GSP Applications in Network Neuroscience

Prakhar Gupta*
*ECE 396 Final Project

***Abstract*—This paper is the report for an exploratory research project for ECE 396. It covers an introduction to Network Neuroscience and Graph Signal Processing. It also goes over preprocessing steps, inferring graph structure from data and possible applications of GSP in this area. Results indicate that Phase Locking Value (PLV) is a a better metric than correlation for functional connectivity analysis. While results are fairly limited due to a lack of data, a lot of promising avenues for future work are discovered.**

***Index Terms*—Graph Signal Processing, Network Neuroscience, Functional Connectivity,**

I. Introduction

The human brain has been subject to research interest for multiple centuries now. Brain connectomics is a relatively new area of study that has sprung with the rise of imaging techniques such as EEG and MRI. Brain connectomics is the study of interactions of networks in the brain[1] - in hopes of shedding light on clinical disorders such as dementia, and schizophrenia, and just better our understanding of how the brain operates.

There exist primarily 2 forms of Brain Networks (BNs) - Structural and Functional Networks. These networks consist of vertices that correspond to specific brain regions, and measuring edges using a pair-wise computation[2]. Edges in structural networks are derived from the physical structure and layout of the brain (physical distance, connected surface area, etc). Functional Network edges are evaluated based on the statistical relation between levels of activity in different regions of the brain - typically using Blood Oxygenation Levels (BOLD) in fMRI, or more recently, EEG signals (with the aid of source localization techniques).

Traditional connectomics involves measuring graph-theoretical metrics of these networks, encapsulating global and local characteristics. These include path efficiency, average node degree, small-worldness, etc. These metrics, however, only provide insight into the average behavior of the network in a given time slice and fail to provide information about network dynamics.

Graph-Signal Processing (GSP) offers a powerful tool for this problem, as it takes into account brain activity as well as the topology of the underlying network structure.

II. GSP Fundamentals

Table I covers some commonly used notations in GSP. $\mathbf{W}_{ij}$ is the weight of the edge between the *i*'th and *j*'th vertices in the graph. The degree matrix $\mathbf{D}_{ii} = \sum_{j=1}^{N} \mathbf{W}_{ij}$, which is the sum of all edges for each node. The Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{W}$, is an important operator in GSP, and approximates the Laplacian operator in the graph-domain, and represents the 'smoothness' of a graph[3]. The eigenvectors of $\mathbf{L}$ make up $\mathbf{U}$. A graph signal is simply a set of values, each of which are mapped to a node on the graph.

Symbol	Explanation
W	Weighted Adjacency Matrix
D	Weighted Degree Matrix
L	Graph Laplacian
U	Laplacian Eigenvectors (Graph Fourier Basis)
x	Graph signal

TABLE I
GSP Notation.

Together, these basic units form the fundamentals of GSP, giving us an effective toolkit to manipulate data with.

III. Preprocessing Steps

Obtaining graph signals from a set of time-series signals was a fairly involved task with several steps, particularly to reduce noise and perform accurate source localization. The raw data consisted of 100s worth of 64-channel EEG recordings (at 512Hz), and the final processed data was 100s worth of 102-node source space signals (at 512Hz).

A. Dropping Noisy Artifacts

First, bad channels in the original EEG recordings were identified by comparing the autocorrelation of the signals received at each node. Additional electrodes were marked as noisy after a visual inspection. These channels were then dropped and interpolated using the remaining data.

Next, we performed ICA on the signals in order to identify the labels for sources in accordance to the SCCN ICLabel toolbox. This allowed us to identify and minimize contributions from artifacts such as eye-blinking, heartbeats, and muscle twitches. This ensures that a large amount of activity observed in the reconstructed signals is largely from brain components.

B. Source Localization (eLORETA)

Exact Low-Resolution Electromagnetic Tomography (eLORETA) is a technique used to estimate the internal activity of the brain, using surface-level data. The resulting brain source data can then be used to obtain a graph signal.

eLORETA involves solving a forward and inverse problem. The forward problem is computing the potentials on the scale produced by 'dipole models' in the brain, and the inverse allows us to estimate signals in the source space. We used a FreeSurfer template[4] of the head, which encapsulates information about the location of dipoles as well as the non-uniform conductivity of fields inside the human skull.

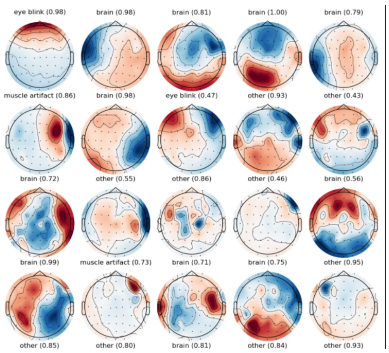

Fig. 1. 2D visualization of ICA components on head topology (top-view). Only components with a probability ¿0.80 of being '*brain*' or '*other*' were used in signal reconstruction.

This procedure yielded about 40,960 dipoles per hemisphere, which were then averaged into 102 signals, each corresponding to a 'parcel' or region in the brain. These averages were computed by grouping dipoles in certain region specified by a map (Schaefer's Atlas[6]).

After having completed all the preprocessing steps, we now have a graph signal $\mathbf{x}$ which has the shape $(102, 51200)$ - 102 nodes, each with 512Hz $\times$ 100s worth of samples. This only leaves generating the $\mathbf{W}$ in order to start applying GSP techniques. The code used for preprocessing can be found here. A bulk of these preprocessing steps were performed using the MNE Python Toolkit[7]

IV. Graph Structure Inference

Inferring graph structure from data is non-trivial. However, doing so allows us to minimize any prior assumptions about the topology of the network we are dealing with. This is well-suited to functional connectivity, since brain-network characteristics vary from individual to individual, and are affected by things like age, sex, psychiatric conditioning[8], etc. Additionally, there exist short-term brain 'states', that exhibit different behavior and network dynamics. As such using a predetermined network structure for GSP purposes is likely to degrade the quality of information that can be obtained. The only assumption we can make about the network structure is that it would follow the *Wattz-Strogatz Model*, due to the small-world structure of functional brain networks.

There exists a technique known as *Multidimensional Scaling*[3] that is used to convert graphs to Euclidean spaces. This technique can be expressed as an equation containing the difference matrix (Δ), where each element(Δ_{ij}) is the Euclidean distance between 2 nodes.

$$\Delta = \mathbf{W} + w(\mathbf{1}_n\mathbf{1}_n^T - I_n - \mathbf{W})$$

This problem is invertible under a few conditions, that $\mathbf{W}$ is symmetric and non-negative. As such I use a pairwise metric to generate Δ and then perform *Inverse Multidimensional Scaling* to obtain $\mathbf{W}$, effectively inferring a graph structure from the data.

$$\Delta_{ij} = e^{(\frac{-(1-\rho(x_i,x_j))^2}{2\sigma_1^2})} \times e^{(\frac{-(1-d(v_i,v_j))}{2\sigma_2^2})}$$

Here $\rho(x_i, x_j)$ is a pairwise metric to represent functional connectivity between 2 nodes *i* and *j*. I used one of the following functions - Phase Locking Value, Amplitude Coupling, Granger Causality, and Correlation. The absolute value of these were used to meet the invertibility criterion. $d(v_i, v_j)$ represents the structural connectivity of 2 nodes, and is the physical distance between the barycenters of each parcel/node. σ_1 and σ_2 are scaling variables.

It is important to note that this work **does not** account for the brain state transitions. A Hidden Markov Model to splice the original time-series source signals about their transition points should fix this, but this could not be implemented due to a lack of computing resources. As such, all data was randomly sliced into a fixed number of states (15) randomly, which is sure to affect results and introduce inconsistencies.

Finally, we have obtained both the graph signal $\mathbf{x}$ and the underlying graph structure through $\mathbf{W}$, and we can begin applying GSP techniques.

V. GSP Applications

A. Graph Sparsification

Due to the nature of the steps used for graph structure inference, we end up with a perfectly connected graph with exactly 5151 edges. This is undesirable since a) GSP computations would take significantly longer, and (b) it becomes difficult to inspect the 'meaningful' connections in the graph.

As such I attempted to perform a spectral graph-sparsification of sorts, using the graph Fourier basis. The algorithm used was as follows

- Decompose L: $L = U\Lambda U^T$
- Project Signal onto Graph Fourier Basis: $\bar{x_t} = Ux_t$
- Use the 40 strongest components of the projection to make $\bar{U}$
- Reconstruct L from the top 50 eigenvectors: $\bar{L} = \bar{U}\bar{\Lambda}\bar{U}^T$

However, this algorithm failed in the last step, which is reconstructing $\bar{W}$ from $\bar{L}$. $\bar{L} = \sum_{j=1}^{N} \bar{W}_{ij} - \bar{W}$, ends up being a difficult calculation. As such this approach was dropped in favor of the Spielman-Srivastava Sparsification algorithm[9].

This facilitated a reduction in edges by over an order of magnitude.

B. Signal Diffusion

GSP Operators can be used to study the diffusion of a graph signal over a network[3]. This can be estimated using

$$\mathbf{x}_{\rho+1} = (\mathbf{I} + \alpha\mathbf{L})^{-1}\mathbf{x}_\rho$$

Given that we have all values of x_ρ, we can evaluate the performance of this algorithm for different values of α using the square error over 100 sample iterations. Interestingly, it seems that the error increases linearly with α (as seen in Fig. 3), indicating that the diffusion model does not hold true for neural activity. Another interesting point to note is that this error is the same for $\mathbf{W}$s derived from different metrics. Perhaps alternative operators or better network inference improves results.

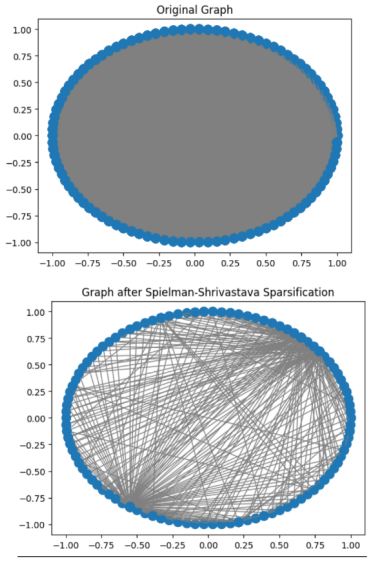

Fig. 2. Graph Visualization before and after sparsification. Note the 2 main clusters in the sparse graph

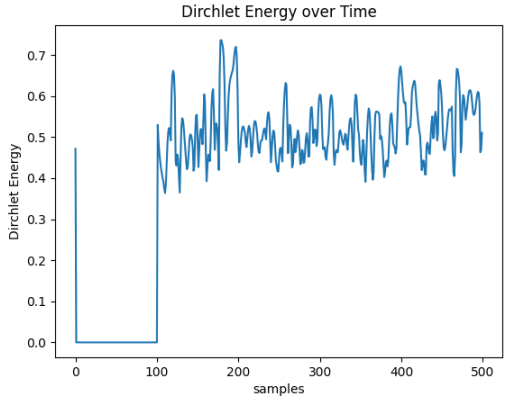

Fig. 4. Dirchlet Energy Over about half a second

Metric	Avg. Edges	Avg. Diffusion Error	Avg. DE
PLV	386.87	101.77	0.245
Correlation	476.47	101.77	0.418

TABLE II
FUNCTIONAL CONNECTIVITY METRICS AND THEIR PERFORMANCE

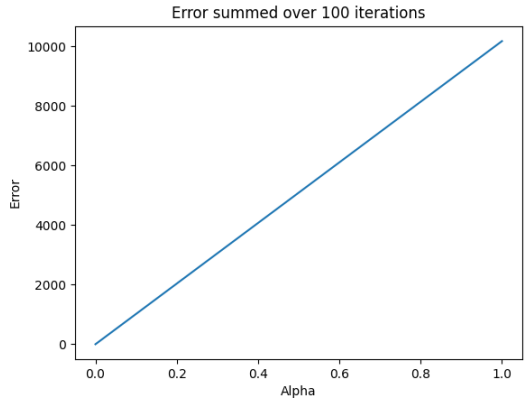

Fig. 3. Diffusion Error over 100 iterations for a sample graph signal

C. Connectivity Dynamics

Past work on the dynamics of functional connectivity networks has relied on sliding window analysis. GSP techniques offer an alternative approach to studying network dynamics, with the aid of Dirichlet Energy (DE)[9].

$$E(G) = \sum_{i,j=1}^{102} W_{ij}(x_i - x_j)^2$$

This measure quantifies the agreement between the graph signal. If it is large and positive, that means that there is a discrepancy between the connectivity and amplitude of signals. Conversely, if the DE is low, then the weights and signal amplitudes are in agreement.

Interestingly, there exists a modular DE (MDE) that also captures local-level behavior in sub-modules. I will not be using MDE due to having to split a graph into sub-graphs, which brings in another instance of the graph structure problem.

As seen in Fig 4, the graph has a very low DE in the first 100 samples, after which it increases manyfold. This is possibly indicative of a short-term state fluctuation around the 100th sample - which would explain the sudden dissonance in connectivity and graph signals. HMM integration might yield more consistent results. Additionally, looking for sharp changes in DE levels might be a lower-cost alternative to using HMMs to approximate brain-state transitions.

D. Performance of Functional Connectivity Metrics

Both (A) Sparsification and (B) Diffusion would change with a change in $\mathbf{W}$. As such, comparing different metrics used to create $\mathbf{W}$ would help shed light on what information is more 'accurate'. Table II shows the average performance of the previous algorithms with graphs constructed using different functional connectivity metrics. These averages are computed using the average performance for 10 EEG recordings from different subjects.

It would appear that PLV outperforms Correlation due to a lower DE, which would imply better coherence with the network structure, but this cannot be verified without correctly implementing the time-varying graph structures as discussed at the end of IV.

VI. CONCLUSION

This paper covers very limited applications of GSP in network neuroscience. Due to the varied number of factors that affect results - a few more advanced studies need to be completed, that can reliably account for factors such as brain state transitions, functional connectivity metrics, and more. The volume of preprocessing steps that needed to be completed in order to start applying GSP techniques also significantly hindered progress.

GSP Applications in this field have been fairly limited in the past and I believe this is a good first step. Further work would investigate the application of Hidden Markov Models to create a time-varying graph structure and compare performance to a DE-change approach mentioned in V(D). There is also interesting work to be done in applying spectral graph sparsification techniques to approximate a tree that represents the network. Additional work can be done using Graph Filters and feature reduction using projection.

ACKNOWLEDGMENTS

I would like to acknowledge the Mobility and Fall Prevention Research Laboratory for providing the Raw EEG data that was used in this study. I would also like to thank the National Center for Supercomputing Applications for providing computing resources. I would also like to acknowledge the maintainers of MNE and PyGSP[5].

REFERENCES

1 Adamovich, T., Zakharov, I., Tabueva, A. et al. The thresholding problem and variability in the EEG graph network parameters. Sci Rep 12, 18659 (2022). https://doi.org/10.1038/s41598-022-22079-2

2 Aviyente, Selin, and M Delgado. “Chapter 31: GSP On Neouronal Networks.” Cooperative and Graph Signal Processing, Elseiver, 2018, pp. 799–816.

3 Sejdic, Ervin and L Stankovic. Vertex-Frequency Analysis of Graph Signals. Springer, 2019.

4 Fischl B. FreeSurfer. Neuroimage. 2012 Aug 15;62(2):774-81. doi: 10.1016/j.neuroimage.2012.01.021. Epub 2012 Jan 10. PMID: 22248573; PMCID: PMC3685476.

5 PYGSP, “Towards Signal Processing on Graphs”, Swiss National Science Foundation

6 Alexander Schaefer, Ru Kong, et Al, Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI, Cerebral Cortex, Volume 28, Issue 9, September 2018, Pages 3095–3114, https://doi.org/10.1093/cercor/bhx179

7 A. Gramfort, M. Luessi, E. Larson, et Al. 2013, MNE software for processing MEG and EEG data.

8 Stevner, A.B.A., Vidaurre, D., Cabral, J. et al. Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep. Nat Commun 10, 1035 (2019). https://doi.org/10.1038/s41467-019-08934-3

9 Spielman, Daniel A., and Nikhil Srivastava. “Graph sparsification by effective resistances.” Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing, 17 May 2008, https://doi.org/10.1145/1374376.1374456.