bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Highlights

Incorporating outlier information into diffusion MR tractogram filtering for robust structural brain connectivity and microstructural analyses

Viljami Sairanen,Mario Ocampo-Pineda,Cristina Granziera,Simona Schiavi,Alessandro Daducci

- We present a novel augmentation to tractogram filtering method that accounts for subject motion related signal dropout artefacts in diffusion weighted images.
- Our method is validated with realistic Monte-Carlo whole brain simulations and evaluated with in vivo infant data.
- We show that even if data has 10% of motion corrupted slices our method is capable to mitigate their effect in structural brain connectivity analyses and microstructural mapping.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Incorporating outlier information into diffusion MR tractogram filtering for robust structural brain connectivity and microstructural analyses

Viljami Sairanen[a,b,c,*], Mario Ocampo-Pineda[a], Cristina Granziera[b], Simona Schiavi[a] and Alessandro Daducci[a]

[a]*Department of Computer Science, University of Verona, Verona, Italy*

[b]*Translational Imaging in Neurology, Department of Medicine and Biomedical Engineering, University Hospital Basel and University of Basel, Neurologic Clinic and Policlinic, Basel, Switzerland*

[c]*BABA Center, Pediatric Research Center, Department of Clinical Neurophysiology, Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland*

ARTICLE INFO

Keywords:
Magnetic Resonance Imaging
Diffusion
Tractography
Filtering
COMMIT
Outlier
Structural Connectivity
Microstructure
Movement

ABSTRACT

The white matter structures of the human brain can be represented via diffusion tractography. Unfortunately, tractography is prone to find false-positive streamlines causing a severe decline in its specificity and limiting its feasibility in accurate structural brain connectivity analyses. Filtering algorithms have been proposed to reduce the number of invalid streamlines but the currently available filtering algorithms are not suitable to process data that contains motion artefacts that are typical in clinical research. We augmented the Convex Optimization Modelling for Microstructure Informed Tractography (COMMIT) filtering algorithm to adjust for signal drop-out artifacts due to subject motion present in diffusion-weighted images. We demonstrate with comprehensive Monte-Carlo whole brain simulations and in vivo infant data that our robust algorithm is capable to properly filter tractography reconstructions despite these artefacts. We evaluated the results using parametric and non-parametric statistics and our results demonstrate that if not accounted for, motion artefacts can have severe adverse effect in the human brain structural connectivity analyses as well as in microstructural property mappings. In conclusion, the usage of robust filtering methods to mitigate motion related errors in tractogram filtering is highly beneficial especially in clinical studies with uncooperative patient groups such as infants. With our presented robust augmentation and open-source implementation, robust tractogram filtering is readily available.

1. Introduction

Diffusion-weighted magnetic resonance imaging (dMRI) of the human brain (Basser et al., 1994) has various applications ranging from early clinical stroke diagnostics (Horsfield and Jones, 2002) to investigations of the microstructural properties of the tissue (Alexander et al., 2019; Novikov et al., 2019) and structural brain connectivity mapping (Griffa et al., 2013; Delettre et al., 2019; Zhang et al., 2021). The latter two are gaining popularity in clinical research (Kamiya et al., 2020) to investigate various brain diseases and neurological conditions of adults (Fieremans et al., 2013; Benitez et al., 2014) and development of the growing brain in children and adolescents (Genc et al., 2017; Huber et al., 2019). Furthermore, with the latest advances in automatic brain segmentation with tools like Infant FreeSurfer (Zöllei et al., 2020), it is likely that the amount of brain connectivity studies of infants (Kunz et al., 2014; Pannek et al., 2018; Pecheva et al., 2019) will grow in the near future too.

The clinical dMRI research comes with its own puzzles to solve, with one most difficult being the patient motion. The subject motion can be unavoidable when imaging infants or patients in discomfort or pain, resulting in complex missing data problems (Andersson et al., 2016; Sairanen et al., 2017, 2018). Therefore, the processing of the motion corrupted images requires specialized algorithms and robust methods to minimize motion induced bias in the results. While robust modeling has been considered in the contexts of diffusion and kurtosis tensor estimations (Chang et al., 2005, 2012; Tax et al., 2015) as well as in higher order models (Pannek et al., 2012) that could be used for tractography purposes, it has not been investigated thoroughly in the context of the brain structural connectivity analyses.

Structural brain connectivity analyses are based on the rapidly developing dMRI tractography (Basser et al., 2000) algorithms that represent the brain white matter structures with streamlines. These streamlines can be used to investigate which gray matter regions might have a structural link. In general, the tractography algorithms are sensitive but they lack specificity and they find great number of false streamlines connections (Thomas et al., 2014; Maier-Hein et al., 2017). This means that two gray matter regions could be linked by tractography streamlines despite that the brain tissue does not form a true structural link. This is a known issue in structural connectivity analyses (Drakesmith et al., 2015; Zalesky et al., 2016; Yeh et al., 2020) to which tractogram filtering has been proposed as one solution. Tractogam filtering can be achieved with different approaches, one being the Convex Optimization Modelling for Microstructure Informed Tractography (COMMIT) (Zhang et al., 2021) which we will use in this study to demonstrate possible effects of subject motion to the tractogram filtering and microstructural mapping as well as how it can be accounted

*Corresponding author
viljami.sairanen@helsinki.fi (V. Sairanen)
ORCID(s): 0000-0002-8833-7155 (V. Sairanen)
https://twitter.com/@ViljamiSairanen (V. Sairanen)

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and corrected for.

There are three alternative post-scan approaches to address the outliers caused by the subject motion. The first approach is to find outliers in dMRI data manually or automatically with statistical methods or deep-learning and simply exclude the artefactual dMRI data or even the whole subject from the analysis (Oguz et al., 2014; Samani et al., 2019). The second approach is to use a model to predict what the measurements should look like, locate the outliers based on differences to model predictions and replace them with these predictions if differences are deemed large enough (Lauzon et al., 2013; Andersson et al., 2016). The third approach is to detect the outliers, but instead of replacing or completely excluding them, their weight is reduced in all subsequent model estimation steps (Sairanen et al., 2018).

Manual outlier detection can be laborious and excluding whole subjects from clinical studies with relatively small number of participants might not be the optimal choice. The outlier replacement approach relies on the quality and robustness of the chosen model and method to represent the measured dMRI signal. If multiple dMRI measurements are corrupted by motion artifacts, this initial modeling and prediction step can fail altogether (Sairanen et al., 2018). Even in the best case, the replaced data points are simply interpolations based on the chosen model and the data points used in the modeling therefore it cannot increase the available information but leads to increased error propagation due to subsequent model fittings. The third approach, on the contrary, enables quantifying the amount of the motion corrupted data and versatile subsequent modeling and analysis options therefore being optimal for our purposes.

While weighted and robust modeling has been implemented before, they have mostly been used outside the scope of tractogram filtering. For example, in diffusion tensor modeling weighted linear least squares is typically the fastest and most robust estimator (Veraart et al., 2013; Tax et al., 2015; Sairanen et al., 2018). Robust modeling has been proposed for higher order models as well (Pannek et al., 2012). In the context of tractogram filtering, weighted cost functions have been introduced earlier in e.g. SIFT (Smith et al., 2013, 2015), but it has only been evaluated with voxels affected by partial voluming. SIFT algorithm states that their 'processing mask' is 'the square of the estimated white matter partial volume fraction' - which indeed should be beneficial in the case of partial voluming. However, the approach in SIFT does not account for outliers that are randomly occurring in the measurements as our newly proposed augmentation to COMMIT does.

In this work, we propose a robust augmentation to the COMMIT filtering algorithm (Daducci et al., 2015) that accounts for the unreliability of the original measurements. We detail the theoretical changes to the algorithm as well as provide open-source code[1] of its implementation. To evaluate the method, we use the data from the Human Connectome Project (HCP) (Van Essen et al., 2013) as a base for thorough Monte-Carlo simulations which emulate various motion induced artifacts in synthetic whole brain data with Rician noise. Synthetic data provide the necessary baseline that can be used to highlight the bias arising from subject motion in structural connectivity analyses as well as how well it can be amended using our robust augmentation.

In the context of this study, the measurement unreliability is associated with outliers due to subject motion. However, it can readily be utilized to correct for measurements that are affected by partial voluming, as our preliminary results have demonstrated earlier (Sairanen et al., 2021). This proposed augmentation to tractogram filtering provides the necessary update for them to be practically usable in clinical research.

2. Materials and Methods

2.1. Implementation

We augmented the original cost function of COMMIT (Daducci et al., 2015) algorithm with a voxelwise weighting factor $\mathbf{W}$ that can be used to downweight data points that have decreased reliability due to subject motion or any other reason. The original COMMIT is based on a minimization of the difference between the original measurements and a forward model prediction. The forward model prediction is calculated by fitting a chosen microstructural model for each streamline in every voxel. COMMIT assigns a weight to each streamline that tells how much that streamline contributes to the predicted signal. These streamline contribution weights are iteratively updated until the difference between the measurements and this prediction converges to a minimum. Any streamline with contribution of zero is then removed as an implausible streamline (i.e. not compatible with the measured signal). If the original measurement is artefactual due to subject motion or any other reason, the algorithm could converge to an incorrect solution. To decrease the weight of these artefactual data points, we propose the robust cost function shown in eq. 1. Our proposed idea is further illustrated in Fig. 1 with a simple toy example.

$$\underset{\hat{\mathbf{x}}\geq 0}{\operatorname{argmin}} \left\|\mathbf{W}\left(\mathbf{A}\hat{\mathbf{x}}-\mathbf{y}\right)\right\|_2^2 . \tag{1}$$

The robust cost function in eq. 1 is intended to be used with outlier detection with tools such as SOLID (Sairanen et al., 2018). SOLID detects slicewise outliers based on robust statistical analysis of the original dMRI data and can be used either to exclude outliers or downweight them depending how strong outliers are. This downweighting scheme is likely a better option to outlier replacement that is proposed in earlier studies (Lauzon et al., 2013; Andersson et al., 2016). If the outlier is replaced with a prediction from a tensor or a gaussian model, then COMMIT would try to minimize the difference from those model predictions to its own model prediction. Since these models can be different and therefore capture different details of the dMRI signal, it is more straightforward to use robust modeling with the proposed weighted cost function.

[1] https://github.com/daducci/COMMIT

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

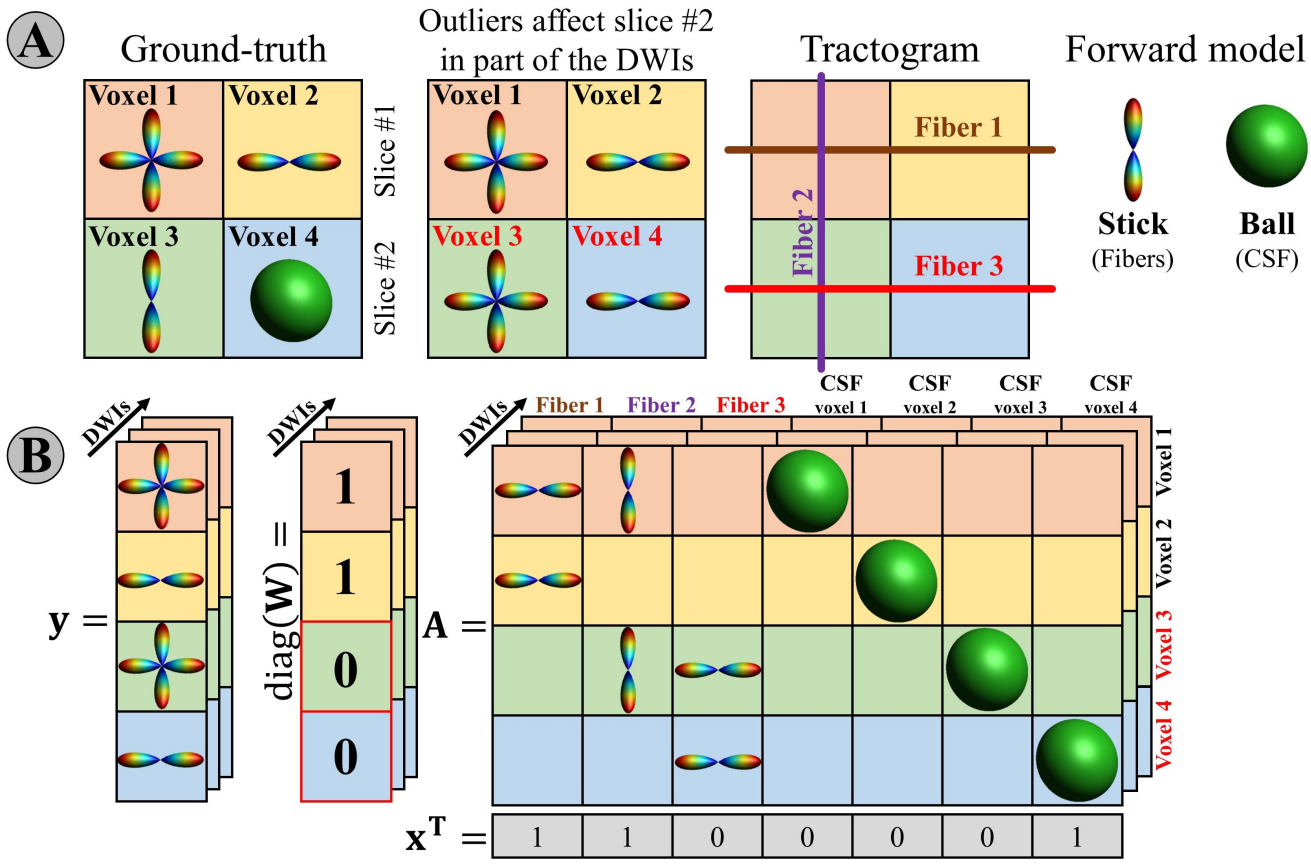

Figure 1: A toy example to illustrate how the COMMIT framework was augmented with the reliability weights. (**A**) A ground-truth synthetic phantom consists of two slices with both having two voxels i.e. a 2x2 grid. A part of the dMRI measurements are affected by slicewise outliers at the slice 2 position therefore fiber orientations in voxels 3 and 4 become biased and resulting tractogram contains implausible streamline Fiber 3. (**B**) The data, also the corrupted measurements, are in the vector **y**. Corresponding ODFs are overlaid on the voxel positions. The diagonal vector of the reliability weight matrix **W** illustrates how the data in the affected dMRI signals is downweighted during modeling. Artefactual measurements are highlighted with red boundaries and 0 weights whereas normal data have the full weight of 1. Same spatial voxel position in different dMRI volumes can have different weights which is demonstrated with the 3rd 'dMRI data' dimension. The design matrix **A** depicts how each streamline is used in the forward modeling as a stick along with additional compartments for e.g. CSF. The streamline contributions **x** are shown at the bottom demonstrating the expected behaviour of the augmented COMMIT.

2.2. Simulations

To investigate the outlier effect on the tractogram filtering, we developed a comprehensive Monte-Carlo simulation pipeline delineated in Fig. 2. Simulations are based on T1-weighted and dMRI data from the HCP subject 100308 which were processed with current state-of-the-art methods (Van Essen et al., 2013). We do not expect or imply that this ground-truth connectivity matrix depicted in Fig. 3 would represent the true structural connections in a human brain. It simply provides us the necessary ground-truth connectivity that we can use to evaluate the noise and outlier effects in the Monte-Carlo simulations with more realistic picture of the whole brain than typical fiber phantoms.

Ground-truth data

We segmented the T1-weighted HCP data with FreeSurfer (Fischl, 2012) to obtain 85 regions-of-interests (ROIs) based on the Desikan-Killiany atlas (Desikan et al., 2006). Instead of the full brainstem, we used only its inferior part of medulla as the last ROI. We used these brain segments to compute the ground-truth connectivity matrix as well as to ensure that we used only the connecting streamlines in our analyses.

To calculate a whole brain tractogram from the HCP dMRI data, we used the anatomically constrained probabilistic tractography (iFOD2) (Tournier et al., 2010; Smith et al., 2012) implemented in MRTrix3 software (Tournier et al., 2019). We used the white matter mask as a seed region for three million streamlines. The tracking parameters were left to their default values and we removed all non-connecting stream-

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

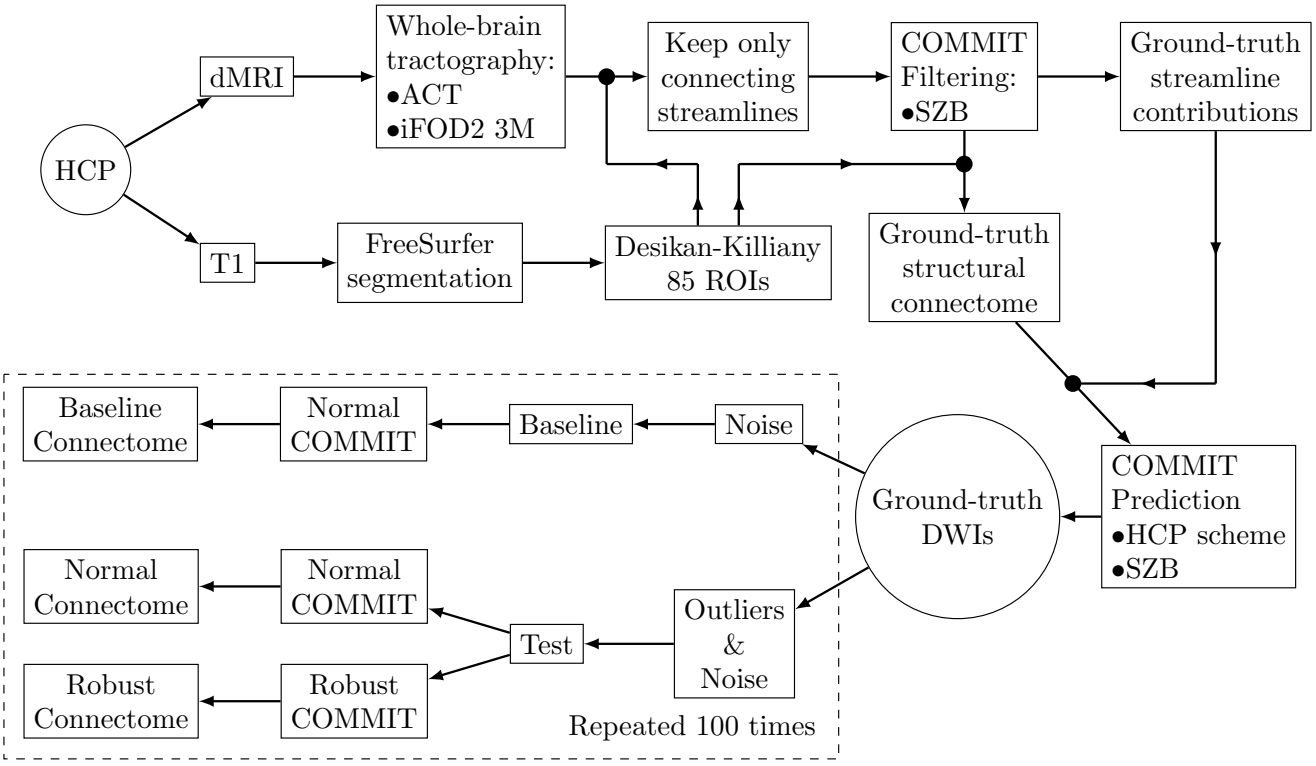

Figure 2: A flow chart describing how the whole-brain simulations were obtained from the HCP dataset. The dMRI and T1-weighted data were used to obtain the ground-truth connectome from which the ground-truth dMRI signals were predicted using COMMIT forward modeling. The ground-truth data were used to perform 100 Monte-Carlo simulations to evaluate the effects of noise and outliers to the structural brain connectome.

lines based on the 85 ROI segmentation of T1-image.

For tractogram filtering, we used the original COMMIT (Daducci et al., 2015) as the data did not contain any slice-wise outliers. The used forward model was the stick-zeppelin-ball (SZB) with $1.7 \cdot 10^{-3} mm^2/s$ for parallel stick and zeppelin diffusivities, $0.61 \cdot 10^{-3} mm^2/s$ for perpendicular zeppelin diffusivity, and $1.7 \cdot 10^{-3} mm^2/s$ and $3.0 \cdot 10^{-3} mm^2/s$ for the isotropic (ball) diffusivities as model parameters (Panagiotaki et al., 2012). The filtered tractogram was used to form the ground-truth connectivity matrix with the information from T1-segmentation (Fig. 3). Moreover, we combined this information with the final streamline contributions to form the synthetic whole brain prediction of dMRI data using the HCP's three-shell gradient scheme. This produced 270 noise free diffusion-weighted whole brain images that we used as a ground-truth for our Monte-Carlo simulations.

Monte-Carlo design

Our Monte-Carlo simulations were based on the ground-truth synthetic whole brain dMRI data obtained from HCP subject. We split the simulations in two groups: Baseline and Test. Baseline group provides the means to evaluate the pure noise effects on the connectome whereas Test group provides the means to evaluate the outlier effects.

In Baseline group, random Rician noise was added before repeating the normal COMMIT filtering with the original non-filtered but connecting streamlines. The Rician noise had signal-to-noise ratio of 20 based on the non-diffusion weighted signal which is roughly similar with signal-to-noise ratios in clinical research. We used the same filtering parameters that were used to form the ground-truth data. This process was repeated to obtain 100 whole brain baseline images and connectomes.

In Test group, outliers were introduced to the data before adding the same Rician noise that was used for the Baseline group. Test group was filtered with both the normal COMMIT as well as the proposed robust COMMIT using the same streamlines and parameters that were used for the Baseline group. This process was repeated to obtain 100 whole brain test images with outliers and corresponding connectomes from normal and robust filtering methods.

The outlier selection for the Test group was done by replacing axial slices with signal decrease outliers in an interleaved manner to 9 (10%) of the dMRI data per shell. The chosen number is not completely arbitrary as it is reported to be approximately the maximum amount of corrupted data that robust pre-processing tools can tolerate (Andersson et al., 2016; Sairanen et al., 2018). To maximize the missing data problem(Sairanen et al., 2017), the sampling of outliers in q-space was done by picking the first gradient

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

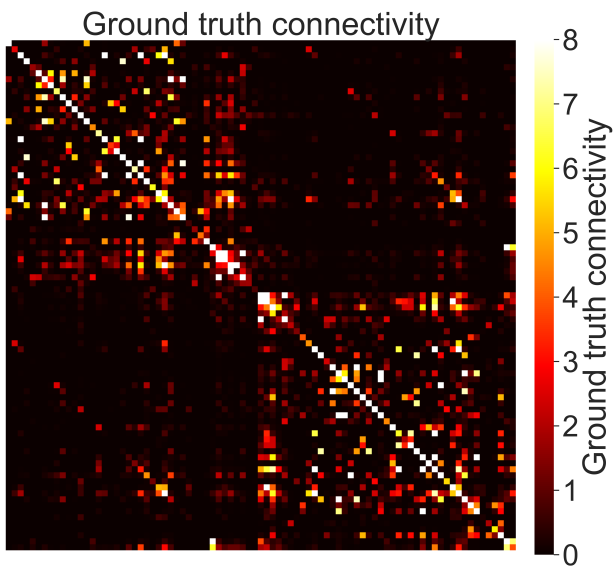

Figure 3: The ground-truth connectivity matrix used in this study was based on one subject. While, this connectivity matrix might not represent the real human brain connections, it provided the necessary ground-truth control level for our Monte-Carlo simulations. Any deviations from this connectivity matrix observed in the simulations would be due to noise, outliers, or both.

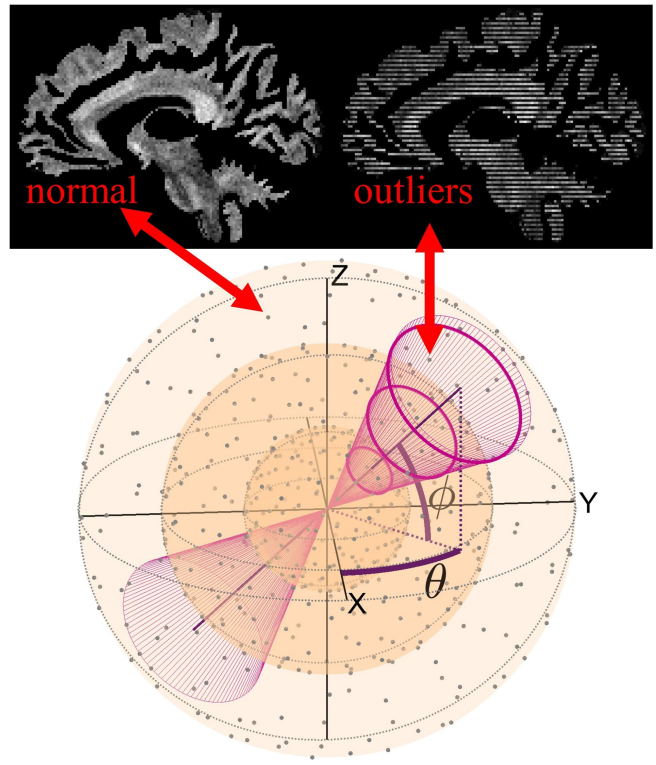

Figure 4: An illustration how a part of the gradient directions were selected from the q-space to alter the corresponding dMRI data with slicewise outliers. The initial direction (θ, ϕ) of was selected randomly after which the opening angle of the selection cone was increased until the specified amount of outliers from each shell remained inside it. This approach ensures the maximal gap in the q-space sampling therefore having a high chance to find error prone schemes.

direction randomly and the rest were selected based on the smallest angular distances from the first one as illustrated in Fig. 4. This was performed for each shell separately.

Statistical analysis

We investigated global brain connectivity as well as individual connections using analysis of variance (ANOVA) accompanied by Tukey's honestly significant difference (HSD) test and non-parametric Friedman's test accompanied by two-sample Kolmogorov-Smirnov tests. The reason for having these different test statistics is that outliers can lead to skewed and long tailed distributions that might not be correctly investigated solely by parametric tests. For example, two-sample Kolmogorov-Smirnov test is necessary as the outlier effect can only be studied by comparing Test groups to Baseline group. This is because Rician noise has a non-zero positive average therefore it likely causes a bias that deviates even Baseline group from the ground-truth values.

While we report p-values from these tests, we argue that the effect sizes are more interesting as they describe how different the tested groups are. The effect sizes are measured using Cohen's D for parametric tests and Kolmogorov-Smirnov statistic for non-parametric tests. The test statistics we employ are widely used and they provide information about average differences and differences in the shapes of the Monte-Carlo simulated distributions. For details about these tests, we recommend any textbook that covers parametric and non-parametric statistics such as Sheskin's handbook (Sheskin, 2004).

2.3. in vivo measurements

Infant data

We obtained preliminary data from an on-going infant study to evaluate our method with in vivo measurements. T1-weighted image and dMRI data were obtained with 3T MRI Siemens Skyra system (Erlangen, Germany) with a 32 channel head coil. The dMRI acquisition consisted of 13 non-diffusion weighted images that were interspersed between 60 diffusion-weighted images with b-value of $750 s/mm^2$ and 74 diffusion-weighted images with b-value of $1800 s/mm^2$ each with uniquely oriented gradients. Bipolar gradient scheme was used to minimize geometrical distortions due to eddy currents. The image resolution was isotropic $2mm$ with 80 $\times$ 80 $\times$ 44 imaging matrix. The in-plane acceleration factor was 2 (SENSE) and multi-band acceleration factor was 2. Only anterior-posterior phase encoded images were acquired as the reverse phase encoding required manual adjustment during the scan which was deemed infeasible at the used clinical scan environment. The use of infant data in this work was approved by the relevant Ethics Committee of the Helsinki University Hospital.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Infant analyses

We used ExploreDTI (Leemans et al., 2009) with SOLID-plugin (Sairanen et al., 2018) to simultaneously detect slice-wise outliers and to correct for subject motion and eddy currents as well as registered the data to anatomical T1-image to correct for geometrical distortions. Additionally, we used Gibbs ringing correction (Perrone et al., 2015). We did not correct for signal drift (Vos et al., 2017) as it was not observed in the measurements.

Processing of this data was limited to specific computers in the hospital network which prevented memory demanding tasks such as segmentation with infant Freesurfer (Zöllei et al., 2020). Therefore, we opted for a simpler analysis that consisted of comparison between normal and robust filtering method outputs instead of full connectivity analyses. The T1-image contrast of this subject was not suitable for white and grey matter segmentations. We obtained a WM mask from multi-shell multi-tissue constrained spherical deconvolution (Jeurissen et al., 2014) implemented in MRTrix3 (Tournier et al., 2019) and used that as a seed mask for probabilistic whole-brain tractography (iFOD2) (Tournier et al., 2010) to generate three million streamlines.

We filtered the generated streamlines with normal COMMIT (Daducci et al., 2015) and the proposed robust COMMIT_r to evaluate the improvements in the overall fit from root mean squared error (RMSE) maps as well as to see the impact of outliers in intracellular and isotropic signal fractions. We used the stick-ball model for both filtering methods with the following parameters: $1.7 \cdot 10^{-3} mm^2/s$ for parallel signal diffusivity, and $1.7 \cdot 10^{-3} mm^2/s$ and $3.0 \cdot 10^{-3} mm^2/s$ for the isotropic signal diffusivities.

3. Results

3.1. Simulations

We investigated the effects of noise to the structural brain connectivity by comparing Baseline group. Test groups could not be directly compared to the ground-truth due to noise bias. Therefore, outlier effects were investigated by comparing Test groups to Baseline group instead. We evaluated these differences in both global connectivity matrix score and individual connection strengths.

Global Connectivity

The global connectivity difference was defined as the sum of the root mean squared differences between the upper triangle of the connectivity matrices from Monte-Carlo groups and the corresponding ground-truth values. The results of this comparison calculated over all 100 Monte-Carlo simulations are shown in Fig. 5 with violin plots. The noise effect on the global connectivity (Baseline) is shown with the first violin from the left, the outlier effect is shown in the middle, and the outlier effect with robust filtering is shown on the right.

Both, Baseline and robust COMMIT_r produced similar global results with differences ranging from 0.2 to 0.24. This demonstrates that on average, the proposed robust filtering method is capable to mitigate the outlier effect. On the contrary, the results from normal COMMIT ranged from 0.2 to 0.3 demonstrating that outliers can have a much stronger effect than noise on the global connectivity values.

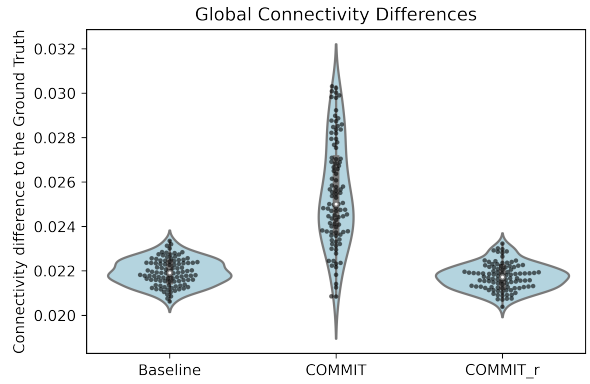

Figure 5: The differences between the ground-truth connectome and each Monte-Carlo simulated connectome depict the global increase in the variability of connectivity values due to outliers. The Baseline group did not have any outliers therefore the difference from the ground-truth is due to noise only. The Test group with outliers was filtered with both normal COMMIT and robust COMMIT_r. The normal COMMIT differences were up to 35% higher than the Baseline connectivity differences whereas the robust COMMIT_r was not affected by the outliers as much and produced very similar distribution with the Baseline.

Parametric statistical analysis. The global connectivity differences with ANOVA detail that the group averages were statistically different with p-value less than 0.01. Tukey's HSD test depicted that normal COMMIT had significantly different mean to both Baseline and COMMIT_r results with p-value less than 0.01. We evaluated the effect size with Cohen's D. The effect size between Baseline and normal COMMIT was 2.0 and the difference between robust COMMIT_r to COMMIT was 2.2. The difference between Baseline and COMMIT_r was not statistically significant with p-value of 0.56 and had a smaller effect size of 0.4. These effect sizes indicate that in our highly realistic scheme with 10% of outliers, the average bias caused by outliers quickly increases and compromises the connectivity analyses if data is not processed robustly.

Non-parametric statistical analysis. The Friedman's test reported also p-value less than 0.01 therefore providing additional support for the graphical analysis and ANOVA results. We applied the two-sample Kolmogorov-Smirnov test to detect which of the distributions were different. Comparison between Baseline and normal COMMIT as well as COMMIT_r and COMMIT resulted in statistically significant differences between distributions with p-values less than 0.01. Respective effect sizes measured with Kolmogorov-Smirnovs statistic were 0.84 and 0.85. The difference between Baseline and robust COMMIT_r was not statistically significant with p-value slightly above 0.05 and effect size of 0.19 being

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

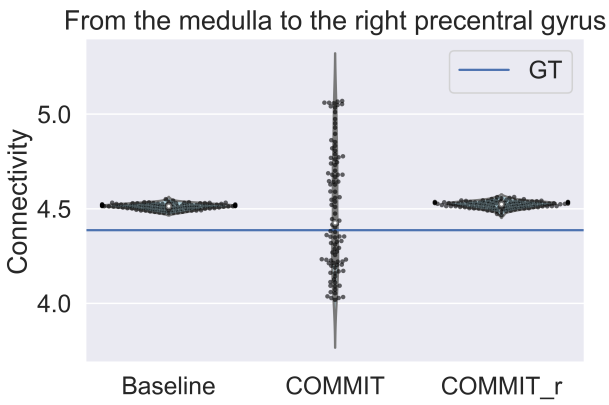

Figure 6: The differences between the ground-truth connectivty and each Monte-Carlo simulated connectivity from medulla to right precentral gyrus. The noise effect on this connection causes Baseline to deviate from the ground-truth value. The outlier effect is far more prominent than noise in normal COMMIT results but leads to a situation where the group average is closer to the ground-truth value than for the other two methods. This is the reason why comparisons between Monte-Carlo simulations were performed against Baseline instead of the ground-truth.

nearly 80% smaller compared to the non-robust counterpart.

Individual Connectivity

We investigated the element-wise differences between the Monte-Carlo connectivity matrices with parametric and non-parametric statistics as complementary information to the global results. The three violin plots in Fig. 6 depict the connectivity values from medulla to the right precentral gyrus. These streamlines are visualised in Fig. 7 and are likely a part of the corticospinal tract and therefore a known true connection.

The noise effect results in a systematic over estimation of the connectivity strength as depicted by Baseline in Fig. 6. However, outliers have a more random effect depending on the affected dMRI measurements. This can either decrease or increase the connectivity strength and can counteract the noise effect. Therefore, group comparisons against Baseline were more meaningful than comparisons against the known ground-truth value. For example, in this case the normal COMMIT produces an average connectivity strength that is closer to the ground-truth than Baseline despite the distribution is wider. Therefore, a simple observation of root mean squared difference between the Monte-Carlo group and ground-truth could be misleading.

Parametric statistical analyses. The connectivity-wise differences between Baseline and normal COMMIT as well as Baseline and robust COMMIT_r are shown in Fig. 8. The color map indicates the effect size measured with Cohen's D. Only elements that were deemed significantly different (p-value less than 0.05) based on ANOVA and Tukey's HSD were drawn. The comparison between Baseline and normal

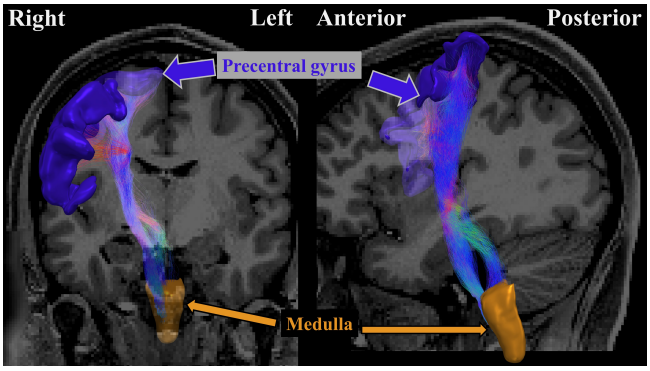

Figure 7: Illustration of a likely true structural connection between medulla and right precentral gyrus. Shape of the tractogram is nearly vertical therefore perpendicular to the introduced slice-wise outliers in the axial plane.

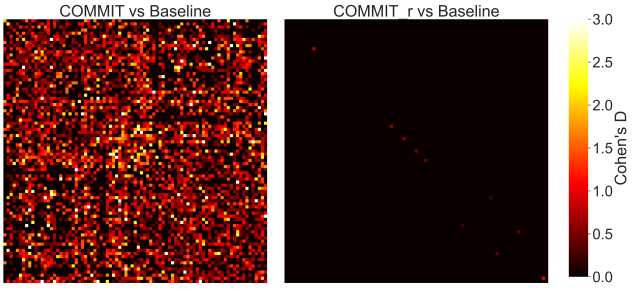

Figure 8: Differences due to outliers measured using parametric statistics. The comparison of the group averages of normal COMMIT and Baseline on the left and robust COMMIT_r and Baseline on the right. The color scale indicates the effect size calculated with Cohen's D. Results illustrate that the normal COMMIT is severely affected by outliers as differences to Baseline are large with effect sizes up to 5. On the contrary, the robust COMMIT_r can ameliorate the outlier effects producing similar results with the Baseline with overall smaller effect sizes.

COMMIT resulted in more elements with significant differences than the comparison between Baseline and COMMIT_r. The effect sizes between Baseline and normal COMMIT ranged from 0 up to 5 indicate that outliers can have strong adverse effects on specific connectivity matrix elements. The overall smaller effect sizes between Baseline and robust COMMIT_r highlight that our augmentation is well capable to mitigate the outlier effects even on individual connectivity level.

Non-parametric statistical analysis. The connectivity-wise distributional differences between Baseline and normal COMMIT as well as Baseline and robust COMMIT_r are shown in Fig. 9. The color map indicates the effect size measured with Kolmogorov-Smirnov statistic. Only elements that were deemed statistically significantly different (p-value less than 0.05) based on Kolmogorov-Smirnov tests were drawn. Similar to the parametric counterpart, the differences between Baseline and normal COMMIT were again more

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

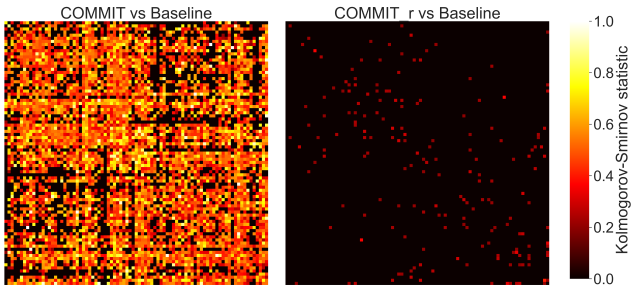

Figure 9: Differences due to outliers measured using non-parametric statistics. The comparison of the two-sample Kolmogorov-Smirnov test of normal COMMIT and Baseline on the left and robust COMMIT_r and Baseline on the right. The color scale indicates the effect size calculated with Kolmogorov-Smirnov statistic. Results illustrate that the connectivity specific distributions of normal COMMIT differ greatly from Baseline distributions due to outliers skewing them and leading to large effect sizes. The robust COMMIT_r produces similar distributions with Baseline with far less affected with effect sizes.

frequent than differences between Baseline and robust COMMIT_r. Also the effect sizes between Baseline and normal COMMIT ranged from 0 to nearly 1 which is the maximum of the used statistic. This indicates that outliers can lead to very large distributional differences. The differences between Baseline and robust COMMIT_r remained relatively small with effect sizes ranging from 0 to 0.2.

3.2. in vivo measurements

Besides tractogram filering, we calculated the intracellular and isotropic signal fractions calculated using the COMMIT framework (Daducci et al., 2015) and the proposed robust COMMIT_r. Fig. 10 shows the results for outlier detection, RMSE, and signal fraction maps obtain from the infant data. On average, the amount if missing data i.e. how much confidence in fitting was decreased per slice position ranged from 5% to 19%.

The RMSE map of normal COMMIT was clearly affected by the outliers resulting in visible stripes in the image. On the contrary, the COMMIT_r RMSE map that describes the robust cost function does not have such stripes therefore the fitting is not affected by outliers. The difference RMSE map visualises the stripy pattern more prominently and ranges from 0 to 30%. The outlier effect on intracellular and isotropic signal fractions was less prominent in visual analysis i.e. less or no stripes. However, the difference between normal COMMIT and robust COMMIT_r depicts that the differences ranged from -10% to +10% even in regions that were less affected by outliers for intracellular signal fraction. For isotropic signal fraction the differences ranged from -7% to +7%.

4. Discussion

We demonstrated that tractogram filtering is severely affected by subject motion artefacts and that with our proposed robust augmentation these effects can be mitigated. In clinical research with uncooperative patients such as infants, it is highly likely that motion to some degree occurs during scan. This leads to corrupted measurements which should not affect any modeling methods applied to the data. To best of our knowledge, this is the first time that motion related outliers are considered in the context of tractogram filtering therefore this update is crucial to enable tractogram filtering in clinical research.

The reason why we evaluated the proposed augmented cost function with simulated brains instead of real brain data was simply to ensure that nothing else in the relatively long dMRI processing pipeline might affect the results. For example, it is currently unknown issue, how outliers affect constrained spherical deconvolution based probabilistic tractographies. While there have been proposals for robust higher order model estimators (Pannek et al., 2012), such are not widely available. Furthermore, developing and evaluation of robustness of currently available constrained spherical deconvolution tractography algorithms are beyond the main scope of this study.

Comparison to other filtering methods

While similar weighted cost function as in eq. 1 has been proposed before in SIFT filtering algorithm (Smith et al., 2013), those have been designed and tested to account for partial voluming related artefacts - not subject motion. The main difference in these artefact types is that partial voluming affects all dMRI data whereas subject motion affects only part of the dMRI data randomly. Therefore, adjusting for partial voluming requires one three-dimensional reliability image whereas adjusting for subject motion requires four-dimensional reliability image as the measurement reliability must be accounted for each dMRI data separately. This difference in the implementations of the algorithms also makes the accurate comparison of them fall outside the scope of this study. While outlier replacement (Lauzon et al., 2013; Andersson et al., 2016) could seem beneficial to correct for subject motion, they actually cannot increase the available information but they could lead to difficulties in the interpretation and comparison of the obtained results (Sairanen et al., 2018).

Correcting for artefacts

Our proposed algorithm (Fig. 1 can also be used to adjust for partial voluming but the necessity of that depends on the forward model used in COMMIT. For example, with ball and sticks model, voxels containing cerebrospinal fluid or gray matter can be described with an increased contribution from a ball compartment therefore the contribution of a stick compartment could be correct even without additional reliability weighting. If reliability weights are used, then the estimate for ball compartment would likely be improved but that should not still affect the filtered tractogram.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

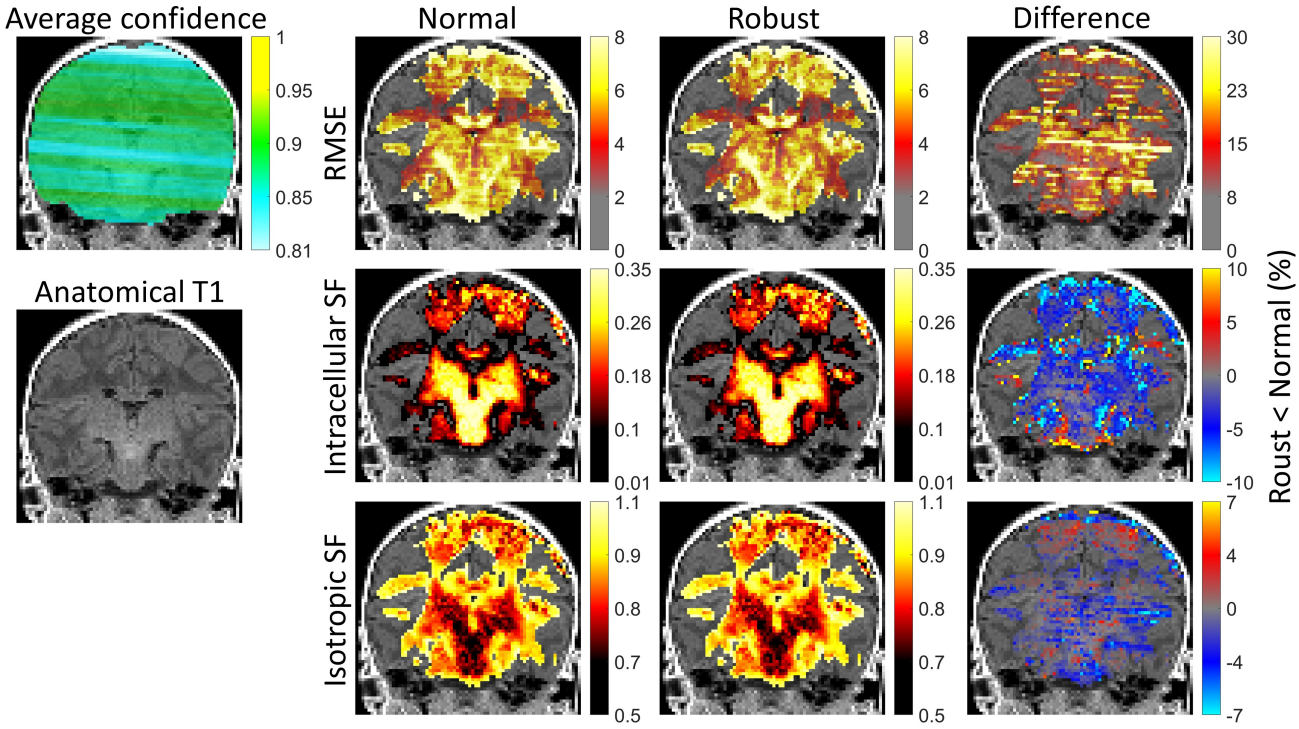

Figure 10: Summary of the in vivo tests. Coronal images are used to visualize the slicewise artefacts that typically occur in axial plane. Due to subject motion correction, the original axial plane can be rotated therefore the final outlier slices can also be oblique. An average confidence map and T1-image are shown on the left to visualize the artefactual regions in the measurements. The three columns on the right show results from normal COMMIT, robust COMMIT_r and their difference. The first row details the results for root mean squared error, the second row for intracellular signal fractions, and the third row for isotropic signal fractions.

With motion induced artefacts, the outliers cause anisotropic signal deviations (Sairanen et al., 2017) affecting only part of the dMRI datas. Therefore, COMMIT cannot adjust for those deviations simply by increasing the contribution of the ball compartment as the the deviations are not isotropic over dMRI measurements. This is demonstrated in Fig. 10 where normal COMMIT obtains incorrect estimates for isotropic signal fraction maps i.e. ball compartments. Issue propagates causing also incorrect estimates for intracellular signal fraction maps i.e. stick contributions. Therefore, a local motion artefact can have a global adverse effect in tractography filtering if not accounted for.

Computational speed

We evaluated our algorithm with the CPU version of COMMIT using a PC with ten 3.6 GHz cores. The 300 whole brain tractogram filterings with HCP like data in our Monte-Carlo simulations required approximately a week to run. Thus, one whole brain tractogram filtering took approximately 30 minutes.

Statistical analysis

The global connectivity difference (Fig. 5) between Baseline and robust COMMIT_r assessed with two-sample Kolmogorov-Smirnov was nearly statistically significant with p-value slightly above 0.05. It is possible that the amount of simulated outliers (%10) was already reaching the limit after which the missing data problem becomes too severe even for robust modeling methods. This could also be related to sample size being so large that Kolmogorov-Smirnov test finds any differences statistically significant despite having relatively small effect sizes.

A more in-depth analysis of the connection from medulla to the right precentral gyrus (Figs. 6 and 7) revealed that ANOVA failed to find statistically significant differences between the groups with a p-value of 0.4. On the contrary, the non-parametric Friedman's tests indicated that differences existed between the groups with a p-value less than 0.01. The effect sizes measured using Kolmogorov-Smirnov statistic between Baseline and normal COMMIT was 0.54 and between Baseline and robust COMMIT_r was 0.33. Both differences were found statistically significant by the Kolmogorov-Smirnov test with p-values less than 0.01.

In summary, it remains unsolved what test statistic would be the most suitable to analyse such data that is affected by outliers in anisotropic manner. We used two alternative approaches to evaluate the differences in group averages (ANOVA) and group distributions (Kolmogorov-Smirnov). Average

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

based analyses are likely inefficient to locate all differences arising from outliers in the data whereas non-parameteric test can be even too sensitive to label finding significant. Therefore, instead of statistical significance, the obtained effect sizes are likely more meaningful results.

Where to go from here?

We considered only post-scan motion corrections in this study because during-scan corrections should be able to produce data that does not need these correction algorithms. The problem with during-scan corrections is their limited availability due to external hardware requirements or still experimental software. Due to the long time span of tens of years required to advance MRI technology in clinical use, it is unlikely that these during-scan correction methods would be so widely available in clinical research centers that post-scan corrections such as our proposal are rendered obsolete any time soon. While the post-scan corrections are more like a remedy to the symptom instead of cure to the cause, novel studies on clinical patients and even infants are increasingly proposed and carried out therefore the need for robust tools is current and cannot wait decades for hardware based solutions.

5. Conclusion

We proposed a augmentation to a tractogram filtering algorithm COMMIT that renders it robust towards subject motion caused outliers in the measurements. This addition is necessary for conducting tractogram filtering in clinical research where subject motion is often unavoidable. While robust data processing has been implemented before in the context of diffusion tensor and higher order model estimations, it has not been previously implemented for tractogram filtering. We used highly realistic whole brain Monte-Carlo simulations and successfully demonstrated that our augmentation is capable to accurately map the structural brain connectivity in the presence of such outliers in the data. We also demonstrated that if this correction is not done, the structural connectivity estimates can become extremely biased. With this update any clinical study investigating structural connectomics of children or uncooperative patient populations can robustly perform their analyses without the need to exclude subjects with outliers from them.

Acknowledgements V.S. was supported by the Brain Research Foundation Verona. M.O-P., C.G., S.S. and A.D have no relevant financial or non-financial interests to disclose.

References

Alexander, D.C., Dyrby, T.B., Nilsson, M., Zhang, H., 2019. Imaging brain microstructure with diffusion MRI: practicality and applications. NMR in Biomedicine 32, e3841. URL: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3841, doi:https://doi.org/10.1002/nbm.3841. _eprint: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.3841.

Andersson, J.L., Graham, M.S., Zsoldos, E., Sotiropoulos, S.N., 2016. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. NeuroImage 141, 556–572. doi:10.1016/j.neuroimage.2016.06.058.

Basser, P.J., Mattiello, J., LeBihan, D., 1994. MR diffusion tensor spectroscopy and imaging. Biophysical Journal 66, 259–267. URL: http://www.ncbi.nlm.nih.gov/pubmed/8130344, doi:10.1016/S0006-3495(94)80775-1. publisher: The Biophysical Society.

Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magnetic Resonance in Medicine 44, 625–632. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/1522-2594%28200010%2944%3A4%3C625%3A%3AAID-MRM17%3E3.0.CO%3B2-O, doi:https://doi.org/10.1002/1522-2594(200010)44:4<625::AID-MRM17>3.0.CO;2-O. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1522-2594%28200010%2944%3A4%3C625%3A%3AAID-MRM17%3E3.0.CO%3B2-O.

Benitez, A., Fieremans, E., Jensen, J.H., Falangola, M.F., Tabesh, A., Ferris, S.H., Helpern, J.A., 2014. White matter tract integrity metrics reflect the vulnerability of late-myelinating tracts in Alzheimer's disease. NeuroImage: Clinical 4, 64–71. URL: https://www.sciencedirect.com/science/article/pii/S2213158213001496, doi:10.1016/j.nicl.2013.11.001.

Chang, L.C., Jones, D.K., Pierpaoli, C., 2005. RESTORE: Robust estimation of tensors by outlier rejection. Magnetic Resonance in Medicine 53, 1088–1095. URL: http://doi.wiley.com/10.1002/mrm.20426, doi:10.1002/mrm.20426. publisher: Wiley Subscription Services, Inc., A Wiley Company.

Chang, L.C., Walker, L., Pierpaoli, C., 2012. Informed RESTORE: A method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. Magnetic Resonance in Medicine 68, 1654–1663. doi:10.1002/mrm.24173.

Daducci, A., Palù, A.D., Lemkaddem, A., Thiran, J., 2015. COMMIT: Convex Optimization Modeling for Microstructure Informed Tractography. IEEE Transactions on Medical Imaging 34, 246–257. doi:10.1109/TMI.2014.2352414. conference Name: IEEE Transactions on Medical Imaging.

Delettre, C., Messé, A., Dell, L.A., Foubet, O., Heuer, K., Larrat, B., Meriaux, S., Mangin, J.F., Reillo, I., de Juan Romero, C., Borrell, V., Toro, R., Hilgetag, C.C., 2019. Comparison between diffusion MRI tractography and histological tract-tracing of cortico-cortical structural connectivity in the ferret brain. Network Neuroscience 3, 1038–1050. URL: https://doi.org/10.1162/netn_a_00098, doi:10.1162/netn_a_00098.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980. URL: https://www.sciencedirect.com/science/article/pii/S1053811906000437, doi:10.1016/j.neuroimage.2006.01.021.

Drakesmith, M., Caeyenberghs, K., Dutt, A., Lewis, G., David, A.S., Jones, D.K., 2015. Overcoming the effects of false positives and threshold bias in graph theoretical analyses of neuroimaging data. NeuroImage 118, 313–333. URL: https://www.sciencedirect.com/science/article/pii/S1053811915003912, doi:10.1016/j.neuroimage.2015.05.011.

Fieremans, E., Benitez, A., Jensen, J.H., Falangola, M.F., Tabesh, A., Deardorff, R.L., Spampinato, M.V.S., Babb, J.S., Novikov, D.S., Ferris, S.H., Helpern, J.A., 2013. Novel White Matter Tract Integrity Metrics Sensitive to Alzheimer Disease Progression. American Journal of Neuroradiology 34, 2105–2112. URL: http://www.ajnr.org/content/34/11/2105, doi:10.3174/ajnr.A3553. publisher: American Journal of Neuroradiology Section: Brain.

Fischl, B., 2012. FreeSurfer. NeuroImage 62, 774–781. URL: https://www.sciencedirect.com/science/article/pii/S1053811912000389, doi:10.1016/j.neuroimage.2012.01.021.

Genc, S., Malpas, C.B., Holland, S.K., Beare, R., Silk, T.J., 2017. Neurite density index is sensitive to age related differences in the developing brain. NeuroImage 148, 373–380. URL: https://www.sciencedirect.com/science/article/pii/S105381191730023X, doi:10.1016/j.neuroimage.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2017.01.023.

Griffa, A., Baumann, P.S., Thiran, J.P., Hagmann, P., 2013. Structural connectomics in brain diseases. NeuroImage 80, 515–526. URL: https://www.sciencedirect.com/science/article/pii/S1053811913004035, doi:10.1016/j.neuroimage.2013.04.056.

Horsfield, M.A., Jones, D.K., 2002. Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - a review. NMR in Biomedicine 15, 570–577. URL: http://doi.wiley.com/10.1002/nbm.787, doi:10.1002/nbm.787. publisher: John Wiley & Sons, Ltd.

Huber, E., Henriques, R.N., Owen, J.P., Rokem, A., Yeatman, J.D., 2019. Applying microstructural models to understand the role of white matter in cognitive development. Developmental Cognitive Neuroscience 36, 100624. URL: https://www.sciencedirect.com/science/article/pii/S1878929318301348, doi:10.1016/j.dcn.2019.100624.

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. URL: https://www.sciencedirect.com/science/article/pii/S1053811914006442, doi:10.1016/j.neuroimage.2014.07.061.

Kamiya, K., Hori, M., Aoki, S., 2020. NODDI in clinical research. Journal of Neuroscience Methods 346, 108908. URL: https://www.sciencedirect.com/science/article/pii/S0165027020303319, doi:10.1016/j.jneumeth.2020.108908.

Kunz, N., Zhang, H., Vasung, L., O'Brien, K.R., Assaf, Y., Lazeyras, F., Alexander, D.C., Hüppi, P.S., 2014. Assessing white matter microstructure of the newborn with multi-shell diffusion MRI and biophysical compartment models. NeuroImage 96, 288–299. URL: https://www.sciencedirect.com/science/article/pii/S1053811914002183, doi:10.1016/j.neuroimage.2014.03.057.

Lauzon, C.B., Asman, A.J., Esparza, M.L., Burns, S.S., Fan, Q., Gao, Y., Anderson, A.W., Davis, N., Cutting, L.E., Landman, B.A., 2013. Simultaneous Analysis and Quality Assurance for Diffusion Tensor Imaging. PLoS ONE 8. URL: http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0061737&type=printable, doi:10.1371/journal.pone.0061737.

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, 3537–3537. URL: http://www.mendeley.com/research/exploredti-a-graphical-toolbox-for-processing-analyzing-and-visualizing-diffusion-mr-data/%5Cnhttp://www.exploredti.com/ref/ExploreDTI_ISMRM_2009.pdf.

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., He, R., 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. URL: https://www.nature.com/articles/s41467-017-01285-x, doi:10.1038/s41467-017-01285-x. number: 1 Publisher: Nature Publishing Group.

Novikov, D.S., Fieremans, E., Jespersen, S.N., Kiselev, V.G., 2019. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. NMR in Biomedicine 32, e3998. URL: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3998, doi:https://doi.org/10.1002/nbm.3998. _eprint: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.3998.

Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., Johnson, H.J., Styner, M., 2014. DTIPrep: quality control of diffusion-weighted images. Frontiers in Neuroinformatics 8, 4–4. URL: http://journal.frontiersin.org/article/10.3389/fninf.2014.00004/abstract, doi:10.3389/fninf.2014.00004.

Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C., 2012. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. NeuroImage 59, 2241–2254. URL: https://www.sciencedirect.com/science/article/pii/S1053811911011566, doi:10.1016/j.neuroimage.2011.09.081.

Pannek, K., Fripp, J., George, J.M., Fiori, S., Colditz, P.B., Boyd, R.N., Rose, S.E., 2018. Fixel-based analysis reveals alterations is brain microstructure and macrostructure of preterm-born infants at term equivalent age. NeuroImage: Clinical 18, 51–59. URL: https://www.sciencedirect.com/science/article/pii/S2213158218300032, doi:10.1016/j.nicl.2018.01.003.

Pannek, K., Raffelt, D., Bell, C., Mathias, J.L., Rose, S.E., 2012. HOMOR: Higher Order Model Outlier Rejection for high b-value MR diffusion data. NeuroImage 63, 835–842. URL: http://ac.els-cdn.com/S1053811912007331/1-s2.0-S1053811912007331-main.pdf?_tid=a24e94d6-66e5-11e7-8034-00000aab0f02&acdnat=1499852391_e4092d37b6ed2c857b0f48df076fe8d2, doi:10.1016/j.neuroimage.2012.07.022.

Pecheva, D., Tournier, J.D., Pietsch, M., Christiaens, D., Batalle, D., Alexander, D.C., Hajnal, J.V., Edwards, A.D., Zhang, H., Counsell, S.J., 2019. Fixel-based analysis of the preterm brain: Disentangling bundle-specific white matter microstructural and macrostructural changes in relation to clinical risk factors. NeuroImage: Clinical 23, 101820. URL: https://www.sciencedirect.com/science/article/pii/S2213158219301706, doi:10.1016/j.nicl.2019.101820.

Perrone, D., Aelterman, J., Pižurica, A., Jeurissen, B., Philips, W., Leemans, A., 2015. The effect of Gibbs ringing artifacts on measures derived from diffusion MRI. NeuroImage 120, 441–455. URL: https://linkinghub.elsevier.com/retrieve/pii/S105381191500573X, doi:10.1016/j.neuroimage.2015.06.068.

Sairanen, V., Kuusela, L., Sipilä, O., Savolainen, S., Vanhatalo, S., 2017. A novel measure of reliability in Diffusion Tensor Imaging after data rejections due to subject motion. NeuroImage 147, 57–65. URL: http://ac.els-cdn.com/S1053811916306796/1-s2.0-S1053811916306796-main.pdf?_tid=68194a88-557c-11e7-a000-00000aacb35e&acdnat=1497938027_3755fe60a57d3726c9123661f80f7e9f, doi:10.1016/j.neuroimage.2016.11.066.

Sairanen, V., Leemans, A., Tax, C.M.W., 2018. Fast and accurate Slicewise OutLIer Detection (SOLID) with informed model estimation for diffusion MRI data. NeuroImage 181, 331–346. doi:10.1016/j.neuroimage.2018.07.003.

Sairanen, V., Ocampo-Pineda, M., Granziera, C., Schiavi, S., Daducci, A., 2021. Enhancing reliability of structural brain connectivity with outlier adjusted tractogram filtering, in: IEEE 18th International Symposium on Biomedical Imaging (ISBI), IEEE.

Samani, Z.R., Alappatt, J.A., Parker, D., Ismail, A.A.O., Verma, R., 2019. QC-Automator: Deep Learning-Based Automated Quality Control for Diffusion MR Images. Frontiers in Neuroscience 13, 1456. doi:10.3389/fnins.2019.01456.

Sheskin, D.J., 2004. Handbook of parametric and nonparametric statistical procedures. 2 ed., Chapman & hall/CRC. Boca Raton, FL.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2012. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. NeuroImage 62, 1924–1938. URL: https://www.sciencedirect.com/science/article/pii/S1053811912005824, doi:10.1016/j.neuroimage.2012.06.005.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2013. SIFT: Spherical-deconvolution informed filtering of tractograms. NeuroImage 67, 298–312. URL: https://www.sciencedirect.com/science/article/pii/S1053811912011615, doi:10.1016/j.neuroimage.2012.11.049.

Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2015. SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. NeuroImage 119,

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.09.447697; this version posted June 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

338–351. URL: https://www.sciencedirect.com/science/article/pii/S1053811915005972, doi:10.1016/j.neuroimage.2015.06.092.

Tax, C.M., 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. URL: http://www.ncbi.nlm.nih.gov/pubmed/24687400, doi:10.1002/mrm.25165.

Thomas, C., Ye, F.Q., Irfanoglu, M.O., Modi, P., Saleem, K.S., Leopold, D.A., Pierpaoli, C., 2014. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proceedings of the National Academy of Sciences 111, 16574–16579. URL: https://www.pnas.org/content/111/46/16574, doi:10.1073/pnas.1405672111. publisher: National Academy of Sciences Section: Biological Sciences.

Tournier, J.D., Calamante, F., Connelly, A., 2010. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions , 1.

Tournier, J.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A., 2019. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. NeuroImage 202, 116137. URL: https://www.sciencedirect.com/science/article/pii/S1053811919307281, doi:10.1016/j.neuroimage.2019.116137.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn Human Connectome Project: An overview. NeuroImage 80, 62–79. URL: https://www.sciencedirect.com/science/article/pii/S1053811913005351, doi:10.1016/j.neuroimage.2013.05.041.

Veraart, J., Sijbers, J., Sunaert, S., Leemans, A., Jeurissen, B., 2013. Weighted linear least squares estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls. NeuroImage 81, 335–346. URL: http://www.ncbi.nlm.nih.gov/pubmed/23684865%5Cnhttp://ac.els-cdn.com/S1053811913005223/1-s2.0-S1053811913005223-main.pdf?_tid=55f8cfae-2b2a-11e6-9dbe-00000aab0f01&acdnat=1465137331_fea523c85254148e1704c9a54f576827, doi:10.1016/j.neuroimage.2013.05.028.

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. Magnetic Resonance in Medicine 77, 285–299. URL: http://doi.wiley.com/10.1002/mrm.26124, doi:10.1002/mrm.26124.

Yeh, C.H., Jones, D.K., Liang, X., Descoteaux, M., Connelly, A., 2020. Mapping Structural Connectivity Using Diffusion MRI: Challenges and Opportunities. Journal of Magnetic Resonance Imaging , 1–17doi:10.1002/jmri.27188.

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L.L., van den Heuvel, M.P., Breakspear, M., 2016. Connectome sensitivity or specificity: which is more important? NeuroImage 142, 407–420. URL: https://www.sciencedirect.com/science/article/pii/S1053811916302841, doi:10.1016/j.neuroimage.2016.06.035.

Zhang, F., Daducci, A., He, Y., Schiavi, S., Seguin, C., Smith, R., Yeh, C.H., Zhao, T., O'Donnell, L.J., 2021. Quantitative mapping of the brain's structural connectivity using diffusion MRI tractography: a review. arXiv:2104.11644 [q-bio] URL: http://arxiv.org/abs/2104.11644. arXiv: 2104.11644.

Zöllei, L., Iglesias, J.E., Ou, Y., Grant, P.E., Fischl, B., 2020. Infant FreeSurfer: An automated segmentation and surface extraction pipeline for T1-weighted neuroimaging data of infants 0–2 years. NeuroImage 218, 116946. URL: https://www.sciencedirect.com/science/article/pii/S1053811920304328, doi:10.1016/j.neuroimage.2020.116946.