LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-796377

Final Report: Brain Connectomics: Opportunities for High-Performance Computing

A. D. Kaplan

November 7, 2019

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

FINAL REPORT
Brain Connectomics: Opportunities for High-Performance Computing
Alan Kaplan (18-FS-009)

IM release #LLNL-JRNL-XXXXXX

Abstract
The human brain is estimated to contain approximately 100 trillion neural connections. This complex map of connectivity (human connectome) underlies cognitive processes and informs disease states, but until recently, studying the human connectome has been hampered by the difficulty of collecting and processing suitable data for analysis. In this context, an initiative denominated The Human Connectome Project (HCP) was created to acquire and collect the largest, most cohesive set of brain data to accelerate neuroscience research. Recently the HCP released a magnetic resonance (MR) dataset of brain imagery with unprecedented scale and organization, together with high quality spatial and temporal resolution; however, fully taking advantage of this high-resolution brain imagery to estimate neurological connectivity networks requires high-performance computing (HPC). In this project, we explored the feasibility of using HPC resources at Lawrence Livermore to effectively estimate brain connectivity networks from this large set of imagery. Specifically, we addressed one of the main current challenges in brain connectomics, which is to advance the computation of individual structural and functional connectomes. To do this, we hosted the latest version of the HCP dataset and developed HPC software to compute computationally expensive tasks required to generate connectomes on the entire HCP cohort (~1200 subjects). This project has supported the NNSA goal of advancing the scientific, technical, and engineering competencies that are the foundation of the NNSA mission. Specifically, by strengthening our ability to compute with complex biological data and by constructing the efficient, datacentric software required to successfully achieve our goals, this project enhances the Laboratory's core competencies in HPC, simulation, and data science, as well as bioscience and bioengineering.

Background and Research Objectives
While many openly available MR datasets exist, the sample sizes are generally too small to increase confidence that any results are generalizable from one study to another, or from one cohort to another. This requires the development of careful and consistent experimental protocols, not only to ensure that the same exact imaging parameters were used from subject to subject, but also to standardize the collection of demographic and health information for all subjects.

In order to address this deficiency, a consortium form Washington University in St. Louis, the University of Minnesota, and Oxford University called the Human Connectome Project (HCP) was formed [Glasser, Smith 2016]. A dataset has been openly released as part of the HCP, which consists of both diffusion and functional MR imagery. The 2017 release contains structural and functional data collected from 1,200 healthy adult subjects. A unified protocol was used to record these scans, resulting in a homogeneous dataset.

Our objective is to demonstrate the capabilities of HPC for computing brain connectivity,

centered around the following goals:

1. Demonstrate high resolution tractography for the computation of structural connectivity,
2. Demonstrate fast functional connectivity computations.

The final connectome computation for both structural and functional versions requires and anatomical atlas. In this project, we used 3 different atlases:

- Freesurfer Desikan-Killiany [Desikan 2006];
- Freesurfer updated [Destrieux 2010];
- Glasser [Glasser, Coalson 2016].

Scientific Approach and Accomplishments

Structural connectivity refers to the extraction of mesoscopic axonal connections between neurons, or tractography. This is a computationally expensive process that leverages diffusion imaging, where image contrast is obtained from the asymmetric nature of water diffusion around myelinated axons (diffusion anisotropy). This step is vital in almost any type of brain imagery analysis, because it provides the roadmap on which activity (function) is measured. Any following computations are limited by this step, so it is important that it is performed carefully and at the highest resolution possible.

We leveraged an established tractography software package called MRtrix to compute high resolution parcellation form the diffusion imagery. When performing tractography, a critical problem is that there may be multiple neuronal fibers intersecting within a single diffusion MR voxel. In fact, this is commonly the case. As opposed to older techniques, that are only capable of estimating a single fiber per voxel, this approach can delineate multiple fibers per voxel. The output of the algorithm is a probabilistic representation, where each tract has some likelihood of connectivity associated with it, instead of relying on a firm deterministic estimate. Figure 1 shows the pipeline that we assembled to compute structural connectivity. With reference to the tools used, the steps involved were:

- Compute tissue segmented regions (5ttgen fsl);
- Estimate fiber orientation distributions using spherical deconvolution (dwi2fod) [Tournier 2004];
- Anatomically constrained tractography using iFOD2 (tckgen) [Smith 2012];
- Perform spherical-deconvolution informed filtering of tractograms (tcksift) [Smith 2013].

A key parameter in this algorithm is the number of random seeds used. The number of seeds is a critical value on which the resulting connectivity resolution depends. We performed streamline tractography using 25 million seeds per subject using a dynamic seeding methodology which oversamples under-represented regions. With this approach we were not able to restart tractography to push beyond the 25 million seed count. However, we also implemented scripting software that can run a uniformly sampled version of tractography which can concatenate multiple tractography runs. This allows for a large number of streamlines to be computed. Using this approach, we computed 40 million streamlines for a subject of approximately 450 unrelated subjects. Figure 2 shows an example of the tracts for one subject. Our computational output includes 27 structural connectomes per subject (3 parcellations, 3 measures, and 3 sessions).

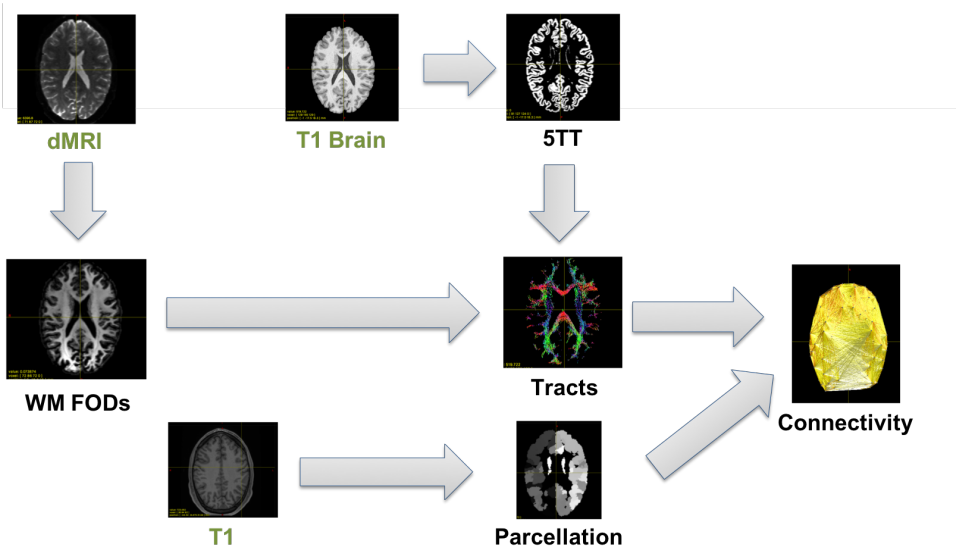

Figure 1. Flow chart of the structural connectivity pipeline. The components in the diagram are: *dMRI* (diffusion MRI imaging data), *WM FODs* (white matter fibre orientation distributions), *T1 Brain* (T1 MRI image with the brain extracted), *5TT* (5 tissue type image segmentation), *Tracts* (streamline tracts resulting from tractography), *T1* (T1 MRI image), *Parcellation* (volumetric anatomical parcellation), and *Connectivity* (anatomical region-region structural connectivity measures).

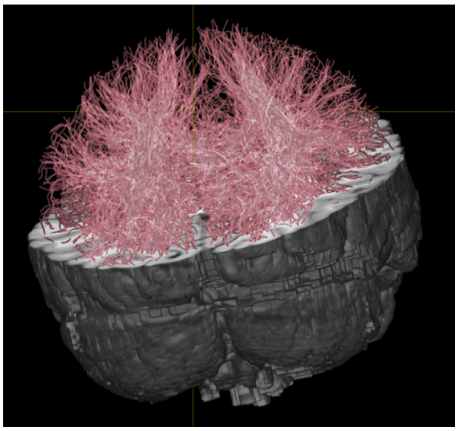

Figure 2. Example of white matter tracts (red) computed from the diffusion MRI along with the corresponding partial T1 structural MRI brain volume.

Functional connectivity quantifies the spatial co-occurrence of neuronal function. The imaging in this case is performed with Blood Oxygenation Level-Dependent (BOLD) contrast, which measures differences in oxygenated vs. deoxygenated blood. Areas of the brain that are more active have greater blood oxygen levels, and thus we are able to measure which areas of the brain are active. Existing computational techniques are predominantly based on the notion of pairwise similarity. A similarity measure, such as empirical correlation or Granger causality, is computed for all voxel or region pairs within the imagery, possibly across many subjects. A subset of all pairs is then determined by thresholding the computed similarity measures, resulting in a connectivity network. The resulting network may then be analyzed

to determine brain properties of interest, hub-like regions, pathological conditions, etc.

As opposed to structural MR imaging, many volumes are acquired for functional imaging over the course of approximately 1 hour (1200 volumes for each 15 minute run). In this project, we will focus on computing functional connectivity using pairwise correlations of the spatial BOLD signals. A correlation matrix is computed over every subject, visit (i.e. test-retest) and task (i.e. resting-state, working-memory, etc). Figure 3 shows the basic steps involved in computing functional connectivity. Our computational output results in 54 functional connectomes per subject (3 parcellations, 9 sessions, 2 gradient directions).

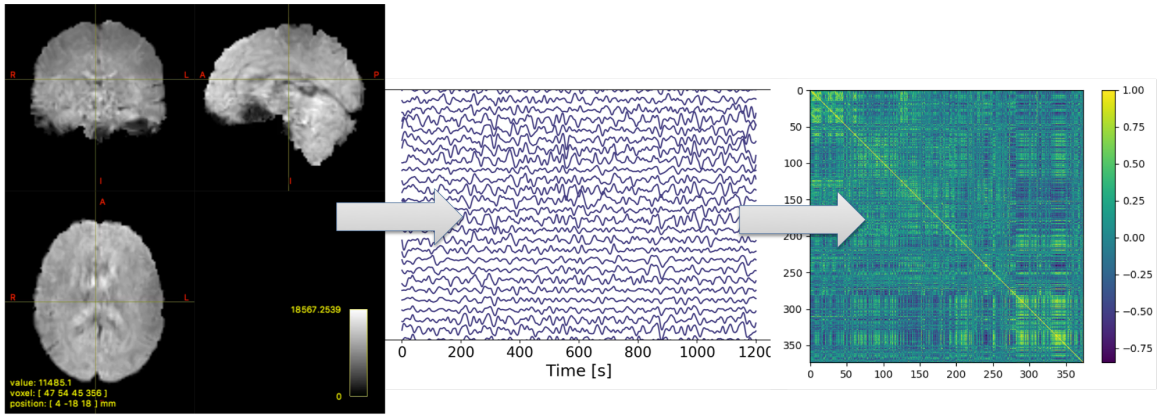

Figure 3. Diagram showing computation of the functional connectivity matrix. Anatomically parcellated timeseries are produced from the fMRI data, from which the connectivity matrix is computed.

Impact on Mission

This project will strengthen capabilities within the *High-Performance Computing, Simulation and Data Science* core competency, because it addresses computing network topology at scale from large amounts of data. The ability to compute with complex biological data and construct efficient data centric software is required to successfully complete this project. In addition, this project will pave the way for new scientific and neural engineering advances (*Bioscience and Bioengineering* core competency). The *Chemical and Biological* mission focus area may benefit from an increased understanding of brain function because of this project.

Conclusion

This work has paved the way for future work in the area of computational neuroimaging. We have identified additional potential collaborators in the space. These collaborations would leverage the data produced, software, and expertise developed in this project. The opportunity to quantify the structural connections inside the brain at a high level of detail are available to us. Instead of working with pairwise measures that describe the connectivity between a few hundred anatomical regions, we now have the capability to express complex geometric networks in the brain. This was not possible without the use of high-performance computing systems. We now turn to the development of novel machine learning and computer vision approaches to model and characterize the complexity that we have unearthed from the imagery. As we look to expand into this exciting and important area of research, this project will have served to determine foundational feasibility of the

computational methods.

We have obtained permission to release the derived data that were computed on this project, which we anticipate will be of great utility to the scientific community. This upcoming data release will serve to introduce our capabilities to the neuroscience community as well as contribute to the DOE mission of accelerating Artificial Intelligence through unique and challenging datasets.

References

Desikan, Rahul S., Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, et al. 2006. "An Automated Labeling System for Subdividing the Human Cerebral Cortex on MRI Scans into Gyral Based Regions of Interest." *NeuroImage* 31 (3): 968–80.

Destrieux, Christophe, Bruce Fischl, Anders Dale, and Eric Halgren. 2010. "Automatic Parcellation of Human Cortical Gyri and Sulci Using Standard Anatomical Nomenclature." *NeuroImage* 53 (1): 1–15.

Glasser, Matthew F., Stephen M. Smith, Daniel S. Marcus, Jesper L. R. Andersson, Edward J. Auerbach, Timothy E. J. Behrens, Timothy S. Coalson, et al. 2016. "The Human Connectome Project's Neuroimaging Approach." *Nature Neuroscience* 19 (9): 1175–87.

Glasser, Matthew F., Timothy S. Coalson, Emma C. Robinson, Carl D. Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, et al. 2016. "A Multi-Modal Parcellation of Human Cerebral Cortex." *Nature* 536 (7615): 171–78.

Smith, Robert E., Jacques-Donald Tournier, Fernando Calamante, and Alan Connelly. 2012. "Anatomically-Constrained Tractography: Improved Diffusion MRI Streamlines Tractography through Effective Use of Anatomical Information." *NeuroImage* 62 (3): 1924–38.

Smith, Robert E., Jacques-Donald Tournier, Fernando Calamante, and Alan Connelly. 2013. "SIFT: Spherical-Deconvolution Informed Filtering of Tractograms." *NeuroImage* 67 (February): 298–312.

Tournier, J-Donald, Fernando Calamante, David G. Gadian, and Alan Connelly. 2004. "Direct Estimation of the Fiber Orientation Density Function from Diffusion-Weighted MRI Data Using Spherical Deconvolution." *NeuroImage* 23 (3): 1176–85.

Publications and Presentations

Kaplan, Alan D, 2019. "Large Scale Connectome Generation." Presentation at Society for Brain Mapping & Therapeutics.