

EEG-estimated functional connectivity, and not behavior, differentiates Parkinson's patients from health controls during the Simon conflict task

1st Xiaoxiao Sun
Columbia University
New York, U.S.A
xiaoxiao.sun@columbia.edu

2nd Chongkun Zhao
Columbia University
New York, U.S.A
cz2715@columbia.edu

3rd Sharath Koorathota
Columbia University
New York, U.S.A
sk4172@columbia.edu

4th Paul Sajda
Columbia University
New York, U.S.A
psajda@columbia.edu

Abstract—Neural biomarkers that can classify or predict disease are of broad interest to the neurological and psychiatric communities. Such biomarkers can be informative of disease state or treatment efficacy, even before there are changes in symptoms and/or behavior. This work investigates EEG-estimated functional connectivity (FC) as a Parkinson's Disease (PD) biomarker. Specifically, we investigate FC mediated via neural oscillations and consider such activity during the Simons conflict task. This task yields sensory-motor conflict, and one might expect differences in behavior between PD patients and healthy controls (HCs). In addition to considering spatially focused approaches, such as FC, as a biomarker, we also consider temporal biomarkers, which are more sensitive to ongoing changes in neural activity. We find that FC, estimated from delta (1-4Hz) and theta (4-7Hz) oscillations, yields spatial FC patterns significantly better at distinguishing PD from HC than temporal features or behavior. This study reinforces that FC in spectral bands is informative of differences in brain-wide processes and can serve as a biomarker distinguishing normal brain function from that seen in disease.

Index Terms—Cognitive Control, Brain Oscillation, Functional Connectivity, Phase Synchronization, Parkinson's Disease

I. INTRODUCTION

With the advent of high-density neural recording tools (e.g., Neuropixels), there is an increasing interest in investigating neural activity in localized brain regions [1]. An alternative approach is to consider brain-wide activity, albeit at a coarser temporal scale (e.g., not measuring spiking activity), since the technology does not exist for such spatial coverage at the resolution of neuronal spikes, particularly for humans. Many brain-wide analyses have instead focused on lower temporal resolution neural oscillations (1-40Hz). Such oscillations are believed to link neural activity across disparate regions, potentially binding activity in a way that yields our integrated and coherent perception of the world [2].

Recent studies have shown brain oscillations drive functional connectivity (FC) across cortical and subcortical networks [3], [4]. Moreover, the specific spatiotemporal properties of such FC appear to correlate with normal cognitive,

This work was supported by a Vannevar Bush Faculty Fellowship from the US Department of Defense (N00014-20-1-2027) and a Center of Excellence grant from the Air Force Office of Scientific Research (FA9550-22-1-0337). All authors are affiliated with the Department of Biomedical Engineering. Dr. Sajda is also affiliated with the Department of Electrical Engineering, Radiology, and Data Science Institute.

emotional, and perceptual functions as well as neuropathologies [5], [6]. For instance, investigating differences in FC between disorders such as schizophrenia and bipolar disorder has provided insights into possible disease mechanisms [7]. This growing body of research has highlighted the potential utility of FC as a biomarker for identifying cognitive function and diagnosing brain disorders.

In this study, we investigate EEG-estimated FC as a potential biomarker for Parkinson's disease (PD) when both PD patients and healthy controls (HCs) are engaged in the Simon conflict task. Using different machine learning (ML) models, we systematically assess the relative importance of temporal and spatial-focused features in classifying PD vs. HC. Our findings show that spatial features represented by EEG-estimated FC are extremely good at classifying PD vs. HC, while models focusing on temporal features perform substantially worse. Notably, neither spatial-FC features nor temporal features are particularly good at classifying trial conditions (i.e., congruent vs incongruent trials). This suggests that FC is a potential spatial biomarker for PD that is not simply an artifact of brain signals that differentiate task performance or task condition. Based on this finding, we also investigated the significance of functional connectivity patterns using a statistical approach – the surrogate test. The results reflect a similar interpretation, evidenced by the higher number of electrode pairs that remain relevant after applying a threshold for group comparisons, as opposed to a reduced number of pairs for trial condition comparisons.

II. RELATED WORK

A. Brain Oscillations and Cognitive Coordination

Neuronal oscillations are believed to play various roles in brain function [4]. Because neuronal oscillations can change the dynamic interactions between brain regions, they have been proposed as a mechanistic gate for routing information on a fast time scale [8]. Oscillations in different frequency bands are believed to be essential for flexible and coordinated brain function, with recurrent connections between excitatory and inhibitory cell populations producing resonant oscillation frequencies that depend on the decay time constants in the neural populations [9].

B. Simon Conflict Task

The Simon conflict task is a widely employed cognitive paradigm designed to investigate the influence of conflicting information on response inhibition and decision-making processes. The task engages participants in a visual-spatial conflict by presenting stimuli in a way that creates a misalignment between the spatial location of the stimulus and the required response [10]. It has been shown valuable for probing the mechanisms underlying cognitive control, attentional processing, and the brain's ability to resolve conflicting information [11]. The task requires flexible and coordinated processing that is typically characterized by the term "cognitive control".

C. Functional Connectivity

Oscillatory neuronal activity may provide a mechanism for dynamic network coordination and be the basis for what is typically called functional connectivity (FC). FC is believed to exist, and even change, while an individual is at rest or engaged in a task. FC is thought to be crucial for governing information processing and functional execution across diverse brain areas [4], [12]. Population differences in FC have also been investigated as being potential biomarkers of neurological disease and psychiatric illness [7], [13]–[15]. In many cases, machine learning is used to characterize features of FC and classify individuals or populations.

D. Machine Learning Models

a) Random Forest (RF) Models: Random Forest (RF) is a widely employed machine learning method for classification problems [16]. Based on the ensemble learning paradigm, RF leverages the power of decision trees by constructing an ensemble of them and subsequently combining their predictions [17]. By aggregating the outcomes of multiple decision trees through a voting mechanism, RF models offer robustness against noise and outliers, making it a versatile tool in various domains.

b) Long Short-Term Memory (LSTM) Networks: A Long Short-Term Memory (LSTM) network is a type of recurrent neural network (RNN) architecture that has gained prominence in deep learning due to its ability to capture and model long-range dependencies in sequential data [18]. Unlike traditional RNNs, LSTM utilizes specialized gating mechanisms, including input, forget, and output gates, which enable them to selectively store and update information over extended sequences, thus mitigating the vanishing gradient problem [19]. Owing to their proficiency in handling sequential data of varying lengths and complexities, LSTMs have found widespread applications across various fields, including natural language processing, speech recognition, time series forecasting, and neural data analysis.

c) Convolution Neural Networks (CNN): Convolution Neural Networks (CNNs) have revolutionized the field of computer vision and image analysis due to their exceptional ability to extract hierarchical features from visual data [20]. The architecture's success can be attributed to its ability to learn discriminative features from data through training. These

neural networks, inspired by the human visual system, consist of multiple layers of learnable filters that perform convolution operations on input images, enabling them to capture local and global patterns [21].

III. METHODOLOGY

A. Experimental Dataset

We investigated functional connectivity within the context of the Simon conflict task using an open-source EEG dataset [22]. In the task, a stimulus is presented on either the left or right side of the screen and participants are instructed to press the left key when the stimulus is yellow/red and the right key when it is cyan/blue [11]. A congruent trial is when the stimuli match the side of the screen where the response hand is located, while an incongruent trial is when the stimuli appear on the opposite side of the screen from the response hand (i.e., spatially congruent/incongruent, see Fig. 1a).

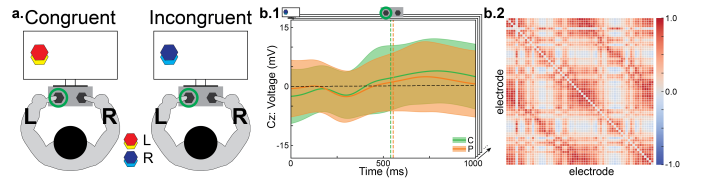


Fig. 1. Overview of the experimental setup and the extracted temporal/spatial features. **a.** Illustration of the Simon conflict task (see text for details). **b.1** Epoched 1-second temporal feature after the stimulus onset, specifically the voltage at electrode Cz within the delta band for incongruent trials. The average response time (RT) is 549.6 ± 113.1 ms for HCs and 563.2 ± 156.2 ms for PD patients in the incongruent trials (highlighted with the vertical dashed line separately, see Fig. 3a.1 and a.2 for more details about RT). **b.2** A corresponding spatial feature from the same time window. The difference in functional connectivity (represented by Pearson Correlation) between HC and PD patients is plotted as an example.

TABLE I
EXPERIMENT DATA SUMMARY

# of Trials	Patient (49)	Control (26)	Total	Average ^a
Congruent	7422	4456	11878	158±27
Incongruent	6655	4235	10890	145±34
Total	14077	8691	22768	-

^aAverage within each subject by trial condition.

Our analysis includes a subset of 75 subjects from the open-source dataset, encompassing 49 patients with PD and 26 HCs (see experiment summary in Table I). EEG was recorded using a 64-channel BrainVision system. The EEG was high-pass filtered at 0.1Hz and down-sampled to 500Hz. Electrode Pz was used as the reference, and electrode Fpz was used as the ground. EEG activity at electrode FT9 was excluded since a specialized cap without FT9 was employed, resulting in 63 channels for further analysis (see TABLE II). Additional details about subject recruitment and data recording are available in their latest paper [11].

B. Temporal Biomarkers: Band-passed electrode time series (BPT)

Since low-frequency oscillations are modulated by conflict, attention, and timing and are also known to be specifically impaired in PD [23], [24], we focus our analysis of temporal biomarkers in delta (1 to 4Hz) and theta (4 to 7Hz) bands of the EEG. Bidirectional filtering was applied to the raw

EEG data using a bandpass filter (Matlab2023a) to attenuate frequencies outside the delta and theta bands. Event-related potentials (ERPs) were extracted with respect to the onset of the stimuli (marked as 0 seconds), spanning from -0.5 seconds before to 1.5 seconds after the stimulus. Baseline removal was done based on the entire epoch. To align the brain activity with the behavioral response, we exclusively considered the 1-second interval following the stimulus onset (see Fig. 1b.1 and b.2), as the maximum response time (RT) typically falls within this time frame (see Fig. 3a.1 and a.2). This extracted 1-second band-passed time series (BPT) of all electrodes were then used as temporal feature inputs for ML models (see TABLE II).

C. Spatial Biomarkers: Functional Connectivity (FC)

FC can be quantified using multiple different metrics, each possessing its own strengths and limitations [25]. Here, we employed three non-directed methods for FC measurement, encompassing amplitude-based, phase-based, and amplitude-phase-mixed approaches.

1) *Pearson Correlation (PC)*: The simplest metric for non-directed interactions is the Pearson correlation coefficient, which measures the linear relationship between two random variables [26]. PC focuses on the dynamic connection between regions represented by their voltage (i.e., amplitude) changes.

$$PC(x, y) = \frac{\sum_{i=1}^n (x(\omega, i) - \bar{x})(y(\omega, i) - \bar{y})}{\sqrt{\sum_{i=1}^n (x(\omega, i) - \bar{x})^2 \sum_{i=1}^n (y(\omega, i) - \bar{y})^2}} \quad (1)$$

where $x(\omega, i)$ ($y(\omega, i)$) are the values of x (y) in a sample i of target frequency band ω (i.e., theta, alpha, etc.); n is the number of samples; \bar{x} (\bar{y}) is the mean of x (y).

2) *Phase Locking Value (PLV)*: An alternative metric to measure FC is to evaluate the synchronization between cortical areas (e.g., between electrodes from EEG recordings). PLV estimates how the relative phase between signals is distributed over the unit circle [27]. PLV is a scalar value that ranges from [0, 1], where 1 indicates the strongest phase synchronization between regions/electrodes [28].

$$PLV(x, y) = \frac{|\frac{1}{n} \sum_{i=1}^n 1_x(\omega, i) 1_y(\omega, i) e^{i(\phi_x(\omega, i) - \phi_y(\omega, i))}|}{\sqrt{(\frac{1}{n} \sum_{i=1}^n 1_x^2(\omega, i)) (\frac{1}{n} \sum_{i=1}^n 1_y^2(\omega, i))}} \quad (2)$$

where $\phi_x(\omega, i)$ ($\phi_y(\omega, i)$) is the phase of x (y) obtained by the Hilbert transform in a sample i of targeted frequency band ω ; n is the number of samples.

3) *Phase Coherence (COH)*: A measurement similar to PLV that represents FC with phase synchronization is COH. It is basically an extension of PLV. Instead of only considering the phase after the Hilbert transform, COH is computed from both the amplitude and the phase [25].

$$COH(x, y) = \frac{|\frac{1}{n} \sum_{i=1}^n A_x(\omega, i) A_y(\omega, i) e^{i(\phi_x(\omega, i) - \phi_y(\omega, i))}|}{\sqrt{(\frac{1}{n} \sum_{i=1}^n A_x^2(\omega, i)) (\frac{1}{n} \sum_{i=1}^n A_y^2(\omega, i))}} \quad (3)$$

where $A_x(\omega, i)$ ($A_y(\omega, i)$) is the amplitude and $\phi_x(\omega, i)$ ($\phi_y(\omega, i)$) is the phase of x (y) obtained by Hilbert transform

in a sample i of targeted frequency band ω ; n is the number of samples.

D. Comparing Machine Learning Models

We used the RF model as a benchmark model for comparing classification performance and assessing the relative effectiveness of temporal (i.e., voltage changes of ERPs) and spatial (i.e., FC) features. LSTMs were used to analyze temporal features, and CNNs used for spatial features, given that their designs matched these dimensions. We also compared the spatial feature weights extracted from different models to gain deeper insight into the importance of FC features across different ML models. By employing this diverse set of ML architectures, our aim was to comprehensively explore the features embedded within the neural data (see Table II). The dataset (see Table I) was divided into three subsets: 64% for training, 16% for validation, and the remaining 20% for testing. We considered the imbalanced dataset distribution to ensure robust and accurate results and applied a 10-fold cross-validation during training.

TABLE II
MODEL ARCHITECTURES SUMMARY

Model	Input Shape	Layers	Para Size
RF ^a	(31500,)	-	-
RF ^b	(1953,)	-	-
LSTM1 ^c	(63, 500)	(256, 256)-256	3.51MB
LSTM2	(63, 500)	(256, 128)-256	2.13MB
LSTM3	(63, 500)	(128, 128)-256	1.01MB
LSTM4	(63, 500)	(128, 64)-256	647.00KB
LSTM5	(63, 500)	(64, 64)-256	327.00KB
LSTM6	(63, 500)	(64, 32)-256	214.50KB
CNN1 ^d	(63, 63)	(32)-256	26.29MB
CNN2	(63, 63)	(16)-256	13.15MB
CNN3	(63, 63)	(16)-128	6.58MB
CNN4	(63, 63)	(16)-64	3.29MB
CNN5	(63, 63)	(16)-32	1.65MB
CNN6	(63, 63)	(8)-32	842.69KB

^aInput shape is transformed into a flattened representation (31500=63×500).

^bInput shape is transformed into a flattened representation, and repeated values have been removed (1953=63×62/2). ^cThe number within parentheses represents the dimensionality of the LSTM layer, and the value following the hyphen(-) indicates the size of the dense layer. ^dThe number within parentheses represents the dimensionality of the CNN layer, and the value following the hyphen indicates the size of the dense layer.

For all the FC approaches, we use the BPT data to compute the FCs. Thus, both the BPT and FC use the same time window of EEG and frequency bands, however, BPT maintains the time-series information while FC collapses it into a spatial connectivity measure across the brain.

E. Surrogate Tests

Since phase-related functional connectivity metrics such as PLV and COH are usually highly sensitive and can yield spurious results (e.g., two narrow-band Gaussian white noise processes can generate a moderate phase-related index), we employed the surrogate test to detect the statistically significant patterns in FC represented by PLV and COH [30]. The essence of the surrogate test is to test statistical significance from an artificially generated null distribution of surrogate signals that simulate the Fourier Power Spectrum of the

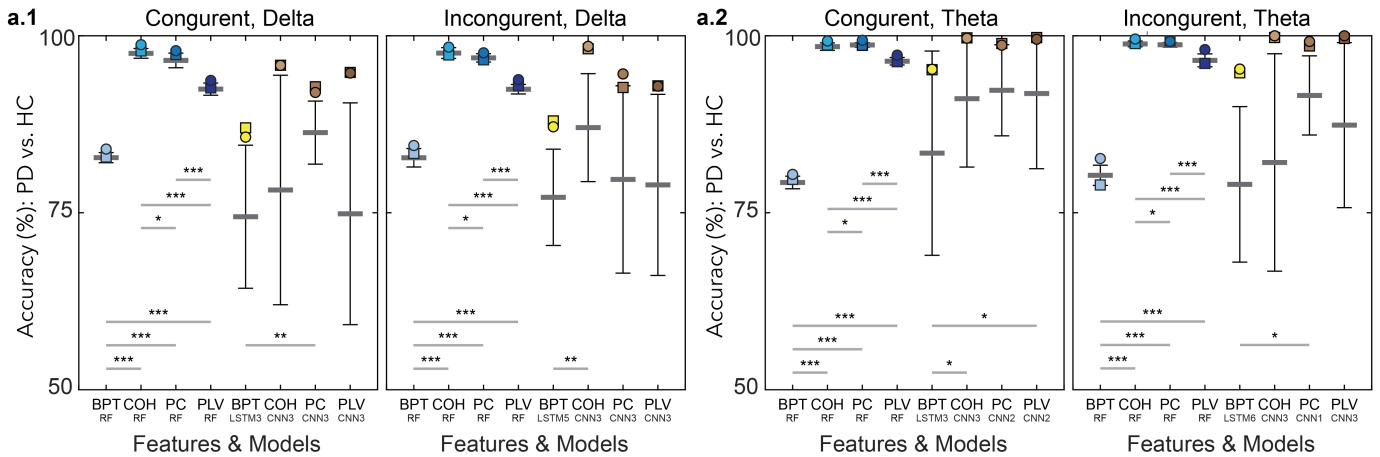


Fig. 2. Prediction accuracy (HC vs. PD) from ML models. **a.1** Model performance with features in the delta band (1-4Hz). Among the LSTM and CNN models trained and evaluated (see Table II), we show the model having the highest test accuracy (highlighted with square markers). The bar plots are generated using 10-fold cross-validation results (mean \pm standard deviation). Each model's highest validation accuracy is highlighted with a circle marker. We used the Mann-Whitney-Wilcoxon test given the small sample sizes ($n=10$) [29]. The significance level of the test result (multi-comparison un-corrected) is indicated by a black asterisk, where *** indicates significance under a 99.9% confidence interval, ** indicates significance under a 99% confidence interval, and * indicates significance under a 95% confidence interval. **a.2** is similar to **a.1**, except that models are evaluated in the theta band (4-7Hz).

original signal. To generate this null distribution, we use an unwrapped Fourier transform algorithm [31], which randomizes the phases by multiplying each instantaneous phase by $e^{i\phi}$, where ϕ is independently chosen from a uniform distribution $U([0, 2\pi])$.

We simulated 100 surrogate signals of the original signal for each trial and computed the corresponding surrogate FC in PLV(COH). The statistical significance of each PLV(COH) measurement was assessed by comparing it to the distribution of surrogate PLVs (COHs) ($p < 0.05$). Repeating this method for all electrode pairs for one trial returned a subset of statistically significant electrode pairs, represented by a feature matrix where 1 indicates a significant pair and 0 indicates an insignificant pair.

IV. RESULTS

A. Model Prediction Efficiency and Accuracy

We investigated two binary classification tasks: group differentiation (PD vs. HC), which is our primary goal, and trial condition differentiation (congruent vs. incongruent). The reason for the latter is that we wanted to check that any potential biomarker of group differentiation was not simply due to detecting differences in how the two populations (PD and HC) did the task. We found for all models and both biomarker types (BPT and FC), there was no ability to discriminate between the trial conditions (mean accuracy 51%, $p = 0.564$). We did, however, find significant discrimination of group (PD vs HC) using the biomarkers, across all ML model types (see Table II). We saw this discrimination both in the congruent trial and incongruent trial data, further indicating that the biomarkers were not artifacts of task condition differences but rather intrinsic disease states.

The spatial-focused features (i.e., FC) consistently outperformed temporal features (i.e., BPT) in both the delta and theta bands. Within the delta band, the amplitude-phase-mixed FC metric, i.e., COH, showed the highest validation and test

accuracy using the RF model among congruent trials and the CNN model among incongruent trials (see Fig. 2a.1). Amplitude-based prediction using the theta band performed better than the delta band (87% in delta vs 95% in theta with the LSTM). The amplitude-based FC metric, i.e., PC, results in the highest validation and test accuracy using the RF model for the data from the congruent trials and the incongruent trials (see Fig. 2a.2). For CNN models, COH-derived FC usually generates the best performance, though when considered statistically, there were no significant performance differences between the different FC metrics. Although similar test accuracies are obtained in both RF and CNN models given spatial features as input, the RF model showed more robust performance across validation folds, reflected in the smaller variance in validation accuracy.

We used the best model (i.e., has the highest test accuracy) obtained from trial-level prediction of PD vs HC to predict the group condition at the subject-level. We did this by computing the fraction of trials that the subject was characterized as PD vs HC, and we call this the subject-level probability. We also compared the results to just using behavior in the task (classifying by RT). Fig. 3b.1 shows the probabilities of one subject being classified as a PD patient with FC calculated in the delta band. Notably, both temporal and spatial features effectively distinguish between different group conditions. However, the models utilizing spatial features have better accuracy, with the subject-level probability scores being closer to 0 for HC subjects and closer to 1 for PD patients. This distinction underscores the efficacy of spatial features in effectively differentiating between groups on an individual level. Analyses within the theta band yielded similar results (see Fig. 3b.2).

B. Evaluation of Spatial Biomarker Patterns using Surrogate Tests

We performed the surrogate test analysis to compare the spatial FC patterns between HC and PD computed for different

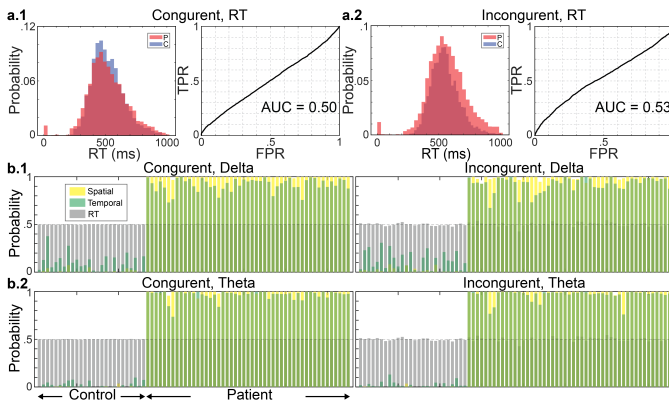


Fig. 3. Classification probability given different biomarkers. **a.1** Histogram of RT and receiver operating characteristic (ROC) curve of HC vs PD classification based on reaction time (RT). **a.2** is similar to **a.1**, but for the incongruent trials. **b.1** Subject-level probabilities of classifying HC vs PD. Both spatial FC (yellow) and temporal BPT (green) biomarkers outperform RT (gray), with 26 controls having probabilities less than 0.5 and 49 patients having probabilities greater than 0.5. The spatial FC biomarkers are clearly more accurate than the temporal BIT biomarkers across the population.

trial conditions (congruent or incongruent). We first computed a significance mask across the FC dimensions using a $p < 0.05$ criterion under the null hypothesis created by the surrogate data. We applied this mask to all trials for each condition and then computed the percentage of trials that passed this test against the null (see Section III-E). We considered two thresholds, 60% and 80%. For example, a 60% threshold in Fig. 4 represents FC patterns that pass the surrogate test for more than 60% of all trials. We compared the FC network patterns between groups (HC vs. PD). We see that most of the dimensions of the FC biomarkers are common across the congruent and incongruent trials (purple color Fig. 4). This is consistent with our previous findings which showed that the ML classifiers could not distinguish trial type using the FC biomarker –i.e. there is no difference between the congruent or incongruent trial types. In addition, we see that the theta band derived FC has more robust FC dimensions relative to those derived from the delta band–i.e. more dimensions survive the surrogate test at the 80% threshold in theta than for delta.

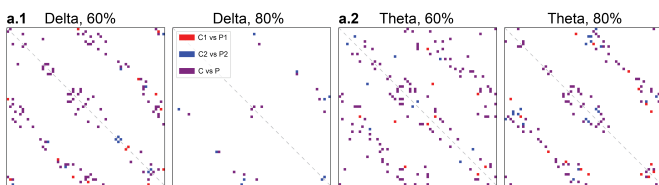


Fig. 4. The surrogate test identified significantly ($p < 0.05$) different functional connectivity (FC) patterns under different thresholds (60% and 80%). **a.1** Different FC patterns between groups within the delta band. Each color represents a comparison within one trial condition (red: congruent, labeled as 1 in the legend, e.g., C1 means HC-congruent; blue: incongruent, labeled as 2 in the legend, purple: shared in both conditions). **a.2** follows the layout of **a.1**, except that it is computed based on the theta band.

V. DISCUSSION AND CONCLUSION

This work investigated spatial and temporal biomarkers of Parkinson’s disease (PD) derived from EEG data recording while participants performed the Simon conflict task. We

found that spatial biomarkers, specifically those constructed to represent functional connectivity, were superior to temporal biomarkers in classifying PD vs. HC. Noteworthy was that the FC and BPT temporal features could not be used to classify trial conditions (congruent vs. incongruent), suggesting that the underlying biomarkers were not task-specific but related to fundamental differences in brain function across the groups and independent of the task. We did not have data to test whether PD vs. HC classification was possible with these biomarkers in resting state EEG, or during another task, so it remains to be seen if the performance of the task is critical for inducing brain states that differentiate the two populations. Given that our analysis was event-locked (to the stimulus presented in this case), it does suggest that a task is desirable for aligning the neural data to construct the biomarkers.

We further explored the significance of the FC biomarkers via a statistical surrogate test with results consistent with our ML results, namely that the FC biomarkers capture spatial information that differentiates PD from HC and which is not simply a result of differences in task performance or trial condition. Our work supports the potential diagnostic value of spatial features, specifically functional connectivity derived from EEG measures.

REFERENCES

- [1] M. J. Kahana, “The cognitive correlates of human brain oscillations,” *Journal of Neuroscience*, vol. 26, no. 6, pp. 1669–1672, 2006.
- [2] L. M. Ward, “Synchronous neural oscillations and cognitive processes,” *Trends in cognitive sciences*, vol. 7, no. 12, pp. 553–559, 2003.
- [3] M. D. Fox, R. L. Buckner, M. P. White, M. D. Greicius, and A. Pascual-Leone, “Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate,” *Biological psychiatry*, vol. 72, no. 7, pp. 595–603, 2012.
- [4] M. Bonnefond, S. Kastner, and O. Jensen, “Communication between brain areas based on nested oscillations,” *neuro*, vol. 4, no. 2, 2017.
- [5] M. A. Smith and A. Kohn, “Spatial and temporal scales of neuronal correlation in primary visual cortex,” *Journal of Neuroscience*, vol. 28, no. 48, pp. 12591–12603, 2008.
- [6] G. Northoff, S. Wainio-Theberge, and K. Evers, “Spatiotemporal neuroscience—what is it and why we need it,” *Physics of life reviews*, vol. 33, pp. 78–87, 2020.
- [7] B. Birur, N. V. Kraguljac, R. C. Shelton, and A. C. Lahti, “Brain structure, function, and neurochemistry in schizophrenia and bipolar disorder—a systematic review of the magnetic resonance neuroimaging literature,” *NPJ schizophrenia*, vol. 3, no. 1, p. 15, 2017.
- [8] F. Varela, J.-P. Lachaux, E. Rodriguez, and J. Martinerie, “The brainweb: phase synchronization and large-scale integration,” *Nature reviews neuroscience*, vol. 2, no. 4, pp. 229–239, 2001.
- [9] K. Kang, M. Shelley, J. A. Henrie, and R. Shapley, “Lfp spectral peaks in v1 cortex: network resonance and cortico-cortical feedback,” *Journal of computational neuroscience*, vol. 29, no. 3, pp. 495–507, 2010.
- [10] J. R. SIMON and J. D. WOLF, “Choice reaction time as a function of angular stimulus-response correspondence and age,” *Ergonomics*, vol. 6, no. 1, pp. 99–105, 1963.
- [11] A. Singh, R. C. Cole, A. I. Espinoza, J. R. Wessel, J. F. Cavanagh, and N. S. Narayanan, “Evoked mid-frontal activity predicts cognitive dysfunction in parkinson’s disease,” *Journal of Neurology, Neurosurgery & Psychiatry*, 2023.
- [12] S. P. Pantazatos, J. R. McIntosh, G. T. Saber, X. Sun, J. Dooze, J. Faller, Y. Lin, J. B. Teves, A. Blankenship, S. Huffman, *et al.*, “The timing of transcranial magnetic stimulation relative to the phase of prefrontal alpha eeg modulates downstream target engagement,” *Brain Stimulation*, vol. 16, no. 3, pp. 830–839, 2023.
- [13] Y. Du, Z. Fu, and V. D. Calhoun, “Classification and prediction of brain disorders using functional connectivity: promising but challenging,” *Frontiers in neuroscience*, vol. 12, p. 525, 2018.

- [14] Y. Tian and A. Zalesky, "Machine learning prediction of cognition from functional connectivity: Are feature weights reliable?," *NeuroImage*, vol. 245, p. 118648, 2021.
- [15] B. Bratić, V. Kurbalija, M. Ivanović, I. Oder, and Z. Bosnić, "Machine learning for predicting cognitive diseases: methods, data sources and risk factors," *Journal of medical systems*, vol. 42, pp. 1–15, 2018.
- [16] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.
- [17] M. Belgiu and L. Drăguț, "Random forest in remote sensing: A review of applications and future directions," *ISPRS journal of photogrammetry and remote sensing*, vol. 114, pp. 24–31, 2016.
- [18] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.
- [19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [20] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [21] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning cnn models for disease detection in plants using image segmentation," *Information Processing in Agriculture*, vol. 7, no. 4, pp. 566–574, 2020.
- [22] A. Singh, R. Cole, A. Espinoza, J. R. Wessel, J. C. jcavanagh@unm.edu, and N. Narayanan, "Dataset: Simon-conflict task.," OpenNeuro, 2023.
- [23] J. F. Cavanagh and M. J. Frank, "Frontal theta as a mechanism for cognitive control," *Trends in cognitive sciences*, vol. 18, no. 8, pp. 414–421, 2014.
- [24] A. Singh, R. C. Cole, A. I. Espinoza, A. Evans, S. Cao, J. F. Cavanagh, and N. S. Narayanan, "Timing variability and midfrontal 4 hz rhythms correlate with cognition in parkinson's disease," *NPJ Parkinson's disease*, vol. 7, no. 1, p. 14, 2021.
- [25] A. M. Bastos and J.-M. Schoffelen, "A tutorial review of functional connectivity analysis methods and their interpretational pitfalls," *Frontiers in systems neuroscience*, vol. 9, p. 175, 2016.
- [26] M. R. Arbabshirani, E. Damaraju, R. Phlypo, S. Plis, E. Allen, S. Ma, D. Mathalon, A. Preda, J. G. Vaidya, T. Adali, *et al.*, "Impact of autocorrelation on functional connectivity," *Neuroimage*, vol. 102, pp. 294–308, 2014.
- [27] G. Niso, R. Bruña, E. Pereda, R. Gutiérrez, R. Bajo, F. Maestú, and F. Del-Pozo, "Hermes: towards an integrated toolbox to characterize functional and effective brain connectivity," *Neuroinformatics*, vol. 11, no. 4, pp. 405–434, 2013.
- [28] X. Sun, J. Doose, J. Faller, J. McIntosh, G. T. Saber, S. Huffman, S. Pantazatos, H. Yuan, R. Goldman, T. Brown, *et al.*, "Increased entrainment and decreased excitability predict efficacious treatment of closed-loop phase-locked rtms for treatment-resistant depression," *medRxiv*, pp. 2023–10, 2023.
- [29] M. P. Fay and M. A. Proschan, "Wilcoxon-mann-whitney or t-test? on assumptions for hypothesis tests and multiple interpretations of decision rules," *Statistics surveys*, vol. 4, p. 1, 2010.
- [30] L. Xu, Z. Chen, K. Hu, H. E. Stanley, and P. C. Ivanov, "Spurious detection of phase synchronization in coupled nonlinear oscillators," *Physical Review E*, vol. 73, no. 6, p. 065201, 2006.
- [31] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, "Testing for nonlinearity in time series: the method of surrogate data," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 77–94, 1992.