributions coming from AD, MD and NuFO. Indeed, since PC2 reflects the local complexity at each voxel, measures like NuFO will vary depending on the underlying structural architecture and therefore taking the average value across the bundle may lead to a summary statistic that is hard to interpret. In contrast, the values of PC1 (or AFD) remain relatively constant over bundles and thus are less impacted by calculating the bundle average. This may also explain why the first principal component derived from bundle-averages was found to show significant correlations with age in white matter bundles, whereas the original DTI measures did not. We also note that sex differences were not accounted for due to our relatively small sample size but may play a role in some of the differences we observed across pathways (Seunarine et al., 2016). Moreover, studies of larger size would be better powered to test whether the resulting PCs are linked with other factors over and above the participant's age, such as sex, IQ and cognition. Yet, our results highlight the sensitivity of PC1 and PC2 as a composite measure by 1) showing significant correlation with age in regions where other measures did not, and 2) reflecting effects captured by the other measures.

Restricted (or hindered) diffusion is primarily caused by dense packing of axons and their cell membranes (Beaulieu, 2002). Other tissue properties such as myelination and local complexity can also affect the degree of hindrance or restrictance measured at each voxel (Vos et al., 2012). In the current study, an increase in PC1 may indicate higher coherence of the underlying white matter bundles for our older subjects, in comparison with the younger ones. Previous studies have demonstrated that dMRI measures can be sensitive to age-related differences, and those are often associated with an increased microstructural organisation (for review, see Lebel et al. (2017)). Given the well-established role of the CST in supporting motor performance, our finding of increased hindrance with age in typically developing children is in line with previous research that showed that brain maturation varies across different pathways, with commissural and projection tracts reaching maturation by early adolescence while association pathways develop over a longer time period (Geeraert et al., 2018). Interestingly, PC2 also captured an increase in orientational dispersion for that same region (which was either marked by an increase in NuFO or decrease in MD, Fig. 4). The fact that those changes appear near the cortex, a region usually contaminated by partial volume effects, highlights the role of MSMT-CSD in achieving adequate fODFs representation at the boundary between gray matter and white matter (Jeurissen et al., 2014).

In light of our results, we stress that the proposed framework was applied to study neuro-developmental changes related to age only, and therefore, results should be interpreted with care in any other context. The framework should be used as a general approach to data reduction of

MRI-derived measures. However, this does not suggest that all studies will benefit from this form of data reduction since our results may not be generalizable to other populations. Thus, we recommend that our proposed framework is applied to each new data set to discover the most appropriate PCA loadings. Indeed, it is entirely plausible that in a disease population or data collected with different acquisition parameters or hardware, the optimal factor loadings would be different than the ones found here.

4.2. Choice of diffusion measures

A growing interest in utilising advanced dMRI measures to study the human brain motivated us to investigate the shared relationship between DTI and HARDI measures (De Santis et al., 2014). Ultimately, the signals captured from a white matter bundle are coming from the same substrate i.e., from the same axons, myelin and other cellular inclusions, so some degree of correlation is likely to be observed. Redundancy between the different measures (in the sense of correlation) does not however imply that the some of the measures are not useful. In the event that two measures are strongly correlated, any deviation from perfect correlation may reflect that each is capturing subtly different information that may indeed be crucial (for example in understanding disease processes) and thus, regarding one redundant measure to be discarded in favour of the other may not be sound decision. Here, we adopt the term redundancy to refer explicitly to data reduction only.

Being a relatively fast-developing field, dMRI offers a multitude of mathematical models to represent the underlying tissue microstructure (Alexander et al., 2017). Here, we focused on DTI and HARDI measures (Descoteaux, 2015). dMRI measures are generally sensitive to differences in tissue microstructure that can potentially be linked to fibre properties such as myelination and axon density (Scholz et al., 2014). Despite the fact that the specific interpretation of these measures remains controversial (Jones et al., 2013), DTI and HARDI measures are routinely used by neuroscientists and clinicians to gain insights into white matter properties. The findings reported here are in line with existing evidence suggesting that HARDI measures may be more specific than DTI for the detection of differences in tissue microstructure (Tournier et al., 2011; Jeurissen et al., 2013; Cousineau et al., 2017). Our results suggest that combining the sensitivity of DTI and the specificity of HARDI has the potential for compromise between the two techniques. Other macrostructural measures (e.g., bundle volume or mean length; Lebel et al. (2012, 2008, 2017); Geeraert et al. (2018); Girard et al. (2014)) have been used to study brain development and may also provide complementary features that could be added in the proposed framework. Other measures such as rotationally invariant spherical harmonic features (RISH; Mirzaalian et al. (2015); Caruyer and Verma (2015); Zucchelli et al. (2018)) could also be introduced in the current framework, with the main advantage of representing more directly the diffusion signals rather than relying on various microstructural models. Ultimately, the key challenge resides in knowing what measure (or combination of; De Santis et al. (2014)) provides the best value in terms of scanning and processing time. To help with the planning of future studies and based on our observations, we present some recommendations for data analysis.

DTI and HARDI: DTI measures can nowadays be easily be derived from a conventional 30 directions protocol acquired at $b = 1000\ s/mm^2$ in approximatively 5 min. Here, all DTI measures were derived using only the $b = 1200\ s/mm^2$. HARDI measures such as AFD (Raffelt et al., 2012; Dell'Acqua et al., 2013) and NuFO (Dell'Acqua et al., 2013) are also readily derived using a standard 3T MRI scanner. Indeed, CSD can usually be performed on data acquired with a minimum of 30–45 directions at $b = 1000\ s/mm^2$ (Dell'Acqua and Tournier, 2018; Alexander and Barker, 2005; Tournier et al., 2013) or even 21 directions if the quality is acceptable (Chenot et al., 2018a). Moreover, going beyond single b-value acquisitions will provide a better estimation of partial volume effects and better characterisation of various tissue types that will subsequently improve HARDI reconstructions in those areas (Jeurissen et al., 2014; Chamberland et al., 2018).

Tract-averaging and along-tract profiling: In the context of along-tract profiling, age-related effects might be more pronounced when performing group-wise comparisons (e.g., young vs old, patients vs controls; Yeatman et al. (2014)) rather than directly looking at a single cross-sectional change in tissue microstructure. Indeed, theses changes might be too subtle to detect, especially considering that the age range of our participants falls on the *inflection point* of the developmental curve (Lebel and Beaulieu, 2011). Moreover, one may consider first looking at the profile along each tract of interest and ask the following: are there any benefits in sub-segmenting the profile into finer portions? Admittedly, if the measure of interest remains stable along the pathways, a conventional tract average is probably better suited than looking at a constant profile at multiple points. Depending on the research hypothesis, the use of a more permissive approach such as false discovery rate (FDR) correction could be considered to assess differences along multiple adjacent bundle segments. Another possible approach to analyse the multi-dimensionality of the data resides in functional data analysis (FDA; Ramsay (2005); Ferraty and Vieu (2006); Wang et al. (2016)), a statistical method that operates on continuous or discrete functions (e.g., tract profiles; Goodlett et al. (2009)) and that takes into account the spatial interdependency of each segment. In addition, the cluster patterns shown in Fig. 1 may suggest that performing a PCA on each bundle separately might result in different decompositions. We presume this approach would be more appropriate in the context of a bundle-specific analysis (or hypothesis-driven approach). However, by doing so, comparisons between bundles become impossible, as there is no guarantee that e.g., the first PC of a given bundle could be matched with the first PC of another bundle.

Lastly, we acknowledge that the data acquisition and processing employed in this study were performed on unique hardware; however, this should not discourage future users to adopt the current framework for their own data analysis. Again, one should not infer that we are recommending that the factor loadings presented here should be used *off-the-shelf*; rather, we are recommending that the general framework presented here is applied to the specific application, acquisition, and hardware so that the data reduction is tailored. Yet, advances in data harmonisation (Tax et al., 2019) show great promise in bridging the gap between state-of-the-art and standard data acquisition schemes.

4.3. Future perspectives

Over the last few years, a wide range of supervised and unsupervised learning applications based on feature extraction has emerged, ranging from individual classifiers for specific brain disorders (Wang et al., 2010; Chu et al., 2012) to data predictors of brain function (Chen et al., 2009; Franke et al., 2012; Casanova et al., 2011; Kucukboyaci et al., 2014). In the field of functional MRI, independent component analysis (ICA) is a successful example of unsupervised dimensionality reduction that allows the extraction of temporally segregated resting-state networks (Beckmann et al., 2009). In all cases, data reduction approaches facilitate a stream to analyse and interpret the increasingly large multi-dimensional data generated by new methodological models. Admittedly, despite PCA being one of the most commonly-used tool for data reduction, it can also over-fit data (Kramer, 1991) and potentially require multiple post-processing regression steps to explore the link between the resulting components and the observations. Similar to PCA, a canonical correlation analysis (Hotelling, 1936) may help in finding the link between the correlated measures and observations by extracting their joint information. Other non-linear dimensionality reduction techniques based on manifold learning such as isometric feature mapping (Tenenbaum, 1998) or locally linear embedding (Roweis and Saul, 2000) may also better disentangle the measurement space. However, one has to be cautious when applying advanced models of dimensionality reduction to medical imaging as it is often a trade-off between model accuracy and interpretability. Indeed, although these techniques may result in better disentangling of the manifold space, this often comes at the expense of

generating complex and less interpretable features that cannot be related to brain tissue microstructure.

Overall, data representation frameworks such as the one presented here can become fundamental in advancing the application of diffusion models in health and disease. The proposed framework may open new avenues for examining brain microstructure in general and other related lines of research, especially if complimented with other modalities such as measures derived from quantitative magnetisation transfer (Rovaris et al., 2003; Cercignani and Bouyagoub, 2018). Indeed, with the ever-growing acquisition of large cohorts of subjects, feature extraction techniques may become essential tools for processing multi-dimensional brain imaging datasets (e.g., the Human Connectome Project with >1000 young adults scans; Van Essen et al. (2013) or the UK Biobank with its 500,000 participants; Miller et al. (2016)).

Finally, our study may also open new avenues for fibre clustering by leveraging microstructural properties mapped over fibre bundles. Suppl. Fig. 1 shows how different bundles project and cluster in the new reference frame formed by PC1 and PC2. One can observe that PC1 is sensitive to various hindrance level in white matter by disentangling bundles such as the CST (green), genu (blue) and splenium (pink). Conversely, pathways that are known to have many crossing regions such as the AF and the SLF are located on the superior portion of the bi-plot, showing properties of increased orientational dispersion (Suppl. Fig. 1 orange and purple, respectively).

5. Conclusions

In summary, our findings demonstrate that there exist redundancies in measures conventionally derived from dMRI and that those redundancies may be exploited by dimensionality reduction to reduce the risk of Type I errors, arising from multiple statistical comparisons. Our results support the use of data reduction to detect along-tract differences in tissue microstructure. Specifically, the curse of dimensionality and redundancies in statistical analyses were considerably mitigated by extracting principal components that summarise the inter-dependent measures. From an application perspective, a general increase in the first component related to white matter hindrance was found to have a significant correlation with age in various developmentally sensitive pathways, a change that would otherwise remain undetected using conventional approaches.

Funding

MC is supported by the Postdoctoral Fellowships Program from the Natural Sciences and Engineering Research Council of Canada (PDF-502385-2017) and a Wellcome Trust New Investigator Award (to DKJ). ER is supported by a Marshall Sherfield Postdoctoral Fellowship. CMWT is supported by a Rubicon grant from the Netherlands Organisation for Scientific Research (680-50-1527). This work was also supported by a Wellcome Trust Investigator Award (096646/Z/11/Z), a Wellcome Trust Strategic Award (104943/Z/14/Z), and an EPSRC equipment grant (EP/M029778/1).

Acknowledgements

The authors would like to thank Umesh Rudrapatna, Peter Hobden, John Evans, Alison Cooper and Isobel Ward (CUBRIC) for their support with data acquisition. The authors are also thankful to Jean-Christophe Houde (Sherbrooke Connectivity Imaging Lab) for sharing implementation details. The authors are also grateful to the Natural Sciences and Engineering Research Council of Canada (NSERC), the Marshall Sherfield Postdoctoral Fellowship (UK), the Netherlands Organisation for Scientific Research (NWO), the Wellcome Trust (UK) and the Engineering and Physical Sciences Research Council (EPSRC, UK) for supporting this research. Finally, we thank the participants and their families for participating in the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.06.020.

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. Wiley Interdiscip. Rev.: Comput. Stat. 2, 433–459.
Alexander, D.C., Barker, G.J., 2005. Optimal imaging parameters for fiber-orientation estimation in diffusion mri. Neuroimage 27, 357–367.
Alexander, D.C., Dyrby, T.B., Nilsson, M., Zhang, H., 2017. Imaging brain microstructure with diffusion mri: practicality and applications. NMR in Biomedicine.
Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J., Dyrby, T.B., 2010. Orientationally invariant indices of axon diameter and density from diffusion mri. Neuroimage 52, 1374–1389.
Andersson, J.L., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage 20, 870–888.
Andersson, J.L., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging. Neuroimage 125, 1063–1078.
Assaf, Y., Basser, P.J., 2005. Composite hindered and restricted model of diffusion (charmed) mr imaging of the human brain. Neuroimage 27, 48–58.
Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J., 2008. Axcaliber: a method for measuring axon diameter distribution from diffusion mri. Magn. Reson. Med. 59, 1347–1354.
Basser, P.J., Jones, D.K., 2002. Diffusion-tensor mri: theory, experimental design and data analysis–a technical review. NMR Biomed. 15, 456–467.
Basser, P.J., Mattiello, J., LeBihan, D., 1994. Mr diffusion tensor spectroscopy and imaging. Biophys. J. 66, 259–267.
Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. 44, 625–632.
Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system–a technical review. NMR Biomed. 15, 435–455.
Beckmann, C.F., Mackay, C.E., Filippini, N., Smith, S.M., 2009. Group comparison of resting-state fmri data using multi-subject ica and dual regression. Neuroimage 47, S148.
Bellman, R.E., 1961. Dynamic Programming Treatment of the Traveling Salesman Problem.
Bells, S., Cercignani, M., Deoni, S., Assaf, Y., Pasternak, O., Evans, C., Leemans, A., Jones, D., 2011. Tractometry–comprehensive multi-modal quantitative assessment of white matter along specific tracts. In: Proc. ISMRM.
Bourbon-Teles, J., Bells, S., Jones, D.K., Coulthard, E., Rosser, A., Metzler-Baddeley, C., 2017. Myelin breakdown in human huntington's disease: multi-modal evidence from diffusion mri and quantitative magnetization transfer. Neuroscience.
Caruyer, E., Verma, R., 2015. On facilitating the use of hardi in population studies by creating rotation-invariant markers. Med. Image Anal. 20, 87–96.
Casanova, R., Wagner, B., Whitlow, C.T., Williamson, J.D., Shumaker, S.A., Maldjian, J.A., Espeland, M.A., 2011. High dimensional classification of structural mri alzheimer's disease data based on large scale regularization. Front. Neuroinf. 5, 22.
Cattell, R.B., 1966. The scree test for the number of factors. Multivariate Behav. Res. 1, 245–276.
Cercignani, M., Bouyagoub, S., 2018. Brain microstructure by multi-modal mri: is the whole greater than the sum of its parts? Neuroimage 182, 117–127.
Chamberland, M., Tax, C.M., Jones, D.K., 2018. Meyer's loop tractography for image-guided surgery depends on imaging protocol and hardware. Neuroimage: Clin. 20, 458–465.
Chamberland, M., Whittingstall, K., Fortin, D., Mathieu, D., Descoteaux, M., 2014. Real-time multi-peak tractography for instantaneous connectivity display. Front. Neuroinf. 8, 59. https://doi.org/10.3389/fninf.2014.00059.
Chen, K., Reiman, E.M., Huan, Z., Caselli, R.J., Bandy, D., Ayutyanont, N., Alexander, G.E., 2009. Linking functional and structural brain images with multivariate network analyses: a novel application of the partial least square method. Neuroimage 47, 602–610.
Chenot, Q., Tzourio-Mazoyer, N., Rheault, F., Descoteaux, M., Crivello, F., Zago, L., Mellet, E., Jobard, G., Joliot, M., Mazoyer, B., et al., 2018a. A population-based atlas of the human pyramidal tract in 410 healthy participants. Brain Struct. Funct. 1–14.
Chu, C., Hsu, A.L., Chou, K.H., Bandettini, P., Lin, C., Initiative, A.D.N., et al., 2012. Does feature selection improve classification accuracy? impact of sample size and feature selection on classification using anatomical magnetic resonance images. Neuroimage 60, 59–70.
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.
Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E., 1999. Tracking neuronal fiber pathways in the living human brain. Proc. Natl. Acad. Sci. Unit. States Am. 96, 10422–10427.
Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G., 2006. Fiber tract-oriented statistics for quantitative diffusion tensor mri analysis. Med. Image Anal. 10, 786–798.
Cousineau, M., Jodoin, P.M., Garyfallidis, E., Cote, M.A., Morency, F.C., Rozanski, V., GrandMaison, M., Bedell, B.J., Descoteaux, M., 2017. A test-retest study on

Parkinson's ppmi dataset yields statistically significant white matter fascicles. Neuroimage: Clin. 16, 222–233.
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. Hum. Brain Mapp. 37, 989–1004.
De Santis, S., Drakesmith, M., Bells, S., Assaf, Y., Jones, D.K., 2014. Why diffusion tensor mri does well only some of the time: variance and covariance of white matter tissue microstructure attributes in the living human brain. Neuroimage 89, 35–44.
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. Hum. Brain Mapp. 34, 2464–2483.
Dell'Acqua, F., Tournier, J.D., 2018. Modelling White Matter with Spherical Deconvolution: How and Why? NMR in Biomedicine, e3945.
Descoteaux, M., 2015. High angular resolution diffusion imaging (hardi), 1–25. Wiley Encyclopedia of Electrical and Electronics Engineering.
Descoteaux, M., Deriche, R., Knosche, T.R., Anwander, A., 2009. Deterministic and probabilistic tractography based on complex fibre orientation distributions. IEEE Trans. Med. Imaging 28, 269–286.
Descoteaux, M., Deriche, R., Le Bihan, D., Mangin, J.F., Poupon, C., 2011. Multiple q-shell diffusion propagator imaging. Med. Image Anal. 15, 603–621.
Dhollander, T., Raffelt, D., Connelly, A., 2016. Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion mr data without a co-registered t1 image. In: ISMRM Workshop on Breaking the Barriers of Diffusion MRI.
Dziuban, C.D., Shirkey, E.C., 1974. When is a correlation matrix appropriate for factor analysis? some decision rules. Psychol. Bull. 81, 358.
Ennis, D.B., Kindlmann, G., 2006. Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images. Magn. Reson. Med. 55, 136–146.
Ferraty, F., Vieu, P., 2006. Nonparametric Functional Data Analysis: Theory and Practice. Springer Science & Business Media.
Fletcher, P.T., Pizer, S.M., Joshi, S., 2004. Statistical variability in nonlinear spaces: application to shape analysis and DT-MRI. Citeseer.
Franke, K., Luders, E., May, A., Wilke, M., Gaser, C., 2012. Brain maturation: predicting individual brainage in children and adolescents using structural mri. Neuroimage 63, 1305–1312.
Garg, A., Tai, K., 2013. Comparison of statistical and machine learning methods in modelling of data with multicollinearity. Int. J. Model. Identif. Control 18, 295–312.
Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I., Contributors, D., 2014. Dipy, a library for the analysis of diffusion mri data. Front. Neuroinf. 8.
Garyfallidis, E., Côté, M.A., Rheault, F., Sidhu, J., Hau, J., Petit, L., Fortin, D., Cunanne, S., Descoteaux, M., 2018. Recognition of white matter bundles using local and global streamline-based registration and clustering. Neuroimage 170, 283–295.
Geeraert, B.L., Lebel, R.M., Mah, A.C., Deoni, S.C., Alsop, D.C., Varma, G., Lebel, C., 2018. A comparison of inhomogeneous magnetization transfer, myelin volume fraction, and diffusion tensor imaging measures in healthy children. Neuroimage 182, 343–350.
Genc, S., Seal, M.L., Dhollander, T., Malpas, C.B., Hazell, P., Silk, T.J., 2017. White matter alterations at pubertal onset. Neuroimage 156, 286–292.
Genc, S., Smith, R.E., Malpas, C.B., Anderson, V., Nicholson, J.M., Efron, D., Sciberras, E., Seal, M., Silk, T.J., 2018. Development of White Matter Fibre Density and Morphology over Childhood: a Longitudinal Fixel-Based Analysis. bioRxiv, p. 342097.
Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M., 2014. Towards quantitative connectivity analysis: reducing tractography biases. Neuroimage 98, 266–278.
Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al., 2013. The minimal preprocessing pipelines for the human connectome project. Neuroimage 80, 105–124.
Goodlett, C.B., Fletcher, P.T., Gilmore, J.H., Gerig, G., 2009. Group analysis of dti fiber tract statistics with application to neurodevelopment. Neuroimage 45, S133–S142.
Groeschel, S., Tournier, J.D., Northam, G.B., Baldeweg, T., Wyatt, J., Vollmer, B., Connelly, A., 2014. Identification and interpretation of microstructural abnormalities in motor pathways in adolescents born preterm. Neuroimage 87, 209–219.
Hotelling, H., 1936. Relations between two sets of variates. Biometrika 28, 321–377.
Jeurissen, B., Leemans, A., Tournier, J.D., Jones, D.K., Sijbers, J., 2013. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. Hum. Brain Mapp. 34, 2747–2766.
Jeurissen, B., Tournier, J.D., Dhollander, T., Connelly, A., Sijbers, J., 2014. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. Neuroimage 103, 411–426.
Jolliffe, I.T., 2002. Graphical representation of data using principal components. Princ. Compon. Anal. 78–110.
Jones, D., Alexander, D., Bowtell, R., Cercignani, M., Dell'Acqua, F., McHugh, D., Miller, K., Palombo, M., Parker, G., Rudrapatna, U., et al., 2018. Microstructural imaging of the human brain with a 'super-scanner': 10 key advantages of ultra-strong gradients for diffusion mri. Neuroimage 182, 8–38.
Jones, D.K., Catani, M., Pierpaoli, C., Reeves, S.J., Shergill, S.S., O'sullivan, M., Golesworthy, P., McGuire, P., Horsfield, M.A., Simmons, A., et al., 2006. Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. Hum. Brain Mapp. 27, 230–238.
Jones, D.K., Catani, M., Pierpaoli, C., Reeves, S.J., Shergill, S.S., O'sullivan, M., Maguire, P., Horsfield, M.A., Simmons, A., Williams, S.C., et al., 2005a. A diffusion tensor magnetic resonance imaging study of frontal cortex connections in very-late-onset schizophrenia-like psychosis. Am. J. Geriatr. Psychiatry 13, 1092–1099.
Jones, D.K., Knösche, T.R., Turner, R., 2013. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion mri. Neuroimage 73, 239–254.
Jones, D.K., Travis, A.R., Eden, G., Pierpaoli, C., Basser, P.J., 2005b. Pasta: pointwise assessment of streamline tractography attributes. Magn. Reson. Med. 53, 1462–1467.
Kanaan, R.A., Shergill, S.S., Barker, G.J., Catani, M., Ng, V.W., Howard, R., McGuire, P.K., Jones, D.K., 2006. Tract-specific anisotropy measurements in diffusion tensor imaging. Psychiatr. Res. Neuroimaging 146, 73–82.
Kassambara, A., Mundt, F., 2017. Factoextra: extract and visualize the results of multivariate data analyses. URL: https://CRAN.R-project.org/package=factoextra. r package version 1.0.5.
Kellner, E., Dhital, B., Kiselev, V.G., Reisert, M., 2016. Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn. Reson. Med. 76, 1574–1581.
Kindlmann, G., Ennis, D.B., Whitaker, R.T., Westin, C.F., 2007. Diffusion tensor analysis with invariant gradients and rotation tangents. IEEE Trans. Med. Imaging 26, 1483–1499.
Kramer, M.A., 1991. Nonlinear principal component analysis using autoassociative neural networks. AIChE J. 37, 233–243.
Kucukboyaci, N., Kemmotsu, N., Leyden, K., Girard, H., Tecoma, E., Iragui, V., McDonald, C., 2014. Integration of multimodal mri data via pca to explain language performance. Neuroimage: Clin. 5, 197–207.
Lebel, C., Beaulieu, C., 2011. Longitudinal development of human brain wiring continues from childhood into adulthood. J. Neurosci. 31, 10937–10947.
Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., Beaulieu, C., 2012. Diffusion tensor imaging of white matter tract evolution over the lifespan. Neuroimage 60, 340–352.
Lebel, C., Treit, S., Beaulieu, C., 2017. A review of diffusion mri of typical white matter development from early childhood to young adulthood. NMR Biomed., e3778
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.
LeBihan, D., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H., 2001. Diffusion tensor imaging: concepts and applications. J. Magn. Reson. Imaging 13, 534–546.
Metzler-Baddeley, C., Foley, S., De Santis, S., Charron, C., Hampshire, A., Caeyenberghs, K., Jones, D.K., 2017. Dynamics of white matter plasticity underlying working memory training: multimodal evidence from diffusion mri and relaxometry. J. Cogn. Neurosci. 29, 1509–1520.
Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., et al., 2016. Multimodal population brain imaging in the UK biobank prospective epidemiological study. Nat. Neurosci. 19, 1523.
Mirzaalian, H., de Pierrefeu, A., Savadjiev, P., Pasternak, O., Bouix, S., Kubicki, M., Westin, C.F., Shenton, M.E., Rathi, Y., 2015. Harmonizing diffusion mri data across multiple sites and scanners. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 12–19.
Mori, S., Van Zijl, P.C., 2002. Fiber tracking: principles and strategies–a technical review. NMR Biomed. 15, 468–480.
Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J. Classif. 31, 274–295.
Mwangi, B., Tian, T.S., Soares, J.C., 2014. A review of feature reduction techniques in neuroimaging. Neuroinformatics 12, 229–244.
Penke, L., Maniega, S.M., Murray, C., Gow, A.J., Hernández, M.C.V., Clayden, J.D., Starr, J.M., Wardlaw, J.M., Bastin, M.E., Deary, I.J., 2010. A general factor of brain white matter integrity predicts information processing speed in healthy older people. J. Neurosci. 30, 7569–7574.
Pierpaoli, C., Basser, P.J., 1996. Toward a quantitative assessment of diffusion anisotropy. Magn. Reson. Med. 36, 893–906.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.
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.
Ramsay, J., 2005. Functional data analysis. Encycl. Stat. Behav. Sci.
Rovaris, M., Iannucci, G., Cercignani, M., Sormani, M.P., De Stefano, N., Gerevini, S., Comi, G., Filippi, M., 2003. Age-related changes in conventional, magnetization transfer, and diffusion-tensor mr imaging findings: study with whole-brain tissue histogram analysis1. Radiology 227, 731–738.
Roweis, S.T., Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290, 2323–2326.
RStudio Team, 2016. RStudio. Integrated Development Environment for R. RStudio, Inc., Boston, MA. URL: http://www.rstudio.com/.
Sairanen, V., Leemans, A., Tax, C., 2018. Fast and accurate slicewise outlier detection (solid) with informed model estimation for diffusion mri data. Neuroimage 181, 331–346.
Scherrer, B., Schwartzman, A., Taquet, M., Sahin, M., Prabhu, S.P., Warfield, S.K., 2016. Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (diamond). Magn. Reson. Med. 76, 963–977.
Scholz, J., Tomassini, V., Johansen-Berg, H., 2014. Individual differences in white matter microstructure in the healthy brain. In: Diffusion MRI. Elsevier, pp. 301–316.
Seunarine, K., Alexander, D., 2009. Multiple fibers: beyond the diffusion tensor. Chapter 4. In: Diffusion mri: from quantitative measurement to in-vivo neuroanatomy.
Seunarine, K.K., Clayden, J.D., Jentschke, S., Munoz, M., Cooper, J.M., Chadwick, M.J., Banks, T., Vargha-Khadem, F., Clark, C.A., 2016. Sexual dimorphism in white matter developmental trajectories using tract-based spatial statistics. Brain Connect. 6, 37–47.

Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., et al., 2013. Advances in diffusion mri acquisition and processing in the human connectome project. Neuroimage 80, 125–143.
Suryanarayana, U., Parker, G., Roberts, J., Jones, D., 2018. Can we correct for interactions between subject motion and gradient-nonlinearity in diffusion mri?. In: Proc. ISMRM.
Tax, C.M., Grussu, F., Kaden, E., Ning, L., Rudrapatna, U., Evans, J., St-Jean, S., Leemans, A., Koppers, S., Merhof, D., et al., 2019. Cross-scanner and cross-protocol diffusion mri data harmonisation: a benchmark database and evaluation of algorithms. NeuroImage.
Tenenbaum, J.B., 1998. Mapping a manifold of perceptual observations. In: advances in neural information processing systems, pp. 682–688.
Tournier, J., Calamante, F., Connelly, A., et al., 2012. Mrtrix: diffusion tractography in crossing fiber regions. Int. J. Imaging Syst. Technol. 22, 53–66.
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, 1459–1472.
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 Biomed. 26, 1775–1786.
Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. Neuroimage 23, 1176–1185.
Tournier, J.D., Mori, S., Leemans, A., 2011. Diffusion tensor imaging and beyond. Magn. Reson. Med. 65, 1532–1556.
Tuch, D.S., et al., 2002. Diffusion MRI of Complex Tissue Structure. Ph.D. thesis. Massachusetts Institute of Technology.
Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al., 2013. The Wu-minn human connectome project: an overview. Neuroimage 80, 62–79.
Veraart, J., Novikov, D.S., Christiaens, D., Ades-Aron, B., Sijbers, J., Fieremans, E., 2016. Denoising of diffusion mri using random matrix theory. Neuroimage 142, 394–406.
Vos, S.B., Jones, D.K., Jeurissen, B., Viergever, M.A., Leemans, A., 2012. The influence of complex white matter architecture on the mean diffusivity in diffusion tensor mri of the human brain. Neuroimage 59, 2208–2216.
Vos, S.B., Tax, C.M., Luijten, P.R., Ourselin, S., Leemans, A., Froeling, M., 2017. The importance of correcting for signal drift in diffusion mri. Magn. Reson. Med. 77, 285–299.
Wang, J.L., Chiou, J.M., Müller, H.G., 2016. Functional data analysis. Annu. Rev. Stat. Appl. 3, 257–295.
Wang, Y., Fan, Y., Bhatt, P., Davatzikos, C., 2010. High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. Neuroimage 50, 1519–1535.
Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54, 1377–1386.
Wickham, H., et al., 2014. Tidy data. J. Stat. Softw. 59, 1–23.
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.
Yeatman, J.D., Wandell, B.A., Mezer, A.A., 2014. Lifespan maturation and degeneration of human brain white matter. Nat. Commun. 5, 4932.
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, 1000–1016.
Zucchelli, M., Deslauriers-Gauthier, S., Deriche, R., 2018. A closed-form solution of rotation invariant spherical harmonic features in diffusion mri. In: MICCAI-computational Diffusion MRI Workshop 2018.