

ATTRIBUTED GRAPH CLUSTERING FOR FUNCTIONAL CONNECTIVITY NETWORKS

Meiby Ortiz-Bouza and Selin Aviyente

Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI

ABSTRACT

Network science has been successful at characterizing the topology of brain networks, showing alterations to the network structure due to disease and cognitive function. Functional connectivity networks (FCNs) represent different brain regions as the nodes and the dependency between them as the edges of a graph. One of the main tools to characterize the topology of FCNs is community detection. While different community detection methods have been applied to uncover the modular structure of the human brain, they rely only on the connectivity matrix. With the advances in graph signal processing, it is now possible to model the neuroimaging data as graph signals defined on the nodes of the FCN and the functional connectivity network as the underlying graph. In this paper, we present an optimal graph filter design procedure for identifying the community structure in FCNs. The resulting community detection algorithm employs both the connectivity and signal dynamics yielding a more robust community structure. The proposed method is applied to electroencephalogram (EEG) data collected from a study of error monitoring in the human brain.

Index Terms— EEG, functional connectivity, graph signal processing, community detection

1. INTRODUCTION

In the past two decades, the field of complex networks has emerged as a powerful tool to characterize the structure and function of the human brain using a variety of tools from graph theory and network science [1]. These analyses have uncovered the organizational principles of brain networks such as the presence of communities where groups of regions are more strongly connected between each other than with other communities [2, 3]. Network analysis has also been related to behavioral and clinical measures to study development and behavior [4, 5].

While network neuroscience provides an explanation of how connectomes and functional brain activity support behavior, it is also important to analyze and understand the dynamics of functional signals with respect to the network structure. Tools from the field of graph signal processing (GSP) have been recently tailored for this purpose [6]. Concepts of

graph Fourier transform (GFT) and the corresponding notions of graph frequency components and graph filters have been utilized to analyze brain signals. These GSP tools permit the decomposition of a graph signal into pieces that represent different levels of variability and have been used for dimensionality reduction and classification purposes [6, 7, 8, 9].

Inspired by this recent work, in this paper we explore the modularity of brain networks from the perspective of GSP. Modularity is thought to reflect functional and anatomical segregation. Assessment of the brain modular organization may provide a key to understanding the relation between aberrant connectivity and brain disease. Various methods have been deployed to investigate the modular structure of brain networks [10, 11]. Typically, these methods rely on the optimization of a quality function that quantifies the goodness of a network partition against that of an ensemble of randomized networks with similar statistical properties. However, these methods rely only on the adjacency matrix constructed from neuroimaging and do not exploit the time series data associated with each node.

In recent years, several methods have been proposed to label nodes of a graph by combining the node attributes and link information [12, 13]. The most common approach is graph embedding, in particular graph autoencoders [14, 15, 12]. In recent work, we proposed a graph signal processing based approach, GraFiCA [16], where optimal polynomial graph filters are learned for clustering. In the first step, optimal graph partitioning is learned for the given node attributes while in the second step, the best linear graph filter that separates the node attributes from each other is learned for the given partition. Unlike graph autoencoders, the proposed approach learns graph filters that minimize the within-class dissimilarity and maximize the between-class dissimilarity. In this paper, we evaluate the use of GraFiCA in an unsupervised setting for learning the community structure of functional connectivity networks of the brain. In particular, we study the community structure of FCNs following an error response in a Flanker task.

2. BACKGROUND

2.1. Functional Connectivity Networks

In this paper, reduced interference Rihaczek (RID-Rihaczek) time-frequency phase synchrony is used to quantify the func-

This work was supported in part by the NSF under CCF-2006800.

tional connectivity between two brain regions [17]. For a signal $x_i(t)$, the RID-Rihaczek distribution is defined as [17]:

$$C_i(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{(\theta\tau)^2}{\sigma}} e^{j\frac{\theta\tau}{\sigma}} A(\theta, \tau) e^{-j(\theta t + 2\pi f\tau)} d\tau d\theta, \quad (1)$$

where $e^{-\frac{(\theta\tau)^2}{\sigma}}$ is the Choi-Williams kernel [18], $e^{j\frac{\theta\tau}{\sigma}}$ is the kernel function for the Rihaczek distribution [19], and $A(\theta, \tau)$ is the ambiguity function defined as $A(\theta, \tau) = \int_{-\infty}^{\infty} x_i(u + \frac{\tau}{2}) x_i^*(u - \frac{\tau}{2}) e^{j\theta u} du$.

From this complex distribution, one can obtain the instantaneous phase, $\phi_i(t, f)$, and the phase difference between two signals x_i and x_j as $\phi_{i,j}(t, f) = \arg \left[\frac{C_i(t, f) C_j^*(t, f)}{|C_i(t, f)| |C_j(t, f)|} \right]$.

Phase Locking Value (PLV), which quantifies the phase synchrony between x_i and x_j , is defined as the consistency of the phase differences across trials and can be computed as:

$$PLV_{i,j}(t, f) = \frac{1}{R} \left| \sum_{r=1}^R \exp(j\phi_{i,j}^r(t, f)) \right|, \quad (2)$$

where R is the total number of trials and $\phi_{i,j}^r(t, f)$ is the phase difference for the r th trial. Once the pairwise PLV values are computed between all pairs of electrodes, the weighted adjacency matrix corresponding to the FCN can be constructed as the average of $PLV_{i,j}(t, f)$ within the time interval and frequency band of interest.

2.2. Graphs

A graph $\mathcal{G} = (V, E, \mathbf{A})$ is defined by a node set V with $|V| = N$, an edge set E and the adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$. The graph Laplacian is given by $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the diagonal degree matrix defined as $D_{ii} = \sum_j A_{ij}$. The normalized Laplacian matrix $\mathbf{L}_n$ is defined as $\mathbf{L}_n = \mathbf{D}^{-1/2}(\mathbf{D} - \mathbf{A})\mathbf{D}^{-1/2} = \mathbf{I}_N - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} = \mathbf{I}_N - \mathbf{A}_n$, where $\mathbf{I}_N$ is the identity matrix of size N and $\mathbf{A}_n$ is the normalized adjacency matrix. The spectrum of $\mathbf{L}_n$ is composed of the diagonal matrix of the eigenvalues, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$ with $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$, and the eigenvector matrix $\mathbf{U} = [u_1|u_2|\dots|u_N]$ such that $\mathbf{L}_n = \mathbf{U}\Lambda\mathbf{U}^\top$ [20].

2.3. Graph Filtering

Signals defined on the nodes of an attributed graph can be represented as a matrix $\mathbf{F} \in \mathbb{R}^{N \times p}$, where p is the number of attributes for each node. A linear graph filter is described as the linear operator

$$\mathcal{H}(\mathbf{L}) = \sum_{t=0}^{T-1} h_t \mathbf{L}^t = \mathbf{U} \left(\sum_{t=0}^{T-1} h_t \Lambda^t \right) \mathbf{U}^\top,$$

where T is the filter order and h_t 's are the coefficients. A filtered graph signal $\tilde{\mathbf{F}}$ is obtained by filtering $\mathbf{F}$ with $\mathcal{H}(\mathbf{L})$, $\tilde{\mathbf{F}} = \mathcal{H}(\mathbf{L})\mathbf{F} = \mathbf{U}\mathcal{H}(\Lambda)\mathbf{U}^\top\mathbf{F}$, where $\mathcal{H}(\Lambda) = \text{diag}(\mathcal{H}(\lambda_1), \dots, \mathcal{H}(\lambda_N))$.

3. ATTRIBUTED GRAPH CLUSTERING (GraFiCA)

3.1. Problem Formulation

Given a network with K communities, $V_1, V_2, \dots, V_K$, normalized adjacency matrix corresponding to the FCN, $\mathbf{A}_n = \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$, and the graph signal matrix $\mathbf{F} \in \mathbb{R}^{N \times p}$, corresponding to the neurophysiological signals corresponding to each node, we propose to learn the node labels and design polynomial graph filters such that the within-community association is minimized while the between-community distance of the filtered attributes is maximized. The proposed cost function is formulated as

$$\sum_{c=1}^K \frac{1}{\text{vol}(V_c)} \sum_{i,j \in V_c} \|\tilde{F}_i - \tilde{F}_j\|^2 - \gamma \sum_{c=1}^K \frac{1}{\text{vol}(V_c)} \sum_{\substack{i \in V_c \\ j \notin V_c}} \|\tilde{F}_i - \tilde{F}_j\|^2, \quad (3)$$

where the first and second terms quantify the similarity of the filtered node attributes within clusters and between clusters, respectively. $\text{vol}(V_c)$ is the total degree of the nodes in cluster c based on the dissimilarity matrix $\tilde{\mathbf{W}}$ with $\tilde{W}_{ij} = \|\tilde{F}_i - \tilde{F}_j\|^2$. The goal is to minimize the cost function in (3) in terms of both the graph partition, $\{V_1, V_2, \dots, V_K\}$, and the graph filter coefficients h_t 's. An alternating minimization approach is proposed to solve this joint optimization problem.

3.1.1. Clustering

In the first step, given $\tilde{\mathbf{F}}$, we minimize (3) with respect to the graph partition, $\{V_1, V_2, \dots, V_K\}$. It can be shown that the two terms in (3) correspond to the normalized association and normalized cut, respectively. Using the relationship between normalized association and normalized cut, these two terms can be combined and written as $\text{tr}(\tilde{\mathbf{Z}}^\top \mathbf{D}^{-1/2} \tilde{\mathbf{W}} \mathbf{D}^{-1/2} \tilde{\mathbf{Z}})$, where $\tilde{\mathbf{Z}} = \mathbf{D}^{1/2} \tilde{\mathbf{Z}}$ and $\tilde{\mathbf{Z}}^\top \tilde{\mathbf{Z}} = \mathbf{I}$, and $\tilde{\mathbf{Z}}$ is the cluster indicator matrix defined in [16]. For the clustering task, we add a regularization term that quantifies the similarity between the normalized adjacency matrix and the cluster assignment matrix, $\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^\top$. Thus, the problem formulation becomes

$$\underset{\tilde{\mathbf{Z}}, \tilde{\mathbf{Z}}^\top \tilde{\mathbf{Z}} = \mathbf{I}}{\text{minimize}} \text{tr}(\tilde{\mathbf{Z}}^\top \mathbf{D}^{-1/2} \tilde{\mathbf{W}} \mathbf{D}^{-1/2} \tilde{\mathbf{Z}}) + \alpha \|\mathbf{A} - \tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^\top\|_F^2. \quad (4)$$

Rewriting the last term in terms of the trace function, and using $\tilde{\mathbf{Z}}^\top \tilde{\mathbf{Z}} = \mathbf{I}$, we obtain

$$\underset{\tilde{\mathbf{Z}}, \tilde{\mathbf{Z}}^\top \tilde{\mathbf{Z}} = \mathbf{I}}{\text{minimize}} \text{tr}(\tilde{\mathbf{Z}}^\top (\mathbf{D}^{-1/2} \tilde{\mathbf{W}} \mathbf{D}^{-1/2} - 2\alpha \mathbf{A}_n) \tilde{\mathbf{Z}}), \quad (5)$$

which can be solved using the eigenvectors corresponding to the K smallest eigenvalues of $\tilde{\mathbf{W}}_n - 2\alpha \mathbf{A}_n$.

3.1.2. Optimal Filter Design

Once the cluster assignments are obtained, we want to determine the coefficients of the optimal polynomial filter $\mathcal{H}(\Lambda) =$

$\sum_{t=0}^{T-1} h_t \Lambda^t$ for the clustering task for a given filter order T . Using $\tilde{\mathbf{F}} = \mathbf{U} \sum_{t=0}^{T-1} h_t \Lambda^t \mathbf{U}^\top \mathbf{F}$, the cost function in (3) can be rewritten in terms of the filter coefficients h_t 's as follows:

$$\underset{h, h^\top h = \mathbf{I}}{\text{minimize}} \text{tr}(\mathbf{F}^\top \mathbf{U} \sum_{t=0}^{T-1} h_t \Lambda^t \mathbf{U}^\top (\mathbf{B} - \gamma \mathbf{C}) \mathbf{U} \sum_{t=0}^{T-1} h_t \Lambda^t \mathbf{U}^\top \mathbf{F}), \quad (6)$$

where $\mathbf{B}$ and $\mathbf{C}$ are K block diagonal matrices defined as

$$B_{ij} = \begin{cases} \frac{|V_c| - 1}{\text{vol}(V_c)}, & i, j \in V_c, i = j \\ \frac{-1}{\text{vol}(V_c)}, & i, j \in V_c, i \neq j \\ 0, & i \in V_c, j \notin V_c \end{cases}$$

$$C_{ij} = \begin{cases} \frac{N}{\text{vol}(V_c)} - \sum_{k=1}^K \frac{|V_k|}{\text{vol}(V_k)} - \frac{2|V_c|}{\text{vol}(V_c)}, & i, j \in V_c, i = j \\ 0, & i, j \in V_c, i \neq j \\ -\frac{1}{\text{vol}(V_c)} - \frac{1}{\text{vol}(V_b)}, & i \in V_c, j \in V_b \end{cases}$$

The problem in (6) can be solved using the the eigenvectors of matrix $\mathbf{S} \in \mathbb{R}^{T \times T}$ defined as $S_{ij} = \text{tr}(\mathbf{F}^\top \mathbf{U} \Lambda^i \mathbf{U}^\top (\mathbf{B} - \gamma \mathbf{C}) \mathbf{U} \Lambda^j \mathbf{U}^\top \mathbf{F})$, as derived in [16]. Once we obtain $h = [h_0 \ h_1 \ \dots \ h_{T-1}]^\top$, we update $\tilde{\mathbf{F}} = \mathbf{U} \sum_{t=0}^{T-1} h_t \Lambda^t \mathbf{U}^\top \mathbf{F}$ and the cluster assignments based on the filtered attributes using $\tilde{\mathbf{W}}$. Since there are T possible solutions to this system, we choose the h that gives the cluster assignment with the highest evaluation metric. Since we do not have ground truth for the community structure in FCNs, we use modularity as an evaluation metric. This process is repeated until convergence.

3.2. Determining the Number of Clusters

In the selection of the best filter among all eigenvectors of $\mathbf{S}$, GraFiCA presented above assumes that the number of clusters and the ground truth community structure are known *a priori*. In the case of FCNs, this is not true. For this reason, in the proposed application we first set the number of communities to K and then determine the optimal filter parameters, h , as the one that maximizes the modularity of the learned partition, where modularity, Q , is defined as:

$$Q = \sum_{i=1}^N \sum_{j=1}^N (A_{ij} - \gamma P_{ij}) \delta_{g_i, g_j}, \quad (7)$$

where P_{ij} is the expected edge weight between nodes i and j under the null model, g_i is the community of node i , and $\delta_{g_i, g_j} = 1$ if $g_i = g_j$ and 0 otherwise.

3.3. Consensus Community Structure

Once the community structure of the FCNs for a group of subjects are detected, it is often desirable to find a group community structure, which summarizes the shared communities across subjects. In this paper, we propose a group community structure detection method based on multiview graphs. Given L subjects, for each subject we obtain the optimal filter based

on maximizing modularity. The resulting community labels are used to construct the binary community structure matrix, $\mathbf{Z} \in \mathbb{R}^{N \times N}$, where $Z_{ij} = 1$ if nodes i and j are in the same community. This process results in L community structure matrices, which can be modeled as the layers of a multiview graph, where each layer is an undirected, binary graph corresponding to each subject. The consensus community structure can be found from this multiview graph using Spectral Clustering on Multi-Layer graphs (SC-ML) [21] where spectral clustering is applied to a modified Laplacian defined as:

$$\mathbf{L}_{mod} = \sum_{l=1}^L \mathbf{L}^l - \alpha \sum_{l=1}^L \mathbf{U}^l \mathbf{U}^{l\top}, \quad (8)$$

where $\mathbf{L}^l$ is the the normalized graph Laplacian and $\mathbf{U}^l$ is the low-rank subspace embedding of layer l . In this work, we set $\alpha = 0.5$, following the guidelines in [21].

4. RESULTS

4.1. EEG Data

EEG data is collected from a cognitive control-related error processing study where participants were performing a letter version of the speeded reaction Flanker task [22]. Each participant was presented with a string of five letters at each trial. Letters could be congruent (e.g., SSSSS) or incongruent stimuli (e.g., SSTSS) and the participants were instructed to respond to the center letter with a standard mouse. Each trial began with 35ms of flanking stimuli (e.g. SS SS), and followed by the target stimuli (e.g., SSSSS/SSTSS), which were presented for 100 ms (total presentation time is 135ms). Trials were followed by an inter-trial random interval ranging from 1200 to 1700 ms. The trials were conducted to study the Error-Related Negativity (ERN) after an error response, where each trial was one second long. Total number of trials was 480 in which the total number of error trials in different participants varied from 20 to 61.

The EEG was recorded using The ActiveTwo system (BioSemi, Amsterdam, The Netherlands). The international 10/20 system is followed for placement of 64 Ag-AgCl electrodes. The sampling frequency of the data was 512 Hz. After removal of the trials with artifacts, the Current Source Density (CSD) Toolbox [23] was employed to minimize the volume conduction.

As previous studies indicate increased synchronization associated with the ERN in the theta frequency band (4–7 Hz) and 25–75 ms time window [22, 24, 25, 26], all analysis was performed for this time and frequency range. The average phase synchrony corresponding to theta band and 25-75 ms time window were recorded into separate 64×64 connectivity matrices, one for each subject. In this paper, we consider data from 20 participants.

4.2. Community Detection

Using the FCNs corresponding to the theta frequency band and ERN time window, we define the graph signals $\mathbf{F} \in \mathbb{R}^{64 \times 512}$ where for each node we take a 1 second recording including the ERN time window obtained by averaging across trials. For each subject, we learn the optimal filter for $T = 3$ and the corresponding community structure with different values of K ranging from 4 to 10.

In Fig. 1, the average variation of the community structure across 20 subjects is shown with respect to the number of communities for our method and for spectral clustering. From this figure, it can be seen that for GraFiCA $K = 8$ yields the most stable and consistent community structure across subjects. Furthermore, it can be seen that the community structure is more consistent across subjects for GraFiCA compared to regular spectral clustering implying that using both connectivity and signal dynamics for community detection results in more robust communities.

For this reason, we evaluate the community structure for $K = 8$. From Fig. 2, a community corresponding to the lateral prefrontal cortex (IPFC) regions including the right-frontal electrodes F2, F4, F6 can be seen. In addition to this community, there is a community comprised of the central-parietal regions corresponding to medial PFC (mPFC). Another community centers on medial-frontal sites including FCz, consistent with the time-domain ERN component topography. Activity in parietal-occipital regions were characterized with three different clusters. These results are in line with prior work demonstrating increased connectivity between electrodes over mPFC and electrodes over IPFC following the commission of an error [26].

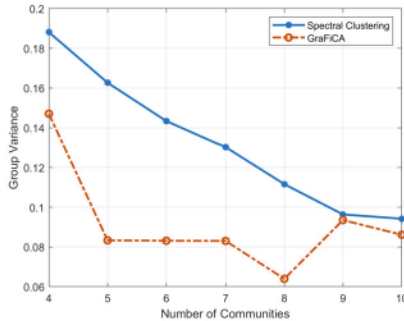


Fig. 1: Comparison of group variance for spectral clustering and GraFiCA.

5. CONCLUSIONS

In this paper, we introduced a GSP based approach to studying the modular structure of FCNs of the brain. In particular, we employed a recently introduced graph filter learning method, GraFiCA, for simultaneous graph filtering and node labeling. The proposed method learns the optimal poly-

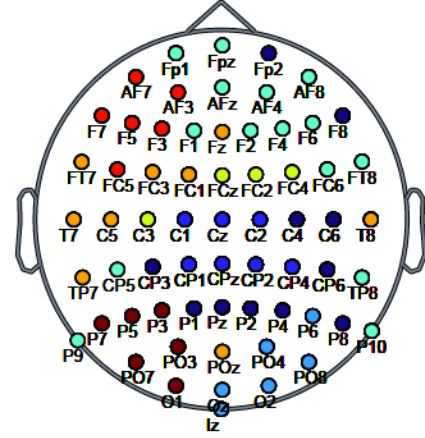


Fig. 2: Consensus community structure for $K = 8$.

nomial filter for discriminating between communities and is evaluated on FCNs constructed from EEG data. The application of the framework to a study of error-related negativity revealed communities aligned with lateral and medial PFC regions known to play a key role in cognitive control. Future work will consider the extension of this framework by considering different filter orders, T , and learning multiple filters to obtain the representation of the graph signals across multiple frequency scales.

6. COMPLIANCE WITH ETHICAL STANDARDS

The study was designed following the experimental protocol and guidelines approved by the Institutional Review Board (IRB) of the Michigan State University (IRB: LEGACY13-144). The data acquisition was performed following the guidelines and regulations established by this protocol. Written and informed consent was collected from each participant before data collection.

7. REFERENCES

- [1] Ed Bullmore and Olaf Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature reviews neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [2] Olaf Sporns and Richard F Betzel, "Modular brain networks," *Annual review of psychology*, vol. 67, pp. 613, 2016.
- [3] Martijn P Van Den Heuvel and Olaf Sporns, "Rich-club organization of the human connectome," *Journal of Neuroscience*, vol. 31, no. 44, pp. 15775–15786, 2011.
- [4] Jonas Richiardi, Sophie Achard, Horst Bunke, and Dimitri Van De Ville, "Machine learning with brain graphs:

predictive modeling approaches for functional imaging in systems neuroscience,” *IEEE Signal processing magazine*, vol. 30, no. 3, pp. 58–70, 2013.

- [5] Danielle S Bassett and Marcelo G Mattar, “A network neuroscience of human learning: potential to inform quantitative theories of brain and behavior,” *Trends in cognitive sciences*, vol. 21, no. 4, pp. 250–264, 2017.
- [6] Weiyu Huang, Thomas AW Bolton, John D Medaglia, Danielle S Bassett, Alejandro Ribeiro, and Dimitri Van De Ville, “A graph signal processing perspective on functional brain imaging,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 868–885, 2018.
- [7] Saurabh Sihag, Sebastien Naze, Foad Taghdiri, Charles Tator, Richard Wennberg, David Mikulis, Robin Green, Brenda Colella, Maria Carmela Tartaglia, and James R Kozloski, “Multimodal dynamic brain connectivity analysis based on graph signal processing for former athletes with history of multiple concussions,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 284–299, 2020.
- [8] Mathilde Ménoret, Nicolas Farrugia, Bastien Padeloup, and Vincent Gripon, “Evaluating graph signal processing for neuroimaging through classification and dimensionality reduction,” in *2017 IEEE Global Conference on Signal and Information Processing (Global-SIP)*. IEEE, 2017, pp. 618–622.
- [9] Liu Rui, Hossein Nejati, and Ngai-Man Cheung, “Dimensionality reduction of brain imaging data using graph signal processing,” in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 1329–1333.
- [10] David Meunier, Renaud Lambiotte, and Edward T Bullmore, “Modular and hierarchically modular organization of brain networks,” *Frontiers in neuroscience*, vol. 4, pp. 200, 2010.
- [11] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al., “Functional network organization of the human brain,” *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.
- [12] Xiaotong Zhang, Han Liu, Qimai Li, and Xiao-Ming Wu, “Attributed graph clustering via adaptive graph convolution,” in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 2019, pp. 4327–4333.
- [13] Jaewon Yang, Julian McAuley, and Jure Leskovec, “Community detection in networks with node attributes,” in *2013 IEEE 13th international conference on data mining*. IEEE, 2013, pp. 1151–1156.
- [14] Thomas N Kipf and Max Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.
- [15] Thomas N Kipf and Max Welling, “Variational graph auto-encoders,” *NIPS Workshop on Bayesian Deep Learning*, 2016.
- [16] Meiby Ortiz-Bouza and Selin Aviyente, “Optimal graph filters for clustering attributed graphs,” *arXiv preprint arXiv:2211.04634*, 2022.
- [17] Selin Aviyente and Ali Yener Mutlu, “A time-frequency-based approach to phase and phase synchrony estimation,” *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3086–3098, 2011.
- [18] Cohen Leon, “Time-frequency analysis: theory and applications,” *USA: Pnentice Hall*, 1995.
- [19] A Rihaczek, “Signal energy distribution in time and frequency,” *IEEE Transactions on information Theory*, vol. 14, no. 3, pp. 369–374, 1968.
- [20] FRK Chung, “Spectral graph theory, providence, ri: Amer,” *Math. Soc*, 1997.
- [21] Xiaowen Dong, Pascal Frossard, Pierre Vanderghenst, and Nikolai Nefedov, “Clustering on multi-layer graphs via subspace analysis on grassmann manifolds,” *IEEE Transactions on signal processing*, vol. 62, no. 4, pp. 905–918, 2013.
- [22] Jason R Hall, Edward M Bernat, and Christopher J Patrick, “Externalizing psychopathology and the error-related negativity,” *Psychological science*, vol. 18, no. 4, pp. 326–333, 2007.
- [23] Craig E Tenke and Jürgen Kayser, “Generator localization by current source density (csd): implications of volume conduction and field closure at intracranial and scalp resolutions,” *Clinical neurophysiology*, vol. 123, no. 12, pp. 2328–2345, 2012.
- [24] Logan T Trujillo and John JB Allen, “Theta eeg dynamics of the error-related negativity,” *Clinical Neurophysiology*, vol. 118, no. 3, pp. 645–668, 2007.
- [25] James F Cavanagh and Michael J Frank, “Frontal theta as a mechanism for cognitive control,” *Trends in cognitive sciences*, vol. 18, no. 8, pp. 414–421, 2014.
- [26] Marcos Bolanos, Edward M Bernat, Bin He, and Selin Aviyente, “A weighted small world network measure for assessing functional connectivity,” *Journal of neuroscience methods*, vol. 212, no. 1, pp. 133–142, 2013.