PROJECT DESCRIPTION

Modeling Latent Neural Dynamics to Decode Brain States and Motor Behavior in Parkinson's Disease

Giedrius Mirklys

Student Number: s1101773

August 3, 2025

1. ABSTRACT

2. PROJECT DESCRIPTION

According to the World Health Organization (2023), rates of Parkinson's disease (PD) increased twice over the last 25 years, with an estimated 8.5 million people affected worldwide in 2019. In PD, the primary pathology is the loss of dopaminergic neurons in the substantia nigra, which substantially disrupts communication within the cortico-basal ganglia-thalamo-cortical loop. This disruption manifests in slowness of movement, rigidity, and tremor, originating primarily from genetic and environmental factors (Ben-Shlomo et al., 2024). Neurophysiologically, Tinkhauser et al. (2017) reported that the prolonged oscillations in the beta frequency band (13-30 Hz) are also indicative of such a disruption, and its power has been correlated with the severity of the motor symptoms, including akinesia and rigidity, whereas in the healthy brain, the beta oscillations typically occur as transient short bursts during motor execution.

Therapeutic strategies aim to correct these disturbances at the network level, reducing the beta frequency synchrony within the structures in basal ganglia and between the motor cortex and the basal ganglia itself (Pauls et al., 2022; Tinkhauser et al., 2017). Pharmacological treatments, primarily Levodopa which increases levels of dopamine, have been shown to suppress prolonged beta oscillations, improving motor performance. For patients with more advanced symptoms, Deep Brain Stimulation (DBS) offers an effective alternative by implanting electrodes in the subthalamic nucleus (STN) within the basal ganglia. DBS is thought to work by imposing a new, high-frequency electrical pattern that acts as an *informational lesion*, overriding the pathological beta rhythm and preventing its propagation through the network (Chiken & Nambu, 2016; McIntyre & Hahn, 2010). Additionally, as highlighted by Y. Wu et al. (2024), the effects of DBS are multifaceted, inducing outcomes like synaptic plasticity and broader neural reorganization. However, the continuous, non-adaptive nature of the stimulation can lead to side effects, which are often attributed to the spread of electrical current to adjacent neural structures not targeted for therapy (Zarzycki & Domitrz, 2020).

Given the complex and noisy relationship between raw neural signals and clinical symptoms, computational models are essential for learning this mapping and translating high-dimensional data into reliable biomarkers that can drive a closed-loop therapeutic approach,

such as *adaptive* DBS (aDBS). The rationale for aDBS is that selectively targeting beta oscillations, where the symptoms are pronounced the most, could enable a more efficient, closed-loop approach to stimulation, potentially reducing side effects (Little et al., 2013). To this end, this thesis will evaluate two state-of-the-art modeling frameworks: the linear Preferential Subspace Identification (PSID) model (Sani et al., 2021) and the non-linear Dissociative Prioritized Analysis of Dynamics (DPAD) model (Sani et al., 2024). While beta oscillations are a leading biomarker, it remains debated whether they are causal to motor symptoms or are an epiphenomenon, and whether patient-specific biomarkers involving other spectral features, such as beta phase-amplitude coupling, might be more effective (Swann et al., 2016; Y. Wu et al., 2024). By comparing the performance of linear and non-linear approaches on this decoding problem, this thesis aims to characterize the capabilities and limitations of these models.

2.1. Preferential Subspace Identification (PSID)

Sani et al. (2021) described two standard approaches to modeling neural dynamics: Neural Dynamic Modeling (NDM) and Representational Modeling (RM). NDM typically learns a latent state that best predicts future neural activity from past neural activity, making it agnostic to behavior. Conversely, RM often models the dynamics of behavior itself, predicting future behavior from past behavior, and then relates this to neural activity, making it agnostic to the intrinsic dynamics of the neural signals. Neither approach is explicitly designed to isolate the dynamics that are shared between neural activity and behavior. This leads to the challenge of separating behaviorally relevant dynamics, which are the neural patterns that co-vary with and are predictive of behavior, from the vast background of behaviorally irrelevant dynamics related to other internal states and cognitive processes. Indeed, B.-S. Wu et al. (2025) elaborated that neural populations in motor-related areas, especially STN, are known to multiplex signals for numerous variables simultaneously, making it critical to disentangle the specific dynamics related to the behavior of interest.

To address it, Sani et al. (2021) developed Preferential Subspace Identification (PSID), a method that directly targets these shared dynamics. The main principle of PSID is that by training a model to predict future behavioral outputs from past neural activity, the model is forced to discover and represent the behaviorally relevant neural dynamics. To formalize this, PSID adopts a linear time-invariant (LTI) state-space model structure, which provides a tractable and interpretable foundation:

$$\begin{cases} \mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k = C_y\mathbf{x}_k + \mathbf{v}_k \\ \mathbf{z}_k = C_z\mathbf{x}_k + \boldsymbol{\epsilon}_k \end{cases}$$

Here, k is the time index; $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the unobserved, low-dimensional latent state that evolves according to the state transition matrix A; $\mathbf{y}_k \in \mathbb{R}^{n_y}$ is the observed neural activity from n_y channels, generated via the observation matrix C_y; and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ is the observed behavioral variable, generated via the observation matrix C_z. The terms $\mathbf{w}_k$, $\mathbf{v}_k$, and $\boldsymbol{\epsilon}_k$ represent state, neural observation, and behavioral residuals, respectively.

The PSID algorithm learns the model parameters and latent states through a non-iterative, closed-form procedure. It begins by constructing matrices of past neural activity ($\mathbf{Y}_p$) and future behavior ($\mathbf{Z}_f$). The "preferential" step is computing the orthogonal projection of future behavior onto past neural activity,

$$\hat{\mathbf{Z}}_f = \mathbf{Z}_f \mathbf{Y}_p^T (\mathbf{Y}_p \mathbf{Y}_p^T)^{-1} \mathbf{Y}_p$$

which isolates the component of future behavior that is linearly predictable from the neural signals. This projection is then decomposed using Singular Value Decomposition (SVD) to identify the observability matrix and the behaviorally relevant latent states. From this optimally identified subspace, the system matrices are estimated via linear regression. Specifically, the state transition matrix (A) is found by regressing future latent states onto current latent states, while the observation matrices (C_y, C_z) are found by regressing the observed neural and behavioral data onto the estimated latent state sequence. Finally, a Kalman filter is applied with these estimated parameters to compute the optimal sequence of the latent state $\mathbf{x}_k$. The final output is a low-dimensional, interpretable linear model that captures the neural dynamics most relevant to the behavior of interest.

2.2. Dissociative Prioritized Analysis of Dynamics (DPAD)

While PSID provides a robust linear framework, it is well-established that the brain's computations are fundamentally non-linear. Sani et al. (2024) proposed a new neural dynamics modelling framework called Dissociative Prioritized Analysis of Dynamics (DPAD). Although it may be reminiscent to PSID, but it is mathematically distinct from PSID. DPAD not only *prioritizes* behaviorally relevant dynamics but also explicitly *dissociates* them from irrelevant dynamics. This is achieved by partitioning the model's latent state $\mathbf{x}_k$ into two separate components: a prioritized state, $\mathbf{x}_k^{(1)}$, which is dedicated to predicting behavior, and a non-prioritized state, $\mathbf{x}_k^{(2)}$, which models all remaining neural variance. This dissociation is enforced through a four-step training process. The model is structured such that the behavior, $\hat{\mathbf{z}}_k$, is predicted only from the prioritized latent state, whereas the neural activity, $\hat{\mathbf{y}}_k$, is reconstructed from both states.

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}'(\mathbf{x}_k) + \mathbf{K}'(\mathbf{y}_k) \\ \hat{\mathbf{y}}_k = C_y^{(1)}(\mathbf{x}_k^{(1)}) + C_y^{(2)}(\mathbf{x}_k^{(2)}) \\ \hat{\mathbf{z}}_k = C_z^{(1)}(\mathbf{x}_k^{(1)}) \end{cases}$$

Here, $\mathbf{A}'$, $\mathbf{K}'$, $C_y^{(1)}$, $C_y^{(2)}$, and $C_z^{(1)}$ are now nonlinear functions (e.g. MLPs) that represent the state recursion, neural input, and observation mappings, respectively
First, the entire network is trained by forcing the prioritized section of the RNN and its latent state $\mathbf{x}_k^{(1)}$ to learn only the dynamics predictive of behavior. Second, the weights of this "prioritized" section are frozen. Third, the optimization objective shifts entirely to neural reconstruction, compelling the second, non-prioritized section of the RNN ($\mathbf{x}_k^{(2)}$) to learn and explain the residual neural variance not captured by the first stage. Finally, an optional fine-tuning step adjusts all parameters jointly. This process yields a non-linear model that not only achieves high predictive accuracy but also provides an interpretable, dissociated latent representation of the neural dynamics, making it a suitable framework for exploring the complex dynamics of Parkinson's disease.

3. Experimental Design and Methodology

The core of this thesis involves a comparison of the PSID and DPAD modeling frameworks across the three primary research questions.

3.1. RQ1: Cross-Modal Neural Prediction

Research Question 1: To what extent can PSID and DPAD predict cortical ECoG activity from subcortical LFP recordings, and how well do these predictions generalize across recording sessions for a given patient?

This question investigates the feasibility of neural signal translation. The models will be trained to predict ECoG using only LFP input, focusing on data from the DBS-OFF state. Prediction accuracy will be quantified using the Coefficient of Determination (R^2). For a deeper analysis of the learned relationship, Canonical Correlation Analysis (CCA) and Deep CCA (DCCA) will be employed to quantify the strength of the linear and non-linear coupling, respectively, between the LFP and ECoG signals.

3.2. RQ2: Brain State Classification

Research Question 2: How accurately can PSID and DPAD classify discrete brain states (DBS ON vs. OFF) from LFP/ECoG signals?

This task assesses the models' ability to identify distinct, clinically relevant brain states. The methodology involves using the learned latent states as features for a logistic regression classifier to perform a binary classification of the DBS ON/OFF condition. Performance will be assessed using metrics, including the Area Under the ROC Curve (AUC), the F1-Score, and Balanced Accuracy. To interpret the learned representations, the latent state trajectories will be visualized using t-SNE to confirm class separability. To quantitatively compare the learned representations, Representational Similarity Analysis (RSA) will be performed. Representational Dissimilarity Matrices (RDMs) will be constructed from both the neural data and the models' latent states, and their correlation will be calculated to assess the alignment between the model's internal geometry and the underlying neural geometry.

3.3. RQ3: Continuous Motor Behavior Decoding

Research Question 3: How effectively can PSID and DPAD decode continuous motor behavior (tracing speed) from LFP activity, and what do their respective latent dynamics reveal about the linear vs. nonlinear neural control of movement?

This final question addresses the core challenge for developing future adaptive DBS systems. The models will be trained to decode a continuous behavioral variable, tracing speed, from LFP signals. Decoding performance will be evaluated using Pearson's correlation coefficient. The quantitative link between the models' internal dynamics and the behavior will be established using CCA and DCCA to measure the correlation between the latent state trajectories and the continuous tracing speed data. For the DPAD model, the distinct contributions of its prioritized ($\mathbf{x}^{(1)}$) and non-prioritized ($\mathbf{x}^{(2)}$) subspaces will be explicitly compared. Furthermore, RSA will be used to compare the geometric structure of the latent spaces to the structure of the behavior itself.

3.4. Data and Preprocessing

This project utilizes data from a single Parkinson's disease patient from the Dareplane project, recorded over three separate sessions (Dold et al., 2024). These recordings provide simultaneous 16-channel LFP data from bilateral subthalamic nucleus (STN) electrodes and 4-channel

ECoG data from the primary motor cortex, originally sampled at 22 kHz. High-resolution hand kinematics were also captured, and the tracing speed will be calculated from the coordinate changes through time. The dataset also contains separate blocks of DBS ON and OFF states. Initial preprocessing addresses high-amplitude electrical artifacts in DBS-ON recordings using a template subtraction method (Hammer et al., 2022; Qian et al., 2017). This technique first derives an accurate template of the stereotyped artifact by averaging signal epochs aligned to the known stimulation pulse timings. To account for pulse-to-pulse variations in the artifact's shape, which can corrupt a simple average, this interpolation-based template rebuilding can be employed. The resulting template is then subtracted from the raw signal at each stimulation event to recover the underlying neural activity. The cleaned data are subsequently downsampled to 1000 Hz, band-pass filtered (3-250 Hz, including high gamma), and notch-filtered to remove power-line noise (50 Hz and its harmonics). A Common Average Reference (CAR) reduces common-mode noise, followed by Source Power Comodulation (SPoC) to derive task-specific spatial filters; the use of the latter one is set in the configuration files. Crucially, no handcrafted features such as bandpower are extracted; instead, the models are trained directly on segmented one-second epochs of the preprocessed time-series data. To rigorously assess model generalization across different recording days, a leave-one-session-out cross-validation scheme is employed (three sessions in total). For each of the three folds, models are trained and validated on data from two sessions, with the held-out third session serving as the final test set. Within each fold, the data from the two designated training-and-validation sessions are further subdivided into a training set (e.g., 80%) and a validation set (20%).

Word Count: 1899

4. SCHEDULE

Task	Start Date	End Date	Key Goals / Milestones
Finalizing project proposal	2025-09-01	2025-09-07	
Data Analysis & pipeline scaffolding	2025-09-08	2025-09-28	Pipeline for DBS/ECoG pre-processing
Building full training pipeline	2025-09-29	2025-11-02	Pipeline running a full experiment (load, preprocess, train, evaluate) from a YAML configs.
Drafting Background & Methodology	2025-09-29	2025-11-02	
Baselines model training	2025-11-03	2025-11-16	Generate first-pass results with PSID and a basic DPAD configuration on all RQs.
Main Training & Hyperparameter Tuning	2025-11-10	2025-12-21	Running all the experiments with different configurations.
In-depth Analysis of Initial Results	2025-11-17	2025-12-07	Analyze baseline results. Develop scripts for latent space analysis, create initial figures.
Refining Background & Methods Drafts	2025-12-08	2025-12-21	Incorporate new insights from initial results into the report drafts.
Christmas Break	2025-12-22	2026-01-04	Rest and light work on refining drafts as needed.
Getting full results	2026-01-05	2026-01-18	Finalize all model training and hyperparameter tuning. Get all the results and ensure they are reproducible and documented.
Drafting Results Section	2026-01-05	2026-01-25	Write the main narrative for the Results section using all finalized model outputs.
Drafting Discussion Section	2026-01-19	2026-02-08	Connect results back to the background and discuss implications.
Planning Thesis Presentation	2026-01-26	2026-02-08	
Drafting Introduction & Abstract	2026-02-09	2026-02-15	
Consolidating Full Report	2026-02-15	2026-02-22	Polish all sections, figures, and references into a single, cohesive document.
Finalizing Thesis Presentation/Slides	2026-02-23	2026-03-01	Finalize slides and practice the presentation.
Final Revisions, Unplanned Delays & Submission Prep	2026-03-02	2026-03-30	

5. SCIENTIFIC, SOCIETAL AND/OR TECHNOLOGICAL RELEVANCE

The clinical implications of this research extend directly to adaptive Deep Brain Stimulation (aDBS) for Parkinson's disease, aiming to transform the current standard of continuous stimulation into an intelligent, on-demand therapy. By validating computational models capable of decoding motor performance and classifying clinical states in real-time, this work provides the analytical foundation for closed-loop systems that could significantly enhance therapeutic efficacy while mitigating side effects. Furthermore, the exploration of cross-modal prediction investigates the potential for creating "virtual sensors," which could leverage the rich information from cortical signals without the need for additional invasive surgeries. However, the broader implications of this methodological framework are not confined to Parkinson's disease or motor control. The ability to dissociate and model behaviorally-relevant neural dynamics is a fundamental tool applicable to a wide range of neurological and psychiatric conditions, including epilepsy, depression, and obsessive-compulsive disorder. The same principles could be used to decode seizure precursors, affective states, or cognitive fluctuations, paving the way for adaptive neuromodulation therapies across the clinical spectrum.

6. REFERENCES

Ben-Shlomo, Y., Darweesh, S., Llibre-Guerra, J., Marras, C., San Luciano, M., & Tanner, C. (2024). The epidemiology of Parkinson's disease. *The Lancet*, *403*(10423), 283–292. https://doi.org/10.1016/S0140-6736(23)01419-8

Chiken, S., & Nambu, A. (2016). Mechanism of Deep Brain Stimulation: Inhibition, Excitation, or Disruption? *The Neuroscientist*, *22*(3), 313–322. https://doi.org/10.1177/1073858415581986

Dold, M., Pereira, J., Sajonz, B., Coenen, V. A., Thielen, J., Janssen, M. L., & Tangermann, M. (2024, September 12). *LFP and ECoG data during CopyDraw task with deep brain stimulation - Dareplane data for proof of concept paper* (Version 1). Radboud University. https://doi.org/10.34973/D214-M342

Hammer, L. H., Kochanski, R. B., Starr, P. A., & Little, S. (2022). Artifact Characterization and a Multipurpose Template-Based Offline Removal Solution for a Sensing-Enabled Deep Brain Stimulation Device. *Stereotactic and Functional Neurosurgery*, *100*(3), 168–183. https://doi.org/10.1159/000521431

Little, S., Pogosyan, A., Neal, S., Zavala, B., Zrinzo, L., Hariz, M., Foltynie, T., Limousin, P., Ashkan, K., FitzGerald, J., Green, A. L., Aziz, T. Z., & Brown, P. (2013). Adaptive deep brain stimulation in advanced Parkinson disease. *Annals of Neurology*, *74*(3), 449–457. https://doi.org/10.1002/ana.23951

McIntyre, C. C., & Hahn, P. J. (2010). Network perspectives on the mechanisms of deep brain stimulation. *Neurobiology of Disease*, *38*(3), 329–337. https://doi.org/10.1016/j.nbd.2009.09.022

Pauls, K. A. M., Korsun, O., Nenonen, J., Nurminen, J., Liljeström, M., Kujala, J., Pekkonen, E., & Renvall, H. (2022). Cortical beta burst dynamics are altered in Parkinson's disease but normalized by deep brain stimulation. *NeuroImage*, *257*, 119308. https://doi.org/10.1016/j.neuroimage.2022.119308

Qian, X., Chen, Y., Feng, Y., Ma, B., Hao, H., & Li, L. (2017). A Method for Removal of Deep Brain Stimulation Artifact From Local Field Potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(12), 2217–2226. https://doi.org/10.1109/TNSRE.2016.2613412

Sani, O. G., Abbaspourazad, H., Wong, Y. T., Pesaran, B., & Shanechi, M. M. (2021). Modeling behaviorally relevant neural dynamics enabled by preferential subspace identification. *Nature Neuroscience*, *24*(1), 140–149. https://doi.org/10.1038/s41593-020-00733-0

Sani, O. G., Pesaran, B., & Shanechi, M. M. (2024). Dissociative and prioritized modeling of behaviorally relevant neural dynamics using recurrent neural networks. *Nature Neuroscience*, *27*(10), 2033–2045. https://doi.org/10.1038/s41593-024-01731-2

Swann, N. C., De Hemptinne, C., Miocinovic, S., Qasim, S., Wang, S. S., Ziman, N., Ostrem, J. L., San Luciano, M., Galifianakis, N. B., & Starr, P. A. (2016). Gamma Oscillations in the Hyperkinetic State Detected with Chronic Human Brain Recordings in Parkinson's Disease. *The Journal of Neuroscience*, *36*(24), 6445–6458. https://doi.org/10.1523/jneurosci.1128-16.2016

Tinkhauser, G., Pogosyan, A., Tan, H., Herz, D. M., Kühn, A. A., & Brown, P. (2017). Beta burst dynamics in Parkinson's disease OFF and ON dopaminergic medication. *Brain*, *140*(11), 2968–2981. https://doi.org/10.1093/brain/awx252

World Health Organization. (2023, August). Parkinson disease [Accessed: August 3, 2025]. https://www.who.int/news-room/fact-sheets/detail/parkinson-disease

Wu, B.-S., Ming, M.-Y., & Wu, Y.-W. (2025, February 13). *Mixed Selectivity of Subthalamic Nucleus Neurons in Encoding Motor and Reward Behaviors*. https://doi.org/10.1101/2025.02.12.637797

Wu, Y., Hu, K., & Liu, S. (2024). Computational models advance deep brain stimulation for Parkinson's disease. *Network: Computation in Neural Systems*, 1–32. https://doi.org/10.1080/0954898X.2024.2361799

Zarzycki, M. Z., & Domitrz, I. (2020). Stimulation-induced side effects after deep brain stimulation – a systematic review. *Acta Neuropsychiatrica*, *32*(2), 57–64. https://doi.org/10.1017/neu.2019.35