Exploring the Structure of an Individual's Resting Brain using Topological Data Analysis

Rafi Ayub

Department of Bioengineering, Stanford University

rafiayub@stanford.edu

I. Introduction

The brain is a multitasking machine; while it manages the effortless heartbeats and breaths that keep it alive, it is also able to yield intense focus on reading a paper, performing mathematical calculations, or driving a car. Neuroscience has explored the functional repertoire of the brain by pinpointing the anatomical correlates to hundreds of simple tasks and imaging the evolution of brain activity during cognitive demands. Yet, there is still no certainty on what the brain does when it is at rest, performing no task at all.

Scientists, philosophers, and the everyday thinker posit that the mind wanders, daydreams, ruminates, reflects, and plans. This rich palette of cognitive behaviour has found some basis within neuroimaging. For example, functional MR imaging studies have observed correlations between distant brain regions in spontaneous activity during rest, deemed resting state functional connectivity (FC) [1] [2]. Across a longer time interval of resting state activity, patterns of correlated networks and sub-networks form and dissolve in simulations and in empirical data [3]. In fact, many of these canonical resting state networks (RSNs) have been found across many studies and have corresponded to critical brain functions such as movement, attention, and vision. Interestingly, these networks and connectivity between certain regions may be impaired in neuropsychiatric disorders such as Alzheimer's disease and depression [4]. Clearly, the brain at rest is a goldmine that could yield key insight into its function and dynamics.

Current methods to characterize resting state FC involve timeseries correlations between regions, sliding-window correlations, deconvolution, temporal Independent Component Analysis, and more. However, these methods collapse data over time and fail to capture clinically relevant temporal dynamics. To avoid arbitrarily collapsing fMRI data over space or time, a tool from the field of Topological Data Analysis called Mapper has been proposed. Mapper creates a combinatorial object from a high dimensional dataset that depicts the manifold of the original data. By using metrics from graph theory, clinically and biophysically relevant insight can be captured from a Mapper graph applied to resting state fMRI data. This approach has been previously used to capture task transitions at a faster time scale than other methods and to predict individual task performance [5].

In this study, we propose to use Mapper to explore the structure of RSNs in resting state fMRI data. We will use 84 cleaned scan sessions from the dataset provided by MyConnectome, which consists of structural and functional MR scan sessions, metabolic profiles, mood questionnaires, and daily activity logs of the same subject for about a year. Specifically, we will analyze the community and core-periphery structure of resting state networks. This approach is based on a previous study in [5], in which the authors correlated modularity and "coreness" of nodes with task performance in an individual. Community and core structure will provide us insight into the diversity of network dynamics in the individual's brain. We hypothesize that a more modular, smaller core graph of the patient's brain correlates with a more alert profile for that scan day, and attention related RSNs congregate in the core more with an alert profile than a less alert profile. By characterizing the structure of the brain at rest and its relationship to behaviour, we can become better equipped to predict, diagnose, and treat neuropsychiatric disorders.

II. Related Work

A. Network dysregulation in neuropsychiatric disorders

Neuropsychiatric and behavioural disorders are hypothesized to be linked to macroscale brain network dysregulation. Thus, many studies have applied graph theory metrics to functional connectivity to explore differences in network dynamics between healthy and patient populations. In the study by Xu et al. [6], the team investigated network abnormalities in borderline personality disorder (BPD), which involves symptoms such as affect dysregulation, impaired sense of self, and self-harm behaviours. To this end, they acquired resting state fMRI data from 20 patients with BPD and 10 healthy controls. They created networks for each subject by taking the correlations between each of 82 cortical and subcortical regions and thresholding to yield a graph density of 0.1. These graphs were analyzed using clustering coefficient, characteristic path length, small-worldness, local efficiency, global efficiency, and degree and correlated with clinical symptom scores. Finally, the study used network features in a machine learning classifier to distinguish BPD patients from healthy controls. The team found that BPD patients exhibited increased size of largest connected component, amount of local cliques, clustering coefficient, local efficiency, and small-worldness. These network measures demonstrated high predictive power when implemented with a classifier.

This study is important in demonstrating the potential utility of analyzing network measures of brain activity to

predict mental health clinical symptoms or diagnose neuropsychiatric disorders. Indeed, the study was able to infer behaviours characteristic of BPD from the significant network measure differences. For example, higher levels of local cliquishness at the amygdala and temporal poles may suggest a rapid rise in negative affect that is difficult to regulate in BPD patients. This type of insight is key to understand the mechanisms behind psychiatric illnesses.

However, this study does not take into consideration temporal dynamics and network fluctuations during resting state activity. Although they were able to successfully train an accurate classifier with their static network, dynamic functional connectivity is necessary to elucidate the neurophysiological underpinnings of BPD. We propose that Mapper can capture temporal dynamics and provide a more mechanistic understanding.

B. Resting state fluctuations within a scan

There is some recent work that does indeed explore network fluctuations in resting state brain activity. One such study by Kabbara et al. [7] took dense-EEG data from 20 subjects, created functional networks, and analyzed the dynamics of network measures and nodal properties over 40 seconds. Specifically, they found the nodes with the highest centrality, vulnerability, strength, and clustering coefficient and categorized them by the type of resting state network they belong to (default mode network, dorsal attention network, salient network, auditory network, visual network). They found that the dynamic analysis using sliding windows yielded very different results than the static analysis where one network was created for the entire time interval, again highlighting the need to incorporate temporal dynamics in network analysis of the brain. Interestingly, they found that the most central and vulnerable nodes tend to be from the default mode network (DMN). Even if they belong to other networks, the brain tends to move the central and vulnerable nodes to the DMN. They also found that the provincial and connector hubs were by a large fraction DMN nodes. These results corroborate other studies suggesting that the DMN is the most dominating resting state network, and that the rich club nodes of the human brain tend to belong to the DMN.

Although this is one of the few studies that does explored temporal dynamics, a few areas could be optimized. For instance, tracking network measures across discretized time may not be the best representation of temporal dynamics. Creating a network at each window can be very susceptible to noise and motion artifacts. The dynamics that they observe could potentially be random fluctuations between time windows. The study failed to compare these graphs to random graph models to see if the changes between time intervals were not also evident in null models. It may also be worthwhile to capture the entirety of the recording session into a network so we can observe how the brain traverses the graph over time. This may help us explore the brain state space of resting state activity that we sample from every time we scan or record from a subject, a direction Mapper can help us explore. Although this study demonstrates the brain's reliance on DMN regions, it does not show how the different resting state networks fluctuate over a period of time, which may be important for understanding healthy and pathological resting state brain dynamics. Mapper can provide this single network representation of an entire scan that still incorporates temporal dynamics.

Another drawback is that this study condenses the voxel space into 68 regions. Most studies on functional connectivity also collapse the spatial dimension, but perhaps using data at the level of voxels could create a more accurate representation. Although Mapper projects the data into a lower-dimensional representation, the lens function that is used for projection can be a non-linear dimensionality reduction algorithm that learns the manifold of the input data. Additionally, the partial clustering is done in the original high-dimensional space. Thus, information of the raw voxel data is preserved. We propose that Mapper is a more robust representation of brain dynamics since it does not collapse the data temporally or spatially. More details of the Mapper algorithm will be explained in Methods.

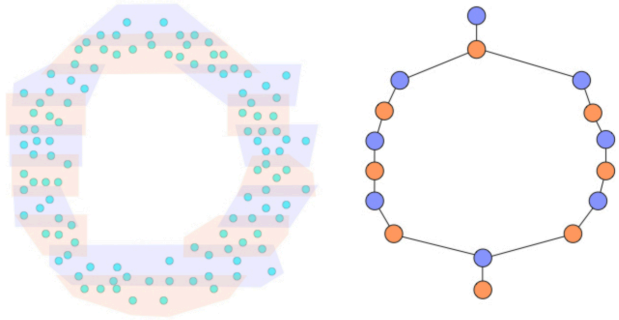

Figure 1: Constructing a simplicial complex by binning the original dataset from the lens function and clustering within the bins. Adapted from https://sauln.github.io/blog/mapper-intro/

III. METHODS

A. Data Collection

The specific protocols are detailed on the MyConnectome website (myconnectome.org/wp/), but will be discussed here briefly. Resting state fMRI scans were performed three times a week (Monday, Tuesday, Thursday), using a multi-band EPI sequence (TR=1.16 ms, TE = 30 ms, flip angle = 63 degrees, voxel size = 2.4 mm X 2.4 mm X 2 mm, distance factor=20%, 68 slices, oriented 30 degrees back from AC/PC, 96×96 matrix, 230 mm FOV, MB factor=4, 10:00 scan length). Gradient echo field maps and spin echo field maps with AP and PA phase encoding were also collected. Behavioral/lifestyle measurements were also collected daily and are detailed in Table 1. Other measurements include sleep, exercise, amount of time outside, blood pressure, pulse, diet, blood sampling, RNA sequencing, and metabolics, though this list in non-exhaustive and the acquisition will not be detailed here. We will also note that on Tuesdays the subject was fasted due to a blood draw that same day, and other days the subject was not fasted. The fMRI scans were preprocessed using fmriprep, an open-source pipeline [8].

Measurement	Instrument
Morning	
Sleep quality	1-7 rating scale
Physical soreness	1-7 rating scale
Blood pressure	
After Each MRI Scan	
Mood	PANAS-X (all 60 questions)
Blood pressure	
Anxiety during scan session	1-7 scale
thoughts during scan	free text
Evening	
Time spent outdoors	Estimated # of minutes
Psoriasis severity	1-7 rating scale
Stress	1-7 rating scale
Gut health	1-7 rating scale
Mood	PANAS-X (short form)
Open-ended report	Free text
Diet	Free text
Alcohol	Floating point

Table 1: Behavioral measurements and the time of day they were recorded.

B. Mapper

Details of the Mapper algorithm are described in [9], but will be briefly discussed here. Essentially, a lens function is applied to the original high-dimensional data to create a low-dimensional representation of the data, called the cover. The datapoints in the cover are binned into overlapping windows. Then, the corresponding original high-dimensional datapoints are clustered based on the binning. These clusters become the nodes of the resultant graph, and edges are defined between nodes when clusters share one or more original datapoints, which is possible due to the overlap. Put very simply, the structure of the resultant graph depicts the similarity of the original datapoints. In this study, we used tSNE, or t-distributed stochastic neighbour embedding [10], for our lens function. We used HDBSCAN [11] as the clusterer. tSNE was chosen because it preserves some of the local structure in the high-dimensional space, since it is a non-linear method. HDBSCAN is a hierarchical clustering algorithm that was used because it does not require the number of clusters to be specified. The number of cubes (bins) was set to 10, and the overlap percentage was set to 60%.

Mapper was applied to each scan session represented by a 518 x 630 (TRs x number of parcels) data matrix. The lens function mapped this to a 518 x 2 matrix. Thus, we are created a Mapper graph in the temporal space, though we are also able to transpose the data matrix and create a graph in the anatomical space, which may provide additional unique insight.

C. Resting state network activations

One of the advantages of Mapper is the ability to annotate nodes with metadata corresponding to the members of each node. This allows us to visualize the localization of certain points of interest. For resting state networks, we can label each member with the network(s) that is/are activated and see how RSNs localize on the Mapper graph.

Each parcel's timeseries was Z-scored. Then, for each RSN, we calculated the mean activity at each timepoint for all the parcels that belong to that RSN. An RSN is considered activated when the mean at a given timepoint surpasses a threshold of $Z > 0.5$. Each member can have multiple network activations, which is, in fact, the most common case. We labelled timepoints with 13 known RSNs, which are described in Table 2.

The resultant graph contains a pie chart for each node, which are proportionally colored by the network activations of the node's members. We can also color the graph by individual RSNs, where the color indicates activation or not.

D. Community Structure

Communities are defined as groups of densely interconnected nodes with sparse connections between groups. We can assign nodes into communities and evaluate the "goodness" of the assignments by calculating a measure known as modularity. Modularity, Q, is defined below, where A is the adjacency matrix of the graph, k is the node degree, m is the total number of edges, and δ returns 1 if both node v and w are in the same community.

$$Q = \frac{1}{(2m)} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{(2m)} \right] \delta(c_v, c_w)$$

We defined communities for each node by the RSN most of its members are labelled by. This allows us to observe how modular resting state networks tend to be.

We ran Louvain community detection on the Mapper graphs to see how well RSNs modularized on their own. In brief, each node is initially assigned to its own community and are reassigned to new communities if the change in modularity is greater than the current modularity. This is repeated until modularity is maximized. Then the communities are compressed into supernodes and the process repeats. The equation for the change in modularity is calculated by the expression below.

$$\Delta Q = \left[\frac{\Sigma_{in} + 2k_{i,in}}{2m} - \left(\frac{\Sigma_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right]$$

Network	Functions	Citation
Default Mode	Emotional processing, self-referential mental activity, recollection	[12]
Dorsal Attention	Covert spatial attention, saccade planning, visual working memory	[13]
Ventral Attention	Attention to unexpected stimuli	[13]
Fronto-parietal	Selection of stimuli for attention	[14]
Cingulo-opercular	Tonic alertness	[15]
Salience	Selection of stimuli for attention, initiation of cognitive control, maintenance of tasks	[16]
Somatomotor	Motor planning and execution, processing sensory input	[17][18]
Visual	Visual perception, processing, attention	[18]
Medial Parietal	Memory	[19]
Parietal Occipital	Visuomotor planning and control	[20]

Table 2: Resting state networks explored in this study and their functions.

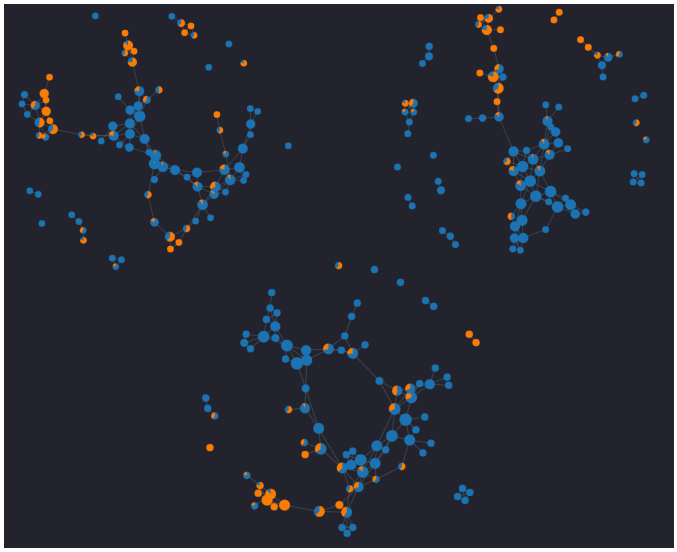

Figure 2: Mapper graphs from three sessions, where orange indicates the presence of timepoints/members with default mode network activation. Seemingly, default mode network timepoints tend localize in the periphery of the network, but it is difficult to say anything with only three samples. Further characterization of the core-periphery structure for each RSN is needed.

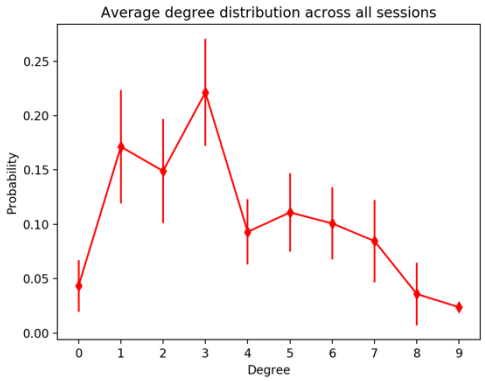

Figure 3: Average degree distribution across all 84 sessions.

IV. Initial Results

We created a Mapper graph for each of the 84 sessions. Sample graphs are shown in Figure 2, annotated by default mode network activation. For the sampled graphs, the default mode network tends to activate near the periphery of the structure. It may be worthwhile to investigate whether some RSNs tend to congregate in the core or the periphery and what that means functionally.

Interestingly, the degree distribution of the Mapper graphs shown in Figure 3 do not follow a clear power law distribution. However, these networks are very small compared to the ones for which the power law distribution is apparent.

We assessed the community structure given by the RSN annotations by comparing the modularity of the assignments with the modularity of the optimal assignments. The Python implementation of the Louvain algorithm allowed a resolution parameter to control for the number of resultant communities, so we altered the resolution until the number of communities matched the RSN assignments. To measure the similarity between the assignments, we calculated the Adjusted Rand Index. For the session depicted in Figure 4, we got a Rand Index of 0.276, indicating that the RSN communities are not the most optimal community organization of the Mapper graphs, although there does seem to be some modular structure. The RSN modularities tended to fall between 0.2 and 0.3 across all sessions with a few outliers, but more analysis needs to be done to correlate this with behavioural metrics for alertness.

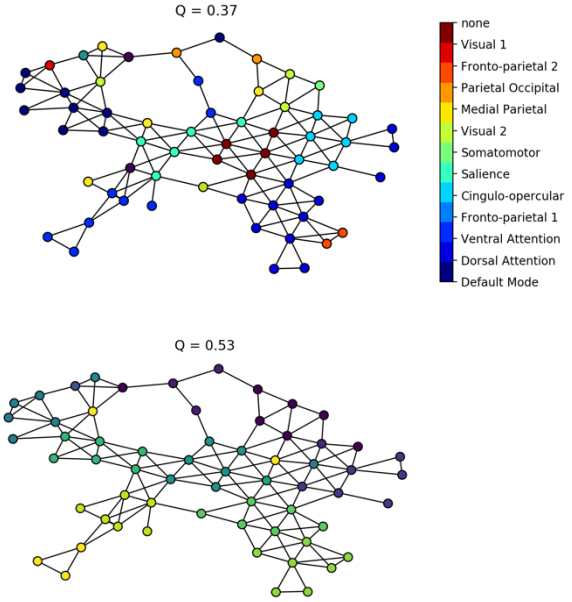

Figure 4: (top) Community assignments from RSN activations and the resulting modularity. Surprisingly, some modular structure is present, though comparison to a null model is needed. (bottom) Optimal community assignments using Louvain's method and a predefined number of communities. The modularity is much higher, indicating RSN communities are not perfectly modular.

V. Next Steps

Our hypothesis is that the network structure is characterized differently between alert and non-alert behavioural profiles in terms of core-periphery and community structure. More analysis needs to performed on relating the RSN communities and modularities to behavioural scores. Additionally, to characterize core structure, we plan to investigate how often each RSN localizes in the core or the periphery of the network. This may tell us which patterns of brain activity tends to be most common or central during the resting state and which tend to be deviations from a baseline, whether intentional or unintentional. Furthermore, we plan to characterize the Mapper graphs with other network measures such as average shortest path length, clustering coefficient, and efficiency to explore questions about the diversity of resting state dynamics. For example, if the resting brain exhibits very monotonous activity, the Mapper graph may have a high clustering coefficient and low diameter. Conversely, if the resting brain is very multifaceted, it may exhibit a lower clustering coefficient and high diameter.

Most importantly, we will explore null models to compare the Mapper graphs to. We plan to use the statistical model proposed by Laumann et al. [21] to generate timeseries data similar to the empirical fMRI and run Mapper on the resulting data matrix. We can use this null Mapper as a benchmark for many of the community and core-periphery characterizations.

Finally, we can create one giant Mapper graph on all 84 sessions and annotate members by session number. This can help extrapolate which sessions are more similar to others and provide some insight into why some sessions deviate, and what behaviours might relate to this.

VI. Acknowledgements

The author would like to acknowledge Caleb Geniesse, who is a partner for this project in another class, PSYCH 267: Brain Networks. He worked on the code for creating all the Mapper visualizations. This project is joint across CS 224W and PSYCH 267, but the work presented here was specifically for this class.

References

[1] Glomb, K., Ponce-Alvarez, A., Gilson, M., Ritter, P., & Deco, G. (2017). Resting state networks in empirical and simulated dynamic functional connectivity. Neuroimage, 159 (November 2016), 388–402. https://doi.org/10.1016/j.neuroimage.2017.07.065

[2] Hansen, E. C. A., Battaglia, D., Spiegler, A., Deco, G., & Jirsa, V. K. (2015). Functional connectivity dynamics: Modeling the switching behavior of the resting state. NeuroImage, 105, 525–535. https://doi.org/10.1016/j.neuroimage.2014.11.001

[3] Deco, G., Jirsa, V. K., & McIntosh, A. R. (2013). Resting brains never rest: Computational insights into potential cognitive architectures. Trends in Neurosciences, 36(5), 268–274. https://doi.org/10.1016/j.tins.2013.03.001

[4] Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. Current Opinion in Neurology, 24(4), 424–430. https://doi.org/10.1097/WCO.0b013e328306f2c5

[5] Saggar, M., Sporns, O., Gonzalez-Castillo, J., Bandettini, P. A., Carlsson, G., Glover, G., & Reiss, A. L. (2018). Towards a new approach to reveal dynamical organization of the brain using topological data analysis. Nature Communications, 9(1), 1–14. https://doi.org/10.1038/s41467-018-03664-4

[6] Xu, T., Cullen, K. R., Mueller, B., Schreiner, M. W., Lim, K. O., Schulz, S. C., & Parhi, K. K. (2016). Network analysis of functional brain connectivity in borderline personality disorder using resting-state fMRI. NeuroImage: Clinical, 11, 302–315. https://doi.org/10.1016/j.nicl.2016.02.006

[7] Kabbara, A., EL Falou, W., Khalil, M., Wendling, F., & Hassan, M. (2017). The dynamic functional core network of the human brain at rest. Scientific Reports, 7(1), 2936. https://doi.org/10.1038/s41598-017-03420-6

[8] Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack RA, Gorgolewski KJ. FMRIPrep: a robust preprocessing pipeline for functional MRI. bioRxiv. 2018. 306951; doi:10.1101/306951

[9] Singh, G., Memoli, F., & Carlsson, G. (2007). Topological Methods for the analysis of high dimensional data sets and 3D object recognition. Eurographics Symposium on Point Based Graphics.

[10] Maaten, L. & Hinton, G. Visualizing data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605 (2008).

[11] McInnes L, Healy J. Accelerated Hierarchical Density Based Clustering In: 2017 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE, pp 33-42. 2017

[12] Raichle, M. E. (2015). The Brain's Default Mode Network. Annual Review of Neuroscience, 38(1), 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

[13] Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. Neuroscientist, 20(2), 150–159. https://doi.org/10.1177/1073858413494269

[14] Ptak, R. (2012). The frontoparietal attention network of the human brain: Action, saliency, and a priority map of the environment. Neuroscientist, 18(5), 502–515. https://doi.org/10.1177/1073858411409051

[15] Sadaghiani, S., & D'Esposito, M. (2015). Functional characterization of the cingulo-opercular network in the maintenance of tonic alertness. Cerebral Cortex, 25(9), 2763–2773. https://doi.org/10.1093/cercor/bhu072

[16] Ham, T., Leff, A., de Boissezon, X., Joffe, A., & Sharp, D. J. (2013). Cognitive Control and the Salience Network: An Investigation of Error Processing and Effective Connectivity. Journal of Neuroscience, 33(16), 7091–7098. https://doi.org/10.1523/JNEUROSCI.4692-12.2013

[17] Sánchez-Castañeda, C., de Pasquale, F., Caravasso, C. F., Marano, M., Maffi, S., Migliore, S., … Squitieri, F. (2017). Resting-state connectivity and modulated somatomotor and default-mode networks in Huntington disease. CNS Neuroscience and Therapeutics, 23(6), 488–497. https://doi.org/10.1111/cns.12701

[18] Heine, L., Soddu, A., Gómez, F., Vanhaudenhuyse, A., Tshibanda, L., Thonnard, M., … Demertzi, A. (2012). Resting state networks and consciousness Alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states. Frontiers in Psychology, 3(AUG), 1–12. https://doi.org/10.3389/fpsyg.2012.00295

[19] Power, J. D., Schlaggar, B. L., & Petersen, S. E. (2014). Studying brain organization via spontaneous fMRI signal. Neuron, 84(4), 681–696. https://doi.org/10.1016/j.neuron.2014.09.007

[20] Hutchison, R. M., Culham, J. C., Flanagan, J. R., Everling, S., & Gallivan, J. P. (2015). Functional subdivisions of medial parieto-occipital cortex in humans and nonhuman primates using resting-state fMRI. NeuroImage, 116, 10–29. https://doi.org/10.1016/j.neuroimage.2015.04.068

[21] Laumann, T. O., Snyder, A. Z., Mitra, A., Gordon, E. M., Gratton, C., Adeyemo, B., … Petersen, S. E. (2017). On the Stability of BOLD fMRI Correlations. Cerebral Cortex, 27(10), 4719–4732. https://doi.org/10.1093/cercor/bhw265