Gaussian Process-based Inference toward Revealing Brain Functional Connectivity

Chen Cui
Dept. of ECE
Stony Brook University
Stony Brook, USA

Sima Mofakham
Dept. of Neurosurgery
Stony Brook University
Stony Brook, USA

Jessica M. Phillips
Wisconsin National Primate Research Center
University of Wisconsin-Madison
Madison, USA

Kurt Butler
Dept. of ECE
Stony Brook University
Stony Brook, USA

Charles B. Mikell
Dept. of Neurosurgery
Stony Brook University
Stony Brook, USA

Yuri B. Saalmann
Dept. of Psychology
University of Wisconsin-Madison
Madison, USA

Petar M. Djurić
Dept. of ECE
Stony Brook University
Stony Brook, USA

***Abstract*—In the field of neuroscience, the task of accurately deciphering brain connectivity from observed data has continued to receive increased attention. In this paper, we address the challenge of inferring candidates for brain functional connectivity using local field potential data, taking into account nonlinear interactions and multiple delays. Our approach leverages Gaussian processes and automatic relevance determination kernels to learn mapping functions from one brain area to another. The resulting learned topology is represented as a directed graph with an adjacency matrix. We validate the approach on both synthetic computational neural datasets and real macaque datasets. The results demonstrate the capability of the method to successfully reveal synthetic graph structures and uncover biologically meaningful pathways in real-world data.**

***Index Terms*—Topology inference, Gaussian processes, brain functional connectivity**

I. Introduction

Undoubtedly, the brain is one of the most complex systems we know of. It is a vast network of 100 billion of neurons connected with more than 100 trillion of connections. One of the biggest mysteries of our time is where and how different cognitive functions, such as working memory, are coded in the brain. It is critical to address this gap to improve our understanding of how the brain works and to develop future therapeutic approaches for neurological disorders. In this paper, we use signal processing tools to illuminate the functional connectivity of the thalamocortical network involved in working memory from recorded local field potentials (LFPs). The results of this work could be used to answer other neuroscience and neurosurgery questions.

Several existing methodological approaches to learning neural connectivity use either covariance or graph Laplacian matrices [6], [20]. However, these methods encounter limitations in effectively capturing unidirectional relationships in the system that generated the data, and we note that these directed relationships are crucial to understanding how information is processed in the brain. Our primary focus is to extract directional information from LFP data about the relationships between brain regions of interest [7]. The LFP data have been recorded in the cerebral cortex and thalamus regions [15].

The propagation of information in the brain is not instantaneous; therefore, models of brain connectivity must account for propagation delays. This raises questions about the applicability of structural equation models (SEMs) [13], which are designed to accommodate simulated influences but may not be suitable for revealing nuanced functional topologies of the brain. Another major approach to recovering a brain topology is based on vector autoregressions (VARs) [9], [18], including structural VARs [2]. These models aim to incorporate delayed dependencies, making them more suitable tools for topology recovery. Furthermore, to capture nonlinearities, which aligns better with reality, the concept of kernel-based Granger causality has been introduced [16], [21].

Some previous studies employ the Gaussian process (GP) in the analysis of brain data. For instance, [19] modeled the LFP as a mixture of GPs to analyze the brain states. In [1], the authors used GPs to model the latent cognitive states underlying observed neural and behavioral data. Additionally, [11] learned the undirected covariance matrix, representing brain functional connectivity. This was achieved by modeling the log-covariance elements as a linear combination of GP latent factors.

In this paper, we address the problem of learning possible functional connectivities between brain regions using GPs [3]–[5], [12]. GPs offer a straightforward means to incorporate prior knowledge into the inference process, simplifying the analysis and interpretations for practitioners [10]. Our method effectively captures nonlinearities and multiple delayed influencing signals while making very mild assumptions about the dependency functions. We use the length scale of the automatic relevance determination (ARD) kernel to indicate the contribution weight of different brain areas to a specific region. The length scales from the ARD kernel have proven successful in feature selection for machine learning problems [22].

We worked on one data set generated by a computational

This work was supported by NSF under Award 2021002 and 2212506.

neural model and on a macaque memory task dataset. For the former, we had the ground truth of the network topology, which we used to evaluate the performance of the proposed method. For the second data set, we had a hypothesis of the topology of the brain network during the studied memory tasks and compared it with the topology obtained using the proposed method.

Our contributions to the paper are as follows: (a) we propose a GP-based method for estimating a topology of a brain network using recorded LFPs. The method measures the strengths of influence on a node by delayed signals from other nodes in the brain and its own delayed signals using the principle of ARD; (b) we demonstrate, with both synthesized and real data, that the proposed method is effective in discovering network topologies.

II. Problem Formulation

A set of LFP signals, denoted as $\{y_{n,t}\}_{n=1:N}^{t=1:T}$, is collected from N electrodes over a time span of T samples, and where n represents the nth electrode. The connectivity between brain areas is represented using a directed graph. This representation uses a binary adjacency matrix, denoted as $\mathbf{A}$, where $a_{i,j} \in \{0,1\}$ is the (i,j)th element corresponding to the presence or absence of a connection between node j and i. A value of 1 for $a_{i,j}$ signifies an edge pointing from node j to node i, while 0 implies no connection between the two nodes. In the context of the brain, a value of 1 for $a_{i,j}$ indicates a functional connection between the areas represented by nodes j and i. We can also define a weighted adjacency matrix, where the entries $w_{i,j} \in [0,1]$ quantify the strengths of these edges.

It is important to note that these connections are subject to a specific delay. Our premise is based on the assumption that brain signals in one area are influenced by signals from other areas with some delays. Mathematically, we model these influences by individual nonlinear functions for each node, which are independent and described as follows:

$$y_{n,t} = f_n(\mathbf{x}_t, \mathbf{w}_n^{[1:\Lambda]}) + \epsilon_n, \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^{N\Lambda} := [y_{1:N,t-1}, \ldots, y_{1:N,t-\Lambda}]^\top$ is a vector composed of past observations from all the nodes, with Λ being the maximum delay. The term ϵ_n represents white Gaussian noise with zero mean and a variance of $\sigma^2_{\epsilon_n}$, f_n is a function associated with node n, and $\mathbf{w}_n^{[1:\Lambda]} := [w_{n,1:N}^{[1]}, \ldots, w_{n,1:N}^{[\Lambda]}]^\top$ are weights associated with each element of the input vector. This formulation encapsulates the interplay between brain regions, their delayed interactions, and the observed signals.

When considering the graph signals of node n across the entire timeline, we can represent them in a vector form as follows:

$$\mathbf{y}_n = \mathbf{f}_n(\mathbf{X}, \mathbf{w}_n^{[1:\Lambda]}) + \boldsymbol{\epsilon}_n, \tag{2}$$

where $\mathbf{y}_n = [y_{n,\Lambda+1}, \ldots, y_{n,T}]^\top$, $\mathbf{X}$ aggregates the input vectors $\mathbf{x}_t$ (the input matrix to all nodes is the same), $t \in \{\Lambda+1, \Lambda+2, \ldots, T\}$, and $\mathbf{f}_n$ is an unknown function. Rather than making deterministic assumptions about linearity or nonlinearity, we propose that the function is drawn from a GP with a zero mean and some kernel κ. This means that the samples of the function come from a zero mean multivariate Gaussian with a covariance matrix $\mathbf{K}_n \in \mathbb{R}^{(T-\Lambda)\times(T-\Lambda)}$, i.e.,

$$\mathbf{f}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_n), \tag{3}$$

where the (i,j)th element $k_{n,i,j}$ of the covariance matrix represents correlation between the ith and jth elements of the output time series and is computed by the kernel function κ. The assumption regarding the function within the framework of GPs entails that the function is smooth, which is a much more moderate assumption compared to a deterministic one.

III. Proposed Solution

We adopt the automatic relevance determination (ARD) kernel for (2) [17]. The kernel is expressed as follows:

$$k_{n,i,j} = \sigma_n^2 \exp\left(-\frac{1}{2}\sum_{m=1}^{N}\sum_{\lambda=1}^{\Lambda} w_{n,m}^{[\lambda]}(y_{m,i-\lambda} - y_{m,j-\lambda})^2\right), \tag{4}$$

where $w_{n,m}^{[\lambda]}$ is the length scale of the kernel. We observe that (4) is employed for feature selection in the field of machine learning [8]. In a more intuitive sense, when $w_{n,m}^{[\lambda]}$ takes on a smaller value, it implies that changes in the mth node within the input graph signal $\mathbf{x}$ have limited influence on the output of node n, for a specific delay λ. For instance, as $w_{n,m}^{[\lambda]}$ approaches zero, the term $y_{m,i-\lambda} - y_{m,j-\lambda}$ nullifies when calculating the kernel. Conversely, a larger value of $w_{n,m}^{[\lambda]}$ indicates that changes in the mth node do affect the output value of the nth node.

Based on this framework, we propose leveraging $w_{n,m}^{[\lambda]}$ to reveal the interaction occurring within the brain from the observed LFP data when a subject engages in a memory task. Details of optimizing the objective function and obtaining the optimal set of hyperparameters $\mathbf{w}_n^{[1:\Lambda]}$, as well as determining the optimal topology based on $\mathbf{w}_n^{[1:\Lambda]}$, can be found in [3].

IV. Numerical Results

We validate the approach on two datasets. The first dataset is a synthetic one that simulates neural activity during a working memory task requiring the sequential application of two task rules of which the first is an abstract rule that instructs whether the orientation or shape of an upcoming visual stimulus is relevant, and the second is a concrete rule that instructs action based on whether orientation or shape is relevant. The second dataset represents real macaque data recorded during the same working memory task. We evaluated the method's performance using the F-score, calculated as follows:

$$\text{F-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{5}$$

where Precision is defined as $\text{TP}/(\text{TP} + \text{FP})$ and Recall, as $\text{TP}/(\text{TP} + \text{FN})$, with TP, FP, and FN denoting true positive, false positive and true negative, respectively.

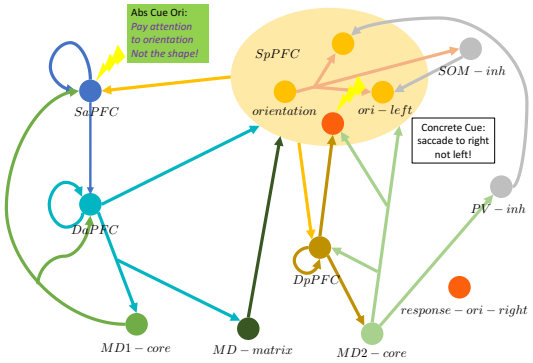

Fig. 1. The network topology of the generative model for the synthetic data.

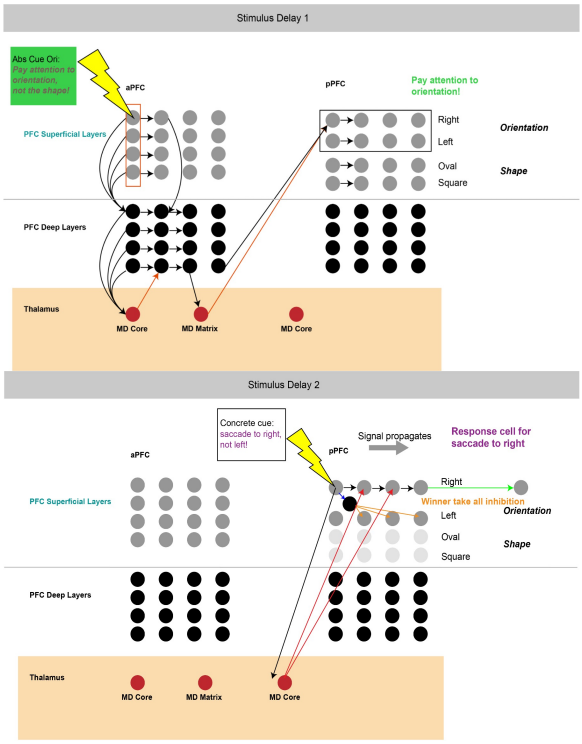

Fig. 2. The flow of neuron engagement. The upper panel shows how the neurons engage in the aPFC area in epoch 1, and the lower one how they do it in the pPFC area in epoch 2.

A. A Synthetic Data Set

The synthetic data set was generated by a computational model that replicates neural interactions that take place when a subject engages in the working memory task. Initially, the focus is on processing orientation cues rather than shapes. In the second epoch, a specific orientation cue is processed. These two task epochs correspond to a shift from left to right in Fig. 1. The flow of information within the model follows the topology illustrated in Fig. 1. The dataset consists of six distinct brain areas, and each epoch is constructed using a total of 500 samples. The anterior prefrontal cortex is delineated into the superficial layer (*SaPFC*) and the deep layer (*DaPFC*); similarly, the posterior prefrontal cortex comprises

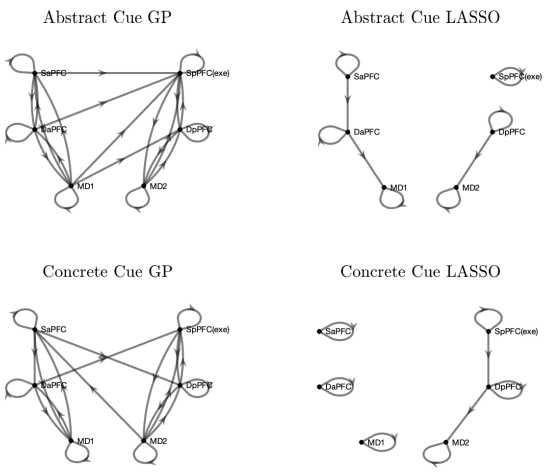

Fig. 3. The graph topologies learned by GP and LASSO during different time epochs, from the synthetic data set.

	abs-GP	abs-LASSO	con-GP	con-LASSO
P	**0.938**	0.438	**0.875**	0.375
R	0.682	**0.778**	0.700	**0.750**
F	**0.790**	0.560	**0.778**	0.500

TABLE I
COMPARISONS OF GP WITH LASSO

the superficial layer (*SpPFC*) and the deep layer (*DpPFC*). Within the thalamus, three distinct areas are identified: *MD1*, *MD-matrix*, and *MD2*.

In generating the synthesized data, we used the Leaky integrate-and-fire model as a spiking neuron model [14]. The neurons' activity is illustrated in Fig. 2. For each area, we had a total of 50 neurons, and we used the average of neurons voltage as our data.

To evaluate the method's performance, we generated histograms illustrating the learned values of $\log(w_{n,m}^{[\lambda]})$ across 100 trials. Further, we calculated the mean of $\log(w_{n,m}^{[\lambda]})$, as shown in Fig. 4, with Λ being three. We computed the F-score values for both the GP and LASSO techniques in both epochs. The results are presented in Table I, where P stands for Precision, R for Recall, F for F-score, and abs-GP and con-GP refer to the performance of the GP for the abstract and concrete epochs, respectively (with analogous designation for the LASSO method). The results show that in Precision and F-score, the GP-based method outperformed LASSO significantly, whereas in Recall its performance was somewhat inferior. The inferred effective topologies are displayed in Fig. 3.

Our observations can be summarized as follows: 1) The pathways *SaPFC* $\rightarrow$ *DaPFC* $\rightarrow$ *MD1* and *SpPFC* $\rightarrow$ *DpPFC* $\rightarrow$ *MD2* were present in both epochs. 2) Self-loops were detected in *DaPFC*, *SaPFC*, *DpPFC* and *SpPFC*. We observed that the *SaPFC* self-loop was stronger during the abstract cue epoch, while the *SpPFC* self-loop was more pronounced during the concrete cue epoch. This follows the pattern

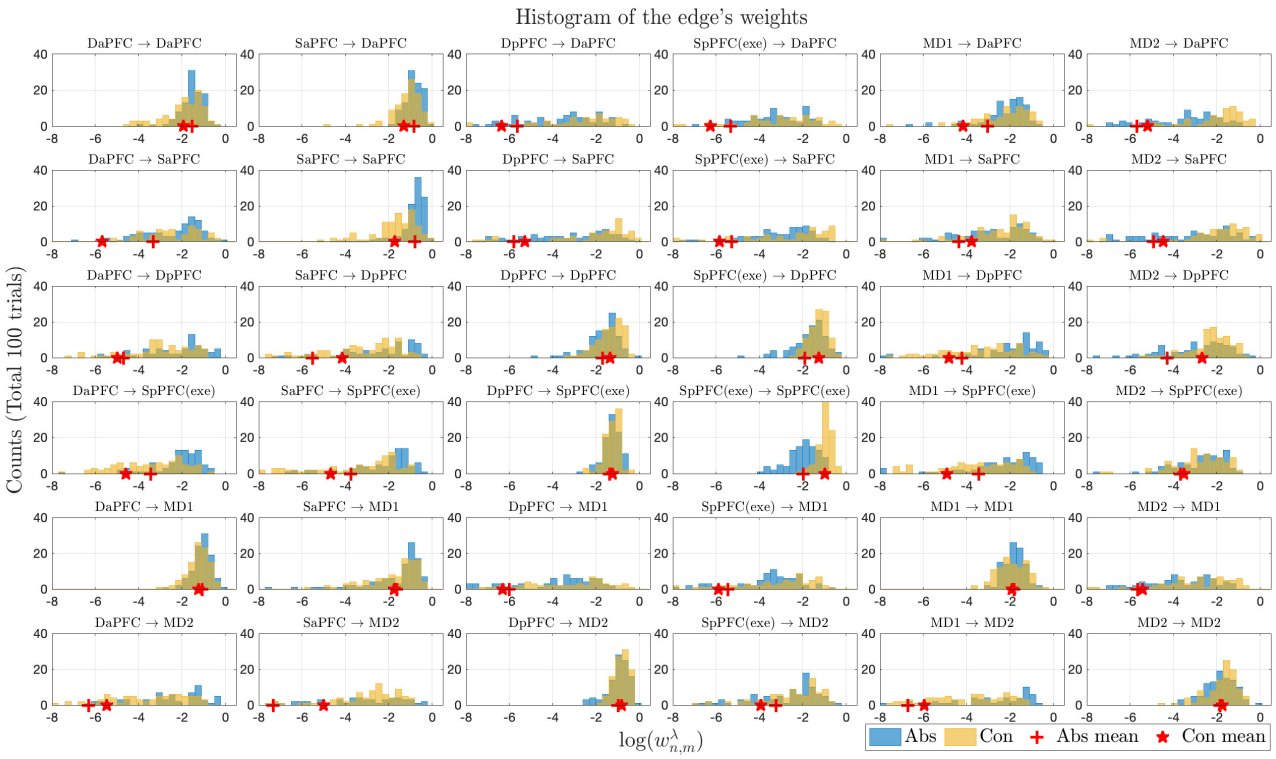

Fig. 4. Histogram of the weights of the edges from 100 trials. The blue color represents the weights of connections in abstract cues. The yellow represents the weights of connection in concrete cue. Each subplot shows the histogram of the weight of each edge. The rows represent the effect, and the columns represent the cause. If the histogram concentrates near 0, this means that the weights of the edge are large.

of stimulus addition, with the stimulus being introduced to *SaPFC* during the abstract cue and to *SpPFC* during the concrete cue epoch. 3) Weak modulation effects were observed from *MD1* $\rightarrow$ *DaPFC, SaPFC* and from *MD2* $\rightarrow$ *DpPFC, SpPFC*. It is important to note that our method is designed for static topology and may face challenges in capturing dynamic, time-varying modulation effects. 4) We identified a pathway from *DpPFC* to *SpPFC*. According to the generative model, there exists a pathway from *DpPFC* to *ori-right*. The *SpPFC* includes cells related to *ori-left*, *ori-right* and *shape*.

B. Macaque Data Set

The data comprised 64 channels from two prefrontal cortical regions (anterior and posterior area 46), which we subdivided into eight subregions (D/S, deep/superficial; d/v, dorsal/ventral): $DaPFC(46d)$, $DaPFC(46v)$, $SaPFC(46d)$, $SaPFC(46v)$, $DpPFC(46d)$, $DpPFC(46v)$, $SaPFC(46d)$ and $SaPFC(46v)$. Our analysis involved two epochs of data from the working memory task: one when the subject was presented the abstract rule cue and another when the subject was shown the concrete rule cue. The abstract cue epoch had 450 samples, and the concrete cue epoch, 480 samples. We hypothesized the following relationships based on strong anatomical connectivity: $SpPFC \rightarrow SaPFC$, $SaPFC \rightarrow DaPFC$, and $DaPFC \rightarrow SpPFC$.

Unlike the results from the synthetic data, the real dataset produced a fully connected graph with connections between all nodes. Therefore, examining the weights w of the connections made more sense than the binary connection map. We calculated the difference in weights between the forward and backward influence strengths (e.g., the weight of *SaPFC*$\rightarrow$*DaPFC* minus the weight of *DaPFC*$\rightarrow$ *SaPFC*) and recorded the number of trials when Δw was greater than 0. This enabled us to determine the number of trials with a stronger forward influence compared to those with a stronger backward influence. Using a threshold of 70%, we identified the topologies shown in Fig. 5. We highlighted the expected paths based on the strongest, direct anatomical connections.

Our observations can be summarized as follows: 1) the pathway $SpPFC \rightarrow SaPFC$ contributed in the abstract rule epoch, consistent with initial sensory information being transmitted to higher levels of the frontal lobe; 2) the pathway $SaPFC(46d) \rightarrow DaPFC(46d)$ contributed in both the abstract and concrete rule epochs, consistent with the processing of rule information; 3) the pathway $DaPFC \rightarrow SpPFC$ contributed in the concrete rule epoch, consistent with information about the appropriate rule-based action being transmitted to brain areas more closely related to action execution; and 4) additional pathways suggesting even weaker anatomical connections can significantly influence information processing in brain networks. Overall, the results support the idea that prefrontal cortical networks are hierarchically organized, with higher levels in anterior prefrontal cortex processing more abstract information, and lower levels in posterior prefrontal

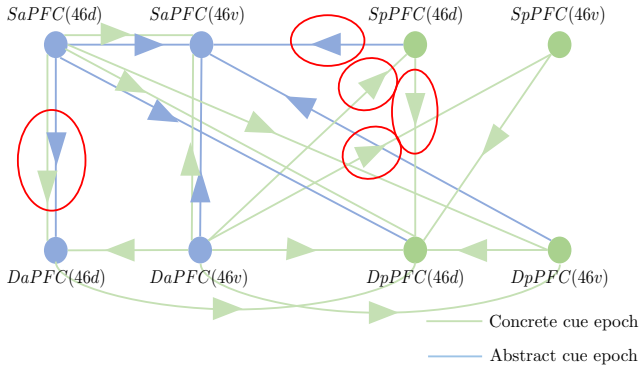

Fig. 5. Connection map for the macaque data. The red circles indicate known anatomical connections.

cortex processing more concrete information, closer to action specification.

V. Conclusion

In this paper, we proposed the processing of recorded LFPs with GPs to infer the topology of a studied brain network. The GPs were based on an ARD kernel that was used to determine the strengths of directed connections between a set of nodes to a given node in the brain. The proposed methodology allows for identifying which delayed signals from other nodes influence the studied node. We tested the method on two sets of data, one synthesized and the other real. Both sets represent neural activity during a memory task. The synthesized data were generated by a computational model of the memory task and therefore, in this experiment we had the ground truth of the topology. The real data were recorded from a macaque's prefrontal cortical regions while the macaque was performing memory tasks. The results from the synthesized data showed very good performance of the proposed method. The results on the macaque data demonstrated a good agreement with our understanding of the operation of the prefrontal cortical networks during memory tasks.

References

[1] G. Bahg, D. G. Evans, M. Galdo, and B. M. Turner. Gaussian process linking functions for mind, brain, and behavior. *Proceedings of the National Academy of Sciences*, 117(47):29398–29406, 2020.

[2] G. Chen, D. R. Glen, Z. S. Saad, J. P. Hamilton, M. E. Thomason, I. H. Gotlib, and R. W. Cox. Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis. *Computers in Biology and Medicine*, 41(12):1142–1155, 2011.

[3] C. Cui, P. Banelli, and P. M. Djurić. Gaussian processes for topology inference of directed graphs. In *2022 30th European Signal Processing Conference (EUSIPCO)*, pages 2156–2160. IEEE, 2022.

[4] C. Cui, P. Banelli, and P. M. Djurić. Topology inference of directed graphs by Gaussian processes with sparsity constraints. *IEEE Transactions on Signal Processing*, 2024.

[5] C. Cui and P. M. Djurić. Inference of time–varying graph topologies via gaussian processes. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 13276–13280. IEEE, 2024.

[6] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst. Learning Laplacian matrix in smooth graph signal representations. *IEEE Transactions on Signal Processing*, 64(23):6160–6173, 2016.

[7] G. T. Einevoll, C. Kayser, N. K. Logothetis, and S. Panzeri. Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience*, 14(11):770–785, 2013.

[8] Y. Grandvalet and S. Canu. Adaptive scaling for feature selection in SVMs. *Advances in Neural Information Processing Systems*, 15, 2002.

[9] L. Hu, M. Guindani, N. J. Fortin, and H. Ombao. A hierarchical bayesian model for differential connectivity in multi-trial brain signals. *Econometrics and Statistics*, 15:117–135, 2020.

[10] M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint arXiv:1807.02582*, 2018.

[11] L. Li, D. Pluta, B. Shahbaba, N. Fortin, H. Ombao, and P. Baldi. Modeling dynamic functional connectivity with latent factor gaussian processes. *Advances in Neural Information Processing Systems*, 32, 2019.

[12] Y. Liu, C. Cui, M. Ajirak, and P. M. Djurić. Estimation of time-varying graph topologies from graph signals. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, 2023.

[13] A. McIntosh and F. Gonzalez-Lima. Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*, 2(1-2):2–22, 1994.

[14] S. Mofakham. *Elucidating the Interplay of Structure, Dynamics, and Function in the Brain's Neural Networks*. PhD thesis, 2016.

[15] S. Mofakham, A. Fry, J. Adachi, P. L. Stefancin, T. Q. Duong, J. R. Saadon, N. J. Winans, H. Sharma, G. Feng, P. M. Djuric, et al. Electrocorticography reveals thalamic control of cortical dynamics following traumatic brain injury. *Communications Biology*, 4(1):1210, 2021.

[16] M. Moscu, R. Borsoi, and C. Richard. Online graph topology inference with kernels for brain connectivity estimation. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1200–1204. IEEE, 2020.

[17] R. M. Neal. *Bayesian Learning for Neural Networks*, volume 118 of *Lecture Notes in Statistics*. Springer, 1996.

[18] C. Schmidt, B. Pester, N. Schmid-Hertel, H. Witte, A. Wismüller, and L. Leistritz. A multivariate Granger causality concept towards full brain functional connectivity. *PLOS One*, 11(4):e0153105, 2016.

[19] K. R. Ulrich, D. E. Carlson, W. Lian, J. S. Borg, K. Dzirasa, and L. Carin. Analysis of brain states from multi-region lfp time-series. *Advances in Neural Information Processing Systems*, 27, 2014.

[20] G. Varoquaux, A. Gramfort, J.-B. Poline, and B. Thirion. Brain covariance selection: better individual functional connectivity models using population prior. *Advances in Neural Information Processing Systems*, 23, 2010.

[21] M. A. Vosoughi and A. Wismüller. Large-scale kernelized Granger causality (lsKGC) for inferring topology of directed graphs in brain networks. In *Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 12036, page 1203603. SPIE, 2022.

[22] T. Wang, H. Huang, S. Tian, and J. Xu. Feature selection for SVM via optimization of kernel polarization with Gaussian ARD kernels. *Expert Systems with Applications*, 37(9):6663–6668, 2010.