Generation of synthetic data for validating tractography-based cortical parcellation and fiber clustering algorithms

Elida Poo[1], Joaquín Molina[1], Jean-François Mangin[2], Cecilia Hernández[1], Pamela Guevara[1]
[1] Faculty of Engineering, Universidad de Concepción, Concepción, Chile
[2] CEA, CNRS, Baobab, Neurospin, Université Paris-Saclay, Gif-sur-Yvette, France,

Introduction. Diffusion Magnetic Resonance Imaging (dMRI) tractography [1] has enabled the study of white matter connectivity and the development of automated methods for both fiber bundle clustering and segmentation and cortical parcellation. However, objective validation of such methods is limited by the lack of anatomical ground truth. We present two tools to address this gap: PhyberSIM [2], a white matter fiber bundle simulator, and a synthetic data generator that produces random cortical parcellations based and their connections [3], to validate tractography-based cortical parcellation (TBCP) methods.

Materials and Methods. PhyberSIM generates fiber bundles using a tubular model parameterized by a centroid, selected from the tractogram of a subject, and five radii along the bundle trajectory, used to generate fibers based on spline curves (Fig. 1). It was validated using bundles from a Deep White Matter (DWM) atlas [4] and employed to evaluate two fiber clustering algorithms, QuickBundles (QB) [5] and FFClust [6] through five classic clustering metrics. The cortical parcellation data simulator creates synthetic data consisting of a geodesic distance-based cortical parcellation based on the cortical mesh of a subject and the connections between the generated parcels. For the simulation of each bundle, it defines a centroid based on the subject's tractogram and adapted end points to fit the shape of the connecting parcels. We generated a database of 20 subjects with 150 parcels per hemisphere that was used to validate and improve a TBCP algorithm based on a two-level fiber clustering [7] (Fig. 4).

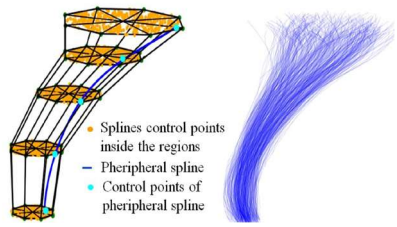

Fig. 1: Schematic of a simulated bundle generated with PhyberSIM (tubular spline model) [2].

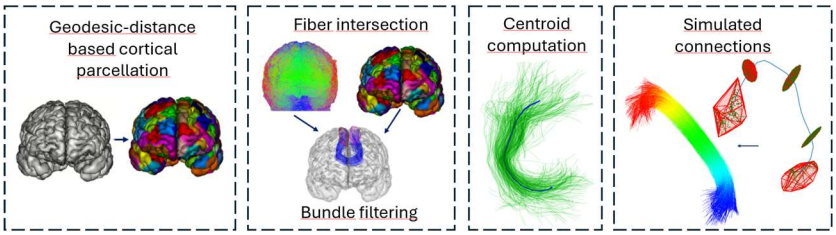

Fig. 2: Schematic of the generator of a random cortical parcellation with its connections for tractography-based cortical parcellation validation [3].

Results. PhyberSIM achieved an average 76.5% overlap with DWM atlas bundles and revealed a good performance for both fiber clustering algorithms, maintaining robustness to different number of simulated bundles and input order permutation. It also detected differences between clustering methods: FFClust tends to over-segment, while QB tends to merge close bundles. Regarding the cortical parcellation data, the tested CBCP algorithm could be improved and for its best parameter configuration, detected 118 parcels for the left hemisphere and 120 parcels for the right hemisphere, based on a DICE > 0.5, with a mean DICE of 0.61.

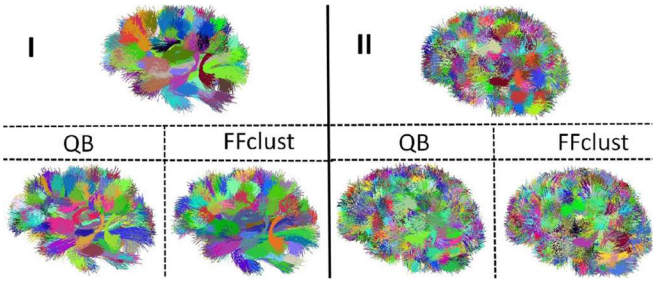

Fig. 3: QB and FFClust results for simulated tractography datasets with 100 bundles (I) and 500 bundles (II) for a distance of 10 mm.

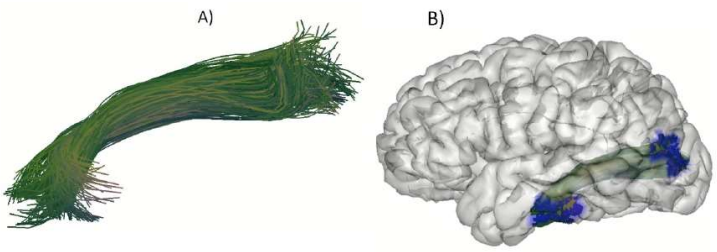

Fig. 4: Example of a simulated bundle generated from two cortical parcels. A) The simulated bundle. B) The simulated bundle (green) connecting a pair of parcels (blue) on the cortical mesh.spline curves.

Conclusions. Together, PhyberSIM and the cortical parcellation simulator offer a comprehensive, controlled validation framework for fiber clustering and tractography-based cortical parcellations algorithms. These tools enable reproducible experiments with known ground truth, providing valuable resources for the dMRI tractography community. Further improvements, such as added noise and flexibility on bundle complexity and configuration are planned for future work.

References

1. Basser, P. J., et al. (2000). Magn. Reson. Med., 44, 625–632.
2. Poo, E. et al. (2024). Front. Neurosci., 18:1396518.
3. Poo, E. et al. (2025). Submitted to Comp. in Biol. and Med.
4. Guevara, P. et al. (2012). NeuroImage, 61(4), 1083–1099.
5. Garyfallidis et al. (2012). Frontiers in Neuroscience 6, 175.
6. Vázquez et al. (2020). NeuroImage, 220, 117070.
7. Vergara et. al. (2021). IEEE EMBC 2021.

Acknoledgments to ANID-Basal AFB240002 and FB210017