INCF GSoC 2022 1.1. Implementation of White Matter Substrates in Disimpy

Google Summer of Code Proposal
Renata Cruz

September 2, 2022

1 Contact Details

Full name:	Renata Bastos da Silveira Pereira da Cruz
Professional e-mail:	renata.cruz@research.fchampalimaud.org
Personal e-mail:	renatacruz95@gmail.com
Phone number:	+351 915 571 740
Location:	Lisbon, Portugal
Neurostars ID:	renata_cruz
Webpage:	https://github.com/renata-cruz

2 Abstract

Diffusion MRI is a non-invasive imaging technique which provides information about the cellular structure of the brain tissue at microscopic scale. Disimpy is an open-source Python package for generating diffusion MRI data through Monte Carlo random walk simulations. With this project, we aim to implement features in Disimpy which could yield further biological significance to diffusion MRI simulations. In particular, we will place our focus on white matter substrates and introduce relevant concepts such as exchange in order to reproduce more realistic physiological conditions. To this end, along with the algorithm implementation, we will provide comprehensive documentation and a tutorial with clear instructions on how to use it.

3 Project Description

Diffusion Magnetic Resonance Imaging (dMRI) is a non-invasive imaging technique which provides information about the cellular structure at microscopic scale, by sensitizing the signal to the displacement of water molecules in the tissue. dMRI is extensively used to characterize brain tissue microstructure, with various applications to study brain development [3], plasticity [11], aging

[9] as well as changes upon disease [5],[1]. In the literature, there have been many different approaches proposed to acquire and analyse the dMRI data, and relate the measurements to the underlying tissue microstructure [6].

Numerical simulations play an important role in understanding the contrast of the measurements, designing new ways of acquiring and optimising the diffusion data, or studying the signal-to-noise ratio (SNR) needed to resolve different tissue features. There are many different approaches to numerical dMRI simulations, including analytical and semi-analytical solutions of the diffusion equations [10], numerical solutions [8] or Monte Carlo (MC) based simulations [4]. An advantage of the latter approach is the flexibility in designing the underlying substrate representing the tissue, as well as the ability to simulate different types of diffusion acquisitions, for instance using different magnetic field gradients to probe the molecular displacement, and have been employed in several (open-source) packages [4],[7]. On the other hand, Monte Carlo simulations can become very slow, especially when designing complicated tissue architectures and simulating a large number of particles ($> 10^6$). Therefore an important effort is made to accelerate MC dMRI simulations, while keeping the complexity of the substrates.

Disimpy is an open-source Python package for generating dMRI data through Monte Carlo random walk simulations [7]. In the current implementation, the package offers simple geometric substrates for the diffusion environment such as cylinders and spheres. For more complex simulations, it makes use of triangular meshes to create any structure required by the user, nevertheless this step requires additional time and proficiency from the user. At the moment, although the default substrates can reflect some gross tissue features, such as anisotropy and restriction inside pores of basic shapes, the underlying features are not complex enough to resemble realistic physiological conditions. Thus, we propose to implement biologically relevant features to Disimpy which can be generated automatically and could yield a more representative experiment for the user.

In this project we will focus mostly on white matter features, which are of great importance for understanding the dMRI contrast in a wide variety of applications. To generate the fibre bundles, we will start by packing 3D cylindrical meshes to which we will add various features and make the bundles more realistic. Specifically, we will incorporate variations of the axon diameter along the bundle [4], including the extreme case of beading that is observed for instance in stroke, as well as the effect of the myelin sheath thickness. Such features will be very helpful to understand the dMRI changes in various diseases, such as multiple sclerosis, and disentangle pathological effects such as axonal loss and demyelination.

Besides adding various geometrical features, we also aim to include in a realistic way the effect of exchange between the intra- and extra-axonal spaces, which is ignored in many MC simulators. An important aspect is to ensure that the exchange rates between the two compartments are constant which is necessary

in the equilibrium condition. This requirement needs a careful implementation of the probability for a particle to cross the membrane which depends on the geometry of the substrate.

To this end, the implementation of this project will result the development of useful and accessible algorithms, relevant for both research and clinical applications.

Project Timeline

Before the Coding Phase

- Read up on the literature and learn more about dMRI simulations;
- Familiarize myself with the Disimpy package by completing the tutorial and studying the documentation and source code;
- Implement a function in Disimpy to generate a standard gradient sequence - Pulsed Gradient Spin Echo (PGSE) - and run the appropriate tests;
- Discuss with the mentors and decide which features to include in the Disimpy package.

During the Coding Phase

This project will be divided into three main phases:

1. Implementation of white matter bundles (2 weeks)

- June 13^{th} - June 20^{th}: I will start by creating 2D circles with gamma distributed radii and subsequently create the cylindrical 3D meshes;
- June 20^{th} - June 27^{th}: Once the cylinder units are implemented, I will pack them to form the fibre bundles.

2. Implementation of axonal changes (4 weeks)

- June 27^{th} - July 4^{th}: First, I will implement myelination to the bundles by adding alternating layers along the fibres;
- July 4^{th} - July 11^{th}: After the myelinated fibres are created, I will include demyelination and myelin sheath thickness to the substrates to simulate different pathological conditions;
- July 11^{th} - July 25^{th}: Then, beading will be introduced as distortions along the cylinders.

3. Implementation of exchange (5 weeks + 1 week for documentation)

- July 25^{th} - August 15^{th}: I will create an algorithm to ensure that the exchange rate between the intra- and extracellular spaces is at equilibrium;

- August 15^{th} - August 29^{th}: Then I will implement permeability to the fibres addressing the different structures - myelin sheath and axonal membrane;
- August 29^{th} - September 5^{th}: The final week will be dedicated to code optimization and documentation.

Minimal Set of Deliverables

By the end of this project, I will have a successful and efficient implementation of a biologically relevant white matter algorithm for Disimpy, comprehensive documentation, and a tutorial with clear instructions on how to use it.

If Time Allows...

If given the opportunity, I will attempt to implement irregular surfaces more representative of real axonal geometry instead of the 3D cylindrical meshes previously described [2].

Mentoring Plan

I plan to maintain an open line of communication primarily via chat (WhatsApp, Slack) or via e-mail, which allows both parties to reach out at any given moment if needed. By the end of each week, I intend to send a detailed e-mail reporting the progress made during that time. Additionally, it would be ideal to schedule a weekly Zoom meeting according to the availability of the participants.

4 Candidate Details

Motivation

By being part of this project, I wish to learn more about software development within the context of dMRI. This will allow me to gain new insights in this field while focusing on the specific problem of computer simulations, which constitutes a new challenge for me as an experimentalist.

In addition, I continuously look for new ways to improve my coding skills and this would be an amazing opportunity to do so while working on a relevant problem with the guidance of the mentors.

Finally, I find it extremely exciting to be able to contribute to the development of open-source tools that could be of use to other people.

Match

I believe I can integrate knowledge from both fields for a better understanding of the problem. At the moment, I work as a research technician in the pre-clinical lab at the Champalimaud Foundation. For the most part, my responsibilities lie

with the acquisition of *in vivo* dMRI datasets. I am also involved in a number of different projects which make use of other MRI techniques such as arterial spin labelling (ASL) and resting state functional MRI (rs-fMRI).

On a different note, I do have prior experience performing data analysis touching a wide range of topics within neuroscience. These include image processing for signal quantification, rodent and fish behavior and metabolic profiles from spectroscopy MR (H1-MRS). With the exception of situations where specific softwares are required, I attempt to conduct all my analysis using Python as a free, open-source programming language. I also have additional training from recently attending a Data Science post-graduation at the University of Lisbon. The curriculum included machine learning, data mining, data processing and other relevant subjects.

As follows, I am confident my experience in dMRI combined with my scientific computing expertise and personal motivation allow me to have a different perspective, well suited for this project. This is the only Google Summer of Code project I will be applying for.

Working Time and Commitments

I intend to dedicate full-time on this project (average of 30+ hours per week as suggested in the GSoC Student Guide). My responsibilities as a technician in the lab will be reduced for the duration of the project allowing a full-time commitment. I do not plan to take holidays during this time.

Past Experience and CV

Please find my CV attached to the last page of this proposal.

References

[1] Rita Alves et al. "Correlation Tensor MRI deciphers underlying kurtosis sources in stroke". In: *NeuroImage* 247 (2022), p. 118833.

[2] Ross Callaghan et al. "ConFiG: Contextual Fibre Growth to generate realistic axonal packing for diffusion MRI simulation". In: *Neuroimage* 220 (2020), p. 117107.

[3] BJ Casey et al. "Imaging the developing brain: what have we learned about cognitive development?" In: *Trends in cognitive sciences* 9.3 (2005), pp. 104–110.

[4] Matt G Hall and Daniel C Alexander. "Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI". In: *IEEE transactions on medical imaging* 28.9 (2009), pp. 1354–1364.

[5] Edward S Hui et al. "Stroke assessment with diffusional kurtosis imaging". In: *Stroke* 43.11 (2012), pp. 2968–2973.

[6] Andrada Ianus et al. “Mapping complex cell morphology in the grey matter with double diffusion encoding MR: A simulation study”. In: *Neuroimage* 241 (2021), p. 118424.

[7] Leevi Kerkelä et al. “Disimpy: A massively parallel Monte Carlo simulator for generating diffusion-weighted MRI data in Python”. In: *Journal of Open Source Software* 5.52 (2020), p. 2527.

[8] Marco Palombo, Daniel C Alexander, and Hui Zhang. “A generative model of realistic brain cells with application to numerical simulation of the diffusion-weighted MR signal”. In: *NeuroImage* 188 (2019), pp. 391–402.

[9] Harry BM Uylings and JM De Brabander. “Neuronal changes in normal human aging and Alzheimer’s disease”. In: *Brain and cognition* 49.3 (2002), pp. 268–276.

[10] C-F Westin et al. “Processing and visualization for diffusion tensor MRI”. In: *Medical image analysis* 6.2 (2002), pp. 93–108.

[11] Robert J Zatorre, R Douglas Fields, and Heidi Johansen-Berg. “Plasticity in gray and white: neuroimaging changes in brain structure during learning”. In: *Nature neuroscience* 15.4 (2012), pp. 528–536.

Renata Bastos da Silveira Pereira da Cruz

Nationality: Portuguese
Date of Birth: 19/12/1995
Address: Lisbon, Portugal

E-mail (Personal): renatacruz95@gmail.com
E-mail (Work): renata.cruz@research.fchampalimaud.org
Mobile (1): +351 915571740

WORK EXPERIENCE & RESEARCH

03/2021 - **Champalimaud Centre for the Unknown** **LISBON, PORTUGAL**

- Research Technician in the Preclinical Lab;
- Responsibilities: assisting project management, designing protocols, data acquisition and analysis;
- Methods: dMRI, ASL, rs-fMRI, Behavioral Experiments, Eletrophysiology;
- Languages: Python, Matlab.

09/2018 - 08/2020 **Champalimaud Centre for the Unknown** **LISBON, PORTUGAL**

- Master Programme Project - 'The Role of the Ventral Telencephalon in Collective Motion'

Study group dynamics of zebrafish under a context in which transgenic individuals suffer neuronal ablation in regions of the forebrain.

- Responsibilities: project management, designing protocols, data acquisition and analysis.
- Methods: Microscopy, Behavioral Experiments, Molecular Assays.
- Languages, Libraries, Tools: Python, NumPy/SciPy, Scikit-learn, Pandas, Matplotlib, Jupyter, Git, Linux.

09/2016 - 07/2017 **Max Planck Institute for Psychiatry- ERAMUS Internship** **MUNICH, GERMANY**

- Neuronal Plasticity, Stress Neurobiology and Neurogenetics Group with collaboration of the Neuroimaging Group - MRS in Stress Animal Models, MEMRI in Contextual Fear Response, in 9.4T scanner.
- Methods: MRS, MEMRI, Behavioral Experiments.
- Responsibilities: assisting project management, designing protocols, data acquisition and analysis.
- Languages, Libraries, Tools: R, MatLab, SPM Software, TopSpin Processing, LCModel, Unix.

01/2016 - 07/2016 **Center for Neurosciences and Cell Biology- Bachelor's Internship** **COIMBRA, PORTUGAL**

- Neuromodulation Group – Behavioral Experiments, Neurochemistry and Electrophysiological Studies
- Final Project- 'Astrocytic Role in Changes in Synaptic Transmission and Plasticity'

EDUCATION

09/2020 - **Post-graduation in Data Science -** University of Lisbon **LISBON, PORTUGAL**

Relevant Coursework: Artificial Intelligence, Advanced Machine Learning, Data Mining, Advanced Databases, Web Applications and Statistics

09/2017 - 03/2021 **MSc in Neuroscience -** University of Lisbon **LISBON, PORTUGAL**

Relevant Coursework: Bayesian Statistics, Programming I, Biostatistics, Neurophysiology, Neuroanatomy, Neuroimaging, Neuroethics, Molecular and Celular Neurobiology and Neuropsychology

09/2013 - 07/2016 **BSc in Biochemistry -** University of Coimbra **COIMBRA, PORTUGAL**

Relevant Coursework: Bioinformatics, Mathematics, Statistics, Laboratory work, Spectroscopy, Differentiation and Development and Neuropsychology

SKILLS

- Excellent at problem solving and critical thinking;
- Good scientific communication skills;
- Strong programming skills with 5 years of experience in scientific computing;
- Programming languages: Python (preferred), R, Matlab, SQL;
- Language Skills: Native in Portuguese; Fluent in English.