**NeuroCluster: A Python toolbox to detect electrophysiological activations related to continuous behavioral signals using time-frequency resolved multiple regression and non-parametric cluster-based permutation testing.**

**# Summary**

Cognitive neurophysiology offers a novel framework for studying cognitive brain-behavior relationships by relating electrophysiological signals to complex behaviors. With the advent of new biotechnologies and neurosurgical practices, large-scale human (and animal) intracranial electrophysiological recordings are becoming widely accessible (CITATION). As a result, cognitive neurophysiologists can design cognitive experiments that leverage both the spatiotemporal resolution of electrophysiological data and the complexity of continuous behavioral variables (EX CITATIONS). Analyzing these data requires sophisticated statistical methods that can interpret multidimensional neurophysiological data and dynamic, continuous behavioral variables. Classical statistical frameworks for analyzing event-related time series data are ill-equipped to manage the high dimensionality and behavioral complexity of cognitive neurophysiology studies. NeuroCluster is an open-source Python toolbox for analysis of multivariate electrophysiological data related to complex, continuous behavioral variables. NeuroCluster introduces a novel statistical approach, which uses non-parametric cluster-based permutation testing to identify time-frequency clusters of oscillatory power modulations that significantly encode time-varying, continuous behavioral variables. It also supports multivariate analyses by allowing for multiple behavioral predictors to model neural activity. NeuroCluster addresses a methodological gap in statistical approaches to relate continuous, cognitive predictors to underlying electrophysiological activity with time and frequency resolution, to determine the neurocomputational processes giving rise to complex behaviors.

**# Statement of Need**

Determining the neurocomputational processes that give rise to human cognition and generate complex behaviors is a fundamental goal of cognitive and systems neuroscience. Cognitive neurophysiologists study the neural underpinnings of latent cognitive processes by relating complex behavioral signals to electrophysiological time series data. Cognitive behavioral signals, which can reflect experimental conditions, participant actions, or underlying cognitive processes, are often continuous and vary over time, especially in human behavioral experiments (CITATION). Computational cognitive models are used to operationalize unobservable cognitive processes and provide estimates of latent cognitive variables, based on participants’ behaviors (Pan et al., 2024). For instance, some cognitive models generate continuous, trial-wise value estimates, like reward prediction errors (RPEs, O’Doherty et al. (2007). Directly linking these cognitive variables to neurophysiological activity offers a dynamic way to study brain-behavior relationships. Unfortunately, the innate complexities of both cognitive behaviors and electrophysiological data (CITATION) presents a significant challenge for neuroscientists using model-based analyses to uncover the neurophysiological signatures of these processes.

NeuroCluster addresses a methodological gap in cognitive neurophysiology, by providing a novel statistical pipeline to relate continuous latent cognitive predictors to underlying electrophysiological activity, with both time and frequency resolution. Non-parametric statistical testing is the standard approach to analyze event-related time series data while controlling for multiple comparisons problems and reducing family-wise error rates (CITATIONS). However, current statistical methods are ill-equipped to interpret complex, cognitive behaviors, nor can they manage the high dimensionality of multi-region intracranial electrophysiological recordings (CITATION). Classic analysis methods relate neuronal activity to discrete behavioral categories, rather than continuous, trial-by-trial behavioral measures using one or two-sample cluster-based permutation tests (CITATION). Neurophysiological activity is typically aggregated by trial-type to perform a two-sample cluster-based permutation test, which tests whether the neuronal encoding patterns differ between two discrete task variables (CITATION). While two-sample permutation tests provide neurophysiological results in the time and frequency domains, they are insufficient for analyses relating neuronal activity to time-varying, continuous behavioral variables. Unfortunately, standard analysis methods capable of relating neural activity to complex, continuous variables sacrifice spectral resolution in either the time (trial-averaged signals) or frequency domains (broadband frequency-averaged signals) (CITATION). Reducing the spatiotemporal resolution of electrophysiological data hinders our ability to define distinct underlying mechanisms of cognitive processes, by eliminating either the temporal profile of a signal containing within and across-region encoding onset, duration, and latency patterns (CITATION) or eliminating the signal’s frequency-specificity, despite the widely accepted theory that oscillatory activity at different frequencies corresponds to distinct neurophysiological mechanisms (CITATION). Additionally, these approaches require neurophysiologists to define *a priori* hypotheses for relevant within-trial epochs and/or frequencies, reducing the generalizability of these analyses (CITATION). NeuroCluster addresses these shortcomings by implementing a novel statistical approach to identify significant clusters of oscillatory power modulations, with time-frequency resolution, related to trial-varying, continuous behavioral variables.

NeuroCluster is an open-source Python toolbox for identification of electrophysiological time-frequency activity related to continuous behavioral variables, using non-parametric cluster-based permutation testing. We demonstrate our approach with human intracranial local field potential data, but NeuroCluster provides functionality for all types of spatiotemporal or spectrotemporal neurophysiological measures (EEG, MEG) (CITATION) and may be applicable to phase-amplitude or phase-phase cross-frequency coupling analyses (CITATION). NeuroCluster is designed to supplement existing Python-based electrophysiological analysis toolboxes (CITATION FOOOF, MNE, eBOSC), such as MNE, which currently offers a cluster-based permutation testing approach for discrete group comparisons (mne.stats.permutation\_cluster\_test) (CITATION MNE). Additionally, NeuroCluster performs analyses with multivariate behavioral data by incorporating multiple predictors to model neural activity (CITATION?). NeuroCluster is amenable to analyses using the same statistical approach for model-based latent cognitive predictors (Pan et al., 2024, O’Doherty et al. 2007), model-free cognitive variables (CITATION), as well as continuous experimental (i.e., perceptual noise; Bang & Fleming (2018)) or behavioral (INTEROCEPTIVE?) (i.e., mood ratings; Blain & Rutledge (2020)) predictors. Our novel statistical method is applicable for numerous analysis goals; the major use cases are performing an initial exploratory analysis to generate specific hypotheses, determine data-driven temporal windows and/or frequencies of interest, or to identify regional patterns of significant clusters within and between subjects. Future directions for NeuroCluster may implement mixed effects regressions, multiple cluster detection, and/or group-level analysis tools (CITATIONS). NeuroCluster addresses a methodological gap in cognitive neurophysiology by implementing a novel statistical framework to relate continuous latent cognitive predictors to underlying time-frequency resolved neurophysiological signals. Directly linking electrophysiological activity to cognitive variables is crucial to understand the neurophysiological mechanisms facilitating complex behaviors and cognition.

**# Documentation**

A diagram of cluster

Description automatically generatedNeuroCluster is accompanied by a detailed tutorial [link to Jupyter notebook] which outlines the workflow (Fig 1) for implementing this approach with time-frequency power estimates from multi-region LFP recording.

**Fig 1.** **NeuroCluster workflow**. This approach involves three key steps: (1) determine cluster statistic in true data, (2) generate a null distribution of cluster statistics by permuting dataset, (3) determine significance of true cluster statistic against null distribution.

Below we outline the statistical approach implemented by NeuroCluster for performing non-parametric permutation-based cluster testing using time-frequency resolved power estimates from neural data estimated using Python-MNE and continuous predictors (i.e., latent cognitive processes, behavior, or experimental conditions). In these example data, we are testing the hypothesis that RPEs are significantly encoded in the electrophysiological signal from a given iEEG channel time-frequency representation (TFR).

1. ***Determine cluster statistic in true data*** 
   1. **Define clusters:** At each time-frequency index, we perform a linear univariate (or multivariate) regression using behaviorally-derived independent variables (e.g., latent cognitive variables, behavioral measures, task conditions) to predict neuronal activity (i.e., power). The coefficient represents the strength and direction of the relationship between each independent variable and the dependent variable. It is estimated from the regression model and reflects how changes in the independent variable are associated with changes in power at the specific time-frequency pair. Pixel-wise regressions are parallelized for speed. For each time-frequency pair, the coefficient for the regressor of interest (the independent variable of primary interest) is extracted from the regression results (**Fig 2A**). A t-statistic is computed for the coefficient to capture how significantly different it is from zero (**Fig 2B**). A significance threshold is applied to the t-statistics of the coefficient for the regressor of interest (**Fig 3**). If the t-statistic for a time-frequency pair exceeds the significance threshold, the pair is deemed significant. Clusters are then defined as adjacent time-frequency pairs where all pairs within the cluster have t-statistics exceeding the threshold, according to the test's desired tails (**Fig 2C**).
   2. **Compute cluster statistics:** For each identified cluster, sum the t-statistics of all time-frequency pairs within the cluster. In a two-tailed test (the default), compute both the maximum and minimum cluster sums (**Fig 2D**).
2. ***Generate null distribution of cluster statistics***
   1. **Permutation procedure:** Labels for the behavioral predictor of interest are shuffled for the desired number of permutations.
   2. **Recalculate cluster statistic:** Steps 1A/1B are repeated to define clusters and compute cluster statistics for each permuted dataset**.**
   3. **Construct null distribution:** The cluster statistics from all permutations are compiled to create a null distribution, representing the distribution of cluster statistics under the null hypothesis (**Fig 2E**). The permuted TFR regressions are also parallelized at the **pixel-level**, while each permutation is performed sequentially. We tested many iterations of these functions with different parallelization approaches and sequential permutation-level computations with pixel-level parallelization within each TFR regression was the fastest method.
3. ***Determine cluster significance*** 
   1. **Compare true cluster statistic to null distribution to compute p-values:**The proportion of cluster statistics in the null distribution falling above (or below) the true cluster statistic(s) determines the p-value associated with the cluster(s) identified in the true data (**Fig 2E**).

***A screenshot of a graph

Description automatically generated***

**Fig 2. NeuroCluster methods. A.**  coefficients for continuous predictor of interest (RPE) predicting power in given time-frequency pair (red outline = maximum positive cluster; blue outline = maximum negative cluster). **B.** T-statistics corresponding with RPE coefficients. C. Clusters as determined using t-critical threshold. **D.** Maximum positive and negative clusters determined by summing t-statistics in identified clusters. **E.** Null distribution of cluster statistics generated by permuting dataset for predictor of interest (100 permutations; red dashed line = true cluster statistic).