



Tractography passes the test: Results from the diffusion-simulated connectivity (disco) challenge

Gabriel Girard^{a,b,c,1,*}, Jonathan Rafael-Patiño^{b,c,1}, Raphaël Truffet^d, Dogu Baran Aydogan^{e,f,g}, Nagesh Adluru^{h,i}, Veena A. Nairⁱ, Vivek Prabhakaranⁱ, Barbara B. Bendlin^j, Andrew L. Alexander^{h,k,l}, Sara Bosticardo^{m,n}, Iliaria Gabusi^{m,o}, Mario Ocampo-Pineda^m, Matteo Battocchio^{m,p}, Zuzana Piskorova^{m,q}, Pietro Bontempi^m, Simona Schiavi^r, Alessandro Daducci^m, Aleksandra Stafiej^s, Dominika Ciupek^t, Fabian Bogusz^s, Tomasz Pieciak^{s,u}, Matteo Frigo^v, Sara Sedlar^v, Samuel Deslauriers-Gauthier^v, Ivana Kojčić^v, Mauro Zucchelli^v, Hiba Laghrissi^{v,ao}, Yang Ji^v, Rachid Deriche^v, Kurt G Schilling^w, Bennett A. Landman^{w,x}, Alberto Cacciola^{y,am,an}, Gianpaolo Antonio Basile^y, Salvatore Bertino^y, Nancy Newlin^x, Praitayini Kanakaraj^x, Francois Rheault^x, Patryk Filipiak^z, Timothy M. Shepherd^z, Ying-Chia Lin^z, Dimitris G. Placantonakis^{aa}, Fernando E. Boada^{ab}, Steven H. Baete^z, Erick Hernández-Gutiérrez^p, Alonso Ramírez-Manzanares^{ac}, Ricardo Coronado-Leija^z, Pablo Stack-Sánchez^{ac}, Luis Concha^{ad}, Maxime Descoteaux^p, Sina Mansour L.^{ae,af}, Caio Seguin^{af,ag,ah}, Andrew Zalesky^{ae,af}, Kenji Marshall^{c,ai}, Erick J. Canales-Rodríguez^c, Ye Wu^{aj,ak}, Sahar Ahmad^{aj}, Pew-Thian Yap^{aj}, Antoine Théberge^p, Florence Gagnon^p, Frédéric Massi^p, Elda Fischì-Gomez^{a,b,c}, Rémy Gardier^c, Juan Luis Villarreal Haro^c, Marco Pizzolato^{c,al}, Emmanuel Caruyer^d, Jean-Philippe Thiran^{a,b,c}

^a CIBM Center for Biomedical Imaging, Switzerland

^b Radiology Department, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland

^c Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

^d Univ Rennes, Inria, CNRS, Inserm, IRISA UMR 6074, Empenn ERL U-1228, Rennes, France

^e A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland

^f Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland

^g Department of Psychiatry, Helsinki University Hospital, Helsinki, Finland

^h Waisman Center, University of Wisconsin-Madison, Madison, WI, United States

ⁱ Department of Radiology, University of Wisconsin-Madison, Madison, WI, United States

^j Department of Medicine, University of Wisconsin-Madison, Madison, WI, United States

^k Department of Medical Physics, University of Wisconsin-Madison, Madison, WI, United States

^l Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, United States

^m Diffusion Imaging and Connectivity Estimation (DICE) Lab, Department of Computer Science, University of Verona, Verona, Italy

ⁿ Translational Imaging in Neurology (ThInk), Department of Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland

^o Department of Advanced Biomedical Sciences, University of Naples Federico II, Naples, Italy

^p Sherbrooke Connectivity Imaging Laboratory (SCIL), Department of Computer Science, University of Sherbrooke, Sherbrooke, QC, Canada

^q Brno Faculty of Electrical Engineering and Communication, Department of mathematics, University of Technology, Brno, Czech Republic

^r Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINO-GM), University of Genoa, Genoa, Italy

^s AGH University of Science and Technology, Kraków, Poland

^t Sano Centre for Computational Personalised Medicine, Kraków, Poland

^u Laboratorio de Procesado de Imagen (LPI), ETSI Telecomunicación, Universidad de Valladolid, Valladolid, Spain

^v Athena Project Team, Centre Inria d'Université Côte d'Azur, France

^w Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center, Nashville, TN, United States

^x Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, United States

^y Brain Mapping Lab, Department of Biomedical, Dental Sciences and Morphological and Functional Images, University of Messina, Messina, Italy

^z Center for Advanced Imaging Innovation and Research (CAI2R), Department of Radiology, NYU Langone Health, New York, NY, United States

* Corresponding author.

E-mail addresses: gabriel.girard@epfl.ch (G. Girard), jonathan.patinolopez@epfl.ch (J. Rafael-Patiño).

¹ These authors contributed equally.

<https://doi.org/10.1016/j.neuroimage.2023.120231>.

Received 2 March 2023; Received in revised form 12 May 2023; Accepted 14 June 2023

Available online 16 June 2023.

1053-8119/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

^{aa} Department of Neurosurgery, Perlmutter Cancer Center, Neuroscience Institute, Kimmel Center for Stem Cell Biology, NYU Langone Health, New York, NY, United States

^{ab} Department of Radiology, Stanford University, Stanford, CA, United States

^{ac} Computer Science Department, Centro de Investigación en Matemáticas A.C., Guanajuato, México

^{ad} Instituto de Neurobiología, Universidad Nacional Autónoma de México, Juriquilla, Querétaro, México

^{ae} Department of Biomedical Engineering, The University of Melbourne, Parkville, Victoria, Australia

^{af} Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne and Melbourne Health, Parkville, Victoria, Australia

^{ag} School of Biomedical Engineering, The University of Sydney, Sydney, Australia

^{ah} Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, United States

^{ai} McGill University, Montréal, QC, Canada

^{aj} Department of Radiology and Biomedical Research Imaging Center (BRIC), The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

^{ak} School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China

^{al} Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark

^{am} Center for Complex Network Intelligence (CCNI), Tsinghua Laboratory of Brain and Intelligence (THBI), Tsinghua University, Beijing, China

^{an} Department of Biomedical Engineering, Tsinghua University, Beijing, China

^{ao} Institut de Biologie de Valrose, Université Côte d'Azur, Nice, France

ARTICLE INFO

Keywords:

Diffusion MRI
Connectivity
Monte carlo simulation
Tractography
Numerical substrates
Microstructure
Challenge

ABSTRACT

Estimating structural connectivity from diffusion-weighted magnetic resonance imaging is a challenging task, partly due to the presence of false-positive connections and the misestimation of connection weights. Building on previous efforts, the MICCAI-CDMRI Diffusion-Simulated Connectivity (DiSCo) challenge was carried out to evaluate state-of-the-art connectivity methods using novel large-scale numerical phantoms. The diffusion signal for the phantoms was obtained from Monte Carlo simulations. The results of the challenge suggest that methods selected by the 14 teams participating in the challenge can provide high correlations between estimated and ground-truth connectivity weights, in complex numerical environments. Additionally, the methods used by the participating teams were able to accurately identify the binary connectivity of the numerical dataset. However, specific false positive and false negative connections were consistently estimated across all methods. Although the challenge dataset doesn't capture the complexity of a real brain, it provided unique data with known macrostructure and microstructure ground-truth properties to facilitate the development of connectivity estimation methods.

1. Introduction

Over the last decade, protocols for diffusion-weighted magnetic resonance imaging (DW-MRI) acquisition, local modelling, tractography algorithms, and connectivity mapping methods have considerably improved (Jeurissen et al., 2017; Sotiropoulos and Zalesky, 2019; Sporns, 2011). However, concerns remain about the reliability of connectivity mapping. International tractography challenges (Côté et al., 2013; Fillard et al., 2011; Maffei et al., 2022; Maier-Hein et al., 2017; Nath et al., 2020) have shown limitations in the ability of tractography to correctly identify binary connectivity and identify white matter pathways consistently. In particular, Maier-Hein et al. (2017) showed that tractography may produce an abundance of false positive connections. Moreover, studies on animal models showed that, albeit tractography can correctly identify connections, the estimated connection weight does not always agree with *ex vivo* tracing data (Ambrosen et al., 2020a; Aydogan et al., 2018; Azadbakht et al., 2015; Delettre et al., 2019; Donahue et al., 2016; Essen et al., 2014; Girard et al., 2020; 2021; van den Heuvel et al., 2015; Jbabdi et al., 2013; Schilling et al., 2019a; Thomas et al., 2014). For instance, Donahue et al. (2016) reported the correlation between *ex vivo* tract tracing data and tractography estimation to be $r = 0.59$, on the intrahemispheric connections the monkey brain. Despite tract tracing being among the best available data to validate diffusion tractography, it is not possible to have the full ground-truth micro- and macro-structure on animal models.

The rich signal from physical MRI phantoms has been used to test and validate methods (Fillard et al., 2011; Schilling et al., 2019b), but their macrostructural complexity is insufficient for quantifying connectivity. Numerical phantoms have also been proposed and demonstrated to be important tools for methods development (Caruyer et al., 2014; Close et al., 2009; Neher et al., 2013), but their biological fidelity for microstructure is limited. Monte Carlo methods (Hall and Alexander, 2009; Lee et al., 2021; Rafael-Patino et al., 2020) can provide realistic microscopic DW-MRI signals, but they are generally limited to a single voxel signals or to a substrate of only a few voxels in size. Recently, Rafael-Patino et al. (2020) proposed a novel diffusion Monte Carlo sim-

ulator able to generate billions of particles. This allows for large-scale substrates with both microscopic and macroscopic complexity, suitable for structural connectivity validation.

The MICCAI-CDMRI 2021 Diffusion-Simulated Connectivity (DiSCo) challenge (Girard et al., 2021) was organized to compare structural connectivity estimation methods using three novel large-scale complex numerical phantoms designed for connectivity assessment (Rafael-Patino et al., 2021a; 2021b). Fourteen teams, adding up to 57 researchers, submitted 111 weighted connectivity matrices estimating the ground-truth connectivity. Results from the challenge are presented below.

2. Methods

2.1. Synthetic data

The three numerical phantoms (training, validation and test phantoms) used for the DiSCo challenge (Rafael-Patino et al., 2021b) are composed of approximately 12,000 numerical tubular fibres. The tubular fibres' outer diameter ranges from $2.0 \mu\text{m}$ to $6 \mu\text{m}$, sampled from a gamma distribution $\Gamma(\kappa, \theta)$, with shape, $\kappa = 0.5$, and scale $\theta = 0.007$. The inner diameter of each fibre ranges from $1.4 \mu\text{m}$ to $4.2 \mu\text{m}$, simulating a fixed g-ratio of 0.7 (Cercignani et al., 2017; Chomiak and Hu, 2009). The numerical fibres connect pairs of Regions of Interest (ROIs) among the 16 ROIs of each phantom (see Fig. 1A). No other numerical compartments were added to the substrates. For the three phantoms, the average percentage of connections with non-zero connection weight is 22.2% among all possible connections (120 pairs of ROIs). The connectivity weight between two ROIs was defined as the sum of the cross-sectional areas of fibres interconnecting the regions. The normalized connection weights range from 0.007 to 0.092 for the three phantoms, resulting with the smallest connection having 7.6% of the weight of the largest one. Those connection weights derive from the numerical phantom initialization parameters described in Rafael-Patino et al. (2021a). Fig. 1B, C, D show the ground-truth synthetic fibre trajectories of the test dataset where fibres are curved and intermingle with other fibres.

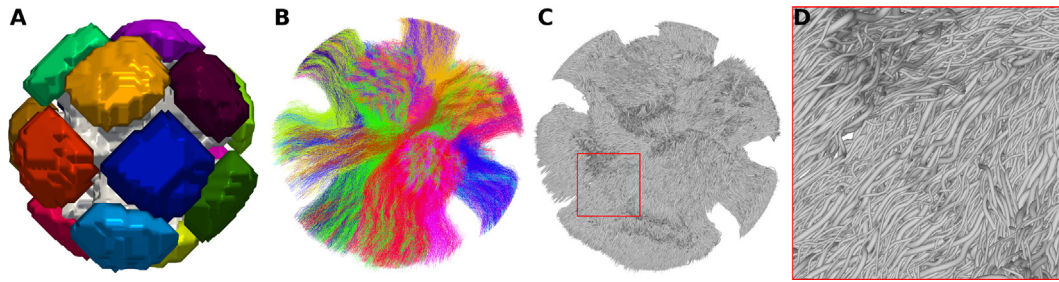


Fig. 1. Ground-truth test dataset composed of 11,032 numerical tubular fibres. (A) 3D rendering showing the synthetic white matter mask (gray) and the 16 ROIs (colors). (B) Trajectories of the fibres of the 26 bundles, each shown using a different color. (C-D) 3D mesh of the outer layer of numerical fibres.

The simulation substrates have an unprecedented volume of 1 cubic millimeter, resulting in an image size of $40 \times 40 \times 40$ voxels of $25 \mu\text{m}$ isotropic resolution. To the best of our knowledge, this is the largest image volume achieved for the Monte Carlo simulation of the DW-MRI signal in complex numerical substrates. Within each voxel, the signal was simulated using Monte Carlo simulations of spin dynamics with a density of one particle per cubic micrometer (Rafael-Patino et al., 2021a; Romascano et al., 2019). Rafael-Patino et al. (2020) showed this was a sufficient number of particles to obtain a robust estimation of the diffusion signal in complex fibre geometries. Particles initiated within the inner diameter of the fibres and outside the outer diameter of the fibres were used to generate the DW-MRI signal. The particles initiated between the outer and inner diameter (in myelin water) were discarded. The voxel-wise intra-tubular volume fraction reaches 52% in the central portion of the numerical phantoms (Rafael-Patino et al., 2021b). The mean voxel-wise fibre diameter is $2.25 \mu\text{m}$ with up to 82 tubular fibres per voxel and up to 5 distinct bundles (Rafael-Patino et al., 2021b).

The DW-MRI protocol is composed of 360 measurements, uniformly distributed over 4 b-value shells (1000, 1925, 3094, 13,191 s/mm^2), as suggested in ActiveAx (Alexander et al., 2010; Daducci et al., 2015), and corrupted with Rician noise with signal-to-noise ratio of 30. The resulting DW-MRI signal is affected by the microscopic properties of the synthetic white matter, such as fibre diameter, packing densities, fibre dispersion and water diffusing around fibres, while also having targeted macroscopic properties like the smoothness of the trajectories and fibres organized in bundles.

2.2. Challenge task

Participating teams had access to one dataset for training, which included the noisy and noiseless DW-MRI signals, the fibre volume fraction map, the label map of the ROIs defining the connectivity endpoints, the synthetic fibre trajectories and their diameter, and the ground-truth connectivity matrix. Additionally, participants had access to one dataset for validation with the noisy DW-MRI signal, label map and ground-truth connectivity matrix. Participating teams were provided with the noisy DW-MRI signal and a label map (ROIs) of the test dataset, and were asked to submit a 16×16 weighted connectivity matrix. Participants were free to select any method to compute the matrix weights. The estimated connection weights between any two pairs of ROIs were compared with the ground-truth total cross-sectional area of the synthetic fibres connecting both ROIs. The teams had to select methods to obtain estimates of the cross-sectional area from their tractography results, such as the proportion or volume of streamlines, or microstructure properties or geometrical features estimated for bundles (Assaf et al., 2008; Daducci et al., 2014; Dimitriadis et al., 2017; Hagmann et al., 2008; 2007; Messaritaki et al., 2019; Smith et al., 2015; Sotiropoulos and Zalesky, 2019; Tournier et al., 2019; Yeh et al., 2021). Teams could submit up to ten connectivity matrices.

2.3. Connectivity evaluation

The Pearson correlation coefficient (r) between the ground-truth matrix and the submitted matrices was used for ranking the teams (Caminiti et al., 2021; Donahue et al., 2016). Moreover, the fraction of valid connectivity weight was computed to compare submissions (Côté et al., 2013; Maier-Hein et al., 2017). This fraction corresponds to the sum of the matrix weights in pairs of regions connected in the ground-truth connectivity matrix divided by the sum of all weights. A Receiver Operating Characteristic (ROC) analysis was also performed (Ambrosen et al., 2020b; Girard et al., 2020; Maffei et al., 2022; Schilling et al., 2019b; Thomas et al., 2014). The true positives (TP) and true negatives (TN) are connections correctly identified as connected and not connected in both the participant matrix and the ground-truth matrix, respectively. The false positives (FP) are connections wrongly identified as connected in the participant matrix. Similarly, the false negatives (FN) are connections erroneously identified as not connected. The ROC curves were constructed by iteratively thresholding the submitted connectivity matrices, starting with a threshold higher than the maximum, thus yielding no pair of ROIs connected, resulting in a specificity ($\frac{TN}{TN+FP}$) of 1 and sensitivity ($\frac{TP}{TP+FN}$) of 0 (all pairs of ROIs not connected in the ground-truth are correctly identified, no ROIs are identified as connected). The threshold is then iteratively reduced until all ROIs are identified as connected, producing a sensitivity of 1. The quicker the sensitivity rises to 1 while the specificity remains high, the better the binary connectivity classification performance of the method. The Area Under the ROC Curve (AUC) summarizes the plot with a number between 0 and 1. The AUC approaches 1 if there are few or no classification errors (a random connectivity matrix would yield an AUC of 0.5). Moreover, we studied the accuracy ($\frac{TP+TN}{TP+TN+FP+FN}$) of the submitted matrices using a threshold selected as 5% of their maximum value. This threshold was fixed following the connectivity weights of the ground-truth matrix.

3. Results

Fourteen teams participated in the DiSCo challenge and submitted 111 connectivity matrices for the test dataset. Fig. 2A shows the Pearson correlation coefficient r between the participant's submitted matrices and the ground-truth connectivity matrix of the validation dataset. Fig. 2B shows the fraction of valid connectivity weight in pairs of connected regions (non-zero connection strength) in the ground-truth connectivity matrix. The best-performing matrix of each team ranges from $r = 0.874$ to $r = 0.973$ (mean $r = 0.950$). The area under the ROC curve (AUC), computed from the submitted matrices and the ground-truth binary connectivity matrix, is reported in Fig. 2C. Fig. 2D shows the accuracy of all methods when thresholding the submitted matrices at 5% of their maximum value. The ground-truth connectivity matrix of the test dataset is shown in Fig. 3, alongside each team's best-performing method (method with maximum r).

Fig. 4 shows the ROC curves for the best-performing methods of each team. The corresponding area under the curve (AUC) is reported in the

Challenge Submission Results

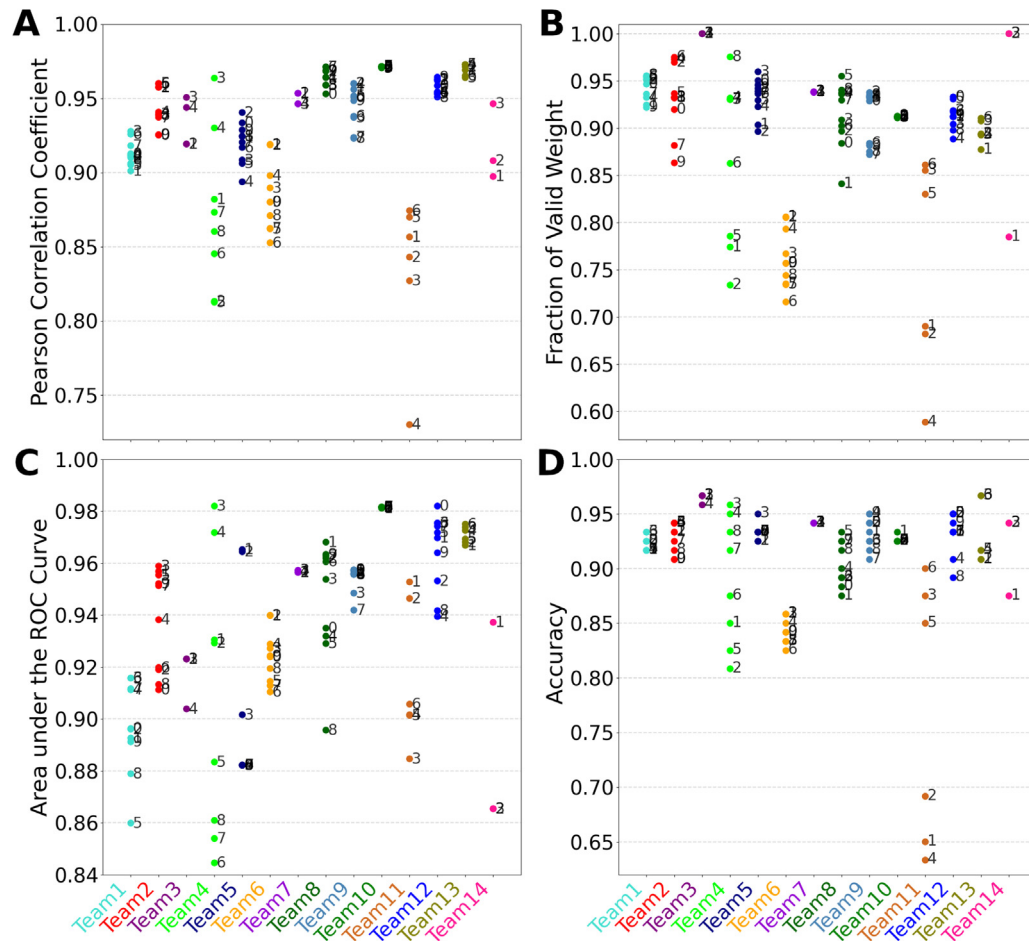


Fig. 2. Challenge submission results of the 14 participating teams (111 submissions). (A) Fraction of valid connectivity weight in pairs of regions connected in the ground-truth connectivity matrix. (B) Pearson correlation coefficient between the participant's submitted matrices and the ground-truth connectivity matrix of the validation dataset. (C) The area under the ROC curve (AUC) computed from the submitted matrices and the ground-truth binary connectivity matrix. (D) The accuracy (fraction of correctly identified pairs of ROIs, out of 120) of the binarised submitted matrices, thresholded at 5% of their maximal value. Numbers indicate the submission indices of each team.

legend, ranging from 0.865 to 0.982 (mean AUC=0.946). Fig. 5 shows the ground-truth binary connectivity matrix (top left) and each team's pairs of ROIs classifications. Matrices were thresholded at 5% of their maximal value. The light green and dark green colours show the true positives and true negatives, respectively. The light red and dark red colours show the false positives and the false negatives, respectively.

The percentage of classification error for each pair of ROIs for all submitted connectivity matrices is shown in Fig. 6. The left subfigure reports the false positive connections. The worst performance is reported for ROIs 5–11 and 4–6 with 73% and 71% of matrices erroneously identifying them as connected. The right subfigure reports the false negative connections, with ROIs 6–9, 4–16, and 3–14 showing the worst classification, with 100%, 97% and 95% of methods erroneously identifying them as not connected, respectively, although connected in the ground-truth. Fig. 7 shows the location of the false positive bundles connecting ROIs 5–11 (blue) and 4–6 (green). Both pairs of ROIs are spatially located next to each other. Fibre ODFs show the corresponding ground-truth numerical fibre distribution. Fig. 8 shows the false negative bundles connecting ROIs 6–9 (green), 4–16 (red), and 3–14 (blue). They are the bundle with the lowest, second lowest and 5th lowest connectivity in the ground-truth weighted connectivity matrix. All three bundles show long and straight configurations going through the centre of the numerical phantom.

Each team's best-performing method processing steps are listed in Table 1. All teams submissions are described in supplementary material.

4. Discussion

The aim of this work was to test tractography algorithms in carefully designed numerical phantoms with intricate connectivity patterns. The challenge was to identify connected pairs of ROIs among 16 ROIs and estimate their connection strength, defined as the cross-sectional area of the synthetic fibres interconnecting them. The DiSCO challenge phantoms were developed to feature challenging configurations found in the human brain, such as branching, crossing, and tortuous trajectories. Although these phantoms don't mimic the anatomy of the human brain, they provide valuable data for studying tractography and connectivity. As such, results obtained on the DiSCO dataset are not directly transferable to real brain data. Rather, they should be used to evaluate the relative performance among connectivity methods. Unlike traditional tractography numerical phantoms that use biophysical models, the DiSCO datasets were obtained from realistic Monte Carlo simulations. This approach allows for a signal with rich microstructure and complex and coherent macrostructure properties, suitable to study properties of connectivity methods.

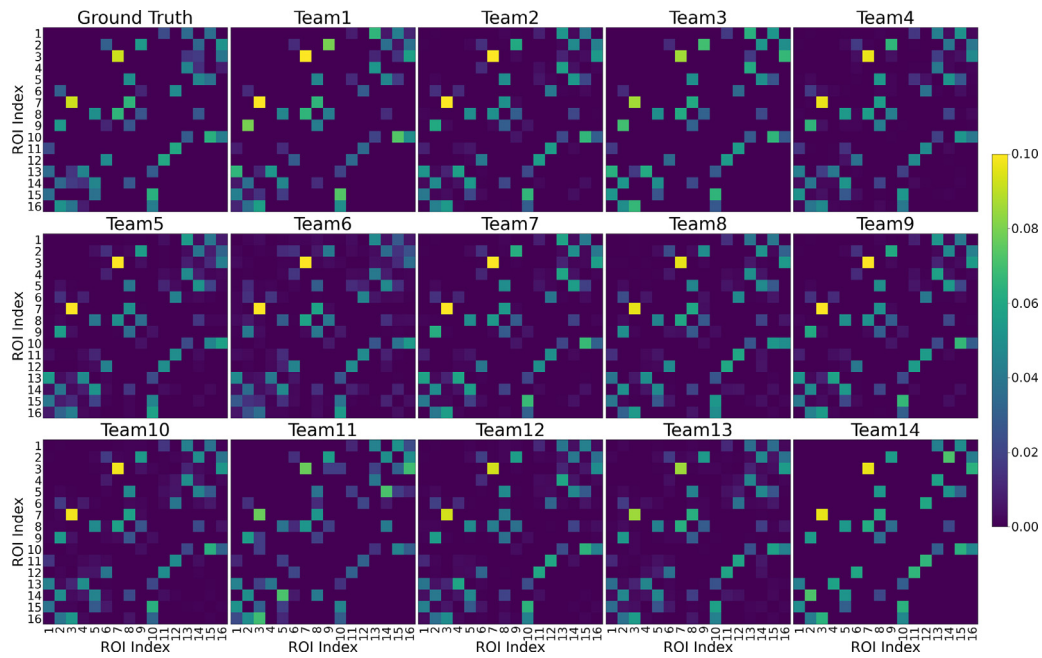


Fig. 3. The test dataset's ground-truth connectivity matrix (top left) and each team's best-performing classification matrices. All matrices are symmetric, and the upper triangular matrices are normalized to sum to one. The 26 non-zero connections of the test dataset have weights ranging from 0.008 to 0.092.

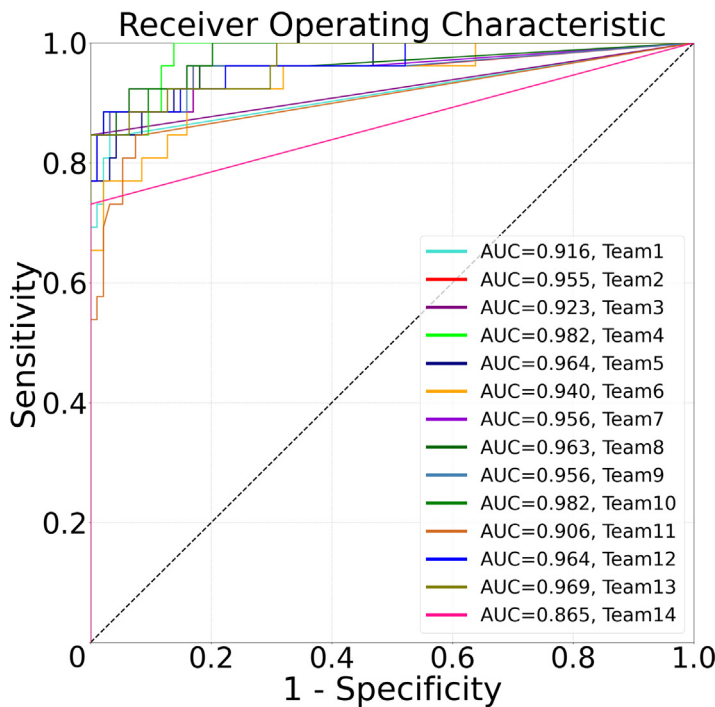


Fig. 4. Receiver Operating Characteristic (ROC) curves of the submitted matrix with the highest correlation for each team. The black dashed line shows the performance of a connectivity matrix with randomly generated weights. The corresponding area under the curve (AUC) is reported in the bottom right panel.

Participating teams did remarkably well, despite the known limitations of diffusion tractography methods (Jbabdi et al., 2015; Jones, 2010). This is shown by the large fraction of connection weight reported in the pair of ROIs connected in the ground-truth matrix (0.89 on average, see Fig. 2A). Methods generally showed high accuracy (average of 0.91) and high AUC (average of 0.95) for the identification of connected/non-connected ROIs (Fig. 2C,D). Overall, the mean Pearson's correlation coefficient across all submissions is $r = 0.95$, with a maximum of $r = 0.973$ (see Fig. 2B). Despite the macroscopic complexity of the numerical phantom, state-of-the-art tractography methods combined with state-of-the-art spherical deconvolution methods can cor-

rectly identify connected ROIs, producing connectivity results predominantly faithful to the numerical substrate.

4.1. Correlation coefficients with the ground-truth weights

The correlation coefficients obtained on numerical data are higher than those reported in brain connectivity studies comparing DW-MRI weights estimation and labelled cell counts from tracing studies in the intraparietal sulcus (Caminiti et al., 2021) ($r = 0.65$) and intrahemispheric (Donahue et al., 2016) ($r = 0.59$) connections. This highlights that the DiSCO numerical substrates oversimplifies the complexity of real

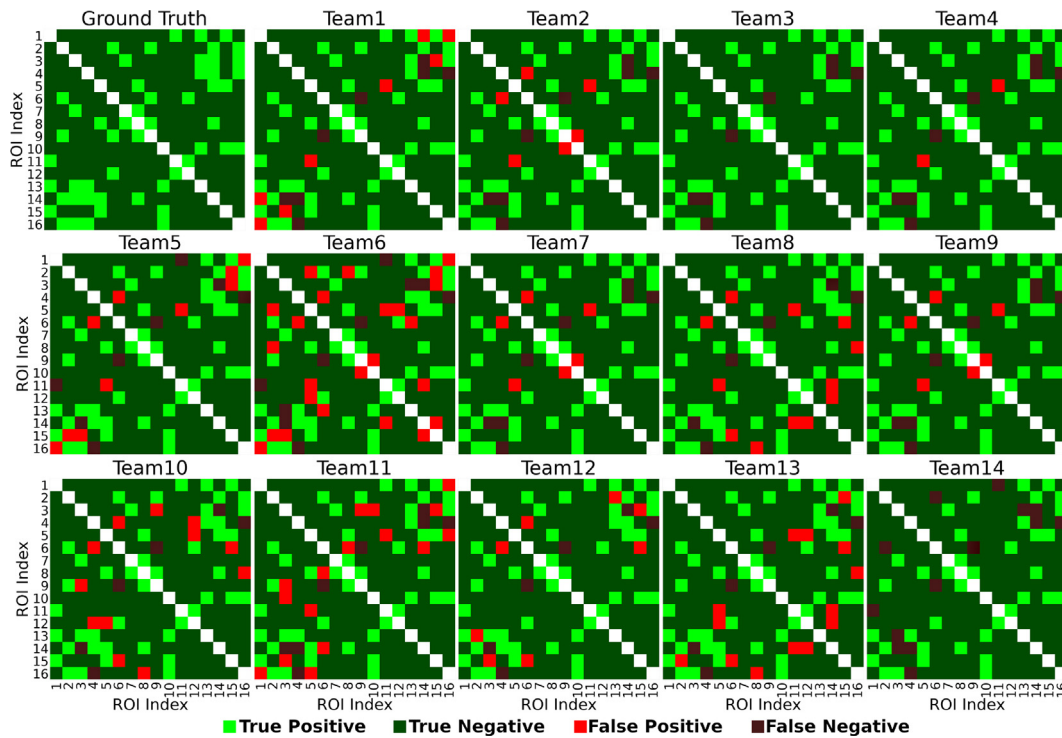


Fig. 5. The test dataset’s ground-truth binary connectivity matrix (top left) and each team’s matrices. All matrices were thresholded at 5% of their maximal value. The light/dark green and light/dark red colours show the true positives/negatives and false positives/negatives, respectively. All matrices are symmetric. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

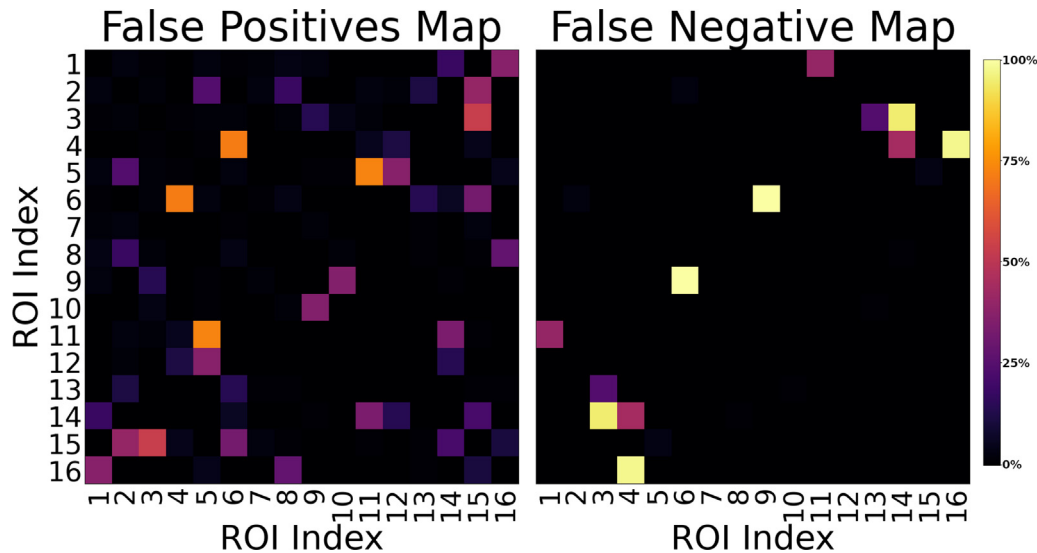


Fig. 6. Percentage of classification error for each pair of ROIs for the submitted matrices (111) and using the threshold at 5% of their maximal value. The left subfigure reports the false positive connections. Regions 5–11 and 4–6 show the worst performance, with 73% (81) and 71% (79) matrices erroneously identifying them connected. The right subfigure reports the false negative connections. Regions 6–9, 4–16, and 3–14 show the worst classification, with 100% (111), 97% (108) and 95% (105) of methods erroneously identifying them as not connected. Both matrices are symmetric.

MRI signals. Indeed, tractography limitations could originate from other factors aside from the diffusion information, such as MRI artifacts (B0 field inhomogeneity, susceptibility, motion, etc) and region-dependent T2 effects (Le Bihan et al., 2006). Despite the complexity achieved with the DiSCo numerical phantoms, real tissue shows a higher heterogeneity (Andersson et al., 2021) that was not reproduced, which may affect the relevance of some findings on biological tissue data. However, it is possible to know the ground-truth connectivity with higher accuracy than tracing studies, including the trajectory and diameter of the numerical

fibres and the voxelwise compartmental volume fractions. Future studies should investigate the effects of MRI artifacts and signal-to-noise ratio on the connectivity estimation. New numerical datasets should be generated with varying numbers of ROIs, ROI sizes, and connectivity strengths. This would allow testing DW-MRI connectivity estimation methods in diverse and complex environments, improving the generalizability of our results. In addition, other evaluation metrics, such as Dice similarity coefficient, could be used to test bundle volume identification using tractography. Moreover, research should be done on com-

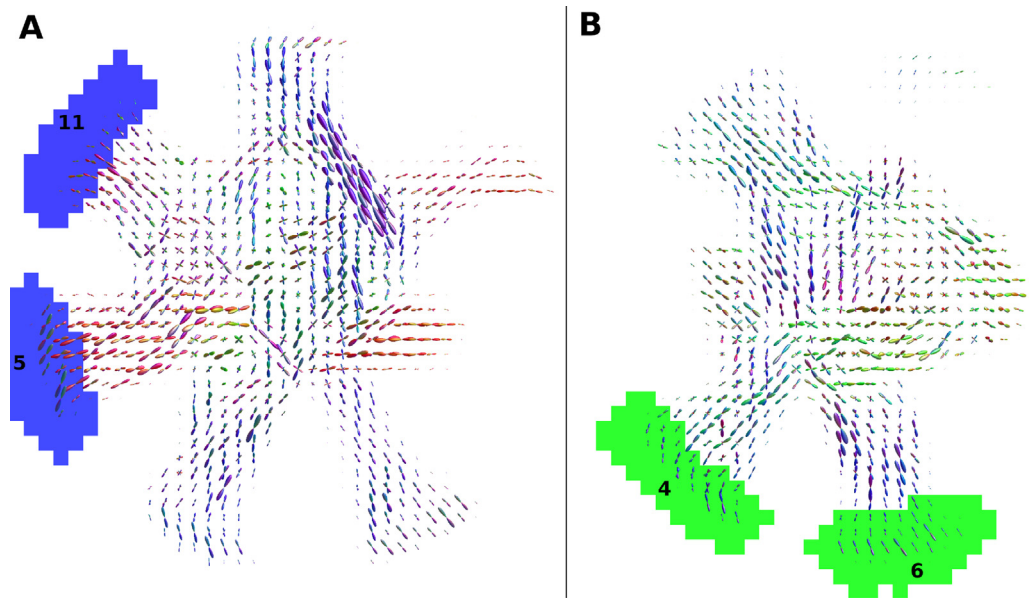


Fig. 7. False positive bundles connecting ROIs 5–11 (A, blue) and 4–6 (B, green). These 2 pairs of regions have been incorrectly identified as connected by 73% and 71% of the submitted matrices, using a threshold at 5% of their maximal value, respectively. Glyphs show the local orientations of the ground-truth tubular fibres intersecting voxels, coloured with their orientation (left-right: red, anterior-posterior: green, superior-inferior: blue). Both pairs of regions are spatially located next to each other. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

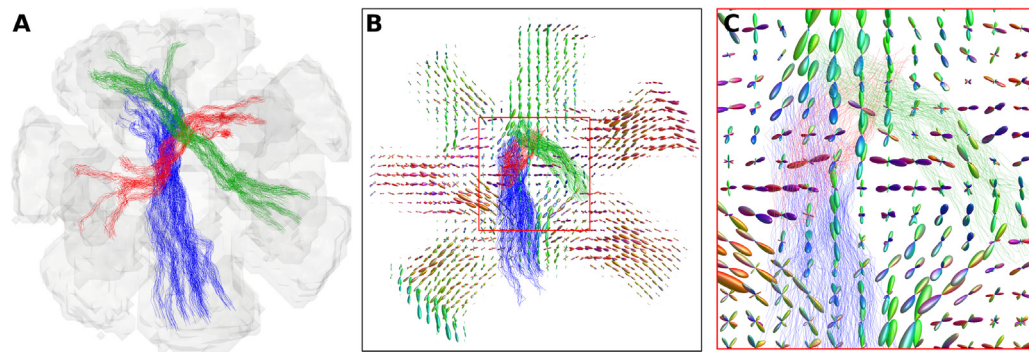


Fig. 8. False negative bundles connecting ROIs 6–9 (green), 4–16 (red), and 3–14 (blue), were erroneously reported non-connected by 100%, 97% and 95% of methods, respectively. A) show a 3D rendering of the ground-truth fibre trajectories of the three bundles. B) and C) show a 2D cross-sectional image of the local orientations of the ground-truth tubular fibres, with fibre segment intersecting the 2D plane. All three bundles show a long and straight configuration going through the centre of the phantom and mixing with the other bundles. Those three bundles are the bundle with the lowest, second lowest and 5th lowest connectivity in the ground-truth weighted connectivity matrix. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

binning tractography performances in a single measurement as different measurements can lead to a change in the ranking of a specific method, sometime in opposite directions.

4.2. Binary classification of the connectivity

The performances of participating teams for binary classification of the connectivity is also higher on the DiSCo numerical phantoms than previously reported results on other synthetic data (Maier-Hein et al., 2017) and real brain (Caminiti et al., 2021; Donahue et al., 2016; Girard et al., 2020). For instance, teams 3 and 14 obtained a specificity of 1, i.e. no false positives (see Figs. 4, 5). Most teams have 3 to 5 false negatives, showing high sensitivity. Team 3 and 4 have the highest accuracy, with 4 and 5 misclassified pairs of ROIs, respectively, out of 120 pairs of ROIs. Moreover, Team 3 and 14's best-performing methods had no false positives, even before applying the thresholding. This was achieved by the teams via thresholding of their matrices before the challenge submission, with a threshold value estimated using the train-

ing dataset. This also suggest the DiSCo substrates, although complex, are oversimplifying real brain connectivity.

Nonetheless, the errors (false positives/negatives) of methods are not randomly distributed among the connections of the numerical substrate. Rather, a subset of bundles is either consistently wrongly connected or wrongly not connected (see Fig. 6). The most frequently reported false negatives are non-dominant bundles with generally low connection strength in the ground-truth matrix (fewer synthetic fibres than other bundles). They also have a straight geometric profile with synthetic fibres crossing with several other bundles in the central partition of the phantom, as shown in Fig. 8. Contrarily, the most frequently reported false positives are bundles connecting adjacent ROIs (see Fig. 7). These bundles are likely the result of two portions of existing bundles wrongly merged due to a low angle crossing and bottlenecks configurations (Girard et al., 2020; Maier-Hein et al., 2017). This may indicate that bundle metrics, such as volume and structural connectivity estimates, may be biased by the shape and size of white matter bundles, rather than being uniform across all of them.

Table 1

Best-performing method for each team. Most of the best-performing methods used DW-MRI signal denoising, multi-shell multi-tissue spherical deconvolution, probabilistic or deterministic tractography, and microstructure informed-tractography filtering approaches. ASI (Wu et al., 2019), AxCaliber (Assaf et al., 2008; Fick et al., 2019), COMMIT (Daducci et al., 2014), COMMIT_{tree} (Ocampo-Pineda et al., 2021), CSD (Tournier et al., 2004; 2019), Deterministic RK4 (Yeh, 2017), iFOD2 (Tournier et al., 2010; 2019), iFOD1 (Tournier et al., 2010; 2019), MPPCA (Veraart et al., 2016), MRDS (Coronado-Leija et al., 2017), msmt-CSD (Jeurissen et al., 2014), ms-fODF (Tran and Shi, 2015), Probabilistic tractography (Garyfallidis et al., 2014), RUMBA-SD (Canales-Rodríguez et al., 2015), Radial DSI (Baete et al., 2016), SD_STREAM (Tournier et al., 2019), SIFT2 (Smith et al., 2015), SR-ASI (Wu et al., 2020), PFT (Girard et al., 2014), PTT (Aydogan and Shi, 2021), U-net fODFs (Sedlar et al., 2021).

	r	Fraction of Valid Streamlines	AUC	Accuracy	Denoising	Local Modelling	Tractography Algorithm	Connectivity Weighting
Team 1	0.945	0.966	0.918	0.942	MPPCA	RUMBA-SD	Probabilistic	counts
Team 2	0.960	0.937	0.955	0.942	MPPCA	msmt-CSD	SD_STREAM	SIFT2
Team 3	0.951	1.000	0.923	0.967	MPPCA	Radial DSI	Deterministic RK4	count, length scaling, thresholding
Team 4	0.964	0.930	0.982	0.958		U-net fODFs	iFOD2	SIFT2
Team 5	0.940	0.900	0.964	0.925		msmt-CSD	iFOD2	SIFT2
Team 6	0.919	0.856	0.940	0.858	MPPCA	ASI	SR-ASI	SIFT2
Team 7	0.954	0.938	0.956	0.942	MPPCA	msmt-CSD	SD_STREAM	counts
Team 8	0.971	0.930	0.963	0.925	MPPCA	msmt-CSD	PTT	COMMIT _{tree}
Team 9	0.960	0.938	0.956	0.942	MPPCA	msmt-CSD	SD_STREAM	SIFT2
Team 10	0.972	0.911	0.982	0.925	MPPCA	msmt-CSD	PTT	SIFT2
Team 11	0.874	0.861	0.906	0.900		CSD	PFT	COMMIT
Team 12	0.964	0.912	0.964	0.942	MPPCA	msmt-CSD	iFOD1	AxCaliber
Team 13	0.973	0.893	0.969	0.917	MPPCA	ms-fODFs	PTT	COMMIT
Team 14	0.946	1.000	0.865	0.942	MPPCA	MRDS	iFOD2	count, thresholding

In this work, we fixed a threshold of 5% of each method maximum connectivity to binarise connectivity matrices. This will inevitably penalise the identification of connections with low weights, where underestimation may lead to the exclusion of connections. This is also the case *in vivo*, when connectivity matrices are binarised. As such, alternative matrix binarisation methods, such as using various fixed thresholds or using thresholds specific to each connection, should be investigated in future work.

4.3. Characteristics of the best-performing methods

The estimated connectivity matrices of the best-performing methods submitted by the teams are shown in Fig. 3, and their corresponding processing methods are listed in Table 1. Most of the team used the MPPCA denoising algorithm (Veraart et al., 2016) before performing the local reconstructions. Although multiple local reconstruction methods (Baete et al., 2016; Canales-Rodríguez et al., 2015; Coronado-Leija et al., 2017; Sedlar et al., 2021; Tournier et al., 2004; Wu et al., 2019) yield a high Pearson correlation coefficient, the multi-shell multi-tissue spherical deconvolution method was the most common (Jeurissen et al., 2014). Various tractography algorithms were selected (Aydogan and Shi, 2021; Garyfallidis et al., 2014; Tournier et al., 2010; 2019; Wu et al., 2020; Yeh, 2017), with the probabilistic streamlines tractography methods being the most common. In particular, the top 3 connectivity methods with the highest Pearson correlation coefficient (r) all used the Parallel Transport Tractography (PTT) algorithm (Aydogan and Shi, 2021). Notably, the method with the highest accuracy used the RK4 deterministic tractography algorithm (Yeh, 2017) combined with the Radial DSI reconstruction (Baete et al., 2016). Moreover, most of the submitted matrices used microstructure-informed tractography (Daducci et al., 2014; Frigo et al., 2021; Smith et al., 2013; 2015) to weigh the connectivity matrices, in particular, the top 3 all used the SIFT2 (Smith et al., 2015) or the COMMIT (Daducci et al., 2014) methods. However, teams using streamline counts or thresholded streamline counts to estimate the connectivity also obtained a high Pearson correlation coefficient, particularly when paired with deterministic tractography algorithms. Future work should target evaluating individual steps (e.g. denoising, local reconstruction, tractography, connectivity weighting methods), fixing the other steps to assess its effects on the connectivity evaluation. Moreover, other methods, not selected by teams, may provide similarly good results and shouldn't be discarded. Rather, results presented here serve as baseline for future method testing and development. Nonetheless, the

geometry of the fibre in DiSCO substrates may favour some methods over others. Hence, conclusions derived from numerical substrates must be challenged against real data.

5. Conclusion

Current tractography and connectivity methods show exceptional performance on the DiSCO datasets. All methods selected by participating teams were able to accurately estimate connectivity weights corresponding to the cross-sectional area of the synthetic fibres connecting the network. Furthermore, they were able to accurately identify the pairs of ROIs interconnected by synthetic fibres. Previous phantoms were designed to validate either tractography or microstructure; we believe that DiSCO phantoms enable an improved assessment of the reliability of quantitative connectivity methods thanks to their microscopic and macroscopic properties. Tractography is capable of accurately solving complex configurations, as demonstrated by this challenge. However, a noticeable gap exists between the challenge results and results in real data or from other validation techniques. As such, the complexity of the numerical substrates should be improved, for instance, by varying the tubular shape of the fibre, increasing the packing density, adding T2 effects and simulating membrane permeability. Moreover, future work should modify the DW-MRI signal by adding MRI artifacts, changing spatial and angular resolutions, as well as varying the acquisition protocol to test tractography in clinically realistic DW-MRI signals. Overall, this work contributes to the growing body of evidence suggesting that tractography research should focus on improving tractography in bottlenecks and other challenging fibre configurations. The DiSCO datasets are available publicly (Rafael-Patino et al., 2021a; 2022) to foster the development of the next generation of structural connectivity methods.

Author contributions

GG, JRP, RF, MP, EC, and JPT conceived and designed the analysis, collected the data, and wrote the initial draft.

DB, NA, VAN, VP, BBB, ALA, SBo, IG, MOP, MB, ZP, PB, SSc, AD, AS, DC, FB, TP, MF, SSe, SD, IK, MZ, HL, YJ, RD, KGS, BAL, AC, GAB, SBe, NN, PK, FR, PF, TS, YCL, DGP, FEB, SHB, EHG, ARM, RCL, PSS, LC, MD, SML, CS, AZ, KM, EJCR, YW, SA, PTY, AT, FG, and FM contributed data and analysis.

All authors reviewed and edited the manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Data availability

The three dataset used in this study is available online at <http://data.mendeley.com/datasets/fgf86jdfg6> (Rafael-Patino et al., 2021a; 2022). The dataset includes the noiseless and corrupted DW-MRI signal (SNR = [10, 20, 30, 40, 50]) at two resolutions (25 μm and 50 μm isotropic voxels). The dataset also includes the 3D mesh used to simulate the signals, the ROIs masks, and the trajectories of the fibres with their corresponding diameter.

Acknowledgements

We acknowledge access to the facilities and expertise of the CIBM Center for Biomedical Imaging, a Swiss research center of excellence funded and supported by Lausanne University Hospital (CHUV), University of Lausanne (UNIL), École polytechnique fédérale de Lausanne (EPFL), University of Geneva (UNIGE) and Geneva University Hospitals (HUG). The calculations have been performed using the facilities of the Scientific IT and Application Support Center of EPFL. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. This project has received funding from the Swiss National Science Foundation under grant number 205320_175974 and Spark grant number 190297. Marco Pizzolato acknowledges the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754462. Erick J. Canales-Rodríguez was supported by the Swiss National Science Foundation (SNSF, Ambizione grant PZ00P2_185814). This research project is part of the MMINCARAV Inria associate team program between Empenn (Inria Rennes Bretagne Atlantique) and LTS5 (École Polytechnique Fédérale de Lausanne - EPFL) that started in 2019. Raphaël Truffet's PhD is partly funded by ENS Rennes. Tomasz Pieciak acknowledges the Polish National Agency for Academic Exchange for the grant PPN/BEK/2019/1/00421 under the Bekker programme and the Ministry of Science and Higher Education (Poland) under the scholarship for outstanding young scientists (692/STYP/13/2018). Dominika Ciupek acknowledges the Ministry of Education and Science (Poland) for the grant MEiN/2021/209/DIR/NN4 under the "Best of the Best 4.0" programme. Dominika Ciupek acknowledges that this work was supported by the European Union's Horizon 2020 research and innovation program under grant agreement Sano No 857533 and the International Research Agendas program of the Foundation for Polish Science No MAB PLUS/2019/13. Dominika Ciupek, Aleksandra Stafiej, Fabian Bogusz and Tomasz Pieciak gratefully acknowledge Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2022/015357. Ye Wu, Sahar Ahmad, and Pew-Thian Yap were supported in part by the United States National Institute of Mental Health (R01MH125479). Ye Wu is supported by National Natural Science Foundation of China (No. 62201265). Patryk Filipiak, Tim Shepherd, Ying-Chia Lin, Dimitris G. Placantonakis, Fernando E. Boada and Steven H. Baete are supported in part by the National Institutes of Health (R01-EB028774, R01-NS082436 and P41-EB017183). Athena Project Team acknowledges that this work was supported by the ERC under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant agreement No 694665:CoBCoM: Computational Brain Connectivity Mapping) and it has been partly supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. Athena Project Team is grateful to Inria Sophia Antipolis - Méditerranée https://wiki.inria.fr/ClustersSophia/Usage_policy "Nef" computation cluster for providing resources and support.

The team from UW-Madison would like to acknowledge the NIH grants U54HD090256, R01NS123378, P50HD105353, R01NS092870, R01EB022883, R01AI117924, R01AG027161, RF1AG059312, P50AG033514, R01NS105646, UF1AG051216, R01NS111022, R01NS117568, P01AI132132, R01AI138647, R34DA050258, and R01AG037639. Andrew Zalesky was supported by research fellowships from the NHMRC (APP1118153). Bennett A. Landman and Kurt G. Schilling were supported by NIH grants 1R01EB017230 and K01EB032898. Maxime Descoteaux and the SCIL participants were supported by NSERC Discovery Grant RGPIN-2020-04818 and institutional research Chair in Neuroinformatics. Alonso Ramírez-Manzanares was partially supported by SNI-CONACYT, México. Luis Concha was partially funded by UNAM-DGAPA (IN204720).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.neuroimage.2023.120231](https://doi.org/10.1016/j.neuroimage.2023.120231)

References

- Alexander, D.C., Hubbard, P.L., Hall, M.G., a. Moore, E., Pitto, M., Parker, G.J.M., Dyrby, T.B., 2010. Orientationally invariant indices of axon diameter and density from diffusion mri. *Neuroimage* 52, 1374–1389. doi:[10.1016/j.neuroimage.2010.05.043](https://doi.org/10.1016/j.neuroimage.2010.05.043).
- Ambrosen, K.S., Eskildsen, S.F., Hinne, M., Krug, K., Lundell, H., Schmidt, M.N., van Gerven, M.A., Mørup, M., Dyrby, T.B., 2020. Validation of structural brain connectivity networks: the impact of scanning parameters. *Neuroimage* 204, 116–207. doi:[10.1016/J.NEUROIMAGE.2019.116207](https://doi.org/10.1016/J.NEUROIMAGE.2019.116207).
- Ambrosen, K.S., Eskildsen, S.F., Hinne, M., Krug, K., Lundell, H., Schmidt, M.N., van Gerven, M.A., Mørup, M., Dyrby, T.B., 2020. Validation of structural brain connectivity networks: the impact of scanning parameters. *Neuroimage* 204, 116–207. doi:[10.1016/J.NEUROIMAGE.2019.116207](https://doi.org/10.1016/J.NEUROIMAGE.2019.116207).
- Andersson, M., Kjer, H.M., Rafael-Patino, J., Pacureanu, A., Pakkenberg, B., Thiran, J.P., Pitto, M., Bech, M., Dahl, A.B., Dahl, V.A., Dyrby, T.B., 2021. Axon morphology is modulated by the local environment and impacts the noninvasive investigation of its structure-function relationship. *Proc. Natl. Acad. Sci. U.S.A.* 117, 33649–33659. doi:[10.1073/PNAS.2012533117](https://doi.org/10.1073/PNAS.2012533117).
- Assaf, Y., Blumenthal-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. doi:[10.1002/mrm.21577](https://doi.org/10.1002/mrm.21577).
- Aydogan, D.B., Jacobs, R., Dulawa, S., Thompson, S.L., Francois, M.C., Toga, A.W., Dong, H., Knowles, J.A., Shi, Y., 2018. When tractography meets tracer injections: a systematic study of trends and variation sources of diffusion-based connectivity. *Brain Struct. Funct.* 223, 2841–2858. doi:[10.1007/s00429-018-1663-8](https://doi.org/10.1007/s00429-018-1663-8).
- Aydogan, D.B., Shi, Y., 2021. Parallel transport tractography. *IEEE Trans. Med. Imaging* 40, 635–647. doi:[10.1109/TMI.2020.3034038](https://doi.org/10.1109/TMI.2020.3034038).
- Azadbakht, H., Parkes, L.M., Haroon, H.A., Augath, M., Logothetis, N.K., de Crespigny, A., D'Arceuil, H.E., Parker, G.J.M., 2015. Validation of high-resolution tractography against in vivo tracing in the macaque visual cortex. *Cerebral Cortex* 25, 4299–4309. doi:[10.1093/cercor/bhu326](https://doi.org/10.1093/cercor/bhu326).
- Baete, S.H., Yutzky, S., Boada, F.E., 2016. Radial q-space sampling for dsi. *Magn. Reson. Med.* 76, 769–780. doi:[10.1002/MRM.25917/ASSET/SUPINFO/MRM25917-SUP-001-SUPINFO01.PDF](https://doi.org/10.1002/MRM.25917/ASSET/SUPINFO/MRM25917-SUP-001-SUPINFO01.PDF).
- Caminiti, R., Girard, G., Battaglia-Mayer, A., Borra, E., Schito, A., Innocenti, G.M., Lupino, G., 2021. The complex hodological architecture of the macaque dorsal intraparietal areas as emerging from neural tracers and dw-mri tractography. *eNeuro* 8. doi:[10.1523/ENEURO.0102-21.2021](https://doi.org/10.1523/ENEURO.0102-21.2021).
- Canales-Rodríguez, E.J., Daducci, A., Sotiropoulos, S.N., Caruyer, E., Aja-Fernández, S., Radua, J., Mendizabal, J.M.Y., Iturria-Molina, Y., Melie-García, L., Alemán-Gómez, Y., Thiran, J.-P., Sarró, S., Pomarol-Clotet, E., Salvador, R., 2015. Spherical deconvolution of multichannel diffusion mri data with non-gaussian noise models and spatial regularization. *PLoS ONE* 10, e0138910. doi:[10.1371/journal.pone.0138910](https://doi.org/10.1371/journal.pone.0138910).
- Caruyer, E., Daducci, A., Descoteaux, M., Christophe Houde, J., Philippe Thiran, J., Verma, R., 2014. Phantoms: a flexible software library to simulate diffusion mr phantoms.
- Côté, M.-A., Girard, G., Boré, A., Garyfallidis, E., Houde, J.-C., Descoteaux, M., 2013. Tractometer: towards validation of tractography pipelines. *Med. Image Anal.* 17, 844–857.
- Cercignani, M., Giuliotti, G., Dowell, N.G., Gabel, M., Broad, R., Leigh, P.N., Harrison, N.A., Bozzali, M., 2017. Characterizing axonal myelination within the healthy population: a tract-by-tract mapping of effects of age and gender on the fiber g-ratio. *Neurobiol. Aging* 49, 109. doi:[10.1016/J.NEUROBIOLAGING.2016.09.016](https://doi.org/10.1016/J.NEUROBIOLAGING.2016.09.016).
- Chomiak, T., Hu, B., 2009. What is the optimal value of the g-ratio for myelinated fibers in the rat CNS? A theoretical approach. *PLoS ONE* 4. doi:[10.1371/JOURNAL.PONE.0007754](https://doi.org/10.1371/JOURNAL.PONE.0007754).
- Close, T.G., Tournier, J.-D., Calamante, F., Johnston, L.A., Mareels, I., Connelly, A., 2009. A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. *Neuroimage* 47, 1288–1300. doi:[10.1016/J.NEUROIMAGE.2009.03.077](https://doi.org/10.1016/J.NEUROIMAGE.2009.03.077).
- Coronado-Leija, R., Ramirez-Manzanares, A., Marroquin, J.L., 2017. Estimation of individual axon bundle properties by a multi-resolution discrete-search method. *Med. Image Anal.* 42, 26–43. doi:[10.1016/J.MEDIA.2017.06.008](https://doi.org/10.1016/J.MEDIA.2017.06.008).

- Daducci, A., Canales-Rodríguez, E.J., Zhang, H., Dyrby, T.B., Alexander, D.C., Thiran, J.-P., 2015. Accelerated microstructure imaging via convex optimization (amico) from diffusion mri data. *Neuroimage* 105, 32–44. doi:10.1016/j.neuroimage.2014.10.026.
- Daducci, A., Palu, A.D., Lemkaddem, A., Thiran, J.-P., 2014. Commit: convex optimization modeling for micro-structure informed tractography. *IEEE Trans. Med. Imaging* 34.
- Delettre, C., Messe, 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 Neurosci. (Cambridge, Mass.)* 3, 1038–1050. doi:10.1101/517136.
- Dimitriadis, S.I., Drakesmith, M., Bells, S., Parker, G.D., Linden, D.E., Jones, D.K., 2017. Improving the reliability of network metrics in structural brain networks by integrating different network weighting strategies into a single graph. *Front. Neurosci.* 11, 694. doi:10.3389/FNINS.2017.00694/BIBTEX.
- Donahue, C.J., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Behrens, T.E., Dyrby, T.B., Coalson, T., Kennedy, H., Knoblauch, K., Essen, D.C.V., Glasser, M.F., 2016. Using diffusion tractography to predict cortical connection strength and distance: a quantitative comparison with tracers in the monkey. *J. Neurosci.* 36, 6758–6770. doi:10.1523/JNEUROSCI.0493-16.2016.
- Essen, D.C.V., Jbabdi, S., Sotiropoulos, S.N., Chen, C., Dikranian, K., Coalson, T., Harwell, J., Glasser, M.F., 2014. Mapping connections in humans and non-human primates: aspirations and challenges for diffusion imaging. *Diffusion MRI* 337–358. doi:10.1016/B978-0-12-396460-1.00016-0.
- Fick, R.H., Wassermann, D., Deriche, R., 2019. The dmipy toolbox: diffusion mri multi-compartment modeling and microstructure recovery made easy. *Front. Neuroinform.* 13, 64. doi:10.3389/FNINF.2019.00064/BIBTEX.
- Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., Reisert, M., Sakaie, K., Tensaouti, F., Yo, T., Mangin, J.-F., Poupon, C., 2011. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion mr phantom. *Neuroimage* 56, 220–234. doi:10.1016/j.neuroimage.2011.01.032.
- Frigo, M., Zucchelli, M., Deriche, R., Deslauriers-Gauthier, S., 2021. Talon: tractograms as linear operators in neuroimaging.
- Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Walt, S.V.D., Descoteaux, M., Nimmo-Smith, I., Contributors, D., 2014. Dipy, a library for the analysis of diffusion mri data. *Front. Neuroinform.* 8, 1–5. <http://www.frontiersin.org/Journal/10.3389/fninf.2014.00008/abstract>.
- Girard, G., Caminiti, R., Battaglia-Mayer, A., St-Onge, E., Ambrosen, K.S., Eskildsen, S.F., Krug, K., Dyrby, T.B., Descoteaux, M., Thiran, J.P., Innocenti, G.M., 2020. On the cortical connectivity in the macaque brain: a comparison of diffusion tractography and histological tracing data. *Neuroimage* 221, 117201. doi:10.1016/j.neuroimage.2020.117201.
- Girard, G., Caruyer, E., Rafael-Patino, J., Pizzolato, M., Truffet, R., Thiran, J.-P., 2021. Diffusion-simulated connectivity challenge. *Zenodo*. <https://doi.org/10.5281/zenodo.4733450>.
- Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M., 2014. Towards quantitative connectivity analysis: reducing tractography biases. *Neuroimage* 98, 266–278. doi:10.1016/j.neuroimage.2014.04.074.
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. *PLoS Biol.* 6, e159. doi:10.1371/journal.pbio.0060159.
- Hagmann, P., Kurlant, G., Gigandet, X., Thiran, P., Wedeen, V.J., Meuli, R., Thiran, J.-P., 2007. Mapping human whole-brain structural networks with diffusion mri. *PLoS ONE* 2, e597. doi:10.1371/journal.pone.0000597.
- Hall, M.G., Alexander, D.C., 2009. Convergence and parameter choice for monte-carlo simulations of diffusion mri. *IEEE Trans. Med. Imaging* 28, 1354–1364. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4797853>.
- van den Heuvel, M.P., de Reus, M.A., Barrett, L.F., Scholtens, L.H., Coopmans, F.M., Schmidt, R., Preuss, T.M., Rilling, J.K., Li, L., 2015. Comparison of diffusion tractography and tract-tracing measures of connectivity strength in rhesus macaque connectome. *Hum. Brain Mapp.* 36, 3064–3075. doi:10.1002/hbm.22828.
- Jbabdi, S., Lehman, J.F., Haber, S.N., Behrens, T.E., 2013. Human and monkey ventral prefrontal fibers use the same organizational principles to reach their targets: tracing versus tractography. *J. Neurosci.* 33, 3190–3201. doi:10.1523/JNEUROSCI.2457-12.2013.
- Jbabdi, S., Sotiropoulos, S.N., Haber, S.N., Essen, D.C.V., Behrens, T.E., 2015. Measuring macroscopic brain connections in vivo. *Nat. Neurosci.* 18, 1546–1555. doi:10.1038/nn.4134.
- Jeurissen, B., Descoteaux, M., Mori, S., Leemans, A., 2017. Diffusion mri fiber tractography of the brain. *NMR Biomed.* e3785. doi:10.1002/nbm.3785.
- 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. doi:10.1016/j.neuroimage.2014.07.061.
- Jones, D.K., 2010. Challenges and limitations of quantifying brain connectivity in vivo with diffusion mri. *Imaging Med.* 2, 341–355. <http://www.futuremedicine.com/doi/abs/10.2217/iim.10.21>.
- Le Bihan, D., Poupon, C., Amadon, A., Lethimonnier, F., 2006. Artifacts and pitfalls in diffusion mri. *J. Magn. Reson. Imaging* 24, 478–488. <http://www.ncbi.nlm.nih.gov/pubmed/16897692>.
- Lee, H.H., Fieremans, E., Novikov, D.S., 2021. Realistic microstructure simulator (rms): monte carlo simulations of diffusion in three-dimensional cell segmentations of microscopy images. *J. Neurosci. Methods* 350, 109018. doi:10.1016/j.jneumeth.2020.109018.
- Maffei, C., Girard, G., Schilling, K.G., Aydogan, D.B., Adluru, N., Zhykka, A., Wu, Y., Mancini, M., Hamamci, A., Sarica, A., Teillac, A., Baete, S.H., Karimi, D., Yeh, F.-C., Yildiz, M.E., Gholipour, A., Bihan-Poudec, Y., Hiba, B., Quattrone, A., Quattrone, A., Boshkovski, T., Stikov, N., Yap, P.-T., de Luca, A., Pluim, J., Leemans, A., Prabhakaran, V., Bendlin, B.B., Alexander, A.L., Landman, B.A., Canales-Rodríguez, E.J., Barakovic, M., Rafael-Patino, J., Yu, T., Rensonnet, G., Schiavi, S., Daducci, A., Pizzolato, M., Fisci-Gomez, E., Thiran, J.-P., Dai, G., Grisot, G., Lazovski, N., Puch, S., Ramos, M., Rodrigues, P., Prkowska, V., Jones, R., Lehman, J., Haber, S.N., Yendiki, A., 2022. Insights from the irontract challenge: optimal methods for mapping brain pathways from multi-shell diffusion mri. *Neuroimage* 257, 119327. doi:10.1016/J.NEUROIMAGE.2022.119327.
- Maier-Hein, K.H., Neher, P.F., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.-C., Lin, Y.-C., Ji, Q., Reddick, W.E., Glass, J.O., Chen, D.Q., Feng, Y., Gao, C., Wu, Y., Ma, J., Renjie, H., Li, Q., Westin, C.-F., Deslauriers-Gauthier, S., González, J.O.O., Paquette, M., St-Jean, S., Girard, G., Rheault, F., Sidhu, J., Tax, C.M.W., Guo, F., Mesri, H.Y., David, 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., Auria, 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. *Nat. Commun.* 8, 1349. doi:10.1038/s41467-017-01285-x.
- Messaritaki, E., Dimitriadis, S.I., Jones, D.K., 2019. Optimization of graph construction can significantly increase the power of structural brain network studies. *Neuroimage* 199, 495–511. doi:10.1016/J.NEUROIMAGE.2019.05.052.
- Nath, V., Schilling, K.G., Parvathaneni, P., Huo, Y., Blaber, J.A., Hainline, A.E., Barakovic, M., Romascano, D., Rafael-Patino, J., Frigo, M., Girard, G., Thiran, J., Daducci, A., Rowe, M., Rodrigues, P., Prkowska, V., Aydogan, D.B., Sun, W., Shi, Y., Parker, W.A., Ismail, A.A.O., Verma, R., Cabeen, R.P., Toga, A.W., Newton, A.T., Wasserthal, J., Neher, P., Maier-Hein, K., Savini, G., Palesi, F., Kaden, E., Wu, Y., He, J., Feng, Y., Paquette, M., Rheault, F., Sidhu, J., Lebel, C., Leemans, A., Descoteaux, M., Dyrby, T.B., Kang, H., Landman, B.A., 2020. Tractography reproducibility challenge with empirical data (traced): the 2017 ismrm diffusion study group challenge. *J. Magn. Reson. Imaging* 51, 234–249. doi:10.1002/jmri.26794.
- Neher, P.F., Stieltjes, B., Laun, F.B., Meinzer, H.-P., Fritzsche, K., 2013. *Fiberfox: a novel tool to generate software phantoms of complex fiber geometries*, Vol. 21.
- Ocampo-Pineda, M., Schiavi, S., Rheault, F., Girard, G., Petit, L., Descoteaux, M., Daducci, A., 2021. Hierarchical microstructure informed tractography. *Brain Connect.* doi:10.1089/brain.2020.0907.
- Rafael-Patino, J., Girard, G., Truffet, R., Pizzolato, M., Caruyer, E., Thiran, J.-P., 2021. The diffusion-simulated connectivity (disco) dataset. *Data Brief* 38, 107429. doi:10.1016/J.DIB.2021.107429.
- Rafael-Patino, J., Girard, G., Truffet, R., Pizzolato, M., Caruyer, E., Thiran, J.-P., 2022. The diffusion-simulated connectivity dataset 2. doi:10.17632/FGF86JDFG6.2.
- Rafael-Patino, J., Girard, G., Truffet, R., Pizzolato, M., Thiran, J.-P., Caruyer, E., 2021. The microstructural features of the diffusion-simulated connectivity (disco) dataset. *Lect. Note. Comput. Sci. (Incl. subser. Lect. Note. Artif. Intell. Lect. Note. Bioinform.)* 13006 LNCS, 159–170. doi:10.1007/978-3-030-87615-9_14.
- Rafael-Patino, J., Romascano, D., Ramirez-Manzanares, A., Canales-Rodríguez, E.J., Girard, G., Thiran, J.-P., 2020. Robust monte-carlo simulations in diffusion-mri: effect of the substrate complexity and parameter choice on the reproducibility of results. *Front. Neuroinform.* 14, 8. doi:10.3389/fninf.2020.00008.
- Romascano, D., Barakovic, M., Rafael-Patino, J., Dyrby, T.B., Thiran, J.P., Daducci, A., 2019. Activeaxadd: toward non-parametric and orientationally invariant axon diameter distribution mapping using pgse. *Magn. Reson. Med.* 83, 2322–2330. doi:10.1002/mrm.28053.
- Schilling, K.G., Gao, Y., Stepniewska, I., Janve, V., Landman, B.A., Anderson, A.W., 2019. Anatomical accuracy of standard-practice tractography algorithms in the motor system – a histological validation in the squirrel monkey brain. *Magn. Reson. Imaging* 55, 7–25. doi:10.1016/j.mri.2018.09.004.
- Schilling, K.G., Nath, V., Hansen, C., Parvathaneni, P., Blaber, J., Gao, Y., Neher, P., Aydogan, D.B., Shi, Y., Ocampo-Pineda, M., Schiavi, S., Daducci, A., Girard, G., Barakovic, M., Rafael-Patino, J., Romascano, D., Rensonnet, G., Pizzolato, M., Bates, A., Fisci, E., Thiran, J.-P., Canales-Rodríguez, E.J., Huang, C., Zhu, H., Zhong, L., Cabeen, R., Toga, A.W., Rheault, F., Theaud, G., Houde, J.-C., Sidhu, J., Chamberland, M., Westin, C.-F., Dyrby, T.B., Verma, R., Rathi, Y., Irfanoglu, M.O., Thomas, C., Pierpaoli, C., Descoteaux, M., Anderson, A.W., Landman, B.A., 2019. Limits to anatomical accuracy of diffusion tractography using modern approaches. *Neuroimage* 185, 1–11. doi:10.1016/J.NEUROIMAGE.2018.10.029.
- Sedlar, S., Papadopoulos, T., Deriche, R., Deslauriers-Gauthier, S., 2021. Diffusion mri fiber orientation distribution function estimation using voxel-wise spherical u-net. *Springer, Cham*, pp. 95–106. doi:10.1007/978-3-030-73018-5_8.
- Smith, R.E., Tournier, J.-D., Calamante, F., Connelly, A., 2013. Sift: spherical-deconvolution informed filtering of tractograms. *Neuroimage* 67, 298–312. 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, 338–351. doi:10.1016/j.neuroimage.2015.06.092.
- Sotiropoulos, S.N., Zalesky, A., 2019. Building connectomes using diffusion mri: why, how and but. *NMR Biomed.* 32. doi:10.1002/NBM.3752.
- Sporns, O., 2011. The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125. <http://www.ncbi.nlm.nih.gov/pubmed/21251014>.
- 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. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16574–16579. doi:10.1073/pnas.1405672111.

- Tournier, J.-D., Calamante, F., Connelly, A., 2010. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions.
- 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. doi:[10.1016/j.neuroimage.2004.07.037](https://doi.org/10.1016/j.neuroimage.2004.07.037).
- 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. doi:[10.1016/j.neuroimage.2019.116137](https://doi.org/10.1016/j.neuroimage.2019.116137).
- Tran, G., Shi, Y., 2015. Fiber orientation and compartment parameter estimation from multi-shell diffusion imaging. *IEEE Trans. Med. Imaging* 34, 2320–2332. doi:[10.1109/TMI.2015.2430850](https://doi.org/10.1109/TMI.2015.2430850).
- 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. doi:[10.1016/j.neuroimage.2016.08.016](https://doi.org/10.1016/j.neuroimage.2016.08.016).
- Wu, Y., Hong, Y., Ahmad, S., Chang, W.T., Lin, W., Shen, D., Yap, P.T., 2020. Globally optimized super-resolution of diffusion mri data via fiber continuity. *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 12267, 260. doi:[10.1007/978-3-030-59728-3_26](https://doi.org/10.1007/978-3-030-59728-3_26).
- Wu, Y., Lin, W., Shen, D., Yap, P.-T., 2019. Asymmetry spectrum imaging for baby diffusion tractography. doi:[10.1007/978-3-030-20351-1_24](https://doi.org/10.1007/978-3-030-20351-1_24).
- Yeh, C.H., Jones, D.K., Liang, X., Descoteaux, M., Connelly, A., 2021. Mapping structural connectivity using diffusion mri: challenges and opportunities. *J. Magn. Reson. Imaging* 53, 1666–1682. doi:[10.1002/JMRI.27188](https://doi.org/10.1002/JMRI.27188).
- Yeh, F., 2017. Diffusion mri reconstruction in dsi studio. Advanced Biomedical MRI Lab, National Taiwan University Hospital. Available online at: <http://dsi-studio.labsolver.org/Manual/Reconstruction>.