

Harnessing Machine Learning to Decode LFPs for Applications in Adaptive Deep Brain Stimulation



Fernando U Rodriguez Plazas

Doctoral Thesis

Nuffield Department of Clinical Neurosciences

MRC Brain Network Dynamics Unit

University of Oxford & Wolfson College

January 2025

I. Declaration

I, Fernando U R Plazas, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in this thesis.

Signed:

A handwritten signature in black ink, appearing to read "Fernando R. Plazas", with a large, stylized flourish underneath.

Date: January 16th, 2025

II. Abstract

Deep brain stimulation (DBS) has advanced the treatment of movement disorders, yet its conventional, open-loop paradigm can struggle to adapt to the fluctuating clinical and behavioural states that define conditions such as Essential Tremor, Parkinson's disease, and Tourette syndrome. This thesis refines and extends adaptive DBS (aDBS) strategies by integrating machine learning (ML) algorithms and neural signal processing, with the aim of contributing towards providing more precise, individualised stimulation.

We begin by highlighting the advantages of deep learning architectures, notably convolutional neural networks (CNNs), for identifying oscillatory patterns in both synthetic and patient-derived local field potentials. By moving beyond narrow single-band control logic, these networks discover spatiotemporal dynamics that extend beyond canonical beta-band power. Despite outperforming traditional feature-based methods in various settings, CNNs continue to face hurdles due to variability in oscillatory patterns across patient cohorts, and the susceptibility of these architectures to artefacts that commonly appear in DBS recordings and impair model generalisation.

Subsequent chapters examine broader neural correlates that may enhance or guide aDBS. Subthalamic recordings in Parkinson's disease reveal distinct spectral signatures for speech versus limb movement—insights that could help alleviate speech-related side effects through tailored stimulation regimes. Shifting focus to Tourette syndrome, the thesis investigates fronto-thalamic connectivity for real-time prediction of upcoming tics. Although alpha-band synchrony between the thalamus and frontal cortex shows promise at a population level, individual-trial classification proves unreliable, illustrating the challenge of translating group-level phenomena into robust real-time decoders.

Building on these findings, we address the practical challenges of deploying CNN-driven models in adaptive stimulation contexts. We show that abrupt changes in amplitude or frequency can destabilise network activations, jeopardizing the stability of the closed loop. To address this, we design a custom time-domain filter to suppress stimulation artefacts while preserving key physiological signals, thereby stabilizing CNN feature extraction and averting feedback loops that otherwise undermine performance.

Collectively, these projects demonstrate the capacity of data-driven aDBS to move beyond a static, one-size-fits-all approach by harnessing enriched oscillatory and connectivity markers. At the same

time, they underscore the necessity of rigorous artefact management, extensive longitudinal datasets, and model interpretability. By merging neural recordings with computational frameworks, this thesis lays groundwork for adaptive DBS systems that fine-tune therapy in real time, adjusting to each patient's evolving neural and clinical profile. Despite the promise of CNNs and connectivity metrics for more tailored neuromodulation, the demands of real-world datasets and model explainability remain central challenges. Nonetheless, these studies affirm that combining advanced data-driven models with careful artefact handling can shape the next generation of aDBS, making stimulation both more precise and adaptable to individual patient's therapeutic needs.

III. Table of Contents

I.	Declaration	i
II.	Abstract	ii
III.	Table of Contents	iv
IV.	List of Figures	vi
V.	List of Supplementary Figures	ix
VI.	List of Tables	x
VII.	Acknowledgements	xi
VIII.	Abbreviations	xii
IX.	Publications Incorporated into This Thesis	xiv
1	General Introduction	1
1.1	Deep Brain Stimulation: A Historical Perspective	2
1.2	Modern DBS: Towards Physiology-Based, Adaptive, and Data-Driven Neuromodulation	7
1.3	Thesis Aims and Roadmap	13
2	General Methods	17
2.1	Ethics	17
2.2	Externalised Electrophysiological Recordings	17
2.3	Training Machine Learning Systems	22
3	Project 1: Exploring the Use of Convolutional Neural Networks for aDBS Applications	33
3.1	Introduction	33
3.2	Methods	40
3.3	Results	53
3.4	Discussion	59
3.5	Conclusion	70
3.6	Summarised Findings	71
4	Project 2: Anticipating Movement Initiation by Decoding Thalamic LFP Signals in Tremor Patients	73
4.1	Introduction	73
4.2	Methods	75
4.3	Results	87
4.4	Discussion	95

4.5	Conclusion	108
4.6	Summarised Findings	109
5	Project 3: Distinguishing Speech from Upper Limb Movements from Subthalamic and Cortical Neural Signals in Parkinson's Disease.....	110
5.1	Introduction	110
5.2	Methods.....	114
5.3	Results.....	123
5.4	Discussion.....	132
5.5	Conclusion	138
5.6	Summarised Findings	139
6	Project 4: Feasibility of Fronto-Thalamic Connectivity-Based Predictive Biomarkers for Tourette Syndrome Tics	141
6.1	Introduction	141
6.2	Methods.....	145
6.3	Results.....	151
6.4	Discussion.....	160
6.5	Conclusion	165
6.6	Summarised Findings	166
7	Project 5: Implementing and Debugging CNN-driven aDBS.....	168
7.1	Introduction	168
7.2	Methods.....	170
7.3	Results.....	180
7.4	Discussion.....	188
7.5	Conclusion	191
7.6	Summarised Findings	192
8	General Discussion.....	193
8.1	General Overview of Results	193
8.2	Machine Learning as a Tool to Study Electrophysiological Data.....	196
8.3	Machine Learning as a Tool to Elucidate and Validate Biomarkers for aDBS	198
8.4	Thesis Limitations	200
8.5	Outlook	203
8.6	Conclusion	206
9	References.....	207
	Supplementary Materials	277

IV. List of Figures

Figure 1.1: From open-loop, continuous Deep Brain Stimulation to closed-loop, adaptive Deep Brain Stimulation.	8
Figure 3.1: Schema of a signal processing pipeline for adaptive Deep Brain Stimulation (aDBS) or Brain-Computer Interface (BCI) applications.	34
Figure 3.2: Feature-based Approach for Neural Decoding.	35
Figure 3.3: End-to-End Deep Learning Approach for Neural Decoding.	36
Figure 3.4: Simulated neural states were modelled as a first-order Markov process	40
Figure 3.5: Oscillations in Beta and Gamma Bands Characterizing Two Simulated States	42
Figure 3.6: Sinusoidal and Sharp Beta Waveforms Distinguishing Two States.....	43
Figure 3.7: Oscillations with Nonlinear Phase Characterizing Two States	43
Figure 3.8: Theta-Gamma Phase-Amplitude Coupling in Two States.....	45
Figure 3.9: Cross-Channel Phase Shift Characterizing Two States.....	46
Figure 3.10: Beta Burst Length Characterizing Two States	46
Figure 3.11: Feature-based models SVM and AR-SVM.....	49
Figure 3.12: Deep learning-based models 1D-CNN, 2D-CNN, and AR-1D-CNN	51
Figure 3.13: Out-of-Sample Performance Across Signal-to-Noise Ratios	53
Figure 3.14: Pairwise Model Performance Comparison.....	57
Figure 3.15: Relative Weight of Manually Extracted vs. Learned Convolutional Features.....	61
Figure 3.16: 1D-CNN Filters and Activations.....	64
Figure 3.17: Effect of Time Windows on Decoding Performance.....	66

Figure 3.18: Estimated Inference-Time Computational Cost.....	68
Figure 4.1: Feature Extraction Pipeline Used for Decoding Pre-Movement Periods from Electrophysiological Signals	82
Figure 4.2: Data Labelling.....	83
Figure 4.3: CNN Architecture.....	85
Figure 4.4: Feat-CNN Architecture.....	86
Figure 4.5: Decoding Performance of Different Models in Detecting Pre-Movement Periods from Neural Data	88
Figure 4.6: Cohort-Level Time-Frequency Decompositions	91
Figure 4.7: Cross-Patient Variability of Movement-Related Modulatory Patterns	92
Figure 4.8: Correlations Between Modulation Indices in Canonical Frequency Bands and Decoding Performance.....	93
Figure 4.9: Decoding Driven by Individual Frequency Bands.....	94
Figure 4.10: Cross-Patient Training and Evaluation.....	95
Figure 4.11: Linear Coupling Between Models Operating on Different Frequency Bands	98
Figure 4.12: Spectral-Only vs. All-Features AUC Values	101
Figure 4.13: Cross-Patient Comparison of AUC Values for the CNN and FeatCNN Models	102
Figure 4.14: Analysis of the Similarity Between the Decoding Performance Yielded by Different Models	104
Figure 5.1: Experimental Paradigm and Lead Localization.....	119
Figure 5.2: Movement-related spectral and functional connectivity modulations in the STN	126
Figure 5.3: Lateralization of speech and upper-limb movement in the STN.....	128

Figure 5.4: Spatial distribution of preferred band-specific oscillatory activity within the STN that distinguishes between rest and task execution	129
Figure 5.5: Theta/alpha - gamma Phase Amplitude Couplings (PAC)	131
Figure 5.6: ROC-AUC Values for Distinguishing Gripping from Speech Periods	132
Figure 6.1: Signal Power Modulation due to Movement-Related Artefacts During Tic Execution.....	151
Figure 6.2: Connectivity-Based Classification of Pre-Tic Windows	153
Figure 6.3: Effect of Additional Frequency Bands and Feature Sets on Classification Performance...	156
Figure 6.4: Time-Shifted Classification Analysis Across the Entire Tic Epoch	158
Figure 6.5: Cross-trial Consistency of Connectivity and Power-based Features	159
Figure 6.6: Sensitivity Comparison Between Trial-Averaged Statistical Tests and Cross-Validated Machine Learning Approaches	163
Figure 7.1: Real-Time CNN-driven aDBS Pipeline.....	173
Figure 7.2: Convolutional Neural Network Activations During Dynamic Range Changes.....	181
Figure 7.3: Effect of the Custom Time-Domain Filter on LFP Signals and their Distribution	184
Figure 7.4 Effect of the Custom Time-Domain Filter on CNN Feature Stability	186
Figure 7.5: Scaling of Spectral Features Across Stimulation Conditions	187

V. List of Supplementary Figures

Supplementary Figure I: Conceptualization of Convolutional Filters as Finite Impulse Response (FIR) Filters.	277
Supplementary Figure II: Effect of Cross-Patient Network Pre-Training on Decoding Performance..	278
Supplementary Figure III: Model Distillation for Neural Decoding.....	280
Supplementary Figure IV: Pairwise Plots of AUCs Values Yielded by Different Feature Sets	281
Supplementary Figure V: Movement-related Spectral and Functional Connectivity Modulation in the Motor Cortex.....	282
Supplementary Figure VI: Spatial distribution of AUC values for speech vs. rest using spectral feature in all or individual frequency bands.....	283
Supplementary Figure VII: Spatial distribution of AUC values for upper limb movement vs. rest using spectral feature in all or individual frequency bands.....	284
Supplementary Figure VIII: Spatial distribution of AUC values for speech vs. upper limb movement using spectral feature in all or individual frequency bands.....	285
Supplementary Figure IX: Feature Progression (alpha band and broadband) for P1.....	286
Supplementary Figure X: Feature Progression (alpha band and broadband) for P2.....	286
Supplementary Figure XI: Feature Progression (alpha band and broadband) for P3.....	287
Supplementary Figure XII: Feature Progression (alpha band and broadband) for P4.	287
Supplementary Figure XIII: Feature Progression (alpha band and broadband) for P5.....	288
Supplementary Figure XIV: Feature Progression (alpha band and broadband) for P6.....	288
Supplementary Figure XV: Feature Progression (alpha band and broadband) for P7.	289
Supplementary Figure XVI: Feature Progression (alpha band and broadband) for P8.....	289

Supplementary Figure XVII: Custom Signal Labelling Software.....	290
Supplementary Figure XVIII: Software Program for real-time CNN-driven aDBS	291

VI. List of Tables

Table 2-1: Comparative Aspects of Feature-Based vs. End-to-End Approaches.....	26
Table 3-1: Average model performance (balanced accuracy score) across SNRs.....	55
Table 3-2: Average Model Performances on LFP-Based Movement Decoding	57
Table 4-1: Clinical Details of Patients	75
Table 4-2: Decoding Performance During the Pre-Movement Period	89
Table 4-3: Benchmarked Time [in milliseconds] Required for an Inference-Time Prediction.....	105
Table 5-1: Demographic and Clinical Characteristics of Patients.....	115
Table 6-1: Root Mean Square (RMS) Signal Power During Tic Execution	152
Table 7-1: Effect of the Custom Time-Domain Filter on Signal Power Across Different Frequency Bands.	185

VII. Acknowledgements

I extend my gratitude to my supervisors, particularly to Prof. Huiling Tan and Dr. Shenghong He, who have guided me with steadfast support and patience throughout the ups and downs of this thesis. Their thoughtful feedback and gentle direction have kept me on course. I also wish to thank my lab colleagues—especially Dr. Thomas Simpson, Dr. Laura Wehmeyer, and Dr. Sahar Yassine—whose help and generosity has shaped every stage of this research, from conceptualising each project and conducting demanding data collection sessions, to tackling the difficulties of analysing neural timeseries data. Beyond that, their positivity, good humour, and encouragement has helped me navigate the challenging moments of my DPhil.

I am likewise grateful for the supportive research environment at the MRC Brain Network Dynamics Unit (MRC BNDU), which fostered the kind of open-minded science that made this thesis possible, and to the Medical Research Council, for providing the financial support that has enabled me to pursue this research.

Foremost, none of this work would have been possible without the generosity of the many patients who, in the name of advancing scientific knowledge and the care of future patients, volunteered for exacting studies during an already challenging period. Their willingness to extend their hospital stays to undergo lengthy recording sessions after highly invasive neurosurgical procedures, often with medication reduced or withdrawn, stands as a testament to their altruism and commitment to find better ways to manage the neurological disorders that ail them. Many sacrificed personal time and comfort so that we could push our research forward, and for that, I owe them an enduring debt of gratitude.

Finally, I would like to thank everyone who has stood by me during these years: my remarkable family, whose unwavering support I am so fortunate to have; and the many friends I've made at Oxford, especially at Wolfson College. Your patience and kindness have sustained me and have infused these years with meaning.

VIII. Abbreviations

aDBS	Adaptive Deep Brain Stimulation
ANOVA	Analysis of Variance
AR	Autoregressive
AUC	Area Under the Curve
BCI	Brain-Computer Interface
BG	Basal Ganglia
CBGTC	Cortico-Basal Ganglia-Thalamo-Cortical
cDBS	Continuous Deep Brain Stimulation
CM	Centromedian
CNN	Convolutional Neural Network
CV	Cross-Validation
DBS	Deep Brain Stimulation
ECoG	Electrocorticography
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ERD	Event-Related Desynchronisation
ERS	Event-Related Synchronisation
ET	Essential Tremor
FDA	Food and Drug Administration
FDR	False Discovery Rate
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FOG	Freezing of Gait
FPR	False Positive Rate
GBDT	Gradient-Boosted Decision Tree
GPI	Globus Pallidus Interna
GPU	Graphical Processing Unit
HFO	High-Frequency Oscillations
HFS	High-Frequency Stimulation
HPC	High-Performance Computing
iCoh	Imaginary Coherence
IIR	Infinite Impulse Response
IPG	Implantable Pulse Generator
JSON	JavaScript Object Notation

KL	Kullback-Leibler
L-DOPA	Levodopa
LFP	Local Field Potential
LOTO	Leave-One-Trial-Out
LR	Logistic Regression
LSTM	Long Short-Term Memory
MEG	Magnetoencephalography
ML	Machine Learning
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
OCD	Obsessive-Compulsive Disorder
OT	Orthostatic Tremor
PAC	Phase-Amplitude Coupling
PD	Parkinson's Disease
PET	Positron Emission Tomography
PF	Parafascicular
PLV	Phase Locking Value
PSI	Phasy Synchrony Index
RMS	Root Mean Squire
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SCC	Subcallosal Cingulate
SD	Standard Deviation
SELU	Scaled Exponential Linear Unit
SID	Stimulation-Induced Dysarthria
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
TH	Thalamus
TPR	True Positive Rate
TRD	Treatment-Resistant Depression
TS	Tourette Syndrome
UPDRS	Unified Parkinson's Disease Rating Scale
VA	Ventral Anterior
ViM	Ventral Intermediate Nucleus
VL	Ventral Lateral
Voi	Ventrooralis Internus
ZI	Zona Incerta

IX. Publications Incorporated into This Thesis

F. Rodriguez, S. He, and H. Tan, ‘The potential of convolutional neural networks for identifying neural states based on electrophysiological signals: experiments on synthetic and real patient data’, *Front. Hum. Neurosci.*, vol. 17, p. 1134599, Jun. 2023, doi: [10.3389/fnhum.2023.1134599](https://doi.org/10.3389/fnhum.2023.1134599).

Other Publications During the PhD:

Rodriguez, F., Krauss, P., Kluckert, J., Ryser, F., Stieglitz, L., Baumann, C., Gassert, R., Imbach, L., & Bichsel, O. (2024). Continuous and Unconstrained Tremor Monitoring in Parkinson’s Disease Using Supervised Machine Learning and Wearable Sensors. *Parkinson’s Disease*, 2024, 1–11.

<https://doi.org/10.1155/2024/5787563>

He, S., West, T. O., **Plazas, F. R.**, Wehmeyer, L., Pogosyan, A., Deli, A., Wiest, C., Herz, D. M., Simpson, T., Andrade, P., Baig, F., Hart, M. G., Morgante, F., FitzGerald, J. J., Barbe, M. T., Visser-Vandewalle, V., Green, A. L., Pereira, E. A., Cagnan, H., & Tan, H. (2024). Cortico-thalamic tremor circuits and their associations with deep brain stimulation effects in essential tremor. *Brain*, awae387. <https://doi.org/10.1093/brain/awae387>

Wiest, C., Simpson, T. G., Pogosyan, A., Hasegawa, H., He, S., **Plazas, F. R.**, Wehmeyer, L., Yassine, S., Guo, X., Shah, R., Merla, A., Perera, A., Raslan, A., O’Keeffe, A., Hart, M. G., Morgante, F., Pereira, E. A., Ashkan, K., & Tan, H. (2024). Stimulation-Evoked Resonant Neural Activity in the Subthalamic Nucleus Is Modulated by Sleep. *Movement Disorders*, mds.30063.

<https://doi.org/10.1002/mds.30063>

Simpson, T. G., He, S., Pogosyan, A., **Plazas, F. R.**, Hart, M. G., Shah, R., Hasegawa, H., Wiest, C., Wehmeyer, L., Yassine, S., Guo, X., Merla, A., Perera, A., Raslan, A., O’Keeffe, A., Welter, M.-L., Ashkan, K., Morgante, F., Andrade, P., ... Tan, H. (2025). *The role of STN beta oscillations on lower extremity muscle activity in Parkinsonian stepping*. Neuroscience.

<https://doi.org/10.1101/2025.01.01.631000>

Wehmeyer, L., Baldermann, J. C., Pogosyan, A., **Plazas, F. R.**, Loehrer, P. A., Bonetti, L., Yassine, S., Mühlen, K. Z., Schüller, T., Kuhn, J., Visser-Vandewalle, V., Tan, H., & Andrade, P. (2024).

Thalamo-frontal connectivity patterns in Tourette Syndrome: Insights from combined intracranial DBS and EEG recordings. <https://doi.org/10.1101/2024.10.09.617413>

1 General Introduction

Deep brain stimulation (DBS) has evolved from a lesion-based experimental procedure to a cornerstone therapy for Parkinson’s disease, essential tremor, and other neurological disorders. Initially conceived as a reversible alternative to inherently irreversible interventions based on ablation of neural tissue, DBS relies on the notion that many conditions stem from dysfunctional brain networks that can be functionally modulated, rather than permanently destroyed. Traditionally, clinicians have delivered continuous stimulation based on stereotactic targeting and clinical heuristics, an approach that is highly effective, yet remains limited by its inability to accommodate real-time changes in patient and brain states.

Recent advances in implantable pulse generators and sensor technologies are reshaping DBS into a more adaptive and patient-specific platform. Newer systems can record local field potentials (LFPs) during stimulation, thus enabling real-time “closed-loop” or “adaptive DBS” (aDBS)¹. In these paradigms, neural signals—commonly referred to as neural “biomarkers” that reflect changes in disease or behavioural states—provide moment-to-moment feedback guiding when, where, how, and how much stimulation is delivered. Such data-driven methods have already shown promise in enhancing therapy while reducing side effects and battery replacement frequencies. Nevertheless, significant challenges persist, including ensuring robust signal acquisition over months or years, developing efficient machine learning algorithms that run on low-power hardware, and addressing ethical and regulatory hurdles tied to chronic brain implants. This thesis aims to contribute to these efforts by developing and evaluating signal processing and machine learning techniques for the robust detection of clinically relevant neural states in adaptive deep brain stimulation (aDBS).

In the next section of this introduction, we chart the historical development of DBS from its early lesion-based origins through its widespread clinical adoption, discuss the transition toward physiology-based neuromodulation, and highlight the growing momentum behind adaptive, data-driven therapies. These sections set the stage for the central aims of this thesis—specifically, the development and

¹ In this thesis, we use the terms *closed-loop* DBS (cl-DBS) and *adaptive* DBS (aDBS) interchangeably.

evaluation of advanced signal processing and machine learning techniques for robust, individualised aDBS.

1.1 Deep Brain Stimulation: A Historical Perspective

1.1.1 The Origins of DBS

The concept of electrically modulating the nervous system dates back to antiquity, when electric eels or rays (e.g., *Torpedo nobiliana*) were sometimes placed on the scalp as a crude method for treating pain or headache (Henderson, 2008; Sironi, 2011). In his first-century work *Compositiones medicamentorum*, Scribonius Largus, who served as physician to the Roman Emperor Claudius, recommended that the torpedo fish be applied to the forehead to relieve headaches (Debru, 2006). These early practices, albeit in a primitive form, established a historical thread of using “neurostimulation” for disorders of the brain and body.

By the eighteenth and nineteenth centuries, systematic experimentation on the effects of electrical current in animal and human tissues began to shape modern neurophysiology. Giovanni Aldini (1762–1834), Luigi Galvani’s nephew, famously applied electrical stimulation to the exposed cortex of recently decapitated criminals, observing muscle twitches and “horrible facial grimaces” (Aldini, 1804; Parent, 2004). Although ethically shocking by modern standards, Aldini’s demonstrations proved that the human cortex was electrically excitable. Subsequent experiments by Luigi Rolando (1773–1831) in 1809, and by David Ferrier (1843–1924) and Robert Bartholow (1831–1904) later in the century, refined these insights by systematically stimulating the exposed cortex of animals and, in Bartholow’s controversial case, an awake patient (Bartholow, 1874; Ferrier, 1873; Patra et al., 2019; Rolando, 1809; Sironi, 2011). The overarching lesson, emergent from all these efforts, was that the brain could indeed be influenced by external electrical currents—and that local stimulation could produce focal effects on movement or sensation (Fritsch, 1870; Gross, 2007; Morgan, 1982).

Beginning in the late nineteenth century, functional neurosurgery arose when Victor Horsley (1857–1916) used surgical interventions, often involving ablation, for treating neurological disorders such as chorea (Horsley, 1890; Vilensky & Gilman, 2002). Horsley’s procedures typically removed portions of cortex to alleviate debilitating involuntary movements (albeit sometimes at the price of inducing paralysis). While these ablative techniques set an early precedent for surgical approaches to movement

disorders, they did not yet employ the concept of “stimulation.” Similarly, early attempts to treat psychiatric illness by surgically modifying the frontal lobes—an approach that would later evolve into the field of “psychosurgery”—also began in this era, eventually culminating in the awarding of the 1949 Nobel Prize in Physiology or Medicine to Portuguese neurologist António Egas Moniz (1874–1955) for his work pioneering the frontal leucotomy (Moniz, 1937; Preul & Dagi, 2017).

Moniz’s procedure is often described (apocryphally, it turns out) as an abrupt inspiration that took place at the 1935 Second International Congress of Neurology in London, where John Fulton and Carlyle Jacobsen presented behavioural changes in chimpanzees—namely more cooperativeness and willingness to accomplish tasks—observed after undergoing frontal lobe ablations (Boettcher & Menacho, 2017). Although Fulton and others disputed this simplified narrative, and the notion that Moniz immediately resolved to operate on the human frontal lobe after being presented with the chimp cases, there is little doubt that a convergence of ideas about newly emerging psychosurgical techniques and the desire to treat otherwise intractable psychiatric illnesses provided the impetus (Mashour et al., 2005). Still, Moniz’s leucotomy reflected an important conceptual shift in neurosurgery: that targeted interventions (whether ablative or stimulatory) on specific brain areas or pathways could meaningfully alter pathological behaviour (Berrios, 1997). From the first lesional procedures performed in the late 1930s by Moniz, through the peak in the late 1940s and early 1950s, approximately 40 000 to 50 000 psychosurgical operations were performed overall, based on very little clinical evidence. Some 4000 of these psychosurgical procedures were performed or supervised by Freeman, a psychiatrist; and many of these used Freeman’s minimal surgical technique: the transorbital lobotomy (Kringelbach & Aziz, 2009; Valenstein, 1974).

From the 1930s onward, additional attempts to treat chronic pain or psychiatric conditions by electrically activating deep brain structures set the stage for modern DBS. For instance, in the 1950s and 1960s, José Delgado (1915–2011) famously implanted radio-equipped so-called “stimocoeivers” in animals and in some human subjects, exploring how electrical impulses applied to the basal ganglia or limbic structures could suppress aggression or modify mood (Delgado, 1969; Horgan, 2005; Sironi, 2011). Although many in the scientific community regarded Delgado’s demonstrations (such as halting a charging bull) as mostly theatrical proofs-of-concept, these demonstrations did show that focal activation of subcortical sites could indeed produce significant behavioural effects. Natalia Petrovna Bekthereva in the Soviet Union and Carl Wilhelm Sem-Jacobsen in Norway, similarly and independently investigating the potential of chronically implanted deep electrodes, reported that “therapeutic electrostimulation” could mitigate hyperkinetic disorders and psychiatric conditions,

though language barriers, limitations in technology, and ethical uncertainties hindered the widespread recognition and adoption of these early developments (Bekhtereva et al., 1963, 1975; Blomstedt & Hariz, 2010; Sem-Jacobsen, 2009).

In parallel, a separate line of inquiry culminated in the introduction of stereotactic neurosurgery by Ernst Spiegel and Henry Wycis in 1947, facilitating more precise approaches to deep-brain targets (Gildenberg, 2005; Spiegel et al., 1947). Their stereotactic frame not only enabled targeted lesions in the thalamus or basal ganglia for movement disorders but also allowed for the placement of chronic stimulating electrodes into deep subcortical nuclei. During intraoperative testing, low-frequency stimulation was soon recognised to exacerbate tremors, whereas high-frequency stimulation (50–100 Hz) consistently suppressed them (Gildenberg, 2005; Schwalb & Hamani, 2008). The gradual shift from ablative procedures (e.g., pallidotomy, thalamotomy) to prolonged intraoperative stimulation, used for target location prior to ablation, gave rise to the field that would eventually become modern DBS (Blomstedt & Hariz, 2010).

Thus, while the notion of using electricity to influence brain function stretches all the way back to antiquity, it was not until the mid-twentieth century that adequate surgical precision, physiological monitoring, and clinical need coalesced to support the development of modern DBS. An important factor in this process was the evolving view that many complex neurological or psychiatric syndromes were not the result of fixed local lesions alone, but rather of dysfunctional neural circuits, i.e. dynamic networks amenable to modulation rather than to tissue destruction alone. This represented a significant conceptual leap: that modulating pathological oscillations or signals might be as useful as ablating tissue in treating conditions like Parkinson’s disease, essential tremor, or pain (Lozano & Lipsman, 2013).

1.1.2 Safety, Efficacy, and the Widespread Adoption of DBS

By the 1960s and 1970s, improvements in imaging, anaesthesia, and operative safety allowed more frequent employment of deep-brain surgical interventions for movement disorders, psychiatric disease, and intractable pain. Initially, the standard approach was to create small, precise lesions (e.g., pallidotomy or thalamotomy) in structures implicated in tremor disorders or Parkinson’s disease (Bergman et al., 1990; Schwalb & Lozano, 2004). These ablative methods could relieve tremor, rigidity, or chorea in certain patients, but came at the cost of permanent destruction of tissue,

sometimes leading to adverse effects if the lesion was oversized or off-target (Gildenberg, 2006; Schwab & Hamani, 2008).

A transition came when neurosurgeons noticed during stereotactic lesioning procedures that high-frequency stimulation of target nuclei—used initially only as a temporary way to confirm or refine lead placement—could itself suppress pathological motor signs in real time, prior to any tissue destruction. In essential tremor or Parkinson’s tremor, for example, stimulating the ventral intermediate nucleus (ViM) of the thalamus briefly removed tremor so effectively that some surgeons realised the same effect as a thalamotomy could be achieved merely by keeping the stimulator on, obviating the need to burn any tissue (Albe-Fessard & Fessard, 1963; Gildenberg, 2005; Pycroft et al., 2018; Tóth & Tomka, 1968). In 1987, a milestone paper by Alim Louis Benabid and colleagues reported that continuous, high-frequency stimulation of the ViM suppressed tremor in patients with Parkinson’s disease and essential tremor (Benabid et al., 1987). This “reversible lesion” concept drew interest and catalysed a broader shift toward exploring chronic DBS to substitute ablative surgery. Implantable stimulators for the spinal cord had been introduced in the 1960s, paving the way for other chronically implanted neurostimulation devices (Pycroft et al., 2018). The convergence of technology and serendipitous clinical observations eventually lead to the study and establishment of pallidal DBS for PD (Siegfried & Lippitz, 1994), and then to STN-DBS (Limousin et al., 1995; Pollak et al., 1993).

Over the next decade, clinical trials firmly established DBS as a durable therapy for movement disorders. In 1997, the U.S. Food and Drug Administration (FDA) approved thalamic DBS for essential tremor and Parkinsonian tremor (Hubble et al., 1996; Koller et al., 1997). In 2002, subthalamic nucleus (STN) DBS and globus pallidus internus (GPi) DBS both gained approval for the treatment of advanced Parkinson’s disease (Follett et al., 2010; Limousin et al., 1998), leading to a rapid expansion of clinical interest. Many patients had been left “off medication” for significant parts of the day due to the issues associated with excessive, long-term dopaminergic therapy and were therefore bradykinetic or rigid. They could now, through round-the-clock stimulation, experience more continuous symptom relief. Other patients benefitted from diminished medication-induced dyskinesias (Castrioto, 2011; Lyons & Pahwa, 2008; Weaver, 2009). Similar outcomes were also documented for primary dystonia, which obtained an FDA Humanitarian Device Exemption in 2003 following studies showing marked improvements in motor scores under pallidal DBS (Isaias et al., 2009; Vidailhet et al., 2005), and for OCD, where FDA approval was granted under humanitarian device exemption in 2009.

The growing success of DBS therapy in movement disorders led researchers to ask whether the same conceptual shift—i.e. managing symptoms by modulating activity in dysfunctional circuits, rather than by creating irreversible ablations—could prove similarly useful in other severe or refractory conditions. In 2009, the FDA extended Humanitarian Device Exemption to deep brain stimulation of the anterior limb of the internal capsule for obsessive-compulsive disorder, responding to open-label, controlled trials suggesting that high-frequency stimulation could lessen compulsions and anxiety (Denys et al., 2010; Greenberg et al., 2006; Mallet et al., 2008). Meanwhile, several studies tested DBS in a number of novel disease areas including Tourette’s syndrome (Baldermann et al., 2016; Vandewalle et al., 1999; Visser-Vandewalle et al., 2003), treatment-resistant depression (Lozano et al., 2008; Mayberg et al., 2005), chronic pain (Kumar et al., 1997), and Alzheimer’s disease (Laxton et al., 2010). Although many of these studies remain exploratory and have not always shown consistent therapeutic benefit, the overall trajectory is clear: a therapy once designed purely for “movement” disorders has evolved into a neuromodulatory paradigm potentially able to restore function across a wide range of so-called “circuitopathies” (Kringelbach et al., 2010; Lozano & Lipsman, 2013; W.-J. Neumann, Horn, et al., 2023).

In the field of DBS, technological development has gone hand in hand with the increasing adoption and improvement of the therapy. Significant developments include the introduction of “directional” or “segmented” electrodes, which allow to more precisely steer current to different areas of the target structure after lead implantation. This increased precision has been shown to increase the therapeutic efficacy by increasing side-effect thresholds (Dembek et al., 2017), albeit potentially at the cost of more onerous programming due to the increased degrees of freedom (Tinkhauser, Pogosyan, et al., 2018). More advanced multi-variate statistical tools, by incorporating data-driven insights, might be able to streamline this process (A. Shah et al., 2023).

The technical foundations of DBS continue to advance today. Neuroimaging tools have refined targeting while enabling surgeons to approach deeper or more challenging structures, such as the posterior subthalamic area for tremor or pedunculopontine nucleus for gait disturbances (Stefani et al., 2007). Data-driven insights into the individualised brain networks that contribute to patient-specific manifestations of neuropsychiatric disease have enabled targeted, symptom-specific modulation of symptomatic networks to improve clinical success in novel areas (Sheth et al., 2022). Real-time physiological feedback during surgery, plus novel computational modelling of electric fields, has enhanced precision and stimulated new research into adaptive and closed-loop DBS devices (Little & Brown, 2012; McIntyre et al., 2004). With more than an estimated 263’000 patients now implanted

worldwide (Johnson et al., 2024), the wide adoption of DBS as both a powerful therapy and an investigational tool attests to a remarkable transformation from ablative neurosurgery to modern, adjustable, and fine-tuned neuromodulation.

Thus, in the span of roughly 50 years, the surgical approach to chronic neurological and psychiatric illness has gone from frontal lobotomies and coarse lesioning to sophisticated, MRI-guided, physiology-based modulation of targeted networks and pathways in the human brain. This journey is fuelled both by clinical need and by a growing understanding that many disorders of the brain—Parkinson’s disease, essential tremor, dystonia, depression, Tourette’s syndrome—are rooted in circuit-level pathophysiology that can be modulated by stimulation (Lozano & Lipsman, 2013).

1.2 Modern DBS: Towards Physiology-Based, Adaptive, and Data-Driven Neuromodulation

Deep brain stimulation (DBS) has evolved from a lesion-inspired, open-loop approach to a neuromodulatory paradigm that continuously modulates activity in neural circuits in real time. Historically, DBS practices were guided by a combination of stereotactic coordinates, lesion-based neurosurgical precedents, and heuristic clinical programming (Blomstedt & Hariz, 2010; Gildenberg, 2005). These methods proved efficacious for treating a subset of movement disorders, but many patients continue to experience non-motor complications, stimulation-induced side effects (e.g., dysarthria, dyskinesias), or suboptimal benefit (Beudel & Brown, 2016; Castrioto, 2011; C. C. Chen, Brücke, et al., 2006; Fasano et al., 2010). Recent technological and theoretical advances are now propelling DBS into a new era—one in which stimulation can be tailored in real time, informed by pathophysiological signals, advanced signal processing, and machine learning (Lozano et al., 2019; Marceglia et al., 2021; W.-J. Neumann, Gilron, et al., 2023; S. Wang et al., 2023).

1.2.1 From Continuous to Adaptive Stimulation

Classical DBS (continuous DBS; cDBS) systems deliver a preset amplitude, frequency, and pulse width continuously, irrespective of fluctuations in the patient’s clinical, behavioural, or medication state. The neurostimulator acts in an open-loop fashion and cannot sense the changing neural signals

in the target networks that it is modulating (Arlotti et al., 2018; Little et al., 2013). Although effective in many cases many, this unchanging mode of stimulation can introduce problems: stimulation remains “always on,” leading to elevated power consumption, repeated battery replacements, potential side effects during periods when high-amplitude stimulation is not necessary, and insufficient control during episodes when higher stimulation might be temporarily warranted (Little & Brown, 2012; Gilron et al., 2021).

By contrast, adaptive DBS (aDBS) harnesses real-time data streams—such as local field potentials (LFPs) (He et al., 2021), cortical electrocorticography (ECoG) (Castaño-Candamil et al., 2020; Opri et al., 2020), or wearable-sensors (Cernera et al., 2021)—and modifies stimulation automatically according to biomarker thresholds and control algorithms (Arlotti et al., 2018; Bronte-Stewart et al., 2020; Little et al., 2013). The fundamental premise is that pathological brain rhythms—oftentimes exemplified by exaggerated beta-band oscillations (13–35 Hz) in Parkinson’s disease (PD)—correlate with symptom severity (Kühn et al., 2006). When the amplitude of these oscillations surpasses a specified threshold, stimulation should intensify to counteract the concomitant symptoms; conversely, when oscillations subside, stimulation should decrease, targeting pathological states more precisely (Little et al., 2013; Rosa, Arlotti, et al., 2017). Early pilot trials using such threshold-based or proportional control have shown improvements in “on time”, reduced dyskinesias, and lower energy consumption (W.-J. Neumann, Gilron, et al., 2023; Velisar et al., 2019). A similar approach to aDBS feedback relies on decoding behavioural (oftentimes movement) states that may elicit symptoms, and delivering stimulation based on the detection of these states (He et al., 2021; S. A. Shah, Tan, et al., 2018; Tan et al., 2019).

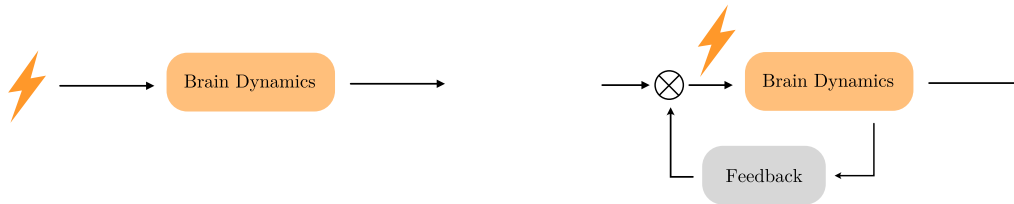


Figure 1.1: From open-loop, continuous Deep Brain Stimulation to closed-loop, adaptive Deep Brain Stimulation.

The success of aDBS depends on several critical steps. First, real-time neural signals must be acquired reliably despite the presence of stimulation artefacts and potential confounds such as motion, electrocardiographic interference, or time-varying neural states (Anidi et al., 2018; W.-J. Neumann et al., 2022; Swann, De Hemptinne, Miocinovic, et al., 2018). Second, robust biomarkers need to be identified or extracted to reflect clinically meaningful transitions in symptom severity (Gilron et al.,

2021; Meidahl et al., 2017). Lastly, fast, adaptive control policies must be implemented in embedded hardware to modulate stimulation with minimal latency (Arlotti et al., 2018; W.-J. Neumann, Gilron, et al., 2023). Despite these complexities, aDBS is increasingly viewed as the future of DBS for movement disorders, with more advanced prototypes now exploring non-motor domains—psychiatry, cognition, addiction—and disorders such as Tourette syndrome (Johnson et al., 2024; Sheth et al., 2022). As will be explored throughout this thesis, developing better sensors, signal processing techniques, and data-driven methods will be important for advancing the clinical efficacy of aDBS.

1.2.2 Sensing, Machine Learning, and Biomarker Discovery

A central theme in any adaptive DBS (aDBS) approach is the need for physiologically meaningful signals—often termed “biomarkers”—that guide stimulation in real time. In Parkinson’s disease (PD), the oscillatory power of subthalamic nucleus (STN) beta activity ($\sim 13\text{--}30$ Hz) has served as the prototype biomarker, as it correlates with bradykinesia and rigidity (Kühn et al., 2006). Researchers have exploited this biomarker by adjusting stimulation intensity in accordance with the magnitude of beta oscillations, effectively mitigating some side effects while retaining clinical benefits compared to standard DBS (Arlotti et al., 2018; Bocci et al., 2021). Although these advances illustrate the utility of closed-loop control, many other prospective biomarkers have emerged to track different pathological states, such as tremor, dyskinesias, or freezing of gait (Bouthour et al., 2019; Bronte-Stewart et al., 2020). The search for diverse and sophisticated signals is motivated by the fact that disorders often arise from distinct and complex circuit anomalies, can fluctuate with medication or sleep, and may therefore not be captured by a single threshold that remains static throughout relevant fluctuations in the state of the patient (Fleming et al., 2022; Meidahl et al., 2017).

Historically, one of the central hurdles for aDBS lay in recording local field potentials (LFPs) in the presence of high-frequency stimulation which, being orders of magnitude higher than neural activity, can cause amplifier saturation and significant artefact, rendering signal processing pipelines to capture biomarkers infeasible (W.-J. Neumann et al., 2022). Several approaches exist to mitigate the effect of the stimulation (and other) artefacts on neural signals (Debarros et al., 2020)—in this thesis, we explore this topic, and contribute to efforts in simultaneous recording and stimulation, in Project 5. Next-generation implants—Medtronic Percept™, Summit™ RC+S, and Newronika AlphaDBS—addressed these limitations through specialised blanking or filtering, enabling simultaneous stimulation and high-quality signal acquisition (Ansó et al., 2022; Jimenez-Shahed, 2021). This shift created what

are effectively “bidirectional” DBS interfaces (Bronte-Stewart et al., 2020), with real-time sensor data feeding into closed-loop algorithms that can modulate stimulation in real time. It also expanded the realm of feature extraction to include, for example, time-domain burst metrics (Anderson et al., 2020; Tinkhauser, Pogosyan, Tan, et al., 2017), cross-frequency coupling (Gilron et al., 2021), as well as other connectivity-based measures (Merk, Peterson, Köhler, et al., 2022). Myriad candidate biomarkers, ranging from Parkinsonian tremor detection (S. A. Shah, Tinkhauser, et al., 2018) to mood and pain decoding (Sani et al., 2018; Shirvalkar et al., 2023), have been explored to guide stimulation—oftentimes using multivariate and complex processing pipelines.

Early implementation of aDBS schemes relied on externalised setups during the postoperative period (Castaño-Candamil et al., 2020; He et al., 2021). These approaches could only be validated during short recording sessions, and questions about their validity—especially through fluctuations in patient state—remained (Little & Brown, 2020). As the opportunities afforded by technology have advanced, an especially promising extension of these paradigms now involves the ability to gather chronic recordings that reflect the full variability of a patient’s daily life—rather than relying exclusively on short-term data captured intraoperatively or in carefully controlled clinical settings. Chronic monitoring has expanded opportunities for biomarker discovery and validation, allowing researchers and clinicians to characterise fluctuations over medication cycles, sleep–wake transitions, or psychosocial contexts and thus identify novel correlates of abnormal motor or psychiatric states (J. Herron, Kullmann, et al., 2024). Combining chronic recordings with machine learning further enriches these systems. These developments align closely with the emerging view that cortical signals can afford richer spectral information and stronger signal-to-noise ratios than deep LFPs, particularly when stimulation artefacts dominate subthalamic or thalamic channels (Merk, Peterson, Lipski, et al., 2022). The superiority of ECoG signals for decoding purposes, when compared to subcortical LFPs, may be explained—amongst other factors—by the more spatially compressed somatotopic organisation of subcortical structures when compared to the cortex. The increased spatial separation of signals in the cortex is better suited for standard electrode arrays (J. Herron, Kullmann, et al., 2024). Studies in essential tremor have documented more accurate movement detection from electrodes over the motor cortices than from leads in the ventral intermediate nucleus, especially during active stimulation (Frączek et al., 2021; Opri et al., 2020), and investigators have started to apply the same approach to mood and obsessive–compulsive phenotypes, targeting orbitofrontal or cingulate circuits as potential sites for capturing higher-amplitude, less artefactual biomarkers (Provenza et al., 2019). By providing objective biomarkers of pathology and guiding the targeting of individualised circuits, biomarker-based

programming and adaptive DBS have the potential to open up new therapeutic possibilities to treat hitherto intractable psychiatric disorders (Widge, 2024).

As next-generation aDBS systems integrate data-driven insights from plentiful longitudinal data, sophisticated machine-learning (ML) algorithms offer a promising route to “intelligent” aDBS (Merk, Peterson, Köhler, et al., 2022). Rather than relying solely on threshold-based detection of one oscillatory band, ML pipelines can incorporate multivariate features derived from invasive neurophysiology—be it STN-LFPs, ECoG signals, or combined cortical and subcortical recordings. The potential payoff more accurate or granular state classification—for instance distinguishing tremor, dyskinesia, or bradykinesia². Distilling these data-driven insights from longitudinally gathered data comes, however, with its own set of challenges. Implementing complex ML solutions in a real-world clinical setting demands careful attention to computational and practical constraints. Features must be computed and processed in real time without overwhelming the implant’s power budget, and cross-validation methods must ensure robust, generalizable performance. The capacity for interpretability is similarly critical, since clinicians must be able to understand or adjust device behaviour if, for instance, stimulation intensifies during inappropriate states. While deep neural networks or ensemble methods may boost accuracy, simpler classifiers—like linear discriminant analysis or random forests—can strike a more practical balance of real-time feasibility and clinical transparency (W. Neumann & Rodriguez-Oroz, 2021; Opri et al., 2020). Some of these trade-offs will be explored throughout the chapters of this thesis.

1.2.3 Toward a Broader Data-Driven Paradigm

Although PD has served as the prototypical case for aDBS, the same conceptual framework can be applied to other movement and neuropsychiatric disorders that are rooted in dysfunctional circuits. For instance, aDBS systems that track cortical signals in the gamma band have been piloted to

² A further application of machine learning (although not explored experimentally in this thesis) is the implementation of schemes to improve real-time adaptability by continuously updating how features are processed and weighted by the ML system in response to subtle changes in the patient’s physiological or behavioural state.

manage dyskinesia (Swann, De Hemptinne, Thompson, et al., 2018). In Tourette syndrome, low-frequency oscillations or fronto-thalamic coherence may serve as prospective biomarkers for tic onset (Cagle et al., 2022; Wehmeyer et al., 2024). Meanwhile, in treatment-resistant depression, investigators are examining whether subcallosal cingulate (SCC) or orbitofrontal LFPs can be used to trigger on-demand stimulation (Scangos et al., 2021; Sheth et al., 2022). Signal processing and machine learning techniques are integral to this data-driven approach. Instead of hand-crafting a single threshold or focusing on a narrow band, ML-based pipelines can assimilate features from across the power spectrum, time-frequency dynamics, or connectivity analyses to develop more robust, individualised biomarker definitions. Combined with tele-programming and cloud analytics, clinicians could systematically optimise aDBS, balancing therapeutic benefit against side effects in real time (W. Chen et al., 2021).

Despite the promise, several obstacles must be resolved. Standardization of artefact-rejection algorithms, clearer “best practices” for extracting stable biomarkers over prolonged periods, and more transparent regulatory frameworks are all pressing. Ethical challenges loom as well: long-term, adaptive modulation of patients’ motor or emotional circuits can pose fundamental questions about identity, autonomy, and unforeseen consequences—particularly if device malfunctions or “brainjacking” attacks disrupt the stimulator’s intended operation (Bouthour et al., 2019; Giordano, 2015). Moreover, at an estimated USD 47,600 per procedure (Bishay et al., 2024), the cost of DBS remains high. This limits access, especially in mental health indications like obsessive-compulsive disorder (OCD), where insurance reimbursement is often denied in spite of ample evidence supporting efficacy and safety (Visser-Vandewalle et al., 2022). Historical bias and scepticism around “psychosurgery” further complicate patient referral and funding, echoing concerns once raised about lobotomy and leucotomy (Kringelbach & Aziz, 2009). In rare conditions, this “crisis of access” is amplified by manufacturer hesitancy to invest in narrow clinical niches, alongside insufficient training opportunities for psychiatrists and psychologists to manage DBS in complex disorders (Visser-Vandewalle et al., 2022). These barriers highlight the need for interdisciplinary advocacy, including rigorous neuroethical oversight, collaborative registries to ensure clinical evidence, and legislative reform to expand insurance coverage and standardise device labelling.

Equally important are broader social and philosophical issues tied to neuromodulation technology. DBS represents a historic success case in translational from basic neurophysiological research to clinical practice (Pycroft et al., 2018) and has therefore relied on extensive animal research—particularly in non-human primates—for exploring hardware and circuit-level insights (Kringelbach et

al., 2010). All of this, however, raises ethical scrutiny over animal welfare and regulatory policies. In addition, the idea of “enhancing” traits rather than merely treating disease can spark debate about medicalization and personal identity, especially if a significant fraction of neurosurgeons would theoretically permit DBS for non-pathological traits (Hariz, 2012). For some patients, the persistent question of “who I am” versus “what the device is making me do” can blur the sense of self and moral agency (Yuste et al., 2017). As artificial intelligence and brain–machine interfaces (BMIs) become more entwined (J. Herron, Kremen, et al., 2024), concerns over data security, privacy, and fair distribution of advanced treatments intensify (Clausen et al., 2017). Altogether, mitigating these ethical and practical hurdles will require a concerted effort by neurologists, neurosurgeons, bioengineers, data scientists, regulatory bodies, and patient advocacy groups.

Modern DBS is transitioning from a uniform, high-frequency stimulation protocol to a dynamic, physiology-based framework in which therapy is delivered on demand according to real-time biomarkers. The combination of next-generation implantable stimulators, robust neural signal processing, and machine learning algorithms opens the possibility of individualised neuromodulation that adapts to changing clinical states. Although robust clinical trial evidence is still forthcoming, the trajectory is clear: harnessing the brain’s own signals to drive precision DBS may yield both improved symptom control and minimised side effects for a wide spectrum of neurological and psychiatric disorders.

1.3 Thesis Aims and Roadmap

Deep brain stimulation (DBS) has progressed from an initially lesion-inspired surgical approach to a sophisticated, physiology-informed therapy for diverse movement and neuropsychiatric disorders that continuously adapts to therapeutic demand. As detailed in the preceding sections, modern DBS increasingly harnesses real-time physiology—particularly local field potentials (LFPs)—to deliver “on-demand” or *adaptive* stimulation. While promising, such adaptive DBS (aDBS) frameworks require reliable biomarkers, robust signal processing pipelines, and principled machine learning (ML) algorithms that can handle real-world variability in patient data and device parameters.

This thesis explores approaches to harness ML and signal processing techniques to decode meaningful brain states from electrophysiological recordings in the context of aDBS. More specifically, its aims are:

- **Establish Data-Driven Strategies for Adaptive DBS**

Rationale: While aDBS has historically relied on single-band (e.g. beta) thresholds or narrowly defined feedback loops, there is growing interest in more flexible, machine learning-driven approaches incorporating more comprehensive brain states.

Aim: Demonstrate how feature-based classifiers and neural networks can detect relevant neural states—both pathological (e.g., tics, dyskinesias) and behavioural (e.g., voluntary movement, speech)—under the constraints of single-trial or real-time decoding, thus guiding stimulation more precisely and potentially mitigating side effects.

- **Evaluate Broader Biomarkers and Frequency Bands:**

Rationale: Beta-band power, although central to Parkinson’s bradykinesia and rigidity, may not capture tremor, dyskinesias, or speech deficits; similarly, local power features alone can miss key connectivity- or phase-based phenomena.

Aim: Investigate a range of candidate biomarkers—spanning, where appropriate based on previous literature, multiple frequency bands, connectivity metrics (e.g., phase synchrony, coherence), and time-domain features—to assess their capacity to track changes in neural circuits relevant for aDBS in tremor, Parkinson’s disease, and Tourette’s syndrome.

- **Elucidate and Harness Neural Correlates of Behaviour and Disease State**

Rationale: Although movement- and speech-related deficits in Parkinson’s disease, Tourette syndrome, dystonia, and other conditions are often clinically distinct, they share a foundation in network-level dysfunctions rather than purely localised oscillatory abnormalities. Characterizing both localised LFP signatures (e.g., for speech vs. movement) and distributed connectivity patterns (e.g., fronto-thalamic coherence) is essential for understanding how pathological circuits underlie motor and non-motor symptoms, and for ensuring that adaptive DBS parameters are attuned to each patient’s unique neural dynamics.

Aim: Develop and apply neural decoding frameworks that integrate local field potential data and network-level connectivity measures across thalamic, subthalamic, and cortical regions. By identifying activity patterns specific to normal behaviours (e.g., speech initiation, limb movement) and to disease states (e.g., tic onset), we aim to enable aDBS strategies capable of distinguishing—and potentially appropriately modulating—these diverse clinical phenomena in real time.

- **Real-Time-Compatible Neural Network Pipelines for Adaptive DBS**

Rationale: Convolutional neural networks (CNNs) and related deep learning methods can effectively capture spatiotemporal patterns in electrophysiological signals, yet practical deployment for adaptive DBS must accommodate stimulation artefacts, dynamic range shifts, and stringent real-time and computational constraints. These challenges necessitate robust artefact mitigation, stable input normalization schemes, and computationally efficient model architectures.

Aim: Investigate the feasibility of neural networks as a real-time decoding approach for aDBS. Implement a comprehensive neural decoding pipeline—addressing artefacts and dynamic range changes, ensuring low-latency inference, and exploring adaptive network architectures—and validate its performance in closed-loop experimental settings. Compare CNN-based methods with traditional feature-based pipelines, clarifying how hybrid solutions or targeted model adaptations can overcome real-world translational hurdles.

These overarching aims guide the chapters of this thesis, each of which delves into specific experimental paradigms, technical developments, and clinical datasets to advance toward aDBS systems that are both physiologically grounded and computationally robust. Below, we outline how the chapters are organised to address these objectives:

- Chapter 2: General Methods. Outlines the common ethical, surgical, and technical foundations underlying all projects in this thesis —covering DBS lead externalization, data acquisition hardware, common preprocessing routines, and the main ML paradigms adopted (feature-based vs. deep learning).
- Chapter 3 (Project 1): Explores how convolutional neural networks can decode distinct oscillatory states from synthetic neural data and from real LFP recordings in tremor patients, benchmarking CNN performance against feature-based classifiers.

- Chapter 4 (Project 2): Investigates the feasibility of detecting and even anticipating upper-limb movements (for intention tremors) via LFP and EEG signals, illustrating the potential for early intervention in aDBS.
- Chapter 5 (Project 3): Examines subthalamic LFP and EEG activity during speech vs. limb movement in Parkinson’s disease, finding group-level modulatory patterns and assessing whether ML methods can segregate these behavioural states and thereby inform speech-sensitive adaptive DBS.
- Chapter 6 (Project 4): Evaluates connectivity-based metrics in Tourette syndrome, probing whether fronto-thalamic coherence can predict tic onset on single trials—an important test of network-level biomarkers for aDBS.
- Chapter 7 (Project 5): Presents the implementation and debugging of a real-time CNN-driven aDBS prototype. Delves into the challenges posed by stimulation artefacts and dynamic range shifts, proposes a custom time-domain filter for artefact suppression, and compares CNN stability to feature-based approaches.
- Chapter 8: General Discussion. Synthesises the main findings, discussing how multi-band analyses, connectivity metrics, and advanced ML pipelines might shape the next generation of adaptive DBS. Reflects on interpretability, patient specificity, and future research directions.

By blending machine learning approaches, analysis of neural signals, and clinical insights across distinct pathologies and patient groups, this thesis aims to contribute towards the development of robust, versatile, and ultimately more effective adaptive DBS strategies.

2 General Methods

2.1 Ethics

All studies described in this thesis have received approval from the relevant ethics bodies, specifically the Ethics Committee of the Medical Faculty of the University of Cologne for any work conducted in Germany, and the South Central–Oxford C Research Ethics Committee (as well as appropriate local Research Ethics Committees) for any work conducted in the United Kingdom. Where applicable, the UK Health Research Authority (HRA) regulations have also been observed. All procedures were conducted in accordance with the Declaration of Helsinki, ensuring that each participant provided written informed consent after being given sufficient time to consider their participation.

To safeguard participant confidentiality, personal identifiers were replaced with anonymised codes. Any identifying paperwork (e.g., consent forms) was stored in secure hospital or university buildings, whereas electronic data was kept on password-protected university servers with restricted access. In line with institutional policies and UK/EU data protection regulations, only members of the research or clinical team who needed to verify study integrity had access to potentially identifiable information. Any proposed multicentre data sharing occurred solely in an anonymised format, in line with each ethics committee’s stipulations. Should participants have wished to withdraw from the study at any time, all further data collection ceased immediately. No participant’s choice to withdraw or not to participate at all affected their standard care.

2.2 Externalised Electrophysiological Recordings

This section outlines the processes through which neural signals were recorded from deep brain stimulation (DBS) leads that were temporarily externalised before the permanent implantable pulse generator (IPG) was connected. Across the different experimental projects included in this thesis, specific target nuclei, recording devices, and electrophysiological modalities varied. The subsections below discuss (i) the surgical approach for lead implantation and externalization, (ii) the hardware and experimental protocols for data capture, and (iii) the data preprocessing routines that followed.

2.2.1 Lead Implantation and Externalisation

Stereotactic Targeting

All patients included in this thesis, following standard procedure, underwent fused MRI and CT imaging to plan DBS lead trajectories based on their clinical diagnoses:

- **Experimental Projects 1–2 (Tremor):** ViM thalamus or zona incerta (ZI) in patients with refractory essential, dystonic, or parkinsonian tremor.
- **Experimental Project 3 (Parkinson's disease):** Bilateral STN for bradykinesia, rigidity, and tremor symptoms in PD.
- **Experimental Project 4 (Tourette's tics):** Medial thalamic targets (centromedian--parafascicular complex [CM-Pf], ventro-oralis internus [Voi], or ventral anterior / ventral lateral [VA/VL] nuclei).

Lead Externalization

Once DBS leads were surgically placed, they were anchored at the burr hole and then externalised through a tunnelling procedure. Instead of immediately connecting them to an implanted pulse generator, the proximal connectors were left external to the skin and covered with sterile dressings. This interim setup—spanning approximately 2–7 days—allowed for direct postoperative recording sessions, clinical imaging, and test stimulation before IPG implantation.

During research recording sessions, each externalised contact was attached to a dedicated connector cable linked to the chosen amplifier. A sterile extension cable was used to interface between the intracranial lead and any external system while preserving the intracranial lead's position. Data collections occurred within the first week after lead placement, prior to IPG implantation.

Several centres have now accumulated experience in postoperative externalization of deep brain stimulation (DBS) leads, showing that it can be carried out safely with no significant increase in infection rates. Multiple retrospective and prospective reports, including large single-centre series and meta-analyses, consistently indicate that the incidence of infections following lead externalization

remains within the lower or mid-range of what has historically been reported for DBS surgery overall, typically between about 2% and 7% of patients.

Rosa and colleagues, for instance, found a 2.8% postoperative infection rate in 105 Parkinson’s patients despite routine externalization for *local field potential* (LFP) recording, concluding that the procedure “did not significantly increase the risk of infection” (Rosa, Scelzo, et al., 2017). Similarly, Feldmann et al. analysed more than 500 patients undergoing two-stage DBS, of which over 70% took part in LFP recordings. Even with extended recording sessions, the authors reported no difference in infections between those who did and did not participate in research recordings, pointing to an overall 3.6% infection rate, “within the lower range of incidences reported in the literature” (Feldmann, Neumann, Faust, et al., 2021). These findings echo a multi-institutional review by Mostofi and colleagues, who likewise documented no increase in hardware-related infections following externalization of leads for research or clinical trial stimulation (Mostofi et al., 2021).

Systematic reviews bolster these observations. One meta-analysis by (Kashanian et al., 2021), encompassing 23 studies and over 1,300 externalised patients, found a pooled infection proportion of approximately 6.9%. This figure aligns with infection rates reported for non-externalised DBS cohorts—an overlap further suggested by (Engel et al., 2018), who call for “unambiguous categories” of complications but note that infection incidences around 3–6% are not uncommon in DBS, irrespective of externalization. Although some teams have observed modestly higher rates in externalised patients—particularly in groups with certain indications such as refractory pain or psychiatric disorders—these differences typically fail to reach statistical significance, especially in large single-centre series (Runge et al., 2024).

Several explanations help account for why lead externalization does not appear to raise infection risk substantially. First, the process is typically coupled to strict perioperative protocols, including sterile dressing changes and limited handling of the external leads (Rosa, Scelzo, et al., 2017). Many centres employ prophylactic antibiotic regimens during the externalization period, along with complete scalp shaving and rigorous wound inspection (Feldmann, Neumann, Faust, et al., 2021). Second, externalization usually persists only for a short interval—often under a week—limiting opportunities for microbial colonization.

From a practical standpoint, externalised leads offer benefits for both patient care and research. Clinicians may test stimulation to confirm target efficacy (particularly relevant in less-established DBS

indications like chronic pain or Tourette syndrome) or carry out LFP recordings in awake patients to deepen understanding of underlying pathophysiology (Mostofi et al., 2021). These data have driven the development of *adaptive DBS* strategies, which rely on a real-time neural feedback signal to titrate stimulation (Rosa, Scelzo, et al., 2017). The shift toward next-generation sensing-enabled IPGs partly aims to obviate the need for externalization in the future, but until these fully replicate the fidelity of external LFP capture, postoperative lead externalization remains an important research tool.

Infections do remain among the most common complications in *all* DBS surgeries, so meticulous wound care and vigilant follow-up are crucial, especially given the possible need for hardware removal in more severe cases. However, from the standpoint of feasibility and patient safety, the evidence indicates that LFP-based or trial-stimulation-based research reliant on externalization of leads can be carried out safely, providing physiological insights without substantially jeopardizing clinical outcomes.

2.2.2 Electrophysiological Signal Acquisition

Amplifier and Data Acquisition

Recording hardware differed across experiments but consistently supported multi-channel, high-resolution sampling:

- **TMSi Porti and Saga Amplifiers (Projects 1–3):** These TMSi devices (Twente Medical Systems International, The Netherlands) were typically operated at sampling rates of 2048–4096 Hz and enabled concurrent recording of multiple channels (e.g., DBS LFP contacts, scalp EEG, EMG). Data were stored offline on secure laboratory computers.
- **Medtronic Percept IPG (Project 4):** In Tourette's patients, a sensing-capable Percept IPG was implanted, streaming LFP data from the subthalamic or thalamic leads at 250 Hz. This permitted longer-term recordings after surgery, albeit at a lower bandwidth, and under more stringent constraints when compared to the externalised recordings with TMSi systems.

Data were saved in either manufacturer-specific (e.g., .json for Medtronic) or custom formats and processed with dedicated software, including Spike2 (Cambridge Electronic Design Limited) or Python-based toolboxes. Some pilot experiments utilised real-time streaming through Lab Streaming Layer (Kothe, 2014) for testing closed-loop algorithms.

EEG Configuration

In addition to subcortical LFPs, scalp EEG was recorded in some projects to monitor cortical oscillations:

- **Projects 2–3:** One or more EEG electrodes were positioned at central or midline locations to capture motor cortical activity.
- **Project 4 (Tourette's):** A more extensive 10–20 EEG layout was applied, utilizing 63 wet Ag/AgCl electrodes, to capture broad cortical activity and identify tic-related oscillatory activity.

Careful grounding and referencing strategies (e.g., a “common average” or a dedicated reference electrode placed on the forearm) were implemented alongside cable management techniques to minimise motion artefacts and extraneous noise.

Peripheral Signals

To contextualise neural recordings with movement or muscle activity, peripheral signals were recorded:

- **EMG:** Surface EMG electrodes were placed over relevant muscles (e.g., forearm extensors in Projects 1–3).
- **Accelerometry/Inertial Sensors:** For Projects 1–2 (tremor), accelerometers attached to wrists or hands quantified tremor severity and indicated movement states. In Project 4 (Tourette's), inertial sensors helped detect episodes of tics.

Behavioural Protocols

A concise overview of behavioural tasks (which are described in detail in the respective chapters):

- **Projects 1–2 (Tremor):** Posture holding, spiral drawing, and gripping to elicit resting, postural, or intention tremor.
- **Project 3 (Parkinson's):** Alternation between reading aloud of words (speech tasks) and isometric gripping to probe cortical/subcortical representations of movement and speech.

- **Project 4 (Tourette's):** Spontaneous tic monitoring for ~30–60 minutes. Clinicians annotated tics in real-time or via video replay.

2.2.3 Data Preprocessing

All raw data underwent a standard preprocessing pipeline to address slow drifts and reduce sources of noise:

1. **High-Pass Filtering:** Typically 0.5–1 Hz cutoff to reduce DC offsets and low-frequency drifts.
2. **Notch Filtering:** Removal of mains interference (50 Hz), including harmonic frequencies if necessary.
3. **Artefact/Trial Rejection:** Signal amplitude thresholds were used to monitor for cable-motion artefacts or high-amplitude spikes. Flagged segments were inspected and manually rejected if contamination was severe.

When DBS stimulation was active, stimulation artefacts frequently contaminated signal portions near the primary frequency (mainly 130 Hz) and its harmonics. The presence of this artefact is commonly recognised as an issue that renders the processing of signals during the delivery of stimulation challenging (W.-J. Neumann, Gilron, et al., 2023). Bipolar referencing of DBS contacts mitigated some stimulation artefacts, yet signal saturations could persist at high amplitudes, especially when using the TMSi Porti system. More advanced artefact suppression and blanking methods were explored in in Project 5.

Collectively, these standardised preprocessing steps provided data of sufficient quality for subsequent offline analyses and machine-learning-driven decoding approaches.

2.3 Training Machine Learning Systems

A critical focus of this thesis is the application of machine learning (ML) techniques to study and decode electrophysiological signals—most notably local field potentials (LFPs) recorded from DBS

leads—in the context of adaptive deep brain stimulation (aDBS). This section summarises why machine learning is well-suited to the analysis of these signals, which specific families of methods were used, and how we addressed practical concerns such as computational overhead, real-time constraints, and model validation.

2.3.1 Rationale for Using Machine Learning in DBS Research

Adaptive DBS paradigms, from a technical perspective, can be assumed to be composed of three main conceptual modules: a sensing module (computes a feedback variable), a control module (interprets this feedback variable to determine a set of instructions for the stimulator), and a stimulation module (delivers the stimulation to the appropriate contact) (Marceglia et al., 2021). The work in this thesis focuses on the sensing module, concretely sensing modules which compute feedback variables from neural data to titrate stimulation.

Classical electrophysiology often relies on trial-averaged analyses, where many repetitions of an event (such as a movement onset) are averaged to isolate consistent neural correlates. While averaging techniques can powerfully detect stable phenomena, they necessarily obscure within-trial variability. In particular, dynamic fluctuations that occur on a single trial basis can be obscured. This becomes problematic in applications like aDBS, which needs real-time decoding of patient-specific neural patterns. Practical aDBS scenarios do not afford the opportunity to average across multiple repetitions to make a therapeutic adjustment to stimulation settings.

Machine learning (ML) methods offer promising ways to overcome these limitations and are well-suited to the single-trial paradigm that aDBS requires:

First, ML frameworks can **handle heterogeneous data**. Neural recordings often vary between patients and sessions. Rather than relying on a single, hand-picked feature—such as alpha power or beta power—ML can learn flexible mappings, either directly from raw signals or from diverse feature sets, to relevant brain states, effectively accounting for inter-individual and inter-session variability. This adaptiveness becomes especially necessary in the multi-patient contexts that future large-scale aDBS systems will face.

Second, ML systems permit the **integration of multiple features and sources of information**. Instead of focusing narrowly on bandpower alone (e.g., alpha or beta), ML pipelines can incorporate a

broad range of descriptors—time-domain statistics, spectral power estimates, phase synchrony measures, or autoregressive parameters—to reveal subtle or nonstationary patterns. Such patterns might be lost or assumed irrelevant in a strictly hand-engineered feature approach. By data-driven weighting of these diverse inputs, ML can automatically highlight whichever combinations best discriminate neural states of interest (e.g. symptom onset, behavioural state, etc.) from background activity.

Third, **ML-based therapy personalization** emerges because machine-learned decoders assign individualised weights to each input feature. When different patients exhibit distinct neural correlates for the same clinical symptom, a single traditional analysis pipeline might fail to capture these differences (without laborious retuning). ML-based models, however, learn patient-specific weighting factors from data, enabling them to adapt to each individual’s idiosyncratic neural signatures. This fosters a more personalised approach to therapy, which is likely necessary for optimising clinical outcomes in DBS.

Fourth, ML models offer real-time viability. An aDBS system must accurately decode the presence or onset of abnormal brain states within a few hundred milliseconds and then adjust stimulation to mitigate unwanted symptoms. Reliable single-trial classification forms the basis of such a closed-loop controller: detecting tremor onset or tic emergence promptly allows the stimulator to respond in a timely manner.

Finally, because ML outputs are typically continuous or probabilistic, clinicians can tune thresholds to balance sensitivity and specificity of detection. This thresholding mechanism determines at which “probability” the model’s output is interpreted as evidence of an abnormal state, such as tremor. By adjusting thresholds, one can optimise energy usage and avoid unnecessary stimulation or side effects while ensuring that critically important events are still detected. This threshold tuning thus becomes an additional dimension for personalizing therapy in clinical practice, complementing other methods for adapting ML decoders to a patient’s unique neural patterns. The relationship between thresholding of continuous outputs of ML-based systems and the consequent titrating of stimulation remains an important consideration in practical aDBS applications (He et al., 2021; Malekmohammadi et al., 2016; Velisar et al., 2019).

Overall, the ML framework in this thesis aims to support two complementary goals. First, there is biomarker discovery, whereby ML assists in uncovering consistent neural signatures that classical

manual approaches might miss. Second, there is state decoding, in which the machine learning pipeline harnesses these markers to identify the neural states of interest needed to drive responsive DBS therapies that only deliver stimulation when needed.

2.3.2 Feature-Based Classifiers vs. End-to-End Networks

Machine learning pipelines in this thesis follow two broad paradigms: feature-based methods and end-to-end neural networks. Each reflects different strategies for data modelling, interpretability, and computational requirements.

Feature-based Methods

In a feature-based approach, electrophysiological signals—often segmented into short windows—are preprocessed to extract a set of meaningful descriptors of instantaneous activity. These **features** might capture broadband time-domain characteristics (such as signal amplitude or kurtosis), signal power in canonical frequency bands (e.g., alpha, low-beta, high-beta, gamma), phase relationships between different channels, or alternative descriptions of oscillatory activity such as cepstral or autoregressive parameters. Because these features condense key information, they can then serve as input to a variety of classifiers.

Within this thesis, simpler **linear models** (e.g., logistic regression) are often used to decode speech, movement, or tic states in projects that prioritise interpretability and rapid training. Logistic regression is straightforward, requiring only a moderate number of hyperparameters and yielding easily interpretable model weights. In cases requiring more flexibility, **non-linear classifiers** such as gradient-boosted decision trees (GBDT) or support vector machines (SVMs) can capture more complex relationships within the same feature space. These non-linear methods are especially helpful when interactions among features (e.g., synergy between distinct frequency bands) become important to discriminate disease states or behaviours.

Neural Networks and End-to-End Learning

An alternative approach is to train a neural network directly on raw electrophysiological (e.g. LFP or EEG) time series, bypassing the need for manual feature engineering. In convolutional neural networks (CNNs), the learned filter kernels act somewhat like trainable bandpass filters—automatically identifying spectral or temporal patterns relevant for classification. For instance, Project 1 harnesses

CNNs that operate on unprocessed LFP signals to discover spatiotemporal patterns related to movement or rest states without requiring explicit bandpower calculations and explores the properties of these networks to capture common modulatory patterns of interest.

Although not used as extensively throughout the thesis, recurrent neural networks (RNNs), such as LSTM architectures, can capture longer time dependencies, such as extended tremor bursts or slow changes in neural state. They are particularly well-suited when the relevant neural phenomena unfold over multiple seconds or involve complex temporal sequences.

Table 2-1: Comparative Aspects of Feature-Based vs. End-to-End Approaches

Approach	Pros	Cons
Feature-based Methods	Generally higher interpretability (e.g., weights in logistic regression, feature importances in GBDT). Lower computational cost and simpler to deploy on embedded hardware.	May miss subtle non-linear interactions or complex spatiotemporal patterns in the signal. Require hand-crafted features, which can be suboptimal if the “true” biomarkers lie outside the chosen feature space.
End-to-End Trained Neural Networks	Potentially higher decoding performance if sufficient training data are available (the network learns to detect relevant patterns) Less manual feature engineering, as the model “learns” filters directly from the raw signals.	Require larger training sets to avoid overfitting. More challenging interpretability. Generally heavier computational footprint at inference time. More sensitive to shifts in the data distribution (e.g., changing stimulation parameters).

2.3.3 Model Training, Computational Footprint, and Real-Time Constraints

One of the core motivations for exploring ML in DBS research is the potential for adaptive or closed-loop stimulation. Such an approach necessitates near-real-time decoding of neural signals and imposes significant requirements on computational efficiency.

In this thesis, model training is primarily performed off-line. Large datasets from recording sessions—often spanning multiple patients, tasks, and electrode configurations—are collected and stored. These data are then transferred to a high-performance computing (HPC) environment equipped with two NVIDIA GPUs, allowing computationally expensive training procedures (e.g., deep network backpropagation or hyperparameter optimization in GBDT) to complete in a reasonable time. Off-line training also provides the freedom to explore multiple architectures, hyperparameters, and cross-validation folds without the pressure of strict real-time deadlines.

For on-device inference, the situation is more constrained. Commercial implantable pulse generators (IPGs) have limited processing power and battery life. Even external controllers, if used for real-time decoding, must operate under tight power and timing budgets. In some projects—particularly Project 5—we tested real-time or “pseudo-real-time” prototypes that feed signals into a CNN or a feature-based classifier and then produce stimulation commands within tens of milliseconds. However, truly embedded deployments often require additional optimizations, such as quantizing neural network weights, using approximate arithmetic, or harnessing specialised hardware (e.g., FPGAs or ASICs).

2.3.4 Evaluating Model Performance

Assessing the quality of ML-driven decoders is crucial for understanding how well these methods can generalise to new data or real-world scenarios. This thesis employs several standard performance metrics and validation procedures.

ROC Curves and AUC

A central tool for assessing classification performance in this thesis is the Receiver Operating Characteristic (ROC) curve, accompanied by the Area Under the ROC Curve (AUC) measure. These provide a threshold-independent way to visualise and compare classifier performance.

An ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) for a series of classification thresholds that vary from 0 to 1. In this thesis, model outputs are almost always generated as continuous (or probabilistic) scores. For instance, logistic regression outputs a weighted sum of features, which is then passed through a sigmoidal (logistic) function to yield a probability in the range $[0,1]$. One must then choose a threshold $\tau \in [0,1]$ such that:

$$f(x) = \begin{cases} 1, & \text{if } p \geq \tau \\ 0, & \text{otherwise} \end{cases}$$

where p is the predicted probability for the “positive” class (e.g., presence of movement, tic, etc.).

The True Positive Rate (TPR or Sensitivity) measures the proportion of actual positive samples correctly identified by the classifier. It is given by

$$TPR = \frac{TP}{TP + FN}$$

where TP (true positives) are cases the model detects correctly as positive, and FN (false negatives) are missed positive cases. The False Positive Rate (FPR) measures the proportion of negative samples that are incorrectly labelled as positive. It is computed as

$$FPR = \frac{FP}{FP + TN}$$

where FP (false positives) are negative samples that were classified as positive, and TN (true negatives) are correctly recognised negatives.

By sweeping through thresholds τ from 0 to 1, one obtains different pairs $(TPR(\tau), FPR(\tau))$, and plotting these points yields the ROC curve.

A high threshold τ means the classifier is more conservative about labelling something as “positive,” so TPR decreases but FPR usually also decreases. This typically corresponds to fewer false alarms but a higher chance of missing true positives. A low threshold τ means the classifier is more liberal about

labelling positives, so TPR increases, but FPR also tends to increase. In many real-world clinical applications—especially in cases where missing a true condition can have adverse consequences—having a higher TPR at the expense of some increase in FPR may be preferable.

The Area Under the ROC Curve (AUC) condenses the entire ROC curve into a single scalar. Geometrically, it is the integral of TPR over FPR:

$$AUC = \int_0^1 TPR(FPR) d(FPR)$$

An AUC of 1.0 indicates a perfect classifier that separates positives from negatives with no overlap: for all thresholds, TPR approaches 1.0 while FPR remains near 0. An AUC of 0.5 indicates random performance, where the classifier cannot distinguish between the two classes better than flipping a fair coin.

AUC scores reflect performance across all possible decision thresholds, preventing one from having to commit to a single threshold upfront. They offer a more robust evaluation in settings where class distributions may shift or where the misclassification costs are not fixed in advance.

Handling Dataset Imbalance

Many real-world clinical datasets exhibit significant class imbalances, whereby the minority class is greatly overshadowed by the majority class. In practice, only a brief portion of the neural signal may correspond to a tremor onset, tic, or other clinically significant episode, while most of the recording remains within a background or baseline state. Because these infrequent episodes are precisely the focus for therapeutic interventions, their reliable detection is crucial. Standard model training, however, can result in under-detection of the minority class if steps are not taken to mitigate the uneven sampling.

Under extreme imbalance, a naive model might exclusively label observations as belonging to the majority class and still produce (deceptively) high overall accuracy, while overlooking critical events — an accuracy of 95% may reflect that 5% of the data are neural states of interest being missed completely. Such shortcomings can lead to suboptimal outcomes in an aDBS context, where ignoring infrequent events can potentially entirely negate the benefits of titrating stimulation on-demand.

Furthermore, certain evaluation metrics, particularly simple accuracy, can be misleading when the minority class makes up only a small percentage of the overall data. Even ROC-AUC, though less sensitive to class imbalance than raw accuracy, may fail to capture the full extent of minority-class under-detection if the model’s thresholding strategy yields few false positives at the expense of detecting almost none of the rare events. Consequently, other measures such as precision, recall, or balanced accuracy are essential to ensure that the classifier preserves sensitivity to the minority class.

To address these concerns, the work in this thesis employs several techniques aimed at giving adequate weight to rarer but important events. One approach involves penalizing misclassifications of the minority class more heavily within the model’s loss function, which can be accomplished in both logistic regressions and neural networks by assigning larger class weights or modifying the standard cross-entropy to place additional emphasis on rare samples. Another strategy, particularly in neural network training, is to oversample or draw the minority class more frequently during mini-batch construction, so that the training is not overshadowed by an excess of “non-event” examples. The resulting increase in exposure helps force the model to learn discriminative boundaries around the minority class.

We believe that measures for imbalance correction prove especially pertinent in an adaptive DBS setting. In scenarios where only brief segments of a recording indicate neural states of interest, failing to recognise those segments might undermine the entire value of an adaptive approach. Excessive false alarms likewise pose a problem, as excessive stimulator adjustments may diminish battery life and create unnecessary stimulation-induced side effects. Striking the right balance between sensitivity and specificity for the minority class remains important. By integrating weighting schemes, specialised sampling, and relevant evaluation metrics, this thesis ensures that ML-based classifiers treat the minority class appropriately and are capable of detecting sporadic neural events in single-trial or near-real-time settings. This attention to imbalance ensures that the pipeline captures the potentially sparse states of neural activity that motivate adaptive stimulation paradigms.

Assessing Classifier Generalisation

Performance statistics and cross-validation schemes in machine learning analyses differ fundamentally from the trial-averaged approach commonly found in electrophysiology. Whereas trial averaging aggregates all data into a single grand mean, machine learning prioritises assessment of how well a

model generalises to unseen data. This distinction is vital in a field marked by substantial variability, including not only inter-trial and inter-session fluctuations but also inter-patient differences. In prospective aDBS systems, where accurate single-trial prediction is required to trigger or modulate therapy, demonstrating robust out-of-sample performance becomes a core objective.

One common strategy, especially for smaller datasets or where fine-grained performance estimates are needed, is to employ Leave-One-Trial-Out cross-validation (LOTO-CV) schemas. In this setup, each trial is successively excluded from the training set and used as the test set. By iterating through the trials, the researcher can, during each iteration, evaluate how the model performs on data unseen by the classifier at training time. This procedure clarifies whether the model captures generalizable correlations as opposed to memorizing spurious relationships within a set of training data. If classification remains accurate when tested on the withheld trial, one can be more confident of the model’s out-of-sample reliability.

For assessing generalisation across different individuals, another variant is Leave-One-Patient-Out cross-validation. Instead of using individual trials as test sets, entire patient datasets are held out in turn, ensuring that the model is trained exclusively on data from the remaining participants. This method probes whether the patterns learned in one group of patients can be transferred to new, previously unseen patients—a question of clinical importance, since it determines whether a single global model might suffice in practice, or whether individualised models are essential to capture patient-specific signal features and noise profiles.

When dealing with time-series data, an additional concern is “data leakage,” which occurs if consecutive windows from a single session appear in both the training and validation sets. Such overlap can allow a model to exploit subtle session-related correlations rather than truly learning the underlying neural markers of interest. Group or block cross-validation schemes prevent this by ensuring that entire blocks of data (for instance, entire recording sessions) are assigned wholly to either the training or test fold. This arrangement simulates real-world deployment of a classifier, where future data are genuinely novel rather than merely new windows from a previously seen session.

Finally, for small or highly imbalanced datasets, permutation testing offers a robust way to establish a significance threshold for classifier performance. Instead of relying solely on parametric assumptions, one generates a null distribution of performance metrics (for example, AUC scores) by randomly shuffling the class labels or relevant time segments. If the model’s actual AUC exceeds the 95th

percentile of this null distribution, it can be concluded that the observed classification accuracy is highly unlikely to have arisen by chance. Such an approach is particularly informative when conventional statistical tools may not be well suited to the idiosyncrasies of limited or highly skewed datasets.

3 Project 1: Exploring the Use of Convolutional Neural Networks for aDBS Applications

The material presented in this chapter was the basis for the following journal publication:

F. Rodriguez, S. He, and H. Tan, ‘The potential of convolutional neural networks for identifying neural states based on electrophysiological signals: experiments on synthetic and real patient data’, *Front. Hum. Neurosci.*, vol. 17, p. 1134599, Jun. 2023, doi: [10.3389/fnhum.2023.1134599](https://doi.org/10.3389/fnhum.2023.1134599).

3.1 Introduction

3.1.1 Signal Processing Approaches for Decoding Neural States from Electrophysiological Signals

Electrophysiological recordings capture neural activity as multi-channel time-series data, revealing oscillatory patterns that reflect underlying neural circuit states. These brain states can indicate various neural, behavioural, or mental conditions, the detection of which can be for both clinical and research applications. For instance, recognizing specific neural patterns can enable the development of brain-computer interfaces (BCIs) that establish alternative communication pathways between the brain and external devices. Additionally, identifying pathological neural states can inform targeted interventions, such as electrical stimulation therapies aimed at modulating aberrant neural activity for therapeutic benefit.

A critical challenge lies in extracting and characterizing the patterns within neural activity that best differentiate distinct neural states. Developing systems capable of recognizing these patterns in real time is essential for applications like adaptive deep brain stimulation (aDBS). In aDBS, stimulation parameters are automatically adjusted based on the identified brain state, optimising therapeutic outcomes while minimising side effects (see Figure 3.1).

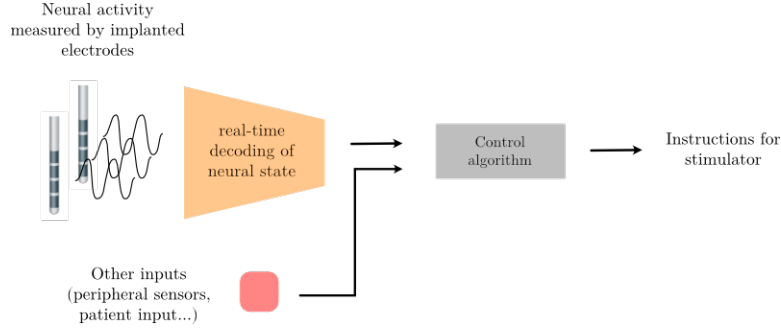


Figure 3.1: Schema of a signal processing pipeline for adaptive Deep Brain Stimulation (aDBS) or Brain-Computer Interface (BCI) applications.

Neural decoding—the process of interpreting neural signals to infer underlying neural states—is typically achieved by training machine learning models such as classifiers, regressors, or clustering algorithms. These models rely on quantifying relevant activity patterns within the signals, which often manifest as synchronous bursts of oscillatory activity at canonical frequencies with specific waveform characteristics (W.-J. Neumann et al., 2022). Therefore, selecting appropriate methods for extracting these patterns—referred to as features—is crucial for effective neural decoding.

In the literature, algorithmic approaches for neural decoding can be broadly classified into two categories: the classical machine learning or feature-based approach, and the end-to-end deep learning approach. Each approach offers distinct methodologies for processing and interpreting electrophysiological data, with unique advantages and limitations.

3.1.1.1 The Classical Machine Learning, or Feature-Based, Approach

The feature-based approach involves manually extracting predefined features from raw neural waveforms using signal processing techniques such as filtering, smoothing, and spectral analysis. These features are chosen based on prior knowledge of the signal characteristics that are believed to be informative about the neural states of interest.

Common methods for feature extraction include spectral power estimation using periodograms and the fast Fourier transform (FFT) (Ahn et al., 2020; Bijanzadeh et al., 2022; He et al., 2021), as well as digital filtering techniques to isolate activity within specific frequency bands (S. A. Shah, Tan, et al., 2018). Time-domain features, such as amplitude and waveform morphology, have also been shown to contribute to decoding performance (Merk, Peterson, Lipski, et al., 2022; Thenaisie et al., 2022). Efforts are ongoing to expand the feature set to encompass additional waveform characteristics,

including waveform shape features, connectivity estimates and non-linear dynamical estimates³ (Merk et al., 2023).

Once the features are extracted, they are input into a classifier that models the statistical relationships between the features and the neural states present in the training data. The classifier learns a mapping function that associates specific combinations of features with the probability of corresponding neural states. At inference time, the classifier uses this learned mapping to predict the most likely neural state based on the features extracted from incoming neural waveforms (see Figure 3.2).

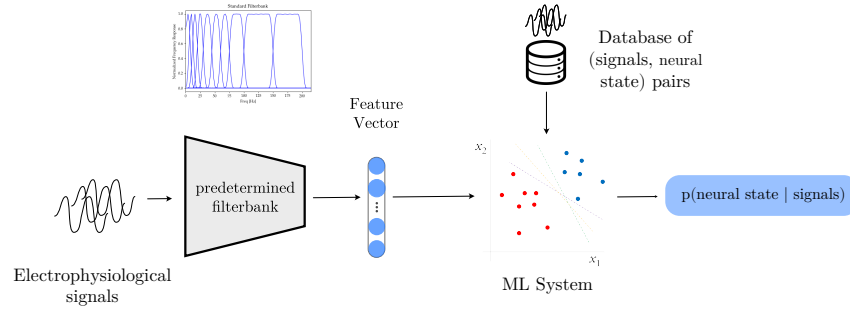


Figure 3.2: Feature-based Approach for Neural Decoding. Electrophysiological signals are processed by a pre-designed filterbank, which extracts features. These manually engineered features are then aggregated into a feature vector, which serves as input to a machine learning classifier or regression model. The model maps the features to corresponding neural or behavioural states, enabling the decoding of neural activity based on predefined signal characteristics.

A key aspect of the feature-based approach is the necessity for careful feature design and selection. The quality and relevance of the extracted features directly impact the classifier's ability to learn meaningful relationships and make accurate predictions. This approach has been extensively utilised in current movement state-driven aDBS implementations (Ahn et al., 2020; He et al., 2021; Merk, Peterson, Köhler, et al., 2022; Tan et al., 2019).

³ We refer to https://github.com/neuromodulation/py_neuromodulation for an open-source, Python-based toolbox that is, at the time of writing, in active development.

Deep learning methods can be incorporated within this framework by applying neural networks to processed features or derivative representations of the raw data. For example, (Golshan et al., 2020) decomposed neural signals into two-dimensional time-frequency images using the continuous wavelet transform and then applied a convolutional neural network (CNN) to these images for decoding.

3.1.1.2 The End-to-End Deep Learning Approach

The end-to-end deep learning approach integrates feature extraction and classification into a single model, typically using convolutional neural networks (CNNs). Unlike the feature-based paradigm, this approach does not rely on manually designed feature extraction pipelines. Instead, the model learns to extract relevant features directly from raw neural waveforms during the training process (see Figure 3.3).

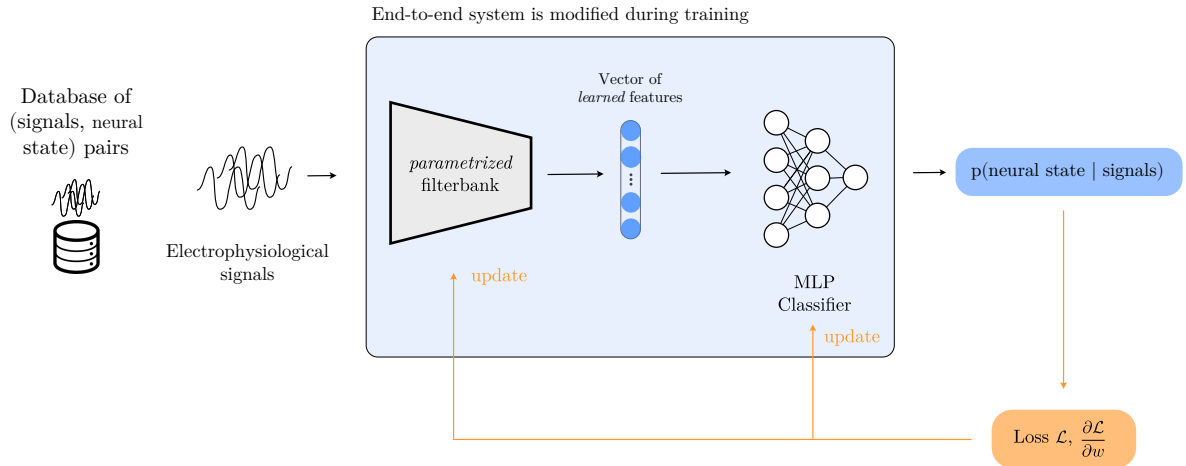


Figure 3.3: End-to-End Deep Learning Approach for Neural Decoding. The end-to-end deep learning approach bypasses manual feature extraction by processing raw electrophysiological signals directly through a neural network, typically a convolutional neural network (CNN). The network consists of multiple layers that learn hierarchical representations from the input data. Convolutional layers extract local patterns and features from the signals and subsequent layers combine these features to identify complex, task-relevant patterns. The network outputs predictions of neural states or intentions, integrating feature extraction and classification within a single, trainable model. This approach harnesses the ability of deep learning to discover and capture structures in the data, potentially enhancing decoding performance.

Convolutional operations are particularly well-suited for processing naturalistic signals like images and waveforms due to their ability to capture local patterns and hierarchical structures. In the context of neural signals, convolutions can effectively identify activity occurring at relevant frequencies. This is

supported by the convolution theorem, which states that convolution in the time domain corresponds to multiplication in the frequency domain: $\mathfrak{F}\{f * g\} = \mathfrak{F}\{f\} \cdot \mathfrak{F}\{g\}$, where $\mathfrak{F}$ denotes the Fourier transform and $*$ the convolution operator (see Supplementary Figure I for a more detailed description of convolutional filters).

The end-to-end approach has demonstrated significant improvements over feature-based methods in domains such as computer vision and speech processing, where signals exhibit complex, high-dimensional structures. In the context of electrophysiology, CNNs have been successfully applied to various applications, including seizure detection, sleep stage classification, and movement decoding (Gao et al., 2021; Golshan et al., 2020; S. M. Peterson et al., 2021; Shoeibi et al., 2021; Supakar et al., 2022).

By employing CNNs, the feature extraction process becomes data-driven, allowing the model to automatically learn optimal representations of the neural signals. This can potentially result in more efficient and accurate decoding by capturing complex patterns that may be overlooked or discarded in manual feature extraction processes. For instance, setting frequency band boundaries or identifying harmonic patterns manually may inadvertently exclude relevant information.

However, the end-to-end deep learning approach presents its own challenges. Designing an appropriate network architecture that is well-suited to the specific decoding task is critical. The model's performance is highly dependent on the architecture's capacity to capture relevant features without overfitting or underfitting the data. Additionally, deep learning models require large amounts of high-quality training data to generalise effectively to new, unseen data. In scenarios where data availability is limited, such as with patient-specific neural recordings, this can be a significant constraint.

Furthermore, deep learning models are often regarded as "black boxes" because they automatically learn complex representations from data without explicit feature design, making it challenging to interpret the features they extract and understand their decision-making processes. This lack of interpretability raises concerns in both research and translational contexts. In research applications, interpretability is essential for gaining insights into the underlying neural mechanisms. Understanding which aspects of the neural signals the model deems important helps neuroscientists validate hypotheses about brain function, identify new biomarkers, and advance fundamental knowledge. Without transparency, it becomes difficult to determine whether the model is capturing meaningful

neural patterns or merely exploiting spurious correlations in the data, which can hinder scientific progress and reduce confidence in the findings derived from such models.

In translational and clinical settings, the interpretability of deep learning models is crucial for predicting failure modes, debugging, ensuring patient safety, and meeting regulatory and ethical requirements. Clinicians and engineers need to understand the modulatory patterns that systems rely on to anticipate how it might fail under different conditions, ensuring reliable behaviour in real-world scenarios. When the system's performance is suboptimal, interpretability also allows developers to identify specific components or features causing issues, enabling targeted improvements and optimizations.

3.1.2 Signal Processing for aDBS

Adaptive deep brain stimulation (aDBS) is an advanced therapeutic technique wherein stimulation parameters are dynamically adjusted based on real-time feedback signals related to brain state or pathology. This approach holds promise for enhancing the efficacy of DBS in treating movement disorders while reducing side effects and expanding the therapeutic window (Guidetti et al., 2021; W. Neumann & Rodriguez-Oroz, 2021).

Feedback signals for aDBS can be extracted directly from neural activity, such as local field potentials (LFPs) recorded from DBS electrode contacts, or from neurophysiological signals obtained from additional electrodes placed in other brain regions. Decoding movement or pathological states from LFPs using machine learning methods enables stimulation to be delivered only when necessary, minimising overall current delivery and potentially reducing side effects associated with continuous stimulation.

Clinical studies have demonstrated that aDBS can maintain therapeutic efficacy comparable to conventional DBS while mitigating side effects affecting speech, gait, or postural stability (He et al., 2021; Little & Brown, 2020). Additionally, aDBS may prevent patients from developing habituation or tolerance to the therapeutic effects of stimulation over time (Peters & Tisch, 2021).

The majority of current aDBS systems rely on the feature-based approach, involving manual extraction of predefined features from neural waveforms and training classifiers to learn the relationships between these features and neural states. While effective, this approach may

inadvertently discard relevant information during feature extraction, potentially limiting decoding performance.

Some studies have begun to explore hybrid methods that combine feature-based and deep learning approaches. For example, deep learning models may be trained on processed features or representations derived from the raw data, such as time-frequency decompositions obtained through wavelet transforms (Golshan et al., 2020; Thenaisie et al., 2022).

In this thesis, we aim to explore the feasibility of end-to-end deep learning approaches, specifically CNNs, for training real-time compatible decoding systems based on LFP waveforms. We compare the performance of these systems with traditional feature-based methods, documenting the advantages and drawbacks of each approach in the context of aDBS.

Our motivation for employing deep learning stems from its successful application in other areas of electrophysiology. For instance, (R. Sun et al., 2022) harnessed deep neural networks to achieve high-fidelity spatiotemporal source reconstruction of high-density EEGs, reducing the need for manual solver tuning. We hypothesise that end-to-end CNNs can similarly capture relevant neural state information from LFP waveforms that might be missed or discarded in manual feature extraction processes.

To test this hypothesis, we designed an end-to-end CNN architecture capable of extracting features from neural signals recorded during short experimental sessions with patients. We first validated our approach using synthetic waveforms with controlled activity patterns and signal-to-noise ratios, allowing us to assess the models' ability to capture specific signal characteristics. Both feature-based classifiers and deep learning-based classifiers were trained to distinguish between simulated states. Subsequently, we applied these models to real patient LFP data recorded from externalised essential tremor patients engaged in movement tasks. The classifiers were trained to distinguish movement periods from rest, and their decoding performance was compared.

The remainder of this study is structured as follows: first we explain how we generate synthetic data to mimic oscillatory patterns that one might find in electrophysiological recordings. We then implement different feature-based and deep learning-based machine learning systems and apply them to the task of decoding states using both synthetic and real patient data. We then benchmark these systems against one another, document respective benefits and drawbacks and conclude this study.

The methods developed here can be applied to both local field potential (LFP) activity measured from DBS electrodes as well as to other types of electrophysiological data including electrocorticography (ECoG), which has also been explored within this context (Merk, Peterson, Lipski, et al., 2022).

3.2 Methods

3.2.1 Dataset I: Synthetic data

To benchmark the ability of different algorithmic approaches to extract information from waveforms containing various oscillatory patterns, we generated synthetic datasets designed to mimic patterns found in neural time-series data modulated by different brain states. The synthetic datasets had a total duration of 20 minutes, with the first 80% (16 minutes) used as training data and the remaining 20% (4 minutes) as validation data. The signals were sampled at 2,048 Hz.

In the synthetic data, neural states were modelled as two-state, first-order Markov process. Each state had a minimum duration of 5 seconds, and the probability of transitioning from that state was randomly drawn from the uniform distribution $\mathcal{U}[0.3, 0.4]$. This configuration led to an average state duration of 10.92 ± 7.94 seconds, introducing stochasticity in the duration of states while maintaining a similar overall prevalence of each state. This approach avoids large class imbalances in the datasets and reflects the in-clinic setting where patients perform trials of movement followed by rest during deep brain stimulation (DBS) training sessions (He et al., 2021).

2-state, 1st order Markov Process:

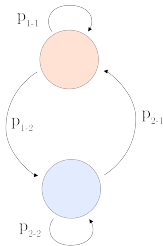


Figure 3.4: Simulated neural states were modelled as a first-order Markov process. Transition probabilities were calibrated to avoid imbalances while preserving some degree of stochasticity.

Each state was characterised by relative differences in the presence of specific oscillatory patterns that have been previously associated with movement or pathological conditions in Parkinson's disease. We generated a total of seven oscillatory patterns, including variations in oscillations at specific frequencies, waveform nonlinearities, oscillatory activity modulated by the phase of another oscillator at a different carrier frequency (phase-amplitude coupling), differences in the duration of transient bursts of synchronised activity, and phase differences between channels. Except for the cross-channel phase-amplitude coupling (PAC) and cross-channel phase shift cases, the patterns were generated within a single channel. The classification task involved distinguishing between the two states based on these patterns.

To simulate realistic neural recordings, 1/f-shaped pink noise was added to the signals at different intensities to modulate the noise content. The pink noise was generated using an open source, Python-based toolbox designed for this purpose⁴. The signal-to-noise ratio (SNR) was defined as:

$$SNR = \frac{rms(signal)}{rms(noise)}$$

where rms denotes the root-mean-square, representing the energy of the signal.

For each oscillatory pattern, we performed SNR sweeps ranging from 0.2 (indicating five times more noise power than signal power) to 2.0 (twice as much signal power as noise power) in increments of 0.05. For each of these 36 noise levels, we generated five datasets, resulting in a total of 180 datasets per oscillatory pattern. This systematic variation allowed us to assess the robustness of the algorithms across different noise conditions.

Below, we detail the seven oscillatory patterns used to characterise the different states in our synthetic datasets.

A. Power Changes in the Beta and Gamma Bands

Modulation in the power of the canonical beta (approx. 13 – 30 Hz) and gamma (approx. 40 – 90 Hz) frequency bands has been widely reported during movements in the cortical-subcortical motor

⁴ Available at <https://github.com/felixpatzelt/colorednoise>

network. To synthetically mimic this pattern, we generated oscillations at these canonical frequencies with differing amplitudes in the two states and summed them together. The resulting frequency spectra exhibited sharp peaks at the relevant frequencies above the noise floor, with different amplitudes characterizing the two states. Specifically, **State 0** was characterised by higher power in the beta band and lower power in the gamma band compared to **State 1**, which had the opposite configuration.

Figure 3.5 illustrates the time-series excerpts and corresponding frequency spectra of the two states, both without noise and with added noise simulated as a Markov process.

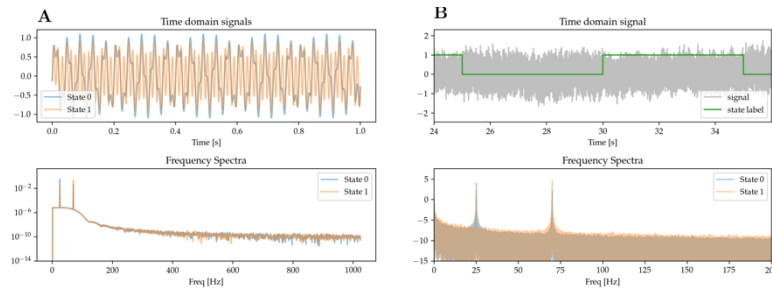


Figure 3.5: Oscillations in Beta and Gamma Bands Characterizing Two Simulated States. (A) Time-series excerpts of the two states without additive $1/f$ noise (top) and their corresponding frequency spectra (bottom). (B) The two states with added noise, simulated as a Markov process (top). The frequency spectra (bottom) show that the two states are distinguished by differences in power at the beta (25 Hz) and gamma (70 Hz) peaks.

B. Beta Waveform Sharpness

The sharpness of the beta waveform has been documented as a marker of circuit pathology in Parkinson's disease, in particular due to it being modulated by beta burst dynamics (Cole et al., 2017; Yeh et al., 2020). We generated sharp beta waveforms as non-sinusoidal, triangular waveforms at beta frequency to represent **State 1**, in comparison to sinusoidal waveforms for **State 0**. The power at the base beta frequency was matched for both states.

In the frequency domain, the triangular waveform's symmetry between positive and negative phases (sharpness at both the peak and trough) results in the cancellation of even harmonics. Consequently, the frequency spectrum is characterised by the presence of odd harmonics for State 1, whereas State 0 exhibits a single peak at the base frequency due to its pure sinusoidal nature.

Figure 3.6 depicts the time-series waveforms and frequency spectra for the two states.

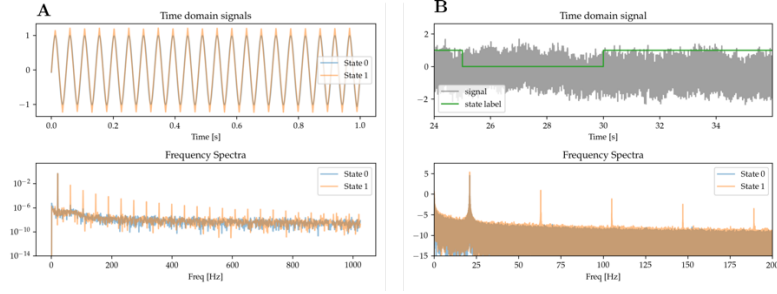


Figure 3.6: *Sinusoidal and Sharp Beta Waveforms Distinguishing Two States.* Time-series excerpts of the sinusoidal waveform (State 0) and the sharp triangular waveform (State 1) with matched power at the base beta frequency (top). The frequency spectra (bottom) show that State 1 has prominent odd harmonics due to waveform sharpness, while State 0 has a single peak at the beta frequency.

C. Non-Linear Phase

Besides waveform sharpness, other kinds of waveform non-linearities (deviations from pure sinusoidality) have been reported as relevant features of neural oscillations (Quinn et al., 2021). We implemented oscillations with nonlinear phase by introducing asymmetry between the durations of the first and second halves of the oscillatory cycle. This resulted in waveforms with fast trough-to-peak (rising) times and slow peak-to-trough (falling) times, producing narrow peaks and wide troughs for **State 1**. **State 0** maintained approximately sinusoidal waveforms.

The asymmetry between positive and negative phases leads to the presence of both odd and even harmonics in the Fourier-derived frequency spectrum. The signals were scaled to equate power at the base beta frequency for both states.

Figure 3.7 presents the time-series data and frequency spectra illustrating the phase nonlinearity.

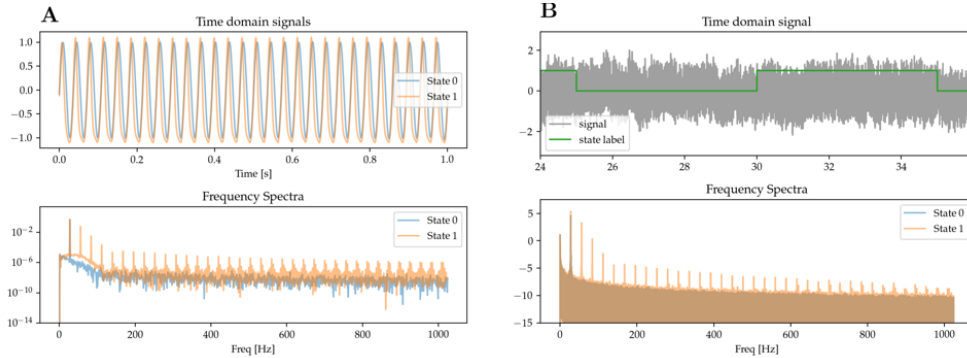


Figure 3.7: *Oscillations with Nonlinear Phase Characterizing Two States.* (Top) Time-series excerpts showing approximately sinusoidal waveforms (State 0) and waveforms with nonlinear phase resulting in sharp peaks and

wide troughs (*State 1*). (Bottom) The frequency spectra indicate the presence of both odd and even harmonics for *State 1* due to phase nonlinearity.

D. Theta-Gamma Phase Amplitude Coupling (PAC)

Interactions between neural oscillations at different frequencies, particularly phase-amplitude coupling, have been observed in various brain regions and linked to cognitive functions and pathological conditions (Canolty et al., 2006a; Tort et al., 2010a; Vivekananda et al., 2021). Increased beta-gamma PAC in the motor cortex is indicative of pathological activity in Parkinson's disease (Meidahl et al., 2019a).

We simulated this feature by generating transient bursts of gamma activity (70 Hz) that were either coupled to the positive phase of an underlying theta oscillation (6 Hz) for **State 1**, or appeared at random times for **State 0**. Overall gamma power was adjusted to be matched between the two states. We also simulated cross-channel PAC, where the gamma activity was coupled to the phase of a theta oscillation in a different channel. The frequency spectra exhibited harmonic microstructure around the gamma peak due to the coupling.

Figure 3.8 illustrates the time-series data and frequency spectra for same-channel and cross-channel PAC.

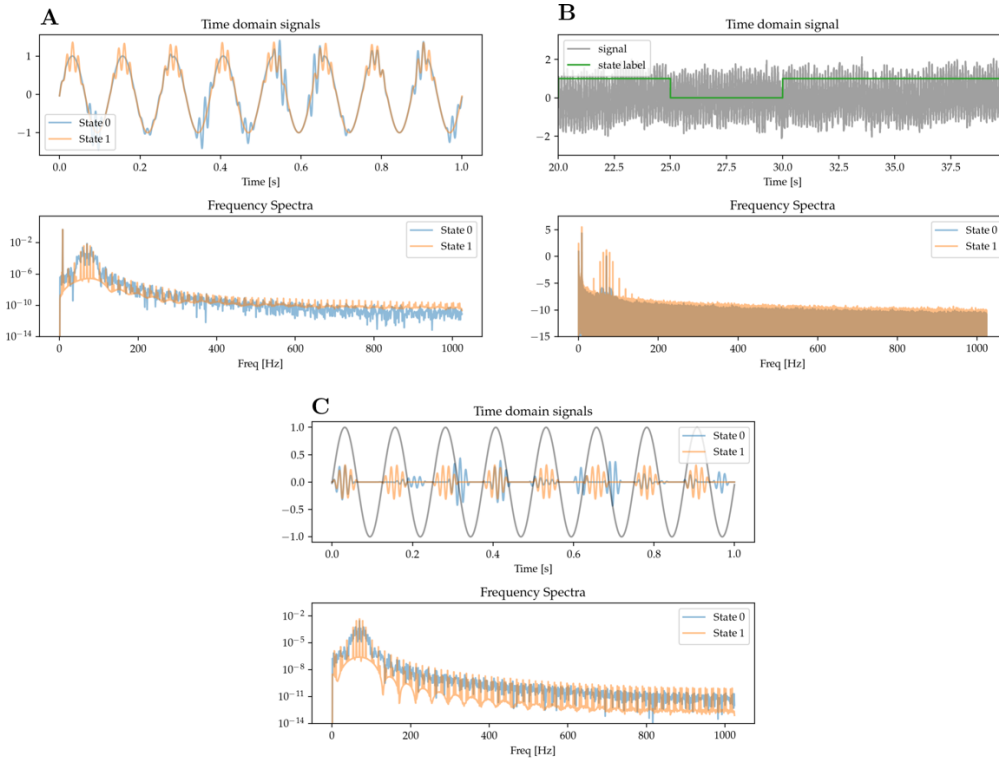


Figure 3.8: *Theta-Gamma Phase-Amplitude Coupling in Two States. (A, B) In same-channel PAC, gamma bursts are either uncoupled (random timing) in State 0 or coupled to the positive phase of a theta oscillation in State 1 (top). The frequency spectra show harmonic microstructure around the gamma peak for State 1 due to PAC (bottom). (C) In cross-channel PAC, gamma bursts in one channel are coupled to the theta phase in a different channel (State 1), whereas in State 0, gamma bursts occur at random times (top).*

E. Cross-Channel Phase Shift

One potentially interesting approach to characterise oscillations across brain regions and across sensor space is to be able to quantify the extent to which the same oscillations take place in different regions with a concomitant time-delay (or phase-shift) between them. Here, we generated signals in two sensor space channels that were phase-shifted by different amounts. The frequency content of the signals is random and there is thus no identifiable pattern in the power spectrum. The classification task involved distinguishing between the two different degrees of phase-shift (see Figure 3.9).

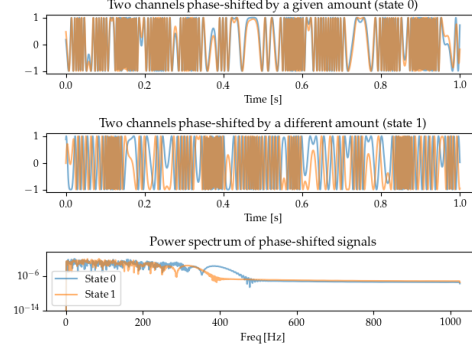


Figure 3.9: *Cross-Channel Phase Shift Characterizing Two States. Time-series excerpts of two channels with signals exhibiting different degrees of phase shift. In State 0, the phase shift between channels is smaller, whereas in State 1, the phase shift is larger. Since the signals are random, the power spectra do not provide distinguishing features between the states.*

F. Beta Burst Length

The role of the temporal dynamics of beta burst in basal ganglia LFPs has been identified as a relevant marker of circuit state and of pathological activity, especially in terms of the length of the bursts (Tinkhauser, Torrecillos, et al., 2018). To model this, we generated randomly occurring burst of activity un the canonical beta frequency. The average lengths for these bursts were different in the two different states: in **State 0**, episodes of increased beta power (beta bursts) were brief in duration, whereas in **State 1**, the beta bursts were longer. We adjusted the amplitude of the bursts to equate beta power at the peak frequency.

Figure 3.10 depicts the time-series waveforms and frequency spectra for the two states.

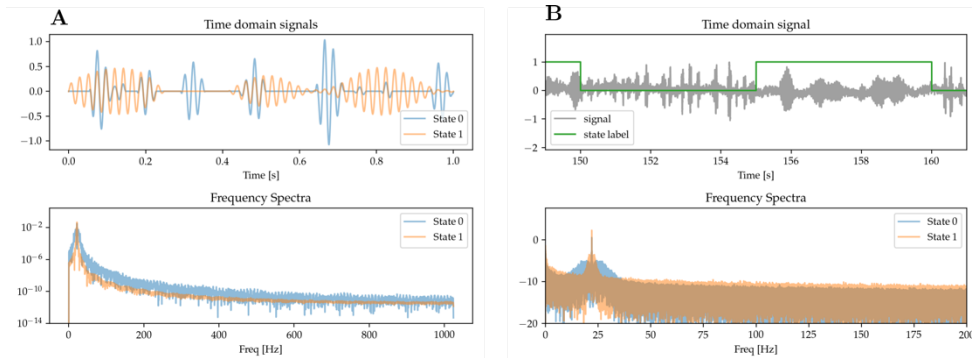


Figure 3.10: *Beta Burst Length Characterizing Two States. Time-series excerpts showing short beta bursts in State 0 and longer beta bursts in State 1 (top). The frequency spectra (bottom) are similar due to matched average power*

at the beta frequency, highlighting that the distinguishing feature relates to the temporal dynamics of the oscillation rather than its spectral power.

By systematically generating synthetic datasets with controlled oscillatory patterns and varying noise levels, we created a comprehensive framework to evaluate the performance of different signal processing and machine learning approaches in decoding neural states. The synthetic data allowed us to assess the algorithms' ability to capture specific features and their robustness to noise, which is critical for applications involving real neural recordings where signal-to-noise ratios can be low and variable.

3.2.2 Dataset II: LFP-based Movement Decoding

In order to evaluate the ability of different systems to decode states not only from synthetic data but also from real patient LFPs, we make use of an existing dataset consisting of LFPs recorded from the bilateral ventralis intermediate nucleus (ViM) of the thalamus in patients with essential tremor (ET). These patients had been externalised after lead implantation surgery and performed three self-paced upper-limb movement tasks. This dataset was first published in (He et al., 2021).

The dataset included recordings from eight patients during three tasks: a pegboard insertion task, a posture holding task, and a rice pouring task. In the pegboard insertion task, patients inserted pegs into a board. For the posture holding task, they held their arms stretched out in front of them. In the rice pouring task, patients poured rice back and forth between two cups. Each movement was performed for approximately 30 seconds, followed by a rest period of about 30 seconds. This cycle was repeated six to ten times per task. Each task was conducted under two conditions: with bilateral 130 Hz DBS stimulation turned on and off at a single preselected stimulating contact. This resulted in an average of 8.4 ± 1.8 minutes (mean $\pm$ standard deviation) of recording for each task per participant.

Recordings were made four or five days after electrode implantation using externalised leads. The dataset included LFPs recorded from all available contacts on the DBS leads in both hemispheres, bipolar electromyography (EMG) measurements of the flexor carpi radialis muscles of both arms, and triaxial accelerometer data from sensors taped to the dorsum of each hand. All signals were sampled at 2,048 Hz using a TMSi Porti amplifier (TMS International, the Netherlands).

As a preprocessing step, all signals were high-pass filtered using a fourth-order Butterworth filter in forward mode with a cut-off frequency of 1 Hz to remove slow drifts and baseline fluctuations. Movement labels were extracted from the accelerometer or EMG signals by thresholding the root-mean-square (RMS) power of the signals of interest, providing binary labels indicating movement or rest periods. All available LFP channels across both hemispheres were used as inputs to the classifiers, offering a comprehensive view of thalamic activity.

Due to the relatively short duration of the datasets—six to ten trials per task—we employed a five-fold cross-validation strategy to evaluate the classifiers' performance. Each dataset was divided into five folds of equal length. In each iteration, four folds were used for training, and the fifth fold served as the validation set. This process was repeated until each fold had served as the validation set once. The data comprised 200 folds in total, corresponding to eight patients, three tasks (some tasks were not recorded for certain patients), and two stimulation conditions. The training sets had an average duration of 6.7 ± 1.9 minutes, and the validation sets averaged 1.7 ± 0.5 minutes. This cross-validation approach ensured that the models were evaluated on unseen data, providing an unbiased assessment of their generalization capabilities.

3.2.3 Decoding Systems

We classify the decoding systems used in this work into one of two categories: feature-based and deep learning-based. Both make use of data to learn the mapping function:

$$f: \text{waveforms} \rightarrow p(\text{state} \mid \text{waveforms}),$$

The two approaches, however, achieve this in different ways, with the feature-based approach requiring one to first derive a feature vector manually from the waveforms, and subsequently using classical machine learning methods to learn the function

$$f_{\text{features}}: \text{features} \rightarrow p(\text{state} \mid \text{features})$$

By contrast, the deep learning approach learns a direct, end-to-end mapping from the waveforms to the state of interest by directly optimising the parameter set θ

$$f_{DL}: \text{waveforms}, \theta \rightarrow p(\text{state} \mid \text{waveforms}, \theta).$$

We implement two feature-based models, implemented using python’s scikit-learn package (Pedregosa et al., 2011), and three deep learning-based models – implemented using PyTorch’s Python API (Paszke et al., 2019). For all these systems, waveforms are split into 500 ms moving window epochs with 100 ms overlap. With this configuration, a system operating in real time would produce one output every 400 ms (equivalent to a rate of 2.5 Hz).

A. Feature-Based Model I: Spectral SVM

This first feature-based method, adapted from (He et al., 2021), computes a periodogram from multichannel waveforms to estimate power in the following frequency bands: δ/θ (2 – 8 Hz), α (8 – 12 Hz), β_{low} (13 – 20 Hz), β_{high} (21 – 30 Hz), γ_1 (31 – 45 Hz), γ_2 (46 – 55 Hz), γ_3 (56 – 75 Hz), γ_4 (76 – 95 Hz), γ_5 (96 – 105 Hz), γ_6 (106 – 145 Hz), γ_7 (146 – 155 Hz), γ_8 (156 – 195 Hz). Each of these features is z-scored with mean and standard deviations values computed from the training feature set, and a support vector machine (SVM) with radial basis functions is trained on the two-class classification problem. An inner cross-validation loop is used to determine the optimal degree of regularization.

B. Feature-Based Model II: Autoregressive Spectral SVM (AR-SVM)

This system operates with the same feature set and classifier specification as the Spectral SVM, but features are concatenated from the preceding 10 epochs, giving the classifier access to the dynamics from the preceding 4 seconds as opposed to solely the last 500 ms, thereby increasing its receptive field.

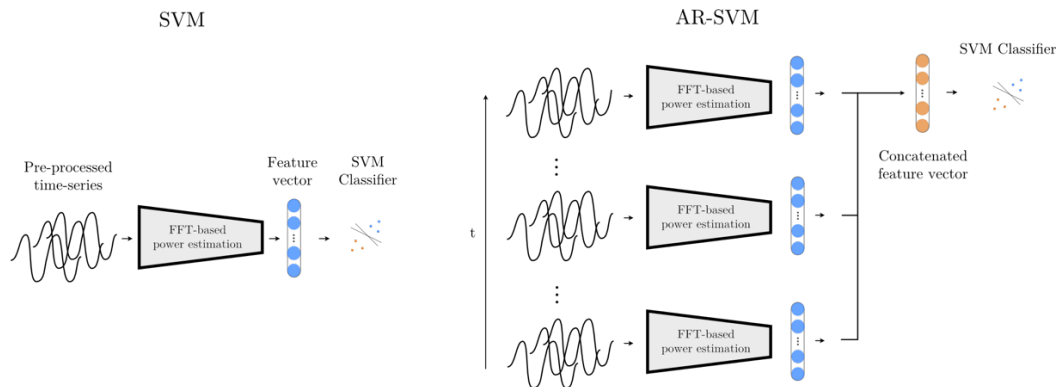


Figure 3.11: Feature-based models SVM and AR-SVM. In the SVM approach (left), pre-processed waveforms are turned into spectral features through FFT-based estimation of power in canonical frequency bands. This feature vector is the input to a support vector machine (SVM) classifier. The AR-SVM approach (right) concatenates

feature-vectors from previous epochs to provide the classifier with enhanced temporal context, i.e. broader access to information from the preceding neural dynamics.

C. Deep Learning-Based Model I: 1D-CNN

With this model, the multichannel waveforms are sequentially processed through 7 layers. Each layer consists of a custom-made non-linear compression operation, a 1D convolutional layer with a 35-element kernel, a Swish nonlinearity⁵ (Ramachandran et al., 2017), an average pooling layer with kernel size of 3, and a batch normalization layer (Ioffe & Szegedy, 2015). After the sequence of convolutional layers, an adaptive mean pooling layer is used to reduce each convolutional output to a single scalar value (i.e. feature), which is in turn fed to a linear layer that projects a linear combination of the features onto a final logit.

D. Deep Learning-Based Model II: Autoregressive 1D-CNN (AR-1D-CNN)

In order to also give our deep learning-based classification systems access to a longer time history (analogously to the AR- SVM), we implement a deep learning-based autoregressive system that has access to the same time course as the AR-SVM. We further augment this system by not only providing it with the features extracted by the learned convolutional filters, but also with the manually extracted features used by the feature-based implementations. To achieve this, the scalar outputs of the 1D-convolutional network (as introduced above) are concatenated with the spectral features used by the SVMs, and this extended feature set is then fed into a recurrent neural network (LSTM with 2 hidden neurons). The output of the last LSTM cell, which incorporates the relevant time-course information, is fed into a linear layer that computes the final logits (see Figure 3.12). At

⁵ The Swish activation function is described by the mathematical formula

$$swish_{\beta}(x) = x \cdot sigmoid(\beta x) = \frac{x}{1 + e^{-\beta x}}$$

When $\beta = 1$ (as used in this work), the Swish function is known as the Sigmoid Linear Unit (SiLU) function. We chose the Swish activation in lieu of a more traditional activation function (such as ReLU) for two main reasons. First, Swish’s smooth, non-monotonic response helps preserve small negative activations rather than zeroing them out, which in practice may improve gradient flow through the architecture and therefore make training more stable. Second, prior work has demonstrated that Swish yields marginal but consistent performance gains on signal-processing tasks—likely because its gentle curvature more faithfully captures the subtle dynamics present in oscillatory neural data (Ramachandran et al., 2017).

inference time, this system works recursively, i.e. with the LSTM layer keeping track of an internal state within which the relevant historical information is stored and updated. As each new epoch of data is passed through the network, the historical data encoded in the LSTM state influences the current prediction, and this internal LSTM state is updated.

E. Deep Learning-Based Model III: 2D-CNN

Most well-known convolutional networks were developed for computer vision applications, and thus take in 2D signals (i.e. images) as inputs. To use these architectures, it is necessary to make use of mappings that project 1D waveforms to 2D space. This can be achieved through timeseries encoding techniques such as time-frequency decompositions or Gramian angular fields (Z. Wang & Oates, 2015). Here, we use a wavelet-based time-frequency decomposition to extract instantaneous spectral power estimates between 8 and 200 Hz at each time point. The number of cycles of the Morlet wavelets are set to one sixth of the wavelet frequency in Hz (e.g. a 42 Hz wavelet will have 7 cycles). These time-frequency decomposed signals, now with two dimensions per channel, are then fed into a ResNet-based convolutional neural network (please see implementation in code for further architectural details).

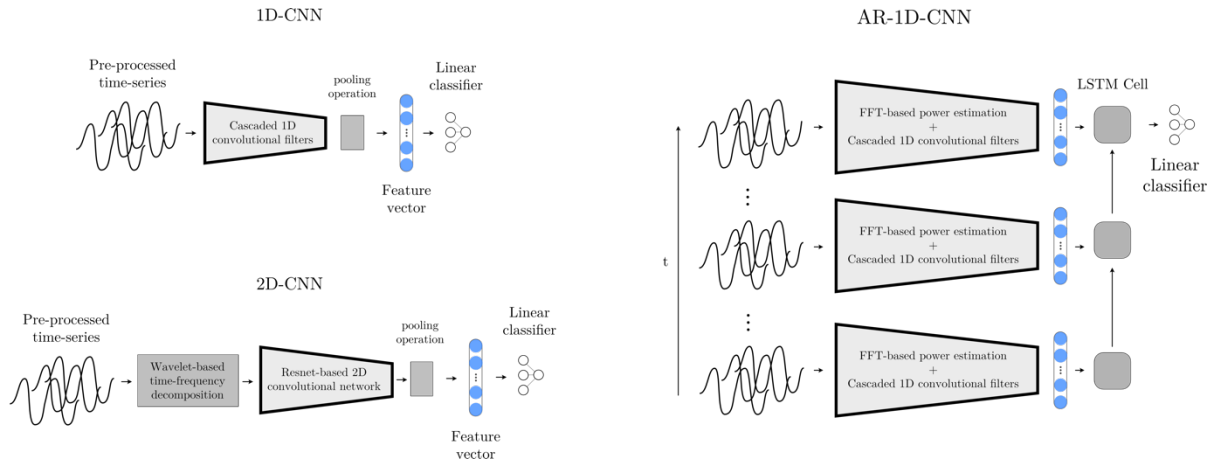


Figure 3.12: Deep learning-based models 1D-CNN, 2D-CNN, and AR-1D-CNN. When using the 1D-CNN approach (top left), the pre-processed multichannel neural time-series are processed by a filterbank of convolutional filters interleaved by non-linearities prior to pooling and linear projection onto a logit. The 2D-CNN (bottom left) works similarly, but the multichannel input time-series are time-frequency decomposed into 2D-signals that resemble images using wavelets. The network consists of 2D kernels that operate simultaneously in the time and frequency dimensions. The AR-1D-CNN (right) works similarly to the 1D-CNN in terms of the way in which it treats individual epochs, but this convolutional feature extraction process is encapsulated within an autoregressive model. This autoregressive model learns relationships contained within sequences of feature vectors. At inference time, the current model output depends not only on the features extracted from the current epoch, but also from the information extracted from

previous epochs. The receptive field of the model in the time dimension is thereby augmented. The classifier in this model is given access to, besides the features extracted by the convolutional layers, the manually extracted spectral features.

The deep learning systems, initialised with random weights for each fold, were trained to minimise the binary cross-entropy loss between the model estimates and the true targets using an Adam optimiser (Kingma & Ba, 2017) with a learning rate of 0.001 that was annealed every 50 epochs by a multiplicative factor of 0.1, and weight decay (l_2 -regularization) factor of $1e-5$. Training ran for a total of 150 epochs, with the model parameters being checkpointed at the end of every epoch and, at the end of training, the checkpoint with the best validation performance as chosen as the final trained network state.

To enhance generalization, data augmentation was implemented utilizing two techniques. First, manifold mixup (Verma et al., 2019; Zhang et al., 2018) was applied in logit space to encourage the model to behave linearly between training examples, which can improve robustness. Second, random 1/f noise was added directly to the input waveforms, with root-mean-square (RMS) power drawn from a uniform distribution between zero and 20% of the RMS power of the signal. This simulates real-world noise conditions and helps the model generalise better to unseen data.

The performance metric used to evaluate all models was the balanced accuracy score (Brodersen et al., 2010), which accounts for class imbalance by equally weighting sensitivity (true positive rate) and specificity (true negative rate). This metric is defined as the arithmetic mean of sensitivity and specificity, and, in terms of the number of samples found in each quadrant of a confusion matrix (TP = true positives, TN = true negatives, FP = false positives, FN = false negatives) is defined as

$$balanced_accuracy = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right)$$

3.2.4 Data and Code Availability

All code (written in Python) used to generate the data, train the models and generate the figure is available at <https://github.com/fer-rplazas/feature-extraction-methods>, and the data is available at <https://data.mrc.ox.ac.uk/lfp-et-dbs> (DOI: [10.5287/bodleian:ZVNyvrw7R](https://doi.org/10.5287/bodleian:ZVNyvrw7R)).

3.3 Results

3.3.1 Dataset I: Synthetic Data

To evaluate the performance of different algorithmic approaches in extracting information from waveforms containing various oscillatory patterns, we generated synthetic datasets for each of the seven oscillatory patterns described in the Methods section. For each pattern, we created five datasets at each of 36 signal-to-noise ratio (SNR) levels ranging from 0.2 to 2.0, resulting in a total of 180 datasets per oscillatory pattern. The SNR levels were chosen to simulate a range of realistic noise conditions, from high noise (low SNR) to low noise (high SNR).

The out-of-sample performance of each model was assessed using the balanced accuracy score, which accounts for class imbalance by equally weighting sensitivity and specificity. Figure 3.13 illustrates the average balanced accuracy scores for all five models across the range of SNRs for each oscillatory pattern. The shaded regions represent the standard deviation across the five independent runs at each SNR level.

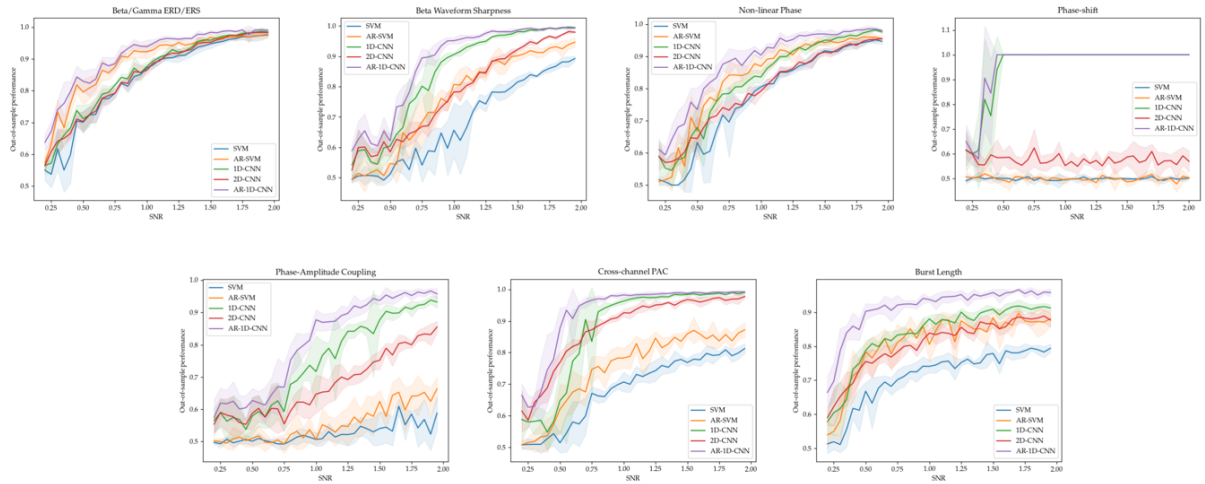


Figure 3.13: Out-of-Sample Performance Across Signal-to-Noise Ratios. For each of the seven oscillatory patterns, the average balanced accuracy score of each model is plotted against the signal-to-noise ratio (SNR).

Shaded regions indicate the standard deviation across five independent datasets at each SNR level.

As anticipated, higher SNRs generally resulted in better model performance across all patterns and models. When distinguishing between states characterised by differences in the relative power of oscillators in the beta and gamma bands, all three non-autoregressive models—the Spectral SVM, 1D-

CNN, and 2D-CNN—performed similarly well. Incorporating autoregressive modelling to increase the receptive field of the models, as seen in the AR-SVM and AR-1D-CNN, further enhanced performance by providing access to a longer time course of data.

Importantly, the 1D-CNN models, which operate directly on raw waveform data without manual feature extraction, performed comparably to the feature-based methods that explicitly use estimates of power in the beta and gamma bands. This suggests that even under noisy conditions and with limited training data, the 1D-CNNs can effectively learn frequency-domain features relevant for decoding.

For states distinguished by the sharpness of the beta waveform, the 1D-CNN models outperformed their feature-based counterparts in both autoregressive and non-autoregressive forms. The 2D-CNN's performance was intermediate between the 1D-CNNs and the feature-based models. This indicates that the 1D-CNNs are particularly adept at capturing waveform shape characteristics directly from the raw data.

In the case of nonlinear phase differences between states, the 1D-CNN models slightly outperformed the feature-based models, with autoregressive versions showing improved performance in both approaches. The 2D-CNN performed similarly to the baseline Spectral SVM. This suggests that 1D-CNNs may be more sensitive to phase-related features not explicitly captured by standard spectral methods.

For states characterised by differences in phase-amplitude coupling, all CNN-based models outperformed the feature-based methods. The performance advantage was more pronounced when PAC occurred within a single channel compared to cross-channel PAC. The 1D-CNN and AR-1D-CNN models were particularly effective, highlighting their ability to capture complex interactions between frequency components.

When distinguishing between degrees of phase shift across channels, the 1D-CNN models achieved near-perfect performance (balanced accuracy approaching 100%) at SNR levels above approximately 0.5. In contrast, the feature-based methods and the 2D-CNN performed at chance level, indicating an inability to learn meaningful relationships for this pattern. This demonstrates the strength of 1D-CNNs in capturing temporal relationships between channels that are not explicitly encoded in traditional spectral features.

In distinguishing between states with long and short beta bursts, the CNN-based models again outperformed the feature-based methods. The autoregressive models, both AR-SVM and AR-1D-CNN, showed improved performance due to their ability to model temporal dependencies. The AR-SVM performed comparably to the non-autoregressive CNNs, suggesting that incorporating temporal information is crucial for capturing burst dynamics.

Table 3-1 summarises the average balanced accuracy scores across all SNR levels for each oscillatory pattern and model. The standard deviations represent the mean of the standard deviations at each SNR level. To quantify the performance differences, we conducted paired t-tests comparing each model's performance to that of the baseline Spectral SVM.

Table 3-1: Average model performance (balanced accuracy score) across SNRs

Oscillatory Pattern	SVM	AR-SVM	1D-CNN	2D-CNN	AR-1D-CNN
Beta/Gamma ERD/ERS	0.84 ± 0.020	0.89 ± 0.018	0.86 ± 0.017	0.85 ± 0.016	0.91 ± 0.017
Beta Waveform Sharpness	0.68 ± 0.032	0.76 ± 0.034	0.84 ± 0.026	0.78 ± 0.018	0.87 ± 0.019
Non-linear Phase	0.78 ± 0.022	0.84 ± 0.026	0.84 ± 0.022	0.80 ± 0.018	0.88 ± 0.022
Phase-Amplitude Coupling (PAC)	0.52 ± 0.026	0.54 ± 0.036	0.73 ± 0.043	0.66 ± 0.035	0.79 ± 0.033
Cross-channel PAC	0.68 ± 0.026	0.75 ± 0.034	0.87 ± 0.035	0.88 ± 0.018	0.93 ± 0.021
Phase-shift	0.50 ± 0.01	0.50 ± 0.016	0.96 ± 0.024	0.58 ± 0.038	0.96 ± 0.016
Burst Length	0.71 ± 0.025	0.81 ± 0.036	0.84 ± 0.021	0.81 ± 0.024	0.91 ± 0.019
Average cross- pattern performance	0.66 ± 0.023 (n/a)	0.72 ± 0.028 (34.55)	0.84 ± 0.027 (41.42)	0.76 ± 0.024 (44.72)	0.89 ± 0.021 (56.64)

(paired t-test vs. SVM)					
----------------------------	--	--	--	--	--

Note: Values are mean balanced accuracy $\pm$ mean of standard deviations across SNR levels. The highest performance for each pattern is highlighted in bold.

Overall, the AR-1D-CNN provided the best decoding accuracy across all simulated scenarios, followed by the 1D-CNN and then the 2D-CNN. The paired t-tests indicated that the performance improvements over the Spectral SVM were statistically significant for all models ($p < 0.001$).

3.3.2 Dataset II: LFP-Based Movement Decoding

To assess model performance on real neural data, we trained the models on datasets recorded from eight participants with essential tremor, each performing two to three different movement tasks. Utilizing a five-fold cross-validation strategy, this resulted in a total of 200 folds (datasets). The tasks involved movements such as pegboard insertion, posture holding, and rice pouring, as detailed in the Methods section.

Figure 3.14 presents pairwise out-of-sample performance comparisons between models for each individual fold. Scatter plots show the performance of two models against each other at each fold, with the diagonal line indicating equal performance. The rightmost column displays the distribution of balanced accuracy scores for each model, revealing that the AR-1D-CNN distribution is skewed toward higher performance.

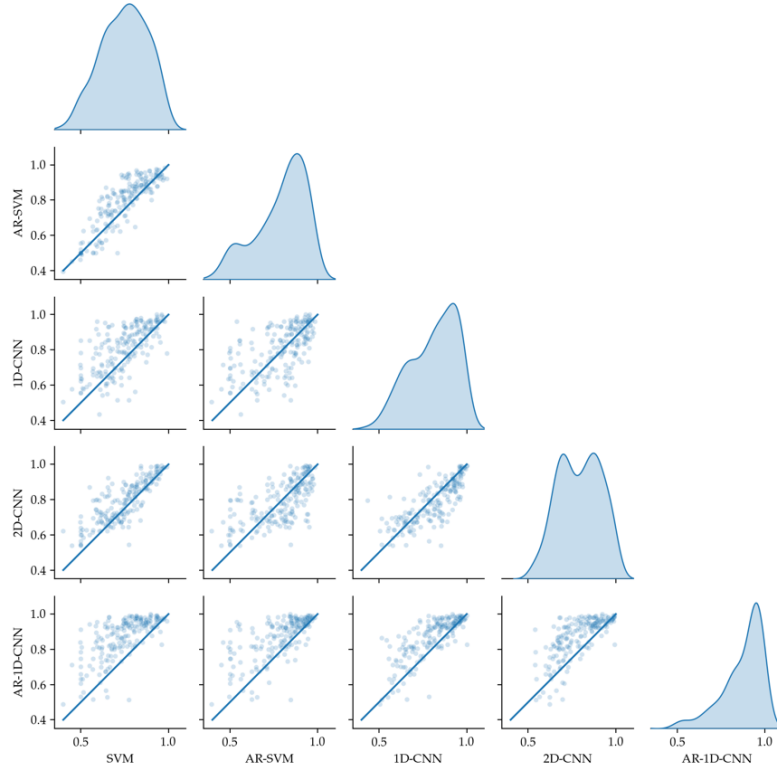


Figure 3.14: Pairwise Model Performance Comparison. The pairwise plots represent the out-of-sample performance of different models across individual cross-validation folds. In the scatter plots, each dot represents the balanced accuracy achieved by two models on a single fold, with the performance of one model on the x -axis and the other on the y -axis. The solid diagonal line indicates equal performance between the two models ($y = x$). Points above the line signify folds where the model on the y -axis outperformed the model on the x -axis, while points below indicate the opposite. The rightmost column displays the distribution of balanced accuracy scores for each model across all folds. The AR-1D-CNN model exhibits the highest overall performance, as evidenced by its distribution being skewed towards higher balanced accuracy values compared to the other models.

Table 3-2 summarises the average model performances, both on an individual fold basis and as cross-fold averages. We also report the results of paired t-tests comparing each model's performance to that of the baseline Spectral SVM.

Table 3-2: Average Model Performances on LFP-Based Movement Decoding

	SVM	AR-SVM	1D-CNN	2D-CNN	AR-1D-CNN
--	-----	--------	--------	--------	-----------

Model performance using individual folds: mean $\pm$ std. deviation (paired t-test vs. SVM)	0.74 ± 0.13 (n/a)	0.79 $\pm$ 0.14 (t=8.13, p< 1e-10)	0.80 $\pm$ 0.13 (t=8.33, p< 1e-10)	0.78 $\pm$ 0.16 (t=8.24, p< 1e-10)	0.87 $\pm$ 0.10 (t=16.7, p< 1e-10)
Cross-fold metrics (mean $\pm$ std) for each model (paired t-test for cross-fold average vs. SVM)	0.75 $\pm$ 0.064 (n/a)	0.79 $\pm$ 0.08 (t=4.96, p=1.38e-5)	0.81 $\pm$ 0.062 (t=5.46, p=2.85e-6)	0.79 $\pm$ 0.054 (t=5.42, p=3.3e-6)	0.87 $\pm$ 0.065 (t=11.17, p=1e-13)

The AR-1D-CNN achieved the highest average balanced accuracy of 0.87 on individual folds, significantly outperforming all other models ($p < 1e-10$ compared to Spectral SVM). The baseline Spectral SVM yielded the lowest performance with an average balanced accuracy of 0.74. The AR-SVM, 1D-CNN, and 2D-CNN exhibited similar performances, with average balanced accuracies of 0.79, 0.80, and 0.78, respectively. Pairwise comparisons indicated that the 1D-CNN slightly outperformed both the AR-SVM ($t = 2.24$, $p = 0.026$) and the 2D-CNN ($t = 2.89$, $p = 0.004$).

Consistent with the results observed in the synthetic data, autoregressive models outperformed their non-autoregressive counterparts, highlighting the importance of incorporating temporal context. Moreover, the deep learning-based models outperformed the feature-based models when the enhanced receptive field due to autoregressive modelling was accounted for.

3.4 Discussion

3.4.1 Understanding Performance Differences Between the Feature-Based and Deep Learning-Based Approaches

In the context of adaptive deep brain stimulation (DBS) and other brain-computer interface (BCI) systems, it might be tempting to conclude that one approach—either feature-based or deep learning-based—is inherently superior to the other. However, we do not advocate this view, despite the fact that, in all our experiments, the AR-1D-CNN model outperformed the other models. We posit that performance differentials can be attributed to specific factors, such as each model's ability to capture, within its feature space, the signal characteristics relevant to a given task. Both feature-based and deep learning-based approaches have their respective advantages and drawbacks, which should be carefully considered when designing BCI or adaptive DBS systems.

To illustrate this point, we examine the performance differences between feature-based and convolution-based models when attempting to distinguish between two specific types of synthetic signals: degrees of phase shift and beta/gamma power changes.

- **Phase Shift:** Consider the task of distinguishing between degrees of phase shift. The spectral SVM, which estimates power in canonical frequency bands using a periodogram, inherently discards phase-related information. Since the extracted feature space does not encode any phase information, it is reasonable to expect that the performance of this model will not surpass chance level when the modulatory pattern of interest is purely phase related. In contrast, the 1D-CNN implements a cascade of finite impulse response (FIR) filters capable of encoding and modelling phase information and thus performs exceptionally well on this task.
- **Beta/Gamma Power Changes:** In this case, all relevant activity occurs within two frequency bands—beta and gamma—that are well captured by the manual feature extraction process. When the information related to the states of interest is adequately captured in the feature space provided to the classifiers, there is no additional benefit to using deep learning methods, which are computationally more costly.

These two waveform patterns represent extremes in the manually extracted feature set's ability to model the signal characteristics of interest. In the phase shift case, changes in the relevant signal characteristics do not cause any alteration in the spectral features; thus, the classifier remains agnostic to these changes. Conversely, for beta/gamma power changes, alterations in the relevant signal characteristics are entirely reflected in the feature space, allowing the classifier to easily capture this information.

Other scenarios—such as waveform sharpness, nonlinear phase, phase-amplitude coupling, and burst length—lie somewhere along this spectrum. These features will also lead to changes in the power spectra and will therefore be partially captured by linear transformation methods such as the fast Fourier transform (FFT), even if the predetermined frequency bands have not been specifically tailored to capture these signal characteristics. In these cases, the spectral SVM shows a limited ability to capture the states of interest. In contrast, convolution-based systems, due to their ability to extract arbitrary features from time series with fewer constraints, tend to outperform the feature-based models.

We further point out that a dichotomous perspective on feature-based and deep learning-based systems can be misleading. Once a convolutional system has been trained, the learned features can be extracted—computed through convolutions or otherwise approximated—and used as inputs in a feature-based method employing an SVM for classification, potentially yielding equivalent results. This implies that when a performance differential between feature-based and deep learning-based methods is observed, a feature set exists that could bring the feature-based performance up to par with that of the deep learning-based methods. However, this does not negate the difficulty of identifying and efficiently constructing such a feature space.

Therefore, we argue that the critical consideration is not the choice of machine learning approach per se (feature-based versus deep learning-based) but rather the design of the feature space. Deep learning methods—also referred to as feature or representation learning—can thus be viewed as automated procedures for feature learning (Goodfellow et al., 2016).

In this work, we are interested in studying the differences between manually extracted feature sets and learned feature sets extracted through convolutions. Our results demonstrate that convolutional neural networks (CNNs) are capable of utilizing various features previously observed in neural activities. In our implementation of the AR-1D-CNN, we provide the classifier with access to both the manually

extracted spectral features and the convolutional features learned during the training process. This raises the question of which feature set is more relevant for performing the classification.

Understanding this could indicate the relative usefulness of the two feature sets.

To investigate this, we first trained the same network on the same data twice: once with and once without the manually extracted spectral features. We then analysed networks trained with the spectral features by selectively upregulating and downregulating the activations corresponding to spectral and convolutional features while measuring the response in activation space. Through this procedure, we determined whether the network had learned to preferentially attend to the manually extracted features or the learned convolutional features, and which feature set contributed more to the decision-making process within the network.

We illustrate this in Figure 3.15, using AR-1D-CNN models trained to distinguish between states of short versus long beta bursts in synthetically generated data. We observed no degradation in performance when the model was not given access to the manually extracted features (Figure 3.15A), indicating that access to the manual feature set is not necessary for the AR-1D-CNN to solve the task or to achieve high performance. Furthermore, activations in response to an upregulation of convolutional feature inputs were higher than those in response to an upregulation of manually extracted features (Figure 3.15B), suggesting that the network had learned to assign greater weight to the convolutional features in its decision-making process.

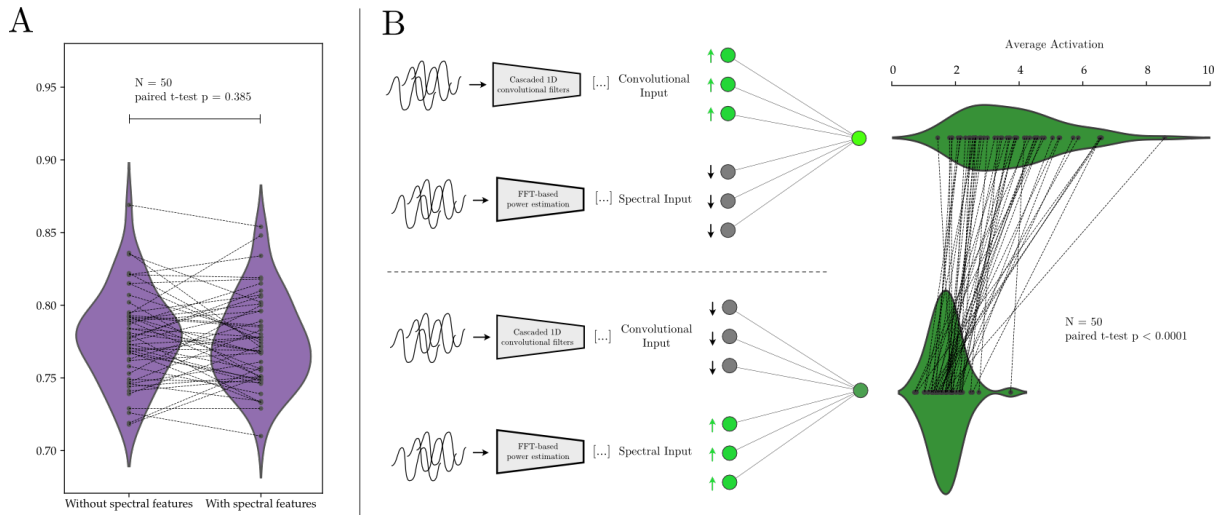


Figure 3.15: Relative Weight of Manually Extracted vs. Learned Convolutional Features. (A) The AR-1D-CNN model was trained on the same data with and without access to the spectral features. The average performance was equivalent in both scenarios (paired t-test, $p = 0.385$), indicating that the manually extracted features were not

essential for the task. (B) After training, we selectively upregulated the inputs corresponding to the convolutional features and the spectral features, measuring the resulting activations. The network showed significantly higher activations in response to the convolutional features compared to the spectral features (activation difference, paired t -test, $p < 0.0001$), demonstrating that the learned convolutional features contributed more to the network's output.

The findings from this study have practical implications. For tasks in which there is a good understanding of the modulatory patterns of interest, and these patterns can be quantified well in real time, manually extracting features and combining them with simple machine learning methods should yield satisfactory performance. In such scenarios, it is unlikely that using deep learning-based methods would provide substantial additional benefit.

Conversely, in tasks where the modulatory patterns of interest are not well known or are difficult to quantify, deep learning-based methods can exploit different idiosyncratic features contained within the oscillatory activities, thereby yielding performance increases, as we have demonstrated in this study. Deep learning models are adept at discovering and harnessing complex, non-obvious patterns in data without explicit feature engineering.

In summary, deep learning-based methods perform well at feature discovery. They do not offer additional benefits when the features of interest are known and can be appropriately captured by a manually designed feature extraction process. However, when the patterns within the data are unknown, not readily quantifiable, or when there are numerous potential features interacting to contribute to decoding—making it challenging to optimally extract all requisite features—deep learning-based methods may outperform feature-based methods.

3.4.2 Exploring the Representations Learned by Convolutional Neural Networks

In this project, we have demonstrated that convolutional neural networks (CNNs) can extract features previously identified as important in electrophysiology, such as power increases in specific frequency bands, phase shifts, and phase-amplitude coupling in neural oscillations. We also reported that features learned by CNNs can lead to higher decoding performance compared to a baseline method using a rudimentary set of manually extracted spectral features. However, we have not yet delved into the potentially novel features that CNNs can learn from data, which may contribute to the observed increase in decoding performance shown in Figure 3.14 and Table 3-2.

Understanding the specific features learned by CNNs is critically important for two primary reasons. First, in the context of BCI systems, adaptive deep brain stimulation (DBS), and clinical neuroscience more broadly, it is essential to comprehend the signal characteristics that autonomous systems depend on and react to. Without this knowledge, it becomes challenging to anticipate and prevent failure modes of these systems, assess their robustness, and evaluate the risks they may pose in real-world settings. Second, elucidating the features learned by machine learning algorithms could offer new insights into the functional physiology and pathophysiology of neural circuits in both healthy and diseased states, providing a novel and valuable data-driven research tool.

Despite the significance of this endeavour, we find a notable lack of well-established methodological tools to systematically address the question of what features CNNs learn from electrophysiological data. This limitation leads us, in agreement with other researchers in the field (Merk, Peterson, Köhler, et al., 2022), to consider neural networks as systems with limited interpretability at this stage. Although considerable research has been conducted to improve the interpretability of neural networks—particularly in the field of computer vision (Chattopadhyay et al., 2018; Desai & Ramaswamy, 2020; Selvaraju et al., 2020; H. Wang et al., 2020)—these methods are not directly transferable to the analysis of electrophysiological time-series data.

Electrophysiological signals possess distinct properties compared to images, such as temporal dynamics, non-stationarity, and the critical importance of phase information. Consequently, interpretability methods developed for spatial data, like images, may not directly apply to time-series analysis. For example, techniques like gradient-weighted class activation mapping (Selvaraju et al., 2020) rely on spatial localization of features in images, which does not translate seamlessly to temporal data where timing and frequency content are the main signal characteristics of interest.

Moreover, neural data exhibit complex multiscale temporal and spectral characteristics. Important features may be embedded in cross-frequency interactions or temporal patterns that span multiple timescales. Interpreting how CNNs capture these features requires methods that can disentangle such intricate dependencies and represent them in a meaningful way.

To illustrate the challenges in interpreting the features learned by CNNs in our context, we conducted an analysis using a 1D-CNN trained to distinguish between periods of synthetically generated short versus long beta bursts. Figure 3.16 depicts the individual kernel elements and their corresponding frequency responses before and after training. Since 1D-CNN kernels can be interpreted as finite

impulse response (FIR) filters, examining their frequency responses can provide some insight into the filtering characteristics learned by the network.

In Figure 3.16A, we observe that during training, the convolutional filters are modified to increase resonance at lower frequencies, which is expected given that the relevant activity occurs in the beta band (approximately 13–30 Hz). The initial random filters evolve to emphasise frequency components associated with the task. However, this analysis provides limited information about the higher-level features learned by the network, as the convolutional layers are interleaved with non-linear activation functions (Swish functions), resulting in complex, non-linear transformations of the input data.

Figure 3.16B presents the outputs of the CNN after the last convolutional layer prior to classification. Through the sequence of convolution operations and non-linearities, the network learns to extract features (represented by the green and pink lines) that activate preferentially during the two different states (short and long beta bursts). A linear combination of these features yields the final classifier output (red line). Despite the linear nature of convolutional filters, the interleaved non-linear operations lead to features that exhibit non-linear behaviour. These learned features are activated by underlying signal characteristics that are themselves non-linear, making it challenging to directly interpret them in terms of traditional electrophysiological measures.

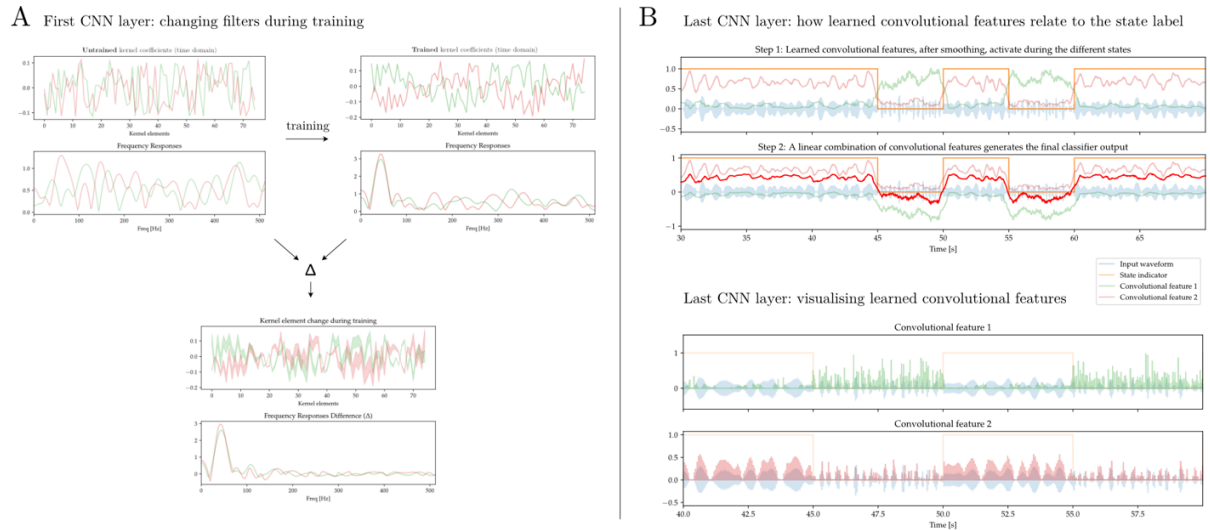


Figure 3.16: 1D-CNN Filters and Activations. (A) Two convolutional filters are shown along with their frequency responses before and after network training (top). The changes in individual kernel elements and frequency responses after training (bottom) indicate that the filters have adapted to emphasise the beta frequency band. (B) Outputs of the CNN after the last convolutional layer prior to classification. Two features (green and pink lines) activate preferentially during different states, and their linear combination yields the classifier output (red line, top). The

features exhibit non-linear behaviour due to the interleaved non-linear activation functions, making direct interpretation challenging (bottom).

The difficulty in interpreting these learned features highlights the limitations of current methods for understanding deep learning models in the context of electrophysiological data. While techniques such as saliency maps, layer-wise relevance propagation, and attention mechanisms have been developed to interpret CNNs in image recognition tasks (Montavon et al., 2018; Samek et al., 2019), their application to time-series data, especially in neuroscience, remains less developed.

3.4.3 Choosing Time-Windows and Autoregressive Parameters

When training classification systems for neural decoding, an important hyperparameter to consider is the length of each epoch used for analysis. Longer epochs can reduce statistical noise and allow for more stable feature estimation, but they come at the cost of decreasing the system's output rate and introducing inertia, which can slow the classifier's response to changes in the underlying signals. Although a target output rate (e.g., 10 classification outputs per second, or a 10 Hz classification rate) can be maintained by overlapping epochs, unwanted inertia or reduced decoding accuracy may still be introduced if the epochs are longer than the brain states we aim to detect, even when overlapping is employed.

In this work, the autoregressive models introduce an additional hyperparameter: the sequence length, defined as the number of trailing epochs used at each classification sample during training of the AR-SVM and AR-1D-CNN. When comparing autoregressive and non-autoregressive models, these parameters can also be conceptualised in terms of whether an epoch should be subdivided into frames—with each frame treated as an element in a sequence within an autoregressive scheme—or whether epochs are best kept as single frames without subdivisions.

To explore the effect of these design parameters on decoding performance, we trained SVMs, 1D-CNNs, AR-SVMs, and AR-1D-CNNs on the task of distinguishing between periods of synthetically generated short versus long beta bursts. Figure 3.17A illustrates that longer time windows are associated with improved decoding performance for both the SVM and 1D-CNN. We attribute this improvement to more stable feature estimation due to reduced noise and better sampling of each estimate. In our experiments, performance degradation was particularly pronounced for time windows shorter than 500 ms. This may be because such short windows cannot effectively capture the bursting

dynamics of oscillations, which require access to enough oscillation cycles to constitute a full long burst (each oscillation at a beta frequency of 20 Hz has a period of 50 ms).

Figure 3.17B shows the effect on decoding performance of subdividing a 5-second epoch into frames for the AR-SVM and AR-1D-CNN. We used two subdivision strategies: overlapping and non-overlapping frames. When using non-overlapping frames, performance degraded rapidly with increasing subdivisions for both the CNN and SVM. This indicates, consistent with the results shown in Figure 3.17A, that the models display a preference for longer time windows. Furthermore, this suggests that autoregressive schemes are not able to seamlessly fuse the information contained in feature vectors from adjacent frames into the equivalent feature vector that would be extracted from a single, longer frame. This observation holds for both the AR-SVM and AR-1D-CNN. We attribute this to improved stability of feature estimation when the frames used to compute each feature are longer.

When adjacent frames are overlapped, this allows for longer individual frames from which to extract features, resulting in more stable feature estimation. With this strategy, the performance degradation associated with subdividing epochs into frames can be mitigated. However, in our experiments, we did not observe any performance benefits from subdivision strategies, whether with overlapping or without, when compared to using whole epochs.

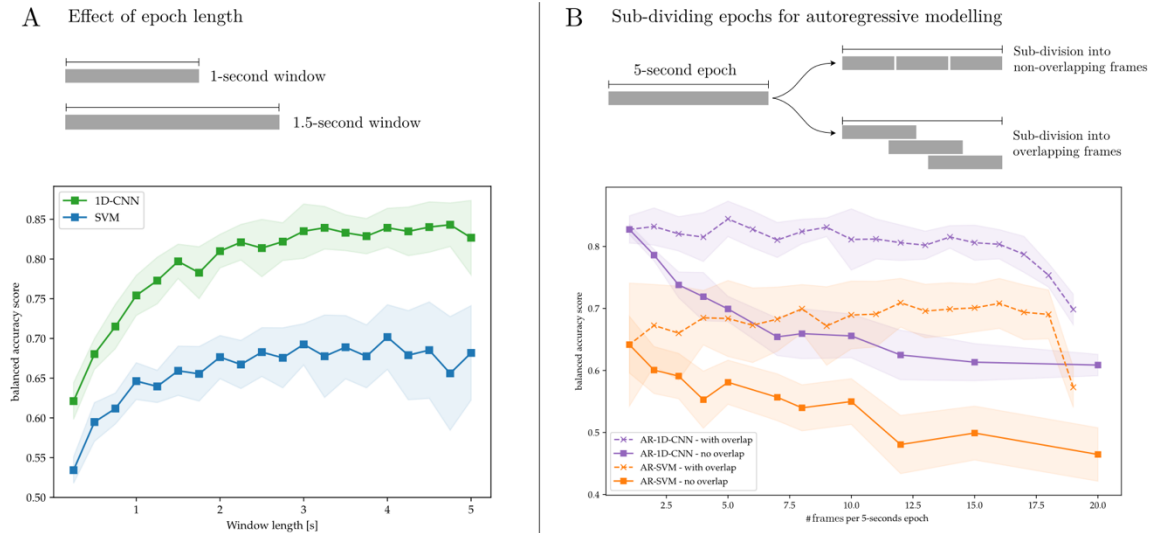


Figure 3.17: Effect of Time Windows on Decoding Performance. (A) Longer time windows are associated with improved decoding performance for both the SVM and 1D-CNN. Overlapping long time windows allow for more stable feature estimation without compromising the classifier's output rate. (B) A 5-second epoch is subdivided into overlapping and non-overlapping frames, which are treated as sequence elements in autoregressive models. Without

frame overlap, performance degrades with increasing subdivision, likely due to decreased feature stability from shorter computation epochs. Overlapping frames help avoid this performance degradation.

While we have experimentally demonstrated the benefits of using longer time windows and attributed the improvement in performance to more stable feature estimation, we have not addressed the potential detrimental effects of using epochs that are too long. The synthetic data used here, modelled as a first-order Markov process, has stochastic yet stable state dynamics with known durations, with each state lasting on average 10.92 ± 7.94 s. This may not be the case in arbitrary real-world data. If the underlying brain state dynamics are complex and rapidly changing (Vidaurre et al., 2017), long epochs might average out relevant state transitions, obscuring important temporal information.

Therefore, the balance between sufficiently long time windows to enable stable feature estimation and short enough windows to model state transitions in neural activities at the appropriate time scale must be carefully considered on a case-by-case basis. Optimal epoch length may depend on the specific characteristics of the neural signals, the temporal dynamics of the brain states of interest, and the requirements of the application, such as the desired responsiveness of an adaptive DBS system.

3.4.4 Inference-Time Computational Cost

A critical consideration for the real-world implementation of adaptive deep brain stimulation (aDBS) systems is ensuring that the proposed algorithms are compatible with the constraints imposed by implantable hardware, such as limited computational resources and restricted power consumption. These constraints pose significant challenges for deploying deep learning (DL) methods in such settings. Deep learning models, including those described in this study, generate predictions by performing a cascade of numerical computations, oftentimes matrix multiplications and convolutions, which are computationally intensive and can be resource-demanding during inference.

In this work, our primary objective was to explore and compare the modelling capabilities of different machine learning-based methods, rather than to optimise performance-to-compute ratios, or to design systems optimised for low computational cost. Consequently, we did not focus on minimising the computational footprint of the models when we implemented them. However, it is important to acknowledge that the computational demands of models can impact their feasibility for deployment in implantable devices.

As shown in Figure 3.18, the 1D-CNNs require approximately one order of magnitude more computational resources at inference time when compared to their SVM counterparts (given the fixed number of features considered here). The 2D-CNNs, in turn, demand another order of magnitude of additional computational resources when compared to the 1D-CNNs. This substantial increase is due to two main factors: the initial wavelet-based time-frequency decomposition required before inputting the data into the network, and the need for deeper network architectures with small kernels to achieve a sufficiently large receptive field in 2D convolutional networks designed for computer vision tasks.

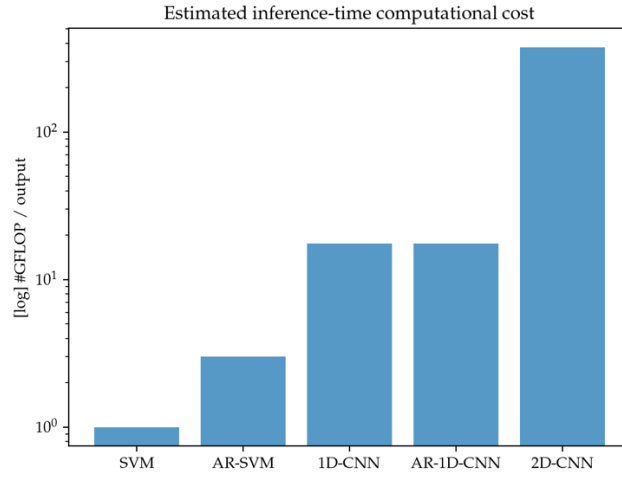


Figure 3.18: *Estimated Inference-Time Computational Cost.* The estimated inference-time computational cost (measured in the number of floating-point operations per decoding output, log scale) is shown for each of the methods implemented in this study. The 2D-CNN exhibits the highest computational cost due to the initial wavelet-based time-frequency decomposition and the deeper network required to achieve a sufficient receptive field with the small kernels used in computer vision architectures.

Techniques exist to reduce the computational footprint of neural network models. One such technique is model of knowledge distillation (Ba & Caruana, 2014; Hinton et al., 2015), which involves training a smaller model using rich logit-space representations learned by a larger model (or an ensemble of large models). Since the logits computed at the output of the large models (prior to applying the *SoftMax* operation) have higher entropy – and therefore contain more information – than binary labels, the student model can learn from these labels directly (see Supplementary Figure IIIA). A sweep of various compression ratios, together with the resulting effect on decoding performance, is visualised in Supplementary Figure IIIB. We implemented this technique to test the feasibility of implementing

smaller versions of the 1D-CNN model that operate on down-sampled versions of the neural time series signals while preserving relative performance.

In future work, it is essential to explore the trade-offs between computational efficiency and model performance, especially in the context of aDBS systems where computational resources are limited. Reconciling the computational requirements of high-performing algorithms with the constraints of implantable hardware should be explicitly explored and relevant trade-offs should be documented.

3.4.5 Limitations and Future Work

Understanding the precise computations performed by a trained convolutional neural network (CNN) to map inputs (e.g. time-series signals in the case of our 1D-CNNs) to outputs (e.g. class labels) remains a challenge. Although numerous feature attribution techniques have been developed for CNNs—primarily in the context of computer vision—they are not readily applicable to systematically analyse entire datasets of electrophysiological time-series data. Moreover, the insights provided by these methods can be non-trivial to interpret, especially when dealing with temporal dynamics and complex neural signals. As a result, many researchers continue to consider CNNs to be, at least partially, “black-box” algorithms.

These attribution methods, often referred to as techniques for investigating or visualizing activations within CNN layers, include approaches like saliency maps, gradient-weighted class activation mapping (Grad-CAM), and layer-wise relevance propagation (Montavon et al., 2018; Samek et al., 2019; Selvaraju et al., 2020). While they have shown promise in providing interpretability in image-based tasks, they do not yet offer systematic approaches to understand features learned by convolutional architectures across time-series datasets. The temporal and spectral complexities inherent in neural data add additional layers of difficulty in interpreting the internal representations learned by CNNs.

We view this lack of insight as a current limitation of using CNNs in research and clinical practice, particularly in neuroscience applications where understanding the underlying neural mechanisms remains an important methodological consideration. In the field of machine learning, there is often an emphasis on optimising predictive performance, with less focus on elucidating the internal workings of algorithms. Within this paradigm, the existence of an actual mapping function $f_{actual}: input \rightarrow output$ is assumed, and the role of the algorithm is to find a function f_{approx} that minimises the error

$E\{f_{actual}(input), f_{approx}(input)\}$. Developing an understanding of how different algorithms process information internally is considered secondary to improving performance.

However, in neuroscience—and especially in translational neuroscience—we advocate for a different approach. Thoroughly understanding the signals that our systems use is crucial for ensuring the robustness and generalization of these systems beyond controlled clinical settings. Without this understanding, it is difficult to anticipate potential failure modes and assess safety. Further research is needed to develop methodological tools that enable a better understanding of the decision-making processes of deep learning algorithms in the context of neural decoding. This includes creating interpretability methods specifically tailored to time-series neural data, which can systematically analyse and visualise the features learned by machine learning systems including CNNs. Additionally, incorporating domain knowledge into model architectures and training processes may enhance interpretability. For example, using physiologically inspired models or constraining the network to learn features that align with known neural phenomena could facilitate a more transparent mapping between learned representations and neural mechanisms.

3.5 Conclusion

In this work, we explored different decoding methods in terms of their ability to identify states from neural oscillations, benchmarking their performance on both simulated data and LFP data recorded from externalised Essential Tremor patients. We were able to show that end-to-end convolutional neural networks can make use of different spectral and time-domain features present in oscillatory activity including power changes, waveform sharpness, phase-amplitude coupling, burst duration, phase shifts, and non-linear phase characteristics. 1D-CNNs trained end-to-end can yield superior performance when compared to feature-based approaches, especially in cases where the underlying features of interest are not well-known or not easily captured by standardised feature extraction pipelines.

3.6 Summarised Findings

The aim of this chapter was to implement and benchmark a class of machine learning models—specifically convolutional neural networks (CNNs)—that, although widely adopted to improve performance in many signal processing applications, are still not widely utilised in the field of neural decoding. We sought to evaluate their effectiveness in identifying neural states from oscillatory brain signals, comparing them to traditional feature-based approaches using both simulated data and real local field potential (LFP) recordings from patients with essential tremor.

The main findings of this chapter are as follows:

- **CNNs can capture relevant signal characteristics:** Convolutional neural networks demonstrated the ability to capture a wide range of relevant signal characteristics inherent in neural oscillations. These include power changes across frequency bands, waveform sharpness, phase-amplitude coupling, burst duration, phase shifts, and nonlinear phase dynamics. By learning directly from raw time-series data, CNNs can model complex relationships that may not be readily apparent or easily extracted through manual feature engineering.
- **Improved neural state decoding performance:** The application of one-dimensional CNNs (1D-CNNs), trained end-to-end, resulted in an average superior neural state decoding performance compared to feature-based methods. This improvement was consistent across both synthetically generated waveforms and real neural recordings. The CNNs effectively utilised the rich temporal and spectral information present in the data to achieve higher accuracy in identifying neural states.
- **Harnessing diverse signal features:** The contributing signal characteristics that these models harnessed were diverse and not limited to predefined features. The CNNs automatically discovered and utilised pertinent features within the oscillatory data without explicit feature design. This autonomous feature learning enables the models to adapt to various signal properties and complexities inherent in neural data.
- **Advantageous in scenarios with unknown or complex features:** The use of CNNs proved particularly advantageous when the characteristics of interest within the oscillatory data were not well-known or were difficult to capture through standardised feature extraction.

pipelines. In such cases, the ability of CNNs to learn representations directly from raw data allowed for improved decoding performance where feature-based methods might struggle due to incomplete feature sets that failure to capture the modulatory patterns of interest.

4 Project 2: Anticipating Movement Initiation by Decoding Thalamic LFP Signals in Tremor Patients

4.1 Introduction

Since receiving FDA approval for the treatment of Essential Tremor (ET) in 1997, Deep Brain Stimulation (DBS) of the ventral intermediate nucleus (ViM) of the thalamus has established itself as an effective therapy for the treatment of tremor disorders in patients with medication-refractory symptoms (Dallapiazza et al., 2019; Ferreira et al., 2019; Kremer et al., 2021). In a meta-analysis comprising 1,714 Essential Tremor (ET) patients, VIM-DBS improved tremor scores by a mean 61.3% at 20 months follow-up, showing a sizeable therapeutic effect in terms of tremor suppression (Lu et al., 2020).

Despite its efficacy for tremor management, VIM-DBS is associated with several adverse effects, including stimulation-induced side effects that often impact speech and postural stability, as well as a gradual loss of therapeutic efficacy over time (Alomar et al., 2017; Barbe et al., 2011; Becker et al., 2020; Fasano & Helmich, 2019; Pahwa et al., 2006; Paschen et al., 2019; Peters & Tisch, 2021; Petry-Schmelzer et al., 2021). To address these challenges, non-continuous or adaptive DBS (aDBS) schemes have been proposed as a means to mitigate stimulation-related side effects and thereby increase the therapeutic window (Cernera et al., 2021, 2024; Meidahl et al., 2017; Opri et al., 2020).

A critical aspect in the development of aDBS systems is the selection of an appropriate feedback signal to modulate stimulation in real time. Various signals have been explored for driving aDBS in ET patients, including muscle activity recorded from surface electromyography (Cernera et al., 2021; Graupe et al., 2010; Yamamoto et al., 2013), scalp electroencephalography (EEG) (J. Herron et al., 2015), and signals recorded from intracranial electrodes (Ferleger et al., 2020; J. A. Herron et al., 2017; Opri et al., 2020). However, there is a growing interest in harnessing modulatory patterns found in neural signals recorded from the implanted DBS electrodes themselves to avoid the need for supplementary implants or external devices (Priori et al., 2013). By capturing the neural correlates of movement through local field potentials (LFPs) measured from the thalamus, it is possible to detect movement states and utilise this information to drive an aDBS system (Tan et al., 2019). This

approach has been successfully implemented in real-time in an acute clinical setting in prior studies (Buijink et al., 2022; Castaño-Candamil et al., 2020; Ferleger et al., 2020; He et al., 2021).

In the context of adaptive deep brain stimulation (aDBS) for patients with intention tremors—tremors that manifest during voluntary movement—the timing of stimulation initiation can be critically important. Early triggering of stimulation, specifically initiating stimulation prior to or at the very onset of voluntary movement, holds significant therapeutic potential. By delivering stimulation before the tremor becomes clinically apparent, it may be possible to prevent the tremor from emerging altogether. This pre-emptive approach contrasts with reactive stimulation paradigms that activate in response to tremor after it has already manifested, which may inherently involve a delay sufficient for symptoms to impact the patient. Early stimulation could modulate the neural circuits involved in tremor generation at a stage when they are more susceptible to intervention, potentially suppressing the pathological oscillations that give rise to tremor activity. By intercepting the neural processes leading to tremor, this strategy could reduce or eliminate tremor episodes before they begin, thereby enhancing therapeutic outcomes. Achieving this, however, requires systems capable of decoding the onset of movement in real time by detecting neural correlates present in the period leading up to movement execution. Early detection of movement from neural signals has been demonstrated, with electrocorticography (ECoG) showing higher decoding performance than LFP signals recorded from the subthalamic nucleus in Parkinson's disease patients (Köhler et al., 2024).

In this study, we aim to benchmark our ability to train machine learning models to decode early correlates of upper limb movement during the pre-movement period, using LFP signals recorded from thalamic DBS electrodes, and comparing the decoding performance against that based on scalp EEG electrodes. To this end, we develop and implement several classification algorithms, utilizing both pre-defined features, such as power estimates in specific frequency bands, and features derived through deep learning techniques. We compare the performance of these models and analyse their outputs to understand the neural correlates that inform their predictions at the cohort level.

The remainder of this chapter is structured as follows: we first introduce the dataset, feature extraction pipeline, and classification algorithms used in this study. We then present the decoding performance achieved by these systems and interpret the neural correlates they capture to make inferences. Finally, we discuss the trade-offs of the different approaches and their implications for early movement-driven aDBS in tremor disorders.

4.2 Methods

4.2.1 Participants and Experimental Protocol

Data were collected from 11 patients (5 female) diagnosed with movement disorders and scheduled for deep brain stimulation (DBS) surgery targeting tremor management. The cohort included nine patients with Essential Tremor (ET), one with Orthostatic Tremor (OT), and one with tremor-dominant Parkinson's Disease (PD). Detailed clinical information for each participant is provided in Table 4-1. All procedures were conducted in accordance with the Declaration of Helsinki and approved by the relevant local ethics committees. Informed written consent was obtained from all participants prior to their inclusion in the study.

Table 4-1: Clinical Details of Patients

Patient ID	Diagnosis	Centre	DBS System	Sex	Age [years]	Disease Duration [years]	Predominant Symptoms Before Surgery
S1	ET	SGH	Abb	F	77	21	Tremor, gait ataxia, tremor worse on right, upper limb and voice tremor
S2	ET	SGH	Abb	M	70	8	Tremor, upper limb, with right worse than left, lower limb tremor
S3	ET	SGH	Abb	F	62	45	Tremor, upper limb tremor left worse than right, voice tremor
S4	ET	SGH	Abb	M	70	5	Tremor, upper limb left worse than right
S5	ET	UHC	Med	F	58	15	Tremor in both hands, left hand worse than right
S6	ET	UHC	Med	M	72	10	Tremor in both hands (stronger in right hand), some head tremor
S7	ET	SGH	Med	M	64	20	Tremor
S8	ET	SGH	Med	M	71	16	Tremor

S9	ET	SGH	Med	F	59	50+	Tremor
S10	PD	UHC	Med	M	68	6	Rigidity and tremor in right hand and arm
S11	OT	UHC	Med	F	61	9	Orthostatic and action tremor in both hands

Abbreviations: ET = Essential Tremor; PD = Parkinson's Disease; OT = Orthostatic Tremor; SGH = St. George's Hospital; UHC = University Hospital Cologne; Abb = Abbott Infinity system; Med = Medtronic SenSight system.

Patients underwent stereotactic neurosurgery for bilateral implantation of DBS leads targeting the ventral intermediate nucleus (ViM), zona incerta (ZI), or posterior subthalamic area (PSA) for the treatment of tremor. For research purposes, the DBS leads were temporarily externalised for up to six days post-implantation, allowing for direct recording of neural signals before implantation of the implantable pulse generator (IPG).

Four patients (S1–S4) were implanted with the Abbott Infinity DBS system (Abbott Laboratories), which features segmented directional leads. Each lead contains eight contacts, with the two central levels comprising three directional segments each. During recordings, the three directional contacts at each level were fused to function as a single ring-mode contact to prevent amplifier saturation. This configuration resulted in four ring-mode contacts per hemisphere. Local field potentials (LFPs) were amplified and sampled in monopolar mode using a TMSi Porti amplifier (TMS International), with a sampling rate of 2,048 Hz. Seven patients (S5–S11) were implanted with Medtronic SenSight directional leads (Medtronic plc), also featuring segmented contacts at the two central levels. For these patients, all contacts, including individual directional segments, were recorded in monopolar mode, yielding eight contacts per hemisphere. LFP signals were acquired using a TMSi SAGA amplifier (TMS International) at a sampling rate of 4,096 Hz. A reference electrode was attached to the wrist of each patient to serve as a common reference point for monopolar recordings across all modalities.

Scalp EEG recordings were obtained using electrodes placed at positions Cz, C3, C4, CPz, CP3, and CP4 (following the international 10–20 system), encompassing somatosensory, sensorimotor, and motor cortical areas. EEG signals were recorded synchronously with LFP signals. Surface electromyography (EMG) was recorded using bipolar electrodes placed over the forearm flexor and extensor muscles. EMG signals provided precise information about upper limb movement onset, aiding in the temporal alignment of neural and behavioural data.

During recording sessions, patients performed a series of upper limb motor tasks designed to elicit voluntary movement and engage different muscle groups: rice pouring task (patients poured rice back and forth between two cups), pegboard insertion task (patients inserted pegs into a pegboard), and foam ball gripping task (patients tightly gripped a foam ball with one hand). Different tasks were used to test the generalisability of the proposed method.

Each task was performed multiple times with rest intervals to prevent fatigue. Instructions were standardised, and patients were encouraged to perform movements at a comfortable pace. All signals—including thalamic LFPs, scalp EEGs, and surface EMGs—were recorded using the same amplifier with custom-developed software tailored for high-fidelity electrophysiological data acquisition. The software synchronised data streams across modalities and annotated timestamps for task events. Recorded data were digitised and stored securely for offline processing and analysis, adhering to data protection and privacy regulations.

4.2.2 Data Preprocessing and Labelling

The electromyography (EMG) signals recorded from the forearm flexor and extensor muscles were utilised to determine the timing of movement initiation. For lateralised movements, such as the one-handed gripping task, EMG traces from the arm executing the movement were selected. For non-lateralised movements, which patients executed with both arms, EMG signals recorded from both arms were included for the subsequent analysis.

To enhance the detection of movement onset, the EMG signals underwent a series of preprocessing steps. Raw EMG signals were first high-pass filtered using a causal fourth-order Butterworth filter implemented in second-order sections. To minimise ringing artefacts due to filter warm-up, the initial filter state was adjusted appropriately. All filtering was performed in a forward-only (causal) manner to preserve the temporal integrity of the signals. After high-pass filtering, the magnitude of the EMG signals was computed to obtain the envelope of muscle activity, and z-scored to standardise the amplitude across trials and participants. These processed EMG traces were visually inspected to ensure clear delineation of movement execution, and traces that did not distinctly define and isolate movement onset were discarded to maintain data quality. For tasks involving both arms, the EMG signals from each arm were averaged to create a composite signal representing overall muscle activity. This signal was then smoothed, and a threshold was set at the earliest points at which a deviation

from baseline activity was detected (i.e. movement initiation). The timing of each movement initiation was determined as the moment at which the smoothed EMG trace crossed that threshold. These timings were subsequently reviewed manually to ensure movement was detected at the earliest possible time point. This labelling served as the ground truth for subsequent analyses.

To prevent the introduction of unwanted phase lags between the movement signals and the neural signals, the same high-pass filter applied to the EMG data was also applied to the thalamic local field potential (LFP) and electroencephalography (EEG) signals. This approach ensured temporal alignment across the different kinds of signals. Additionally, to eliminate power line interference and its harmonics, notch filters were applied at 50 Hz, 100 Hz, and 150 Hz using a causal filter design with second-order sections. It should be acknowledged that the high-pass filtering performed during the pre-processing of EEG/LFPs signals may remove activity taking place at very low frequencies, including readiness potentials that have been shown to contribute towards predicting the onset of movement (Köhler et al., 2024).

The movement onset times defined from the EMG signals were used to segment the LFP and EEG data into epochs corresponding to individual movement trials. For participants who performed multiple movement-related tasks (gripping, rice pouring, and pegboard insertion), the epochs from all tasks were combined for analysis.

4.2.3 Time-Resolved Characterization of Oscillatory Activity

To train neural decoders capable of detecting pre-movement periods, we implemented a feature extraction pipeline designed to capture relevant information within electrophysiological signals in a time-resolved manner. The extracted features aim to provide a comprehensive characterization of the relevant oscillatory activity at any given time point.

In order to decode the movement state at each time point, features were extracted from LFP and EEG signals. To this end, signals were first segmented into short-term windows. From each window, various features were extracted and subsequently concatenated along the time dimension, offering a time-resolved representation of the oscillatory activity contained within the signals. This overlapping windowing procedure is analogous to the short-term Fourier transform (STFT), a technique used to compute time-frequency decompositions. To avoid making strong assumptions about the optimal window length for capturing and representing the oscillatory dynamics of interest, we extracted

features using two different window lengths: 0.5 seconds and 1 second. These features were provided to the classifier independently, granting the optimization algorithm access to features extracted over these two different time scales and allowing the classifier to weight them appropriately based on their contribution to detecting the state of interest.

To ensure that our analysis pipeline remained fully causal—that is, no information from future time points was used to make assessments at the present time point—the windows were computed in a trailing manner, i.e. utilizing only data points leading up to the current time point. We set the distance between contiguous windows at 20 milliseconds, which, in a real-time setting, corresponds to a system that outputs estimations at a rate of 50 Hz. This parameter allowed us to define the granularity of the state estimation pipeline, and a short distance between contiguous windows enabled us to capture and train our classifiers with the short-term fluctuations that may be encountered during real-time application.

At each time point, and for each trailing window length, we extracted several features to characterise the oscillatory activity within the electrophysiological signals comprehensively.

F. Power in Canonical Frequency Bands

For each time window of electrophysiological data, we computed a periodogram, which was subsequently log-transformed. Prior to conversion to the frequency domain, a Hann window was applied to the time-domain signals to minimise spectral leakage and mitigate edge artefacts that could distort the spectral estimates.

From the resulting periodogram, we extracted the mean and standard deviation of power within predefined canonical frequency bands. These bands were carefully selected to capture distinct oscillatory activities associated with various cognitive and physiological states. In our configuration, the bands included theta (4–8 Hz), alpha (8–12 Hz), low beta (13–20 Hz), high beta (20–30 Hz), gamma (30–60 Hz, 60–80 Hz, 80–100 Hz), and high-frequency activity (200–500 Hz).

G. Time-Domain Statistics

Time-domain statistics are features that capture the characteristics of the signal directly from its temporal representation. These features provide crucial insights into the underlying neural processes and are often used in the preprocessing or analysis stages of electrophysiological data. The primary time-domain features extracted included the first four statistical moments of the signal distribution as well as the Hjorth parameters.

The **first four statistical moments** are:

- **Mean (First Moment):** Average value of the signal over time. Provides a measure of the central tendency, the overall level of activity within the signal, as well as a measure of low-frequency activity.
- **Variance (Second Moment):** Measures the signal's dispersion around the mean, reflecting the signal's power or energy.
- **Skewness (Third Moment):** Quantifies the asymmetry of the signal's amplitude distribution. A positive skew indicates a distribution with a longer tail on the right side, whereas a negative skew indicates a longer tail on the left. This measure detects deviations from normality in the signal's amplitude.
- **Kurtosis (Fourth Moment):** Describes the "tailedness" or peakedness of the signal distribution. High kurtosis indicates the presence of infrequent, extreme deviations from the mean, often associated with transients or bursts in the signal.

The **Hjorth parameters** are a set of three descriptors useful in the analysis of non-stationary signals, such as LFP, EEG, or MEG data. They provide a compact representation of the signal's dynamic properties:

- **Activity:** Corresponds to the variance of the signal, representing the overall signal power. It reflects the amplitude of the signal fluctuations and is directly related to the signal's energy content.
- **Mobility:** Defined as the square root of the variance of the first derivative of the signal divided by the variance of the signal itself. It quantifies the mean frequency or the rate of change in the signal, with higher mobility indicating faster signal fluctuations.
- **Complexity:** Measures the signal's waveform complexity relative to a pure sine wave, calculated as the ratio of the mobility of the first derivative of the signal to the mobility of the

signal itself. It reflects the degree of variability in the frequency content of the signal, with higher values indicating more complex, non-sinusoidal waveforms.

Together, these time-domain features provided a comprehensive characterization of the electrophysiological signals, capturing essential aspects of the signal's amplitude, variability, and temporal structure. They are commonly used in various applications, including brain-computer interfaces, cognitive state monitoring, and the identification of pathological activity such as epileptic seizures (Alawee et al., 2023; Hag et al., 2023; Hjorth, 1970).

H. Other Features

We also extracted autoregressive (AR) parameters and cepstral coefficients to capture additional signal dynamics.

For the autoregressive parameters, we employed AR modelling to describe a signal as a linear combination of its 10 previous values. The AR parameters were estimated by solving the Yule-Walker equations, which relate the autocorrelation function of the signal to the AR coefficients.

The cepstral coefficients were derived through cepstral analysis, which involves taking the inverse Fourier transform of the logarithm of the signal's spectrum, providing a representation in the *quefrequency* domain. This technique captures the spectral envelope and harmonic structures of the signal, revealing periodicities and resonances that may not be apparent in the original spectrum. By focusing on the rate of change across different spectral components, cepstral coefficients provide a fine-grained representation of variation in the frequency spectrum. This representation is useful for identifying repeating patterns or harmonic relationships in the oscillatory data and may complement the information provided by power estimates in canonical frequency bands. The cepstrum was segmented into ten bands, and the power within these bands was computed.

Once all features were extracted, they were normalised by z-scoring using causal running statistics obtained from the previous 10 seconds of data. This normalization ensured that the features were on a comparable scale and that the analysis pipeline remained compatible with a real-time setting by avoiding introducing information from future time points (Merk, Peterson, Lipski, et al., 2022).

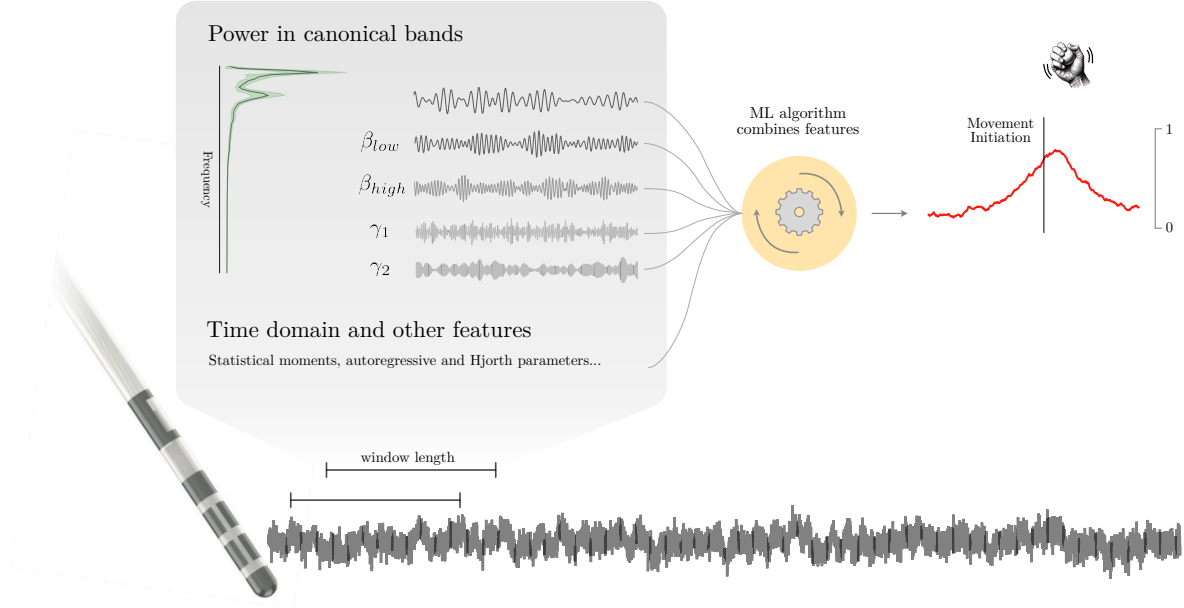


Figure 4.1: Feature Extraction Pipeline Used for Decoding Pre-Movement Periods from Electrophysiological Signals. Raw local field potential (LFP) and electroencephalography (EEG) signals are first pre-processed and segmented into overlapping trailing windows of 0.5 and 1-second lengths, updated every 20 milliseconds. Within each window, features are extracted, and the extracted features are normalised using causal running statistics and concatenated to form a feature vector. These feature vectors are used as inputs to machine learning classifiers designed to detect pre-movement periods in real time.

4.2.4 Classification Pipeline and Evaluation of Decoding Performance

To train classifiers capable of detecting pre-movement periods, we labelled time windows that ended within 500 milliseconds prior to movement initiation as belonging to the pre-movement period.

Determining *a priori* the exact time window during which neural patterns associated with movement intention or preparation occur is challenging. For our initial analysis aimed at identifying potential pre-movement neural activity, we selected this 500 ms window preceding movement onset. Previous findings from intracranial electrocorticography (ECoG) recordings suggest that pre-movement neural correlates in subcortical LFP signals may be found within this interval (Köhler et al., 2024). In subsequent analyses, we performed a time-resolved, moving-window analysis to precisely determine when pre-movement correlates become identifiable and how neural modulation increases as the onset of movement approaches.

Windows ending within the interval $[-4, -2]$ seconds relative to movement onset were labelled as belonging to the *rest period*. Our classifiers were trained to distinguish between these two periods based on the features extracted from the corresponding electrophysiological time series. This paradigm focuses on detecting neural signatures preceding movement execution and allows for variability in the period of interest.

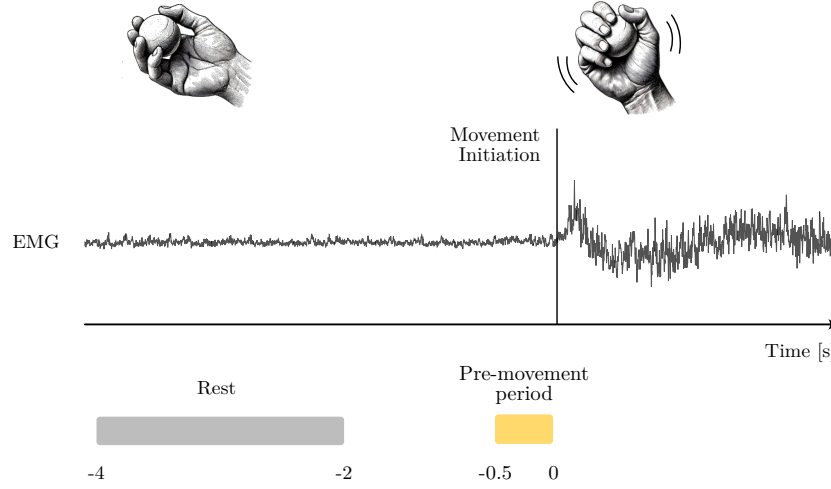


Figure 4.2: Data Labelling. Movement initiation is determined based on EMG activity. Features are extracted using a strictly causal pipeline, ensuring that only features computed from preceding samples are used for estimations at the present time.

The performance of the classifiers was evaluated using a leave-one-trial-out cross-validation scheme. In each iteration, one trial was reserved as validation data to benchmark the classifier's performance, while the remaining trials were used to train the classifier by setting its weights and parameters. This process was repeated iteratively, with each trial serving as the validation set once and as part of the training data in all other iterations. This approach ensures that all trials are used for out-of-sample testing, providing a robust estimate of the classifier's generalization performance.

During each iteration, we computed the area under the Receiver Operating Characteristic curve (ROC-AUC; abbreviated AUC) for the predictions made by the trained classifier on the validation data. This metric was then averaged across all folds of the cross-validation to obtain a single cross-validated performance score. To account for class imbalances—specifically, the longer duration of the rest period compared to the pre-movement period—we adjusted the ROC curve by differentially weighting samples based on the period from which they were extracted. This weighting corrects for

the disproportionate number of samples from the rest period. This procedure was performed individually for each EEG and LFP channel.

We implemented and benchmarked the following classification architectures:

4.2.4.1 *Logistic Regression (LR)*

Logistic Regression is a simple yet powerful classification algorithm that models the probability of a binary outcome using a linear combination of input features. The logistic regression model computes a linear combination of the features and applies the logistic (sigmoid) function to produce an estimate of the probability that the input features belong to the positive class $p(y = 1 \mid \text{features})$.

In our implementation, during each iteration of the leave-one-trial-out cross-validation, we performed an internal three-fold cross-validation on the training data to determine the optimal level of regularization (e.g., L1 or L2 penalty terms). During training, sample weights were adjusted to correct for class imbalance by giving more weight to samples from the minority class. Using the selected hyperparameters, a final logistic regression model was trained on the training data and evaluated on the validation trial for the current fold.

4.2.4.2 *Gradient-Boosted Decision Trees (GBDT)*

Gradient-Boosted Decision Trees are an ensemble learning method that builds a predictive model by sequentially combining multiple weak learners—typically decision trees—with each new tree aiming to correct the errors of the preceding ensemble. This iterative process is guided by gradient descent optimization, minimising a specified loss function such as the negative log-likelihood.

In our implementation, we tuned the GBDT model to find optimal hyperparameters (including the number of trees, learning rate, and tree depth) through an internal cross-validation procedure at each iteration of the leave-one-trial-out scheme. Specifically, we utilised a grid search to explore a range of hyperparameter values, selecting the combination that yielded the best cross-validated performance on the training data. Once the optimal hyperparameters were determined, the GBDT model was trained on the entire training set of the current fold and evaluated on the validation trial. Sample weighting was applied during training to account for dataset imbalance.

Our primary motivation for benchmarking GBDTs alongside logistic regression was to assess the potential benefits of capturing non-linear relationships between features and labels. If significant non-linear relationships existed in the data, GBDTs were expected to better capture and harness these patterns compared to the linear logistic regression model.

4.2.4.3 Convolutional Neural Network (CNN)

Convolutional Neural Networks are deep learning models that automatically learn hierarchical feature representations from raw input data through multiple layers of convolutional filters. Unlike the feature-based methods described above, CNNs take minimally pre-processed electrophysiological time series as input, without relying on manually extracted features.

In our CNN architecture, the input time series were processed through a cascade of convolutional layers interleaved with non-linear activation functions, batch normalization layers (to correct for scaling differences and improve training stability), and pooling layers (to reduce the dimensionality of activations). The convolutional filters were trained to extract relevant information content from the signals pertinent to the classification task. The training process involved the use of Adam, an adaptive gradient descent algorithm that iteratively adjusted the filter coefficients to minimise the cross-entropy loss between the model's predictions and the true labels.

At each stage of the network, the processed signals (referred to as "activations") were filtered, scaled, combined, and passed through non-linear activation functions. After the final convolutional layer, descriptive statistics were computed from the activations, producing scalar features that were input into a fully connected linear classifier (i.e. perceptron), yielding the final output logits.

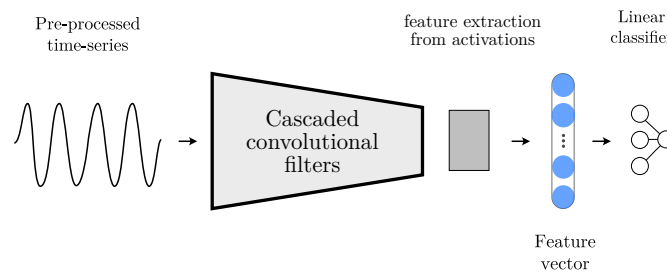


Figure 4.3: CNN Architecture. Minimally pre-processed electrophysiological timeseries signals are processed through a cascade of convolutional filters. Activations are filtered, scaled, combined, and passed through non-linear

activation functions as they progress through the different layers of the network. Final activations are summarised using descriptive statistics, which serve as input features for a linear classifier, producing the final logit.

4.2.4.4 Convolutional Neural Network with Manually Extracted Features (FeatCNN)

To harness both automatically learned features and manually extracted features, we implemented a hybrid architecture that combines the CNN described above with the manually extracted features from our feature extraction pipeline. After the time series data were processed by the convolutional part of the network, the resulting convolutional features were concatenated with the manually extracted features. This combined feature set was then fed into the fully connected linear layers at the end of the network.

This approach allowed us to compare the performance of a model driven solely by learned convolutional features (CNN) with a model that integrates both manually extracted features and learned features (FeatCNN). By evaluating both architectures, we aimed to assess whether the addition of manually extracted features could enhance classification performance.

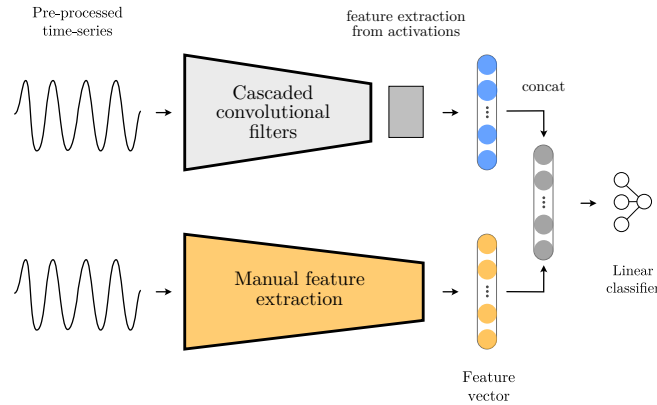


Figure 4.4: Feat-CNN Architecture. This model combines a Convolutional Neural Network (CNN) and the manual feature extraction pipeline into a single architecture.

For both convolutional network architectures, the electrophysiological time series signals were downsampled to 512 Hz from the original sampling rates (2048 Hz or 4096 Hz) prior to being passed to the network. The cascade of convolutional filters consisted of 6 layers, each with a convolutional filter comprising 45 kernel elements and a stride of 2 samples, a batch normalization layer, and a SiLU

activation function (Ramachandran et al., 2017). During training, we employed data augmentation techniques to enhance model generalization. Specifically, we performed data augmentation in logit space by linearly combining samples using the *Manifold Mixup* technique (Verma et al., 2019). This approach regularises the model by encouraging smoother decision boundaries and works by synthetically generating samples through interpolation of existing samples in logit space.

To address dataset imbalance during training, we applied a weighted sampling technique. Samples from the minority class (pre-movement period) were oversampled to ensure that the classifier received a balanced representation of both classes during training iterations.

Further details on the motivation, implementation, and suitability of these convolutional neural network models can be found in our prior work (Rodriguez et al., 2023), as well as in Section 3.2.3 of this thesis.

4.3 Results

4.3.1 Decoding of Pre-Movement Periods

We evaluated the performance of four classification models—Logistic Regression (LR), Gradient-Boosted Decision Trees (GBDT), Convolutional Neural Networks (CNN), and a hybrid CNN incorporating manually extracted features (FeatCNN)—in distinguishing rest periods from pre-movement periods using both thalamic local field potentials (LFPs) and scalp electroencephalography (EEG) signals.

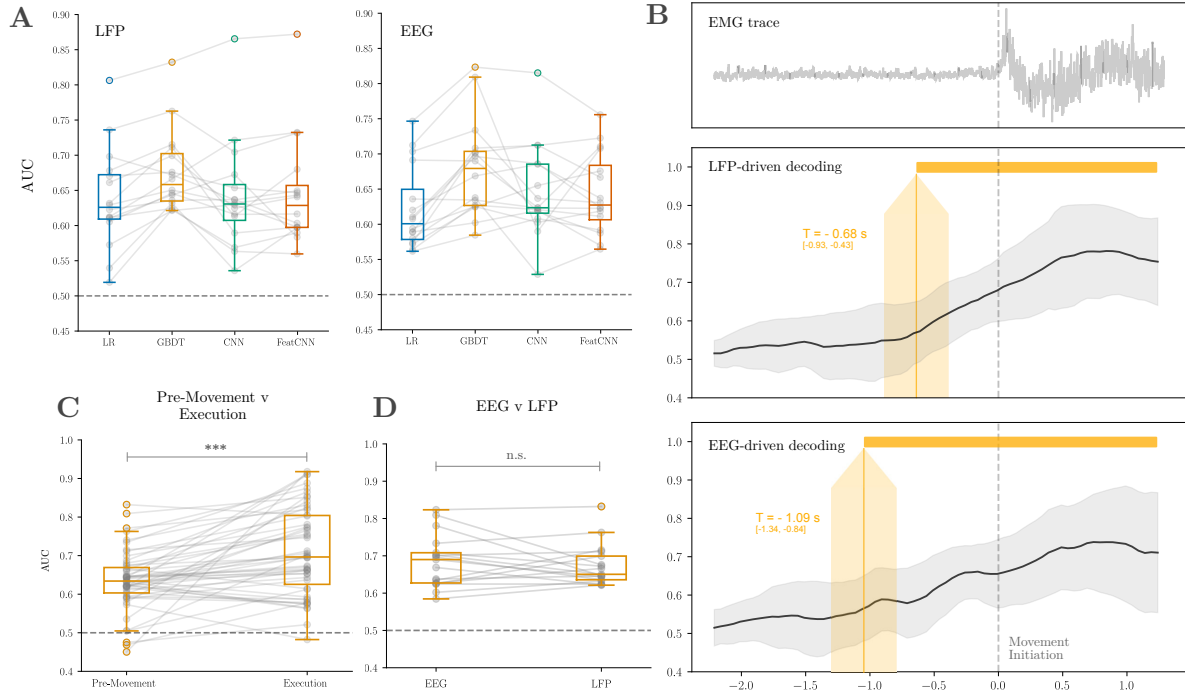


Figure 4.5: Decoding Performance of Different Models in Detecting Pre-Movement Periods from Neural Data. **(A)** *Classification Model Performance:* Boxplots of the Area Under the Receiver Operating Characteristic Curve (AUC) values for each classification model—Logistic Regression (LR), Gradient Boosted Decision Trees (GBDT), Convolutional Neural Network (CNN), and Feature-CNN (FeatCNN)—in distinguishing rest from pre-movement periods using Local Field Potential (LFP) and Electroencephalography (EEG) data. The boxplots represent the distribution of AUC values across participants. Individual dots outside the whiskers denote patients that are outliers. **(B)** *Time-Resolved Decoding Performance:* Plots showing the time-resolved AUC values for logistic regression models trained to detect pre-movement periods at various time windows relative to movement onset. Movement onset ($t = 0$ s) is marked by the vertical dashed line. The top subplot displays the cross-participant average Electromyography (EMG) trace, aligning the timing of muscle activation. The middle (LFP data) and bottom (EEG data) subplots illustrate the progression of mean AUC values over time. The solid black lines represent the mean AUC across participants, and the grey shaded areas indicate ± 1 standard deviation. The vertical shaded region (centred around the vertical yellow line) highlights the earliest time span at which decoding performance becomes significantly higher than chance level ($AUC = 0.5$), demonstrating the models' ability to predict movement onset in advance. **(C)** *Pre-Movement vs. Movement Execution Decoding:* Comparison of AUC values for GBDT models trained to detect pre-movement periods (from -500 ms to 0 ms relative to movement onset) versus movement execution periods (0 ms to 500 ms) using LFP data. This comparison assesses the models' effectiveness in decoding neural activity immediately before movement initiation versus during movement execution. **(D)** *Comparison Between Signal Modalities:* AUC values for GBDT models utilizing LFP and EEG data, illustrating the difference in decoding performance between the two neural signal modalities.

Figure 4.5A illustrates the area under the Receiver Operating Characteristic curve (AUC) values for each model using LFP and EEG data. Higher AUC values indicate better decoding performance. Mean AUC values and the corresponding 95% confidence intervals are presented in Table 4-2.

Table 4-2: Decoding Performance During the Pre-Movement Period. Mean AUC value [95% CI] for decoding the [-500ms, 0] pre-movement period with different models. Higher values indicate better decoding performance.

	LR	GBDT	CNN	FeatCNN
LFP	0.64 [0.60, 0.67]	0.68 [0.65, 0.70]	0.64 [0.60, 0.68]	0.65 [0.61, 0.68]
EEG	0.62 [0.59, 0.65]	0.68 [0.65, 0.71]	0.65 [0.61, 0.68]	0.64 [0.62, 0.67]

Paired t -tests were conducted to compare the performance of the models, with corrections for multiple comparisons applied using the False Discovery Rate (FDR) method. For the LFP-driven models, GBDTs significantly outperformed the other models. Specifically, the comparison between LR and GBDT ($p = 0.002$), between CNN and GBDT ($p = 0.035$), and between FeatCNN and GBDT ($p = 0.033$) indicated significant differences. No significant differences were found among the other three models (LR, CNN, FeatCNN).

For the EEG-driven models, GBDTs also showed superior performance compared to LRs ($p = 0.0003$). However, the differences between GBDT and CNN ($p = 0.13$) and between GBDT and FeatCNN ($p = 0.03$) were not statistically significant after correcting for multiple comparisons. Here again, no significant differences were observed among the other three models.

To determine the earliest time point at which movement onset could be predicted, we trained logistic regression models using a sliding 500 ms window, shifted in 25 ms increments from -2.25 seconds up to 1.5 seconds relative to movement onset. This resulted in AUC values at each time step, providing a time-resolved measure of the model's ability to distinguish pre-movement periods from rest, as shown in Figure 4.5B.

Statistical analyses revealed that, for LFP-driven models, the first time window where the AUC was statistically significant above chance (paired t -test, $p < 0.05$) was the $[-0.93 \text{ s}, -0.43 \text{ s}]$ interval, centred at 680 ms before movement onset. For EEG-driven models, the earliest significant window was $[-1.32 \text{ s}, -0.84 \text{ s}]$, centred at 1,090 ms before movement onset.

We compared the decoding performance of GBDT models trained to detect pre-movement periods ($[-500 \text{ ms}, 0]$) with those trained to detect movement execution periods ($[0, 500 \text{ ms}]$) using LFP data. Figure 4.5C illustrates this comparison. For LFP-driven models, the AUC values for movement execution were significantly higher than for pre-movement decoding. A paired t -test yielded a t -statistic of 14.12 and a p -value less than 0.001, with an average AUC increase of 15.3% when comparing movement execution to pre-movement decoding. For EEG-driven models, the AUC difference was also significant, with a t -statistic of 9.02 and a p -value below 0.001, and an average relative AUC increase of 14.5%.

We assessed whether there was a consistent difference in decoding performance when using LFP signals from the thalamus versus EEG signals recorded from the scalp. Paired t -tests revealed no significant differences in AUC values between LFP and EEG across all models. The uncorrected p -values were 0.55 for LR, 0.47 for GBDT, 0.80 for CNN, and 0.87 for FeatCNN. The AUC values for GBDT models using LFP and EEG data are depicted in Figure 4.5D.

4.3.2 Cross-Patient Variability in Modulatory Patterns

Regardless of their architecture, classification models harness changes in the information content of electrophysiological signals—often manifested as synchronization or desynchronization of oscillatory activity in canonical frequency bands—to drive decoding performance. To visualise the average modulatory patterns across patients, we computed cohort-level time-frequency decompositions (see Figure 4.6). These were calculated by using resonator IIR filters with a quality (Q) factor of 20, rectifying the resulting narrowband activity, and smoothing it using a 250ms kernel. Baseline activity levels were computed for each frequency bin from the rest period, defined as $[-4, -2]$ seconds relative to movement onset, and used to normalise the activities at this frequency. Additionally, we computed the average, cross-patient time-frequency decompositions separately for the top and bottom 50th

percentile performers based on decoding accuracy (AUC) to explore the differences in modulatory patterns associated with different levels of classification performance.

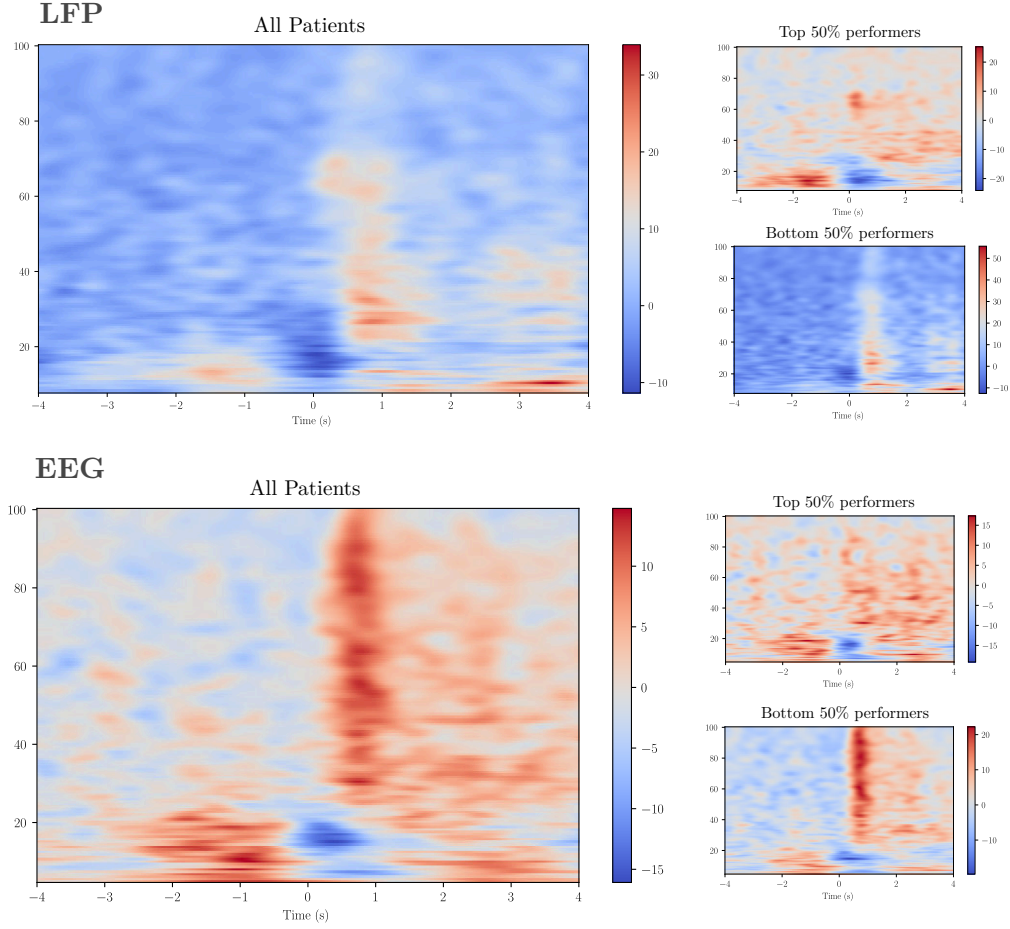


Figure 4.6: Cohort-Level Time-Frequency Decompositions. Represents oscillatory activity across patients around the onset of movement ($T = 0$). Plots depict the modulation of power (percentage change) relative the baseline activity computed from $[-4, -2]$.

To investigate the specific modulatory patterns associated with the pre-movement period, we computed modulation indices for individual frequency bands. The modulation index for a specific band was calculated by comparing the power spectral density (PSD) estimates in that band during the pre-movement period with those during the rest period, and averaging this modulation across trials:

$$MI_{band} = \frac{1}{n_{trials}} \sum_{trial} \frac{PSD_{Pre-Movement}^{trial,band} - PSD_{Rest}^{trial,band}}{PSD_{Rest}^{trial,band}}$$

Modulation indices for each subject and band, along with the mean and 95% confidence intervals, are depicted in Figure 4.7A for both LFP and EEG signals. A negative modulation index corresponds to

desynchronization (decreased oscillatory activity during the pre-movement period when compared to rest) in the respective frequency band, while a positive modulation index indicates synchronization (increased activity). Figure 4.7B shows time-frequency decompositions from individual patients, illustrating the variability in modulatory patterns observed across the cohort.

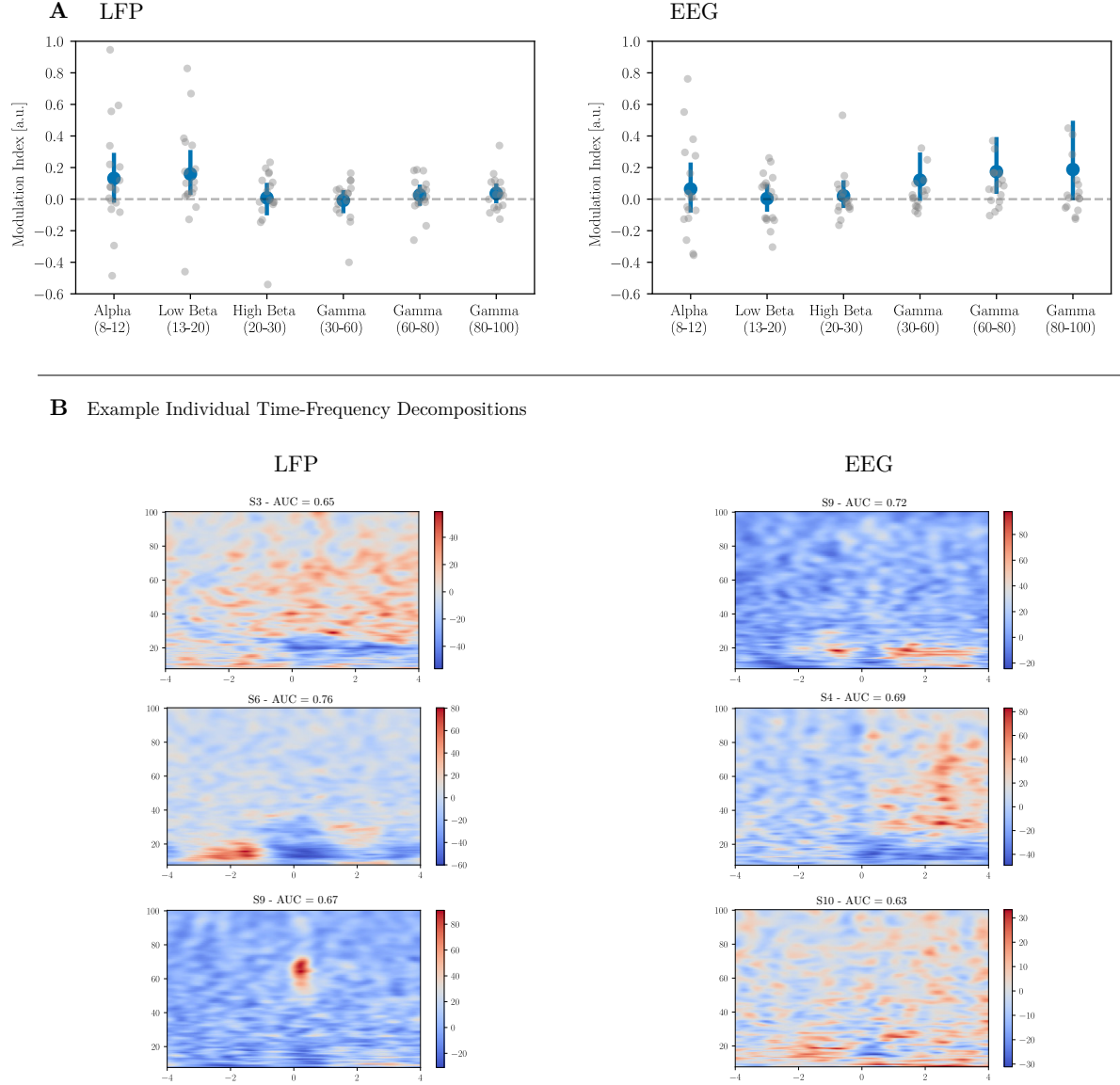


Figure 4.7: Cross-Patient Variability of Movement-Related Modulatory Patterns. (A) Modulation indices for each band. Each dot represents an individual patient, and the blue bars represent the cross-patient mean and 95% confidence intervals. Negative modulation indices represent desynchronization (decreased activity) during the pre-movement periods, whereas positive indices represent synchronization (increased activity) relative to rest. (B) Individual time-frequency decomposition for a selection of patients, showcasing various modulatory patterns that are present in the dataset.

We investigated the relationship between modulation of oscillatory activity in canonical frequency bands and the ability of our trained systems to detect pre-movement periods. We employed modulation indices to quantify changes in power spectral density (PSD) within specific frequency bands during pre-movement periods compared to rest. We then correlated these modulation indices with the decoding performance, measured by the area under the ROC curve (AUC) values of logistic regression (LR) models.

Figure 4.8 displays linear regressions between AUC values of LR models and corresponding modulation indices in canonical bands for both LFP and EEG signals. None of these regressions yielded p -values below the statistical significance threshold of 0.05 (without correction for multiple comparisons being applied).

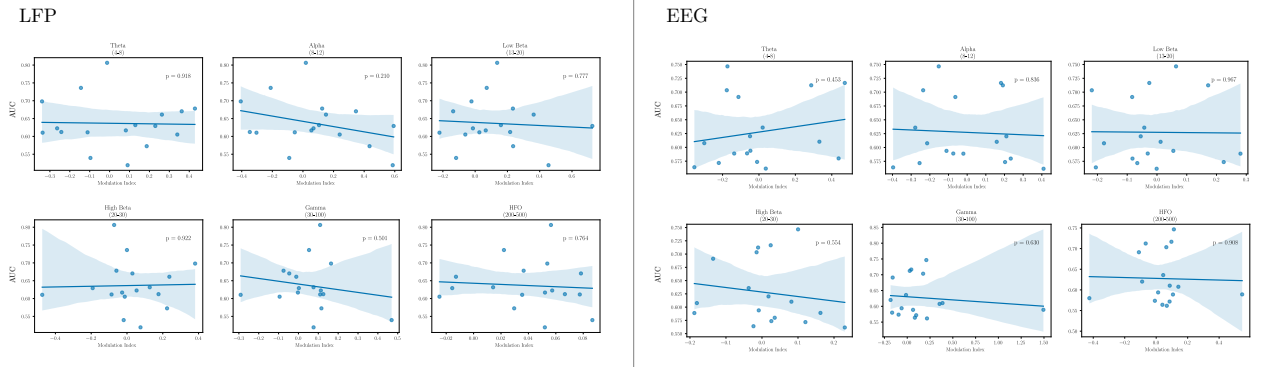


Figure 4.8: Correlations Between Modulation Indices in Canonical Frequency Bands and Decoding Performance. None of the canonical bands is significantly correlated with decoding performance.

To further elucidate the relative contribution of specific bands towards driving decoding performance, we trained LR models using features extracted from a single, individual canonical band, as opposed to simultaneously using all canonical bands to drive a model. This approach allowed us to benchmark the impact of each frequency band on decoding performance across the cohort against the all-band model introduced previously.

Figure 4.9 depicts the differences in AUC scores between models trained on individual bands and the model with access to all bands. Across all frequency bands, relying on only a single band resulted in an average AUC decrease of $18.8\% \pm 8.5\%$ (mean $\pm$ standard deviation) compared to the all-band model. The smallest performance loss was observed in the low beta band (13–20 Hz), with an average

AUC decrease of $15.25\% \pm 9.9\%$. The largest loss occurred in the high-frequency oscillation (HFO) band (200–500 Hz), with an average decrease of $23.01\% \pm 6.22\%$.

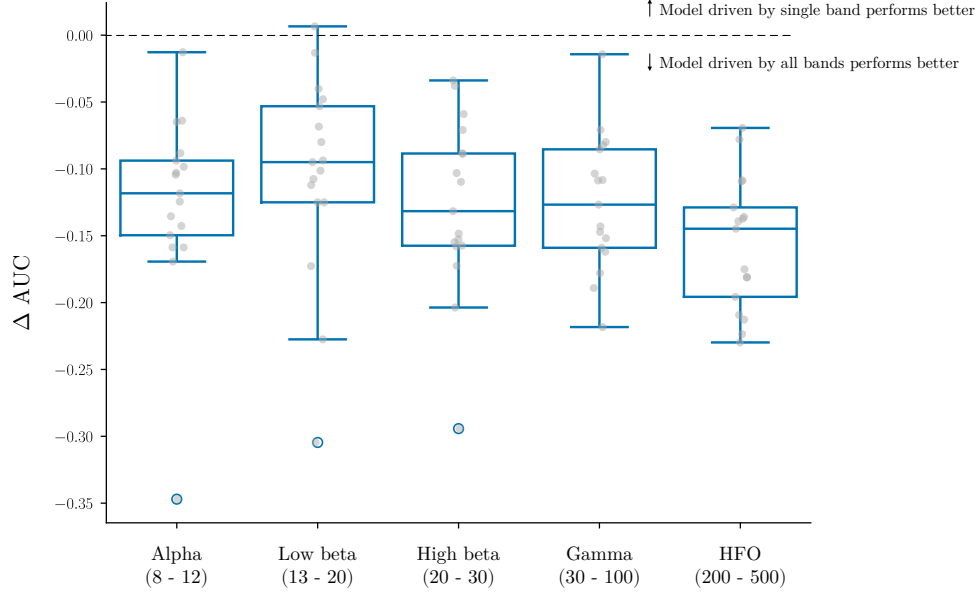


Figure 4.9: Decoding Driven by Individual Frequency Bands. The distributions showcase the loss in decoding performance (difference in AUC value) when models are trained using features from a single canonical frequency band, compared to a model trained using all bands. Negative values indicate lower AUC for single-band models relative to the all-band model.

To assess the generalisability of modulatory patterns across patients –i.e. the ability to make predictions based on patterns learned from one patient on data from another patient–, we trained logistic regression models under different training conditions and compared their out-of-sample decoding performance. Specifically, we evaluated three training scenarios:

1. **Patient-Specific Training:** Models were trained exclusively on data from the patient of interest.
2. **Combined Training:** Models were trained on data from the patient of interest supplemented with data from other patients.
3. **Cross-Patient Training:** Models were trained exclusively on data from other patients, without including any data from the patient of interest.

We then evaluated the performance of these models on the patient-specific test data. For both LFP and EEG signals, including data from other patients during training led to a degradation in decoding performance compared to patient-specific training ($p < 0.001$ in both cases; see Figure 4.10). Furthermore, models trained exclusively on data from other patients performed significantly worse than those trained on combined data (LFP: $p = 0.011$; EEG: $p = 0.0017$). These results suggest that modulatory patterns associated with pre-movement periods exhibit substantial inter-patient variability, limiting the generalisability of models across patients.

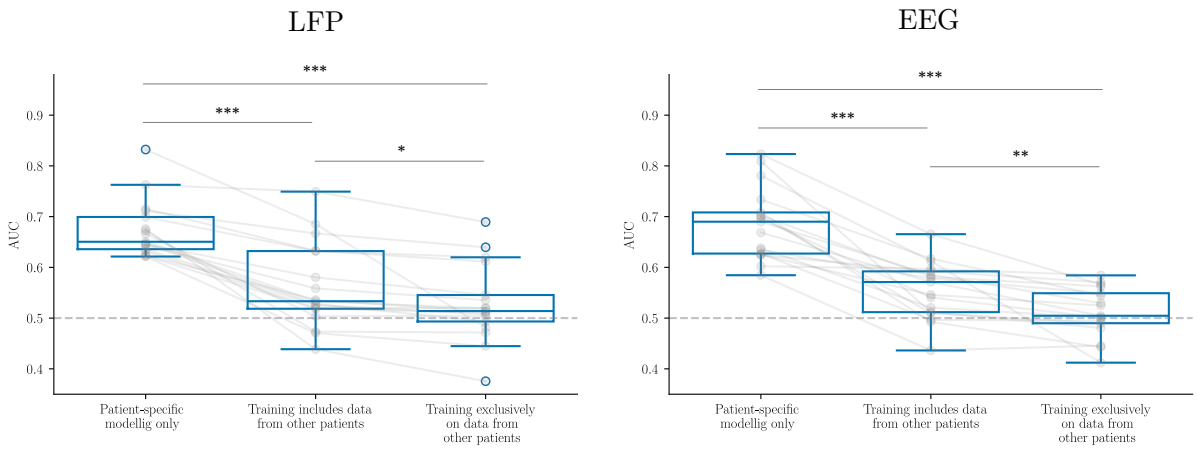


Figure 4.10: Cross-Patient Training and Evaluation. Logistic Regression models were evaluated on patient-specific test data but trained under three different conditions: (1) using only patient-specific training data, (2) using combined training data that included both the patient's data as well as data from other patients, and (3) using only training data from other patients. Decoding performance decreased when data from other patients was included during training, indicating limited generalisability of modulatory patterns across patients.

4.4 Discussion

4.4.1 Decoding Performance is Driven by a Broad Spectrum of Frequencies

Machine learning has gained relevance as a powerful tool for investigating the neural correlates of behaviour and pathology, as well as for quantifying the strength of coupling between neural oscillatory

phenomena and associated behavioural or pathological states. By training multivariate models to predict behavioural or pathological states from the neural oscillatory data, we can assess the extent to which the patterns of neural activity are associated with these states. A key advantage of machine learning models lies in their ability to adjust weights based solely on the information contained within the data, without relying on assumptions made by the researcher or technician. This data-driven approach is particularly desirable in translational electrophysiology, since it enables the development of systems that are tailored to individual patients and account for patient-specific idiosyncrasies.

However, this adaptability also presents challenges. The same capacity of machine learning algorithms to adapt to individual cases can make it non-trivial to interpret the modulatory patterns that a model is harnessing to make its predictions. Understanding the heuristics that models rely upon is crucial for several translationally relevant considerations, including debugging, improving model robustness, and predicting the potential failure modes that may occur outside of a controlled clinical environment. Moreover, the ability to systematically aggregate and analyse the modulatory patterns that these algorithms rely on is essential for identifying shared phenomena among patients and gain insights into underlying functional or pathological neural processes. This capability to interpret the heuristics learned by models is necessary if machine learning is to be considered a suitable methodological approach in neuroscience research.

In this study, we investigated the relationship between modulation of activity in canonical frequency bands and the ability of our trained systems to detect pre-movement periods. Modulation indices quantified changes in oscillatory activity in specific frequency bands during the pre-movement period and were correlated AUC values from logistic regression models. None of the bands yielded a significant correlation between modulation and decoding performance. This lack of significant relationships suggests that no single frequency band consistently drives decoding performance across patients. We note that this might be related to the small sample size of this study, and that increased patient numbers might reveal more consistent (and better statistically powered) relationships between decoding performance and modulation in specific frequency bands. Here, contributions to decoding appear to be distributed across several frequency bands and may vary among individuals instead. Models driven by all canonical frequency band consistently outperformed models driven by a single band.

To further investigate whether contributions from different frequency bands are independent, we conducted an analysis examining the relationship between the decoding performance of models trained

on single frequency bands and that of the all-band model. Specifically, we aimed to determine whether the AUC values from frequency-specific models could predict the AUC values of the all-band model, thereby assessing the extent to which individual bands contribute uniquely to the overall decoding performance.

We performed linear regression analyses where the AUC values obtained from models trained on individual frequency bands served as independent variables, and the AUC values from the all-band model were the dependent variable. By examining the strength and significance of these relationships, we could infer whether certain frequency bands had unique contributions to the decoding task that were not captured by other bands.

Our analyses revealed that the AUC values from individual-band models were indeed linearly related to those of the all-band model, explaining a statistically significant portion of the variance ($p < 0.05$). For instance, the AUC values from the low beta band model (13–20 Hz) explained 20% of the variance in the AUC values of the all-band model ($R^2 = 0.20$). Similarly, the gamma band model (30–100 Hz) explained 17% of the variance ($R^2 = 0.17$). These significant correlations indicate that both the low beta and gamma bands contribute meaningfully to the decoding performance achieved by the all-band model.

Importantly, we observed that the AUC values from the low beta and gamma band models were poorly correlated with each other ($R^2 = 0.05$), suggesting that these frequency bands capture different aspects of the neural signals relevant to the decoding task. This low correlation implies that the contributions from the low beta and gamma bands are relatively independent of one another.

To assess the combined effect of these bands, we constructed a multiple regression model using the AUC values from both the low beta and gamma band models as independent variables to predict the AUC of the all-band model. This combined model explained 48% of the variance in the all-band model's AUC ($R^2 = 0.48$), a substantial increase when compared to the variance explained by each band individually. The significant increase in explained variance when both bands are considered together indicates that the low beta and gamma bands provide complementary information that, when combined, enhances decoding performance.

These findings suggest that contributions from different frequency bands are both individually significant as well as relatively independent. The independence of these contributions implies that different frequency bands capture distinct neural dynamics associated with pre-movement periods. By

harnessing the complementary information from multiple frequency bands, models can achieve superior decoding performance compared to using any single band alone. Our results highlight the importance of incorporating a broad spectrum of frequency bands in decoding models to capture the complex neural signatures underlying motor preparation and execution. The independent and additive contributions from different bands support the notion that neural processes involved in movement are distributed across multiple oscillatory activities.

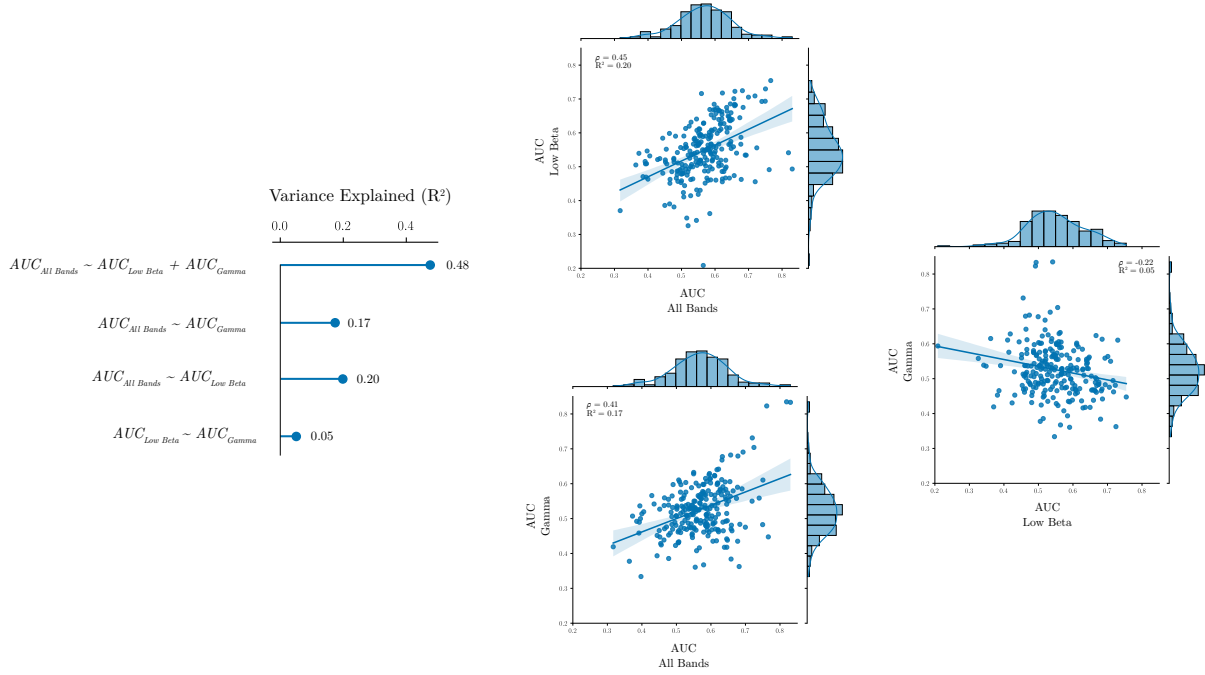


Figure 4.11: Linear Coupling Between Models Operating on Different Frequency Bands. The scatter plots and linear regression lines shows the relationships between decoding performance (AUC values) of models trained on individual frequency bands and the all-band model. (top) All-band model vs. low beta band model. (bottom) All-band model vs. gamma band model. (right) Low beta band model vs. gamma band model.

Collectively, our results demonstrate that decoding performance is not driven by a single frequency band but rather benefits from the simultaneous inclusion of multiple bands. Specifically, we report that:

- **No Single Band Dominates Decoding Performance:** Even though activity in the low beta band may yield better cross-patient decoding performance than activities in any other band (if only a single frequency band is used to drive the decoding), there is a lack of significant relationships between modulation in any single frequency band and decoding performance, suggesting that contributions are distributed across different bands.

- **Enhanced Performance with Multi-Band Models:** Decoding performance increases when using features from all frequency bands simultaneously, compared to models utilizing only individual bands.
- **Independent Contributions from Multiple Bands:** Individual bands, particularly the low beta and gamma bands, contribute independently to the overall decoding performance of the all-band model.

Our findings align with previous literature highlighting the importance of multi-band analyses in decoding neural signals. In (Ahn et al., 2020), authors used microelectrodes recordings from the subthalamic nucleus of Parkinson’s Disease patients to decode a proxy for motor performance. They reported superior decoding performance when models were trained on a broad spectrum of frequencies compared to individual bands, consistent with our observations. Similarly, (Khawaldeh et al., 2020) and (He et al., 2021) emphasised the simultaneous contributions of multiple frequency bands in subthalamic and thalamic LFPs, as evidenced by the weights assigned by machine learning algorithms to features from various bands. (Khawaldeh et al., 2022) further demonstrated that including a broad range of frequency bands significantly improved the prediction of clinical movement impairment scores in Parkinson’s disease patients, compared to using only the canonical beta band.

For the specific task studied here—the detection of periods leading up to movement execution—we find that utilizing multiple frequency bands enhances decoding performance. These results support the multi-band, machine learning-based approach to neural decoding that has been proposed in the literature (Merk, Peterson, Köhler, et al., 2022; W. Neumann & Rodriguez-Oroz, 2021; W.-J. Neumann et al., 2019).

Our study also underscores the potential limitations of focusing solely on canonical frequency bands or predefined biomarkers. The variability in modulatory patterns across patients suggests that individualised models, which account for patient-specific neural dynamics across multiple frequency bands, may offer better performance. This highlights the importance of personalised and data-driven approaches in developing adaptive deep brain stimulation systems and potentially in other neurotechnological interventions.

We note that this finding may contrast with reports from recent literature, which suggest that, given sufficiently large population and dataset sizes, it might be possible to train machine learning models that generalise and retain performance across patients (Yin et al., 2024). Given larger patient numbers than the ones available for this study, this might be achievable.

4.4.2 Comparing Manually Extracted Features and Deep Learning-Based Features

An essential consideration when applying machine learning techniques to translational electrophysiology is the optimal characterization of relevant activity taking place within the recorded neural signals through feature extraction. In Chapter 3 of this thesis, we benchmarked and compared several signal processing approaches to optimise the feature extraction process. The goal of the feature extraction pipeline is to capture and highlight relevant information efficiently while minimising noise and spurious artefacts. Changes to the feature extraction pipeline, however, can also serve as a research tool: by systematically adding or removing features from the classification pipeline, we can estimate their relevance for solving the decoding task. For example, when assessing the relative contributions of different frequency bands to decoding performance, we modified the features input into the classifiers and analysed the corresponding changes in decoding accuracy. This method allowed us to determine the strength of each frequency band's relationship with the behavioural state of interest, providing insights into which neural oscillations are most informative for predicting pre-movement periods.

In this context, we investigated whether including additional types of features beyond power spectral features (power estimates in canonical bands) would enhance decoding performance. Our results indicate that spectral features alone provide classifiers with most of the relevant information needed for accurate decoding, and that the addition of other feature types does not contribute significantly to the system's performance. This finding is illustrated in Figure 4.12, which displays the correlation between the AUC values yielded by a model using spectral features across all frequency bands and an equivalent model that incorporates additional features (such as time-domain statistics, autoregressive coefficients, and others) alongside the spectral features (we refer to Section 4.2.3 for information on feature composition and the pipeline used to compute the features).

Including these additional features led to a slight increase in performance in a minority of cases; however, this difference was not statistically significant across the cohort ($p > 0.05$). The AUC values from these two models explained 62% of each other's variance, representing the highest correlation observed between any two models or feature sets in this study (further pairwise comparisons are depicted in Supplementary Figure IV). This strong correlation suggests that spectral features capture most of the relevant information required for decoding pre-movement periods.

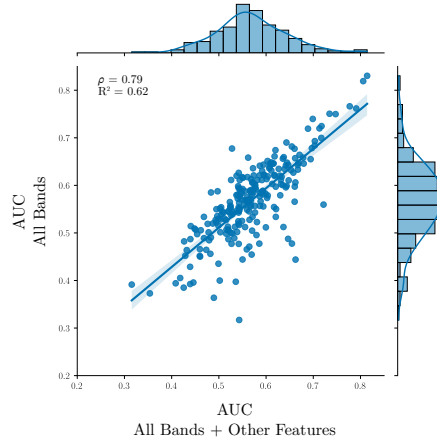


Figure 4.12: Spectral-Only vs. All-Features AUC Values. The scatter plot and regression line illustrate the relationship between decoding performance (AUC values) of a model driven solely by spectral features across all frequency bands, and a model utilizing additional features (time-domain statistics, autoregressive coefficients, and others) together with the spectral features. The high correlation between the decoding performance values suggests that adding the additional features does not significantly improve decoding performance.

One of the advantages of using parameterised deep learning systems, such as convolutional neural networks (CNNs), is their ability to bypass the manual feature extraction process. While approaches based on manually extracted features require careful selection of features—such as choosing specific frequency bands for spectral analysis—the optimal set of features may vary between patients, making it challenging to generalise and potentially limiting performance.

Deep learning approaches accept raw data inputs and process them through a series of learned mathematical operations, in our case, convolutions applied to time-domain signals. These operations are parameterised and can capture a broad spectrum of signal characteristics. The parameters are optimised during the training process to minimise a loss function related to the classification task.

This data-driven learning of feature representations eliminates the need for manual implementation of signal processing pipelines to reduce high-dimensional time-series data to meaningful feature sets.

In this study, we investigated whether deep learning systems could learn feature representations that capture the information necessary to decode pre-movement periods, and whether these learned features provide any advantages over manually extracted features. To this end, we implemented two deep learning-based systems: a standard CNN and a hybrid model referred to as FeatCNN. The CNN processes signals through a cascade of convolutional filters whose coefficients are learned during training, as described in Section 4.2.4. The FeatCNN incorporates a CNN as a subcomponent but additionally uses manually extracted features, thereby combining the two feature sets to make predictions.

By comparing the performance of these two systems, we aimed to determine whether manually extracted features contain relevant information that CNNs cannot capture, and conversely, whether CNNs can capture all pertinent information without the need for manual features. Figure 4.13 presents the paired comparison between the AUC values of the CNN and FeatCNN models. The results show no significant difference in decoding performance between the two models, suggesting that the addition of manually extracted features does not provide a substantial benefit over the CNN's learned features ($p > 0.05$).

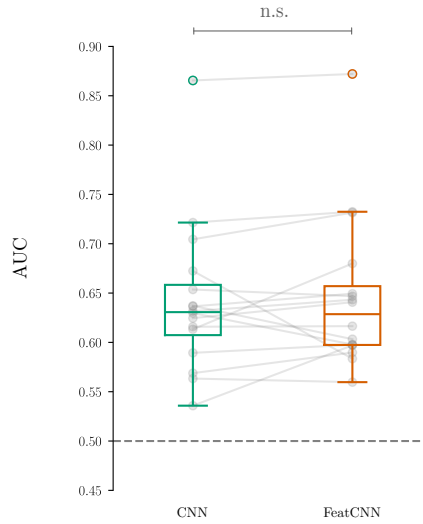


Figure 4.13: Cross-Patient Comparison of AUC Values for the CNN and FeatCNN Models. Each paired set of points represents an individual patient. No significant difference in decoding performance was observed, indicating that the addition of manually extracted features does not enhance the CNN's ability to decode pre-movement periods.

The lack of significant difference in performance between the CNN and FeatCNN indicates that the CNN is capable of capturing the relevant information content from the time-series signals through its learned convolutional filters, including modulatory patterns in relevant frequency bands or their correlates. This finding suggests that CNNs can effectively learn feature representations that are as informative as manually extracted features commonly used in adaptive DBS applications.

However, this does not imply that deep learning-based systems and systems based on manually extracted features are equivalent in how they process data or make inferences. To explore this further, we analysed the pairwise correlations between the performance metrics (AUC values) of the four models: LR, GBDT, CNN, and FeatCNN. Our analysis revealed stronger correlations within the deep learning-based systems (CNN and FeatCNN) and within the feature-based systems (LR and GBDT) than across the two categories. Specifically, 54% of the variance was shared between the CNN and FeatCNN models, and 52% between the LR and GBDT models. In contrast, the CNN shared only 17% and 18% of the variance with the GBDT and LR models, respectively.

Moreover, the correlations between the FeatCNN and the feature-based models (LR and GBDT) were higher (40% and 33%, respectively) than those between the CNN and the feature-based models, suggesting that the inclusion of manually extracted features in the FeatCNN model makes it operate more similarly to the feature-based systems. Nonetheless, the FeatCNN's coupling remains tighter with the CNN than with the LR and GBDT models. These relationships are visualised in Figure 4.14A, which depicts the pairwise correlations between the models, and Figure 4.14B, which presents a two-dimensional embedding where the distances between points represent the proportion of variance explained.

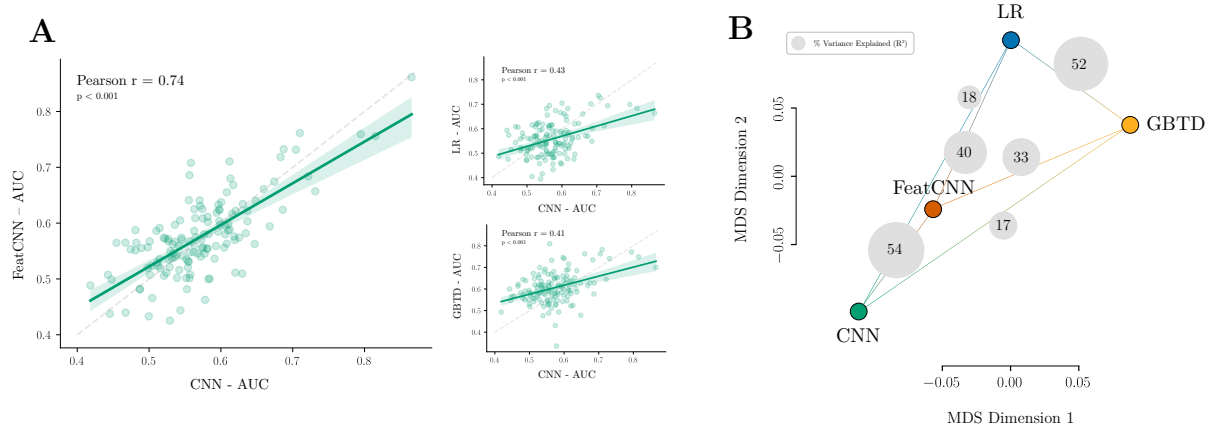


Figure 4.14: Analysis of the Similarity Between the Decoding Performance Yielded by Different Models. (A) Pairwise correlations between the AUC values of the models (higher correlation values indicate greater similarity in model performance). (B) Two-dimensional embedding representing the relationships between models based on the proportion of shared variance in AUC values. The distances between points reflect the degree of similarity in decoding performance across patients.

These findings suggest that, while the deep learning-based and feature-based models may achieve similar levels of decoding performance, they may rely on different aspects of the neural signals or process the information differently to make predictions. The deep learning models may capture complex, non-linear relationships and higher-order features that are not explicitly represented in the manually extracted feature sets. Conversely, feature-based models rely on predefined features that may be more interpretable but potentially less flexible in capturing individual patient variability.

The mechanisms underlying the differences between deep learning-based and feature-based models, as well as the potential implications for translational applications, remain areas for further research, and understanding these differences is crucial for developing reliable and interpretable models for clinical use.

4.4.3 Inference-Time Computational Cost

In the context of adaptive deep brain stimulation (aDBS), signal processing pipelines that incorporate machine learning-based components present translationally exciting opportunities (W.-J. Neumann et al., 2019). However, the computational environment within which aDBS can be feasibly implemented

outside of the clinic remains constrained due to limitations in hardware resources and power consumption (Ansó et al., 2022). Therefore, efforts to develop algorithms intended for aDBS applications should take these considerations into account to ensure that the systems are practical for real-world deployment.

Estimating the computational complexity of algorithms implemented in high-level scripting languages such as Python or MATLAB is non-trivial, as these implementations may differ significantly from implementations optimised for the embedded systems with specialised hardware that would be used in real-world application. Nonetheless, providing reference values for the computational requirements can offer valuable insights into the feasibility of different algorithms in computationally constrained environments.

To assess the computational demands of our models, we benchmarked the time required to make a single inference-time prediction—that is, to process time-series signals and produce a model output—on a dedicated workstation. The server was configured with an AMD EPYC 7402 24-core CPU, 503 GB of system memory, and two NVIDIA GeForce RTX 3090 GPUs (each with 24 GB of onboard RAM). All benchmarks were performed under a lightly loaded operating system to approximate real-time deployment conditions. This setup provides a realistic upper bound on the computational efficiency of each algorithm during online operation.

The feature extraction step, when using the full feature set, requires 187 milliseconds (ms) per prediction. Within this, the estimation of spectral power features—which drive most of the performance of the feature-based models (see Section 4.4.2) —takes only 5 ms. The time estimates for the forward pass of the different algorithms are outlined in Table 4-3.

Table 4-3: Benchmarked Time [in milliseconds] Required for an Inference-Time Prediction

	LR	GBDT	CNN	FeatCNN
<i>Full Feature Set</i>				
Manual Feature Extraction	187	187	N/A	187
Model Inference Time	0.06	0.2	1.0	1.25
<i>Total Time</i>	187.06	187.2	1.0	188.25
<i>Spectral Feature Set Only</i>				
Manual Feature Extraction	5.0	5.0	N/A	5.0
Model Inference Time	0.06	0.2	1.0	1.25

<i>Total Time</i>	5.06	5.2	1.0	6.25
-------------------	------	-----	------------	------

Abbreviations: LR = Logistic Regression; GBDT = Gradient-Boosted Decision Trees; CNN = Convolutional Neural Network; FeatCNN = CNN with Manually Extracted Features; N/A = Not Applicable.

When using the full feature set, which includes time-domain statistics, autoregressive coefficients, cepstral coefficients, and spectral features, the feature extraction step dominates the computational cost, requiring 187 ms per prediction. The logistic regression (LR) and gradient-boosted decision tree (GBDT) models have minimal inference times of 0.06 ms and 0.2 ms, respectively, resulting in total inference times of approximately 187 ms. In contrast, the CNN model, which operates directly on the raw time-series data without manual feature extraction, has a total inference time of 1.0 ms. The FeatCNN model, which combines the CNN with manually extracted features, requires both the feature extraction time (187 ms) and the CNN inference time (1.25 ms), leading to a total inference time of approximately 188 ms.

When using only the spectral feature set—which, as previously discussed, captures most of the relevant information—the feature extraction time is significantly reduced to 5 ms. Under this configuration, the total inference times for the LR and GBDT models decrease to approximately 5 ms, whereas the CNN's inference time remains at 1.0 ms, as it does not rely on manual feature extraction. The FeatCNN model's total inference time reduces to approximately 6.25 ms.

4.4.4 Limitations of this Study

In this study, we benchmarked the ability to decode upper limb movements in the milliseconds preceding movement execution by implementing and testing various machine learning algorithms and feature sets. While our findings provide valuable insights, several limitations should be acknowledged.

First, a significant limitation of our study is the relatively small sample size of our patient cohort ($n = 11$), which may constrain the generalisability of our findings. In electrophysiological research, inter-patient variability is a critical factor that can substantially affect the performance of machine learning models. This variability may arise from a range of factors including differences across patients in individual brain anatomy, electrode placement, individual pathophysiology, and neural signal characteristics. In our study, we observed considerable heterogeneity in the modulatory patterns of

neural activity across patients, which led to variability in decoding performance and a decrease in performance when using cross-patient models.

The poor performance of models trained across patients compared to individualised, patient-specific decoders underscores the challenges associated with developing generalised machine learning systems in electrophysiology. Personalised models, tailored to the unique neural signatures of an individual patient, tend to perform better because they can capture the specific patterns and idiosyncrasies of that patient's neural activity. In contrast, cross-patient models must contend with a diverse range of neural patterns present across different individuals, which can dilute the features critical for accurate decoding in any given patient.

Including a larger and more diverse patient population in future studies would help address this limitation by increasing statistical power and enabling the identification of robust patterns that generalise well across individuals. Larger datasets provide more examples of the variability present in neural signals, allowing machine learning models to learn more generalised representations that are less sensitive to individual differences. Previous studies have shown that with sufficient data, cross-subject models can be achieved in certain tasks (Yin et al., 2024).

Moreover, researchers have proposed techniques to mitigate inter-patient variability, such as cross-patient data normalization and domain adaptation methods. These techniques aim to align neural signals across patients by transforming them into a common feature space or adjusting for individual-specific factors, thereby reducing the impact of individual differences and improving the performance of cross-patient models. For instance, cross-patient data normalization has been implemented using EEG and ECoG data to enhance the generalisability of decoding models (S. M. Peterson et al., 2021). Despite these advances, such methods may not fully account for all sources of variability, and their effectiveness can depend on the nature of the data and the specific application.

Second, we employed limited machine learning interpretability techniques on our models—which included logistic regression, gradient-boosted decision trees, and convolutional neural networks. The lack of significant correlations between modulation indices in specific frequency bands and decoding performance suggests that complex, non-linear relationships may exist in the data that are not fully understood. Employing explainable machine learning approaches or incorporating interpretability methods could provide deeper insights into the neural mechanisms underlying movement preparation and execution.

Third, all analyses were performed offline, despite being implemented with a real-time compatible and fully causal pipeline. This means that we lack performance metrics gathered during real-time adaptive deep brain stimulation (aDBS). Assessing the models in real-time settings is crucial to evaluate their practical utility and effectiveness in clinical environments. Future work should focus on real-time implementation and testing to determine the feasibility and reliability of these decoding systems in actual aDBS applications.

Additionally, while we investigated the use of both thalamic local field potentials (LFPs) and scalp electroencephalography (EEG) signals, we did not extensively explore the potential differences and complementarities between these modalities. Further research could delve into how combining LFP and EEG data might enhance decoding performance or provide more robust biomarkers for movement prediction.

Finally, the experimental tasks used to elicit movements were specific and may not encompass the full range of motor activities encountered in daily life. This could limit the applicability of our models to more naturalistic settings. Expanding the repertoire of tasks and incorporating more ecologically valid movement paradigms could improve the models' relevance and utility in real-world scenarios.

4.5 Conclusion

We demonstrated that early neural correlates of upper limb movement within thalamic LFP and EEG signals can be harnessed to decode pre-movement periods prior to movement execution. Decoding performance was statistically significant from chance at approximately 680 ms before movement onset for LFPs and 1.09 s for EEGs. Individualised, patient-specific decoders outperformed cross-patient models, consistent with the observed variability in modulatory patterns across subjects. Multiple frequency bands contributed independently to decoding performance, highlighting the importance of incorporating information from a broad range of frequencies.

To optimise decoding performance, our findings emphasise the necessity of individualised approaches that account for patient-specific neural dynamics and harness multi-band spectral information. These insights have important implications for the development of adaptive neurostimulation therapies and brain-machine interfaces, suggesting that personalised, data-driven models can enhance the effectiveness and reliability of neural decoding systems in clinical applications.

4.6 Summarised Findings

The objective of this chapter was to evaluate the feasibility of decoding pre-movement neural activity from thalamic local field potentials (LFPs) and scalp electroencephalography (EEG) signals in patients with tremor disorders. By harnessing early neural correlates of upper limb movement, we aimed to predict movement initiation prior to execution, which holds significant potential for enhancing adaptive neurostimulation therapies.

The main findings of this study are as follows:

- **Decoding of Pre-Movement Neural Activity is Possible:** We demonstrated that both thalamic LFPs and EEG signals can contain discernible neural signatures preceding movement onset. Decoding performance became statistically significant from chance at approximately 680 milliseconds before movement onset for LFPs and 1.09 seconds for EEGs. This indicates that these neural signals have the potential to be used to anticipate voluntary movement initiation before its onset.
- **Superior Performance of Individualised Decoders:** Individualised, patient-specific decoders consistently outperformed cross-patient models. In settings with limited data availability, this underscores the substantial inter-patient variability in neural modulatory patterns and highlights the importance of tailoring decoding models to individual neural dynamics to achieve optimal performance.
- **Contribution of Multiple Frequency Bands:** Analysis revealed that multiple frequency bands contributed independently to decoding performance. This emphasises the necessity of incorporating information across a broad spectrum of frequencies rather than relying on a single frequency band. Such an approach captures a more comprehensive representation of the neural processes involved in movement preparation.

5 Project 3: Distinguishing Speech from Upper Limb Movements from Subthalamic and Cortical Neural Signals in Parkinson's Disease

5.1 Introduction

5.1.1 The Neural Correlates of Speech Production

Speech production is a complex and dynamic behaviour that requires integrating cognitive processes with precise sensorimotor control. From the perspective of motor control, speech production involves the temporally coordinated activation of over 100 muscles across the larynx, lips, tongue, rib cage, and abdomen (Tremblay et al., 2016). This intricate orchestration enables the rapid and fluid articulation that is necessary for spoken language. Understanding the neural mechanisms that underpin this coordination is essential for advancing theoretical models of speech production as well as for developing clinical interventions for speech disorders.

The exploration of the neural substrates underlying speech production has a rich history dating back to the 19th century (Rutten, 2022). Early neuroscientific models were profoundly influenced by the work of Paul Broca, who identified the posterior inferior frontal gyrus—now known as Broca's area—as critical for spoken language through his studies of patients with aphasia (Broca, 1861). These early models adopted a *localist* perspective, attributing complex behaviours to specific, localised brain regions. Rooted in the contemporaneous practice of phrenology, this viewpoint was prevalent in the neuroscientific thought of that era (Dronkers et al., 2007; Lieberman, 2000).

Over subsequent decades, advancements in neuroimaging and electrophysiological techniques have reshaped our understanding of the neural basis of speech production. Modern models have transitioned toward from a *localist* to a *connectionist* framework, emphasizing the relevance of interactions within distributed neural networks that encompass multiple cortical and subcortical structures. This shift reflects a broader recognition that complex behaviours—including speech—emerge from dynamic interactions across brain regions, rather than from activity taking place in isolated loci (Yuste, 2015).

Research aiming to elucidate the neural correlates of speech production has predominantly focused on the neocortex. Extensive studies have mapped the neural representation of speech articulators within cortical areas, providing detailed insights into the cortical control of speech movements (Bouchard et al., 2013; Mesgarani et al., 2014). In recent years, these cortical patterns have been harnessed to develop brain-machine interfaces capable of generating synthetic speech by decoding neural activity directly from the cortex (Anumanchipalli et al., 2019; Card et al., 2024; Metzger et al., 2023; Moses et al., 2021). These advancements underscore the potential of cortical signals both for understanding the neural mechanisms underlying speech and for developing translational clinical applications for such as the restoration of communication.

However, the neural correlates of speech production extend beyond the neocortex to include critical interactions with subcortical structures. Of particular interest are the basal ganglia, and specifically the subthalamic nucleus (STN), given its central role in motor control and its status as a surgical target in the treatment of Parkinson's disease (PD). The basal ganglia have been implicated in various aspects of motor planning, initiation, and learning, and accumulating evidence suggests they play a significant role in the neural network underlying speech production.

Theoretical models propose that the basal ganglia contribute to the sequencing and timing of speech through mechanisms of state feedback control and motor gating (Bohland et al., 2010; Houde & Nagarajan, 2011). Lesion studies provide clinical evidence supporting this role: damage to basal ganglia structures can result in speech impairments such as dysarthria and speech apraxia, highlighting their functional importance in speech production (Speedie et al., 1993).

Electrophysiological recordings offer direct evidence of basal ganglia involvement during speech production tasks. Single-unit recordings have identified neuronal activity in the basal ganglia correlated with speech articulation and timing (Lipski et al., 2018a, 2023; Tankus & Fried, 2019; Vissani et al., 2023). Similarly, local field potential (LFP) recordings from basal ganglia structures reveal oscillatory patterns associated with various speech processes, including phonation and syllable production (Chrabaszc et al., 2019a, 2021a; Hebb et al., 2012a; Hell et al., 2023).

Intraoperative studies have further elucidated the role of the STN in speech. Recordings have demonstrated hyperdirect pathways connecting the STN and cortical language regions, suggesting a mechanism for rapid cortico-subcortical communication during speech production (Jorge et al., 2022). Additionally, these studies have shown the propagation and transfer of task-related information within this distributed network during speech tasks (Vissani et al., 2023; A. R. Weiss et al., 2023).

Moreover, investigations into the pathophysiology of PD—a neurodegenerative disorder characterised by both the dysfunction and loss of dopaminergic neurons within the basal ganglia (Cramb et al., 2023)—have highlighted the development of speech-related symptoms such as hypokinetic dysarthria and hypophonia (Lieberman, 2009; Rusz et al., 2024; Silveri, 2021). These symptoms reflect impaired motor control and timing function in the basal ganglia. Research on pharmacological interventions that modulate basal ganglia activity—such as dopamine replacement therapy—and their effect on speech production further support the basal ganglia's involvement in speech mechanisms (Cai et al., 2024; Norel et al., 2020; Tykalova et al., 2022).

5.1.2 STN-DBS and Speech

In the clinical management of Parkinson's disease, deep brain stimulation of the subthalamic nucleus (STN-DBS) has emerged as a highly effective treatment for motor symptoms such as rigidity, bradykinesia, and tremor (Dembek et al., 2019). However, speech dysfunction remains a significant and challenging symptom in PD. Impairments such as hypophonia (reduced speech volume), monopitch (reduced pitch variation), and hypokinetic dysarthria affect up to 90% of PD patients, severely impacting their quality of life (Logemann et al., 1978; Polychronis et al., 2019).

While STN-DBS can alleviate some motor aspects of speech by improving general motor function, it has been associated with deleterious effects on both motor and cognitive components of speech production. Speech degradation due to neurostimulation, commonly referred to as stimulation-induced dysarthria⁶ (SID), is recognised as a frequent and debilitating side effect of STN-DBS (Aldridge et al., 2016; Haarmann et al., 2022; Kluin et al., 2022; Krack et al., 2003; Okun et al., 2012; Skodda, 2012; Tripoliti et al., 2011; Tsuboi et al., 2015; Wertheimer et al., 2014).

⁶ The term *dysarthria* is an umbrella term used to describe “a group of speech disorders resulting from disturbances in muscular control over the speech mechanism due to damage of the central or peripheral nervous system. It designates problems in oral communication due to paralysis, weakness, or incoordination of the speech musculature” (Darley et al., 1969; Moreau & Pinto, 2019).

The impact of STN-DBS on speech production is notably variable among patients. Postoperative speech outcomes can range from improvement (Jorge et al., 2020a) to severe impairment (Okun et al., 2012; Tanaka et al., 2015, 2020). Several factors contribute to this variability, including patient-specific characteristics such as disease duration, baseline speech intelligibility, and cognitive status (Tripoliti et al., 2014). The precise positioning of DBS electrodes within the STN also plays a crucial role; stimulation of certain subregions, particularly those adjacent to speech-related pathways, can exacerbate stimulation-induced speech deficits (Bobin et al., 2024a; Jorge et al., 2020a). Stimulating cerebellothalamic and pyramidal tract fibres has been linked to dysarthric symptoms (Jergas et al., 2024).

Adjusting stimulation parameters offers an opportunity to mitigate speech-related side effects of stimulation. These parameters can include the hemisphere being stimulated (Bobin et al., 2024b; Santens et al., 2003; E. Q. Wang et al., 2006), the choice of active contacts within the STN (Picillo et al., 2016), and the configuration of stimulation settings such as amplitude (Bobin et al., 2024b; M. J. Hammer et al., 2010; Törnqvist et al., 2005; Tripoliti et al., 2008), frequency (Fabbri et al., 2019; Grover et al., 2019; Morello et al., 2020), and pulse width (Fabbri et al., 2021). Modifying stimulation settings can influence the extent to which speech is affected by DBS, and thereby offer a potential avenue for individualised therapy optimization.

5.1.3 The Potential of Adaptive DBS to Ameliorate Speech-Related Side Effects

The capacity to adjust DBS parameters to modulate speech outcomes suggests that an adaptive DBS (aDBS) approach could be beneficial in reducing SID. By incorporating a real-time speech detection system, stimulation parameters could be dynamically adjusted based on this information to minimise adverse effects during the production of speech while maintaining relative therapeutic efficacy for the control of other motor symptoms through stimulation. This approach aligns with the broader goals of precision medicine, which aims to tailor interventions to individual patient needs.

Previous studies have demonstrated that aDBS targeting pathological oscillations can reduce SID (Little et al., 2016), indicating that responsive stimulation strategies may alleviate speech side effects more effectively than conventional continuous, open-loop stimulation. Additionally, recent advancements in closed-loop DBS systems have highlighted the feasibility of integrating neural biomarkers to guide stimulation in real time (Arlotti et al., 2018; Gilron et al., 2021; W.-J. Neumann,

Gilron, et al., 2023). Implementing a speech-specific aDBS protocol could enhance the quality of life for PD patients by preserving speech function while control of motor symptoms is maintained.

5.1.4 Aims of this Study

The primary aim of this study is to evaluate the feasibility of distinguishing speech from upper limb movement using subthalamic nucleus (STN) local field potential (LFP) signals recorded from implanted deep brain stimulation (DBS) electrodes in patients with Parkinson's disease (PD). By recording electroencephalography (EEG) and STN-LFP signals from externalised STN-DBS patients during tasks involving speech production and upper limb movements, we aim to characterise the neural signatures associated with each type of movement. This involves analysing the neural time-series data to identify modulatory patterns and examining the spectral, temporal, and spatial features of STN-LFPs linked to both speech and limb movements. Furthermore, we seek to determine the distinguishing features between speech-related and limb movement-related neural activity within the STN, as well as their spatial characteristics. By comparing the neural correlates of both movements, we intend to identify unique biomarkers that can effectively differentiate between these two types of motor activity. This differentiation is crucial for developing decoding algorithms capable of discriminating speech from limb movements based on STN-LFP signals.

To achieve this, in this work we will harness machine learning techniques to develop and implement decoding algorithms and quantify their ability to distinguish speech from upper limb movement. Our focus is on creating algorithms compatible with real-time deployment in adaptive DBS systems. By accomplishing these objectives, this study aims to contribute towards the development of speech-specific adaptive DBS protocols that reduce stimulation-induced side effects affecting speech.

5.2 Methods

5.2.1 Patients and Surgical Procedure

A total of 17 patients (four females) with idiopathic Parkinson's disease (PD) participated in this study after being enrolled in three different centres: King's College Hospital (KCH), London, UK

(n=6), St George’s Hospital (SGH), London, UK (n=7) and University Hospital Cologne (UHC), Cologne, Germany (n=6). The average age was 63.1 ± 1.9 years (mean $\pm$ SEM), with average disease duration of 10.1 ± 1.5 years. All patients underwent bilateral implantation of DBS electrodes targeting the STN as part of their therapeutical high frequency stimulation for PD. Experiments and recordings were carried out between 1 and 6 days after lead implantation, and before the implantation of the neurostimulator in a second surgery. At this stage, the DBS leads were still externalised and could be used to record STN-LFPs. The study was approved by local ethics committees and all patients provided written informed consent according to the declaration of Helsinki. The clinical details of the patients are summarised in Table 5-1.

DBS systems from two manufacturers were used: Medtronic Inc. Neurological Division, USA or Boston Scientific, USA. In 16 patients, DBS electrodes were octopolar directional leads with a 1-3-3-1 configuration distributed on 4 levels, with levels 2 and 3 each containing three segmented contact for directional recording and stimulation. In one patient, leads were octopolar non-directional distributed on 8 levels and targeting both the STN and the ventralis intermedius nucleus (ViM) of the thalamus. The localization of the leads was assessed by co-registering pre-operative MRI with post-operative CT scans and normalizing them into the Montreal Neurological Institute (MNI) space using the semiautomatic approach implemented in the Lead-DBS toolbox (Horn et al., 2019; Horn & Kühn, 2015) (see Figure 5.1B). MNI-defined coordinates of macroelectrodes were extracted in bipolar mode and were used for further analyses.

Table 5-1: Demographic and Clinical Characteristics of Patients

<i>Patient ID</i>	<i>Surgery and recording centre</i>	<i>Gender (M/F)</i>	<i>Age (years)</i>	<i>Disease duration (years)</i>	<i>Time of recording (days post-Op)</i>	<i>Predominant symptoms</i>	<i>DBS Lead</i>	<i>DBS target</i>
1	UHC	M	71	6	1	Tremor	Medtronic Sensight (directional)	STN

2	UHC	F	67	6	1	Tremor	Boston Scientific Cartesia D-Leads	STN
3	KCH	M	54	-	5	Tremor, rigidity, bradykinesia	Medtronic Sensight (directional)	STN
4	SGH	F	63	9	5	Akinetic-rigid	Medtronic Sensight (directional)	STN
5	UHC	M	75	6	1	Tremor, rigidity	Medtronic Sensight (directional, Percept IPG)	STN
6	KCH	M	61	5	5	Tremor	Medtronic Sensight (directional)	STN
7	KCH	M	58	12	5	Rigidity, bradykinesia	Medtronic Sensight (directional)	STN
8	SGH	M	66	10	4	Akinetic-rigid	Boston Scientific (directional)	STN
9	KCH	M	57	5	5	Tremor, bradykinesia, FOG	Medtronic Sensight (directional)	STN

10	SGH	M	63	13	5	Akinetic-rigid	Boston Scientific (directional)	STN
11	SGH	M	70	9	5	Tremor	Boston Scientific Octode (non-directional)	STN + ViM
12	SGH	F	62	12	5	Akinetic-rigid	Medtronic SenSight (directional)	STN
13	SGH	M	69	30	4	Tremor	Boston Scientific (directional)	STN
14	KCH	F	67	12	5	Dyskinesia, tremor, rigidity	Medtronic Sensight (directional)	STN
15	UHC	M	67	6	1	Tremor, rigidity	Medtronic Sensight (directional)	STN
16	SGH	M	61	14	6	Tremor	Boston Scientific (directional)	STN
17	KCH	M	42	7	4	Dyskinesia, tremor, rigidity, bradykinesia	Medtronic Sensight (directional)	STN

M=Male, F=Female, UHC=University Hospital of Cologne, KCH=King's College Hospital, SGH=St Georges Hospital, STN=Subthalamic Nucleus, ViM=ventralis intermedius nucleus.

5.2.2 Experimental Protocol

During the experiments, patients were comfortably seated in an armchair. The experimental protocol comprised a 2-minutes eyes-open rest recording, a single syllable speech production task, and an upper limb gripping task. For the speech production task, patients were asked to read 120 single syllable words—this paradigm was adapted from (Chrabaszcz et al., 2019a)—at either a loud volume or at the volume corresponding to normal speaking during casual conversation. The visual stimuli consisted of, at the beginning of each new trial, a white volume icon against a black computer screen, lasting for 1 second. This indicated the loudness level (either low or high volume) at which patients had to read out the syllable that was going to be presented subsequently. After this preparatory cue, the screen remained black for 1 second before the single syllable word was presented for two seconds. Participants would read the word aloud, producing speech. After this, the intertrial interval until the next preparatory cue was set to 2 seconds. A schematic of the experimental procedure is shown in Figure 5.1A. The stimuli were created using Psychophysics Toolbox extensions (Brainard, 1997) and played as an mp4 video. The second task consisted of performing a total of 120 trials of upper limb movement by gripping a foam ball in response to a verbal cue by the experimenter, alternating between left and right hands (60 grips per side). Patients were instructed to hold the grip for three seconds with a minimum of 3-second intertrial interval.

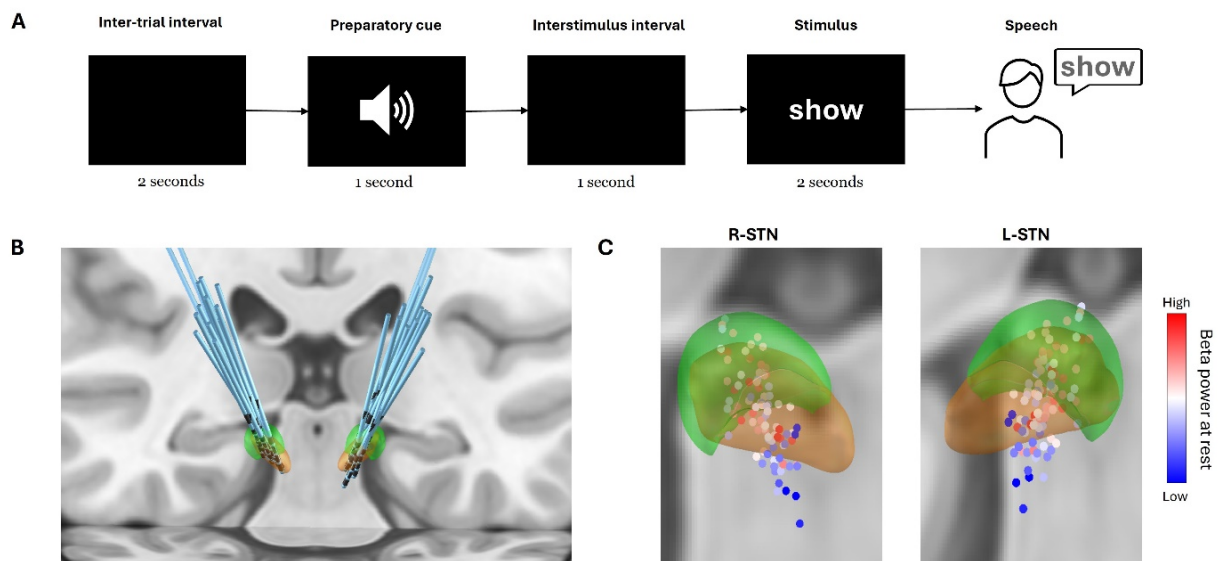


Figure 5.1: *Experimental Paradigm and Lead Localization.* (A) *Experimental paradigm of speech production task.* (B) *Reconstruction of DBS leads into the standard Montreal Neurological institute (MNI) space using the LeadDBS toolbox. Leads are shown to be clustered in the STN (shown in orange) and close to the sweet spot (shown in green) that has been suggested to provide optimal motor improvement for PD patients with DBS (Dembek et al., 2019).* (C) *Point-cloud representation of bipolar contacts plotted in MNI space based on the strength of beta power during rest. Contacts with the strongest beta power cluster around the motor improvement "sweet spot".*

5.2.3 Data Recordings

Monopolar Local Field Potentials (LFP) and Electroencephalography (EEG) were recorded from both hemispheres with a common ground electrode placed on the forearm of the patient. Six EEG electrodes (C3, C4, Cz, CP3, CP4 and CPz) were placed according to the international 10–20 system to cover the primary motor area of the cortex (M1). To synchronise the presentation of the cue in the speech production task, we recorded the signal from a light-sensitive photodiode attached to the corner of the presentation laptop. The stimuli were programmed to emit a brightness pulse under the sensor at the onset of the preparatory and word cue. Participants responses were recorded using an omnidirectional microphone placed on a tripod, positioned at approximately 0.5 meters and 45° angle from the participant. Muscles activity for the gripping task were recorded by placing surface electromyography (EMG) bipolar electrodes on two upper-limb muscles on both sides (the forearm flexor and extensor muscles). Triaxial accelerometers were also taped to the index fingers of both hands to record hand movements. All signals were recorded and amplified using a TMSi SAGA system (TMS International, Netherlands) at 4096 Hz sampling rate.

5.2.4 Electrophysiological Data Processing

Data analysis was performed offline using Spike 2 software (Cambridge Electronic Design, Cambridge, UK), custom-written MATLAB Scripts (version 2023a; Mathworks) and the open-source Fieldtrip toolbox (Oostenveld et al., 2011). Events for speech onset and grip onset were defined visually according to the DC removed, rectified and smoothed (0.1 s) signals of the microphone and EMGs, respectively. Self-paced gripping trials having an intertrial interval of less than two seconds were excluded from the analysis. Due to limited number of trials for the gripping task ($n=13$ grips per side), patient 1 was excluded from the comparative analyses. This resulted in a final sample size of 17 patients (34 STN) for speech production task and 16 patients (32 STN) for the upper-limb movement task.

EEG signals were re-referenced to Cz and LFP bipolars were computed by subtracting adjacent channels from different levels. This resulted in 9 or 7 bipolar contact pairs for directional leads and non-directional leads, respectively, in each hemisphere. LFP and EEG signals were band-passed filtered between 1-100 Hz and band-stopped filtered between 48-52 Hz using a fourth-order zero-phase Butterworth filter, then down-sampled to 256 Hz. Movement-related epochs of 5 seconds were subsequently created around trial onsets of each task $[-2.5 \text{ s}, 2.5 \text{ s}]$. All trials underwent visual inspection, and those containing artefacts were excluded. The average number of trials was 102.7 ± 4.4 (mean $\pm$ SEM) [range 62-119] for the speech production task, 44 ± 3.6 [range 13-60] for the right-hand grips and 46.5 ± 3.65 [range 13-71] for the left-hand grips.

Next, time-series epochs were decomposed into the time-frequency domain, covering frequencies from 1 to 100 Hz with a linear resolution of 1 Hz. This was achieved by applying a second-order IIR peaking resonator digital filter with a 2 Hz bandwidth and a 3 dB attenuation level, followed by the absolute value of the Hilbert transform. To account for movement-related power changes in the STN and motor cortex, the time-frequency power of each bipolar channel in each trial was z-score normalised relative to its baseline power in the pre-onset window, ranging from -2 to -1 seconds before movement onset. Normalised power was then averaged across trials. In acknowledgment of the fact that not all electrode contacts could be expected to lie in the STN, we selected for each task the bipolar channel that showed the greatest event-related desynchronization (ERD) in beta frequency band (13-35 Hz) i.e. the largest movement-related suppression in beta power. If different bipolar contacts were selected for both tasks in each STN, their power/connectivity were averaged for the comparative analysis. Upper-limb movement from the contralateral hand were considered when assessing differences between

tasks. Power modulation over time for different frequency bands (delta-theta: 1-7 Hz, alpha: 8-12 Hz, low-beta: 13-20 Hz, high-beta: 21-35 Hz, and gamma: 40-100 Hz) were also calculated by averaging time-frequency power across the respective frequency ranges.

5.2.5 Functional Connectivity Analysis

To quantify the functional connectivity between STN-LFPs and M1-cortical EEGs, we calculated the Phase Synchronization Index (PSI) (Lachaux et al., 1999) between the selected STN-LFP bipolar contact and each bipolar EEG. Specifically, for each trial and frequency bin between 1 and 100 Hz, PSI was estimated over time by averaging the phase angle differences using a sliding-window technique with a fixed window length of 300 ms centred around the sample at timepoint t and moving in steps of one sample ($3.9 \text{ ms} = 1/\text{sampling rate}$).

To assess task-related connectivity changes, PSIs for each trial were z-score normalised relative to the pre-onset baseline connectivity in the $[-2, -1\text{s}]$ window, and subsequently averaged across trials, following the approach used in the power analysis.

5.2.6 Statistical Analysis

Statistical analyses were performed using custom-written scripts in both MATLAB (version 2023a; Mathworks) and Python. Non-parametric cluster-based permutation test was used to assess significant time-frequency and connectivity modulations compared with baseline as well as to evaluate the significance of movement-related spectral/connectivity changes between tasks (speech vs. gripping) or between hemispheres (left vs. right, contralateral vs. ipsilateral). This test adjusts for multiple comparisons by randomly reassessing condition labels for STNs and correcting p-values using suprathreshold clusters ($p < 0.05$) (Maris & Oostenveld, 2007). Pearson's correlation was used to test whether the degree of modulation is related to the voice level during speech or muscle activity during upper-limb movement.

5.2.7 Phase-Amplitude Coupling (PAC) Analysis

We investigated phase-amplitude coupling (PAC) between cortical EEGs in the theta/alpha band and STN-LFPs in the gamma band using Tort's Modulation Index (MI) method (Tort et al., 2010b) for

rest, speech, and upper-limb movement. For this analysis, signals were band-passed filtered between 1-120 Hz. Rest signals were pre-processed similarly to task signals described previously, then divided into 4-second epochs for PAC calculations. For task data, we focused on a 1-second movement execution interval, by selecting epochs from -250 ms to 750 ms relative to movement onset. A two-way least squares FIR filter was applied to band-pass the EEG signals into low-frequency phase components (4-13 Hz in 1 Hz steps with a 2 Hz bandwidth) and the LFP signals into high-frequency amplitude components (40-120 Hz in 5 Hz steps with a 10 Hz bandwidth). The Hilbert transform was used to extract the instantaneous phase of the low-frequency signals and the amplitude of the high-frequency signals. These phase/amplitude signals were then concatenated across trials, and a phase-amplitude histogram was created using 18 frequency bins of 20 degrees each as recommended in (Tort et al., 2010b). The MI was estimated by calculating the Kullback-Leibler divergence, which measures the difference in entropy between the observed phase-amplitude distribution and a uniform distribution. To normalise the MI, we computed z-scores based on 200 surrogates generated by randomly shuffling the amplitude time blocks. The final z-scored PAC, computed for multiple frequencies of EEG theta/alpha phases and LFP gamma amplitudes, was represented as a comodulogram. This procedure was repeated for each STN bipolar and EEG channel, and the resulting PAC values were averaged to produce the final PAC representation for each task.

5.2.8 ML-Based Analysis

We employed patient-specific logistic regression models to evaluate the ability of subthalamic nucleus signals to distinguish speech from upper limb movements. Each trial’s electrophysiological data underwent the preprocessing routines described previously, ensuring that artefacts were minimised and that segments of interest (trials of speech or gripping) were aligned around their respective onsets. Once these time-locked segments were obtained, we derived spectral features across the frequency bands that exhibited meaningful power modulation or connectivity changes relative to baseline and extracted the features corresponding to their execution periods from -250 ms to 750 ms around movement onset. These features mainly comprised oscillatory power in various sub-bands (see Chapter 4.2.3 for further details).

Patient-specific classifiers were trained for each patient to discriminate between the signals corresponding to either speech or upper-limb movement. For cross-validation, we opted against a leave-one-trial-out scheme because of the unequal number of trials for gripping and speech available

for some patients. A LOTO CV approach requires constructing validation folds that contain both classes that one aims to discriminate. Since the number of trials available for each of the behaviours was different, a LOTO CV implementation was not directly possible. Instead, we used a bootstrapped method that randomly selected a single speech trial and a single upper limb movement trial to comprise the validation set during each fold—leaving the remaining trials in the training set. This procedure was iterated 300 times to ensure that the combinations of speech and movement trials in the validation folds was representative. Classifiers were also trained to distinguish speech from rest, and upper limb movement from rest. For these scenarios, features extracted from the -250 ms to 750 ms segment (relative to motor onset) would be labelled as execution, and features in the -3.0 s to -750 ms would be labelled as rest. A leave-one-trial-out cross-validation scheme was used to evaluate performance for distinguishing rest from task execution.

5.3 Results

5.3.1 Movement-related power modulation of speech and upper-limb movement in the STN and motor cortex

STN-LFP activity exhibited significant time-frequency modulation during speech production and upper-limb movement. During speech, there was a significant decrease in z-scored spectral power within the beta frequency band (13-30 Hz) occurring approximately between [-0.5s, 0.5s] around speech onset ($p < 0.005$) followed by a significant beta power rebound starting 1 second after speech onset ($p < 0.005$). Significant increases in z-scored power were also observed in the delta-theta frequency bands (1-7 Hz) during the interval [-0.5s, 1s] and in the broadband gamma frequencies (40-100 Hz) from -0.35s onward ($p < 0.005$) (Figure 2-A). During upper-limb movement, significant power desynchronization was observed in alpha (8-12 Hz) and beta (13-30 Hz) frequency bands ($p < 0.01$) between [-0.2s, 0.8s] along with significant increases in delta-theta power and in broadband gamma power ($p < 0.01$) (Figure 5.2B).

Considering the differences in execution duration and onset precision between the two tasks, we compared power changes across frequencies during movement execution by averaging the z-scored power over time for each frequency bin within a 1-second interval of maximal power modulation in the

STN ([-0.5s, 0.5s] for speech and [-0.2s, 0.8s] for upper-limb movement). Cluster-based permutation tests showed significant differences in power modulation between tasks across the [1-26 Hz] frequency range ($p < 0.001$). Notably, upper-limb movement showed stronger and more widespread desynchronization, covering both the alpha and beta frequency bands, whereas speech-related desynchronization was limited to the beta band (Figure 5.2C). These results remained significant when considering a common 0.5 second after onset interval ([0s, 0.5s], $p < 0.01$). Note that the pronounced suppression of beta power during upper-limb movement compared to speech, was unlikely related to the strength of muscle activity when performing hand gripping, as no significant correlation was observed between beta activity in the STN and the contralateral rectified EMGs (Left STN - Right EMGs: Pearson's $R = -0.034$, $p = 0.98$; Right STN - Left EMGs: Pearson's $R = -0.15$, $p = 0.57$).

In the motor cortex, speech production was associated with a significant decrease in power in the alpha (8-12 Hz) and low-beta (13-20 Hz) bands, along with an increase in broadband gamma power (40-100 Hz) following speech onset ($p < 0.01$). Regarding upper-limb movement, similar modulation patterns as in STN were observed, but with significant desynchronization across a broader frequency range, including alpha, low-beta, and high-beta bands ($p < 0.01$). These findings were further supported by a cluster-based permutation test on the average power per frequency bin over the 1-second interval of maximal power modulation for EEGs ([-0.25s, 0.75s] for speech and [0, 1s] for upper-limb movement), as the analysis revealed a significant difference between tasks in sub-ranges of the beta band (16-28 Hz, $p < 0.05$), with greater power reductions observed during upper-limb movement compared to speech (Supplementary Figure V).

5.3.2 Movement-related STN-M1 connectivity during speech and upper-limb movement

Functional connectivity between STN-LFPs and each cortical EEG channel was estimated using the Phase Synchronization Index (PSI), z-scored to the baseline, and then averaged across EEG channels to assess STN-M1 time-frequency connectivity changes. During speech, there was a significant decrease in connectivity in the alpha and low-beta (13-20 Hz) frequency bands ($p < 0.05$), contrasted by an increase in connectivity in the delta-theta and sub-ranges of gamma band (40-70 Hz, $p < 0.05$, Figure 2-D). For upper-limb movement, STN-M1 connectivity modulation showed a reduction in connectivity in alpha and in sub-ranges of beta band (13-26 Hz, $p < 0.05$), along with increased connectivity in the delta-theta and sub-ranges of high-beta and gamma frequencies (26-50 Hz, $p < 0.05$, Figure 5.2E).

A cluster-based permutation test on the averaged connectivity across frequency bins, within the same intervals of maximal power modulation used for power analysis, revealed a significant difference in STN-M1 connectivity changes in a sub-range of the high-beta band (27-32 Hz, $p < 0.01$). Specifically, speech exhibited decreased connectivity, while upper-limb movement showed increased connectivity in this range (Figure 5.2F).

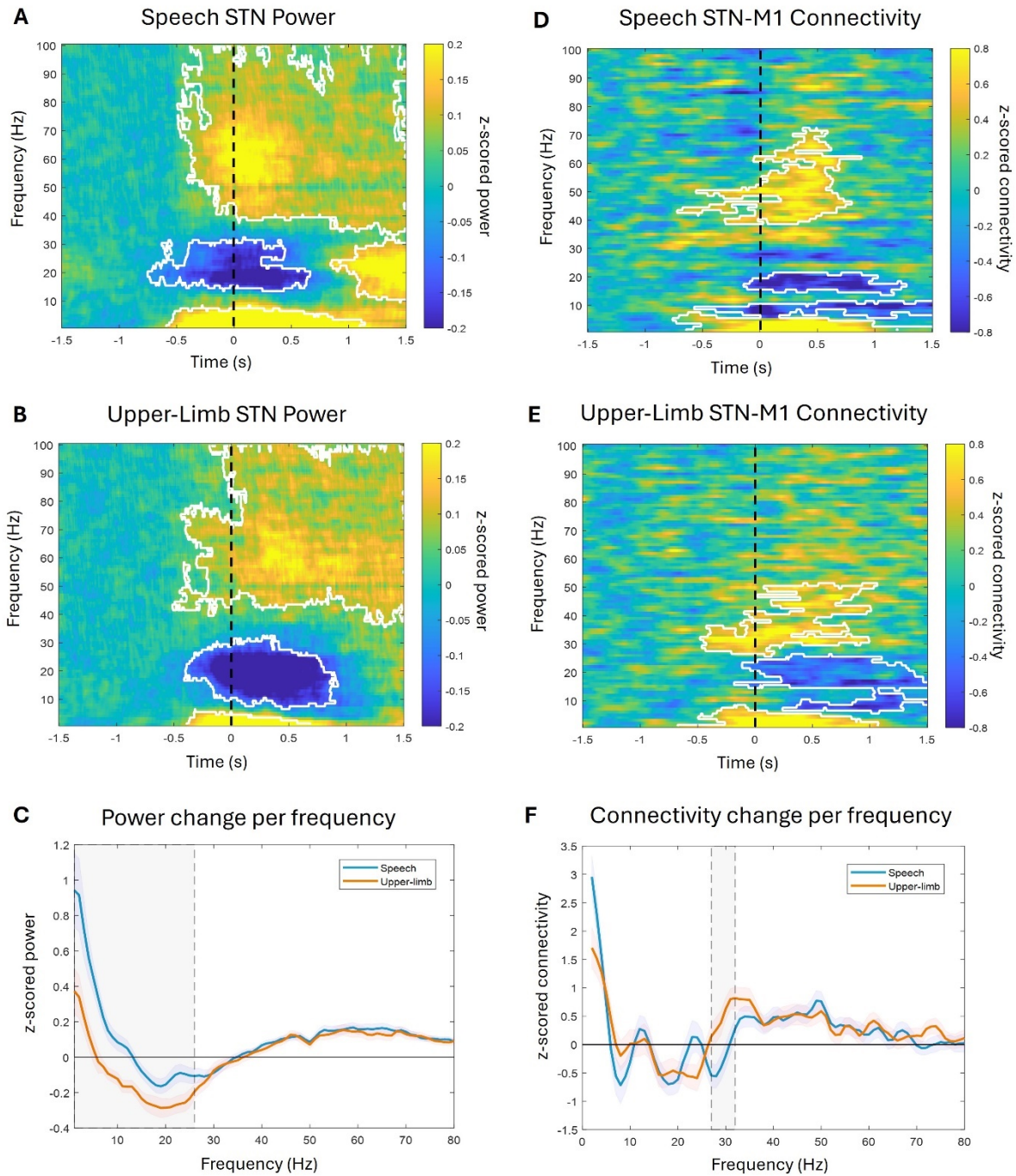


Figure 5.2: Movement-related spectral and functional connectivity modulations in the STN. Grand average of z-scored spectral-power relative to baseline across all STNs during **A)** speech production ($n=34$) and **B)** upper-limb movement ($n=32$). **C)** Differences in movement-related averaged z-scored spectral changes relative to baseline per frequency bin between speech production task (blue) and upper-limb movement (orange) during 1s of maximum modulation. Grand average of z-scored STN-M1 functional connectivity relative to baseline during **D)** speech production and **E)** upper-limb movement. **F)** Differences in movement-related averaged z-scored functional connectivity changes relative to baseline

per frequency bin between speech production task (blue) and upper-limb movement (orange) during 1s of maximum modulation. Clusters with significant power/ connectivity modulations (cluster-based permutation test, $p < 0.05$) are marked with white contour. Gray bands marked the frequency range of significant differences in modulation between speech and upper-limb movement using cluster-based permutation analysis ($p < 0.05$). Values in C and F are represented as mean $\pm$ SEM.

5.3.3 Lateralization of speech and upper-limb movement in the STN and motor cortex

To assess the lateralization of the STN during speech production and upper-limb movement, task-related modulations were compared between hemispheres, left-STN vs. right-STN for speech ($n=17$) and contralateral-STN vs. ipsilateral-STN for upper-limb movement ($n=32$). No significant differences were observed between the left and right STN during speech, as both hemispheres showed comparable beta power suppression around movement onset (Figure 5.3A,C). In contrast, the STN contralateral to upper limb movement exhibited greater and more prolonged beta suppression than the ipsilateral STN ($p < 0.05$, Figure 5.3B,D).

A similar analysis was conducted on the motor cortex, with the left M1 represented by the C3 and CP3 EEG channels, and the right M1 by the C4 and CP4 channels. Similar to the STN, speech production did not exhibit lateralization in the motor cortex, as there were no significant differences in power modulation between the left and right hemispheres. However, during upper-limb movement, power modulations between contralateral and ipsilateral M1 were comparable, indicating a lack of lateralization in the motor cortex. Further investigations were also carried out to examine the lateralization of functional connectivity between STN and the motor cortex for both tasks. No significant differences were observed when comparing STN-contralateral-M1 and STN-ipsilateral-M1 connectivity, suggesting a lack of lateralization of STN-M1 functional connectivity for both tasks.

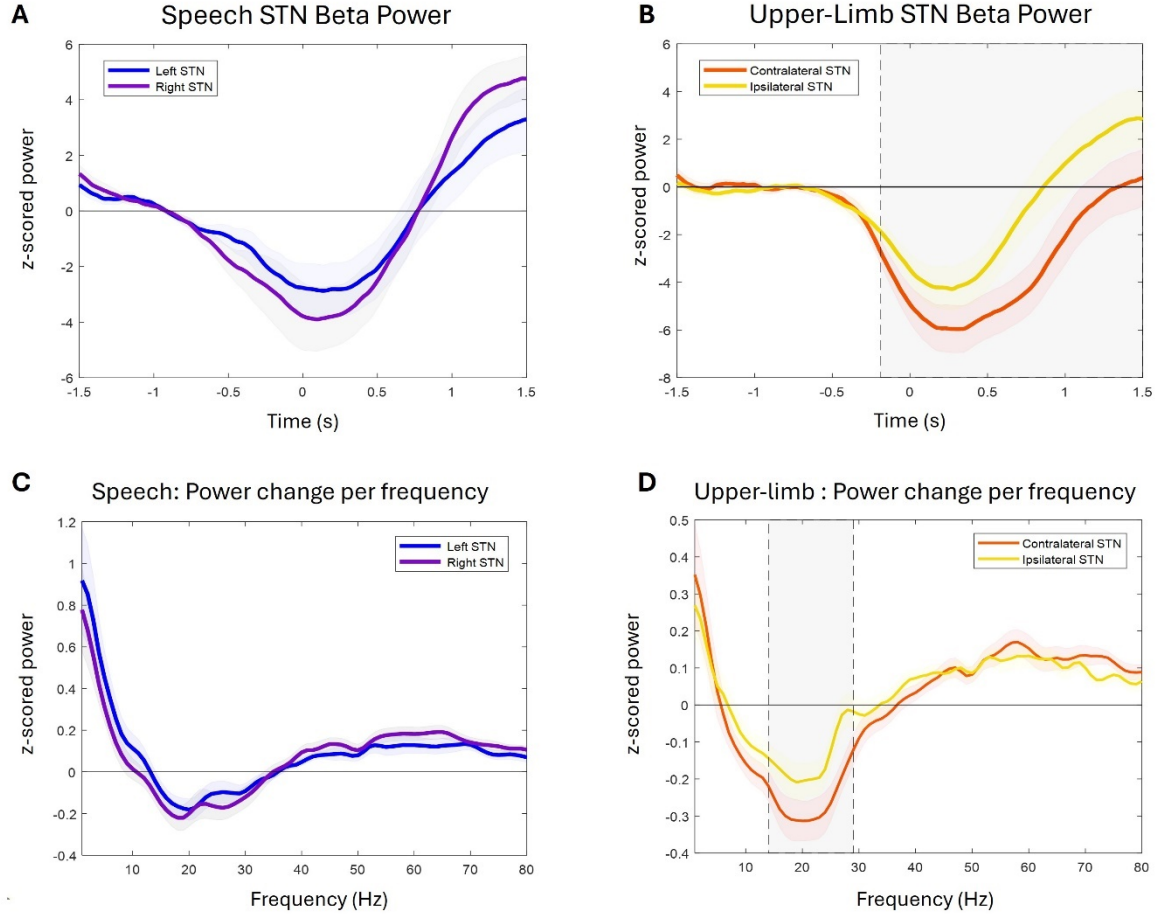


Figure 5.3: Lateralization of speech and upper-limb movement in the STN. Average z-scored beta power in time for (A) left (dark blue) vs. right (purple) STN ($n=17$) during speech production and (B) contralateral (red) vs. ipsilateral (yellow) STN ($n=32$) during upper-limb movement. Differences in movement-related averaged z-scored spectral changes relative to baseline per frequency bin during 1s of maximum modulation between (C) left vs. right STN for speech production task and (D) contralateral vs. ipsilateral hemispheres for upper-limb movement. Gray bands marked the time interval and frequency range of significant differences between compared STNs using cluster-based permutation analysis ($p<0.05$). Values are represented as mean $\pm$ SEM.

5.3.4 Spatial distribution of band-specific oscillations distinguishing rest from speech and from upper-limb movement within the STN

To determine the spatial distribution of bipolar contacts that performed best in distinguishing speech/hand gripping from rest, or speech from hand gripping, we used the AUC values obtained from the previous analysis and mapped them into the MNI-space (see Supplementary Figure VI - VIII). We

observed similar distribution patterns across frequency bands and conditions. Notably, clusters of high AUC values were highly overlapped with the cluster of bipolar contacts that exhibited the highest beta power during rest. This led us to adopt a combined frequency band visualization approach, which provided a clearer representation of the spatial distribution of task-specific oscillatory activity within the STN. To do so, we compared AUC values derived from individual frequency bands (theta, alpha, beta, or gamma) and assigned each bipolar contact to the frequency band with the highest AUC. Bipolar contacts with the highest AUC values in the theta and alpha bands were combined into a single group, as these bands had shown similar performance during the previous analysis. When distinguishing speech from rest, 19% of the bipolar contacts showed the highest AUC values in the gamma band, particularly in the ventro-posterior part of the STN. In contrast, 41% of the contacts, primarily in the ventro-anterior region, exhibited dominance in the theta and alpha bands. Beta oscillations with 40% of contacts showed a more uniform distribution throughout the STN (Figure 5.4A). For upper-limb movement vs. rest, beta oscillations outperformed other frequencies in 55% of the bipolar contacts, though no distinct spatial distribution patterns were observed in any frequency band (Figure 5.4B).

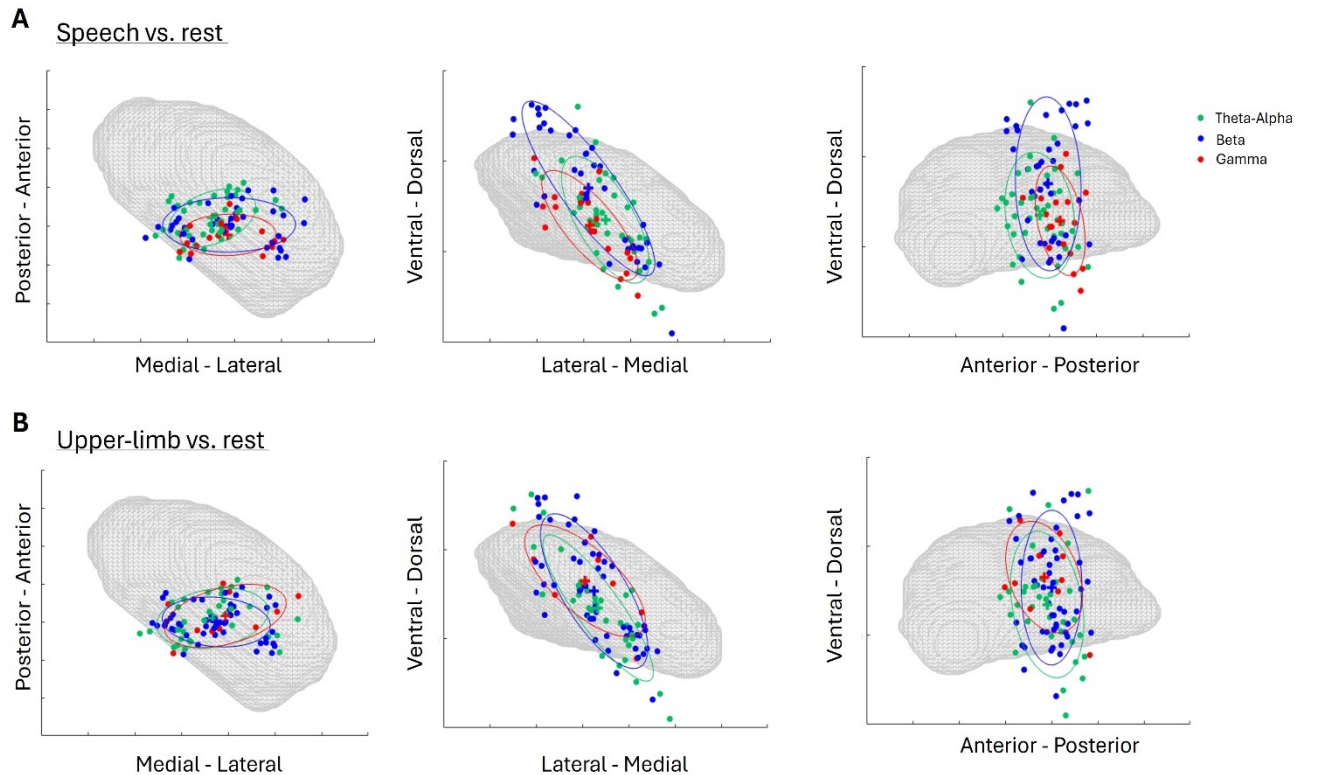


Figure 5.4: Spatial distribution of preferred band-specific oscillatory activity within the STN that distinguishes between rest and task execution. Outcome of comparing AUC values obtained from spectral

features of different frequency bands (theta/alpha, beta, gamma) for (A) speech production vs. rest and (B) upper-limb movement vs. rest. Bipolar contacts are assigned to the group with highest AUC value and groups are plotted into the STN with three different views. Theta/alpha group is represented by green dots, beta group with blue dots, and gamma group with red dots.

5.3.5 Phase amplitude coupling between cortical EEG and STN-LFPs during speech and upper-limb movement

We investigated whether the amplitude of gamma oscillations in the STN was locked to the phase of theta/alpha oscillations in the motor cortex during rest, speech, and upper-limb movement. The grand-average PAC comodulograms across all recordings (n=32 STN) are presented in Figure 5.5A. In both movement-related tasks, PAC between the motor cortex and the STN was strongest between the theta phase (4-7 Hz) and the higher gamma amplitude (60-120 Hz). The difference in PAC between speech and upper-limb was represented by subtracting their corresponding PACs and a permutation-based cluster analysis was used to assess the statistical significance between the two tasks. The analysis revealed that speech exhibited significantly higher theta (4-7 Hz) to high gamma (70-110 Hz) PAC than upper-limb movement ($p < 0.001$, Figure 5.5B). Both tasks showed higher z-scored MI values in these frequencies compared to rest, which was confirmed by a post-hoc Wilcoxon signed rank test (rest vs. upper-limb, $p < 0.01$; rest vs. speech, $p < 0.001$; upper-limb vs. speech, $p < 0.01$, Figure 5.5C). While weaker alpha (8-12 Hz) to gamma (40-120 Hz) PAC was observed during speech compared to upper-limb movement (Figure 5.5B), the difference was not statistically significant.

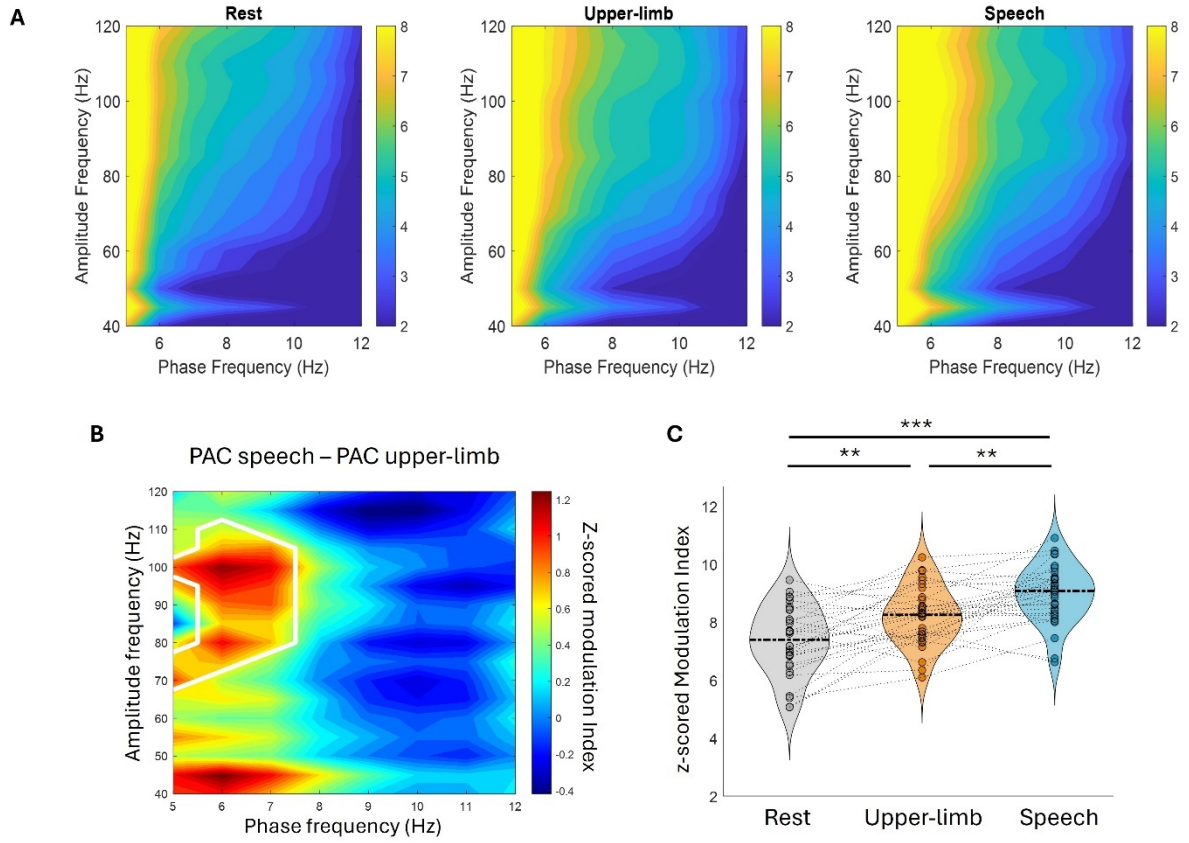


Figure 5.5: Theta/alpha - gamma Phase Amplitude Couplings (PAC). **A)** Group-level averaged PAC comodulograms between cortical EEG and STN-LFPs during rest, speech production and upper-limb movement. **B)** Difference in PACs between movement-related tasks obtained by subtracting PAC of speech from PAC of upper-limb movement. Significant differences obtained from cluster-based permutation test ($p < 0.05$) are marked by white contour. **C)** Violin plot representing z-scored modulation index calculated for theta phase (4-7 Hz) and high-gamma amplitude (70-110 Hz) for rest, upper-limb movement and speech. Post-hoc Wilcoxon: ** $p < 0.01$, *** $p < 0.001$.

5.3.6 Decoding Performance

We investigated the performance of logistic regression classifiers trained to, within patient, distinguish timeseries excerpts corresponding to gripping from those corresponding to speech. Classifiers trained with all canonical bands yielded an average AUC of 0.74 ± 0.11 (mean $\pm$ std. deviation). Classifiers trained on activity from a single canonical bands yielded lower decoding performance: 0.69 ± 0.09 for the theta band, 0.63 ± 0.09 for the alpha band, 0.63 ± 0.07 for the low beta band, 0.65 ± 0.09 for the high beta band, 0.60 ± 0.08 for the gamma band, and 0.59 ± 0.09 for high frequency oscillations (see Figure 5.6).

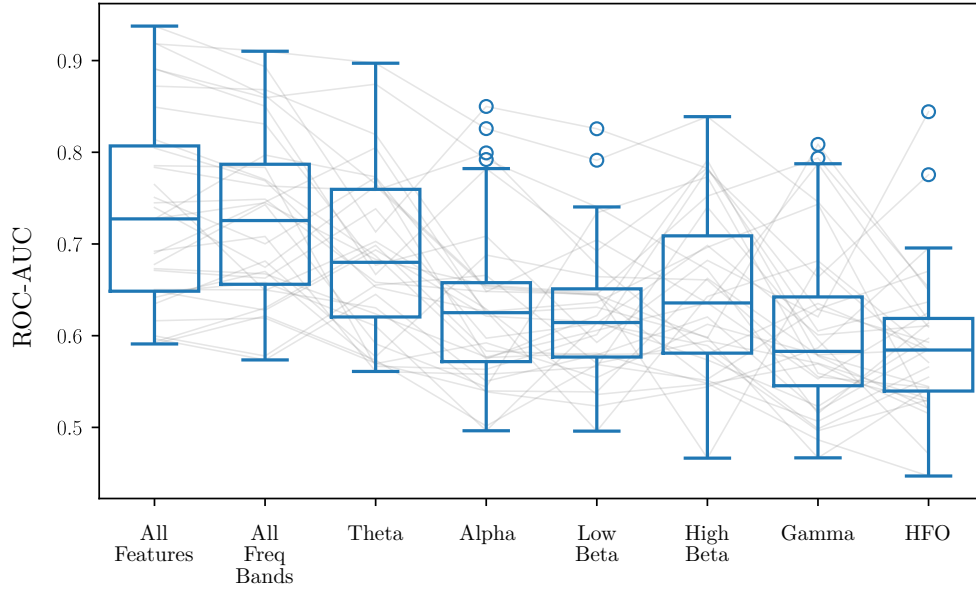


Figure 5.6: ROC-AUC Values for Distinguishing Gripping from Speech Periods. Each datapoint corresponds to one hemisphere. Boxplots represent AUC distributions for different feature sets (namely different canonical bands) used to drive classification. The highest classification performance, with an ROC-AUC of 0.74 ± 0.11 (mean $\pm$ std. deviation) is achieved when using the complete feature set (All Features).

5.4 Discussion

In this study, we aimed to investigate the electrophysiological correlates across the STN and motor cortex of PD patients associated with the execution of two distinct kinds of motor activity: speech production and upper-limb movement. We aimed to harness these correlates as part of a machine learning pipeline to distinguish speech from upper-limb movement execution based on these electrophysiological signatures. To this end, we analysed simultaneous STN-LFPs from externalised DBS leads and cortical EEG signals recorded while PD patients performed these tasks. Our findings revealed discriminatory oscillatory and connectivity patterns, with upper-limb movement exhibiting (1) stronger and broader STN desynchronization in the alpha and beta bands, (2) stronger STN-M1 coupling in the high-beta band, and (3) weaker M1-theta-phase to STN-gamma-amplitude coupling, compared to speech production. Speech was also characterised by a lack of lateralization between hemispheres, unlike upper-limb movement, which showed more prolonged modulation in the contralateral STN than in the ipsilateral STN. Spatial analyses revealed that, during speech

production, focal gamma oscillatory activity was localised in the ventro-posterior part of the STN. during upper-limb movement, oscillatory activity across all frequency bands remained spatially dispersed. Using spectral features and a logistic regression model driven by activity in canonical frequency bands, we were able to successfully differentiate speech from upper-limb movement with an average AUC of 0.74 (SD: 0.1).

5.4.1 Sustained spectral patterns in STN/motor cortex for speech and upper-limb movement

As no previous studies have directly compared the STN/motor cortex event-related oscillatory activity and functional connectivity for speech production vs. upper-limb movement, we must compare our observed patterns with previous studies reporting modulation of both tasks separately. Our findings highlighted similar patterns of STN activity during speech and upper-limb movement but with a notable difference in the scope of desynchronization. Specifically, upper-limb movement was characterised by a broader desynchronization across both alpha (8–12 Hz) and beta (13–30 Hz) bands, while speech production showed a more selective beta-band desynchronization. This distinction aligns with existing research showing that alpha and beta desynchronization in the STN are consistent markers of active motor states (Lofredi et al., 2018; Lipski et al., 2017; Kempf et al., 2007; Androulidakis et al., 2007; Geng et al., 2018), mainly when performing upper-limb movement (Anzak et al., 2012; Tan et al., 2013, 2016; Tinkhauser et al., 2019), while beta desynchronization in STN-LFPs and single-unit firing activity have been shown to also characterise speech production (Chrabaszcz et al., 2019b, 2021b; Hebb et al., 2012b; Lipski et al., 2018b; Watson & Montgomery, 2006). In general, beta-band desynchronization has been linked to the release from maintaining static behaviour in movements (Alhourani et al., 2020; Holt et al., 2019; Lipski et al., 2017; Müller et al., 2003), a process also shown to occur during speech production (Chrabaszcz et al., 2019b; Hebb et al., 2012b; Lipski et al., 2018b). In addition to the desynchronization patterns, we observed increased gamma-band activity (40–100 Hz) during both tasks, aligning with previous studies linking gamma power increases to both motor planning and execution phases – as articulatory and motor precision in speech – in the motor cortex (Cheyne & Ferrari, 2013; Chrabaszcz et al., 2019b; Crone et al., 1998, 2011; Szurhaj et al., 2006) and in the STN (Androulidakis et al., 2007; Chrabaszcz et al., 2019b, 2021b; Dastolfo-Hromack et al., 2022; Fischer et al., 2017, 2020; Lipski et al., 2018b). We also observed a stronger M1 beta desynchronization during upper-limb movement compared to speech. However, we should note an important limitation of this study, which is the restricted cortical EEG

coverage, which limits our ability to examine task-specific differences in language-related cortical areas, such as the sensorimotor cortex, auditory cortex, and inferior frontal gyrus. This constraint is due to challenges in EEG placement post-surgery, where scars and bandages hinder access to these regions.

5.4.2 STN-Motor Cortex Functional Connectivity for Speech and Upper-Limb Movement

Research indicates that several cortical areas, including the motor cortex, project directly to the STN via a "hyperdirect" pathway, as initially described by (Nambu et al., 1996) and (Kelly & Strick, 2004). The existence of this pathway in humans has been confirmed through recent studies using evoked potentials in premotor, supplementary motor, and primary motor regions following STN stimulation (Brunenberg et al., 2012; W. Chen et al., 2020; Jorge et al., 2020b; Miocinovic et al., 2018). Our findings revealed established event-related STN-M1 functional connectivity for both tasks, which exhibited similar general synchronization patterns to those observed in spectral power of STN and motor cortex separately. Similar coupling patterns have been reported in phase-locking analysis studies of contralateral upper-limb movement, where movement-related decoupling in beta-band oscillations between the STN and motor cortex occurred around movement onset (Cao et al., 2024; Lalo et al., 2008; Lipski et al., 2017), and increased cortico-subcortical synchronization were reported in the gamma band (Fischer et al., 2020; Lalo et al., 2008; Litvak et al., 2012). In speech, however, few studies have examined STN-motor cortex connectivity. Manes and colleagues (Manes et al., 2013) reported significant unilateral co-activation between the STN and the supplementary motor cortex in their meta-analysis, while (Dastolfo-Hromack et al., 2022) found a correlation between high STN-precentral gyrus synchronization in the beta band and impaired articulatory precision in speech. Building on these findings, the observed differences in STN-motor cortex high-beta (26–32 Hz) connectivity in our study—where decoupling after movement onset was characteristic of speech and coupling characterised upper-limb movement— may indicate a frequency-specific modulation in the STN-motor cortex circuit that supports task-specific motor control demands. This frequency-specific difference aligns with Lipski and colleagues' findings of suppressed low-beta (13–20 Hz) spike phase-locking, rather than high-beta, between the STN and motor cortical areas during contralateral hand movement (Lipski et al., 2017).

5.4.3 Frequency-specific spatial segregation of speech and upper-limb movement in the STN

Our findings revealed that, when differentiating either speech or upper-limb movement from rest, or speech from upper-limb movement, no discernible spatial distribution emerged in any individual frequency band. The highest performing contacts in all bands highly overlapped with the cluster of highest beta power at rest. Proximity to the sensorimotor part of the STN is shown to be relevant in clinical contexts, as closer contacts have been associated with improved stimulation efficacy and better clinical outcomes (C. C. Chen, Pogossyan, et al., 2006a; Horn et al., 2017; Lei et al., 2022; Tinkhauser, Pogossyan, Debove, et al., 2017). Thus, our results suggest that proximity to sensorimotor region may also provide the best performance in differentiating between speech and upper-limb movement, regardless of the frequency of the oscillations. Additionally, we detected a focal spatial segregation of gamma oscillations for speech—but not for upper-limb movement—in the ventro-posterior part of the STN. This is consistent with Chrabaszcz and colleagues work, who reported that high-gamma activity during speech correlated with recording sites along both the ventral-dorsal and anterior-posterior axes (Chrabaszcz et al., 2019b). The focality of gamma oscillations, compared to slower rhythms like beta and alpha, has also been observed in decision-making and cognitive tasks (Donner et al., 2009; X.-J. Wang, 2010), unlike movement-related gamma oscillations during upper-limb movement which were shown to be more dispersed across the STN without a clear spatial pattern (Geng et al., 2018).

5.4.4 Cortico-STN PAC during Speech and Upper-Limb Movement

Cross-frequency PAC has been identified as a prominent pathophysiological feature of PD within both the STN (de Hemptinne et al., 2013; López-Azcárate et al., 2010; Meidahl et al., 2019b; Ozturk et al., 2020; Shreve et al., 2017; van Wijk et al., 2016) and the sensorimotor cortex (de Hemptinne et al., 2013; Gong et al., 2021, 2022; Yin et al., 2022). PAC functions as an integrative mechanism in which the phase of low-frequency oscillations temporally modulates the amplitude of high-frequency oscillations (Canolty et al., 2006b; Canolty & Knight, 2010; de Hemptinne et al., 2013; Fujisawa & Buzsáki, 2011). Our study shows that theta power (4–7 Hz) in the motor cortex is more coupled with STN high-gamma power (70–110 Hz) during speech than during upper-limb movement. Building on previous research on PAC in speech, the coordination between theta phase and gamma power oscillations has been identified as an essential neural mechanism in the auditory cortex, where theta phase tracks syllabic rhythms and modulates phonemic processing over time via gamma oscillations

(Giraud & Poeppel, 2012; Hyafil et al., 2015; Lizarazu et al., 2019, 2023). Consequently, this could suggest that such mechanisms may extend beyond the auditory cortex processing and reflect the engagement of cortico-basal ganglia networks in speech production. By contrast, studies investigating PAC in PD during limb movements have primarily reported abnormal patterns of alpha-gamma and beta-gamma PAC associated with impairments in motor initiation and execution, as seen in tasks like walking and finger-tapping (AuYong et al., 2018; Farokhniaee et al., 2024; Gong et al., 2022; Yin et al., 2022). Although our analysis focused only on the coupling of theta-alpha phases with gamma amplitude and on the movement execution interval, the non-significant differences in alpha-gamma PAC between speech and upper-limb movements may suggest that alpha-gamma coupling in the cortico-STN network support a general motor processing function rather than a task-specific specialization.

5.4.5 Limitations and Future Work

Although this work provides insights into contrasting neural correlates of speech and upper-limb movement in Parkinson’s disease, several limitations must be acknowledged. Using a single syllable reading task offered a degree of experimental control, yet prosodic variations and spontaneous vocalizations are largely absent from such an approach. These short vocalisations cannot capture the characteristics of everyday naturalistic speech, including features like intonation, rapid turn-taking, or conversational adjustments. The smaller number of trials available for the gripping task also limited statistical power when comparing it directly with the speech condition, and it also leaves open questions about how generalisable the findings in this work may be for other forms of limb movement.

An additional limitation stems from our restricted EEG montage, which—due to the practical constraints of working with externalized patients—focused on the primary motor region. Although this configuration successfully captured movement-related oscillations, it excluded language-specific areas such as Broca’s region and the superior temporal gyrus. Activity in these sites likely reflects speech planning and auditory feedback processes that could provide valuable signals for reliable automatic speech detection that is actionable from the perspective of the design of aDBS systems.

Our analyses also focused on laboratory-based tasks recorded during the postoperative period when patients were still externalised. Although this setup was invaluable for collection of signals from the subthalamic nucleus, it is nonetheless a constrained clinical environment that substantially differs from

real-life conditions in which speech occurs amidst variable medication states, emotional contexts, and other relevant environmental factors. Prolonged ambulatory recordings, collected while patients engage in activities of daily living, could offer an understanding of how subthalamic nucleus signals support the continuous modulation of speech and limb movements across a range of naturalistic contexts.

Finally, although the logistic regression-based decoding approach achieved promising classification of speech and limb movement, the results presented here remain preliminary. Our analyses of power modulation and connectivity have primarily focused on group-level correlates, and the classifiers trained thus far serve only as a proof-of-concept.

An important consideration for future decoding efforts is how hemispheric lateralization of STN activity might influence classification performance. In our current analysis, we pooled features from both hemispheres without explicitly testing whether the decoder relied on differences between the two sides. Yet, as shown in this work, speech-related desynchronization in the STN is generally bilateral, whereas upper-limb movement produces a contralateral-dominant suppression. A naïve classifier could exploit this simple inter-hemispheric contrast to distinguish speech from movement without tapping into the richer, mechanistically relevant features we wish to harness.

To probe this possibility, one straightforward approach would be to train parallel decoding pipelines using only contralateral leads or only ipsilateral leads and compare their performance. If a model trained solely on contralateral contacts still differentiates speech from limb movement with high accuracy, it would suggest that lateralization per se plays a dominant role. Conversely, a large performance drop would indicate that other spatial or spectral features are carrying the discriminative signal. Complementarily, examining the feature weights or interpretability maps of our logistic regression (or future deep-learning) models could reveal the relative contributions of each hemisphere's signals. These analyses would indicate whether the decoder implicitly privileges hemispheric asymmetries or truly integrates more nuanced temporal and frequency-specific patterns.

It is worth noting, however, that even if lateralization proves to be a strong cue, this does not necessarily undermine the translational value of our findings. From a clinical standpoint, what matters is the reliability and robustness of the biomarkers under real-world conditions. As long as hemisphere-based desynchronization remains a stable signature of upper-limb movement—and its absence reliably indicates the production of speech—speech—an adaptive DBS system that leverages this feature could still function effectively. In fact, an ideal aDBS controller might combine multiple complementary

signals—lateralization differences alongside focal gamma bursts or distinctive connectivity patterns—to maximize detection accuracy and minimize false positives.

Ultimately, by explicitly dissecting the role of lateralization in model performance, we can ensure that our decoding strategies rest on well-understood, mechanistically meaningful neural features. This clarity will not only strengthen the interpretability and safety of future closed-loop protocols but will also guide the design of more sophisticated algorithms that seamlessly integrate spatial, spectral, and temporal dimensions of STN activity.

In future work, we will need to capture and harness the neural patterns observed in this work within individualised machine learning pipelines, systematically benchmark decoding performance, and identify the key drivers of classification performance. This will be necessary to evaluate how these neural correlates might be deployed in speech-specific adaptive DBS protocols. Decoding subthalamic signals for real-time adjustments in stimulation will also likely require refinements in artefact management, optimization of model architectures, and validation in ecologically valid settings.

5.5 Conclusion

In this chapter, we examined how subthalamic and cortical signals in people with Parkinson’s disease differentiate speech production from upper-limb movement. By analysing data from externalised deep brain stimulation (DBS) leads and scalp electroencephalography (EEG), we identified distinct yet partly overlapping patterns of oscillatory modulation in the STN and motor cortex. Both speech and gripping were accompanied by beta-band desynchronization in subthalamic local field potentials, but upper-limb movement showed a broader and stronger desynchronization that extended into the alpha range. Connectivity in the high-beta band was stronger for upper limb movement, whereas speech relied more selectively on beta-band modulations. Speech showed no clear lateralization in the STN, whereas upper limb gripping elicited asymmetric contralateral suppression of STN beta power.

A preliminary spectral feature-based logistic regression decoder successfully distinguished speech from upper-limb movement, suggesting that task-specific time-frequency characteristics can be leveraged for machine learning classification. It is important to emphasise that these classification outcomes remain preliminary. They reflect proof-of-concept findings rather than a fully optimised pipeline for

individualised decoding. Spatial analysis suggested that bipolar contacts near the sensorimotor zone of the STN—marked by strong beta-band power at rest—generally performed best in discriminating tasks, regardless of frequency band. A subset of focal gamma oscillations in the ventro-posterior STN was unique to speech, potentially reflecting specialised articulatory motor demands.

While these observations point to meaningful ways in which speech and limb movements diverge in their spatial and spectral neural correlates, they are still based on data collected in a constrained clinical environment. For future adaptive DBS approaches, it will be crucial to build on these preliminary results by incorporating broader individualization, refining artefact management, and benchmarking decoding performance in more naturalistic conditions. As speech deficits remain a major challenge for subthalamic DBS, the possibility of tuning stimulation parameters selectively during speech—based on features such as gamma or beta oscillations—warrants further exploration. The promise of a speech-specific aDBS paradigm hinges on gathering more extensive evidence regarding the robustness of these biomarkers, as well as validating how precisely they can guide real-time stimulation in broader clinical and everyday settings.

5.6 Summarised Findings

- **Movement-Related Beta Desynchronization:** Both speech and upper-limb movement involve clear beta-band suppression in the STN and motor cortex, but limb gripping elicits broader desynchronization that includes alpha frequencies.
- **Distinct Connectivity Profiles:** While both tasks show increased delta/theta- and gamma-band power, as well as STN–M1 synchronization, upper limb movement elevates high-beta connectivity between the subthalamic nucleus and cortex. Speech, by contrast, exhibits reduced high-beta coupling and no strong lateralization in the STN.
- **Spatial Distribution and Task-Specific Oscillation:** Proximity to sensorimotor regions of the STN (dorsolateral STN) correlates with higher classification performance, suggesting these regions encode the strongest task signatures. Speech-related gamma activity is focal in ventro-posterior subregions, whereas limb movement patterns are more diffusely distributed across alpha, beta, and

gamma.

- **Machine Learning Classification and Implications for Adaptive DBS:** Preliminary logistic regression classifiers incorporating spectral features from the STN can distinguish speech from limb movement with an average AUC above 0.7, suggesting that task-specific signals in subthalamic LFPs may be decodable for real-time adaptive DBS. However, these results are still preliminary and require further analysis to capture and benchmark all relevant modulatory patterns, as well as to determine their practical clinical utility—particularly in tailoring stimulation parameters to address speech-related side effects in Parkinson’s disease.

6 Project 4: Feasibility of Fronto-Thalamic Connectivity–Based Predictive Biomarkers for Tourette Syndrome Tics

6.1 Introduction

Tourette Syndrome (TS), first documented by Georges Albert Édouard Brutus Gilles de la Tourette in 1885, is a neurodevelopmental condition characterised by the presence of motor and vocal tics that typically emerge in early childhood or adolescence (McNaught & Mink, 2011). Epidemiological studies estimate a prevalence of TS of approximately 0.3–0.9% in school-aged children, with symptom onset generally between ages 5 and 7, and peak severity in late childhood (ages 8–12). Although many patients experience attenuation of tics by late adolescence, a subset continues to be affected into adulthood, with up to 23% of affected patients continuing to have persistent and sometimes severe tics in adulthood and only 17% experiencing complete remission (Johnson et al., 2023; Swain et al., 2007). Tics range from simple, brief movements (such as blinking) to more complex behaviours and are often preceded by premonitory urges. In TS, the presence of comorbid psychiatric conditions is common—most frequently attention-deficit/hyperactivity disorder (ADHD) and obsessive-compulsive disorder (OCD)—and can exacerbate functional impairment as well as impact the quality of life of patients (Hirschtritt et al., 2015; Swain et al., 2007).

6.1.1 Involvement of the Cortico-Basal Ganglia-Thalamic-Cortical (CBGTC) Circuit in the Pathophysiology of TS

Tourette Syndrome (TS) is increasingly recognised as a network-level disorder involving altered structure and functioning of cortico-basal ganglia-thalamic-cortical (CBGTC) circuits (McNaught & Mink, 2011). Under normal conditions, these circuits integrate excitatory cortical inputs with inhibitory signals in the basal ganglia and thalamus, regulating the selection, initiation, and inhibition of motor commands (Brown & Marsden, 1998; Mink, 2003). In TS, disruption of these gating mechanisms may give rise to excessive or ill-timed movement and vocalizations, which manifest clinically as tics (Caligiore et al., 2017; Mink, 2001).

Extensive evidence underscores the central role that CBGTC dysfunction plays in TS pathophysiology (Caligiore et al., 2017; Leckman et al., 2010). Dysregulated dopaminergic transmission in the basal ganglia has been linked to tic exacerbation (Shale et al., 1986; Singer et al., 2002), whereas dopamine-receptor antagonists (such as haloperidol and aripiprazole) have been shown to alleviate tic severity (Shapiro & Shapiro, 1968; Yoo et al., 2011). Neuroimaging data further reinforce the concept of CBGTC network abnormalities in TS. Functional MRI (fMRI) studies indicate that voluntary tic suppression is associated with heightened activity in the prefrontal cortex, basal ganglia, and thalamus—an effect that is inversely correlated with tic severity (B. S. Peterson et al., 1998). Similarly, (Worbe et al., 2012) reported that premotor, sensorimotor, parietal, and cingulate cortices, along with the medial thalamus and insula, show functional disruptions that correlate with tic severity and complexity. Complementary positron emission tomography (PET) studies reveal altered striatal–frontal connectivity (Jeffries, 2002) as well as distinct metabolic patterns in basal ganglia and frontal structures linked to tics and obsessive–compulsive symptoms (Pourfar et al., 2011), underscoring multiple facets of CBGTC involvement in TS pathophysiology.

Additional evidence of distributed CBGTC network dysfunction comes from event-related fMRI data. (Bohlhalter, 2006) found that paralimbic and sensory association regions (including the anterior cingulate, insula, and supplementary motor area) activate just before tic onset, while sensorimotor regions are recruited at tic onset. Other studies similarly highlight that tic-related activity often spans neocortical, paralimbic, and subcortical regions (Stern et al., 2000). Dopaminergic modulation with levodopa can normalise some of these atypical activation patterns (Hershey et al., 2004), highlighting the role of dopamine-dependent pathways in the CBGTC network in TS.

6.1.2 Deep Brain Stimulation (DBS) as a Treatment for Tourette Syndrome

For individuals with medically refractory TS, deep brain stimulation (DBS) has emerged as a promising therapeutic intervention. DBS involves the surgical implantation of electrodes into deep brain structures such as the thalamus or globus pallidus internus (GPi), two key nodes in the CBGTC network implicated in tic generation. By delivering focal electrical stimulation to these targets, DBS seeks to modulate dysregulated activity in the circuit, potentially restoring more adaptive network dynamics (Lozano & Lipsman, 2013). Indeed, thalamic DBS has been shown to modulate dopamine transmission in the striatum—suggesting a mechanistic pathway for the therapeutic effect of DBS in

TS (Kuhn et al., 2012; Rusheen et al., 2023; Vernaleken et al., 2009)– as well as to modulate oscillatory activity in cortical motor areas beyond the thalamus (Schüller et al., 2024).

A growing body of clinical data attests to the therapeutic potential of DBS for TS, with systematic reviews and meta-analyses suggesting that DBS can reduce tic severity by up to 50% in carefully selected patients (Baldermann et al., 2016; Casagrande et al., 2019; Martinez-Ramirez et al., 2018; Wehmeyer et al., 2021). Nevertheless, outcomes can vary significantly across patients, highlighting key challenges related to patient selection, electrode placement, stimulation parameters, and stimulation-induced side effects, and adverse events (Cagle et al., 2020; Casagrande et al., 2019; Martinez-Ramirez et al., 2018; Schrock et al., 2015; Servello et al., 2010). Identifying reliable, network-level biomarkers of TS could help enhance the efficacy and precision of DBS by informing patient screening, optimising programming, and potentially paving the way for adaptive stimulation paradigms. This goal aligns with recent advances in adaptive DBS (aDBS) in other neurological conditions, notably Parkinson’s Disease, where – by more precisely targeting pathological neural states – stimulation that is dynamically adjusted based on real-time physiological feedback has already shown promise in improving therapeutic efficacy while reducing stimulation-induced side effects (Little et al., 2016; Piña-Fuentes et al., 2020; Rosa, Arlotti, et al., 2017). Translating this approach to TS requires the discovery and validation of robust biomarkers that reliably track TS symptoms, i.e. tic-related neural states (Cagle et al., 2022; Marceglia et al., 2017; Okun et al., 2024)–for which pilot studies and implementations are already underway (Molina et al., 2018).

6.1.3 Neural Correlates of Tourette’s Tics in the CBGTC Network: Oscillatory Power and Connectivity

Intracranial electrophysiological recordings have advanced our understanding of the neural substrates underlying tics in Tourette Syndrome (TS), revealing specific local oscillatory signatures within the CBGTC network. Chronic recordings from DBS leads and subdural cortical strips have shown elevated low-frequency (1–10 Hz) activity in the centromedian–parafascicular (CM-PF) thalamic nucleus during complex tics, particularly during tics that are prolonged and involve multiple body regions (Shute et al., 2016). Similarly, both pallidal and thalamic local field potentials (LFPs) show heightened theta- and beta-band power correlated with tic severity, with prolonged theta bursts in subcortical structures predicting worse presurgical motor tic ratings (Neumann et al., 2018). Moreover, low-frequency oscillations in the CM thalamus have been associated with tics but not

voluntary movements (Cagle et al., 2020), pointing to their relevance in pathological processes specifically.

While these findings highlight local power alterations, a growing body of research underscores the importance of network-level synchrony and functional connectivity in characterizing pathological neural states (Voytek & Knight, 2015). In fact, the effectiveness of DBS in TS patients has been directly linked to imaging-based functional connectivity profiles in specific brain networks (Baldermann et al., 2024; Johnson et al., 2020). This perspective mirrors broader trends in neuroscience-based investigations of pathophysiology, where diseases such as depression, Parkinson's disease, obsessive-compulsive disorder, and Alzheimer's disease are increasingly understood as network phenomena, rather than disorders affecting focal brain regions (Sullivan et al., 2021). For example, Fox and colleagues demonstrated that therapeutic efficacy of brain stimulation—both invasive and noninvasive—can be predicted by the functional connectivity of targeted nodes within the same brain network (Fox et al., 2014). They proposed that a lack of connectivity indicates sites where stimulation is ineffective, and that the sign of connectivity determines whether excitatory or inhibitory stimulation is optimal. Such insights have driven a paradigm shift toward optimising therapy by targeting specific brain networks rather than isolated regions, with applications ranging from essential tremor (Dallapiazza et al., 2019) to minimally conscious states (Yamamoto et al., 2010).

In TS, this network-centric perspective has gained prominence through recent studies that shift the focus from localised power alterations to emphasise the critical role of network-level synchrony and functional connectivity in defining pathological neural states. Early electroencephalography (EEG) studies showed that voluntary tic suppression is accompanied by elevated alpha-band (8–12 Hz) coherence between frontal and sensorimotor areas, suggesting a compensatory role for frontal executive networks in providing top-down inhibitory control (Serrien, 2004). In (Bour et al., 2015), authors demonstrated that thalamo-cortical coherent activity is temporally locked to spontaneous tics but not to mimicked tics, highlighting the role of network-level pathology in tic generation. Because this modulation of coherent activity across regions emerges just before tic onset, it may provide an early electrophysiological signal suitable for anticipatory, closed-loop interventions.

More recently, Wehmeyer et al. (2024) coupled implanted thalamic LFP recordings with high-density EEG to identify a distinct thalamo-frontal alpha network whose connectivity strength inversely correlated with TS symptom severity. This alpha connectivity decreased prior to tic onset, suggesting network-level inhibitory deficits in tic generation. These effects were specific to connectivity, rather

than local power, emphasizing the importance of network-based biomarkers for real-time detection and intervention. Large-sample EEG studies in children support this viewpoint, showing that tic suppression enhances alpha connectivity in distributed cortical networks (Morand-Beaulieu et al., 2023). These convergent findings from multiple modalities and patient cohorts highlight alpha-band connectivity as an important factor in modulating both tic suppression and playing a role in the underlying neural processes generating tics.

6.1.4 The Aims of this Study

Building on the dataset and findings reported in Wehmeyer et al. (2024)—which demonstrated a temporally specific modulation of thalamo-frontal connectivity before tic onset—this study aims to determine whether these connectivity markers can be used to generate an electrophysiology-based signal to detect incoming tics in real time. To this end, we derive time-resolved, real-time-compatible measures of thalamo-frontal connectivity in the time-period preceding tic generation, and train subject-specific machine learning models to identify these pre-tic changes in neural activity. We then assess the predictive accuracy of these classifiers and compare the performance yielded by connectivity-based features when compared to conventional measures such as local oscillatory power.

We ultimately seek to evaluate the feasibility of translating the finding findings from (Wehmeyer et al., 2024) into adaptive (closed-loop) DBS paradigms for TS. Demonstrating that fronto-thalamic connectivity modulation reliably precedes tics on a clinically relevant timescale would be a significant step toward real-time neuromodulation approaches capable of targeting and mitigating tic-related activity. Through this work, we aim to refine mechanistic insights into tic generation and contribute towards the development of personalised, feedback-driven therapeutic approaches for DBS in TS.

6.2 Methods

6.2.1 Patient Cohort and Data Acquisition

Eight individuals with Tourette Syndrome (TS), who had undergone bilateral deep brain stimulation (DBS) implantation for severe tics at the University Hospital of Cologne between 2009 and 2022,

participated in this study. All subjects received the Medtronic Percept™ PC implantable pulse generator (IPG), which features data streaming capabilities, either at the time of their initial DBS lead implantation or upon later battery replacement. The surgical targets were the centromedian nucleus–ventrooralis internus (CM-Voi) or the ventral anterior/ventral lateral (VA/VL) nuclei of the thalamus, and the leads were quadripolar or directional DBS leads. Further details on the surgical procedures, patient demographics, and lead targeting strategies can be found in (Wehmeyer et al., 2024). Written informed consent was obtained from all participants prior to data collection, in accordance with the Declaration of Helsinki, and the study was approved by the local Ethics Committee of the Medical Faculty of the University of Cologne (No. 21–1351).

During each experimental session, patients were seated comfortably in a quiet room. Several data streams were recorded simultaneously:

- **Electroencephalography (EEG):** Sixty-three Ag/AgCl electrodes were arranged according to the extended 10–20 system, referenced to Cz and grounded at FPz. An ActiChamp (Brain Products GmbH, Germany) amplifier sampled EEG signals at 5 kHz, and electrode impedances were maintained below 10 k Ω .
- **Electrooculography (EOG):** Horizontal and vertical EOG channels were used to monitor eye movements.
- **Accelerometry:** Lightweight sensors were attached to body regions prone to tics to capture movement data.
- **Video and Audio:** Synchronised video and audio recordings were obtained to enable offline identification of motor and vocal tics. A trained clinical rater inspected these recordings to label each tic’s onset and offset.

Local field potentials (LFPs) were acquired from the implanted Percept™ PC IPG using BrainSense™ Streaming mode. Bilateral LFP recordings were configured at a sampling rate of 250 Hz by selecting bipolar contacts (0 and 3) for sensing, while middle contacts (1 and 2) served as possible stimulation sites; stimulation itself was switched off at least two minutes before data collection. The raw LFP signals were exported in JavaScript Object Notation (JSON) format and processed offline with an open-source toolbox (Thenaisie et al., 2021) to correct for possible missing data points. A fourth-order Butterworth filter was then applied to band-limit the signals between 1 and 100 Hz.

In parallel, the EEG data were down-sampled from 5 kHz to 250 Hz to match the LFP sampling rate and filtered between 1 and 100 Hz (fourth-order Butterworth IIR filter). To precisely align EEG and LFP signals, short bursts of subthreshold stimulation were delivered at the beginning and end of each recording session, producing visible artefacts in both modalities. These artefacts served as synchronization markers, achieving an average timing mismatch of 7.29 ± 6.19 ms (mean $\pm$ std. deviation; range: 0–32 ms).

6.2.2 Tic Labelling and Data Preprocessing

Motor and vocal tics were identified by an experienced clinical rater who reviewed the time-synchronised video and audio recordings. EOG traces and accelerometry data were consulted to confirm these markers in cases of subtle or ambiguous movements. The onset of each distinct tic was labelled in the EEG/LFP time series, and data epochs surrounding each tic were extracted for subsequent analyses. To avoid overlapping events, tics separated by fewer than two seconds were treated as a single tic sequence, with the earliest onset marking the start of the sequence. Any epoch in which a voluntary movement appeared in close proximity to the event of interest was excluded. Final analyses therefore relied on well-defined epochs for each individual tic across all patients.

This dataset—consisting of synchronised EEG, LFP, EOG, accelerometry, and video/audio streams—was used to explore neural oscillatory features predictive of tic onset, focusing primarily on putative connectivity markers in thalamo-frontal circuits. Subsequent signal processing steps related to connectivity-based feature extraction and classification procedures are described in the sections below.

6.2.3 Time-Resolved Computation of Connectivity Features

To assess thalamo-frontal connectivity during the pre-tic period, we extended the signal processing and feature extraction pipeline introduced in Chapter 4.2.3 to incorporate time-resolved measures of Phase Synchrony Index (PSI), Phase Locking Value (PLV), and Imaginary Coherence (iCoh). These three metrics, widely validated in previous research, were chosen for their ability to capture dynamic oscillatory interactions within the cortico-basal ganglia-thalamo-cortical (CBGTC) network. By analyzing how connectivity evolves over time, we aimed to detect transient changes in network synchrony that might serve as biomarkers for anticipatory tic detection.

6.2.3.1 Phase Synchrony Index

The **Phase Synchrony Index** (PSI) quantifies the consistency of phase differences between two oscillatory signals over time. It provides a robust measure of phase-based connectivity that is independent of signal amplitude, making it particularly suitable for analysing neural interactions. For two signals $X(t)$ and $Y(t)$, with instantaneous phases $\phi_X(t)$ and $\phi_Y(t)$, PSI is defined as:

$$\text{PSI} = 1 - \sqrt{\frac{\langle \sin^2(\Delta\phi(t)) \rangle}{\langle \cos^2(\Delta\phi(t)) \rangle + \langle \sin^2(\Delta\phi(t)) \rangle}}$$

where $\Delta\phi(t) = \phi_X(t) - \phi_Y(t)$ is the phase difference, and $\langle \cdot \rangle$ denotes the temporal average over a sliding window. PSI ranges from 0 (no phase synchrony) to 1 (perfect phase synchrony).

PSI-derived measures have been applied in various contexts to associate oscillatory connectivity with pathological states. For instance, (Kawano et al., 2021) demonstrated the utility of PSI in evaluating post-stroke motor impairments, while (F. Wang et al., 2019) used it to improve quantitative assessments of consciousness levels in patients recovering from ischemic stroke. Additionally, (Dauwels et al., 2010) highlighted its potential for early detection of dementia-associated cognitive impairment.

6.2.3.2 Phase Locking Value

The **Phase Locking Value** (PLV) is a widely used metric that quantifies the stability of phase relationships between two signals. Unlike the PSI, the PLV measures the consistency of phase differences without differentiating between specific phase offsets. For two signals $X(t)$ and $Y(t)$, the PLV is computed as:

$$\text{PLV}_{X,Y}(t) = \left| \frac{1}{T} \sum_{t=1}^T e^{-i\Delta\phi(t)} \right|$$

where $\Delta\phi(t) = \phi_X(t) - \phi_Y(t)$, and T represents the number of time points in the analysis window. PLV ranges from 0 (no phase locking) to 1 (perfect phase locking).

PLV has been demonstrated to be a robust marker of neural connectivity in a range of applications. For example, it has been shown to predict cognitive decline in Alzheimer's dementia (Ruiz-Gómez et al., 2019), forecast epileptic seizures (Myers et al., 2016), and identify imagined movements for brain-computer interface applications (Wei et al., 2007).

6.2.3.3 Imaginary Coherence

Imaginary Coherence (iCoh) isolates the coupling between two signals that occurs with a nonzero phase lag and is designed to reduce spurious connectivity caused by volume conduction or common reference effects. For two signals $X(t)$ and $Y(t)$, iCoh is defined as:

$$\text{iCoh}_{X,Y}(f) = \frac{\text{Im}(S_{XY}(f))}{\sqrt{S_{XX}(f)S_{YY}(f)}}$$

where $S_{XY}(f)$ is the cross-spectral density between $X(t)$ and $Y(t)$, and $\text{Im}(\cdot)$ denotes the imaginary part. $S_{XX}(f)$ and $S_{YY}(f)$ are the auto-spectral densities of $X(t)$ and $Y(t)$, respectively. iCoh ranges from 0 (no coupling) to 1 (maximum nonzero phase-lagged coherence).

Imaginary coherence has been applied to diverse clinical contexts, highlighting its relevance in identifying network-level disruptions. For example, Hinkley et al. (2010) used iCoh to assess cognitive impairments in schizophrenia, while (Hohlefeld et al., 2014) demonstrated its utility in detecting cross-hemispheric coupling between subthalamic nuclei in Parkinson’s disease. Additionally, (Black et al., 2018) identified altered task-related connectivity patterns in patients with autism spectrum disorders.

All connectivity metrics were computed in a time-resolved manner using a causal sliding window approach with a 1000 ms window length and 20 ms step size. This allowed us to capture dynamic changes in connectivity during the pre-tic period while providing high temporal resolution for identifying transient neural signatures. Thalamo-frontal channel pairs were analyzed (L-Th–Fz and R-Th–Fz), and data from both hemispheres were concatenated to drive the machine learning models.

6.2.4 Machine Learning Pipeline

To identify neural markers predictive of the pre-tic period, we trained logistic regression classifiers to distinguish 500-ms pre-tic intervals from baseline activity occurring earlier in the trial. Because substantial movement artefacts occur following tic onset (see Section 6.3.1), these analyses focused exclusively on pre-tic dynamics. We selected logistic regression for its interpretability and its capacity to generate a continuous prediction score over time: during training, the model learns a linear mapping from each input feature to a weight, indicating that feature’s relative contribution to distinguishing pre-tic from baseline activity. Applying this linear mapping across all time points yields a dynamic estimate of how closely the observed neural signals resemble the pre-tic state.

Classifier performance was evaluated using a leave-one-trial-out (LOTO) cross-validation (CV) procedure in which each trial was iteratively omitted from training and reserved for validation. This approach ensures that every trial serves as an unseen test sample exactly once, providing a robust assessment of generalization. We used the area under the receiver operating characteristic curve (AUC) as our primary performance metric. To address class imbalance between pre-tic and baseline windows, the logistic regression model was weighted according to the proportion of samples in each class.

To establish a baseline for statistical significance, we performed permutation testing by circularly shifting each trial’s time dimension by a random amount. This process preserved the distribution of features but disrupted meaningful temporal relationships. The logistic regression model was trained and tested on these shuffled data using the same LOTO-CV approach, thereby generating a null distribution of AUC values across 1,000 permutations. We considered the true AUC significant if it exceeded the 95th percentile of the null distribution in a one-tailed test (95% confidence level).

6.2.5 Assessing Cross-Trial Variability

To measure how closely each trial’s features align with the trial-averaged pattern, we conducted a bootstrapped, permutation-based test to quantify how consistent across trials the progression of features is. Specifically, for each feature x and trial i , we computed a single-trial time series $x_i(t)$ and a cross-trial average $\bar{x}(t)$. We then calculated the correlation coefficient r_i between $x_i(t)$ and $\bar{x}(t)$ from which we derived r_i^2 (the proportion of variance in the single-trial trajectory that is explained by the cross-trial mean). The average amount of explained variance across all trials was then compared against a permutation-based null distribution obtained by circularly shifting each trial along the time axis 1000 times.

6.3 Results

6.3.1 Post-Tic Artefacts and Rationale for Focusing on the Pre-Tic Window

Simultaneous electroencephalographic (EEG) and local field potential (LFP) recordings revealed substantial motion artefacts that coincided with tic execution. To quantify the extent of artefact contamination, we calculated a running root-mean-square (RMS) value (250-ms window) for raw EEG and LFP signals (Figure 6.1). Compared with a baseline interval (-2.5 to -0.5 s relative to tic onset), signal power in the first second following tic onset increased markedly across multiple channels. Table 6-1 summarises these artefact-related power changes for the frontal EEG (Fz) channel and bilateral thalamic LFPs channels (left thalamus [L-Th] and right thalamus [R-Th]). While the severity of artefact contamination varied across patients, increases in RMS power were frequently several hundred percent above baseline in subcortical LFPs, underscoring the pronounced impact of tic-related movements on the signals.

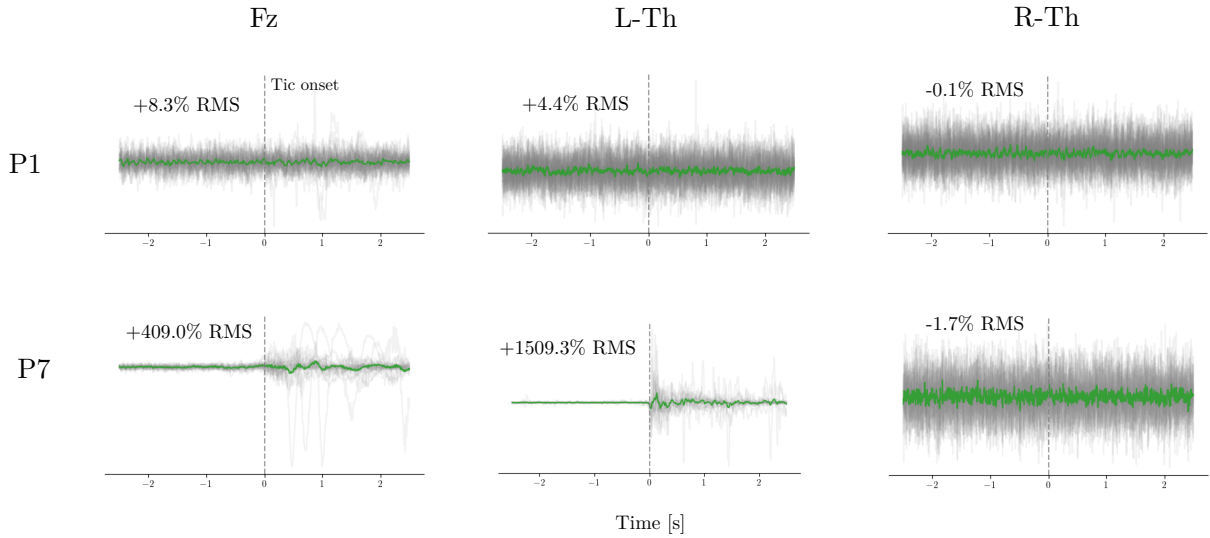


Figure 6.1: Signal Power Modulation due to Movement-Related Artefacts During Tic Execution. Representative EEG (Fz) and LFP (left and right thalami) traces for two patients, showing raw signals (with corresponding percentage changes) in root mean square (RMS) values during tic execution compared to baseline.

Because motion artefacts significantly contaminate electrophysiological signals during tic execution, subsequent analyses focused on the pre-tic interval (-500 ms to tic onset). This shift ensures that the identified features are not confounded by large-amplitude artefacts and aligns with the translational goal of identifying predictive markers for real-time adaptive neuromodulation. By concentrating on the

time window preceding tic onset, we sought to capture network-level changes that might act as early indicators of impending tics.

Table 6-1: Root Mean Square (RMS) Signal Power During Tic Execution. Percentage increase in RMS signal power during the first second post-tic onset compared to baseline (-2.5 to -0.5 s) for EEG (Fz), left thalamus (L-Th), and right thalamus (R-Th). Values represent mean (standard deviation) across all recorded tics for each patient. The table highlights wide inter-patient variability and pronounced artefact contamination, especially in LFP recordings.

Patient ID	Fz	L-Th	R-Th
P1	8.3 (24.2)	4.4 (9.9)	-0.1 (9.3)
P2	10.6 (18.9)	46.1 (197.6)	89.0 (40.7)
P3	21.6 (16.4)	6.5 (10.9)	1.1 (31.0)
P4	-2.1 (10.7)	500.8 (288.8)	23.2 (65.2)
P5	-5.0 (37.1)	745.9 (548.1)	45.6 (20.1)
P6	-2.6 (12.4)	148.7 (62.2)	63.4 (67.7)
P7	409 (63.9)	1509.3 (850.8)	-1.7 (3.4)
P8	41.6 (15.5)	624.3 (487.1)	108.1 (98.7)

6.3.2 Connectivity-Based Classification of Pre-Tic Intervals

To determine whether thalamo-frontal connectivity could reliably identify the pre-tic period, we trained logistic regression classifiers to differentiate 500-ms intervals preceding tic onset (pre-tic) from earlier baseline segments (-2.5 to -0.75 s). Three connectivity metrics—Phase Synchrony Index (PSI), Phase Locking Value (PLV), and Imaginary Coherence (iCoh)—were computed in a time-resolved manner between the frontal EEG channel (Fz) and each of the bilateral LFP recordings (L-Th, R-Th). We evaluated these connectivity measures at a 250-Hz sampling rate, using a 1000-ms sliding window that advanced in 20-ms step increments. Trial-averaged plots of the time-resolved connectivity features for individual patients are provided in Supplementary Figures V–XII.

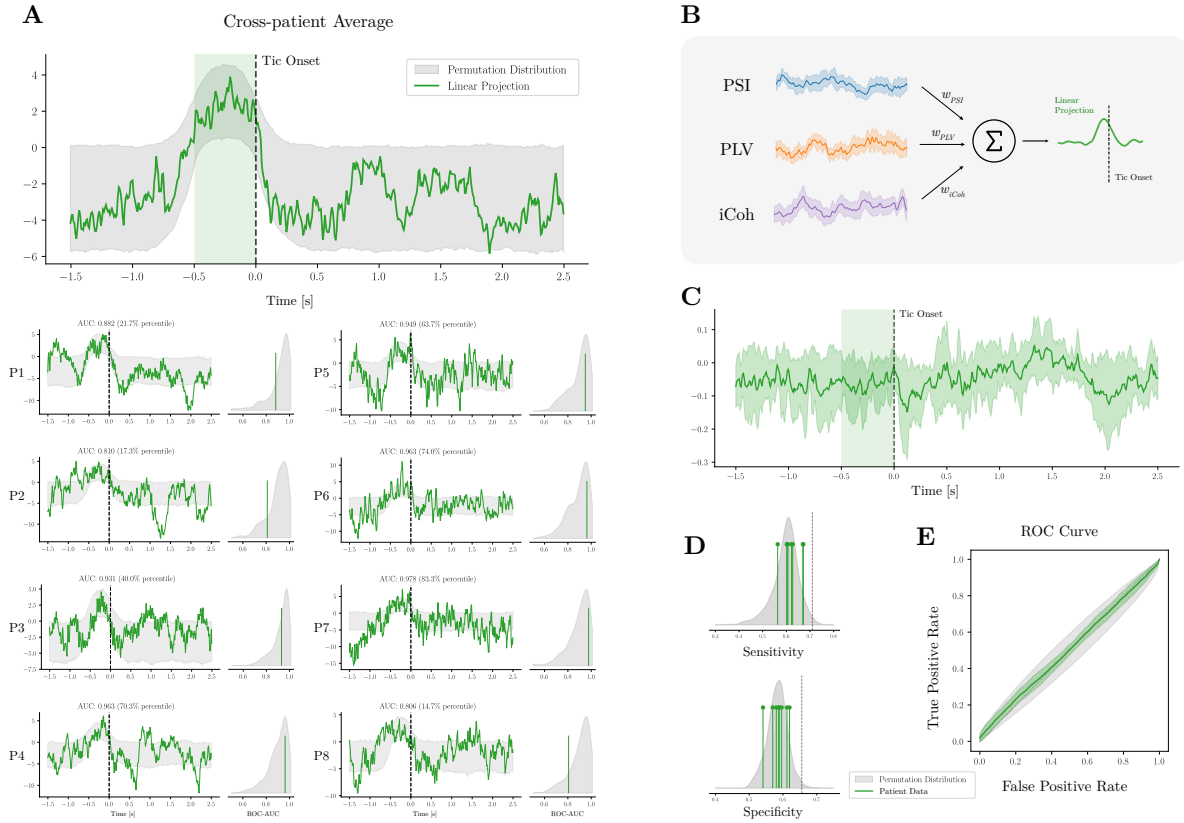


Figure 6.2: Connectivity-Based Classification of Pre-Tic Windows!Unexpected End of Formula. (A) Time-resolved linear mapping derived from a classifier trained to distinguish the 500-ms pre-tic interval from a rest baseline. The top panel shows the cross-patient average projection, while the bottom panels display individual patient trajectories. In each individual patient plot, a kernel density estimate of permutation AUC values is shown in grey, with a vertical green line marking the actual AUC achieved on the unshuffled data. Although some trial-averaged outputs exhibit high nominal AUCs, none exceed the permutation-based threshold. (B) Schematic of the logistic regression model that combines PSI, PLV, and iCoh features to classify pre-tic vs. baseline segments. (C)

Cross-patient out-of-sample projection obtained under a leave-one-trial-out cross-validation (LOTO-CV) approach. The green line represents the average linear projection across patients, and the shaded region indicates cross-patient variability. Notably, there is no pronounced elevation in this projection during the pre-tic window (−500 to 0 ms), suggesting that the learned mapping does not provide predictive power for tic onset in unseen data. (D) Sensitivity and specificity values for the LOTO-CV analysis, compared against the permutation-based null distribution (grey). The vertical dashed grey line indicates the 95th percentile of the cross-patient permutation results, and individual patient scores are shown by the green vertical lines. The individual classifier’s sensitivity and specificity remain below chance level. (E) Cross-patient ROC curve (green line) for the LOTO-CV analysis, with the green shaded region denoting cross-patient variance. The grey shaded region represents the permutation-based chance distribution. The overlap between the true classifier performance and the permutation envelope demonstrates that no significant above-chance classification was achieved.

Figure 6.2 summarises our classification results. In Figure 6.2A, a logistic regression model was trained on trial-averaged data to distinguish the pre-tic window from a baseline interval. The top panel shows the cross-patient average linear projection across time, while the bottom panels display individual patient trajectories. A kernel density estimate (grey) indicates the permutation-based AUC distribution for each patient, and a vertical green line marks the actual AUC achieved on the unshuffled data. Although nominal AUC values in some cases were high (0.81–0.98), none surpassed the 95th percentile of the permutation distribution after randomising each trial’s temporal structure. Thus, despite the classifier retrospectively fitting the trial-averaged data, no robust connectivity features could be confirmed once temporal dependencies were disrupted.

To gauge out-of-sample performance, we adopted a more rigorous leave-one-trial-out (LOTO) cross-validation approach. Figure 6.2C depicts the cross-patient average of the linear projection learned in these LOTO analyses; the green line shows how the model output evolves over time, with the shaded region illustrating variability across subjects. Notably, there is no clear modulation—or increase—in this projection during the 500-ms window before tic onset, suggesting the mapping learned on unseen data does not meaningfully track pre-tic state changes. In Figure 6.2D and E, we evaluate the classifier’s sensitivity, specificity, and overall AUC relative to permutation-based chance levels. Specifically, in Figure 6.2D, vertical dashed grey lines represent the 95th percentile of the permutation distribution for sensitivity and specificity, and green vertical lines mark individual patient scores. All remain below the threshold for significance. In Figure 6.2E, the cross-patient receiver operating characteristic (ROC) curve (green line) sits within the permutation envelope (grey), again indicating no above-chance discrimination.

Collectively, these results show that while connectivity-based features (PSI, PLV, iCoh) can retrospectively distinguish pre-tic intervals when averaged across trials, the effects do not generalise under LOTO cross-validation or permutation testing. Neither single-subject models nor cross-patient analyses yielded a consistent pre-tic signature strong enough to outperform chance once temporal structure was randomised or samples were withheld in testing.

6.3.3 Additional Features and Frequency Bands

Given that earlier work (Wehmeyer et al., 2024) documented an average fronto-thalamic connectivity modulation before tic onset, we investigated whether extended feature sets—harnessing features beyond connectivity alone—could improve detection of the pre-tic interval. In particular, we considered multiple frequency ranges (delta/theta: 2–8 Hz, alpha: 8–12 Hz, beta: 12–30 Hz, gamma: 30–100 Hz) and combined them with power spectral and time-domain features described in Chapter 4.2.3. For each frequency band, we formed two types of feature sets:

1. **Conn:** Models using only connectivity-based features (PSI, PLV, and iCoh), computed between frontal (Fz) and each thalamic channel (L-Th, R-Th).
2. **Conn+:** Models extending the connectivity feature set with additional spectral and time-domain metrics derived from the same frequency band.

These augmented sets were intended to capture a broader range of neural signatures potentially relevant to tic generation. As illustrated in Figure 6.3, individual patient AUC values are plotted for each feature set-frequency band combination, with grey dots indicating patient-specific classification results under leave-one-trial-out (LOTO) cross-validation. The green error bars represent 95% confidence intervals around the cross-patient mean AUC, none of which exceeded 0.5—confirming that no feature combination surpassed the 95th percentile of permutation-based chance thresholds. Supplementary Table I provides a detailed breakdown of each patient’s AUC scores and permutation distributions under these different feature sets.

An example of how these results map onto earlier findings is seen in the *Alpha Conn* model: this condition represents connectivity-only features extracted in the alpha band (8–12 Hz), effectively being identical to the frequency-specific results shown in Figure 6.2 (Panels C, D, and E). There, the logistic regression classifier focused on alpha-band connectivity between Fz and bilateral thalamic LFP channels. While earlier studies highlighted alpha-band connectivity changes before tic onset

(Wehmeyer et al., 2024), our present analysis showed that, under LOTO cross-validation and permutation testing, even alpha-only connectivity features fell short of above-chance classification for single-trial tic detection.

Overall, despite the broader range of features, none of the expanded sets improved pre-tic detection to a statistically significant level. AUC values spanned 0.32 to 0.71 (mean: 0.50, SD: 0.07), indicating that both connectivity-only (Conn) and connectivity-plus-power/time-domain (Conn+) approaches did not confer a decisive advantage. These results highlight that while certain frequency bands—particularly alpha—may show group-level modulations, they do not generalise to consistent, above-chance classification performance at the single-trial level under the current data and modelling conditions.

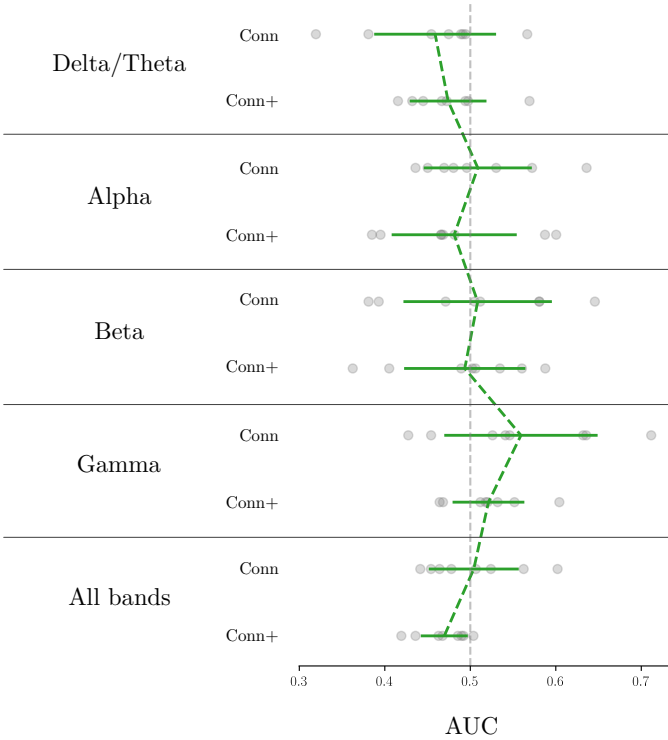


Figure 6.3: Effect of Additional Frequency Bands and Feature Sets on Classification Performance. Each point shows an individual patient’s area under the receiver operating characteristic curve (AUC) for the specified feature set, with grey dots indicating patient-specific classification results. Conn refers to models using connectivity features (PSI, PLV, iCoh) only, while Conn+ incorporates connectivity, time-domain, and spectral power metrics. Feature extraction was performed in one of four frequency bands—delta/theta (2–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), or gamma (30–100 Hz)—as indicated on the x-axis. Green error bars represent the 95% confidence interval of the cross-patient mean AUC; none exceeds 0.5, indicating no feature set yields above-chance

performance. All reported individual AUC values were evaluated against a permutation-based chance distribution, and none surpassed the significance threshold.

6.3.4 Time-shifted Analysis Across the Tic Epoch

Although the previous sections focused on the specific pre-tic window (−500 ms to tic onset), there was no guarantee that this interval alone would capture all potentially discriminative neural signatures. To address this, we performed a time-shifted analysis in which a 500-ms classification window was systematically moved in 20-ms increments across the entire tic epoch. This approach ensures that any segment of the tic epoch with potentially predictive features—whether immediately before tics or elsewhere in the recording—could be evaluated.

As shown in Figure 6.4, we computed the area under the ROC curve (AUC) at each time shift for both the cross-patient average (top panel) and individual patients (bottom panels). While certain time intervals and certain patients exhibit AUCs higher (or lower) than 0.5, none of these fluctuations consistently exceed the permutation-based threshold at the group level. In other words, although some individual participants show transient intervals where classification performance is nominally above chance, these intervals do not generalise across patients once averaged. Consequently, no 500-ms segment within the tic epoch emerges as reliably predictive of tic onset under our current data and methods.

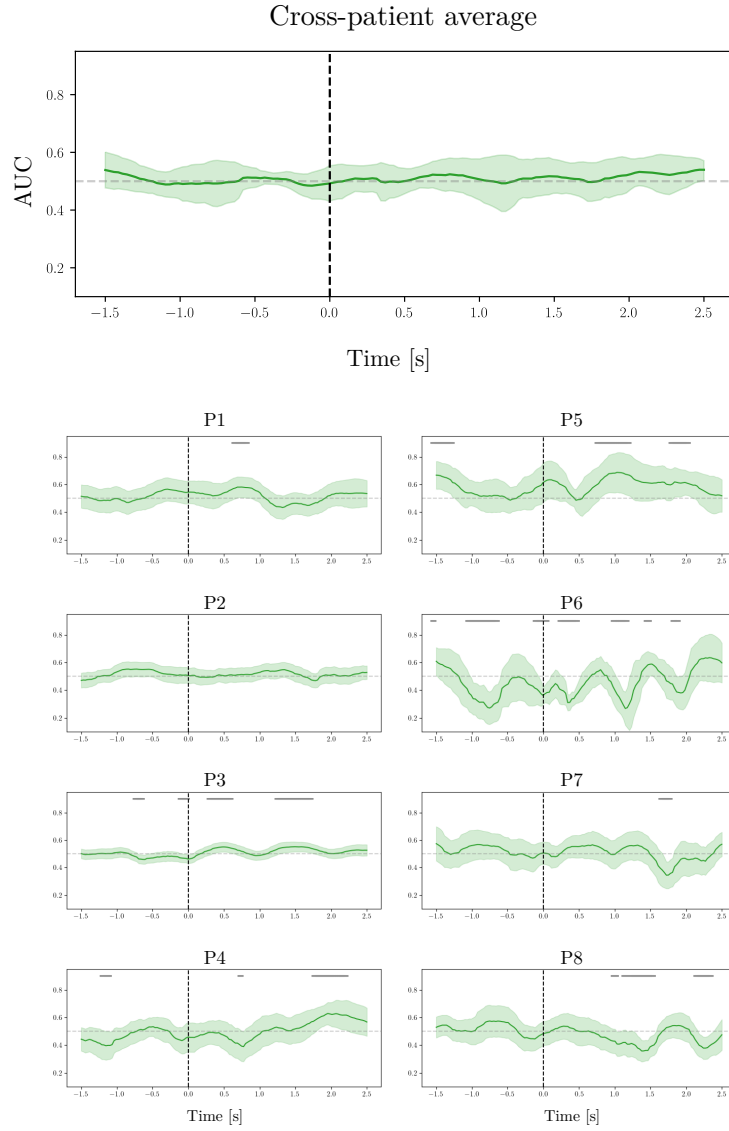


Figure 6.4: Time-Shifted Classification Analysis Across the Entire Tic Epoch. The top panel displays the cross-patient average AUC values (green line) for a sliding 500-ms classification window shifted in 20-ms increments. The shaded region in the top panel indicates the cross-patient 95% confidence interval. The bottom panels show individual patient results, each with its own 95% confidence interval computed across trials. Although certain patients exhibit intervals where the AUC is distinctly above or below 0.5, none surpass permutation-based chance thresholds on a consistent, cross-patient basis.

Collectively, these analyses indicate that machine learning classifiers trained on time-resolved connectivity measures did not outperform chance-level classification in detecting imminent tics, despite prior evidence of group-level connectivity changes in the pre-tic period (Wehmeyer et al., 2024). Expanding the feature space to include spectral power and time-domain metrics similarly failed to yield above-chance performance. These findings suggest that—under the current conditions—no robust, generalizable neural signature of tic onset was identifiable in the pre-tic window. The

implications of these negative findings, along with possible methodological considerations and possible future directions, are elaborated in the Discussion section below.

6.3.5 Cross-Trial Variability

Permutation-based tests revealed that 3 out of 48 (6.25%) connectivity-based features, and 13 out of 24 (54%) of power-based features, showed above-threshold cross-trial consistency. The variance in connectivity-based features explained by the cross-trial mean trajectory was $5.2\% \pm 4.1\%$ (mean $\pm$ std. dev.), and $15.2\% \pm 17.1\%$ for power-based features.

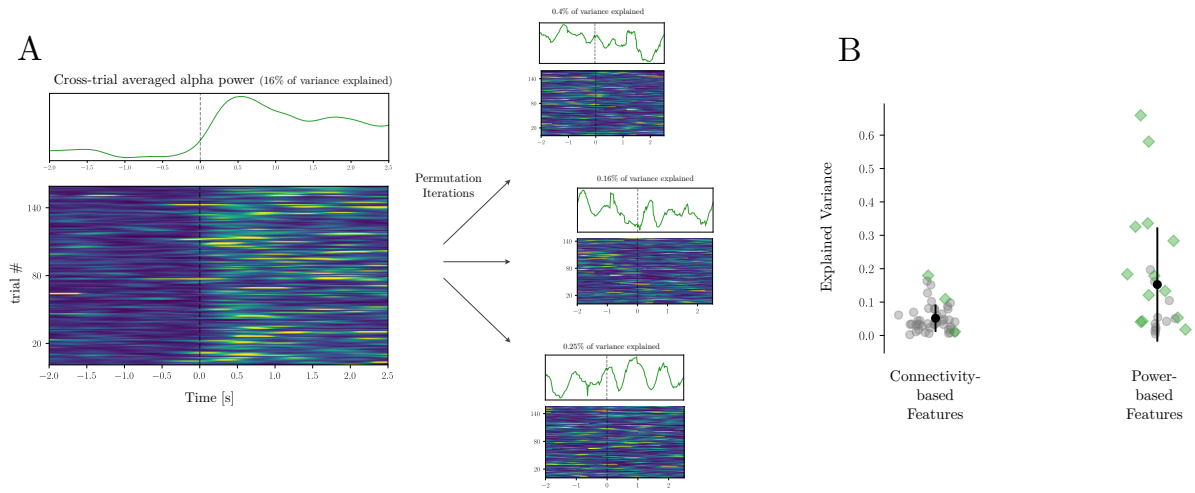


Figure 6.5: Cross-trial Consistency of Connectivity and Power-based Features. Panel A shows an example alpha-band power feature, comparing individual trial trajectories and the cross-trial average (green line, above). A permutation test was used to assess whether each trial's feature consistently tracked the mean trajectory more than expected by random chance (achieved by circularly shifting each trial's time axis). Panel B illustrates how much variance the cross-trial average feature trajectory explains for connectivity-based and power-based signals. Green diamonds represent features exceeding the permutation threshold, indicating greater trial-level consistency than chance. Overall, connectivity features (PSI, PLV, iCoh) exhibit minimal explained variance, often below 10%, while power features show higher but still modest consistency. This limited inter-trial consistency likely contributes to the difficulty in reliably classifying pre-tic intervals on a single-trial basis.

6.4 Discussion

6.4.1 Overview of Main Findings

This study investigated the feasibility of using oscillatory connectivity markers in fronto-thalamic circuits to predict the onset of Tourette Syndrome (TS) tics in real time. Building on prior work demonstrating average group-level modulation of phase synchrony before tic onset (Wehmeyer et al., 2024), we employed logistic regression classifiers and a rigorous leave-one-trial-out cross-validation (LOTO) scheme to assess whether connectivity markers, combined with other spectral or time-domain features, could discriminate the 500 ms pre-tic interval from an earlier rest period. Despite exploring diverse connectivity (PSI, PLV, iCoh) as well as further time domain and power-based metrics, our classifiers did not exceed permutation-based chance thresholds in any tested interval, including the immediate pre-tic window (−500 ms to 0 ms). These negative results persisted even when we shifted the classification window across the broader tic epoch, suggesting that no stable, generalizable neural signatures of imminent tic onset emerged under these conditions.

Our findings contrast with prior reports of alpha-band modulations before tic onset, which had suggested that pre-tic decreases in fronto-thalamic connectivity might reflect network-level deficits in inhibitory control (Wehmeyer et al., 2024), and we did not observe above-chance discrimination in out-of-sample classification. This discrepancy emphasises the challenges of translating group-level phenomena—where small yet consistent changes can be discerned—into a robust real-time biomarker applicable on a single-trial basis. In line with other negative findings in clinical electrophysiology, our results highlight the need for evaluating biomarkers under conditions that approximate real-world variability and real time-compatible classification demands.

6.4.2 The Potential Influence of Cross-Trial Variability

An important factor that likely contributed to the below chance classification performance achieved in this chapter is the high degree of cross-trial variability in both connectivity and power-based features. While average analyses (such as those reported in Wehmeyer et al., 2024) show group-level modulation of fronto-thalamic connectivity prior to tic onset, these consistent changes did not translate into robust single-trial markers in our study. In electrophysiological experiments, small, relatively stable

modulations can be detected when data are aggregated across trials. However, individual trials may deviate substantially from the mean pattern, distancing real-time classification designed to predict tics on a per-trial basis from observed average effects.

To explore how closely each trial's features align with the trial-averaged pattern, we conducted a bootstrapped, permutation-based test that quantified the variance each trial's feature trajectory shares with the overall mean trajectory. For connectivity metrics (e.g., PSI, PLV, iCoh), we found that the trial-average projection typically explained only a modest fraction of the variance in single-trial data. The small number of features that exceeded permutation thresholds suggests that only a limited subset of connectivity patterns is consistently reproduced on a per-trial basis. Power-based metrics showed somewhat stronger consistency, but a significant fraction of that variance may still reflect the presence of movement artefacts in the signals (see Section 6.3.1 above). Consequently, no single set of features was able to capture a stable, repeatable pre-tic signature across the dynamic neural conditions within which tics arise.

High inter-trial variability is not unique to TS; it is frequently reported in clinical and translational electrophysiological studies and is exacerbated by several confounding factors. Medication states (Tinkhauser, Pogosyan, Tan, et al., 2017), circadian fluctuations (van Rheede et al., 2022), and differing levels of voluntary tic suppression can alter underlying network activity, potentially obscuring the modulatory patterns that precede a tic. In a small dataset, these fluctuations can potentially dominate any latent signal of interest, preventing machine learning (ML) classifiers from generalising.

The causes of the variable signals across trials observed in this study remain unclear and cannot be investigated with the available data. They could relate to (1) variability/noise related to the measurement technique (e.g. electrodes are not placed to capture relevant oscillatory activity), or (2) to variability in the underlying physiological process. As mentioned in the Chapter 1.2.2, cortical signals from invasive ECoG recordings have in several studies shown superior signal-to-noise and concomitant superior decoding performance when compared to subcortical LFPs (Merk, Peterson, Lipski, et al., 2022; Opri et al., 2020). This has led to a growing interest in harnessing these cortical signals for the purposes of aDBS (J. Herron, Kullmann, et al., 2024).

Strategies to address the variability include collecting larger numbers of tics per participant, reducing confounds, and applying more advanced modelling approaches that adapt to session-specific, contextual information. New devices with enhanced sensing capabilities, although more constrained in

their signal acquisition and processing capabilities than equipment utilised for externalised recordings, have opened up new avenues in biomarker discovery (Gilron et al., 2021) and might help address some of the challenges encountered here. Ultimately, our results underscore that robust real-time detection of imminent tics may require tackling the inherent variability of human neural data as part of classifier design.

6.4.3 Sensitivity and Variability: Comparing Trial-Averaged Statistics with Machine Learning-Based Approaches

Our methodological framework underscores the distinction between trial-averaged analyses—ubiquitous in neuroscience for detecting small, consistent group-level effects—and machine learning (ML) pipelines that aim to identify real-time biomarkers in variable data. Traditional trial-averaged approaches (e.g., t-tests, ANOVA) excel at revealing subtle changes in central tendencies, as they aggregate data across repetitions and effectively “smooth out” noisy fluctuations. However, this smoothing process assumes homogeneity of neural signals across trials and can obscure the transient dynamics crucial for real-time applications such as adaptive deep brain stimulation (aDBS).

In contrast, ML-based methods—particularly those using cross-validation, such as leave-one-trial-out (LOTO)—provide an estimate of how well a classifier generalises to unseen data. This approach is vital in clinical settings where tics (or other pathological events) occur unpredictably. Yet to be successful, ML classifiers depend on sufficient consistency in neural signals. If connectivity or power features exhibit high inter-trial variability, the model struggles to identify a unifying pattern that generalises beyond the training samples.

Figure 6.6 conceptualises this trade-off between traditional statistics and ML approaches. When distribution variance is moderate, a t-test can achieve statistical significance ($p < 0.05$) by detecting small, uniform shifts across trials. Conversely, logistic regression requires relatively lower variance to achieve comparable performance because its generalization is hindered by large trial-to-trial fluctuations. This reduced power of machine learning-based testing has been reported on before and attributed, amongst other factors, to the inefficient use of data that cross-validated pipelines make when computing performance statistics (Rosenblatt et al., 2021). While trial-averaged methods are thus highly sensitive to small, consistent group-level effects, they may not capture the dynamic, context-dependent signals needed for real-time classification.

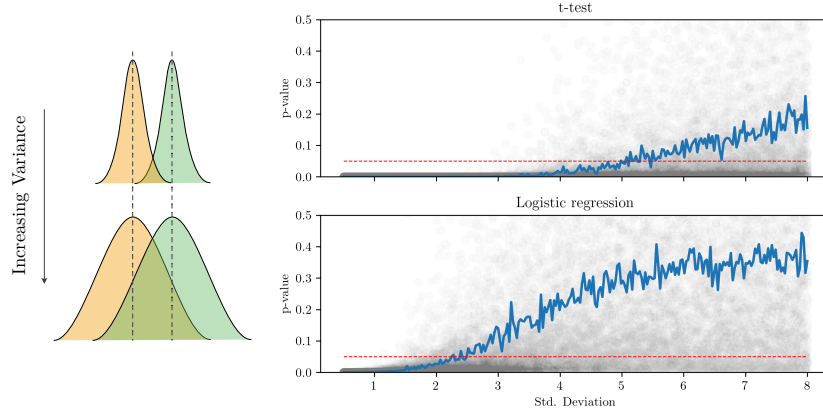


Figure 6.6: Sensitivity Comparison Between Trial-Averaged Statistical Tests and Cross-Validated Machine Learning Approaches. The t -test remains sensitive to subtle yet consistent effects, even under higher variances, because data aggregation preserves stable shifts in mean responses. By contrast, the logistic regression classifier demands lower variance to match this sensitivity, as it must generalise across individual trials rather than detect differences in a consolidated dataset. The figure exemplifies the complementary strengths and weaknesses of each method for detecting neural signatures.

Hypothesis-driven investigations that aim to identify stable group-level phenomena are well served by trial-averaged approaches, which excel at identifying small yet consistent effects. However, real-time interventions such as aDBS require robust single-trial detection mechanisms that must contend with inherent variability in the appearance of the neural correlates. Although ML frameworks can dynamically integrate a wide array of features, they are prone to overfitting in small, high-dimensional datasets. This highlights a need for larger data volumes, careful model regularization, and rigorous cross-validation schemes to validate the generalisability of models. ML-based approaches may also miss subtle but reproducible group-level modulations if those effects are overshadowed by trial-specific noise. In practice, combining elements of both methodologies may be optimal—harnessing trial-averaged insights to guide feature selection or interpret group-level patterns, while reserving ML-based pipelines for classification at the single trial level.

In summary, while trial-averaged methods provide a foundation for discovering subtle, reproducible electrophysiological effects in TS, ML-based analyses are better aligned with the goals of real-time tic detection and adaptive neuromodulation. Both strategies are constrained by the high inter-trial variability in TS data, yet they address this variability in different ways.

6.4.4 Limitations

The analysis conducted in this study is subject to a number of limitations that might relate to the negative findings of this study:

1. **Limited dataset size:** our dataset included eight individuals with Tourette Syndrome (TS) who underwent deep brain stimulation (DBS), the total number of usable tics per patient remained limited. This constraint stemmed from practical considerations, including the exclusion of epochs overlapping with other tics or voluntary movements and the logistics of conducting multiple recording sessions. Small sample sizes and heterogeneous patient populations can obscure underlying neural patterns, particularly when clinical phenotypes (e.g., tic severity, tic distribution) differ markedly across individuals. Expanding the cohort and increasing the number of recorded tics per subject—preferably over multiple sessions—could bolster the statistical power of future analyses, allowing more reliable conclusions about the consistency and predictive utility of electrophysiological markers. Longitudinal data recordings enabled by devices with enhanced sensing capabilities, as opposed to in-clinic recordings, have already contributed insights into pathophysiology and candidate biomarkers in other areas of aDBS research (Feldmann et al., 2022; Gilron et al., 2021; van Rheede et al., 2022). They might also contribute towards generating the richer TS datasets that might be needed to elucidate and validate robust electrophysiological markers of TS pathology.
2. **Artefactual contamination and residual noise:** despite focusing the analysis of electrophysiological signals on the pre-tic time interval, subtle motion artefacts or other confounding factors (e.g., changes in arousal or attention) may still influence the signals. Notably, the LFP channels proved particularly susceptible to large amplitude increases, and partial contamination might remain even in the pre-tic window.
3. **Real-Time feasibility and computational constraints:** Although our study adopted a time-resolved feature extraction and classification pipeline that was designed to be compatible with real-time application, we did not deploy it in a real-time, in-clinic neuromodulatory setting. Implementing these methods in clinical practice, and with a translational goal in mind, would likely require the use of specialised hardware capable of computing connectivity and spectral features with minimal latency. Additionally, the computational load of real-time classification must be balanced against battery and resource constraints in implanted devices. Currently

ongoing advances in neural signal processing hardware and optimised software algorithms may address these limitations and facilitate the practical realization of closed-loop DBS for TS.

6.4.5 Implications for Adaptive Neuromodulation

Although the present study did not yield a reliable real-time biomarker of tic onset, our negative findings nonetheless offer important insights for future adaptive DBS (aDBS) research in TS. First, they emphasise the need to systematically account for high inter-trial variability in neural signals. Strategies like model personalization—training separate classifiers for each patient or using ML systems that adapt to the patient state at the time at which adaptive stimulation adjustments are being made—may be more effective than attempting to deploy a single universal model. Functional modulation of neural activity, if it affects the signals that are used for aDBS feedback, has in fact already been shown to interfere with the effectiveness of the closed-loop approach (Busch et al., 2024). Second, alternative neural features (e.g., high-frequency oscillations, phase-amplitude coupling, or advanced connectivity measures with improved artefact rejection) warrant exploration. Third, the dynamic nature of tic pathophysiology may necessitate integrating contextual or behavioural data (e.g., states of urge intensity, voluntary suppression efforts, or comorbid ADHD/OCD fluctuations) to improve prediction accuracy.

Future studies, enabled by a new generation of DBS devices, could benefit from larger cohorts and longitudinal designs, capturing repeated recordings over weeks or months. Such longitudinal data could improve modelling of individual variability and clarify whether progressive adaptation of the aDBS controller is needed. With those richer data, additional computational techniques—like recurrent neural networks or hidden Markov models—may better capture transient, multi-step dynamics characteristic of tic onset. Finally, specialised hardware capable of real-time connectivity computation, combined with prospective validation in actual aDBS paradigms, would be required to advance the translation from theoretical connectivity-based biomarkers to clinical practice.

6.5 Conclusion

In this chapter, we investigated whether connectivity-based markers in fronto-thalamic circuits could predict imminent tic onset in Tourette Syndrome (TS). Building on earlier evidence that group-level

changes in alpha-band synchrony may precede tics, we systematically evaluated connectivity metrics (PSI, PLV, iCoh) and combined them with spectral power and time-domain features to train machine learning classifiers. Despite observing average modulations in trial-aggregated data, this group-level modulation did not generalise to reliable, single-trial detection of tics as evaluated by leave-one-trial-out cross-validation and permutation testing.

Multiple factors likely contributed to these negative findings. The available dataset, though comprising multiple DBS-implanted patients, remained modest in size, limiting the statistical power of classification models. Substantial inter-trial variability as well as motion artefacts in the recorded signals introduced additional noise that likely diminished the ability to capture consistent pre-tic patterns on a single-trial basis. Further work should investigate whether larger, longitudinal datasets have the potential to elucidate biomarkers that can enable tic detection at the individual trial level.

6.6 Summarised Findings

- **Pre-tic classification performance:** Neither connectivity metrics (PSI, PLV, iCoh) nor expanded feature sets (incorporating power and time-domain features) achieved above-chance accuracy in distinguishing the pre-tic interval from baseline, failing to surpass permutation-based chance levels.
- **Artefactual contamination:** Pronounced movement artefacts were detected during tic execution, justifying a focus on the pre-tic window (−500 ms to 0 ms) for analyses – as well as highlighting the challenges of real-world clinical recordings.
- **Cross-trial variability:** In spite of the presence of average, group level modulatory patterns, significant variability across trials in both connectivity and power features likely constrained the performance of machine learning classifiers, as no single consistent neural signature was detectable on a per-trial basis.
- **Trial-averaged vs. machine learning approaches:** Traditional trial-averaged analyses can reveal subtle group-level changes, but machine learning pipelines that aim for real-time compatible, single-trial predictions are hampered by variable signals and small datasets.

7 Project 5: Implementing and Debugging CNN-driven aDBS

7.1 Introduction

In Chapter 3 of this thesis, we explored the application of convolutional neural networks (CNNs) as a signal processing tool for extracting information from neural time-series signals. Specifically, we demonstrated that various CNN architectures could effectively detect relevant neural states from simulated waveforms and from local field potential (LFP) recordings in patients with essential tremor (ET) during movement and rest. In several cases, these CNNs outperformed traditional feature-based approaches, highlighting their potential for neural decoding tasks.

Building upon these findings, we sought to adapt the CNN-based pipeline for real-time clinical applications, particularly for implementing adaptive deep brain stimulation (aDBS) systems. The goal was to develop a closed-loop system capable of modulating stimulation based on the LFP signals from a patient in real time.

To achieve this, we developed a real-time pipeline consisting of four steps:

1. **Data Acquisition:** We recorded LFP signals from implanted DBS electrodes and corresponding movement sensor data while patients performed trials of prescribed motor tasks. This setup allowed us to collect synchronised neural and behavioural data necessary for training the decoding models.
2. **Data Preprocessing and Labelling:** We developed a custom software tool to preprocess the LFP signals and process the movement sensor data. This tool facilitated the generation of movement state labels by analysing the sensor data, with functionality to fine-tune the labelling parameters as needed.
3. **Model Training:** The preprocessed LFP signals and corresponding labels were sent to a high-performance computing server equipped with graphical processing units (GPUs) to expedite the training of neural networks. We trained multiple CNN models, selecting the one with the best validation performance for deployment in the real-time setting. This model

would be downloaded and prepared for inference on the CPU of the laptop available during the experiment.

4. **Real-Time Inference and Closed-Loop Control:** The trained CNN was used as part of a real-time system to decode movement states from streaming LFP signals. Incoming signals were streamed from the amplifier, buffered, and preprocessed using a pipeline compatible with real-time constraints. The model's output was then used to drive a DBS pulse generator through custom software, closing the feedback loop.

Despite successful implementation in controlled laboratory settings, including successful real-time decoding experiments conducted in a saline tank, we encountered challenges during clinical experimental sessions with patients. The CNN-based system did not achieve the expected performance levels observed during development.

Analysis of the experimental data revealed that changes in the stimulation state would induce significant alterations in the dynamic range of the recorded LFP signals. These alterations led to instability in the CNN outputs. This instability propagated through the closed-loop system, triggering further changes in stimulation state and preventing the system from stabilizing.

This phenomenon is consistent with reports in the literature highlighting the difficulties of recording and processing deep brain signals during simultaneous stimulation and signal acquisition (Ansó et al., 2022; L. H. Hammer et al., 2022; Lio et al., 2018). The presence of stimulation artefacts and changes in signal characteristics pose challenges for real-time neural decoding systems, particularly those relying on machine learning models sensitive to the distribution of inputs.

Upon returning to the laboratory, we conducted analyses to identify the root cause of these issues. We traced the instability back to numerical problems related to changes in the statistical distribution—specifically, the dynamic range—of the LFP signals when the stimulation state switched. The CNNs, being sensitive to the statistical properties of their input data, were adversely affected by these changes, leading to unreliable predictions. The closed-loop nature of the system, where the output of the CNN would in turn affect the signals that it would subsequently receive as input, perpetuated the instability.

To address these challenges, we explored and implemented signal processing solutions aimed at mitigating the effects of dynamic range changes on the CNNs' performance. These implemented

custom time-domain filtering techniques to remove stimulation-induced artefacts and stabilise the input signals, as well as exploring the use of frequency-domain representations to reduce sensitivity to amplitude fluctuations.

In this thesis chapter, we describe our experiences implementing CNN-driven aDBS systems in a clinical setting. First, we detail the procedure developed for efficiently training and deploying CNNs for inference in real time. Next, we examine how changes in stimulation parameters affect the CNNs' ability to process incoming signals and contrast this with the use feature-based methods in real-time applications. We then present an approach to alleviate the stability issues arising from changes in dynamic range, namely a custom time-domain filter to eliminate the stimulation artefact from the incoming signals. Finally, we discuss the implications of our findings, comparing the performance of CNNs with traditional feature-based methods and highlighting relevant considerations for implementing aDBS schemes based on neural networks.

7.2 Methods

7.2.1 Real-Time CNN-Driven Adaptive Deep Brain Stimulation Pipeline

To translate the results of our convolutional neural network (CNN)-based neural decoding into a real-time clinical application, we developed a pipeline for implementing CNN-driven adaptive deep brain stimulation (aDBS). This pipeline encompassed data acquisition during experimental sessions, data preprocessing and labelling, model training, and real-time inference to drive the DBS stimulator in a closed-loop fashion.

During the experimental sessions, patients with externalised deep brain stimulation leads were instructed to perform trials of upper-limb movements. While the patients executed these movements, we simultaneously recorded local field potential (LFP) signals from the externalised leads of the implanted electrodes, along with peripheral movement data captured by accelerometers and electromyography (EMG) sensors placed on relevant muscles. This simultaneous recording of neural and movement signals provided a synchronised datasets that could be used for training a neural decoder capable of mapping LFP patterns to movement states.

After completing the recording sessions, we transferred the collected time-series data to a custom-made software tool designed to facilitate efficient data preprocessing and labelling. This tool allowed us to select relevant channels from the recorded data, preprocess the LFP signals, and process the accelerometry and EMG data to extract movement labels. The processing of the peripheral sensor signals was parameterised, enabling adjustments through the user interface to optimise the labelling process. A screenshot of this software tool is provided in Supplementary Figure XVII.

Once the data were preprocessed and labelled, the time-series signals were securely transferred to a remote server using a secure file transfer protocol⁸. The server was equipped with two NVIDIA RTX 3090 graphics processing units (GPUs) (NVIDIA Corporation, USA), providing the necessary computational resources for training deep neural networks efficiently. Upon receiving the data, an automated workflow was triggered to train various CNN models, as described in Chapter 3.2.3, with a range of hyperparameter configurations. This training process involved exploring different network architectures, learning rates, regularization techniques, and optimization algorithms to identify the model that achieved the best performance on validation data. The training progress was monitored using a neural network training tracking tool, Weights & Biases (Biewald, 2020), which facilitated experiment management and hyperparameter optimization. The trained model with the best validation performance was then downloaded back to the local machine for deployment.

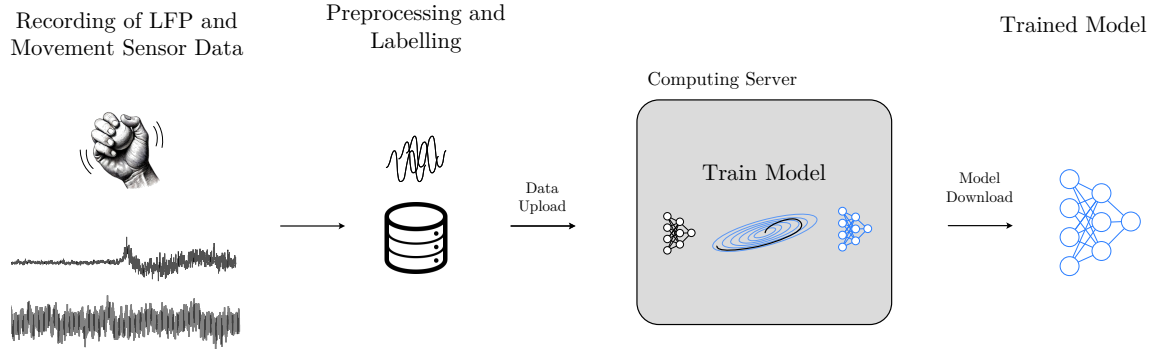
With the trained model available locally, we initiated a session in a separate custom-made program designed for real-time inference and closed-loop control. This program implemented the necessary functionality to load the trained CNN models for real-time inference, connect to an incoming LFP data stream, buffer and preprocess incoming chunks of data, format the data appropriately for input into the CNN, and parse the CNN output to guide the DBS stimulator. A screenshot of this software tool is displayed in Supplementary Figure XVIII. The real-time streaming of signals was implemented using the open-source networking ecosystem Lab Streaming Layer (LSL) (Kothe, 2014), which provided a robust framework for time-synchronised data acquisition and transmission. The integration of LSL ensured the low-latency data streaming necessary for the real-time operation of the closed-loop aDBS systems.

⁸ Secure Copy Protocol (scp): <https://man.openbsd.org/scp.1>

Incoming LFP signals were continuously streamed from the amplifier, buffered in real time, and preprocessed using a pipeline compatible with real-time constraints. Preprocessing steps included filtering to remove noise and artefacts, normalization to maintain consistent signal amplitude, and segmentation to prepare data windows suitable for CNN input. The preprocessed data were then formatted into the appropriate tensor shapes and data types required by the CNN model. The model processed these inputs to generate predictions about the patient's movement state, indicating whether the patient was at rest or performing a movement.

The CNN output was parsed and utilised to drive the DBS pulse generator. The program sent control signals to the stimulator through a data stream, adjusting the stimulation state. This closed-loop control allowed the system to modulate stimulation in response to the patient's neural state. During this study, we utilised two types of stimulators: a clinical-grade stimulator for use during patient sessions and a research-grade stimulator for laboratory experiments and simulations. The integration of these stimulators required implementation of custom communication protocols to enable compatibility with our real-time pipeline.

A Training



B Inference

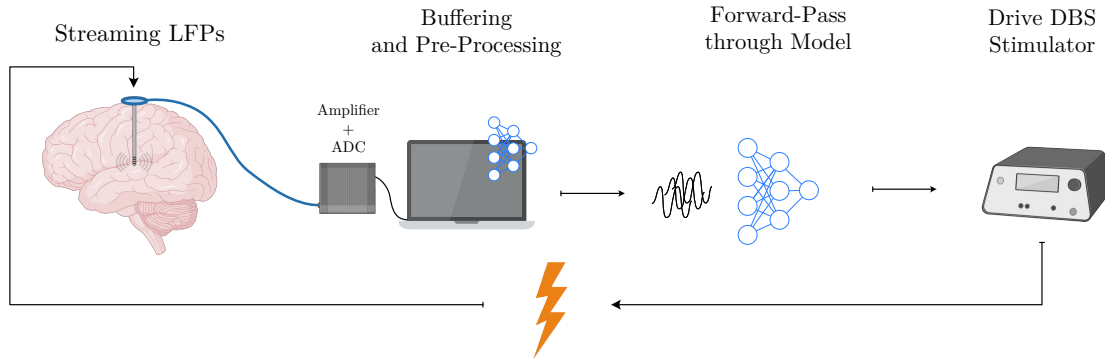


Figure 7.1: Real-Time CNN-driven aDBS Pipeline. (A) Training Phase. Local field potential (LFP) signals are recorded from an externalised patient performing trials of an upper-limb movement task. The raw LFP signals are preprocessed and movement labels are generated by processing data from peripheral sensors (accelerometers and electromyography). The processing of peripheral sensor data is parameterised, allowing adjustments to optimise label accuracy. The preprocessed LFP signals and the corresponding movement labels form a labelled dataset, which is transferred to a remote server for training. On the server, a convolutional neural network (CNN) model is trained to decode movement states from the LFP signals. The trained CNN model is returned for deployment in the real-time system. (B) Inference Phase. Incoming LFP signals are continuously streamed from the patient's implanted electrodes using a real-time data streaming framework. The signals are buffered and preprocessed in real time, applying the same preprocessing steps used during training to ensure compatibility with the CNN's input requirements. The preprocessed signals are formatted appropriately and fed into the trained CNN model, which processes the data to decode the patient's current movement state. The CNN's output is parsed and utilised to guide the deep brain stimulator.

7.2.2 Stabilising CNN Activations Across Changes in Stimulation State

The deployment of machine learning systems in adaptive deep brain stimulation (aDBS) systems presents unique challenges, particularly when operating across changes in stimulation state. One of the primary issues arises from stimulation artefacts that significantly alter the dynamic range of recorded local field potential (LFP) signals, leading to instabilities in feature space and subsequent decoding performance degradation. This challenge is particularly acute for deep learning systems that are sensitive to the distribution of their inputs, such as convolutional neural networks. Addressing this challenge is crucial for the reliable operation of aDBS systems that harness deep learning architectures for neural decoding.

Several strategies have been proposed in the literature to mitigate the effects of stimulation artefacts and stabilise neural decoder performance during changes in stimulation state. One common approach is *blanking*, which involves temporarily pausing the decoding pipeline—effectively halting the aDBS operation—for a predefined period after each change in stimulation state. This *blanking period* allows the LFP signals to settle and return to a baseline state that is consistent with the conditions under which the decoding pipeline was trained. In (He et al., 2021), this method, in combination with a two-model approach, was employed in a clinical setting, demonstrating its practicality for managing stimulation-induced signal perturbations.

However, blanking has inherent limitations, especially for decoding systems that rely on temporal information from extended periods into the past, such as those incorporating autoregressive models or networks with large receptive fields. In Chapter 3.4.3, we established that models that harness longer temporal contexts can enhance decoding performance by capturing relevant neural dynamics. Blanking interrupts the continuity of the data stream, forcing the system to re-buffer data after each stimulation state change, which undermines the benefits of models that depend on uninterrupted temporal information. Consequently, there is a need for approaches that enable continuous operation across stimulation state changes without compromising the integrity of the input signals or the stability of the neural decoder.

Another class of strategies focuses on alleviating or eliminating the stimulation artefact from the recorded LFP signals, thereby minimising its impact on the decoder's performance. Various methods have been developed for artefact removal, including signal processing techniques and algorithmic approaches designed to isolate and subtract the artefact component. However, many of these methods

are tailored for offline analysis and are not suitable for real-time applications (Allen et al., 2010; L. H. Hammer et al., 2022; Hashimoto et al., 2002; Lio et al., 2018; Qian et al., 2017; Y. Sun et al., 2014). As a result, they are not applicable to aDBS systems that require immediate responsiveness to neural signals.

(P. Chen et al., 2022) proposed a novel periodic artefact removal algorithm designed for deep brain stimulation applications. Their method models the stimulation artefact as a periodic signal and employs optimization techniques to estimate and subtract it from the recorded LFP data. The algorithm is robust to missing data and can operate when the stimulation frequency exceeds the Nyquist frequency. While this approach shows promise, it requires buffering chunks of data to estimate the artefact accurately. This buffering introduces latency that may be incompatible with the timing requirements of real-time aDBS systems, where minimising latency is critical for effective closed-loop control.

To overcome these limitations, we sought methods that can effectively suppress stimulation artefacts in real time with minimal latency. (Nie et al., 2021) and (J. M. Weiss et al., 2019) developed an artefact removal technique based on irregular sampling of signals. The fundamental premise of their algorithm is that the stimulation artefact has a predictable shape and occurs at consistent intervals corresponding to the stimulation pulses. By accurately detecting these "polluted intervals"—the time periods during which the artefact dominates the signal—it is possible to eliminate or interpolate over them, effectively isolating the underlying neural activity.

The method presented in (Nie et al., 2021) involves sampling the neural signals at irregular intervals to avoid the periods contaminated by the artefact. By reconstructing the signal from the non-contaminated samples, they can achieve artefact suppression without the need for computationally intensive algorithms. This approach is advantageous for real-time applications because it minimises processing delays and preserves the temporal continuity of the neural signals essential for decoding.

However, their system was originally designed for signals sampled at high frequencies (e.g., 44 kHz), which are not commonly available in implantable devices or research amplifiers used in deep brain stimulation studies. High sampling rates facilitate precise identification of the artefact intervals and accurate reconstruction of the neural signals. In practical settings, especially with commercially available hardware, we often operate at lower sampling frequencies (i.e. under 8k Hz). Lower sampling

rates pose challenges for the detection and elimination of stimulation artefacts due to reduced temporal resolution.

To address this issue, we adapted the method from (Nie et al., 2021) to be compatible with lower sampling frequencies. Our adaptation involves refining the detection of polluted intervals for environments with low sampling frequencies. By doing so, we aim to achieve effective artefact suppression in real time without incurring significant latency, thereby generating signals that don't stabilise CNN activations through changes in stimulation state.

In the following section, we outline the implementation of our adapted system, detailing the custom time-domain filtering techniques employed to minimise the effect of the stimulation artefact and discuss the algorithmic modifications made to accommodate signals sampled at lower frequencies.

7.2.2.1 Custom Time-domain Filtering of Signals to Minimise the Effect of the Stimulation Artefact

In this section, we introduce our custom time-domain filtering algorithm designed to minimise the effect of stimulation artefacts on the recorded local field potential (LFP) signals. This algorithm is based on the work presented by (Nie et al., 2021), but has been adapted to operate effectively at lower sampling frequencies (e.g., 4,096 Hz) commonly used in clinical settings and with research amplifiers.

Our algorithm is implemented in the Julia programming language (Bezanson et al., 2017) and is optimised for real-time applications by providing output on a sample-by-sample basis with minimal latency. It aims to detect and suppress portions of the signal contaminated by stimulation artefacts while preserving the intrinsic neural activity essential for accurate decoding by convolutional neural networks (CNNs) and other algorithms.

A. Algorithm Overview

The algorithm distinguishes between intrinsic neural signals and stimulation artefacts by dynamically computing a threshold based on running statistics of the incoming signal. When a sample exceeds this threshold, it is identified as part of a stimulation artefact, and the algorithm enters a *suppression period* (also referred to as "blinking mode") for a predetermined number of samples. During this period, the algorithm outputs predictions from an autoregressive (AR) model that is continuously

updated using previous signal samples. This approach effectively replaces the contaminated samples with estimates based on prior neural activity.

When not in suppression mode, incoming samples are considered part of the intrinsic neural signal and are returned without modification. The continuous updating of running statistics and the AR model ensures that the algorithm adapts to changes in signal characteristics over time.

At lower sampling frequencies, accurately detecting stimulation artefacts becomes challenging due to aliasing effects and reduced temporal resolution. The high-frequency components of the stimulation pulses can be aliased into lower frequencies, causing artefact samples to appear at varying amplitudes. Therefore, relying solely on the amplitude of individual samples is insufficient for reliable detection. Our algorithm addresses this by incorporating adaptive thresholds and running statistics that consider the broader context of the signal.

B. Filter Implementation

The pseudocode of the algorithm is presented below. The algorithm operates in two main phases: the filter warm-up phase and the real-time filtering phase. During the filter warm-up phase, the initial parameters that comprise the state of the filter are updated. While the filter is stabilizing, the input samples are returned without processing, effectively performing an identity transformation. Once the filter state has stabilised, the algorithm transitions to the real-time filtering phase, where it processes incoming samples according to the defined filtering strategy.

Algorithm 1 Custom Time-Domain Filtering to Minimize Stimulation Artifact

```
1: Initialization:
2: Initialize buffer  $\mathbf{X}$  of past samples with length  $N_{\text{buffer}}$ 
3: Initialize autoregressive (AR) predictor of order  $p$ .
4: Initialize running statistics: median  $M$ , quantiles  $\mathbf{Q} = [Q_{20}, Q_{80}]$ 
5: Initialize quantile velocities  $\mathbf{V} = [V_M, V_{Q_{20}}, V_{Q_{80}}]$  to zero
6: Define momentum parameters  $\mu$  and learning rate  $\alpha$ 
7: Set suppression flag to False

8: Filter Warm-Up:
9: while buffer  $\mathbf{X}$  is not full do
10:   Acquire next sample  $x_t$ 
11:   Append  $x_t$  to buffer  $\mathbf{X}$ 
12:   Initialize running statistics  $M$ ,  $\mathbf{Q}$  using initial samples
13:   Update AR predictor coefficients using buffer  $\mathbf{X}$ 

14: Real-Time Filtering:
15: for each new incoming sample  $x_t$  do
16:   Update buffer  $\mathbf{X}$  with  $x_t$ .
17:   Update running statistics using momentum:
18:     Compute gradient  $\nabla$ 
19:     Update velocities:  $\mathbf{V} \leftarrow \mu \cdot \mathbf{V} + \alpha \cdot \nabla$ 
20:     Update statistics:  $\begin{bmatrix} M_t \\ Q_{20,t} \\ Q_{80,t} \end{bmatrix} \leftarrow \begin{bmatrix} M_{t-1} \\ Q_{20,t-1} \\ Q_{80,t-1} \end{bmatrix} + \mathbf{V}$ 
21:   if in suppression period then
22:     Decrement suppression counter  $S_{\text{count}}$ 
23:     if  $S_{\text{count}} = 0$  then
24:       Exit suppression period
25:     Predict output  $y_t$  using AR predictor
26:   else if  $x_t > Q_{80} + \delta$  or  $x_t < Q_{20} - \delta$  then ▷ Threshold crossing
27:     Enter suppression period
28:     Set suppression counter  $S_{\text{count}} = N_{\text{suppression}}$ 
29:     Generate output  $y_t$  using AR predictor
30:   else
31:     Update AR predictor coefficients using buffer  $\mathbf{X}$ .
32:     Set output  $y_t = x_t$ 
33:   Output  $y_t$ 
```

The parameters of this algorithm are:

- N_{buffer} : Length (in number of samples) of the buffer holding past samples.
- p : Order of the autoregressive (AR) model.
- δ : Multiplicative factor threshold relative to the quantiles of the input signal to compute the dynamic threshold
- $N_{\text{suppression}}$: Number of samples to suppress once an artefact is detected (equivalent to the pulse width).
- μ, α : Momentum parameters to determine the magnitude of the updates

In our algorithm, we utilise momentum-based updates to efficiently compute the running statistics of the incoming signal in real time. Specifically, we apply this technique to update the median (M) and

the 20th and 80th percentiles (Q_{20} and Q_{80}), which are necessary to compute the threshold that may indicate the presence of stimulation artefacts.

Momentum updates are based on optimization methods in machine learning, where they help accelerate convergence and smooth out fluctuations during training. In our context, momentum allows the running statistics to adapt quickly to changes in the signal while maintaining stability, thereby enhancing the robustness of artefact detection.

For each incoming sample $\mathbf{x}_t$, we first compute the gradient between the new sample and the current estimate of each statistic:

$$\nabla_t = x_t - \theta_t$$

where θ_t represents the current estimate of the statistic at time t (either M_t , $Q_{20,t}$, or $Q_{80,t}$). We then update the velocity (momentum term) for each statistic using:

$$v_t = \mu \cdot v_{t-1} + \alpha \cdot \nabla_t,$$

where v_t is the velocity at time t , μ is the momentum coefficient (between 0 and 1) controlling the influence of previous updates, and α is the learning rate determining the step size of the update.

Finally, we update the statistic by adding the velocity to the current estimate:

$$\theta_{t+1} = \theta_t + v_t.$$

This method allows the running statistics to respond swiftly to significant changes in the signal, such as those introduced by stimulation artefacts, while smoothing out minor fluctuations and reducing the impact of transient noise. The momentum coefficient μ helps to incorporate past information, providing a balance between sensitivity and stability, whereas the learning rate α controls how aggressively the statistics are updated based on new data.

7.3 Results

7.3.1 CNN Activations Across Changes in the Dynamic Range of Input Signals

During our real-time experiments, once the feedback loop was closed— i.e. the output provided by the trained convolutional neural network (CNN) was used to guide the deep brain stimulator—the system became unstable. This instability manifested as repetitive self-triggering of the stimulation, leading to changes in stimulation parameters that did not rely on decoding the neural from deep brain LFP signals and using this information to modulate the stimulation. Upon investigation, we identified that the root cause of this instability was the change in the dynamic range of the recorded local field potential (LFP) signals when the stimulation was activated.

The activation of the stimulator introduces significant artefacts into the LFP signals (i.e. the stimulation artefact), which alters their statistical properties. Specifically, the presence of stimulation artefacts resulted in a substantial increase in the amplitude of the recorded signals, effectively changing the scale of the inputs fed into the CNN during operation. Since CNNs are sensitive to the statistical distribution of their inputs, these changes adversely affected the network's ability to generate stable and reliable features for decoding.

To illustrate this effect, we analysed the activations within the CNN during periods of dynamic range changes induced by stimulation. Figure 7.2 presents this analysis, showing both the input signal and the corresponding standard deviation of the features computed by the CNN's convolutional layers.

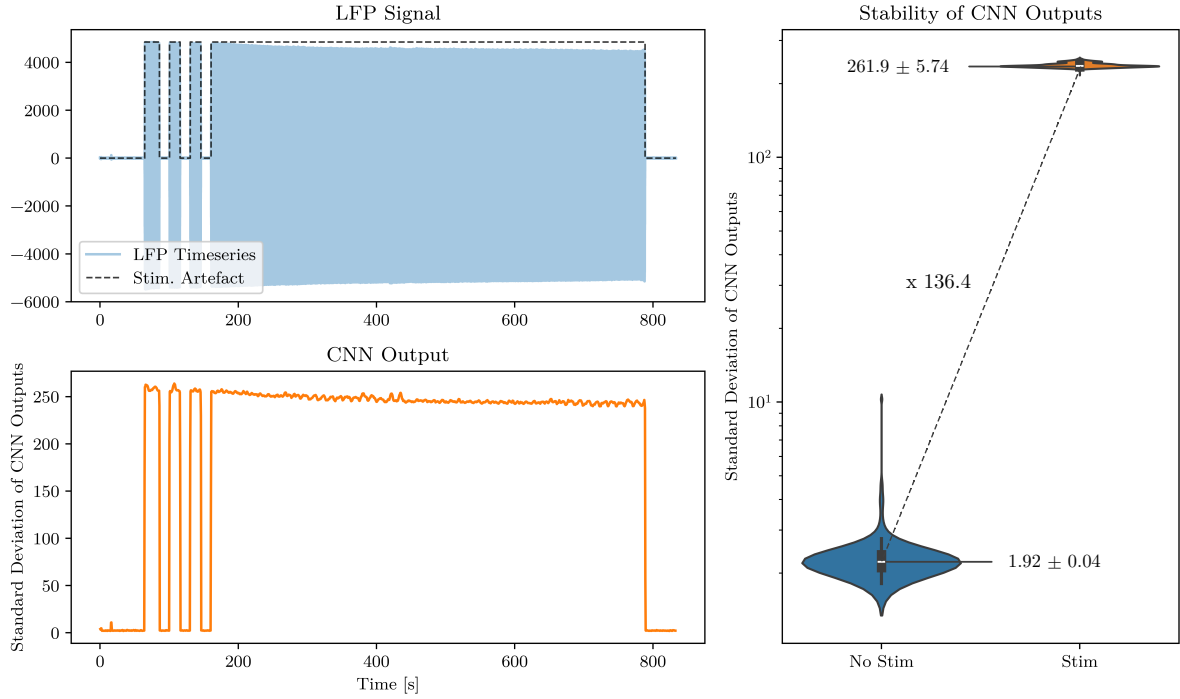


Figure 7.2: Convolutional Neural Network Activations During Dynamic Range Changes. (Top) Input LFP signal with switches in stimulation state. The signal amplitude increases significantly when stimulation is turned on due to the introduction of stimulation artefacts. (Bottom & Right) Standard deviation of the features computed by the CNN's convolutional layers over time. The standard deviation increases dramatically when stimulation is active, changing by an average factor of 136.4 compared to the non-stimulation periods. This indicates a substantial change in feature scale, leading to instability in the network's output.

In the top panel of Figure 7.2, we depict an input LFP signal, highlighting the periods when stimulation is switched on and off. The transitions between stimulation states are clearly visible, with noticeable changes in signal amplitude. When the stimulation is turned on, the amplitude of the LFP signal increases sharply due to the superimposed stimulation artefact. This artefact arises from the electrical activity introduced by the stimulation pulses, which contaminates the recorded neural signals.

The bottom and right panels of Figure 7.2 shows the standard deviation of the features computed by the CNN's convolutional layers over the same time period. We use the standard deviation as a measure of the scale of the feature values extracted by the network. When the stimulation is inactive, the standard deviation of the features remains relatively stable and low, reflecting consistent input signal characteristics that can be processed by the neural network. However, when the stimulation is activated, there is a sizeable increase in the standard deviation of the features. Quantitatively, the

standard deviation changes by an average factor of 136.4 between non-stimulation and stimulation periods.

The substantial change in feature scaling indicates that the CNN extracts features on different scales depending on the stimulation state. Consequently, CNNs trained on data from a specific stimulation condition struggle to generalise when the input distribution changes. The network’s internal representations are optimised for input signals within a certain dynamic range and statistical distribution. Significant deviations due to changes in dynamic range impair the CNN’s ability to extract meaningful features and generate accurate predictions.

This instability arises because the CNN’s activations – and therefore its outputs – are sensitive to the scale of the input data. In the presence of stimulation artefacts, amplified input signals lead to disproportionately large activations in the convolutional layers. These exaggerated activations can saturate the network’s neurons, causing the outputs to no longer discriminate between different neural states and resulting in unreliable predictions.

Furthermore, the unstable output from the CNN feeds back into the system, causing inappropriate adjustments to the stimulation state. For instance, if the CNN incorrectly predicts a movement state due to the altered input scale, it triggers the stimulator to change its state, which in turn further modifies the input signal characteristics. This creates a feedback loop where the CNN's instability is exacerbated by the very changes it induces in the system, preventing the closed-loop system from reaching a stable operating point.

This phenomenon highlights a significant challenge in deploying CNNs for real-time adaptive neurostimulation: the need to ensure robustness to variations in input signal characteristics, particularly those induced by the system's own actions. In the context of deep brain stimulation, the stimulator's operation can directly affect the recorded neural signals through artefacts and changes in dynamic range. Neural decoding pipelines must be designed to handle such variations.

7.3.2 Time-Domain Filtering

In this section, we evaluate the effectiveness of the custom time-domain filter introduced in this work in mitigating stimulation artefacts in local field potential (LFP) signals and its impact on the stability

of convolutional neural network (CNN) features used for neural decoding. We present an analysis of how the filter affects the time-series signals, the distribution of signal amplitudes, the power spectral density across relevant frequency bands, and the CNN feature activations.

To assess the filter's effect on the recorded LFP signals, we analysed the time-series data before and after applying the filter, both in the presence and absence of stimulation artefacts. Figure 7.3A displays representative segments of the LFP signals under these conditions. In the presence of stimulation artefacts (left panels), the unfiltered signals exhibit large amplitude excursions corresponding to the artefacts introduced by the stimulation pulses. After applying the custom filter, these extreme values are effectively suppressed, resulting in a signal distribution that closely resembles the intrinsic neural activity.

In contrast, when stimulation artefacts are absent (right panels of Figure 7.3A), the time-series signals before and after filtering are virtually identical, indicating that the filter minimally affects the underlying neural signals in the absence of artefacts.

To quantitatively assess the impact on signal distribution, we examined the amplitude histograms of the LFP signals before and after filtering (Figure 7.3B). In the presence of stimulation artefacts, the unfiltered signals display a distribution with heavy tails, reflecting the high-amplitude artefacts. The filtered signals, however, exhibit a distribution with reduced skewness and kurtosis, closely resembling the distribution observed in artefact-free conditions. The Kullback-Leibler (KL) divergence between the distributions before and after filtering was significantly reduced when the filter was applied (from 0.112 to 0.0015), demonstrating that the filter effectively normalises the signal distribution without distorting the intrinsic neural activity.

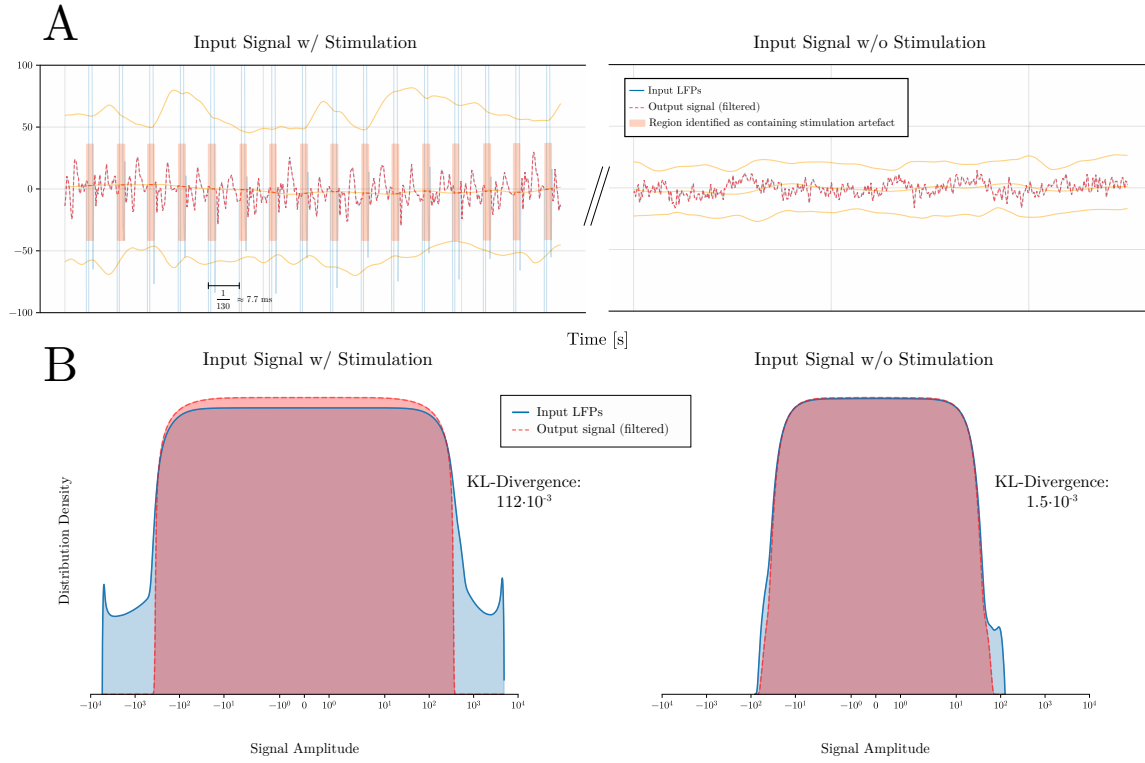


Figure 7.3: Effect of the Custom Time-Domain Filter on LFP Signals and their Distribution. (A) Time-series LFP signals before (blue) and after (red, dashed) applying the custom filter, displayed for conditions with stimulation artefacts (left panels) and without artefacts (right panels). In the presence of stimulation artefacts, the unfiltered signals exhibit large amplitude spikes corresponding to the stimulation pulses, which are effectively suppressed by the filter. When stimulation artefacts are absent, the filtered and unfiltered signals are nearly identical, indicating minimal impact on the intrinsic neural activity. (B) Probability density functions (PDFs) of the signal amplitudes before (blue) and after (red) filtering (shown on a logarithmic scale to highlight differences in the tails of the distributions). With stimulation artefacts present (left panel), the unfiltered signal displays a distribution with heavy tails due to high-amplitude artefacts. The filter significantly reduces these extreme values, resulting in a distribution similar to the artefact-free condition. The KL divergence between the unfiltered and filtered distributions decreases from 0.112 (with artefacts) to 0.0015 (without artefacts), demonstrating the filter's effectiveness in normalizing the signal distribution.

To further evaluate the filter's performance, we analysed its impact on the power spectral density (PSD) of the LFP signals across relevant frequency bands. Table 7-1 summarises the power measurements before and after filtering in three frequency bands: around the stimulation frequency (127–133 Hz), the first harmonic (257–263 Hz), and the physiological signal band (2–125 Hz).

The filter achieved a substantial reduction in power at the stimulation frequency and its first harmonic, decreasing the power by 99.70% and 99.84%, respectively. This indicates that the filter effectively suppresses the stimulation artefacts concentrated in these frequency ranges. Importantly, in

the physiological signal band (2–125 Hz), the filter only reduced signal power by 5.6%, demonstrating that it largely preserves the intrinsic neural signals critical for decoding.

Table 7-1: Effect of the Custom Time-Domain Filter on Signal Power Across Different Frequency Bands.

Frequency Band (Hz)	Input LFP Power	Filtered Signal Power	Power Reduction (%)
127 – 133 (Stimulation Frequency)	97,155.35	295.98	99.7
257 – 263 (First Harmonic)	319,985.49	497.66	99.84
2 – 125 (Putative Physiological Signal)	374.67	353.67	5.6

A critical aspect of deploying CNNs in real-time neural decoding is ensuring the stability of the network's feature activations across varying stimulation conditions. To assess how the custom filter influences the CNN's internal representations, we examined the standard deviation of the features computed by the CNN's convolutional layers before and after applying the filter.

Figure 7.4A displays the filtered LFP signal alongside the standard deviation of the CNN features over time. Prior to filtering, the presence of stimulation artefacts caused significant fluctuations in the feature activations, with the standard deviation increasing dramatically during periods of stimulation. This instability led to a multiplicative scaling factor of approximately 136.4 between the features computed during non-stimulation and stimulation periods, rendering the network's output unreliable.

After applying the custom filter, the standard deviation of the CNN features remained consistent across stimulation state changes, with a reduced scaling factor of approximately 1.27. Figure 7.4B illustrates the distribution of the CNN feature standard deviations before and after filtering, highlighting the significant improvement in feature stability. The filter effectively mitigated the impact of the stimulation artefacts on the network's activations, ensuring that the CNN operates on inputs of consistent scale and maintains reliable decoding performance.

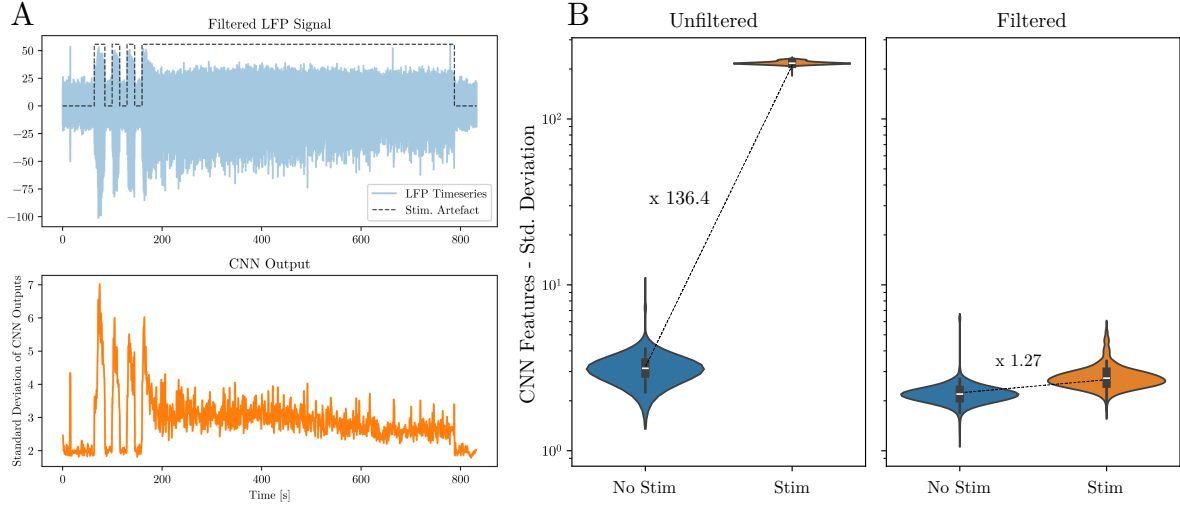


Figure 7.4 Effect of the Custom Time-Domain Filter on CNN Feature Stability. (A) Filtered LFP signal (top panel) and the corresponding standard deviation of the CNN features over time (bottom panel). The application of the custom filter results in consistent feature activations across changes in stimulation state, with minimal fluctuations in the standard deviation. (B) Histograms of the standard deviation of CNN features before (left subplot) and after (right subplot) applying the filter. Prior to filtering, the CNN features exhibited a substantial difference in scaling between non-stimulation and stimulation periods, with a multiplicative factor of 136.4. After filtering, this difference is reduced to a factor of 1.27, indicating that the features are now on the same order of magnitude regardless of the stimulation state. This stabilization is crucial for the reliable functioning of the CNN in a closed-loop aDBS system.

The results demonstrate that the custom time-domain filter effectively suppresses stimulation artefacts in LFP signals without significantly altering the intrinsic neural activity. By dynamically identifying and replacing artefact-contaminated samples using an autoregressive predictor, the filter preserves the essential features of the neural signals required for accurate decoding.

The substantial reduction in power at the stimulation frequency and its harmonic confirms the filter's efficacy in eliminating artefacts concentrated in those bands. The minimal impact on the physiological signal band ensures that the neural information pertinent to the patient's movement state is retained, which is critical for the CNN's decoding performance.

Moreover, stabilizing the CNN's feature activations across changes in stimulation state addresses the instability issues observed when deploying the neural network in a closed-loop system. The filter's ability to maintain consistent input signal scaling prevents the propagation of erratic activations through the network, thereby enhancing the reliability of the CNN's output. This stability is essential

for the effective operation of the aDBS system, as it ensures that stimulation adjustments are based on accurate and consistent neural decoding.

7.3.3 Stability of Feature-based Methods

We investigated the scaling of spectral features across changes in stimulation condition. Figure 7.5 illustrates the time series and corresponding distributions of spectral features extracted from the LFP signals during periods with and without stimulation. Notably, the majority of spectral features remained stable (i.e. within the same order of magnitude) across stimulation conditions, except for features that include the stimulation frequency (130 Hz) or its harmonics in its power estimation.

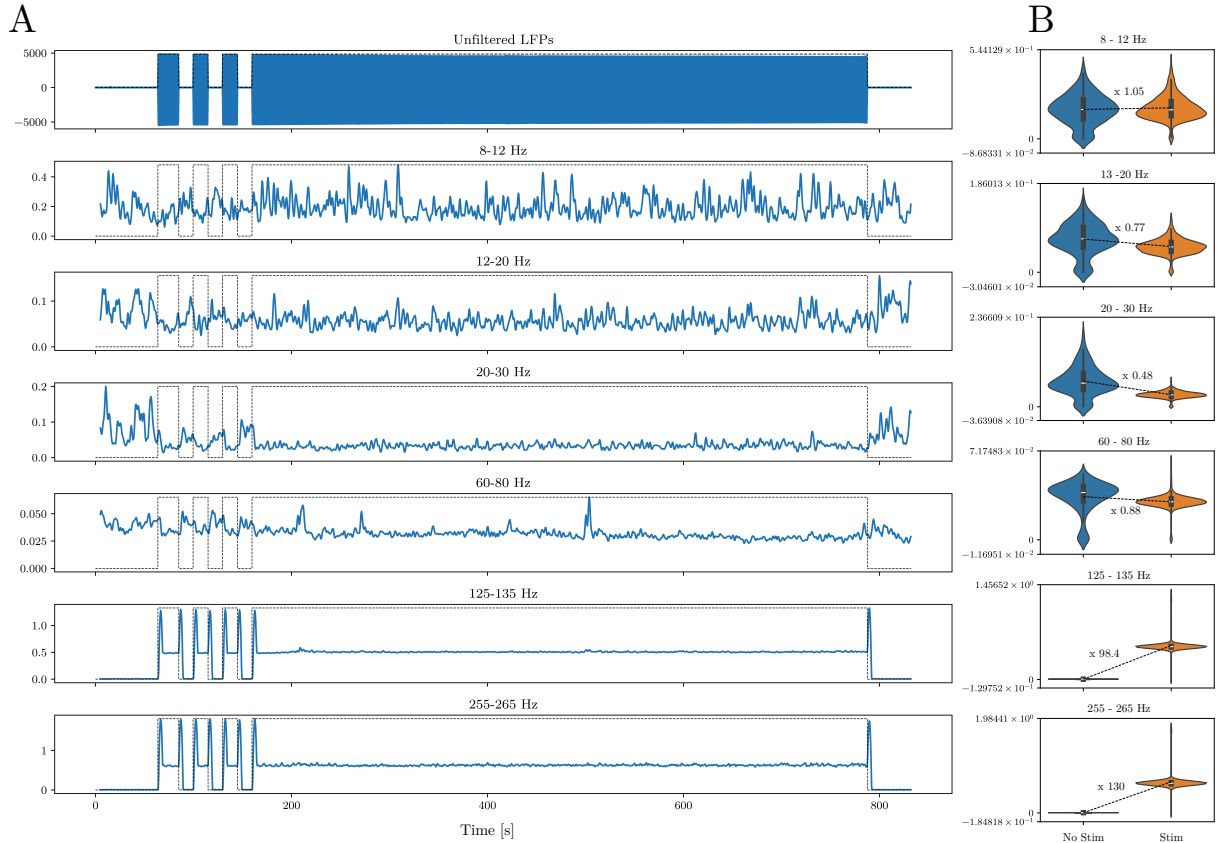


Figure 7.5: Scaling of Spectral Features Across Stimulation Conditions. (A) Time series of spectral features extracted from the LFP signals during periods with and without stimulation artefacts. The features represent the logarithm of the spectral power in canonical frequency bands. The majority of the spectral features exhibit consistent amplitudes across different stimulation conditions, indicating stability in their scaling across stimulation conditions. Only the features in the frequency bands around the stimulation frequency (130 Hz) and its

harmonics (260 Hz) show significant deviations due to the presence of stimulation artefacts. (B) Distribution of the spectral features during stimulation and non-stimulation periods. The distributions reveal that, aside from the bands corresponding to the stimulation frequency and its harmonics, the spectral features maintain similar distributions across different conditions. This suggests that feature-based methods are less affected by scaling issues introduced by stimulation artefacts, except in the specific frequency bands where the artefacts are present.

7.4 Discussion

In this section, we developed a real-time pipeline for adaptive deep brain stimulation (aDBS) driven by convolutional neural networks (CNNs). We addressed the challenge posed by stimulation artefacts, which significantly alter the distribution of local field potential (LFP) signals and impair the stability of CNNs when used in closed-loop systems. By introducing a custom time-domain filter capable of operating in real time with minimal latency, we effectively eliminated stimulation artefacts from the signals, stabilizing the CNN activations and ensuring reliable neural decoding.

This work highlights the importance of mitigating stimulation-induced signal variations to maintain the performance of machine learning models in neurostimulation applications. While the custom filter significantly improved the robustness of CNN model, it remains relevant to how explore traditional decoding approaches, such as feature-based methods, compare to CNNs in terms of their stability and sensitivity to stimulation artefacts.

7.4.1 Comparison with Feature-Based Methods

Feature-based methods for neural decoding involve extracting predefined characteristics from the signals – such as spectral power in specific frequency bands – and using these features as inputs to classifiers or regression models. These methods are often considered more interpretable and may inherently suffer less from scaling issues caused by changes in stimulation state.

The stability of spectral features across stimulation conditions suggests that feature-based methods may be inherently more robust to changes in signal scaling induced by stimulation artefacts. Since these methods rely on aggregated measures of signal characteristics within specific frequency bands,

they are less sensitive to transient changes in signal amplitude that can disproportionately affect CNN activations.

However, the features corresponding to the stimulation frequency and its harmonics do exhibit scaling differences due to the artefacts. To address this, our custom time-domain filter can be integrated into the feature extraction pipeline. By suppressing the artefacts in these frequency bands, the filter can ensure that the spectral features accurately represent the underlying neural activity without contamination from stimulation-induced artefacts.

7.4.2 Limitations

While the proposed custom time-domain filter demonstrates effectiveness in suppressing stimulation artefacts and stabilizing convolutional neural network (CNN) activations, several limitations must be acknowledged.

The filter has not been systematically tested on a wide range of deep brain stimulation (DBS) datasets. DBS signals can vary significantly between patients due to differences in neural anatomy, disease pathology, electrode placement, and stimulation parameters. Factors such as varying stimulation frequencies, pulse widths, and amplitudes can influence the characteristics of the artefacts and the underlying neural signals. Comprehensive testing on larger and more varied datasets is essential to validate the filter’s robustness and adaptability in real-world applications.

The filter relies on several parameters – such as threshold levels, suppression durations, momentum coefficients, and learning rates – that require tuning to ensure optimal performance. These parameters were manually adjusted based on empirical observations in the current study. However, manual tuning can be time-consuming and may not be feasible in dynamic clinical environments. The filter’s performance may degrade if the parameters are not appropriately set for a given patient’s signal characteristics or stimulation settings. Developing autotuning schemes or adaptive algorithms that can automatically adjust these parameters in real time would enhance the filter’s practicality.

While the custom filter addresses the issue of scaling instability in CNNs caused by stimulation artefacts, alternative strategies may exist to dynamically stabilise neural network activations without external filtering. Incorporating custom normalization layers that adjust the activations during training and inference to maintain consistent scaling within the neural network, could mitigate

sensitivity to changes in input distribution and potentially reduce the impact of artefacts on network performance.

Another avenue involves the use of activation functions that are robust to input scaling, such as self-normalizing neural networks with scaled exponential linear units (SELU) (Klambauer et al., 2017). Additionally, training the network with data augmentation techniques that simulate variations in signal scaling or incorporating regularization methods could improve the network’s resilience to input fluctuations. Recent literature in the field of neural signal processing has explored adaptive filtering and deep learning techniques that jointly learn to suppress artefacts while performing classification or regression tasks. For example, adversarial training methods (Goodfellow et al., 2016) can enhance robustness to input perturbations by exposing the network to adversarial examples during training.

While the filter is designed to preserve the physiological signal content, there is a risk that it may inadvertently attenuate or distort features of the neural signals that are relevant for decoding. The suppression of artefacts may overlap with high-frequency components of the neural activity that carry important information. Evaluation of the filter’s impact on signal features across the entire frequency spectrum is necessary to ensure that critical neural information is not lost.

7.4.3 Implications for Adaptive DBS Implementations Based on Convolutional Neural Networks

The application of convolutional neural networks (CNNs) in adaptive deep brain stimulation (aDBS) applications holds promise due to the ability of these model to learn complex patterns directly from raw neural signals. However, the unique challenges posed by DBS—particularly stimulation artefacts that alter the distribution of neural time-series data—can impair the stability and performance of CNN-based decoders.

Stimulation artefacts introduce substantial changes in the dynamic range and statistical properties of recorded signals, leading to instability in CNN activations and unreliable outputs. CNNs are inherently sensitive to the statistical characteristics of their inputs; thus, changes in signal scaling caused by artefacts can degrade their performance. This instability is especially problematic in closed-loop systems, where the decoders output influences stimulation parameters, potentially creating feedback loops that exacerbate signal distortions. Incorporating artefact suppression techniques, such

as the custom time-domain filter we propose and implement, can help stabilise the input signals and maintain consistent CNN activations across different stimulation states.

We contrast this with feature-based methods, which are generally less sensitive to scaling issues caused by stimulation artefacts. By extracting predefined features like spectral power in specific frequency bands, these methods reduce the impact of transient signal variations on decoder performance. CNNs, however, have the potential to uncover novel neural patterns and biomarkers not captured by traditional feature-based methods.

Moving forward, the implications for CNN-based aDBS implementations include the need for robust preprocessing pipelines to suppress artefacts and optimise network architectures to mitigate sensitivity to input scaling. Exploring hybrid models that combine the strengths of CNNs and feature-based methods may offer a balanced approach, harnessing the advanced pattern recognition capabilities of CNNs while maintaining stability and interpretability.

While CNNs present exciting opportunities for advancing aDBS therapies, their successful implementation requires careful consideration of the challenges posed by stimulation-induced signal variations. Balancing the advantages of CNNs with the stability and interpretability of feature-based methods may provide an effective path towards developing reliable aDBS systems.

7.5 Conclusion

In this section, we presented a real-time pipeline for adaptive deep brain stimulation (aDBS) that integrates data recording, signal processing, model training, and closed-loop operation with a stimulator in the loop. We identified significant challenges related to the stability of convolutional neural network (CNN) models when encountering changes in the distribution of neural signals, such as those occurring during switches in stimulation conditions. To address these issues, we implemented a custom time-domain filter that suppresses stimulation artefacts, thereby enhancing the stability of CNN feature activations. Our results demonstrate that this filter not only stabilises the CNN outputs but also preserves the integrity of the physiological signals, enabling reliable and robust aDBS performance. Furthermore, we compared CNN-based methods with traditional feature-based approaches, highlighting the advantages and limitations of each in the context of adaptive neuromodulation systems.

7.6 Summarised Findings

- **Development of a CNN-Based aDBS Pipeline:** We implemented a complete pipeline for adaptive deep brain stimulation based on Convolutional Neural Networks (CNNs), which includes data recording and processing, model training and preparation for inference, and real-time closed-loop signal processing with an integrated stimulator.
- **Identification of Stability Challenges in CNN Models:** We documented the challenges associated with the stability of convolutional neural networks when exposed to changes in the distribution of neural signals, particularly during transitions between different stimulation conditions. These changes can lead to significant scaling issues that impair the performance of CNN-based decoders.
- **Implementation of a Custom Time-Domain Filter:** To mitigate the numerical instability caused by stimulation artefacts, we designed and implemented a custom time-domain filter capable of operating in real time with minimal latency. This filter effectively suppresses stimulation artefacts, enhancing the stability of the features computed by CNNs and ensuring consistent decoding performance.
- **Comparison with Feature-Based Methods:** We contrasted the performance of CNN-based methods with traditional feature-based approaches. Our findings suggest that while CNNs require careful handling of signal scaling due to their sensitivity to input distribution changes, feature-based methods inherently offer greater stability across varying stimulation conditions.

8 General Discussion

Each chapter of this thesis has its own focused discussion, addressing specific technical or experimental findings. In this concluding section, we aim to bring those discussions together. By synthesizing results from five distinct but complementary projects, we aim to make some holistic remarks about the role of data-driven methods in adaptive deep brain stimulation (aDBS), the challenges limiting the real-world applicability of the work performed in this thesis and discuss how our findings point to future directions for clinically robust adaptive neuromodulation. We will revisit the thesis aims, introduced in Chapter 1, and show how they were addressed across the individual studies. We then consider the broader implications for machine learning (ML)-based systems in aDBS, including their potential as a tool to study neural data and to discover and validate biomarkers for aDBS. We conclude by examine relevant limitations and propose some outlook for this rapidly evolving field.

8.1 General Overview of Results

Over the course of this thesis, we sought to explore data-driven solutions to advance adaptive DBS (aDBS) and thereby address some of the clinical limitations of continuous deep brain stimulation (cDBS). Although cDBS is a well-established therapy for conditions like Parkinson's disease (PD) and essential tremor (ET), its inability to adapt to fluctuating symptom states can result in suboptimal symptom control, battery drain, and potential side effects (Little & Brown, 2012; Meidahl et al., 2017). To accomplish this, we formulated four major aims. First, we wanted to establish data-driven strategies for adaptive DBS that exceed the commonly used single-band threshold approach. Second, we sought to evaluate biomarkers and frequency bands beyond the “beta-only” perspective, including cortical signals and network-level metrics. Third, we aimed to elucidate and harness neural correlates of both normal behaviour and disease processes, spanning movement, speech, and tic generation. Lastly, we explored how to develop real-time-compatible neural network pipelines that could feasibly run under the stringent requirements of an aDBS device.

Throughout five distinct but complementary studies, we consistently highlighted how dynamic pathological activity requires a flexible, feedback-driven approach to DBS. Each study simulated or

directly tested the feasibility of real-time methods, building, where appropriate, causal signal processing pipelines that rely only on past data and carefully balance computational complexity. By investigating multiple recording sites—cortical and subcortical—we underscored that no single region or frequency band captures the entire scope of pathological activity (S. M. Peterson et al., 2021). Moreover, these efforts revealed that while complex neural networks, specifically convolutional neural networks (CNNs), can uncover intricate patterns in electrophysiological signals, their power comes with the need for substantial training data, artefact handling, and interpretability tools (Li et al., 2017; Rodriguez et al., 2023).

Project 1 directly tackled the first aim by showing how convolutional neural networks (CNNs) can outperform simpler feature-based models in decoding oscillatory signals. Our results aligned with the broader literature (Livezey & Glaser, 2021; Merk, Peterson, Köhler, et al., 2022), demonstrating that CNNs learn nuanced spatiotemporal patterns from synthetic and real LFP data in ways that single-band threshold logic overlooks. Nonetheless, we found that classic feature-engineering methods can perform comparably in well-defined tasks for which the relevant correlates are well known and can be captured through feature engineering—underscoring a key trade-off between computational complexity and the potential discovery of less obvious features.

Project 2 helped us pursue the second aim more thoroughly by examining whether multiple frequency bands and recording sites (subthalamic LFP, scalp EEG) could detect movement intention in Parkinson’s disease. We showed that while moderate accuracy could be achieved, large inter-patient variability and artefact-related factors limit straightforward generalization. The findings supported the notion (J. Herron, Kullmann, et al., 2024; Merk, Peterson, Lipski, et al., 2022) that cortical signals can prove just as valuable—or even more so—than subcortical signals, thereby advocating for a more expansive view of which neural signatures might guide aDBS.

Project 3 took a different angle, focusing on the third aim by delineating distinct subthalamic signals for speech vs. upper limb movement in Parkinson’s disease. These data indicate that speech production may be associated with subtly different oscillatory patterns than those underlying upper limb movement and suggest that it might be possible to capture these for the purposes of adapting stimulation to specific movement states for the purposes of optimising therapy. This insight reaffirms the premise that aDBS systems must differentiate between diverse clinical states—such as speech, gait, or tremor—rather than suppressing subthalamic oscillations unselectively. Were these speech-related patterns harnessed in a real-time pipeline, clinicians might mitigate side effects related to

speech difficulties, a point emphasised in (Wilkins et al., 2023), where authors advocate for symptom-specific neuromodulation.

Project 4 proceeded to test the relevance and potential use of connectivity-based metrics—framed in our aims as exploring biomarkers for aDBS—in Tourette syndrome. Although we saw encouraging group-level differences in fronto-thalamic coherence before tic onset (Wehmeyer et al., 2024), translating them into single-trial predictors proved elusive. This underscores one of the central challenges for aDBS: group-level biomarkers may not always be reliable at the individual-event or trial scale, particularly in disorders where symptom manifestations vary widely. Our findings here dovetailed with prior cautions expressed in (Meidahl et al., 2017) and (Watts et al., 2020) about the pitfalls of applying broad statistical markers to real-time, patient-specific contexts.

Project 5 focused on another of the aims of this thesis—building real-time-compatible neural network-based aDBS pipelines—by addressing a specific practical barrier: stimulation artefacts that destabilise neural network stability and therefore decoding performance. We developed a custom time-domain filter that effectively constrained sudden amplitude shifts due to changes in stimulation settings. The feasibility of using CNN-based decoders in an online setting was strengthened by this, although challenges remain for the clinical adoption of embedded neural-network driven aDBS, including artefact management, model calibration, and hardware constraints (Rodriguez et al., 2023; Zhu et al., 2020).

Overall, these projects together highlight how data-driven machine learning approaches have the potential to open new frontiers for adaptive DBS. They demonstrate that CNNs and related methods discover valuable patterns in high-dimensional neural data, but also require extensive datasets, robust artefact handling, and a proper interpretability framework—points previously noted in the literature (Li et al., 2017; Merk, Peterson, Köhler, et al., 2022). Simple threshold-based control loops remain useful in certain domains, especially where a single, well-characterised pathological oscillator prevails, but the deeper nuance gleaned from advanced ML is necessary when dealing with complex multi-symptom conditions. Hence, by explicitly grappling with our four stated aims—data-driven classification beyond single-band thresholds, broader biomarkers and connectivity metrics, behaviours-specific decoding, and robust real-time CNN pipelines—this thesis contributes towards carving out an interdisciplinary path toward a new generation of adaptive, multimodal neuromodulation strategies.

8.2 Machine Learning as a Tool to Study Electrophysiological Data

Machine learning has emerged as an important methodological approach in neuromodulation research, serving both as a real-time decoding engine for adaptive DBS (aDBS) and as a powerful research tool for investigating fundamental aspects of brain function (Glaser et al., 2020; Livezey & Glaser, 2021; Richards et al., 2019). Its influence now spans nearly many facets of deep brain stimulation therapy, including surgical targeting, biomarker discovery, postoperative device programming, and continuous feedback control (Germann et al., 2024; Wilkins et al., 2023; Wong et al., 2023). This adoption reflects the inherently high-dimensional, nonlinear, and often non-stationary nature of electrophysiological signals—which can include subcortical local field potentials (LFPs) as well as cortical signals such as ECoG or EEG—for which traditional linear methods may be insufficient to capture the intricate structure in these datasets, prompting researchers to embrace more flexible and data-driven approaches.

(Watts et al., 2020) highlight how machine learning solutions have been productively introduced at multiple decision points in DBS—ranging from candidate selection, where ML can identify which patients are likely to benefit most (e.g., by analysing large cohorts and detecting presurgical risk factors), to programming optimization, surgical targeting, and even understanding the mechanisms by which stimulation modifies pathological circuits. In this broader context, decoding neural signals is only one application: machine learning can help classify symptom severity remotely via wearable sensors, suggest initial parameter settings that might expedite the months-long programming process, or even assist in identifying electrophysiological markers indicative of suboptimal lead location.

When focusing on real-time decoding, data-driven strategies have become attractive precisely because of their ability to model nonlinear transformations (Li et al., 2017). By conceptualizing a neural decoder as a function that maps raw electrophysiological signals to continuous or categorical variables—such as motor trajectories, speech articulation states, or tremor severity—deep networks have been shown to surpass more rigid, linear approaches (Livezey et al., 2019). This advantage is especially pronounced when sufficient training data are available and when the target phenomena are complex or poorly characterised by simple heuristics. Nonetheless, as cautioned by authors in (Richards et al., 2019), successful deep learning frameworks often require large numbers of parameters, which can become difficult to interpret or justify in smaller clinical datasets. Neural decoding thus

becomes a balancing act, where deep networks offer the potential for expressive power but may demand greater data resources, specialised hardware, and advanced interpretability methods (Merk, Peterson, Köhler, et al., 2022).

In recent years, a series of methodological advances has demonstrated the promise of deep learning for analysing electrophysiological data in diverse clinical and experimental contexts (Anumanchipalli et al., 2019; Glaser et al., 2020; Schirrneister et al., 2017). With regard to speech decoding from human sensorimotor cortex, for instance, (Livezey et al., 2019) showed that deep networks could identify subtle cross-frequency couplings and hierarchical latent structures that were overlooked by linear or shallow methods. Their findings showcase how nonlinear methods can yield both higher prediction accuracy and reveal novel representations within neural signals. These sorts of insights resonate with attempts in aDBS to identify spectral correlates, spatiotemporal signatures, or connectivity-based markers that might be predictive of disease states or clinically relevant behavioural states, yet difficult to isolate through simple analytical tools.

Parallel discussions by authors in (W.-J. Neumann et al., 2019) envision “intelligent aDBS” that capitalises on precisely these strengths. By combining subcortical and cortical recordings in a hierarchical deep network, the system could parse multiple oscillatory features and behaviours, then dynamically adapt stimulation amplitude or frequency to mitigate various symptoms—tremor, dyskinesia, gait freezing, or speech deficits. Rather than employing a single-band threshold, such intelligent aDBS would leverage the sort of rich, data-driven representation exemplified by modern deep networks. Through multicentre collaborations designed to gather large-scale training data, as Neumann et al. argue, the field could refine these models to a point where they reliably capture context-sensitive markers of pathology, thereby enabling symptom-specific stimulation regimes.

A study by (Alagapan et al., 2023) on depression provides an example of how advanced machine learning can exceed simple classification tasks and illuminate both therapeutic effects and circuit-level mechanisms. By deploying an explainable AI framework to discriminate “sick” from “stable response” states in patients receiving subcallosal cingulate DBS, the authors identified a “spectral discriminative component” of the local field potentials that signalled the transition to improved mood. Such a biomarker offered high classification accuracy while also revealing the electrophysiological underpinnings of clinical improvement, thus bridging robust decoding performance with deeper mechanistic insights. This speaks to a more general theme in neural decoding, where advanced

learning architectures can uncover unknown frequency bands or coupling patterns that link directly to specific clinical features.

Nevertheless, deploying machine learning in aDBS requires overcoming substantial challenges: data collection must be extensive enough to prevent overfitting (Watts et al., 2020), and real-time usage of large networks often depends on specialised hardware or energy-hungry computation (Zhu et al., 2020). Interpreting these complex models remains difficult, particularly in low-sample clinical settings, and any data-driven pipeline must handle artefacts stemming from stimulation itself as well as the non-stationarity of disease states over time. In that sense, the highly successful systems described in previous studies (Alagapan et al., 2023; Livezey et al., 2019; W.-J. Neumann et al., 2019) illustrate not only future possibilities but also the meticulous design and large-scale data support needed to ensure reliable, stable performance.

Taken together, these examples reinforce the conclusion that machine learning’s role in neuromodulation likely will extend beyond a tool for feedback control. It provides an analytic and hypothesis-generating framework that can reveal hidden aspects of neural computation, validate new biomarkers, and suggest targeted manipulations to mitigate disease symptoms more effectively. As ML-driven pipelines mature, they promise to clarify not just whether a biomarker is viable for real-time aDBS, but also why the biomarker arises in the circuit, which frequency bands or interactions are most salient, and how we might reconfigure stimulation parameters to exploit these insights. To achieve that promise, researchers must embrace collaborative data-sharing, refined interpretability methods, and carefully managed computational resources, building a robust foundation for truly intelligent, next-generation neuromodulation therapies (Merk, Peterson, Köhler, et al., 2022).

8.3 Machine Learning as a Tool to Elucidate and Validate Biomarkers for aDBS

Because adaptive deep brain stimulation (aDBS) hinges on the timely detection of a reliable physiological marker, identifying and rigorously validating this biomarker is very important to ensure reliable function (Bouthour et al., 2019; Groppa et al., 2023). When feedback signals are robust, aDBS can dynamically adjust stimulation parameters based on ongoing clinical states, modulating energy usage and potentially limiting side effects. Yet, as underscored by both our own experiments and

reports from the literature, reliably capturing such neural correlates can be challenging (Merk, Peterson, Köhler, et al., 2022; Wendt et al., 2022). While oscillatory features like subthalamic beta power have historically been correlated with Parkinsonian symptoms at the population level (W. Neumann et al., 2016), in this thesis we have shown how the existence of group-level correlations might not necessarily generalise to signals suitable for real-time control on an individual event basis.

Machine learning offers a systematic means of evaluating whether a given biomarker is suitable to drive an aDBS system, particularly in high-dimensional settings where manual feature selection risks overlooking crucial nonlinear or cross-frequency interactions (Little & Brown, 2012; Merk, Peterson, Köhler, et al., 2022). By applying cross-validation and validation on individual trials—methods that we employed throughout this thesis—ML-based pipelines can stress-test candidate biomarkers under realistic conditions, revealing whether they hold steady across medication states, task demands, or inter-patient variability. Our unsuccessful attempts to use front-thalamic alpha-band connectivity modulations for predicting Tourette's tics at the single-trial level highlight how a signal with strong average correlations can still falter as a real-time trigger. In this sense, ML frameworks not only confirm whether a biomarker is repeatable and robust but also indicate whether more data sources (e.g., cross-frequency features, connectivity matrices, or peripheral sensor streams) might be needed to achieve acceptable performance (S. Wang et al., 2023).

Equally important, machine learning can serve as a powerful engine for biomarker discovery, moving beyond the single-band or single-region viewpoint. Instead of limiting analysis to known signatures in a canonical frequency band, data-driven models can sift through a vast feature space to unearth novel signals, such as intricate cross-frequency couplings or transient waveform shapes (Livezey & Glaser, 2021; W. Neumann & Rodriguez-Oroz, 2021). These hidden patterns may prove especially valuable when the usual narrow-band markers are insufficient to capture pathological states. Project 1 of this thesis, for instance, showed how data-driven classifiers can adapt to nuanced signal features, sometimes capturing oscillatory sharpness or waveform asymmetry that hand-engineered methods overlook. Harnessing this flexibility, machine learning could integrate recordings from subthalamic targets, cortical sensors, and even wearable trackers to produce composite biomarkers that are more predictive of fluctuating clinical phenomena (W.-J. Neumann et al., 2019).

Still, machine learning cannot replace a sound physiological rationale. Identifying a correlation between a neural feature and a clinical state does not guarantee that this feature should drive aDBS. Ideally, the chosen biomarker should reflect core disease mechanisms and be consistent with

established neuroscience theories (Bouthour et al., 2019). Moreover, many high-capacity models function as "black boxes," detecting correlations in the data without clarifying why those relationships arise or how generalizable they are across contexts (Merk, Peterson, Köhler, et al., 2022). Interpretable methods or post-hoc analyses of network weights are therefore crucial, not only to provide confidence in clinical use but also to refine experimental hypotheses (Alagapan et al., 2023). Clinicians must understand the reliability and failure modes of any data-driven biomarker if they are to trust it for real-time adjustments to stimulation parameters, ensuring patient safety and meeting regulatory requirements. Likewise, high inter-patient variability—common in both movement disorders and psychiatric disorders—calls for adaptive and personalised modelling. While ML approaches are well-suited to customization, they can also overfit laboratory data, especially if the training dataset does not adequately represent the heterogeneity of the target population (Watts et al., 2020).

Ultimately, leveraging machine learning to discover and validate new biomarkers for aDBS requires a balance between automation and mechanistic insight. On the one hand, data-driven algorithms can efficiently mine large, multimodal datasets to identify complex neural signatures that can outperform single-band metrics. On the other hand, without careful oversight, they risk learning spurious associations that overfit training data and fail to generalize (Wendt et al., 2022). By grounding the use of these methods in domain-driven validation, using carefully curated datasets, and favouring interpretable modelling strategies, we can harness ML's potential to guide a new generation of aDBS systems—ones that dynamically adapt to moment-to-moment variations in patient state and deliver more precise and effective neuromodulation (W. Chen et al., 2021).

8.4 Thesis Limitations

Despite the insights gained through the studies presented in this thesis, several important limitations became evident, many of which trace back to the practical constraints associated to conducting aDBS research during the immediate postoperative period with externalised patients. While this approach grants unique access to subcortical signals in the days immediately after lead implantation, it also introduces potential confounds that may obscure the steady-state neural dynamics most relevant for long-term adaptive DBS (aDBS).

A primary concern is that neural recordings in the perioperative window—in the studies of this thesis within one week of lead insertion surgery—may be substantially affected by the surgical procedure itself. Multiple groups have documented acute, surgery-related alterations in local brain physiology, sometimes referred to as the “stun” or “insertional” effect (C. C. Chen, Pogosyan, et al., 2006b; Mann et al., 2009; Thuberg et al., 2021). (Mann et al., 2009) showed that in Parkinson’s disease, the mere act of placing electrodes in subthalamic (STN) or pallidal (GPi) targets can acutely improve motor symptoms before any stimulation is delivered, suggesting that microlesions or transient oedema may dampen pathological oscillations. Similar observations have emerged in epilepsy, where lead placement in the anterior nucleus of the thalamus produced short-term seizure frequency reductions that subsequently dissipated without active stimulation (Thuberg et al., 2021). Moreover, (Sure et al., 2023) reported that postoperative changes in resting-state networks (RSNs) can persist for weeks and overshadow stable baseline connectivity patterns. Consequently, signals recorded during this limited postoperative period may provide an inflated or misleading picture of potential biomarkers and their clinical relevance.

Another issue stems from the fact that these short-term recordings cannot capture how signals evolve once patients leave the operating room, receive chronic stimulation, and undergo everyday medication adjustments. Although certain oscillatory features, such as STN beta-band activity, have shown some stability over months or years (Giannicola et al., 2012), collecting such longitudinal data with externalised leads is logistically unfeasible. The models developed in this thesis thus remain largely untested under long-term conditions characterised by ongoing stimulation, medication cycles, disease progression, and potential plastic changes in circuit activity, relevant concerns for aDBS systems (Little & Brown, 2012; Meidahl et al., 2017; Watts et al., 2020). We cannot be certain that the algorithms and pipelines proposed here would sustain the same accuracy and reliability once patients return home and use the stimulator for extended periods.

A further limitation involves the challenge of building machine learning (ML) decoders that generalise across patients. In Projects 1 and 2, we explored cross-patient training with the expectation that a model trained on multiple individuals might transfer beneficial knowledge. Instead, performance generally plateaued or worsened, implying that even seemingly similar signals (e.g., STN-LFPs in different Parkinson’s disease patients) can diverge considerably because of variations in electrode location, local physiology, or medication status. Perioperative confounds, including insertional effects, may amplify this variability, complicating attempts to learn stable, shared features. Meanwhile, studies in the literature (Merk et al., 2023; S. M. Peterson et al., 2021; Yin et al., 2024) have

demonstrated successful cross-patient decoders, but typically require larger, heterogeneous datasets, more elaborate architectures (e.g., Hilbert-based transforms, specialised neural networks), or advanced normalization. By contrast, our modest sample sizes and restricted postoperative window likely hindered such strategies. Early postoperative data, intrinsically shaped by acute surgical artefacts (Sure et al., 2023; Thuberg et al., 2021), might also fail to align with long-term signals acquired weeks or months later, leading to a mismatch that further reduces the benefits of cross-patient pooling.

The short postoperative period additionally introduces “collisions” between medication changes and the surgery itself. Many DBS trials adjust or withdraw antiparkinsonian medication to evaluate lead placement under standardised conditions, yet the postoperative neural signals could reflect not only the mechanical effects of the electrode insertion but also suboptimal or transient medication schedules (Sure et al., 2023). Consequently, the ML-based classifiers we trained may overfit to transient properties that do not match patients’ typical at-home states.

Sample size constraints compound these issues. Even though we implemented cross-validation techniques to maximise generalisability, the low number of participants in most projects limited our ability to disentangle acute surgical artefacts from bona fide biomarkers of relevant neural states. The considerable day-to-day variability of clinical conditions (for example, periods dominated by tremor as opposed to bradykinesia) cannot be captured by brief, postoperative lab sessions alone, especially if the signals are still recovering from invasive lead insertion. In addition, real-time testing in this timeframe faces several practical obstacles, such as patient fatigue, the narrow interval before IPG implantation, and the need for strict sterile protocols around externalised leads.

Recent findings illustrate that even less expected or more subtle factors can skew neural biomarkers when data is collected over short durations. (Feldmann et al., 2023) demonstrated how a festive event, in this case Christmas Eve, led to a significant reduction in subthalamic beta power in chronic recordings taken from a handful of implanted Parkinson’s disease patients. Although seemingly anecdotal, this underscores that even non-clinical, social factors—essentially real-life experiences—can cause measurable shifts in putative biomarkers. In a short-term, postoperative study, it may be impossible to gauge how dynamic external factors (whether social, emotional, or circadian) modulate the signals used for aDBS control. Such fluctuations highlight that biomarkers are not static but can reflect patients’ emotional or environmental states—an aspect that short, lab-based recordings may fail to capture or generalise.

By contrast, studies such as that of (Gilron et al., 2021) highlight what is possible when neural signals are recorded over longer intervals in real-world settings. Streaming multi-channel cortical and subcortical signals from fully implanted devices over many months not only reveals patient-specific fluctuations that occur outside controlled conditions, but also enables the derivation of more stable and comprehensive biomarkers. This evidence underscores the potential shortcomings of relying on short, perioperative data that cannot reflect full-day or multi-month variability, including medication patterns, sleep effects, or normal daily activities. The ephemeral nature of the insertional effect, combined with ongoing shifts in neural signals as patients resume their habitual routines, means that short-term data can easily overfit transient phenomena, leaving us uncertain about how well any “postoperative” model would fare in genuinely chronic conditions.

Taken collectively, these observations reinforce that the neural activity recorded in the immediate postoperative period, though valuable for feasibility tests, may not embody the stable neural characteristics on which chronic aDBS depends (Meidahl et al., 2017; Wilkins et al., 2023). Future research could harness new-generation IPGs featuring continuous, at-home data acquisition (e.g., Medtronic Percept™ or similar platforms) to follow the temporal trajectory of neural signals for days, weeks, or even months (Gilron et al., 2021). In doing so, investigators can more fully capture the true variability induced by medication cycles, everyday behaviours, circadian rhythms, and other relevant contextual variables (Feldmann, Neumann, Krause, et al., 2021; W.-J. Neumann et al., 2019; van Rheede et al., 2022). Complemented by larger cohorts, harmonised protocols across multiple sites, and advanced ML algorithms able to integrate multi-modal data (including cortical ECoG and subcortical LFP), these prolonged, ecologically valid recordings stand to yield more robust biomarkers that genuinely track pathological states over time. Ultimately, bridging the gap between acute, postoperative recordings and longer-term, real-world measures remains crucial for translating proof-of-concept machine learning pipelines into reliable, patient-tailored neuromodulation strategies.

8.5 Outlook

As adaptive deep brain stimulation (aDBS) continues to evolve, a confluence of technical, clinical, and ethical considerations suggests that its future is likely to be both transformative and complex. Recent advances in implant hardware, for instance, have yielded more flexible and precise control options for stimulation, smaller devices with rechargeable batteries, and—importantly—built-in sensing capabilities

that bring chronic, real-world neurophysiology within reach (Frey et al., 2022; Gilron et al., 2021). At the same time, emerging algorithms and software frameworks are poised to leverage these long-term data streams to provide automated, data-driven approaches to parameter tuning, remote telemedicine, or even entirely new stimulation paradigms (Johnson et al., 2024). In parallel, the field is steadily moving away from the one-size-fits-all model of continuous open-loop stimulation toward neuromodulation approaches that operate more like "brain circuit interventions" (W.-J. Neumann, Horn, et al., 2023), calibrated specifically to the evolving pathophysiology of each patient.

One significant frontier lies in the refinement of closed-loop control strategies that incorporate the dynamic, time-dependent nature of both neural and clinical states (Wendt et al., 2022; Wilkins et al., 2023). Although early aDBS designs commonly relied on single-threshold heuristics tied to one oscillatory band (for example, subthalamic beta power in Parkinson's disease) (Little et al., 2013), there is growing consensus that more complex control policies--ranging from dual-threshold designs (Bronte-Stewart et al., 2020; Velisar et al., 2019) to fully time-varying or reinforcement learning architectures--can more effectively capture symptom variability. Additionally, closed-loop approaches are maturing beyond mere amplitude adjustments: new IPGs in development or under early trials may support the automated adaptation of contact choice, frequency, or pulse width (Frey et al., 2022). As next-generation technology emerges, the challenge will be to evaluate systematically how such multidimensional adaptivity affects both clinical outcomes and side-effect profiles in larger patient cohorts (W.-J. Neumann et al., 2019; Widge, 2024).

A second critical trend is the steadily increasing sophistication of machine learning (ML) methods that can handle, interpret, and integrate diverse data streams. Early aDBS often centred on subcortical local field potentials (LFPs), but modern frameworks incorporate cortical ECoG, EEG, wearable sensors, and imaging-derived biomarkers (Merk, Peterson, Köhler, et al., 2022; Wilkins et al., 2023). Research has already shown that combining multiple modalities can improve decoding of pathological states and motor behaviours (Merk et al., 2023; Widge, 2024). In addition, "intelligent aDBS" could automatically detect and adapt to changes in user behaviour, mood, or environmental context, setting the stage for advanced neuromodulation that not only reacts to disease fluctuations (W.-J. Neumann et al., 2019) but also capitalises on windows of enhanced plasticity (Widge, 2024). Nevertheless, challenges remain in embedding sophisticated ML into an implantable device that must balance computational load, power constraints, and the user's safety, while still preserving the interpretability clinicians need to trust the system (Zhu et al., 2020).

Moving forward, harnessing the full potential of these platforms will likely demand strategies for remote connectivity and telemedicine, allowing patients to receive therapy updates or parameter changes seamlessly over secure networks (Frey et al., 2022). This approach may further democratise care by enabling smaller clinics or general neurologists to manage advanced DBS devices through user-friendly, AI-driven interfaces (Johnson et al., 2024). Yet alongside these opportunities, a growing body of work has warned of the necessity to safeguard data privacy and to maintain robust device security, given the heightened vulnerabilities of remote access (Pugh et al., 2018).

Equally pivotal are the clinical and conceptual steps needed to integrate these innovations into actual practice. Larger-scale registries, multicentre data-sharing consortia, and standardised programming workflows have been repeatedly advocated (Frey et al., 2022; Wilkins et al., 2023) to unify best practices, expedite discovery, and prevent fragmentation across different centres. The use of connectomic and imaging-based strategies, combined with physiological biomarkers, may soon yield "personalised connectomics" that better matches each patient to a specific DBS target and adaptive algorithm (W.-J. Neumann, Horn, et al., 2023). Ongoing prospective clinical trials, such as ADAPT-PD for adaptive subthalamic stimulation in Parkinson's disease, will help clarify how these new methods translate into improved outcomes and reduced side effects (Frey et al., 2022; Johnson et al., 2024).

Finally, as these advanced approaches develop, the field must remain cognizant of the ethical questions that come with increasingly autonomous devices. Neural implants that adapt their stimulation based on behavioural or circuit-level feedback could, in principle, affect personality, agency, or sense of self (Giordano, 2015; Yuste et al., 2017). Addressing these questions requires interdisciplinary collaborations, clear standards for informed consent, and rigorous regulatory frameworks. Despite these challenges, the emerging technological ecosystem--encompassing hardware, software, artificial intelligence, and clinical knowledge--hints that the next decade may see aDBS evolve from proof-of-concept experimental devices into widely used, precision-tailored therapies for numerous neurological and psychiatric conditions (Frey et al., 2022; Neumann et al., 2023; Wilkins et al., 2023). By continuing to integrate network-level insights, advanced ML, and a patient-centred perspective, aDBS stands poised to redefine neuromodulation's role in modern medicine.

8.6 Conclusion

The evolution of deep brain stimulation from a traditional, open-loop approach to one that is dynamic and fully adaptive represents a fundamental shift in how we treat neurological and psychiatric disorders. By selectively delivering stimulation based on real-time physiological signals, aDBS has the potential to deliver more precise therapy, reducing unwanted side effects and conserving battery life while maintaining or even enhancing clinical benefits. The work presented in this thesis contributes to that trajectory by developing and validating machine learning pipelines capable of automatically extracting clinically relevant features, managing stimulation artefacts, and adapting to each patient's unique patterns of neural activity. Although these methods have been tested only within restricted time windows, with relatively few patients, and in controlled environments, the outcomes nonetheless underscore the feasibility of machine learning-based aDBS systems.

Equally important, this thesis demonstrates that simple threshold-based approaches may overlook valuable information, and that richer, data-driven strategies can reveal novel biomarkers beyond conventional oscillatory bands. Moreover, because effective aDBS must function in ecologically valid scenarios, integrating cortical and subcortical signals becomes essential. When harnessed in a sophisticated control loop, machine learning offers a powerful means of capturing the complexity of pathological circuits while still responding to real-time changes in clinical states. Building on these results will require coordinated efforts to obtain longer-term data in naturalistic settings, refine the interpretability of deep models, and ensure safe deployment of computationally demanding methods on implanted hardware.

Despite the remaining barriers, including questions of interpretability, validation in fully implanted systems, and unresolved ethical concerns, clear momentum favours the convergence of advanced machine learning with next-generation neuromodulation devices. Over the coming years, approaches that fuse multimodal signal acquisition, automated biomarker discovery, and adaptive controller design hold great promise for alleviating symptom burden in movement disorders, Tourette syndrome, and other conditions. By synthesizing insights from electrophysiology, neural engineering, and machine learning, we move closer to a future in which aDBS can be tailored to the individual in a seamless, data-driven, and ethically responsible manner--offering new therapeutic horizons for patients worldwide.

9 References

- Ahn, M., Lee, S., Lauro, P. M., Schaeffer, E. L., Akbar, U., & Asaad, W. F. (2020). Rapid motor fluctuations reveal short-timescale neurophysiological biomarkers of Parkinson's disease. *Journal of Neural Engineering*, 17(4), 046042. <https://doi.org/10.1088/1741-2552/abaca3>
- Alagapan, S., Choi, K. S., Heisig, S., Riva-Posse, P., Crowell, A., Tiruvadi, V., Obatusin, M., Veerakumar, A., Waters, A. C., Gross, R. E., Quinn, S., Denison, L., O'Shaughnessy, M., Connor, M., Canal, G., Cha, J., Hershenberg, R., Nauvel, T., Isbaine, F., ... Rozell, C. J. (2023). Cingulate dynamics track depression recovery with deep brain stimulation. *Nature*, 622(7981), 130–138. <https://doi.org/10.1038/s41586-023-06541-3>
- Alawee, W. H., Basem, A., & Al-Haddad, L. A. (2023). Advancing biomedical engineering: Leveraging Hjorth features for electroencephalography signal analysis. *Journal of Electrical Bioimpedance*, 14(1), 66–72. <https://doi.org/10.2478/joeb-2023-0009>
- Albe-Fessard, D., & Fessard, A. (1963). Thalamic Integrations and their Consequences at the Telencephalic Level. In *Progress in Brain Research* (Vol. 1, pp. 115–154). Elsevier. [https://doi.org/10.1016/S0079-6123\(08\)60591-9](https://doi.org/10.1016/S0079-6123(08)60591-9)
- Aldini, G. (1804). *Essai théorique et expérimental sur le Galvanisme avec une série d'expériences*. Imprimerie de Fournier fils.
- Aldridge, D., Theodoros, D., Angwin, A., & Vogel, A. P. (2016). Speech outcomes in Parkinson's disease after subthalamic nucleus deep brain stimulation: A systematic review. *Parkinsonism & Related Disorders*, 33, 3–11. <https://doi.org/10.1016/j.parkreldis.2016.09.022>

- Alhourani, A., Korzeniewska, A., Wozny, T. A., Lipski, W. J., Kondylis, E. D., Ghuman, A. S., Crone, N. E., Crammond, D. J., Turner, R. S., & Richardson, R. M. (2020). Subthalamic Nucleus Activity Influences Sensory and Motor Cortex during Force Transduction. *Cerebral Cortex* (New York, N.Y.: 1991), 30(4), 2615–2626. <https://doi.org/10.1093/cercor/bhz264>
- Allen, D. P., Stegemöller, E. L., Zadikoff, C., Rosenow, J. M., & MacKinnon, C. D. (2010). Suppression of deep brain stimulation artifacts from the electroencephalogram by frequency-domain Hampel filtering. *Clinical Neurophysiology*, 121(8), 1227–1232. <https://doi.org/10.1016/j.clinph.2010.02.156>
- Alomar, S., King, N. K. K., Tam, J., Bari, A. A., Hamani, C., & Lozano, A. M. (2017). Speech and language adverse effects after thalamotomy and deep brain stimulation in patients with movement disorders: A meta-analysis. *Movement Disorders*, 32(1), 53–63. <https://doi.org/10.1002/mds.26924>
- Anderson, R. W., Kehnemouyi, Y. M., Neuville, R. S., Wilkins, K. B., Anidi, C. M., Petrucci, M. N., Parker, J. E., Velisar, A., & Brontë-Stewart, H. M. (2020). A novel method for calculating beta band burst durations in Parkinson’s disease using a physiological baseline. *Journal of Neuroscience Methods*, 343, 108811. <https://doi.org/10.1016/j.jneumeth.2020.108811>
- Androulidakis, A. G., Kühn, A. A., Chu Chen, C., Blomstedt, P., Kempf, F., Kupsch, A., Schneider, G.-H., Doyle, L., Dowsey-Limousin, P., Hariz, M. I., & Brown, P. (2007). Dopaminergic therapy promotes lateralized motor activity in the subthalamic area in Parkinson’s disease. *Brain*, 130(2), 457–468. <https://doi.org/10.1093/brain/awl358>
- Anidi, C., O’Day, J. J., Anderson, R. W., Afzal, M. F., Syrkin-Nikolau, J., Velisar, A., & Bronte-Stewart, H. M. (2018). Neuromodulation targets pathological not physiological beta bursts

during gait in Parkinson's disease. *Neurobiology of Disease*, 120, 107–117.

<https://doi.org/10.1016/j.nbd.2018.09.004>

Ansó, J., Benjaber, M., Parks, B., Parker, S., Oehr, C. R., Petrucci, M., Gilron, R., Little, S., Wilt, R., Bronte-Stewart, H., Gunduz, A., Borton, D., Starr, P. A., & Denison, T. (2022).

Concurrent stimulation and sensing in bi-directional brain interfaces: A multi-site translational experience. *Journal of Neural Engineering*, 19(2), 026025.

<https://doi.org/10.1088/1741-2552/ac59a3>

Anumanchipalli, G. K., Chartier, J., & Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature*, 568(7753), 493–498. <https://doi.org/10.1038/s41586-019-1119-1>

Anzak, A., Tan, H., Pogosyan, A., Foltynie, T., Limousin, P., Zrinzo, L., Hariz, M., Ashkan, K., Bogdanovic, M., Green, A. L., Aziz, T., & Brown, P. (2012). Subthalamic nucleus activity optimizes maximal effort motor responses in Parkinson's disease. *Brain: A Journal of Neurology*, 135(Pt 9), 2766–2778. <https://doi.org/10.1093/brain/aws183>

Arlotti, M., Marceglia, S., Foffani, G., Volkmann, J., Lozano, A. M., Moro, E., Cogiamanian, F., Prenassi, M., Bocci, T., Cortese, F., Rampini, P., Barbieri, S., & Priori, A. (2018). Eight-hours adaptive deep brain stimulation in patients with Parkinson disease. *Neurology*, 90(11). <https://doi.org/10.1212/WNL.0000000000005121>

AuYong, N., Malekmohammadi, M., Ricks-Oddie, J., & Pouratian, N. (2018). Movement-Modulation of Local Power and Phase Amplitude Coupling in Bilateral Globus Pallidus Interna in Parkinson Disease. *Frontiers in Human Neuroscience*, 12, 270. <https://doi.org/10.3389/fnhum.2018.00270>

- Ba, L. J., & Caruana, R. (2014). Do deep nets really need to be deep? *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, 2654–2662.
<https://doi.org/10.5555/2969033.2969123>
- Baldermann, J. C., Petry-Schmelzer, J. N., Schüller, T., Mahfoud, L., Brandt, G. A., Dembek, T. A., Van Der Linden, C., Krauss, J. K., Szejko, N., Müller-Vahl, K. R., Ganos, C., Al-Fatly, B., Heiden, P., Servello, D., Galbiati, T., Johnson, K. A., Butson, C. R., Okun, M. S., Andrade, P., ... Barbe, M. T. (2024). A critical role of action-related functional networks in Gilles de la Tourette syndrome. *Nature Communications*, 15(1), 10687. <https://doi.org/10.1038/s41467-024-55242-6>
- Baldermann, J. C., Schüller, T., Huys, D., Becker, I., Timmermann, L., Jessen, F., Visser-Vandewalle, V., & Kuhn, J. (2016). Deep Brain Stimulation for Tourette-Syndrome: A Systematic Review and Meta-Analysis. *Brain Stimulation*, 9(2), 296–304.
<https://doi.org/10.1016/j.brs.2015.11.005>
- Barbe, M. T., Liebhart, L., Runge, M., Pauls, K. A. M., Wojtecki, L., Schnitzler, A., Allert, N., Fink, G. R., Sturm, V., Maarouf, M., & Timmermann, L. (2011). Deep brain stimulation in the nucleus ventralis intermedius in patients with essential tremor: Habituation of tremor suppression. *Journal of Neurology*, 258(3), 434–439. <https://doi.org/10.1007/s00415-010-5773-3>
- Bartholow, R. (1874). ART. I.–experimental investigations into the functions of the human brain. *The American Journal of the Medical Sciences (1827-1924)*, 134, 305.
- Becker, J., Thies, T., Petry-Schmelzer, J. N., Dembek, T. A., Reker, P., Mücke, D., Grice, M., Visser-Vandewalle, V., Fink, G. R., & Barbe, M. T. (2020). The effects of thalamic and posterior subthalamic deep brain stimulation on speech in patients with essential tremor – A

prospective, randomized, doubleblind crossover study. *Brain and Language*, 202, 104724.

<https://doi.org/10.1016/j.bandl.2019.104724>

Bekhtereva, N. P., Bondartchuk, A. N., Smirnov, V. M., Meliutcheva, L. A., & Shandurina, A. N.

(1975). Method of Electrostimulation of the Deep Brain Structures in Treatment of Some Chronic Diseases. *Stereotactic and Functional Neurosurgery*, 37(1–3), 136–140.

<https://doi.org/10.1159/000102727>

Bekhtereva, N. P., Grachev, K. V., Orlova, A. N., & Iatsuksl, null. (1963). Utilization of multiple

electrodes implanted in the subcortical structure of the human brain for the treatment of hyperkinesia. *Zhurnal Nevropatologii I Psikhiatrii Imeni S.S. Korsakova (Moscow, Russia: 1952)*, 63, 3–8.

Benabid, A. L., Pollak, P., Louveau, A., Henry, S., & De Rougemont, J. (1987). Combined

(Thalamotomy and Stimulation) Stereotactic Surgery of the VIM Thalamic Nucleus for Bilateral Parkinson Disease. *Stereotactic and Functional Neurosurgery*, 50(1–6), 344–346.

<https://doi.org/10.1159/000100803>

Bergman, H., Wichmann, T., & DeLong, M. R. (1990). Reversal of Experimental Parkinsonism by

Lesions of the Subthalamic Nucleus. *Science*, 249(4975), 1436–1438.

<https://doi.org/10.1126/science.2402638>

Berrios, G. E. (1997). The origins of psychosurgery: Shaw, Burckhardt and Moniz. *History of*

Psychiatry, 8(29), 061–081. <https://doi.org/10.1177/0957154X9700802905>

Beudel, M., & Brown, P. (2016). Adaptive deep brain stimulation in Parkinson's disease.

Parkinsonism & Related Disorders, 22, S123–S126.

<https://doi.org/10.1016/j.parkreldis.2015.09.028>

- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1), 65–98.
- Biewald, L. (2020). *Experiment tracking with weights and biases*. <https://www.wandb.com/>
- Bijanzadeh, M., Khambhati, A. N., Desai, M., Wallace, D. L., Shafi, A., Dawes, H. E., Sturm, V. E., & Chang, E. F. (2022). Decoding naturalistic affective behaviour from spectro-spatial features in multiday human iEEG. *Nature Human Behaviour*, 6(6), Article 6.
<https://doi.org/10.1038/s41562-022-01310-0>
- Bishay, A. E., Lyons, A. T., Koester, S. W., Paulo, D. L., Liles, C., Dambrino, R. J., Feldman, M. J., Ball, T. J., Bick, S. K., Englot, D. J., & Chambliss, L. B. (2024). Global Economic Evaluation of the Reported Costs of Deep Brain Stimulation. *Stereotactic and Functional Neurosurgery*, 102(4), 257–274. <https://doi.org/10.1159/000537865>
- Black, M. H., Almabruk, T., Albrecht, M. A., Chen, N. T., Lipp, O. V., Tan, T., Bolte, S., & Girdler, S. (2018). Altered Connectivity in Autistic Adults during Complex Facial Emotion Recognition: A Study of EEG Imaginary Coherence. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2752–2755.
<https://doi.org/10.1109/EMBC.2018.8512802>
- Blomstedt, P., & Hariz, M. I. (2010). Deep brain stimulation for movement disorders before DBS for movement disorders. *Parkinsonism & Related Disorders*, 16(7), 429–433.
<https://doi.org/10.1016/j.parkreldis.2010.04.005>
- Bobin, M., Sulzer, N., Bründler, G., Staib, M., Imbach, L. L., Stieglitz, L. H., Krauss, P., Bichsel, O., Baumann, C. R., & Frühholz, S. (2024a). Direct subthalamic nucleus stimulation influences

- speech and voice quality in Parkinson's disease patients. *Brain Stimulation*, 17(1), 112–124.
<https://doi.org/10.1016/j.brs.2024.01.006>
- Bobin, M., Sulzer, N., Bründler, G., Staib, M., Imbach, L. L., Stieglitz, L. H., Krauss, P., Bichsel, O., Baumann, C. R., & Frühholz, S. (2024b). Direct subthalamic nucleus stimulation influences speech and voice quality in Parkinson's disease patients. *Brain Stimulation*, 17(1), 112–124.
<https://doi.org/10.1016/j.brs.2024.01.006>
- Boettcher, L. B., & Menacho, S. T. (2017). The early argument for prefrontal leucotomy: The collision of frontal lobe theory and psychosurgery at the 1935 International Neurological Congress in London. *Neurosurgical Focus*, 43(3), E4. <https://doi.org/10.3171/2017.6.FOCUS17249>
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural Representations and Mechanisms for the Performance of Simple Speech Sequences. *Journal of Cognitive Neuroscience*, 22(7), 1504–1529. <https://doi.org/10.1162/jocn.2009.21306>
- Bohlhalter, S. (2006). Neural correlates of tic generation in Tourette syndrome: An event-related functional MRI study. *Brain*, 129(8), 2029–2037. <https://doi.org/10.1093/brain/awl050>
- Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441), 327–332.
<https://doi.org/10.1038/nature11911>
- Bour, L. J., Ackermans, L., Foncke, E. M. J., Cath, D., Van Der Linden, C., Visser Vandewalle, V., & Tijssen, M. A. (2015). Tic related local field potentials in the thalamus and the effect of deep brain stimulation in Tourette syndrome: Report of three cases. *Clinical Neurophysiology*, 126(8), 1578–1588. <https://doi.org/10.1016/j.clinph.2014.10.217>

- Bouthour, W., Mégevand, P., Donoghue, J., Lüscher, C., Birbaumer, N., & Krack, P. (2019). Biomarkers for closed-loop deep brain stimulation in Parkinson disease and beyond. *Nature Reviews Neurology*, 15(6), 343–352. <https://doi.org/10.1038/s41582-019-0166-4>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Broca, P. (1861). Nouvelle observation d’aphémie produite par une lésion de la moitié postérieure des deuxième et troisième circonvolutions frontales. *Bulletins de la Société Anatomique de Paris*, 36, 398–407.
- Brodersen, K. H., Ong, C. S., Stephan, K. E., & Buhmann, J. M. (2010). The Balanced Accuracy and Its Posterior Distribution. *2010 20th International Conference on Pattern Recognition*, 3121–3124. <https://doi.org/10.1109/ICPR.2010.764>
- Bronte-Stewart, H. M., Petrucci, M. N., O’Day, J. J., Afzal, M. F., Parker, J. E., Kehnemouyi, Y. M., Wilkins, K. B., Orthlieb, G. C., & Hoffman, S. L. (2020). Perspective: Evolution of Control Variables and Policies for Closed-Loop Deep Brain Stimulation for Parkinson’s Disease Using Bidirectional Deep-Brain-Computer Interfaces. *Frontiers in Human Neuroscience*, 14, 353. <https://doi.org/10.3389/fnhum.2020.00353>
- Brown, P., & Marsden, C. (1998). What do the basal ganglia do? *The Lancet*, 351(9118), 1801–1804. [https://doi.org/10.1016/S0140-6736\(97\)11225-9](https://doi.org/10.1016/S0140-6736(97)11225-9)
- Brunenberg, E. J. L., Moeskops, P., Backes, W. H., Pollo, C., Cammoun, L., Vilanova, A., Janssen, M. L. F., Visser-Vandewalle, V. E. R. M., ter Haar Romeny, B. M., Thiran, J.-P., & Platel, B. (2012). Structural and resting state functional connectivity of the subthalamic nucleus: Identification of motor STN parts and the hyperdirect pathway. *PloS One*, 7(6), e39061. <https://doi.org/10.1371/journal.pone.0039061>

- Buijink, A. W. G., Piña-Fuentes, D. A., Stam, M. J., Bot, M., Schuurman, P. R., van den Munckhof, P., van Rootselaar, A. F., de Bie, R. M. A., & Beudel, M. (2022). Thalamic local field potentials recorded using the deep brain stimulation pulse generator. *Clinical Neurophysiology Practice*, 7, 103–106. <https://doi.org/10.1016/j.cnp.2022.03.002>
- Busch, J. L., Kaplan, J., Habets, J. G. V., Feldmann, L. K., Roediger, J., Köhler, R. M., Merk, T., Faust, K., Schneider, G.-H., Bergman, H., Neumann, W.-J., & Kühn, A. A. (2024). Single threshold adaptive deep brain stimulation in Parkinson's disease depends on parameter selection, movement state and controllability of subthalamic beta activity. *Brain Stimulation*, 17(1), 125–133. <https://doi.org/10.1016/j.brs.2024.01.007>
- Cagle, J. N., Deeb, W., Eisinger, R. S., Molina, R., Opri, E., Holland, M. T., Foote, K. D., Okun, M. S., & Gunduz, A. (2020). Lead Repositioning Guided by Both Physiology and Atlas Based Targeting in Tourette Deep Brain Stimulation. *Tremor and Other Hyperkinetic Movements*, 10(1), 18. <https://doi.org/10.5334/tohm.140>
- Cagle, J. N., Okun, M. S., Cernera, S., Eisinger, R. S., Opri, E., Bowers, D., Ward, H., Foote, K. D., & Gunduz, A. (2022). Embedded Human Closed-Loop Deep Brain Stimulation for Tourette Syndrome: A Nonrandomized Controlled Trial. *JAMA Neurology*, 79(10), 1064. <https://doi.org/10.1001/jamaneurol.2022.2741>
- Cai, W., Young, C. B., Yuan, R., Lee, B., Ryman, S., Kim, J., Yang, L., Levine, T. F., Henderson, V. W., Poston, K. L., & Menon, V. (2024). Subthalamic nucleus–language network connectivity predicts dopaminergic modulation of speech function in Parkinson's disease. *Proceedings of the National Academy of Sciences*, 121(22), e2316149121. <https://doi.org/10.1073/pnas.2316149121>

- Caligiore, D., Mannella, F., Arbib, M. A., & Baldassarre, G. (2017). Dysfunctions of the basal ganglia-cerebellar-thalamo-cortical system produce motor tics in Tourette syndrome. *PLOS Computational Biology*, *13*(3), e1005395. <https://doi.org/10.1371/journal.pcbi.1005395>
- Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2006a). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science (New York, N.Y.)*, *313*(5793), 1626–1628. <https://doi.org/10.1126/science.1128115>
- Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2006b). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science*, *313*(5793), 1626–1628. <https://doi.org/10.1126/science.1128115>
- Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, *14*(11), 506–515. <https://doi.org/10.1016/j.tics.2010.09.001>
- Cao, C., Litvak, V., Zhan, S., Liu, W., Zhang, C., Sun, B., Li, D., & van Wijk, B. C. M. (2024). Low-beta versus high-beta band cortico-subcortical coherence in movement inhibition and expectation. *Neurobiology of Disease*, *201*, 106689. <https://doi.org/10.1016/j.nbd.2024.106689>
- Card, N. S., Wairagkar, M., Iacobacci, C., Hou, X., Singer-Clark, T., Willett, F. R., Kunz, E. M., Fan, C., Vahdati Nia, M., Deo, D. R., Srinivasan, A., Choi, E. Y., Glasser, M. F., Hochberg, L. R., Henderson, J. M., Shahlaie, K., Stavisky, S. D., & Brandman, D. M. (2024). An Accurate and Rapidly Calibrating Speech Neuroprosthesis. *New England Journal of Medicine*, *391*(7), 609–618. <https://doi.org/10.1056/NEJMoa2314132>

- Casagrande, S. C., Cury, R. G., Alho, E. J., & Fonoff, E. T. (2019). Deep brain stimulation in Tourette's syndrome: Evidence to date. *Neuropsychiatric Disease and Treatment, Volume 15*, 1061–1075. <https://doi.org/10.2147/NDT.S139368>
- Castañó-Candamil, S., Ferleger, B. I., Haddock, A., Cooper, S. S., Herron, J., Ko, A., Chizeck, Howard. J., & Tangermann, M. (2020). A Pilot Study on Data-Driven Adaptive Deep Brain Stimulation in Chronically Implanted Essential Tremor Patients. *Frontiers in Human Neuroscience, 14*, 541625. <https://doi.org/10.3389/fnhum.2020.541625>
- Castrioto, A. (2011). Ten-Year Outcome of Subthalamic Stimulation in Parkinson Disease: A Blinded Evaluation. *Archives of Neurology, 68*(12), 1550. <https://doi.org/10.1001/archneurol.2011.182>
- Cernera, S., Alcantara, J. D., Opri, E., Cagle, J. N., Eisinger, R. S., Boogaart, Z., Pramanik, L., Kelberman, M., Patel, B., Foote, K. D., Okun, M. S., & Gunduz, A. (2021). Wearable sensor-driven responsive deep brain stimulation for essential tremor. *Brain Stimulation, 14*(6), 1434–1443. <https://doi.org/10.1016/j.brs.2021.09.002>
- Cernera, S., Long, S., Kelberman, M., Hegland, K. W., Hicks, J., Smith-Hublou, M., Taylor, B., Mou, Y., De Hemptinne, C., Johnson, K. A., Cagle, J. N., Moore, K., Foote, K. D., Okun, M. S., & Gunduz, A. (2024). Responsive Versus Continuous Deep Brain Stimulation for Speech in Essential Tremor: A Pilot Study. *Movement Disorders*, mds.29865. <https://doi.org/10.1002/mds.29865>
- Chattopadhyay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 839–847. <https://doi.org/10.1109/WACV.2018.00097>

- Chen, C. C., Brücke, C., Kempf, F., Kupsch, A., Lu, C. S., Lee, S. T., Tisch, S., Limousin, P., Hariz, M., & Brown, P. (2006). Deep brain stimulation of the subthalamic nucleus: A two-edged sword. *Current Biology*, 16(22), R952–R953. <https://doi.org/10.1016/j.cub.2006.10.013>
- Chen, C. C., Pogosyan, A., Zrinzo, L. U., Tisch, S., Limousin, P., Ashkan, K., Yousry, T., Hariz, M. I., & Brown, P. (2006a). Intra-operative recordings of local field potentials can help localize the subthalamic nucleus in Parkinson’s disease surgery. *Experimental Neurology*, 198(1), 214–221. <https://doi.org/10.1016/j.expneurol.2005.11.019>
- Chen, C. C., Pogosyan, A., Zrinzo, L. U., Tisch, S., Limousin, P., Ashkan, K., Yousry, T., Hariz, M. I., & Brown, P. (2006b). Intra-operative recordings of local field potentials can help localize the subthalamic nucleus in Parkinson’s disease surgery. *Experimental Neurology*, 198(1), 214–221. <https://doi.org/10.1016/j.expneurol.2005.11.019>
- Chen, P., Kim, T., Dastin-van Rijn, E., Provenza, N. R., Sheth, S. A., Goodman, W. K., Borton, D. A., Harrison, M. T., & Darbon, J. (2022). Periodic Artifact Removal With Applications to Deep Brain Stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30, 2692–2699. <https://doi.org/10.1109/TNSRE.2022.3205453>
- Chen, W., De Hemptinne, C., Miller, A. M., Leibbrand, M., Little, S. J., Lim, D. A., Larson, P. S., & Starr, P. A. (2020). Prefrontal-Subthalamic Hyperdirect Pathway Modulates Movement Inhibition in Humans. *Neuron*, 106(4), 579–588.e3. <https://doi.org/10.1016/j.neuron.2020.02.012>
- Chen, W., Kirkby, L., Kotzev, M., Song, P., Gilron, R., & Pepin, B. (2021). The Role of Large-Scale Data Infrastructure in Developing Next-Generation Deep Brain Stimulation Therapies. *Frontiers in Human Neuroscience*, 15, 717401. <https://doi.org/10.3389/fnhum.2021.717401>

Cheyne, D. O., & Ferrari, P. (2013). MEG studies of motor cortex gamma oscillations: Evidence for a gamma “fingerprint” in the brain? *Frontiers in Human Neuroscience*, 7.

<https://doi.org/10.3389/fnhum.2013.00575>

Chrabaszcz, A., Neumann, W.-J., Stretcu, O., Lipski, W. J., Bush, A., Dastolfo-Hromack, C. A., Wang, D., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2019a). Subthalamic Nucleus and Sensorimotor Cortex Activity During Speech Production. *The Journal of Neuroscience*, 39(14), 2698–2708.

<https://doi.org/10.1523/JNEUROSCI.2842-18.2019>

Chrabaszcz, A., Neumann, W.-J., Stretcu, O., Lipski, W. J., Bush, A., Dastolfo-Hromack, C. A., Wang, D., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2019b). Subthalamic Nucleus and Sensorimotor Cortex Activity During Speech Production. *The Journal of Neuroscience*, 39(14), 2698–2708.

<https://doi.org/10.1523/JNEUROSCI.2842-18.2019>

Chrabaszcz, A., Wang, D., Lipski, W. J., Bush, A., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2021a). Simultaneously recorded subthalamic and cortical LFPs reveal different lexicality effects during reading aloud. *Journal of Neurolinguistics*, 60, 101019. <https://doi.org/10.1016/j.jneuroling.2021.101019>

Chrabaszcz, A., Wang, D., Lipski, W. J., Bush, A., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2021b). Simultaneously recorded subthalamic and cortical LFPs reveal different lexicality effects during reading aloud. *Journal of Neurolinguistics*, 60, 101019. <https://doi.org/10.1016/j.jneuroling.2021.101019>

- Clausen, J., Fetz, E., Donoghue, J., Ushiba, J., Spörhase, U., Chandler, J., Birbaumer, N., & Soekadar, S. R. (2017). Help, hope, and hype: Ethical dimensions of Neuroprosthetics. *Science*, *356*(6345), 1338–1339. <https://doi.org/10.1126/science.aam7731>
- Cole, S. R., van der Meij, R., Peterson, E. J., de Hemptinne, C., Starr, P. A., & Voytek, B. (2017). Nonsinusoidal Beta Oscillations Reflect Cortical Pathophysiology in Parkinson’s Disease. *The Journal of Neuroscience*, *37*(18), 4830–4840. <https://doi.org/10.1523/JNEUROSCI.2208-16.2017>
- Cramb, K. M. L., Beccano-Kelly, D., Cragg, S. J., & Wade-Martins, R. (2023). Impaired dopamine release in Parkinson’s disease. *Brain*, *146*(8), 3117–3132. <https://doi.org/10.1093/brain/awad064>
- Crone, N. E., Korzeniewska, A., & Franaszczuk, P. J. (2011). Cortical gamma responses: Searching high and low. *International Journal of Psychophysiology*, *79*(1), 9–15. <https://doi.org/10.1016/j.ijpsycho.2010.10.013>
- Crone, N. E., Miglioretti, D. L., Gordon, B., & Lesser, R. P. (1998). Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. *Brain: A Journal of Neurology*, *121* (Pt 12), 2301–2315. <https://doi.org/10.1093/brain/121.12.2301>
- Dallapiazza, R. F., Lee, D. J., De Vloo, P., Fomenko, A., Hamani, C., Hodaie, M., Kalia, S. K., Fasano, A., & Lozano, A. M. (2019). Outcomes from stereotactic surgery for essential tremor. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(4), 474–482. <https://doi.org/10.1136/jnnp-2018-318240>

- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential Diagnostic Patterns of Dysarthria. *Journal of Speech and Hearing Research*, 12(2), 246–269.
<https://doi.org/10.1044/jshr.1202.246>
- Dastolfo-Hromack, C., Bush, A., Chrabaszc, A., Alhourani, A., Lipski, W., Wang, D., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2022). Articulatory Gain Predicts Motor Cortex and Subthalamic Nucleus Activity During Speech. *Cerebral Cortex*, 32(7), 1337–1349. <https://doi.org/10.1093/cercor/bhab251>
- Dauwels, J., Vialatte, F., Musha, T., & Cichocki, A. (2010). A comparative study of synchrony measures for the early diagnosis of Alzheimer’s disease based on EEG. *NeuroImage*, 49(1), 668–693. <https://doi.org/10.1016/j.neuroimage.2009.06.056>
- de Hemptinne, C., Ryapolova-Webb, E. S., Air, E. L., Garcia, P. A., Miller, K. J., Ojemann, J. G., Ostrem, J. L., Galifianakis, N. B., & Starr, P. A. (2013). Exaggerated phase-amplitude coupling in the primary motor cortex in Parkinson disease. *Proceedings of the National Academy of Sciences of the United States of America*, 110(12), 4780–4785.
<https://doi.org/10.1073/pnas.1214546110>
- Debarros, J., Gaignon, L., He, S., Pogosyan, A., Benjaber, M., Denison, T., Brown, P., & Tan, H. (2020). Artefact-free recording of local field potentials with simultaneous stimulation for closed-loop Deep-Brain Stimulation. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 3367–3370.
<https://doi.org/10.1109/EMBC44109.2020.9176665>
- Debru, A. (2006). The power of torpedo fish as a pathological model to the understanding of nervous transmission in Antiquity. *Comptes Rendus. Biologies*, 329(5–6), 298–302.
<https://doi.org/10.1016/j.crv.2006.03.001>

- Delgado, J. M. (1969). *Physical Control of the Mind*. Harper New York.
- Dembek, T. A., Reker, P., Visser-Vandewalle, V., Wirths, J., Treuer, H., Klehr, M., Roediger, J., Dafsari, H. S., Barbe, M. T., & Timmermann, L. (2017). Directional DBS increases side-effect thresholds—A prospective, double-blind trial. *Movement Disorders*, 32(10), 1380–1388. <https://doi.org/10.1002/mds.27093>
- Dembek, T. A., Roediger, J., Horn, A., Reker, P., Oehr, C., Dafsari, H. S., Li, N., Kühn, A. A., Fink, G. R., Visser-Vandewalle, V., Barbe, M. T., & Timmermann, L. (2019). Probabilistic sweet spots predict motor outcome for deep brain stimulation in Parkinson disease. *Annals of Neurology*, 86(4), 527–538. <https://doi.org/10.1002/ana.25567>
- Denys, D., Mantione, M., Figee, M., Van Den Munckhof, P., Koerselman, F., Westenberg, H., Bosch, A., & Schuurman, R. (2010). Deep Brain Stimulation of the Nucleus Accumbens for Treatment-Refractory Obsessive-Compulsive Disorder. *Archives of General Psychiatry*, 67(10), 1061. <https://doi.org/10.1001/archgenpsychiatry.2010.122>
- Desai, S., & Ramaswamy, H. G. (2020). Ablation-CAM: Visual Explanations for Deep Convolutional Network via Gradient-free Localization. *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 972–980. <https://doi.org/10.1109/WACV45572.2020.9093360>
- Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of Choice-Predictive Activity in Human Motor Cortex during Perceptual Decision Making. *Current Biology*, 19(18), 1581–1585. <https://doi.org/10.1016/j.cub.2009.07.066>
- Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130(5), 1432–1441. <https://doi.org/10.1093/brain/awm042>

- Engel, K., Huckhagel, T., Gulberti, A., Pötter-Nerger, M., Vettorazzi, E., Hidding, U., Choe, C., Zittel, S., Braaß, H., Ludewig, P., Schaper, M., Krajewski, K., Oehlwein, C., Mittmann, K., Engel, A. K., Gerloff, C., Westphal, M., Moll, C. K. E., Buhmann, C., ... Hamel, W. (2018). Towards unambiguous reporting of complications related to deep brain stimulation surgery: A retrospective single-center analysis and systematic review of the literature. *PLOS ONE*, *13*(8), e0198529. <https://doi.org/10.1371/journal.pone.0198529>
- Fabbri, M., Natale, F., Artusi, C. A., Romagnolo, A., Bozzali, M., Giulietti, G., Guimaraes, I., Rizzone, M. G., Accornero, A., Lopiano, L., & Zibetti, M. (2021). Deep brain stimulation fine-tuning in Parkinson's disease: Short pulse width effect on speech. *Parkinsonism & Related Disorders*, *87*, 130–134. <https://doi.org/10.1016/j.parkreldis.2021.05.007>
- Fabbri, M., Zibetti, M., Ferrero, G., Accornero, A., Guimaraes, I., Rizzone, M. G., Romagnolo, A., Ferreira, J. J., & Lopiano, L. (2019). Is lowering stimulation frequency a feasible option for subthalamic deep brain stimulation in Parkinson's disease patients with dysarthria? *Parkinsonism & Related Disorders*, *64*, 242–248. <https://doi.org/10.1016/j.parkreldis.2019.04.018>
- Farokhniaee, A., Palmisano, C., Del Vecchio Del Vecchio, J., Pezzoli, G., Volkmann, J., & Isaias, I. U. (2024). Gait-related beta-gamma phase amplitude coupling in the subthalamic nucleus of parkinsonian patients. *Scientific Reports*, *14*(1), 6674. <https://doi.org/10.1038/s41598-024-57252-2>
- Fasano, A., & Helmich, R. C. (2019). Tremor habituation to deep brain stimulation: Underlying mechanisms and solutions. *Movement Disorders*, *34*(12), 1761–1773. <https://doi.org/10.1002/mds.27821>

- Fasano, A., Romito, L. M., Daniele, A., Piano, C., Zinno, M., Bentivoglio, A. R., & Albanese, A. (2010). Motor and cognitive outcome in patients with Parkinson's disease 8 years after subthalamic implants. *Brain*, 133(9), 2664–2676. <https://doi.org/10.1093/brain/awq221>
- Feldmann, L. K., Lofredi, R., Al-Fatly, B., Busch, J. L., Mathiopoulou, V., Roediger, J., Krause, P., Schneider, G., Faust, K., Horn, A., Kühn, A. A., & Neumann, W. (2023). Christmas-Related Reduction in Beta Activity in Parkinson's Disease. *Movement Disorders*, 38(4), 692–697. <https://doi.org/10.1002/mds.29334>
- Feldmann, L. K., Lofredi, R., Neumann, W.-J., Al-Fatly, B., Roediger, J., Bahners, B. H., Nikolov, P., Denison, T., Saryyeva, A., Krauss, J. K., Faust, K., Florin, E., Schnitzler, A., Schneider, G.-H., & Kühn, A. A. (2022). Toward therapeutic electrophysiology: Beta-band suppression as a biomarker in chronic local field potential recordings. *Npj Parkinson's Disease*, 8(1), 44. <https://doi.org/10.1038/s41531-022-00301-2>
- Feldmann, L. K., Neumann, W., Krause, P., Lofredi, R., Schneider, G., & Kühn, A. A. (2021). Subthalamic beta band suppression reflects effective neuromodulation in chronic recordings. *European Journal of Neurology*, 28(7), 2372–2377. <https://doi.org/10.1111/ene.14801>
- Feldmann, L. K., Neumann, W.-J., Faust, K., Schneider, G.-H., & Kühn, A. A. (2021). Risk of Infection after Deep Brain Stimulation Surgery with Externalization and Local-Field Potential Recordings: Twelve-Year Experience from a Single Institution. *Stereotactic and Functional Neurosurgery*, 99(6), 512–520. <https://doi.org/10.1159/000516150>
- Ferleger, B. I., Houston, B., Thompson, M. C., Cooper, S. S., Sonnet, K. S., Ko, A. L., Herron, J. A., & Chizeck, H. J. (2020). Fully implanted adaptive deep brain stimulation in freely moving essential tremor patients. *Journal of Neural Engineering*, 17(5), 056026. <https://doi.org/10.1088/1741-2552/abb416>

- Ferreira, J. J., Mestre, T. A., Lyons, K. E., Benito-León, J., Tan, E., Abbruzzese, G., Hallett, M., Haubenberger, D., Elble, R., Deuschl, G., & on behalf of MDS Task Force on Tremor and the MDS Evidence Based Medicine Committee. (2019). MDS evidence-based review of treatments for essential tremor. *Movement Disorders*, *34*(7), 950–958. <https://doi.org/10.1002/mds.27700>
- Ferrier, D. (1873). Experimental researches in cerebral physiology and pathology. *British Medical Journal*, *1*(643), 457.
- Fischer, P., Lipski, W. J., Neumann, W.-J., Turner, R. S., Fries, P., Brown, P., & Richardson, R. M. (2020). Movement-related coupling of human subthalamic nucleus spikes to cortical gamma. *eLife*, *9*, e51956. <https://doi.org/10.7554/eLife.51956>
- Fischer, P., Pogosyan, A., Herz, D. M., Cheeran, B., Green, A. L., Fitzgerald, J., Aziz, T. Z., Hyam, J., Little, S., Foltynie, T., Limousin, P., Zrinzo, L., Brown, P., & Tan, H. (2017). Subthalamic nucleus gamma activity increases not only during movement but also during movement inhibition. *eLife*, *6*, e23947. <https://doi.org/10.7554/eLife.23947>
- Fleming, J. E., Kremen, V., Gilron, R., Gregg, N. M., Zamora, M., Dijk, D.-J., Starr, P. A., Worrell, G. A., Little, S., & Denison, T. J. (2022). Embedding digital chronotherapy into bioelectronic medicines. *iScience*, *25*(4), 104028. <https://doi.org/10.1016/j.isci.2022.104028>
- Follett, K. A., Weaver, F. M., Stern, M., Hur, K., Harris, C. L., Luo, P., Marks, W. J., Rothlind, J., Sagher, O., Moy, C., Pahwa, R., Burchiel, K., Hogarth, P., Lai, E. C., Duda, J. E., Holloway, K., Samii, A., Horn, S., Bronstein, J. M., ... Reda, D. J. (2010). Pallidal versus Subthalamic Deep-Brain Stimulation for Parkinson's Disease. *New England Journal of Medicine*, *362*(22), 2077–2091. <https://doi.org/10.1056/NEJMoa0907083>

- Fox, M. D., Buckner, R. L., Liu, H., Chakravarty, M. M., Lozano, A. M., & Pascual-Leone, A. (2014). Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. *Proceedings of the National Academy of Sciences*, *111*(41). <https://doi.org/10.1073/pnas.1405003111>
- Frączek, T. M., Ferleger, B. I., Brown, T. E., Thompson, M. C., Haddock, A. J., Houston, B. C., Ojemann, J. G., Ko, A. L., Herron, J. A., & Chizeck, H. J. (2021). Closing the Loop With Cortical Sensing: The Development of Adaptive Deep Brain Stimulation for Essential Tremor Using the Alevia PC+S. *Frontiers in Neuroscience*, *15*, 749705. <https://doi.org/10.3389/fnins.2021.749705>
- Frey, J., Cagle, J., Johnson, K. A., Wong, J. K., Hilliard, J. D., Butson, C. R., Okun, M. S., & de Hemptinne, C. (2022). Past, Present, and Future of Deep Brain Stimulation: Hardware, Software, Imaging, Physiology and Novel Approaches. *Frontiers in Neurology*, *13*, 825178. <https://doi.org/10.3389/fneur.2022.825178>
- Fritsch, G. T. (1870). Über die elektrische Erregbarkeit des Grosshirns. *Arch Anat Physiol*, *37*, 300–332.
- Fujisawa, S., & Buzsáki, G. (2011). A 4 Hz Oscillation Adaptively Synchronizes Prefrontal, VTA, and Hippocampal Activities. *Neuron*, *72*(1), 153–165. <https://doi.org/10.1016/j.neuron.2011.08.018>
- Gao, W., Yu, T., Yu, J.-G., Gu, Z., Li, K., Huang, Y., Yu, Z. L., & Li, Y. (2021). Learning Invariant Patterns Based on a Convolutional Neural Network and Big Electroencephalography Data for Subject-Independent P300 Brain-Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *29*, 1047–1057. <https://doi.org/10.1109/TNSRE.2021.3083548>

- Geng, X., Xu, X., Horn, A., Li, N., Ling, Z., Brown, P., & Wang, S. (2018). Intra-operative characterisation of subthalamic oscillations in Parkinson's disease. *Clinical Neurophysiology*, 129(5), 1001–1010. <https://doi.org/10.1016/j.clinph.2018.01.075>
- Germann, J., Gouveia, F. V., Beyn, M. E., Elias, G. J. B., & Lozano, A. M. (2024). Computational Neurosurgery in Deep Brain Stimulation. In A. Di Ieva, E. Suero Molina, S. Liu, & C. Russo (Eds.), *Computational Neurosurgery* (Vol. 1462, pp. 435–451). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64892-2_26
- Giannicola, G., Rosa, M., Servello, D., Menghetti, C., Carrabba, G., Pacchetti, C., Zangaglia, R., Cogiamanian, F., Scelzo, E., Marceglia, S., Rossi, L., & Priori, A. (2012). Subthalamic local field potentials after seven-year deep brain stimulation in Parkinson's disease. *Experimental Neurology*, 237(2), 312–317. <https://doi.org/10.1016/j.expneurol.2012.06.012>
- Gildenberg, P. L. (2005). Evolution of Neuromodulation. *Stereotactic and Functional Neurosurgery*, 83(2–3), 71–79. <https://doi.org/10.1159/000086865>
- Gildenberg, P. L. (2006). Evolution of Basal Ganglia Surgery for Movement Disorders. *Stereotactic and Functional Neurosurgery*, 84(4), 131–135. <https://doi.org/10.1159/000094844>
- Gilron, R., Little, S., Perrone, R., Wilt, R., De Hemptinne, C., Yaroshinsky, M. S., Racine, C. A., Wang, S. S., Ostrem, J. L., Larson, P. S., Wang, D. D., Galifianakis, N. B., Bledsoe, I. O., San Luciano, M., Dawes, H. E., Worrell, G. A., Kremen, V., Borton, D. A., Denison, T., & Starr, P. A. (2021). Long-term wireless streaming of neural recordings for circuit discovery and adaptive stimulation in individuals with Parkinson's disease. *Nature Biotechnology*, 39(9), 1078–1085. <https://doi.org/10.1038/s41587-021-00897-5>

- Giordano, J. (2015). A Preparatory Neuroethical Approach to Assessing Developments in Neurotechnology. *AMA Journal of Ethics*, 17(1), 56–61.
<https://doi.org/10.1001/virtualmentor.2015.17.1.msoc1-1501>
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517.
<https://doi.org/10.1038/nn.3063>
- Glaser, J. I., Benjamin, A. S., Chowdhury, R. H., Perich, M. G., Miller, L. E., & Kording, K. P. (2020). Machine Learning for Neural Decoding. *Eneuro*, 7(4), ENEURO.0506-19.2020.
<https://doi.org/10.1523/ENEURO.0506-19.2020>
- Golshan, H. M., Hebb, A. O., & Mahoor, M. H. (2020). LFP-Net: A deep learning framework to recognize human behavioral activities using brain STN-LFP signals. *Journal of Neuroscience Methods*, 335, 108621. <https://doi.org/10.1016/j.jneumeth.2020.108621>
- Gong, R., Mühlberg, C., Wegscheider, M., Fricke, C., Rumpf, J.-J., Knösche, T. R., & Classen, J. (2022). Cross-frequency phase-amplitude coupling in repetitive movements in patients with Parkinson's disease. *Journal of Neurophysiology*, 127(6), 1606–1621.
<https://doi.org/10.1152/jn.00541.2021>
- Gong, R., Wegscheider, M., Mühlberg, C., Gast, R., Fricke, C., Rumpf, J.-J., Nikulin, V. V., Knösche, T. R., & Classen, J. (2021). Spatiotemporal features of β - γ phase-amplitude coupling in Parkinson's disease derived from scalp EEG. *Brain: A Journal of Neurology*, 144(2), 487–503.
<https://doi.org/10.1093/brain/awaa400>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

- Graupe, D., Basu, I., Tuninetti, D., Vannemreddy, P., & Slavin, K. V. (2010). Adaptively controlling deep brain stimulation in essential tremor patient via surface electromyography. *Neurological Research*, 32(9), 899–904. <https://doi.org/10.1179/016164110X12767786356354>
- Greenberg, B. D., Malone, D. A., Friehs, G. M., Rezai, A. R., Kubu, C. S., Malloy, P. F., Salloway, S. P., Okun, M. S., Goodman, W. K., & Rasmussen, S. A. (2006). Three-Year Outcomes in Deep Brain Stimulation for Highly Resistant Obsessive–Compulsive Disorder. *Neuropsychopharmacology*, 31(11), 2384–2393. <https://doi.org/10.1038/sj.npp.1301165>
- Groppa, S., Gonzalez-Escamilla, G., Tinkhauser, G., Baqapuri, H. I., Sajonz, B., Wiest, C., Pereira, J., Herz, D. M., Dold, M. R., Bange, M., Ciolac, D., Almeida, V., Neuber, J., Mirzac, D., Martín-Rodríguez, J. F., Dresel, C., Muthuraman, M., Adarmes Gomez, A. D., Navas, M., ... on behalf of the EuroDBS. (2023). Perspectives of Implementation of Closed-Loop Deep Brain Stimulation: From Neurological to Psychiatric Disorders. *Stereotactic and Functional Neurosurgery*, 1–15. <https://doi.org/10.1159/000535114>
- Gross, C. G. (2007). The discovery of motor cortex and its background. *Journal of the History of the Neurosciences*, 16(3), 320–331.
- Grover, T., Georgiev, D., Kalliola, R., Mahlknecht, P., Zacharia, A., Candelario, J., Hyam, J., Zrinzo, L., Hariz, M., Foltynie, T., Limousin, P., Jahanshahi, M., & Tripoliti, E. (2019). Effect of Low versus High Frequency Subthalamic Deep Brain Stimulation on Speech Intelligibility and Verbal Fluency in Parkinson’s Disease: A Double-Blind Study. *Journal of Parkinson’s Disease*, 9(1), 141–151. <https://doi.org/10.3233/JPD-181368>
- Guidetti, M., Marceglia, S., Loh, A., Harmsen, I. E., Meoni, S., Foffani, G., Lozano, A. M., Moro, E., Volkmann, J., & Priori, A. (2021). Clinical perspectives of adaptive deep brain stimulation. *Brain Stimulation*, 14(5), 1238–1247. <https://doi.org/10.1016/j.brs.2021.07.063>

- Haarmann, L., Kalbe, E., Lewis, C. J., Eggers, C., Kühn, A. A., Krug, H., Volkmann, J., Kirsch, A. D., Wojtecki, L., Schnitzler, A., Deuschl, G., Krauss, J. K., Woopen, C., Timmermann, L., & Maier, F. (2022). The Deep Brain Stimulation Impairment Scale: A useful complement in assessment of well-being and functioning in DBS-patients – Results from a large multicentre survey in patients with Parkinson’s disease. *Parkinsonism & Related Disorders*, *99*, 8–15. <https://doi.org/10.1016/j.parkreldis.2022.04.008>
- Hag, A., Al-Shargie, F., Handayani, D., & Asadi, H. (2023). Mental Stress Classification Based on Selected Electroencephalography Channels Using Correlation Coefficient of Hjorth Parameters. *Brain Sciences*, *13*(9), 1340. <https://doi.org/10.3390/brainsci13091340>
- 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>
- Hammer, M. J., Barlow, S. M., Lyons, K. E., & Pahwa, R. (2010). Subthalamic nucleus deep brain stimulation changes speech respiratory and laryngeal control in Parkinson’s disease. *Journal of Neurology*, *257*(10), 1692–1702. <https://doi.org/10.1007/s00415-010-5605-5>
- Hariz, M. (2012). Twenty-five years of deep brain stimulation: Celebrations and apprehensions. *Movement Disorders*, *27*(7), 930–933. <https://doi.org/10.1002/mds.25007>
- Hashimoto, T., Elder, C. M., & Vitek, J. L. (2002). A template subtraction method for stimulus artifact removal in high-frequency deep brain stimulation. *Journal of Neuroscience Methods*, *113*(2), 181–186. [https://doi.org/10.1016/S0165-0270\(01\)00491-5](https://doi.org/10.1016/S0165-0270(01)00491-5)

- He, S., Baig, F., Mostofi, A., Pogosyan, A., Debarros, J., Green, A. L., Aziz, T. Z., Pereira, E., Brown, P., & Tan, H. (2021). Closed-Loop Deep Brain Stimulation for Essential Tremor Based on Thalamic Local Field Potentials. *Movement Disorders*, *36*(4), 863–873.
<https://doi.org/10.1002/mds.28513>
- Hebb, A. O., Darvas, F., & Miller, K. J. (2012a). Transient and state modulation of beta power in human subthalamic nucleus during speech production and finger movement. *Neuroscience*, *202*, 218–233. <https://doi.org/10.1016/j.neuroscience.2011.11.072>
- Hebb, A. O., Darvas, F., & Miller, K. J. (2012b). Transient and State Modulation of Beta Power in Human Subthalamic Nucleus during Speech Production and Finger Movement. *Neuroscience*, *202*, 218–233. <https://doi.org/10.1016/j.neuroscience.2011.11.072>
- Hell, F., Eißner, A., Mehrkens, J. H., & Bötzel, K. (2023). Subthalamic oscillatory activity during normal and impaired speech. *Clinical Neurophysiology*, *149*, 42–50.
<https://doi.org/10.1016/j.clinph.2023.02.166>
- Henderson, J. M. (2008). Peripheral nerve stimulation for chronic pain. *Current Pain and Headache Reports*, *12*(1), 28–31. <https://doi.org/10.1007/s11916-008-0006-5>
- Herron, J. A., Thompson, M. C., Brown, T., Chizeck, H. J., Ojemann, J. G., & Ko, A. L. (2017). Chronic electrocorticography for sensing movement intention and closed-loop deep brain stimulation with wearable sensors in an essential tremor patient. *Journal of Neurosurgery*, *127*(3), 580–587. <https://doi.org/10.3171/2016.8.JNS16536>
- Herron, J., Denison, T., & Chizeck, H. J. (2015). Closed-loop DBS with movement intention. *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, 844–847.
<https://doi.org/10.1109/NER.2015.7146755>

- Herron, J., Kremen, V., Simeral, J. D., Dawes, H., Worrell, G. A., Starr, P. A., Denison, T., & Borton, D. (2024). The convergence of neuromodulation and brain–computer interfaces. *Nature Reviews Bioengineering*. <https://doi.org/10.1038/s44222-024-00187-0>
- Herron, J., Kullmann, A., Denison, T., Goodman, W. K., Gunduz, A., Neumann, W.-J., Provenza, N. R., Shanechi, M. M., Sheth, S. A., Starr, P. A., & Widge, A. S. (2024). Challenges and opportunities of acquiring cortical recordings for chronic adaptive deep brain stimulation. *Nature Biomedical Engineering*. <https://doi.org/10.1038/s41551-024-01314-3>
- Hershey, T., Black, K. J., Hartlein, J. M., Barch, D. M., Braver, T. S., Carl, J. L., & Perlmutter, J. S. (2004). Cognitive-pharmacologic functional magnetic resonance imaging in tourette syndrome: A pilot study. *Biological Psychiatry*, 55(9), 916–925. <https://doi.org/10.1016/j.biopsych.2004.01.003>
- Hinton, G., Vinyals, O., & Dean, J. (2015). *Distilling the Knowledge in a Neural Network* (No. arXiv:1503.02531). arXiv. <http://arxiv.org/abs/1503.02531>
- Hirschtritt, M. E., Lee, P. C., Pauls, D. L., Dion, Y., Grados, M. A., Illmann, C., King, R. A., Sandor, P., McMahon, W. M., Lyon, G. J., Cath, D. C., Kurlan, R., Robertson, M. M., Osiecki, L., Scharf, J. M., & Mathews, C. A. (2015). Lifetime Prevalence, Age of Risk, and Genetic Relationships of Comorbid Psychiatric Disorders in Tourette Syndrome. *JAMA Psychiatry*, 72(4), 325. <https://doi.org/10.1001/jamapsychiatry.2014.2650>
- Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3), 306–310. [https://doi.org/10.1016/0013-4694\(70\)90143-4](https://doi.org/10.1016/0013-4694(70)90143-4)
- Hohlefeld, F. U., Huchzermeyer, C., Huebl, J., Schneider, G. -H., Brücke, C., Schönecker, T., Kühn, A. A., Curio, G., & Nikulin, V. V. (2014). Interhemispheric functional interactions between the

- subthalamic nuclei of patients with Parkinson's disease. *European Journal of Neuroscience*, 40(8), 3273–3283. <https://doi.org/10.1111/ejn.12686>
- Holt, A. B., Kormann, E., Gulberti, A., Pötter-Nerger, M., McNamara, C. G., Cagnan, H., Baaske, M. K., Little, S., Köppen, J. A., Buhmann, C., Westphal, M., Gerloff, C., Engel, A. K., Brown, P., Hamel, W., Moll, C. K. E., & Sharott, A. (2019). Phase-Dependent Suppression of Beta Oscillations in Parkinson's Disease Patients. *Journal of Neuroscience*, 39(6), 1119–1134. <https://doi.org/10.1523/JNEUROSCI.1913-18.2018>
- Horgan, J. (2005). The Forgotten Era of Brain Chips. *Scientific American*, 293(4), 66–73. <https://doi.org/10.1038/scientificamerican1005-66>
- Horn, A., & Kühn, A. A. (2015). Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *NeuroImage*, 107, 127–135. <https://doi.org/10.1016/j.neuroimage.2014.12.002>
- Horn, A., Li, N., Dembek, T. A., Kappel, A., Boulay, C., Ewert, S., Tietze, A., Husch, A., Perera, T., Neumann, W.-J., Reisert, M., Si, H., Oostenveld, R., Rorden, C., Yeh, F.-C., Fang, Q., Herrington, T. M., Vorwerk, J., & Kühn, A. A. (2019). Lead-DBS v2: Towards a comprehensive pipeline for deep brain stimulation imaging. *NeuroImage*, 184, 293–316. <https://doi.org/10.1016/j.neuroimage.2018.08.068>
- Horn, A., Neumann, W., Degen, K., Schneider, G., & Kühn, A. A. (2017). Toward an electrophysiological “sweet spot” for deep brain stimulation in the subthalamic nucleus. *Human Brain Mapping*, 38(7), 3377–3390. <https://doi.org/10.1002/hbm.23594>
- Horsley, V. (1890). Remarks on the surgery of the central nervous system. *British Medical Journal*, 2(1562), 1286.

- Houde, J. F., & Nagarajan, S. S. (2011). Speech Production as State Feedback Control. *Frontiers in Human Neuroscience*, 5. <https://doi.org/10.3389/fnhum.2011.00082>
- Hubble, J. P., Busenbark, K. L., Wilkinson, S., Penn, R. D., Lyons, K., & Koller, W. C. (1996). Deep brain stimulation for essential tremor. *Neurology*, 46(4), 1150–1153. <https://doi.org/10.1212/WNL.46.4.1150>
- Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A.-L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *eLife*, 4, e06213. <https://doi.org/10.7554/eLife.06213>
- Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift* (No. arXiv:1502.03167). arXiv. <https://doi.org/10.48550/arXiv.1502.03167>
- Isaias, I. U., Alterman, R. L., & Tagliati, M. (2009). Deep Brain Stimulation for Primary Generalized Dystonia: Long-term Outcomes. *Archives of Neurology*, 66(4). <https://doi.org/10.1001/archneurol.2009.20>
- Jeffries, K. (2002). The Functional Neuroanatomy of Tourette's Syndrome An FDG PET Study III: Functional Coupling of Regional Cerebral Metabolic Rates. *Neuropsychopharmacology*, 27(1), 92–104. [https://doi.org/10.1016/S0893-133X\(01\)00428-6](https://doi.org/10.1016/S0893-133X(01)00428-6)
- Jergas, H., Petry-Schmelzer, J. N., Hannemann, J. H., Thies, T., Strelow, J. N., Rubi-Fessen, I., Quinting, J., Baldermann, J. C., Mücke, D., Fink, G. R., Visser-Vandewalle, V., Dembek, T. A., & Barbe, M. T. (2024). One side effect: Two networks? Lateral and posteromedial stimulation spreads induce dysarthria in subthalamic deep brain stimulation for Parkinson's

disease. *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp-2024-333434.

<https://doi.org/10.1136/jnnp-2024-333434>

Jimenez-Shahed, J. (2021). Device profile of the percept PC deep brain stimulation system for the treatment of Parkinson's disease and related disorders. *Expert Review of Medical Devices*, 18(4), 319–332. <https://doi.org/10.1080/17434440.2021.1909471>

Johnson, K. A., Dosenbach, N. U. F., Gordon, E. M., Welle, C. G., Wilkins, K. B., Bronte-Stewart, H. M., Voon, V., Morishita, T., Sakai, Y., Merner, A. R., Lázaro-Muñoz, G., Williamson, T., Horn, A., Gilron, R., O'Keeffe, J., Gittis, A. H., Neumann, W.-J., Little, S., Provenza, N. R., ... Wong, J. K. (2024). Proceedings of the 11th Annual Deep Brain Stimulation Think Tank: Pushing the forefront of neuromodulation with functional network mapping, biomarkers for adaptive DBS, bioethical dilemmas, AI-guided neuromodulation, and translational advancements. *Frontiers in Human Neuroscience*, 18, 1320806. <https://doi.org/10.3389/fnhum.2024.1320806>

Johnson, K. A., Duffley, G., Anderson, D. N., Ostrem, J. L., Welter, M.-L., Baldermann, J. C., Kuhn, J., Huys, D., Visser-Vandewalle, V., Foltynie, T., Zrinzo, L., Hariz, M., Leentjens, A. F. G., Mogilner, A. Y., Pourfar, M. H., Almeida, L., Gunduz, A., Foote, K. D., Okun, M. S., & Butson, C. R. (2020). Structural connectivity predicts clinical outcomes of deep brain stimulation for Tourette syndrome. *Brain*, 143(8), 2607–2623. <https://doi.org/10.1093/brain/awaa188>

Johnson, K. A., Worbe, Y., Foote, K. D., Butson, C. R., Gunduz, A., & Okun, M. S. (2023). Tourette syndrome: Clinical features, pathophysiology, and treatment. *The Lancet Neurology*, 22(2), 147–158. [https://doi.org/10.1016/S1474-4422\(22\)00303-9](https://doi.org/10.1016/S1474-4422(22)00303-9)

- Jorge, A., Dastolfo-Hromack, C., Lipski, W. J., Kratter, I. H., Smith, L. J., Gartner-Schmidt, J. L., & Richardson, R. M. (2020a). Anterior Sensorimotor Subthalamic Nucleus Stimulation Is Associated With Improved Voice Function. *Neurosurgery*, 87(4), 788–795.
<https://doi.org/10.1093/neuros/nyaa024>
- Jorge, A., Dastolfo-Hromack, C., Lipski, W. J., Kratter, I. H., Smith, L. J., Gartner-Schmidt, J. L., & Richardson, R. M. (2020b). Anterior Sensorimotor Subthalamic Nucleus Stimulation Is Associated With Improved Voice Function. *Neurosurgery*, 87(4), 788–795.
<https://doi.org/10.1093/neuros/nyaa024>
- Jorge, A., Lipski, W. J., Wang, D., Crammond, D. J., Turner, R. S., & Richardson, R. M. (2022). Hyperdirect connectivity of opercular speech network to the subthalamic nucleus. *Cell Reports*, 38(10), 110477. <https://doi.org/10.1016/j.celrep.2022.110477>
- Kashanian, A., Rohatgi, P., Chivukula, S., Sheth, S. A., & Pouratian, N. (2021). Deep Brain Electrode Externalization and Risk of Infection: A Systematic Review and Meta-Analysis. *Operative Neurosurgery*, 20(2), 141–150. <https://doi.org/10.1093/ons/opaa268>
- Kawano, T., Hattori, N., Uno, Y., Hatakenaka, M., Yagura, H., Fujimoto, H., Nagasako, M., Mochizuki, H., Kitajo, K., & Miyai, I. (2021). Association between aphasia severity and brain network alterations after stroke assessed using the electroencephalographic phase synchrony index. *Scientific Reports*, 11(1), 12469. <https://doi.org/10.1038/s41598-021-91978-7>
- Kelly, R. M., & Strick, P. L. (2004). Macro-architecture of basal ganglia loops with the cerebral cortex: Use of rabies virus to reveal multisynaptic circuits. In *Progress in Brain Research* (Vol. 143, pp. 447–459). Elsevier. [https://doi.org/10.1016/S0079-6123\(03\)43042-2](https://doi.org/10.1016/S0079-6123(03)43042-2)

Kempf, F., Kühn, A. A., Kupsch, A., Brücke, C., Weise, L., Schneider, G.-H., & Brown, P. (2007).

Premovement activities in the subthalamic area of patients with Parkinson's disease and their dependence on task. *European Journal of Neuroscience*, 25(10), 3137–3145.

<https://doi.org/10.1111/j.1460-9568.2007.05536.x>

Khawaldeh, S., Tinkhauser, G., Shah, S. A., Peterman, K., Debove, I., Nguyen, T. A. K., Nowacki, A.,

Lachenmayer, M. L., Schuepbach, M., Pollo, C., Krack, P., Woolrich, M., & Brown, P. (2020).

Subthalamic nucleus activity dynamics and limb movement prediction in Parkinson's disease.

Brain, 143(2), 582–596. <https://doi.org/10.1093/brain/awz417>

Khawaldeh, S., Tinkhauser, G., Torrecillos, F., He, S., Foltynie, T., Limousin, P., Zrinzo, L., Oswal,

A., Quinn, A. J., Vidaurre, D., Tan, H., Litvak, V., Kühn, A., Woolrich, M., & Brown, P.

(2022). Balance between competing spectral states in subthalamic nucleus is linked to motor impairment in Parkinson's disease. *Brain*, 145(1), 237–250.

<https://doi.org/10.1093/brain/awab264>

Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (No. arXiv:1412.6980).

arXiv. <http://arxiv.org/abs/1412.6980>

Klambauer, G., Unterthiner, T., Mayr, A., & Hochreiter, S. (2017). *Self-Normalizing Neural Networks*

(No. arXiv:1706.02515). arXiv. <http://arxiv.org/abs/1706.02515>

Kluin, K. J., Mossner, J. M., Costello, J. T., Chou, K. L., & Patil, P. G. (2022). Motor speech effects

in subthalamic deep brain stimulation for Parkinson's disease. *Journal of Neurosurgery*,

137(3), 722–728. <https://doi.org/10.3171/2021.12.JNS211729>

Köhler, R. M., Binns, T. S., Merk, T., Zhu, G., Yin, Z., Zhao, B., Chikermane, M., Vanhoecke, J.,

Busch, J. L., Habets, J. G. V., Faust, K., Schneider, G.-H., Cavallo, A., Haufe, S., Zhang, J.,

- Kühn, A. A., Haynes, J.-D., & Neumann, W.-J. (2024). Dopamine and deep brain stimulation accelerate the neural dynamics of volitional action in Parkinson's disease. *Brain*, awae219. <https://doi.org/10.1093/brain/awae219>
- Koller, W., Pahwa, R., Busenbark, K., Hubble, J., Wilkinson, S., Lang, A., Tuite, P., Sime, E., Lazano, A., Hauser, R., Malapira, T., Smith, D., Tarsy, D., Miyawaki, E., Norregaard, T., Kormos, T., & Olanow, C. W. (1997). High-frequency unilateral thalamic stimulation in the treatment of essential and parkinsonian tremor. *Annals of Neurology*, 42(3), 292–299. <https://doi.org/10.1002/ana.410420304>
- Kothe, C. A. (2014). *Lab streaming layer (LSL)*. <https://github.com/sccn/labstreaminglayer>
- Krack, P., Batir, A., Van Blercom, N., Chabardes, S., Fraix, V., Ardouin, C., Koudsie, A., Limousin, P. D., Benazzouz, A., LeBas, J. F., Benabid, A.-L., & Pollak, P. (2003). Five-Year Follow-up of Bilateral Stimulation of the Subthalamic Nucleus in Advanced Parkinson's Disease. *New England Journal of Medicine*, 349(20), 1925–1934. <https://doi.org/10.1056/NEJMoa035275>
- Kremer, N. I., Pauwels, R. W. J., Pozzi, N. G., Lange, F., Roothans, J., Volkmann, J., & Reich, M. M. (2021). Deep Brain Stimulation for Tremor: Update on Long-Term Outcomes, Target Considerations and Future Directions. *Journal of Clinical Medicine*, 10(16), 3468. <https://doi.org/10.3390/jcm10163468>
- Kringelbach, M. L., & Aziz, T. Z. (2009). Deep Brain Stimulation: Avoiding the Errors of Psychosurgery. *JAMA*, 301(16), 1705. <https://doi.org/10.1001/jama.2009.551>
- Kringelbach, M. L., Green, A. L., Owen, S. L. F., Schweder, P. M., & Aziz, T. Z. (2010). Sing the mind electric – principles of deep brain stimulation. *European Journal of Neuroscience*, 32(7), 1070–1079. <https://doi.org/10.1111/j.1460-9568.2010.07419.x>

- Kühn, A. A., Kupsch, A., Schneider, G.-H., & Brown, P. (2006). Reduction in subthalamic 8–35 Hz oscillatory activity correlates with clinical improvement in Parkinson's disease. *European Journal of Neuroscience*, 23(7), 1956–1960. <https://doi.org/10.1111/j.1460-9568.2006.04717.x>
- Kuhn, J., Janouschek, H., Raptis, M., Rex, S., Lenartz, D., Neuner, I., Mottaghy, F. M., Schneider, F., Schaefer, W. M., Sturm, V., Gründer, G., & Vernaleken, I. (2012). In Vivo Evidence of Deep Brain Stimulation-Induced Dopaminergic Modulation in Tourette's Syndrome. *Biological Psychiatry*, 71(5), e11–e13. <https://doi.org/10.1016/j.biopsych.2011.09.035>
- Kumar, K., Toth, C., & Nath, R. K. (1997). Deep Brain Stimulation for Intractable Pain: A 15-Year Experience. *Neurosurgery*, 40(4), 736–747. <https://doi.org/10.1097/00006123-199704000-00015>
- Lachaux, J.-P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, 8(4), Article 4. [https://doi.org/10.1002/\(SICI\)1097-0193\(1999\)8:4<194::AID-HBM4>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C)
- Lalo, E., Thobois, S., Sharott, A., Polo, G., Mertens, P., Pogosyan, A., & Brown, P. (2008). Patterns of Bidirectional Communication between Cortex and Basal Ganglia during Movement in Patients with Parkinson Disease. *The Journal of Neuroscience*, 28(12), 3008. <https://doi.org/10.1523/JNEUROSCI.5295-07.2008>
- Laxton, A. W., Tang-Wai, D. F., McAndrews, M. P., Zumsteg, D., Wennberg, R., Keren, R., Wherrett, J., Naglie, G., Hamani, C., Smith, G. S., & Lozano, A. M. (2010). A phase I trial of deep brain stimulation of memory circuits in Alzheimer's disease. *Annals of Neurology*, 68(4), 521–534. <https://doi.org/10.1002/ana.22089>

- Leckman, J. F., Bloch, M. H., Smith, M. E., Larabi, D., & Hampson, M. (2010). Neurobiological Substrates of Tourette's Disorder. *Journal of Child and Adolescent Psychopharmacology*, 20(4), 237–247. <https://doi.org/10.1089/cap.2009.0118>
- Lei, H., Yang, C., Zhang, M., Qiu, Y., Wang, J., Xu, J., Hu, X., & Wu, X. (2022). Optimal Contact Position of Subthalamic Nucleus Deep Brain Stimulation for Reducing Restless Legs Syndrome in Parkinson's Disease Patients: One-Year Follow-Up with 33 Patients. *Brain Sciences*, 12(12), 1645. <https://doi.org/10.3390/brainsci12121645>
- Li, Y., Murias, M., Major, S., Dawson, G., Dzirasa, K., Carin, L., & Carlson, D. E. (2017). Targeting EEG/LFP synchrony with neural nets. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4623–4633.
- Lieberman, P. (2000). *Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought*. Harvard University Press.
- Lieberman, P. (2009). FOXP2 and Human Cognition. *Cell*, 137(5), 800–802. <https://doi.org/10.1016/j.cell.2009.05.013>
- Limousin, P., Krack, P., Pollak, P., Benazzouz, A., Ardouin, C., Hoffmann, D., & Benabid, A.-L. (1998). Electrical Stimulation of the Subthalamic Nucleus in Advanced Parkinson's Disease. *New England Journal of Medicine*, 339(16), 1105–1111. <https://doi.org/10.1056/NEJM199810153391603>
- Limousin, P., Pollak, P., Benazzouz, A., Hoffmann, D., Le Bas, J.-F., Perret, J. E., Benabid, A.-L., & Broussolle, E. (1995). Effect on parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. *The Lancet*, 345(8942), 91–95. [https://doi.org/10.1016/S0140-6736\(95\)90062-4](https://doi.org/10.1016/S0140-6736(95)90062-4)

- Lio, G., Thobois, S., Ballanger, B., Lau, B., & Boulinguez, P. (2018). Removing deep brain stimulation artifacts from the electroencephalogram: Issues, recommendations and an open-source toolbox. *Clinical Neurophysiology*, 129(10), 2170–2185.
<https://doi.org/10.1016/j.clinph.2018.07.023>
- Lipski, W. J., Alhourani, A., Pirnia, T., Jones, P. W., Dastolfo-Hromack, C., Helou, L. B., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2018a). Subthalamic Nucleus Neurons Differentially Encode Early and Late Aspects of Speech Production. *The Journal of Neuroscience*, 38(24), 5620–5631.
<https://doi.org/10.1523/JNEUROSCI.3480-17.2018>
- Lipski, W. J., Alhourani, A., Pirnia, T., Jones, P. W., Dastolfo-Hromack, C., Helou, L. B., Crammond, D. J., Shaiman, S., Dickey, M. W., Holt, L. L., Turner, R. S., Fiez, J. A., & Richardson, R. M. (2018b). Subthalamic Nucleus Neurons Differentially Encode Early and Late Aspects of Speech Production. *The Journal of Neuroscience*, 38(24), 5620–5631.
<https://doi.org/10.1523/JNEUROSCI.3480-17.2018>
- Lipski, W. J., Bush, A., Chrabaszczyk, A., Crammond, D. J., Fiez, J. A., Turner, R. S., & Richardson, R. M. (2023). *Subthalamic nucleus neurons encode syllable sequence and phonetic characteristics during speech*. <https://doi.org/10.1101/2023.12.11.569290>
- Lipski, W. J., Wozny, T. A., Alhourani, A., Kondylis, E. D., Turner, R. S., Crammond, D. J., & Richardson, R. M. (2017). Dynamics of human subthalamic neuron phase-locking to motor and sensory cortical oscillations during movement. *Journal of Neurophysiology*, 118(3), 1472–1487. <https://doi.org/10.1152/jn.00964.2016>

- Little, S., & Brown, P. (2012). What brain signals are suitable for feedback control of deep brain stimulation in Parkinson's disease? *Annals of the New York Academy of Sciences*, 1265(1), 9–24. <https://doi.org/10.1111/j.1749-6632.2012.06650.x>
- Little, S., & Brown, P. (2020). Debugging Adaptive Deep Brain Stimulation for Parkinson's Disease. *Movement Disorders*, 35(4), 555–561. <https://doi.org/10.1002/mds.27996>
- 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>
- Little, S., Tripoliti, E., Beudel, M., Pogosyan, A., Cagnan, H., Herz, D., Bestmann, S., Aziz, T., Cheeran, B., Zrinzo, L., Hariz, M., Hyam, J., Limousin, P., Foltynie, T., & Brown, P. (2016). Adaptive deep brain stimulation for Parkinson's disease demonstrates reduced speech side effects compared to conventional stimulation in the acute setting. *Journal of Neurology, Neurosurgery & Psychiatry*, 87(12), 1388–1389. <https://doi.org/10.1136/jnnp-2016-313518>
- Litvak, V., Eusebio, A., Jha, A., Oostenveld, R., Barnes, G., Foltynie, T., Limousin, P., Zrinzo, L., Hariz, M. I., Friston, K., & Brown, P. (2012). Movement-related changes in local and long-range synchronization in Parkinson's disease revealed by simultaneous magnetoencephalography and intracranial recordings. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(31), 10541–10553. <https://doi.org/10.1523/JNEUROSCI.0767-12.2012>
- Livezey, J. A., Bouchard, K. E., & Chang, E. F. (2019). Deep learning as a tool for neural data analysis: Speech classification and cross-frequency coupling in human sensorimotor cortex. *PLOS Computational Biology*, 15(9), e1007091. <https://doi.org/10.1371/journal.pcbi.1007091>

- Livezey, J. A., & Glaser, J. I. (2021). Deep learning approaches for neural decoding across architectures and recording modalities. *Briefings in Bioinformatics*, 22(2), 1577–1591. <https://doi.org/10.1093/bib/bbaa355>
- Lizarazu, M., Carreiras, M., & Molinaro, N. (2023). Theta-gamma phase-amplitude coupling in auditory cortex is modulated by language proficiency. *Human Brain Mapping*, 44(7), 2862–2872. <https://doi.org/10.1002/hbm.26250>
- Lizarazu, M., Lallier, M., & Molinaro, N. (2019). Phase–amplitude coupling between theta and gamma oscillations adapts to speech rate. *Annals of the New York Academy of Sciences*, 1453(1), 140. <https://doi.org/10.1111/nyas.14099>
- Lofredi, R., Neumann, W.-J., Bock, A., Horn, A., Huebl, J., Siegert, S., Schneider, G.-H., Krauss, J. K., & Kühn, A. A. (2018). Dopamine-dependent scaling of subthalamic gamma bursts with movement velocity in patients with Parkinson’s disease. *eLife*, 7, e31895. <https://doi.org/10.7554/eLife.31895>
- Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and Cooccurrence of Vocal Tract Dysfunctions in the Speech of a Large Sample of Parkinson Patients. *Journal of Speech and Hearing Disorders*, 43(1), 47–57. <https://doi.org/10.1044/jshd.4301.47>
- López-Azcárate, J., Tainta, M., Rodríguez-Oroz, M. C., Valencia, M., González, R., Guridi, J., Iriarte, J., Obeso, J. A., Artieda, J., & Alegre, M. (2010). Coupling between beta and high-frequency activity in the human subthalamic nucleus may be a pathophysiological mechanism in Parkinson’s disease. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(19), 6667–6677. <https://doi.org/10.1523/JNEUROSCI.5459-09.2010>

- Lozano, A. M., & Lipsman, N. (2013). Probing and Regulating Dysfunctional Circuits Using Deep Brain Stimulation. *Neuron*, 77(3), 406–424. <https://doi.org/10.1016/j.neuron.2013.01.020>
- Lozano, A. M., Lipsman, N., Bergman, H., Brown, P., Chabardes, S., Chang, J. W., Matthews, K., McIntyre, C. C., Schlaepfer, T. E., Schulder, M., Temel, Y., Volkmann, J., & Krauss, J. K. (2019). Deep brain stimulation: Current challenges and future directions. *Nature Reviews Neurology*, 15(3), 148–160. <https://doi.org/10.1038/s41582-018-0128-2>
- Lozano, A. M., Mayberg, H. S., Giacobbe, P., Hamani, C., Craddock, R. C., & Kennedy, S. H. (2008). Subcallosal Cingulate Gyrus Deep Brain Stimulation for Treatment-Resistant Depression. *Biological Psychiatry*, 64(6), 461–467. <https://doi.org/10.1016/j.biopsych.2008.05.034>
- Lu, G., Luo, L., Liu, M., Zheng, Z., Zhang, B., Chen, X., Hua, X., Fan, H., Mo, G., Duan, J., Li, M., Hong, T., & Zhou, D. (2020). Outcomes and Adverse Effects of Deep Brain Stimulation on the Ventral Intermediate Nucleus in Patients with Essential Tremor. *Neural Plasticity*, 2020(1), 2486065. <https://doi.org/10.1155/2020/2486065>
- Lyons, K. E., & Pahwa, R. (2008). Deep Brain Stimulation and Tremor. *Neurotherapeutics*, 5(2), 331–338. <https://doi.org/10.1016/j.nurt.2008.01.004>
- Malekmohammadi, M., Herron, J., Velisar, A., Blumenfeld, Z., Trager, M. H., Chizeck, H. J., & Brontë-Stewart, H. (2016). Kinematic Adaptive Deep Brain Stimulation for Resting Tremor in Parkinson's Disease. *Movement Disorders*, 31(3), 426–428. <https://doi.org/10.1002/mds.26482>
- Mallet, L., Polosan, M., Jaafari, N., Baup, N., Welter, M.-L., Fontaine, D., Montcel, S. T. D., Yelnik, J., Chéreau, I., Arbus, C., Raoul, S., Aouizerate, B., Damier, P., Chabardès, S., Czernecki, V., Ardouin, C., Krebs, M.-O., Bardinet, E., Chaynes, P., ... Pelissolo, A. (2008). Subthalamic

- Nucleus Stimulation in Severe Obsessive–Compulsive Disorder. *New England Journal of Medicine*, 359(20), 2121–2134. <https://doi.org/10.1056/NEJMoa0708514>
- Manes, J. L., Parkinson, A. L., Larson, C. R., Greenlee, J. D., Eickhoff, S. B., Corcos, D. M., & Robin, D. A. (2013). Connectivity of the subthalamic nucleus and globus pallidus pars interna to regions within the speech network: A meta-analytic connectivity study. *Human Brain Mapping*, 35(7), 3499. <https://doi.org/10.1002/hbm.22417>
- Mann, J. M., Foote, K. D., Garvan, C. W., Fernandez, H. H., Jacobson, C. E., Rodriguez, R. L., Haq, I. U., Siddiqui, M. S., Malaty, I. A., Morishita, T., Hass, C. J., & Okun, M. S. (2009). Brain penetration effects of microelectrodes and DBS leads in STN or GPi. *Journal of Neurology, Neurosurgery & Psychiatry*, 80(7), 794–798. <https://doi.org/10.1136/jnnp.2008.159558>
- Marceglia, S., Guidetti, M., Harmsen, I. E., Loh, A., Meoni, S., Foffani, G., Lozano, A. M., Volkmann, J., Moro, E., & Priori, A. (2021). Deep brain stimulation: Is it time to change gears by closing the loop? *Journal of Neural Engineering*, 18(6), 061001. <https://doi.org/10.1088/1741-2552/ac3267>
- Marceglia, S., Rosa, M., Servello, D., Porta, M., Barbieri, S., Moro, E., & Priori, A. (2017). Adaptive Deep Brain Stimulation (aDBS) for Tourette Syndrome. *Brain Sciences*, 8(1), 4. <https://doi.org/10.3390/brainsci8010004>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Martinez-Ramirez, D., Jimenez-Shahed, J., Leckman, J. F., Porta, M., Servello, D., Meng, F.-G., Kuhn, J., Huys, D., Baldermann, J. C., Foltynie, T., Hariz, M. I., Joyce, E. M., Zrinzo, L.,

- Kefalopoulou, Z., Silburn, P., Coyne, T., Mogilner, A. Y., Pourfar, M. H., Khandhar, S. M., ... Okun, M. S. (2018). Efficacy and Safety of Deep Brain Stimulation in Tourette Syndrome: The International Tourette Syndrome Deep Brain Stimulation Public Database and Registry. *JAMA Neurology*, 75(3), 353. <https://doi.org/10.1001/jamaneurol.2017.4317>
- Mashour, G. A., Walker, E. E., & Martuza, R. L. (2005). Psychosurgery: Past, present, and future. *Brain Research Reviews*, 48(3), 409–419. <https://doi.org/10.1016/j.brainresrev.2004.09.002>
- Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., Schwalb, J. M., & Kennedy, S. H. (2005). Deep Brain Stimulation for Treatment-Resistant Depression. *Neuron*, 45(5), 651–660. <https://doi.org/10.1016/j.neuron.2005.02.014>
- McIntyre, C. C., Grill, W. M., Sherman, D. L., & Thakor, N. V. (2004). Cellular Effects of Deep Brain Stimulation: Model-Based Analysis of Activation and Inhibition. *Journal of Neurophysiology*, 91(4), 1457–1469. <https://doi.org/10.1152/jn.00989.2003>
- McNaught, K. St. P., & Mink, J. W. (2011). Advances in understanding and treatment of Tourette syndrome. *Nature Reviews Neurology*, 7(12), 667–676. <https://doi.org/10.1038/nrneurol.2011.167>
- Meidahl, A. C., Moll, C. K. E., van Wijk, B. C. M., Gulberti, A., Tinkhauser, G., Westphal, M., Engel, A. K., Hamel, W., Brown, P., & Sharott, A. (2019a). Synchronised spiking activity underlies phase amplitude coupling in the subthalamic nucleus of Parkinson's disease patients. *Neurobiology of Disease*, 127, 101–113. <https://doi.org/10.1016/j.nbd.2019.02.005>
- Meidahl, A. C., Moll, C. K. E., van Wijk, B. C. M., Gulberti, A., Tinkhauser, G., Westphal, M., Engel, A. K., Hamel, W., Brown, P., & Sharott, A. (2019b). Synchronised spiking activity

underlies phase amplitude coupling in the subthalamic nucleus of Parkinson's disease patients.

Neurobiology of Disease, 127, 101–113. <https://doi.org/10.1016/j.nbd.2019.02.005>

Meidahl, A. C., Tinkhauser, G., Herz, D. M., Cagnan, H., Debarros, J., & Brown, P. (2017). Adaptive Deep Brain Stimulation for Movement Disorders: The Long Road to Clinical Therapy: Adaptive DBS Review. *Movement Disorders*, 32(6), 810–819. <https://doi.org/10.1002/mds.27022>

Merk, T., Köhler, R., Peterson, V., Lyra, L., Vanhoecke, J., Chikermane, M., Binns, T., Li, N., Walton, A., Bush, A., Sisterson, N., Busch, J., Lofredi, R., Habets, J., Huebl, J., Zhu, G., Yin, Z., Zhao, B., Merkl, A., ... Neumann, W.-J. (2023). *Invasive neurophysiology and whole brain connectomics for neural decoding in patients with brain implants*. <https://doi.org/10.21203/rs.3.rs-3212709/v1>

Merk, T., Peterson, V., Köhler, R., Haufe, S., Richardson, R. M., & Neumann, W.-J. (2022). Machine learning based brain signal decoding for intelligent adaptive deep brain stimulation. *Experimental Neurology*, 351, 113993. <https://doi.org/10.1016/j.expneurol.2022.113993>

Merk, T., Peterson, V., Lipski, W. J., Blankertz, B., Turner, R. S., Li, N., Horn, A., Richardson, R. M., & Neumann, W.-J. (2022). Electrocorticography is superior to subthalamic local field potentials for movement decoding in Parkinson's disease. *eLife*, 11, e75126. <https://doi.org/10.7554/eLife.75126>

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic Feature Encoding in Human Superior Temporal Gyrus. *Science*, 343(6174), 1006–1010. <https://doi.org/10.1126/science.1245994>

- Metzger, S. L., Littlejohn, K. T., Silva, A. B., Moses, D. A., Seaton, M. P., Wang, R., Dougherty, M. E., Liu, J. R., Wu, P., Berger, M. A., Zhuravleva, I., Tu-Chan, A., Ganguly, K., Anumanchipalli, G. K., & Chang, E. F. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. *Nature*. <https://doi.org/10.1038/s41586-023-06443-4>
- Mink, J. W. (2001). Basal ganglia dysfunction in Tourette's syndrome: A new hypothesis. *Pediatric Neurology*, 25(3), 190–198. [https://doi.org/10.1016/S0887-8994\(01\)00262-4](https://doi.org/10.1016/S0887-8994(01)00262-4)
- Mink, J. W. (2003). The Basal Ganglia and Involuntary Movements: Impaired Inhibition of Competing Motor Patterns. *Archives of Neurology*, 60(10), 1365. <https://doi.org/10.1001/archneur.60.10.1365>
- Miocinovic, S., Hemptinne, C. de, Chen, W., Isbaine, F., Willie, J. T., Ostrem, J. L., & Starr, P. A. (2018). Cortical Potentials Evoked by Subthalamic Stimulation Demonstrate a Short Latency Hyperdirect Pathway in Humans. *Journal of Neuroscience*, 38(43), 9129–9141. <https://doi.org/10.1523/JNEUROSCI.1327-18.2018>
- Molina, R., Okun, M. S., Shute, J. B., Opri, E., Rossi, P. J., Martinez-Ramirez, D., Foote, K. D., & Gunduz, A. (2018). Report of a patient undergoing chronic responsive deep brain stimulation for Tourette syndrome: Proof of concept. *Journal of Neurosurgery*, 129(2), 308–314. <https://doi.org/10.3171/2017.6.JNS17626>
- Moniz, E. (1937). Prefrontal leucotomy in the treatment of mental disorders. *American Journal of Psychiatry*, 93(6), 1379–1385.
- Montavon, G., Samek, W., & Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73, 1–15. <https://doi.org/10.1016/j.dsp.2017.10.011>

- Morand-Beaulieu, S., Wu, J., Mayes, L. C., Grantz, H., Leckman, J. F., Crowley, M. J., & Sukhodolsky, D. G. (2023). Increased Alpha-Band Connectivity During Tic Suppression in Children With Tourette Syndrome Revealed by Source Electroencephalography Analyses. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 8(3), 241–250.
<https://doi.org/10.1016/j.bpsc.2021.05.001>
- Moreau, C., & Pinto, S. (2019). Misconceptions about speech impairment in Parkinson's disease. *Movement Disorders*, 34(10), 1471–1475. <https://doi.org/10.1002/mds.27791>
- Morello, A. N. D. C., Beber, B. C., Fagundes, V. C., Cielo, C. A., & Rieder, C. R. M. (2020). Dysphonia and Dysarthria in People With Parkinson's Disease After Subthalamic Nucleus Deep Brain Stimulation: Effect of Frequency Modulation. *Journal of Voice*, 34(3), 477–484.
<https://doi.org/10.1016/j.jvoice.2018.10.012>
- Morgan, J. P. (1982). The First Reported Case of Electrical Stimulation of the Human Brain. *Journal of the History of Medicine and Allied Sciences*, XXXVII(1), 51–64.
<https://doi.org/10.1093/jhmas/XXXVII.1.51>
- Moses, D. A., Metzger, S. L., Liu, J. R., Anumanchipalli, G. K., Makin, J. G., Sun, P. F., Chartier, J., Dougherty, M. E., Liu, P. M., Abrams, G. M., Tu-Chan, A., Ganguly, K., & Chang, E. F. (2021). Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria. *New England Journal of Medicine*, 385(3), 217–227. <https://doi.org/10.1056/NEJMoa2027540>
- Mostofi, A., Baig, F., Bourlogiannis, F., Uberti, M., Morgante, F., & Pereira, E. A. C. (2021). Postoperative Externalization of Deep Brain Stimulation Leads Does Not Increase Infection Risk. *Neuromodulation: Technology at the Neural Interface*, 24(2), 265–271.
<https://doi.org/10.1111/ner.13331>

- Müller, G. R., Neuper, C., Rupp, R., Keinrath, C., Gerner, H. J., & Pfurtscheller, G. (2003). Event-related beta EEG changes during wrist movements induced by functional electrical stimulation of forearm muscles in man. *Neuroscience Letters*, *340*(2), 143–147.
[https://doi.org/10.1016/S0304-3940\(03\)00019-3](https://doi.org/10.1016/S0304-3940(03)00019-3)
- Myers, M. H., Padmanabha, A., Hossain, G., De Jongh Curry, A. L., & Blaha, C. D. (2016). Seizure Prediction and Detection via Phase and Amplitude Lock Values. *Frontiers in Human Neuroscience*, *10*. <https://doi.org/10.3389/fnhum.2016.00080>
- Nambu, A., Takada, M., Inase, M., & Tokuno, H. (1996). Dual somatotopical representations in the primate subthalamic nucleus: Evidence for ordered but reversed body-map transformations from the primary motor cortex and the supplementary motor area. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *16*(8), 2671–2683.
<https://doi.org/10.1523/JNEUROSCI.16-08-02671.1996>
- Neumann, W., Degen, K., Schneider, G., Brücke, C., Huebl, J., Brown, P., & Kühn, A. A. (2016). Subthalamic synchronized oscillatory activity correlates with motor impairment in patients with Parkinson's disease. *Movement Disorders*, *31*(11), 1748–1751.
<https://doi.org/10.1002/mds.26759>
- Neumann, W., & Rodriguez-Oroz, M. C. (2021). Machine Learning Will Extend the Clinical Utility of Adaptive Deep Brain Stimulation. *Movement Disorders*, *36*(4), 796–799.
<https://doi.org/10.1002/mds.28567>
- Neumann, W.-J., Gilron, R., Little, S., & Tinkhauser, G. (2023). Adaptive Deep Brain Stimulation: From Experimental Evidence Toward Practical Implementation. *Movement Disorders*, *38*(6), 937–948. <https://doi.org/10.1002/mds.29415>

- Neumann, W.-J., Horn, A., & Kühn, A. A. (2023). Insights and opportunities for deep brain stimulation as a brain circuit intervention. *Trends in Neurosciences*, 46(6), 472–487.
<https://doi.org/10.1016/j.tins.2023.03.009>
- Neumann, W.-J., Köhler, R. M., & Kühn, A. A. (2022). A practical guide to invasive neurophysiology in patients with deep brain stimulation. *Clinical Neurophysiology*, 140, 171–180.
<https://doi.org/10.1016/j.clinph.2022.05.004>
- Neumann, W.-J., Turner, R. S., Blankertz, B., Mitchell, T., Kühn, A. A., & Richardson, R. M. (2019). Toward Electrophysiology-Based Intelligent Adaptive Deep Brain Stimulation for Movement Disorders. *Neurotherapeutics*, 16(1), 105–118. <https://doi.org/10.1007/s13311-018-00705-0>
- Nie, Y., Guo, X., Li, X., Geng, X., Li, Y., Quan, Z., Zhu, G., Yin, Z., Zhang, J., & Wang, S. (2021). Real-time removal of stimulation artifacts in closed-loop deep brain stimulation. *Journal of Neural Engineering*, 18(6), 066031. <https://doi.org/10.1088/1741-2552/ac3cc5>
- Norel, R., Agurto, C., Heisig, S., Rice, J. J., Zhang, H., Ostrand, R., Wacnik, P. W., Ho, B. K., Ramos, V. L., & Cecchi, G. A. (2020). Speech-based characterization of dopamine replacement therapy in people with Parkinson’s disease. *Npj Parkinson’s Disease*, 6(1), 12.
<https://doi.org/10.1038/s41531-020-0113-5>
- Okun, M. S., Cagle, J., Gomez, J., Bowers, D., Wong, J., Foote, K. D., & Gunduz, A. (2024). Responsive deep brain stimulation for the treatment of Tourette syndrome. *Scientific Reports*, 14(1), 6467. <https://doi.org/10.1038/s41598-024-57071-5>
- Okun, M. S., Gallo, B. V., Mandybur, G., Jagid, J., Foote, K. D., Revilla, F. J., Alterman, R., Jankovic, J., Simpson, R., Junn, F., Verhagen, L., Arle, J. E., Ford, B., Goodman, R. R., Stewart, R. M., Horn, S., Baltuch, G. H., Kopell, B. H., Marshall, F., ... Tagliati, M. (2012).

- Subthalamic deep brain stimulation with a constant-current device in Parkinson's disease: An open-label randomised controlled trial. *The Lancet Neurology*, 11(2), 140–149.
[https://doi.org/10.1016/S1474-4422\(11\)70308-8](https://doi.org/10.1016/S1474-4422(11)70308-8)
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, 2011, 156869. <https://doi.org/10.1155/2011/156869>
- Opri, E., Cernera, S., Molina, R., Eisinger, R. S., Cagle, J. N., Almeida, L., Denison, T., Okun, M. S., Foote, K. D., & Gunduz, A. (2020). Chronic embedded cortico-thalamic closed-loop deep brain stimulation for the treatment of essential tremor. *Science Translational Medicine*, 12(572), eaay7680. <https://doi.org/10.1126/scitranslmed.aay7680>
- Ozturk, M., Abosch, A., Francis, D., Wu, J., Jimenez-Shahed, J., & Ince, N. F. (2020). Distinct subthalamic coupling in the ON state describes motor performance in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society*, 35(1), 91–100.
<https://doi.org/10.1002/mds.27800>
- Pahwa, R., Lyons, K. E., Wilkinson, S. B., Simpson, R. K., Ondo, W. G., Tarsy, D., Norregaard, T., Hubble, J. P., Smith, D. A., Hauser, R. A., & Jankovic, J. (2006). *Long-term evaluation of deep brain stimulation of the thalamus*. <https://doi.org/10.3171/jns.2006.104.4.506>
- Parent, A. (2004). Giovanni Aldini: From Animal Electricity to Human Brain Stimulation. *Canadian Journal of Neurological Sciences / Journal Canadien Des Sciences Neurologiques*, 31(4), 576–584. <https://doi.org/10.1017/S0317167100003851>
- Paschen, S., Paschen, S., Forstenpointner, J., Forstenpointner, J., Forstenpointner, J., Becktepe, J. S., Becktepe, J. S., Heinzl, S., Heinzl, S., Heinzl, S., Hellriegel, H., Hellriegel, H., Witt, K.,

- Witt, K., Helmers, A., Helmers, A.-K., Deuschl, G., & Deuschl, G. (2019). Long-term efficacy of deep brain stimulation for essential tremor: An observer-blinded study. *Neurology*.
<https://doi.org/10.1212/wnl.00000000000007134>
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 12.
- Patra, D. P., Hess, R. A., Abi-Aad, K. R., Muzyka, I. M., & Bendok, B. R. (2019). Roberts Bartholow: The progenitor of human cortical stimulation and his contentious experiment. *Neurosurgical Focus*, 47(3), E6. <https://doi.org/10.3171/2019.6.FOCUS19349>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 6.
- Peters, J., & Tisch, S. (2021). Habituation After Deep Brain Stimulation in Tremor Syndromes: Prevalence, Risk Factors and Long-Term Outcomes. *Frontiers in Neurology*, 12.
<https://doi.org/10.3389/fneur.2021.696950>
- Peterson, B. S., Skudlarski, P., Anderson, A. W., Zhang, H., Gatenby, J. C., Lacadie, C. M., Leckman, J. F., & Gore, J. C. (1998). A Functional Magnetic Resonance Imaging Study of Tic Suppression in Tourette Syndrome. *ARCH GEN PSYCHIATRY*, 55.
<https://doi.org/10.1001/archpsyc.55.4.326>

- Peterson, S. M., Steine-Hanson, Z., Davis, N., Rao, R. P. N., & Brunton, B. W. (2021). Generalized neural decoders for transfer learning across participants and recording modalities. *Journal of Neural Engineering*, 18(2), 026014. <https://doi.org/10.1088/1741-2552/abda0b>
- Petry-Schmelzer, J. N., Jergas, H., Thies, T., Steffen, J. K., Reker, P., Dafsari, H. S., Mücke, D., Fink, G. R., Visser-Vandewalle, V., Dembek, T. A., & Barbe, M. T. (2021). Network Fingerprint of Stimulation-Induced Speech Impairment in Essential Tremor. *Annals of Neurology*, 89(2), 315–326. <https://doi.org/10.1002/ana.25958>
- Picillo, M., Lozano, A. M., Kou, N., Puppi Munhoz, R., & Fasano, A. (2016). Programming Deep Brain Stimulation for Parkinson’s Disease: The Toronto Western Hospital Algorithms. *Brain Stimulation*, 9(3), 425–437. <https://doi.org/10.1016/j.brs.2016.02.004>
- Piña-Fuentes, D., Van Dijk, J. M. C., Van Zijl, J. C., Moes, H. R., Van Laar, T., Oterdoom, D. L. M., Little, S., Brown, P., & Beudel, M. (2020). Acute effects of adaptive Deep Brain Stimulation in Parkinson’s disease. *Brain Stimulation*, 13(6), 1507–1516. <https://doi.org/10.1016/j.brs.2020.07.016>
- Pollak, P., Benabid, A. L., Gross, C., Gao, D. M., Laurent, A., Benazzouz, A., Hoffmann, D., Gentil, M., & Perret, J. (1993). [Effects of the stimulation of the subthalamic nucleus in Parkinson disease]. *Revue Neurologique*, 149(3), 175–176.
- Polychronis, S., Niccolini, F., Pagano, G., Yousaf, T., & Politis, M. (2019). Speech difficulties in early de novo patients with Parkinson’s disease. *Parkinsonism & Related Disorders*, 64, 256–261. <https://doi.org/10.1016/j.parkreldis.2019.04.026>

- Pourfar, M., Feigin, A., Tang, C. C., Carbon-Correll, M., Bussa, M., Budman, C., Dhawan, V., & Eidelberg, D. (2011). Abnormal metabolic brain networks in Tourette syndrome. *Neurology*, 76(11), 944–952. <https://doi.org/10.1212/WNL.0b013e3182104106>
- Preul, M. C., & Dagi, T. F. (2017). Editorial. London 1935: The frontal lobe, insanity, and a brain surgery. *Neurosurgical Focus*, 43(3), E5. <https://doi.org/10.3171/2017.6.FOCUS17429>
- Priori, A., Foffani, G., Rossi, L., & Marceglia, S. (2013). Adaptive deep brain stimulation (aDBS) controlled by local field potential oscillations. *Experimental Neurology*, 245, 77–86. <https://doi.org/10.1016/j.expneurol.2012.09.013>
- Provenza, N. R., Matteson, E. R., Allawala, A. B., Barrios-Anderson, A., Sheth, S. A., Viswanathan, A., McIngvale, E., Storch, E. A., Frank, M. J., McLaughlin, N. C. R., Cohn, J. F., Goodman, W. K., & Borton, D. A. (2019). The Case for Adaptive Neuromodulation to Treat Severe Intractable Mental Disorders. *Frontiers in Neuroscience*, 13, 152. <https://doi.org/10.3389/fnins.2019.00152>
- Pugh, J., Pycroft, L., Sandberg, A., Aziz, T., & Savulescu, J. (2018). Brainjacking in deep brain stimulation and autonomy. *Ethics and Information Technology*, 20(3), 219–232. <https://doi.org/10.1007/s10676-018-9466-4>
- Pycroft, L., Stein, J., & Aziz, T. (2018). Deep brain stimulation: An overview of history, methods, and future developments. *Brain and Neuroscience Advances*, 2, 2398212818816017. <https://doi.org/10.1177/2398212818816017>
- 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. IEEE Transactions on Neural Systems and Rehabilitation Engineering. <https://doi.org/10.1109/TNSRE.2016.2613412>
- Quinn, A. J., Lopes-dos-Santos, V., Huang, N., Liang, W.-K., Juan, C.-H., Yeh, J.-R., Nobre, A. C., Dupret, D., & Woolrich, M. W. (2021). Within-cycle instantaneous frequency profiles report oscillatory waveform dynamics. *Journal of Neurophysiology*, 126(4), 1190–1208. <https://doi.org/10.1152/jn.00201.2021>
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017). *Searching for Activation Functions* (No. arXiv:1710.05941). arXiv. <http://arxiv.org/abs/1710.05941>
- Ravanelli, M., & Bengio, Y. (2019). *Speaker Recognition from Raw Waveform with SincNet* (No. arXiv:1808.00158). arXiv. <http://arxiv.org/abs/1808.00158>
- Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R. P., De Berker, A., Ganguli, S., Gillon, C. J., Hafner, D., Kepecs, A., Kriegeskorte, N., Latham, P., Lindsay, G. W., Miller, K. D., Naud, R., Pack, C. C., ... Kording, K. P. (2019). A deep learning framework for neuroscience. *Nature Neuroscience*, 22(11), 1761–1770. <https://doi.org/10.1038/s41593-019-0520-2>
- Rodriguez, F., He, S., & Tan, H. (2023). The potential of convolutional neural networks for identifying neural states based on electrophysiological signals: Experiments on synthetic and real patient data. *Frontiers in Human Neuroscience*, 17, 1134599. <https://doi.org/10.3389/fnhum.2023.1134599>
- Rolando, L. (1809). *Saggio sopra la vera struttura del cervello dell'uomo e degl'animali e sopra le funzioni del sistema nervoso di Luigi Rolando..* (No. 2). nella stamperia da SSRM privilegiata.

- Rosa, M., Arlotti, M., Marceglia, S., Cogiamanian, F., Ardolino, G., Fonzo, A. D., Lopiano, L., Scelzo, E., Merola, A., Locatelli, M., Rampini, P. M., & Priori, A. (2017). Adaptive deep brain stimulation controls levodopa-induced side effects in Parkinsonian patients. *Movement Disorders*, 32(4), 628–629. <https://doi.org/10.1002/mds.26953>
- Rosa, M., Scelzo, E., Locatelli, M., Carrabba, G., Levi, V., Arlotti, M., Barbieri, S., Rampini, P., & Priori, A. (2017). Risk of Infection After Local Field Potential Recording from Externalized Deep Brain Stimulation Leads in Parkinson's Disease. *World Neurosurgery*, 97, 64–69. <https://doi.org/10.1016/j.wneu.2016.09.069>
- Rosenblatt, J. D., Benjamini, Y., Gilron, R., Mukamel, R., & Goeman, J. J. (2021). Better-than-chance classification for signal detection. *Biostatistics*, 22(2), 365–380. <https://doi.org/10.1093/biostatistics/kxz035>
- Ruiz-Gómez, S. J., Hornero, R., Poza, J., Maturana-Candelas, A., Pinto, N., & Gómez, C. (2019). Computational modeling of the effects of EEG volume conduction on functional connectivity metrics. Application to Alzheimer's disease continuum. *Journal of Neural Engineering*, 16(6), 066019. <https://doi.org/10.1088/1741-2552/ab4024>
- Runge, J., Nagel, J. M., Blahak, C., Kinfe, T. M., Heissler, H. E., Schrader, C., Wolf, M. E., Saryyeva, A., & Krauss, J. K. (2024). Does Temporary Externalization of Electrodes After Deep Brain Stimulation Surgery Result in a Higher Risk of Infection? *Neuromodulation: Technology at the Neural Interface*, 27(3), 565–571. <https://doi.org/10.1016/j.neurom.2023.08.004>
- Rusheen, A. E., Rojas-Cabrera, J., Goyal, A., Shin, H., Yuen, J., Jang, D.-P., Bennet, K. E., Blaha, C. D., Lee, K. H., & Oh, Y. (2023). Deep brain stimulation alleviates tics in Tourette syndrome via striatal dopamine transmission. *Brain*, 146(10), 4174–4190. <https://doi.org/10.1093/brain/awad142>

- Rusz, J., Dusek, P., Tykalova, T., Novotny, M., Illner, V., Simek, M., Kouba, T., Kryze, P., Zogala, D., Ruzicka, E., Sousa, M., Jorge, A., Nef, T., & Krack, P. (2024). Is speech function lateralised in the basal ganglia? Evidence from de novo Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp-2024-334297. <https://doi.org/10.1136/jnnp-2024-334297>
- Rutten, G.-J. (2022). Broca-Wernicke theories: A historical perspective. In *Handbook of Clinical Neurology* (Vol. 185, pp. 25–34). Elsevier. <https://doi.org/10.1016/B978-0-12-823384-9.00001-3>
- Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K.-R. (Eds.). (2019). *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (Vol. 11700). Springer International Publishing. <https://doi.org/10.1007/978-3-030-28954-6>
- Sani, O. G., Yang, Y., Lee, M. B., Dawes, H. E., Chang, E. F., & Shanechi, M. M. (2018). Mood variations decoded from multi-site intracranial human brain activity. *Nature Biotechnology*, 36(10), 954–961. <https://doi.org/10.1038/nbt.4200>
- Santens, P., De Letter, M., Van Borsel, J., De Reuck, J., & Caemaert, J. (2003). Lateralized effects of subthalamic nucleus stimulation on different aspects of speech in Parkinson's disease. *Brain and Language*, 87(2), 253–258. [https://doi.org/10.1016/S0093-934X\(03\)00142-1](https://doi.org/10.1016/S0093-934X(03)00142-1)
- Scangos, K. W., Khambhati, A. N., Daly, P. M., Makhoul, G. S., Sugrue, L. P., Zamanian, H., Liu, T. X., Rao, V. R., Sellers, K. K., Dawes, H. E., Starr, P. A., Krystal, A. D., & Chang, E. F. (2021). Closed-loop neuromodulation in an individual with treatment-resistant depression. *Nature Medicine*, 27(10), 1696–1700. <https://doi.org/10.1038/s41591-021-01480-w>
- Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., & Ball, T. (2017). Deep learning with

- convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 38(11), 5391–5420. <https://doi.org/10.1002/hbm.23730>
- Schrock, L. E., Mink, J. W., Woods, D. W., Porta, M., Servello, D., Visser-Vandewalle, V., Silburn, P. A., Foltynie, T., Walker, H. C., Shahed-Jimenez, J., Savica, R., Klassen, B. T., Machado, A. G., Foote, K. D., Zhang, J., Hu, W., Ackermans, L., Temel, Y., Mari, Z., ... Tourette Syndrome Association International Deep Brain Stimulation (DBS) Database and Registry Study Group. (2015). Tourette syndrome deep brain stimulation: A review and updated recommendations. *Movement Disorders*, 30(4), 448–471. <https://doi.org/10.1002/mds.26094>
- Schüller, T., Huys, D., Kohl, S., Visser-Vandewalle, V., Dembek, T. A., Kuhn, J., Baldermann, J. C., & Smith, E. E. (2024). Thalamic deep brain stimulation for tourette syndrome increases cortical beta activity. *Brain Stimulation*, 17(2), 197–201. <https://doi.org/10.1016/j.brs.2024.01.011>
- Schwalb, J. M., & Hamani, C. (2008). The History and Future of Deep Brain Stimulation. *Neurotherapeutics*, 5(1), 3–13. <https://doi.org/10.1016/j.nurt.2007.11.003>
- Schwalb, J. M., & Lozano, A. M. (2004). Surgical Management of Tremor: *Neurosurgery Quarterly*, 14(1), 60–68. <https://doi.org/10.1097/00013414-200403000-00010>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2020). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *International Journal of Computer Vision*, 128(2), 336–359. <https://doi.org/10.1007/s11263-019-01228-7>
- Sem-Jacobsen, C. W. (2009). Depth electrographic stimulation and treatment of patients with Parkinson's disease including neurosurgical technique. *Acta Neurologica Scandinavica*, 41(S13), 365–376. <https://doi.org/10.1111/j.1600-0404.1965.tb01899.x>

- Serrien, D. J. (2004). Motor inhibition in patients with Gilles de la Tourette syndrome: Functional activation patterns as revealed by EEG coherence. *Brain*, 128(1), 116–125.
<https://doi.org/10.1093/brain/awh318>
- Servello, D., Sassi, M., Brambilla, A., Defendi, S., & Porta, M. (2010). Long-Term, Post-Deep Brain Stimulation Management of a Series of 36 Patients Affected With Refractory Gilles de la Tourette Syndrome. *Neuromodulation: Technology at the Neural Interface*, 13(3), 187–194.
<https://doi.org/10.1111/j.1525-1403.2009.00253.x>
- Shah, A., Nguyen, T.-A. K., Peterman, K., Khawaldeh, S., Debove, I., Shah, S. A., Torrecillos, F., Tan, H., Pogosyan, A., Lachenmayer, M. L., Michelis, J., Brown, P., Pollo, C., Krack, P., Nowacki, A., & Tinkhauser, G. (2023). Combining Multimodal Biomarkers to Guide Deep Brain Stimulation Programming in Parkinson Disease. *Neuromodulation: Technology at the Neural Interface*, 26(2), 320–332. <https://doi.org/10.1016/j.neurom.2022.01.017>
- Shah, S. A., Tan, H., Tinkhauser, G., & Brown, P. (2018). Towards Real-Time, Continuous Decoding of Gripping Force From Deep Brain Local Field Potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 26(7), 1460–1468. <https://doi.org/10.1109/TNSRE.2018.2837500>
- Shah, S. A., Tinkhauser, G., Chen, C. C., Little, S., & Brown, P. (2018). Parkinsonian Tremor Detection from Subthalamic Nucleus Local Field Potentials for Closed-Loop Deep Brain Stimulation. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2320–2324. <https://doi.org/10.1109/EMBC.2018.8512741>
- Shale, H., Fahn, S., & Mayeux, R. (1986). Tics in a patient with Parkinson's disease. *Movement Disorders*, 1(1), 79–83. <https://doi.org/10.1002/mds.870010111>

- Shapiro, A. K., & Shapiro, E. (1968). Treatment of Gilles de la Tourette's Syndrome with Haloperidol. *British Journal of Psychiatry*, 114(508), 345–350. <https://doi.org/10.1192/bjp.114.508.345>
- Sheth, S. A., Bijanki, K. R., Metzger, B., Allawala, A., Pirtle, V., Adkinson, J. A., Myers, J., Mathura, R. K., Oswalt, D., Tsolaki, E., Xiao, J., Noecker, A., Strutt, A. M., Cohn, J. F., McIntyre, C. C., Mathew, S. J., Borton, D., Goodman, W., & Pouratian, N. (2022). Deep Brain Stimulation for Depression Informed by Intracranial Recordings. *Biological Psychiatry*, 92(3), 246–251. <https://doi.org/10.1016/j.biopsych.2021.11.007>
- Shirvalkar, P., Prosky, J., Chin, G., Ahmadipour, P., Sani, O. G., Desai, M., Schmitgen, A., Dawes, H., Shanechi, M. M., Starr, P. A., & Chang, E. F. (2023). First-in-human prediction of chronic pain state using intracranial neural biomarkers. *Nature Neuroscience*, 26(6), 1090–1099. <https://doi.org/10.1038/s41593-023-01338-z>
- Shoeibi, A., Sadeghi, D., Moridian, P., Ghassemi, N., Heras, J., Alizadehsani, R., Khadem, A., Kong, Y., Nahavandi, S., Zhang, Y.-D., & Gorriz, J. M. (2021). Automatic Diagnosis of Schizophrenia in EEG Signals Using CNN-LSTM Models. *Frontiers in Neuroinformatics*, 15, 777977. <https://doi.org/10.3389/fninf.2021.777977>
- Shreve, L. A., Velisar, A., Malekmohammadi, M., Koop, M. M., Trager, M., Quinn, E. J., Hill, B. C., Blumenfeld, Z., Kilbane, C., Mantovani, A., Henderson, J. M., & Brontë-Stewart, H. (2017). Subthalamic oscillations and phase amplitude coupling are greater in the more affected hemisphere in Parkinson's disease. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 128(1), 128–137. <https://doi.org/10.1016/j.clinph.2016.10.095>
- Siegfried, J., & Lippitz, B. (1994). Bilateral Chronic Electrostimulation of Ventroposterolateral Pallidum. *Neurosurgery*, 35(6), 1126–1130. <https://doi.org/10.1227/00006123-199412000-00016>

- Silveri, M. C. (2021). Contribution of the Cerebellum and the Basal Ganglia to Language Production: Speech, Word Fluency, and Sentence Construction—Evidence from Pathology. *The Cerebellum*, 20(2), 282–294. <https://doi.org/10.1007/s12311-020-01207-6>
- Singer, H. S., Szymanski, S., Giuliano, J., Yokoi, F., Dogan, A. S., Brasic, J. R., Zhou, Y., Grace, A. A., & Wong, D. F. (2002). Elevated Intrasynaptic Dopamine Release in Tourette's Syndrome Measured by PET. *American Journal of Psychiatry*, 159(8), 1329–1336. <https://doi.org/10.1176/appi.ajp.159.8.1329>
- Sironi, V. A. (2011). Origin and Evolution of Deep Brain Stimulation. *Frontiers in Integrative Neuroscience*, 5. <https://doi.org/10.3389/fnint.2011.00042>
- Skodda, S. (2012). Effect of Deep Brain Stimulation on Speech Performance in Parkinson's Disease. *Parkinson's Disease*, 2012, 1–10. <https://doi.org/10.1155/2012/850596>
- Speedie, L. J., Wertman, E., Ta'ir, J., & Heilman, K. M. (1993). Disruption of automatic speech following a right basal ganglia lesion. *Neurology*, 43(9), 1768–1768. <https://doi.org/10.1212/WNL.43.9.1768>
- Spiegel, E. A., Wycis, H. T., Marks, M., & Lee, A. J. (1947). Stereotaxic Apparatus for Operations on the Human Brain. *Science*, 106(2754), 349–350. <https://doi.org/10.1126/science.106.2754.349>
- Stefani, A., Lozano, A. M., Peppe, A., Stanzione, P., Galati, S., Tropepi, D., Pierantozzi, M., Brusa, L., Scarnati, E., & Mazzone, P. (2007). Bilateral deep brain stimulation of the pedunculopontine and subthalamic nuclei in severe Parkinson's disease. *Brain*, 130(6), 1596–1607. <https://doi.org/10.1093/brain/awl346>

- Stern, E., Silbersweig, D. A., Chee, K.-Y., Holmes, A., Robertson, M. M., Trimble, M., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (2000). A Functional Neuroanatomy of Tics in Tourette Syndrome. *ARCH GEN PSYCHIATRY*, 57.
- Sullivan, C. R. P., Olsen, S., & Widge, A. S. (2021). Deep brain stimulation for psychiatric disorders: From focal brain targets to cognitive networks. *NeuroImage*, 225, 117515.
<https://doi.org/10.1016/j.neuroimage.2020.117515>
- Sun, R., Sohrabpour, A., Worrell, G. A., & He, B. (2022). Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics. *Proceedings of the National Academy of Sciences*, 119(31), e2201128119.
<https://doi.org/10.1073/pnas.2201128119>
- Sun, Y., Farzan, F., Garcia Dominguez, L., Barr, M. S., Giacobbe, P., Lozano, A. M., Wong, W., & Daskalakis, Z. J. (2014). A novel method for removal of deep brain stimulation artifact from electroencephalography. *Journal of Neuroscience Methods*, 237, 33–40.
<https://doi.org/10.1016/j.jneumeth.2014.09.002>
- Supakar, R., Satvaya, P., & Chakrabarti, P. (2022). A deep learning based model using RNN-LSTM for the Detection of Schizophrenia from EEG data. *Computers in Biology and Medicine*, 151, 106225. <https://doi.org/10.1016/j.combiomed.2022.106225>
- Sure, M., Mertiens, S., Vesper, J., Schnitzler, A., & Florin, E. (2023). Alterations of resting-state networks of Parkinson's disease patients after subthalamic DBS surgery. *NeuroImage: Clinical*, 37, 103317. <https://doi.org/10.1016/j.nicl.2023.103317>

- Swain, J. E., Scahill, L., Lombroso, P. J., King, R. A., & Leckman, J. F. (2007). Tourette Syndrome and Tic Disorders: A Decade of Progress. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(8), 947–968. <https://doi.org/10.1097/chi.0b013e318068fbcc>
- Swann, N. C., De Hemptinne, C., Miocinovic, S., Qasim, S., Ostrem, J. L., Galifianakis, N. B., Luciano, M. S., Wang, S. S., Ziman, N., Taylor, R., & Starr, P. A. (2018). Chronic multisite brain recordings from a totally implantable bidirectional neural interface: Experience in 5 patients with Parkinson's disease. *Journal of Neurosurgery*, 128(2), 605–616. <https://doi.org/10.3171/2016.11.JNS161162>
- Swann, N. C., De Hemptinne, C., Thompson, M. C., Miocinovic, S., Miller, A. M., Gilron, R., Ostrem, J. L., Chizeck, H. J., & Starr, P. A. (2018). Adaptive deep brain stimulation for Parkinson's disease using motor cortex sensing. *Journal of Neural Engineering*, 15(4), 046006. <https://doi.org/10.1088/1741-2552/aabc9b>
- Szurhaj, W., Labyt, E., Bourriez, J.-L., Kahane, P., Chauvel, P., Mauguière, F., & Derambure, P. (2006). Relationship between intracerebral gamma oscillations and slow potentials in the human sensorimotor cortex. *The European Journal of Neuroscience*, 24(3), 947–954. <https://doi.org/10.1111/j.1460-9568.2006.04876.x>
- Tan, H., Debarros, J., He, S., Pogosyan, A., Aziz, T. Z., Huang, Y., Wang, S., Timmermann, L., Visser-Vandewalle, V., Pedrosa, D. J., Green, A. L., & Brown, P. (2019). Decoding voluntary movements and postural tremor based on thalamic LFPs as a basis for closed-loop stimulation for essential tremor. *Brain Stimulation*, 12(4), 858–867. <https://doi.org/10.1016/j.brs.2019.02.011>
- Tan, H., Pogosyan, A., Anzak, A., Ashkan, K., Bogdanovic, M., Green, A. L., Aziz, T., Foltynie, T., Limousin, P., Zrinzo, L., & Brown, P. (2013). Complementary roles of different oscillatory

- activities in the subthalamic nucleus in coding motor effort in Parkinsonism. *Experimental Neurology*, 248, 187. <https://doi.org/10.1016/j.expneurol.2013.06.010>
- Tan, H., Pogosyan, A., Ashkan, K., Green, A. L., Aziz, T., Foltynie, T., Limousin, P., Zrinzo, L., Hariz, M., & Brown, P. (2016). Decoding gripping force based on local field potentials recorded from subthalamic nucleus in humans. *eLife*, 5, e19089. <https://doi.org/10.7554/eLife.19089>
- Tanaka, Y., Tsuboi, T., Watanabe, H., Kajita, Y., Fujimoto, Y., Ohdake, R., Yoneyama, N., Masuda, M., Hara, K., Senda, J., Ito, M., Atsuta, N., Horiguchi, S., Yamamoto, M., Wakabayashi, T., & Sobue, G. (2015). Voice features of Parkinson's disease patients with subthalamic nucleus deep brain stimulation. *Journal of Neurology*, 262(5), 1173–1181. <https://doi.org/10.1007/s00415-015-7681-z>
- Tanaka, Y., Tsuboi, T., Watanabe, H., Nakatsubo, D., Maesawa, S., Kato, S., Kajita, Y., Sato, M., Oodake, R., Hattori, M., Yamamoto, M., Wakabayashi, T., Katsuno, M., & Sobue, G. (2020). Longitudinal Speech Change After Subthalamic Nucleus Deep Brain Stimulation in Parkinson's Disease Patients: A 2-Year Prospective Study. *Journal of Parkinson's Disease*, 10(1), 131–140. <https://doi.org/10.3233/JPD-191798>
- Tankus, A., & Fried, I. (2019). Degradation of Neuronal Encoding of Speech in the Subthalamic Nucleus in Parkinson's Disease. *Neurosurgery*, 84(2), 378–387. <https://doi.org/10.1093/neuros/nyy027>
- Thenaisie, Y., Lee, K., Moerman, C., Scafa, S., Gálvez, A., Pirondini, E., Burri, M., Ravier, J., Puiatti, A., Accolla, E., Wicki, B., Zacharia, A., Castro Jiménez, M., Bally, J. F., Courtine, G., Bloch, J., & Moraud, E. M. (2022). Principles of gait encoding in the subthalamic nucleus

of people with Parkinson's disease. *Science Translational Medicine*, 14(661), eabo1800.

<https://doi.org/10.1126/scitranslmed.abo1800>

Thenaisie, Y., Palmisano, C., Canessa, A., Keulen, B. J., Capetian, P., Jiménez, M. C., Bally, J. F., Manferlotti, E., Beccaria, L., Zutt, R., Courtine, G., Bloch, J., Van Der Gaag, N. A., Hoffmann, C. F., Moraud, E. M., Isaias, I. U., & Contarino, M. F. (2021). Towards adaptive deep brain stimulation: Clinical and technical notes on a novel commercial device for chronic brain sensing. *Journal of Neural Engineering*, 18(4), 042002. <https://doi.org/10.1088/1741-2552/ac1d5b>

Thuberg, D., Buentjen, L., Holtkamp, M., Voges, J., Heinze, H.-J., Lee, H., Kitay, A.-Y., & Schmitt, F. C. (2021). Deep Brain Stimulation for Refractory Focal Epilepsy: Unraveling the Insertional Effect up to Five Months Without Stimulation. *Neuromodulation: Technology at the Neural Interface*, 24(2), 373–379. <https://doi.org/10.1111/ner.13349>

Tinkhauser, G., Pogosyan, A., Debove, I., Nowacki, A., Shah, S. A., Seidel, K., Tan, H., Brittain, J., Petermann, K., Di Biase, L., Oertel, M., Pollo, C., Brown, P., & Schuepbach, M. (2018). Directional local field potentials: A tool to optimize deep brain stimulation. *Movement Disorders*, 33(1), 159–164. <https://doi.org/10.1002/mds.27215>

Tinkhauser, G., Pogosyan, A., Debove, I., Nowacki, A., Shah, S. A., Seidel, K., Tan, H., Brittain, J.-S., Petermann, K., Biase, L. di, Oertel, M., Pollo, C., Brown, P., & Schuepbach, M. (2017). Directional local field potentials: A tool to optimize deep brain stimulation. *Movement Disorders*, 33(1), 159. <https://doi.org/10.1002/mds.27215>

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>

- Tinkhauser, G., Shah, S. A., Fischer, P., Peterman, K., Debove, I., Nygyuen, K., Nowacki, A., Torrecillos, F., Khawaldeh, S., Tan, H., Pogosyan, A., Schuepbach, M., Pollo, C., & Brown, P. (2019). Electrophysiological differences between upper and lower limb movements in the human subthalamic nucleus. *Clinical Neurophysiology*, 130(5), 727–738.
<https://doi.org/10.1016/j.clinph.2019.02.011>
- Tinkhauser, G., Torrecillos, F., Duclos, Y., Tan, H., Pogosyan, A., Fischer, P., Carron, R., Welter, M.-L., Karachi, C., Vandenberghe, W., Nuttin, B., Witjas, T., Régis, J., Azulay, J.-P., Eusebio, A., & Brown, P. (2018). Beta burst coupling across the motor circuit in Parkinson's disease. *Neurobiology of Disease*, 117, 217–225. <https://doi.org/10.1016/j.nbd.2018.06.007>
- Törnqvist, A. L., Schalén, L., & Rehnström, S. (2005). Effects of different electrical parameter settings on the intelligibility of speech in patients with Parkinson's disease treated with subthalamic deep brain stimulation. *Movement Disorders*, 20(4), 416–423.
<https://doi.org/10.1002/mds.20348>
- Tort, A. B. L., Komorowski, R., Eichenbaum, H., & Kopell, N. (2010a). Measuring Phase-Amplitude Coupling Between Neuronal Oscillations of Different Frequencies. *Journal of Neurophysiology*, 104(2), 1195–1210. <https://doi.org/10.1152/jn.00106.2010>
- Tort, A. B. L., Komorowski, R., Eichenbaum, H., & Kopell, N. (2010b). Measuring Phase-Amplitude Coupling Between Neuronal Oscillations of Different Frequencies. *Journal of Neurophysiology*, 104(2), 1195. <https://doi.org/10.1152/jn.00106.2010>
- Tóth, S., & Tomka, I. (1968). Responses of the Human Thalamus and Pallidum to High Frequency Stimulations. *Stereotactic and Functional Neurosurgery*, 30(1), 17–40.
<https://doi.org/10.1159/000103517>

- Tremblay, P., Deschamps, I., & Gracco, V. L. (2016). Neurobiology of Speech Production. In *Neurobiology of Language* (pp. 741–750). Elsevier. <https://doi.org/10.1016/B978-0-12-407794-2.00059-6>
- Tripoliti, E., Limousin, P., Foltynie, T., Candelario, J., Aviles-Olmos, I., Hariz, M. I., & Zrinzo, L. (2014). Predictive factors of speech intelligibility following subthalamic nucleus stimulation in consecutive patients with Parkinson’s disease. *Movement Disorders*, 29(4), 532–538. <https://doi.org/10.1002/mds.25816>
- Tripoliti, E., Zrinzo, L., Martinez-Torres, I., Frost, E., Pinto, S., Foltynie, T., Holl, E., Petersen, E., Roughton, M., Hariz, M. I., & Limousin, P. (2011). Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease. *Neurology*, 76(1), 80–86. <https://doi.org/10.1212/WNL.0b013e318203e7d0>
- Tripoliti, E., Zrinzo, L., Martinez-Torres, I., Tisch, S., Frost, E., Borrell, E., Hariz, M. I., & Limousin, P. (2008). Effects of contact location and voltage amplitude on speech and movement in bilateral subthalamic nucleus deep brain stimulation. *Movement Disorders*, 23(16), 2377–2383. <https://doi.org/10.1002/mds.22296>
- Tsuboi, T., Watanabe, H., Tanaka, Y., Ohdake, R., Yoneyama, N., Hara, K., Nakamura, R., Watanabe, H., Senda, J., Atsuta, N., Ito, M., Hirayama, M., Yamamoto, M., Fujimoto, Y., Kajita, Y., Wakabayashi, T., & Sobue, G. (2015). Distinct phenotypes of speech and voice disorders in Parkinson’s disease after subthalamic nucleus deep brain stimulation. *Journal of Neurology, Neurosurgery & Psychiatry*, 86(8), 856–864. <https://doi.org/10.1136/jnnp-2014-308043>

- Tykalova, T., Novotny, M., Ruzicka, E., Dusek, P., & Rusz, J. (2022). Short-term effect of dopaminergic medication on speech in early-stage Parkinson's disease. *Npj Parkinson's Disease*, 8(1), 22. <https://doi.org/10.1038/s41531-022-00286-y>
- Valenstein, E. S. (1974). *Brain control*. Wiley.
- van Rheede, J. J., Feldmann, L. K., Busch, J. L., Fleming, J. E., Mathiopoulou, V., Denison, T., Sharott, A., & Kühn, A. A. (2022). Diurnal modulation of subthalamic beta oscillatory power in Parkinson's disease patients during deep brain stimulation. *NPJ Parkinson's Disease*, 8, 88. <https://doi.org/10.1038/s41531-022-00350-7>
- van Wijk, B. C. M., Beudel, M., Jha, A., Oswal, A., Foltynie, T., Hariz, M. I., Limousin, P., Zrinzo, L., Aziz, T. Z., Green, A. L., Brown, P., & Litvak, V. (2016). Subthalamic nucleus phase-amplitude coupling correlates with motor impairment in Parkinson's disease. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 127(4), 2010–2019. <https://doi.org/10.1016/j.clinph.2016.01.015>
- Vandewalle, V., Van Der Linden, C., Groenewegen, H., & Caemaert, J. (1999). Stereotactic treatment of Gilles de la Tourette syndrome by high frequency stimulation of thalamus. *The Lancet*, 353(9154), 724. [https://doi.org/10.1016/S0140-6736\(98\)05964-9](https://doi.org/10.1016/S0140-6736(98)05964-9)
- Velisar, A., Syrkin-Nikolau, J., Blumenfeld, Z., Trager, M. H., Afzal, M. F., Prabhakar, V., & Bronte-Stewart, H. (2019). Dual threshold neural closed loop deep brain stimulation in Parkinson disease patients. *Brain Stimulation*, 12(4), 868–876. <https://doi.org/10.1016/j.brs.2019.02.020>
- Verma, V., Lamb, A., Beckham, C., Najafi, A., Mitliagkas, I., Courville, A., Lopez-Paz, D., & Bengio, Y. (2019). *Manifold Mixup: Better Representations by Interpolating Hidden States* (No. arXiv:1806.05236). arXiv. <http://arxiv.org/abs/1806.05236>

- Vernaleken, I., Kuhn, J., Lenartz, D., Raptis, M., Huff, W., Janouschek, H., Neuner, I., Schaefer, W. M., Gründer, G., & Sturm, V. (2009). Bithalamical Deep Brain Stimulation in Tourette Syndrome Is Associated with Reduction in Dopaminergic Transmission. *Biological Psychiatry*, 66(10), e15–e17. <https://doi.org/10.1016/j.biopsych.2009.06.025>
- Vidailhet, M., Vercueil, L., Houeto, J.-L., Krystkowiak, P., Benabid, A.-L., Cornu, P., Lagrange, C., Tézenas Du Montcel, S., Dormont, D., Grand, S., Blond, S., Detante, O., Pilon, B., Ardouin, C., Agid, Y., Destée, A., & Pollak, P. (2005). Bilateral Deep-Brain Stimulation of the Globus Pallidus in Primary Generalized Dystonia. *New England Journal of Medicine*, 352(5), 459–467. <https://doi.org/10.1056/NEJMoa042187>
- Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48), 12827–12832. <https://doi.org/10.1073/pnas.1705120114>
- Vilensky, J. A., & Gilman, S. (2002). Lesions of the precentral gyrus in nonhuman primates: A pre-Medline bibliography. *International Journal of Primatology*, 23, 1319–1333.
- Vissani, M., Bush, A., Lipski, W. J., Bullock, L., Fischer, P., Neudorfer, C., Holt, L. L., Fiez, J. A., Turner, R. S., & Richardson, R. M. (2023). *Spike-phase coupling of subthalamic neurons to posterior opercular cortex predicts speech sound accuracy*. <https://doi.org/10.1101/2023.10.18.562969>
- Visser-Vandewalle, V., Andrade, P., Mosley, P. E., Greenberg, B. D., Schuurman, R., McLaughlin, N. C., Voon, V., Krack, P., Foote, K. D., Mayberg, H. S., Figee, M., Kopell, B. H., Polosan, M., Joyce, E. M., Chabardes, S., Matthews, K., Baldermann, J. C., Tyagi, H., Holtzheimer, P. E., ... Okun, M. S. (2022). Deep brain stimulation for obsessive–compulsive disorder: A crisis of access. *Nature Medicine*, 28(8), 1529–1532. <https://doi.org/10.1038/s41591-022-01879-z>

- Visser-Vandewalle, V., Temel, Y., Boon, P., Vreeling, F., Colle, H., Hoogland, G., Groenewegen, H. J., & Van Der Linden, C. (2003). Chronic bilateral thalamic stimulation: A new therapeutic approach in intractable Tourette syndrome: Report of three cases. *Journal of Neurosurgery*, 99(6), 1094–1100. <https://doi.org/10.3171/jns.2003.99.6.1094>
- Vivekananda, U., Bush, D., Bisby, J. A., Baxendale, S., Rodionov, R., Diehl, B., Chowdhury, F. A., McEvoy, A. W., Miserocchi, A., Walker, M. C., & Burgess, N. (2021). Theta power and theta-gamma coupling support long-term spatial memory retrieval. *Hippocampus*, 31(2), 213–220. <https://doi.org/10.1002/hipo.23284>
- Voytek, B., & Knight, R. T. (2015). Dynamic Network Communication as a Unifying Neural Basis for Cognition, Development, Aging, and Disease. *Biological Psychiatry*, 77(12), 1089–1097. <https://doi.org/10.1016/j.biopsych.2015.04.016>
- Wang, E. Q., Metman, L. V., Bakay, R. A. E., Arzbaecher, J., Bernard, B., & Corcos, D. M. (2006). Hemisphere-Specific Effects of Subthalamic Nucleus Deep Brain Stimulation on Speaking Rate and Articulatory Accuracy of Syllable Repetitions in Parkinson’s Disease. *Journal of Medical Speech-Language Pathology*, 14(4), 323–334.
- Wang, F., Zhang, X., Hu, F., Li, F., & Tian, Y.-C. (2019). Quantitative Electroencephalography Analysis for Improved Assessment of Consciousness in Cerebral Hemorrhage and Ischemic Stroke Patients. *IEEE Access*, 7, 63674–63685. <https://doi.org/10.1109/ACCESS.2019.2916165>
- Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., Mardziel, P., & Hu, X. (2020). Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks (No. arXiv:1910.01279). arXiv. <http://arxiv.org/abs/1910.01279>

- Wang, S., Zhu, G., Shi, L., Zhang, C., Wu, B., Yang, A., Meng, F., Jiang, Y., & Zhang, J. (2023). Closed-Loop Adaptive Deep Brain Stimulation in Parkinson's Disease: Procedures to Achieve It and Future Perspectives. *Journal of Parkinson's Disease*, 13(4), 453–471.
<https://doi.org/10.3233/JPD-225053>
- Wang, X.-J. (2010). Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews*, 90(3), 1195–1268.
<https://doi.org/10.1152/physrev.00035.2008>
- Wang, Z., & Oates, T. (2015). *Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks*. 7.
- Watson, P., & Montgomery, E. B. (2006). The relationship of neuronal activity within the sensorimotor region of the subthalamic nucleus to speech. *Brain and Language*, 97(2), 233–240.
<https://doi.org/10.1016/j.bandl.2005.11.004>
- Watts, J., Khojandi, A., Shylo, O., & Ramdhani, R. A. (2020). Machine Learning's Application in Deep Brain Stimulation for Parkinson's Disease: A Review. *Brain Sciences*, 10(11), 809.
<https://doi.org/10.3390/brainsci10110809>
- Weaver, F. M. (2009). Bilateral Deep Brain Stimulation vs Best Medical Therapy for Patients With Advanced Parkinson DiseaseA Randomized Controlled Trial. *JAMA*, 301(1), 63.
<https://doi.org/10.1001/jama.2008.929>
- Wehmeyer, L., Baldermann, J. C., Pogosyan, A., Plazas, F. R., Loehrer, P. A., Bonetti, L., Yassine, S., Mühlen, K. Z., Schüller, T., Kuhn, J., Visser-Vandewalle, V., Tan, H., & Andrade, P. (2024). *Thalamo-frontal connectivity patterns in Tourette Syndrome: Insights from combined intracranial DBS and EEG recordings*. <https://doi.org/10.1101/2024.10.09.617413>

- Wehmeyer, L., Schüller, T., Kiess, J., Heiden, P., Visser-Vandewalle, V., Baldermann, J. C., & Andrade, P. (2021). Target-Specific Effects of Deep Brain Stimulation for Tourette Syndrome: A Systematic Review and Meta-Analysis. *Frontiers in Neurology*, 12, 769275.
<https://doi.org/10.3389/fneur.2021.769275>
- Wei, Q., Wang, Y., Gao, X., & Gao, S. (2007). Amplitude and phase coupling measures for feature extraction in an EEG-based brain–computer interface. *Journal of Neural Engineering*, 4(2), 120–129. <https://doi.org/10.1088/1741-2560/4/2/012>
- Weiss, A. R., Korzeniewska, A., Chrabaszcz, A., Bush, A., Fiez, J. A., Crone, N. E., & Richardson, R. M. (2023). Lexicality-Modulated Influence of Auditory Cortex on Subthalamic Nucleus During Motor Planning for Speech. *Neurobiology of Language*, 4(1), 53–80.
https://doi.org/10.1162/nol_a_00086
- Weiss, J. M., Flesher, S. N., Franklin, R., Collinger, J. L., & Gaunt, R. A. (2019). Artifact-free recordings in human bidirectional brain–computer interfaces. *Journal of Neural Engineering*, 16(1), 016002. <https://doi.org/10.1088/1741-2552/aae748>
- Wendt, K., Denison, T., Foster, G., Krinke, L., Thomson, A., Wilson, S., & Widge, A. S. (2022). Physiologically informed neuromodulation. *Journal of the Neurological Sciences*, 434, 120121.
<https://doi.org/10.1016/j.jns.2021.120121>
- Wertheimer, J., Gottuso, A. Y., Nuno, M., Walton, C., Duboille, A., Tuchman, M., & Ramig, L. (2014). The impact of STN deep brain stimulation on speech in individuals with Parkinson’s disease: The patient’s perspective. *Parkinsonism & Related Disorders*, 20(10), 1065–1070.
<https://doi.org/10.1016/j.parkreldis.2014.06.010>

- Widge, A. S. (2024). Closing the loop in psychiatric deep brain stimulation: Physiology, psychometrics, and plasticity. *Neuropsychopharmacology*, 49(1), 138–149.
<https://doi.org/10.1038/s41386-023-01643-y>
- Wilkins, K. B., Melbourne, J. A., Akella, P., & Bronte-Stewart, H. M. (2023). Unraveling the complexities of programming neural adaptive deep brain stimulation in Parkinson's disease. *Frontiers in Human Neuroscience*, 17, 1310393. <https://doi.org/10.3389/fnhum.2023.1310393>
- Wong, J. K., Mayberg, H. S., Wang, D. D., Richardson, R. M., Halpern, C. H., Krinke, L., Arlotti, M., Rossi, L., Priori, A., Marceglia, S., Gilron, R., Cavanagh, J. F., Judy, J. W., Miocinovic, S., Devergnas, A. D., Sillitoe, R. V., Cernera, S., Oehr, C. R., Gunduz, A., ... Okun, M. S. (2023). Proceedings of the 10th annual deep brain stimulation think tank: Advances in cutting edge technologies, artificial intelligence, neuromodulation, neuroethics, interventional psychiatry, and women in neuromodulation. *Frontiers in Human Neuroscience*, 16, 1084782.
<https://doi.org/10.3389/fnhum.2022.1084782>
- Worbe, Y., Malherbe, C., Hartmann, A., Pélérini-Issac, M., Messé, A., Vidailhet, M., Lehericy, S., & Benali, H. (2012). Functional immaturity of cortico-basal ganglia networks in Gilles de la Tourette syndrome. *Brain*, 135(6), 1937–1946. <https://doi.org/10.1093/brain/aws056>
- Yamamoto, T., Katayama, Y., Kobayashi, K., Oshima, H., Fukaya, C., & Tsubokawa, T. (2010). Deep brain stimulation for the treatment of vegetative state. *European Journal of Neuroscience*, 32(7), 1145–1151. <https://doi.org/10.1111/j.1460-9568.2010.07412.x>
- Yamamoto, T., Katayama, Y., Ushiba, J., Yoshino, H., Obuchi, T., Kobayashi, K., Oshima, H., & Fukaya, C. (2013). On-Demand Control System for Deep Brain Stimulation for Treatment of Intention Tremor. *Neuromodulation: Technology at the Neural Interface*, 16(3), 230–235.
<https://doi.org/10.1111/j.1525-1403.2012.00521.x>

- Yeh, C.-H., Al-Fatly, B., Kühn, A. A., Meidahl, A. C., Tinkhauser, G., Tan, H., & Brown, P. (2020). Waveform changes with the evolution of beta bursts in the human subthalamic nucleus. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 131(9), 2086–2099. <https://doi.org/10.1016/j.clinph.2020.05.035>
- Yin, Z., Yu, H., Yuan, T., Smyth, C., Anjum, M. F., Zhu, G., Ma, R., Xu, Y., An, Q., Gan, Y., Merk, T., Qin, G., Xie, H., Zhang, N., Wang, C., Jiang, Y., Meng, F., Yang, A., Neumann, W.-J., ... Zhang, J. (2024). Generalized sleep decoding with basal ganglia signals in multiple movement disorders. *Npj Digital Medicine*, 7(1), 122. <https://doi.org/10.1038/s41746-024-01115-7>
- Yin, Z., Zhu, G., Liu, Y., Zhao, B., Liu, D., Bai, Y., Zhang, Q., Shi, L., Feng, T., Yang, A., Liu, H., Meng, F., Neumann, W. J., Kühn, A. A., Jiang, Y., & Zhang, J. (2022). Cortical phase-amplitude coupling is key to the occurrence and treatment of freezing of gait. *Brain*, 145(7), 2407. <https://doi.org/10.1093/brain/awac121>
- Yoo, H. K., Lee, J.-S., Paik, K.-W., Choi, S.-H., Yoon, S. J., Kim, J. E., & Hong, J. P. (2011). Open-label study comparing the efficacy and tolerability of aripiprazole and haloperidol in the treatment of pediatric tic disorders. *European Child & Adolescent Psychiatry*, 20(3), 127–135. <https://doi.org/10.1007/s00787-010-0154-0>
- Yuste, R. (2015). From the neuron doctrine to neural networks. *Nature Reviews Neuroscience*, 16(8), 487–497. <https://doi.org/10.1038/nrn3962>
- Yuste, R., Goering, S., Arcas, B. A. Y., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., ... Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, 551(7679), 159–163. <https://doi.org/10.1038/551159a>

Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2018). *mixup: Beyond Empirical Risk*

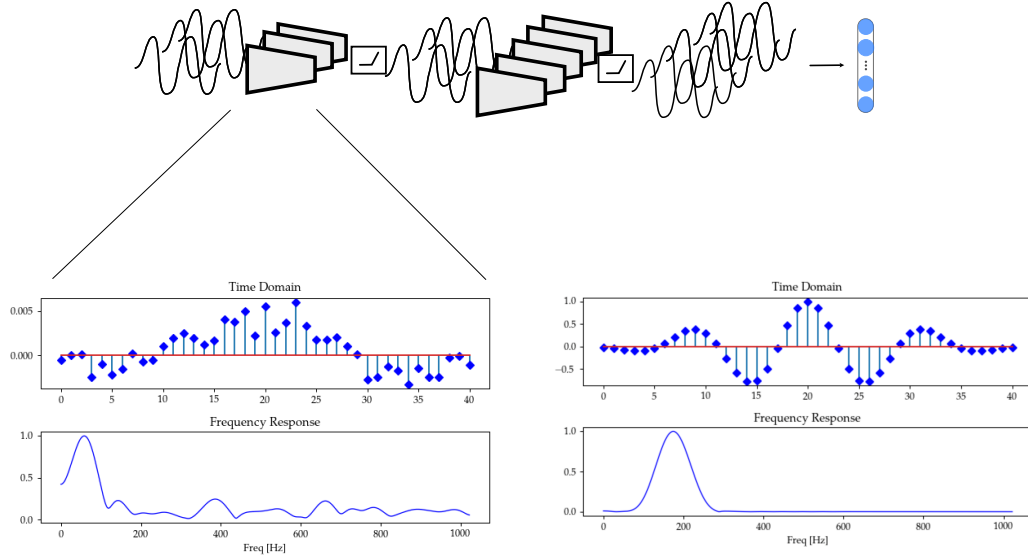
Minimization (No. arXiv:1710.09412). arXiv. <http://arxiv.org/abs/1710.09412>

Zhu, B., Shin, U., & Shoaran, M. (2020). Closed-Loop Neural Interfaces with Embedded Machine

Learning. *2020 27th IEEE International Conference on Electronics, Circuits and Systems*

(ICECS), 1–4. <https://doi.org/10.1109/ICECS49266.2020.9294844>

Supplementary Materials

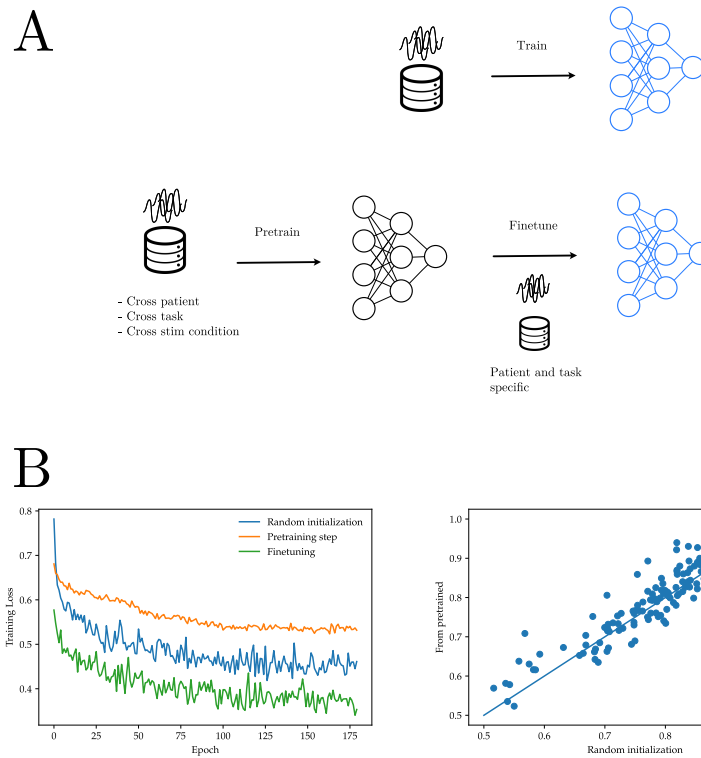


Supplementary Figure I: Conceptualization of Convolutional Filters as Finite Impulse Response (FIR) Filters.

This schematic illustrates how multichannel time-series signals are processed through a cascade of convolutional filters interleaved with nonlinear activation functions. In each convolutional layer, the filters consist of arrays of kernels that are convolved with the input signals. This convolution operation is mathematically equivalent to applying finite impulse response (FIR) filters to the input signals, where the kernel weights serve as the filter coefficients. As a result, the convolutional filters possess well-defined frequency and phase responses, allowing them to act as tuneable band-pass filters in the frequency domain.

By constraining the shape of these kernels, it is possible to influence various filter properties such as frequency selectivity, bandwidth, and phase characteristics. In the bottom right of the figure, we illustrate an approach where the kernel shapes are constrained to specific functional forms. Concretely, the kernel is made to follow the shape

of a sinc function⁹. This method is exemplified by the SincNet architecture proposed in (Ravanelli & Bengio, 2019) for automatic speech recognition applications.



Supplementary Figure II: Effect of Cross-Patient Network Pre-Training on Decoding Performance

(A) Two Approaches to Network Training: The figure illustrates two strategies for training neural decoding networks in the context of aDBS applications.

1. *Direct Individualized Training (Top):* In this approach, a neural network—such as a one-dimensional convolutional neural network (1D-CNN)—is trained exclusively on neural time-series signals and corresponding state labels from a single patient. This individualized training focuses on capturing patient-specific neural patterns associated with different physiological or cognitive states during specific tasks. The network learns to decode neural signals without any prior knowledge from other patients or tasks.
2. *Cross-Patient Pre-Training Followed by Fine-Tuning (Bottom):* This two-stage training strategy begins with **cross-patient pre-training**, where the network's parameters are initially trained on a large,

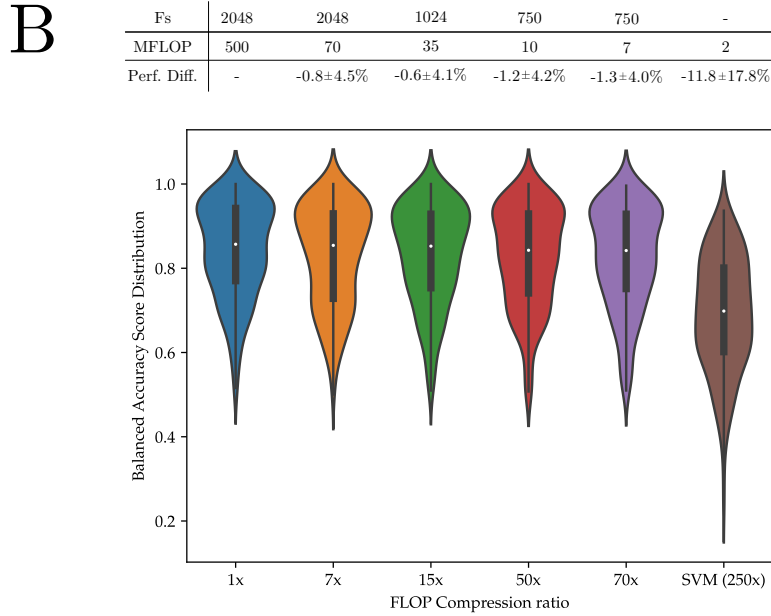
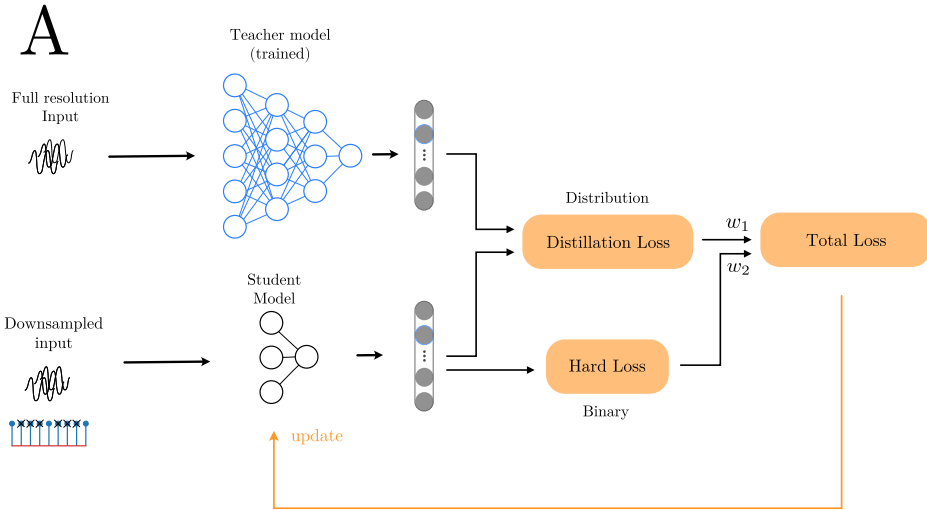
⁹ The sinc function is defined as $\text{sinc}(x) = \sin(x)/x$ for all $x \neq 0$. The Fourier transform of the sinc function is the rectangular function, making it a suitable time-domain function to implement bandpass filters in the frequency domain.

heterogeneous dataset comprising neural signals from multiple patients, tasks, and conditions. The aim is to learn generalizable neural features that are common across different individuals and activities. In the second stage, known as fine-tuning, the pre-trained network is further trained on the individualized patient data. This step allows the network to adapt the previously learned general features to the specific neural patterns of the individual patient, potentially enhancing decoding performance by harnessing both shared and unique neural characteristics.

(B) Impact of Cross-Patient Pre-Training on Decoding Performance:

Our experiments assessed whether the cross-patient pre-training strategy improves decoding performance compared to direct individualized training without pre-training. We conducted extensive evaluations using neural datasets recorded from patients with essential tremor performing various motor tasks.

The results indicate that cross-patient pre-training **does not lead to significant increases in performance** over models trained directly on individualized patient data. Both training approaches yielded comparable decoding accuracies, suggesting that the inclusion of cross-patient data during pre-training did not provide a noticeable advantage in this context.



Supplementary Figure III: Model Distillation for Neural Decoding.

(A) Model Distillation Pipeline:

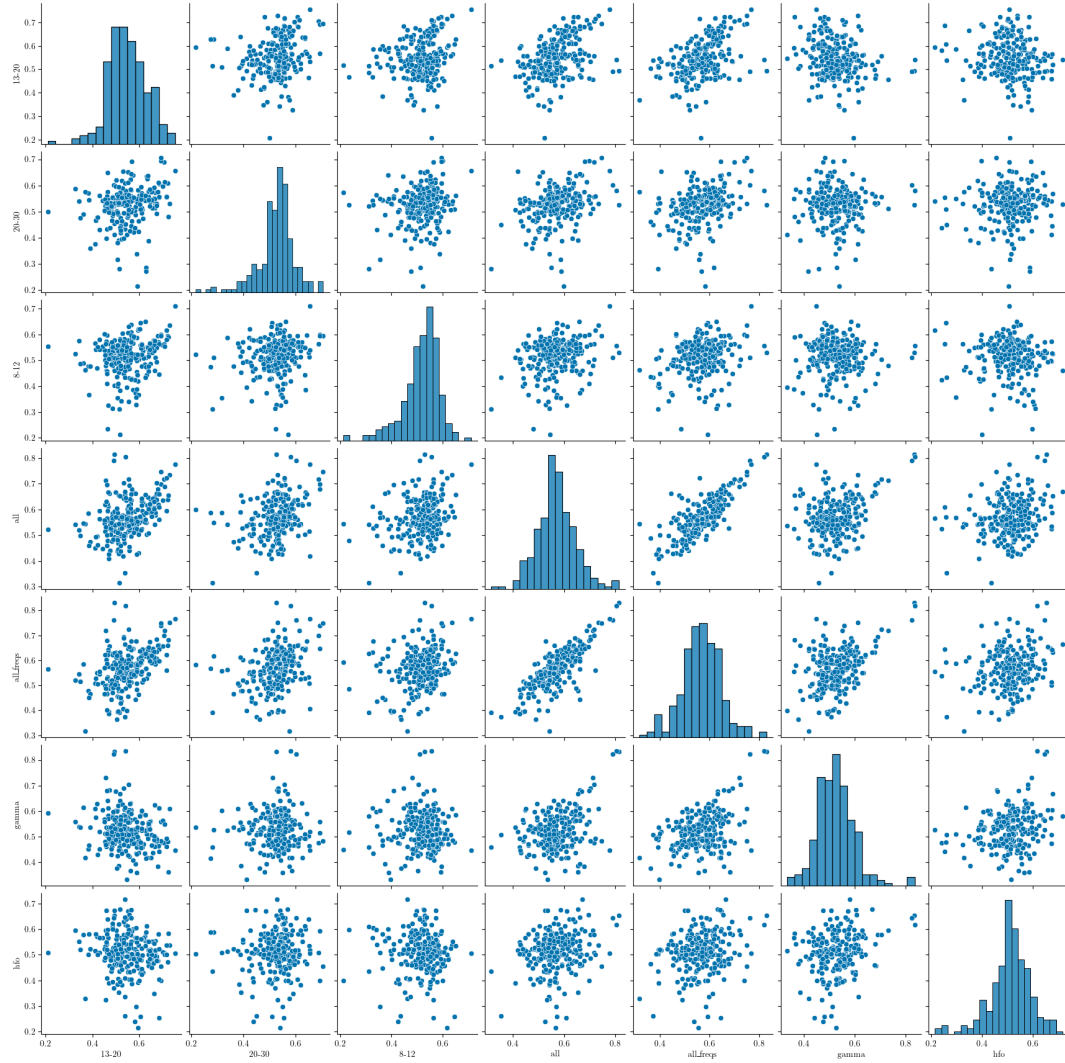
A schematic illustrating model distillation applied to neural decoding models. A large, pre-trained **teacher model** processes high-resolution neural signals and produces high-entropy outputs. A smaller, more efficient **student model** is trained on low-resolution (downsampled) signals. During training, the student model minimizes a combined loss function that includes:

- *Distillation Loss:* Aligning the student's outputs with the teacher's soft outputs.
- *Hard Learning Loss:* Matching the student's outputs to the ground truth labels.

This approach enables the student model to learn essential features from the teacher, achieving comparable performance with reduced computational complexity suitable for deployment on resource-constrained hardware.

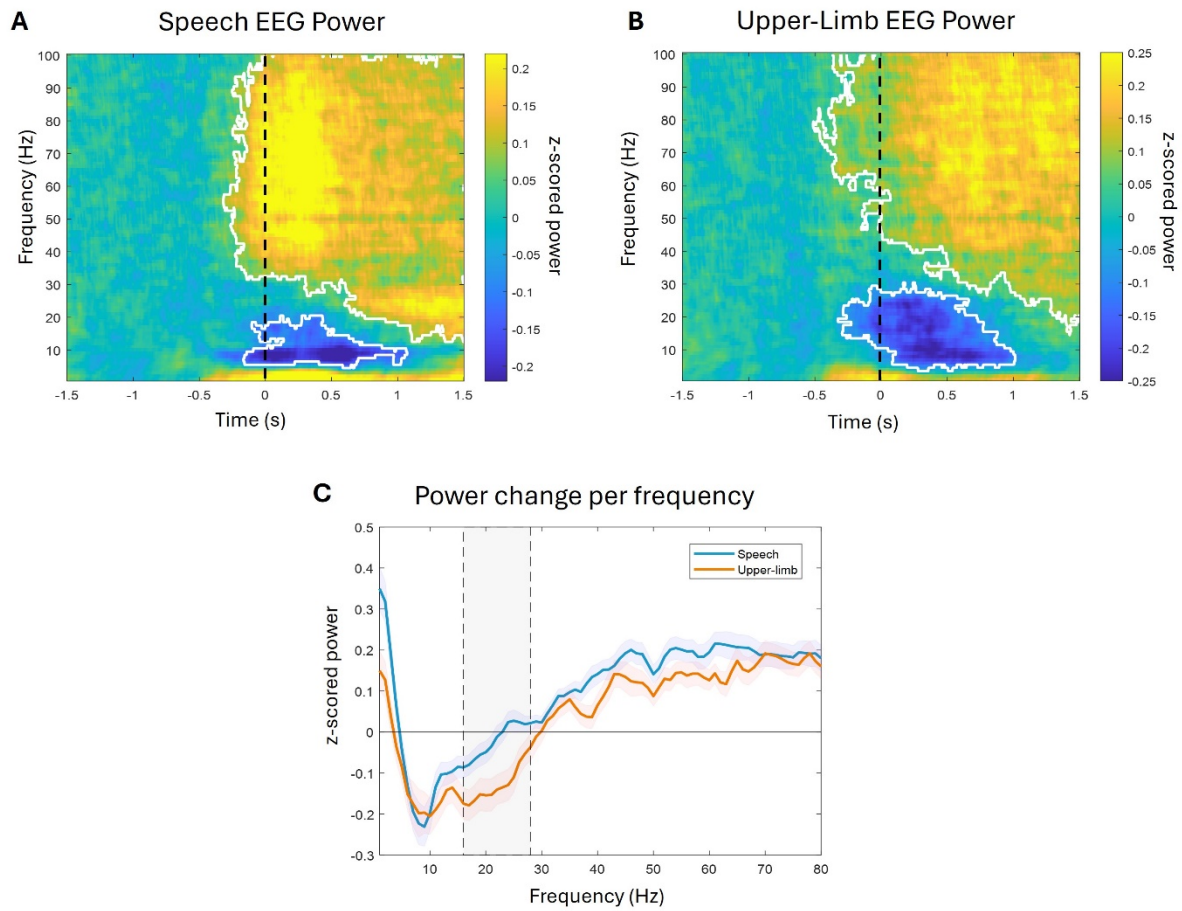
(B) Compression Ratios vs. Decoding Performance:

The relationship between model compression ratios and decoding performance (vs. the uncompressed 1D-CNN model). The student models trained with model distillation only suffer minor decreases in performance (up to $-1.3 \pm 4.0\%$ balanced accuracy). This demonstrates that model distillation techniques can preserve decoding accuracy while reducing the computational footprint of the models, making the neural decoding models better compatible with real-time environments with limited computational resources.

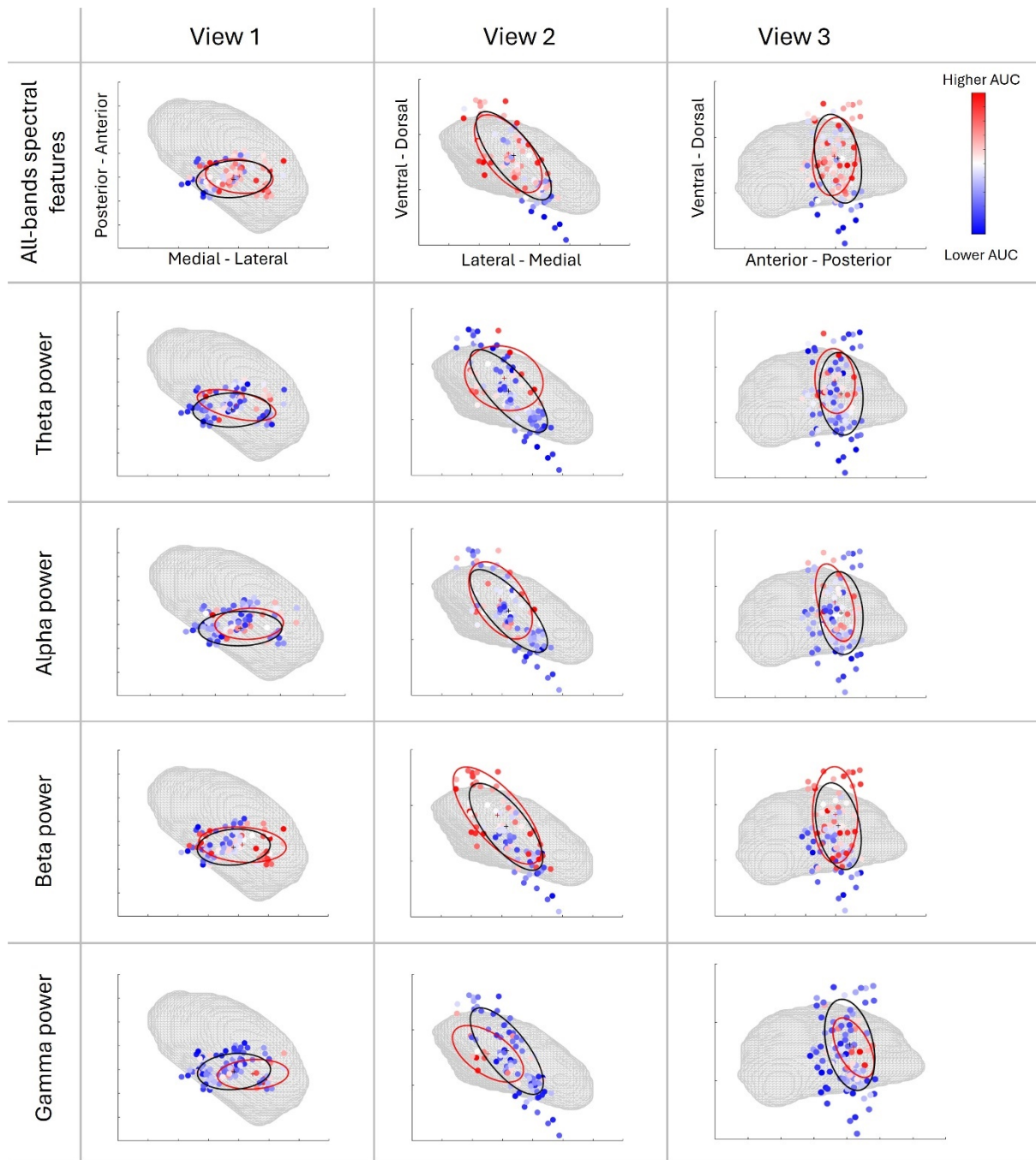


Supplementary Figure IV: Pairwise Plots of AUCs Values Yielded by Different Feature Sets. Feature sets are: alpha band (8 – 12), low beta band (13 – 20), high beta band (20 – 30), gamma band (gamma), high frequency oscillations (HFO), all of the aforementioned frequency bands simultaneously (“all_freqs”), and all frequency bands and

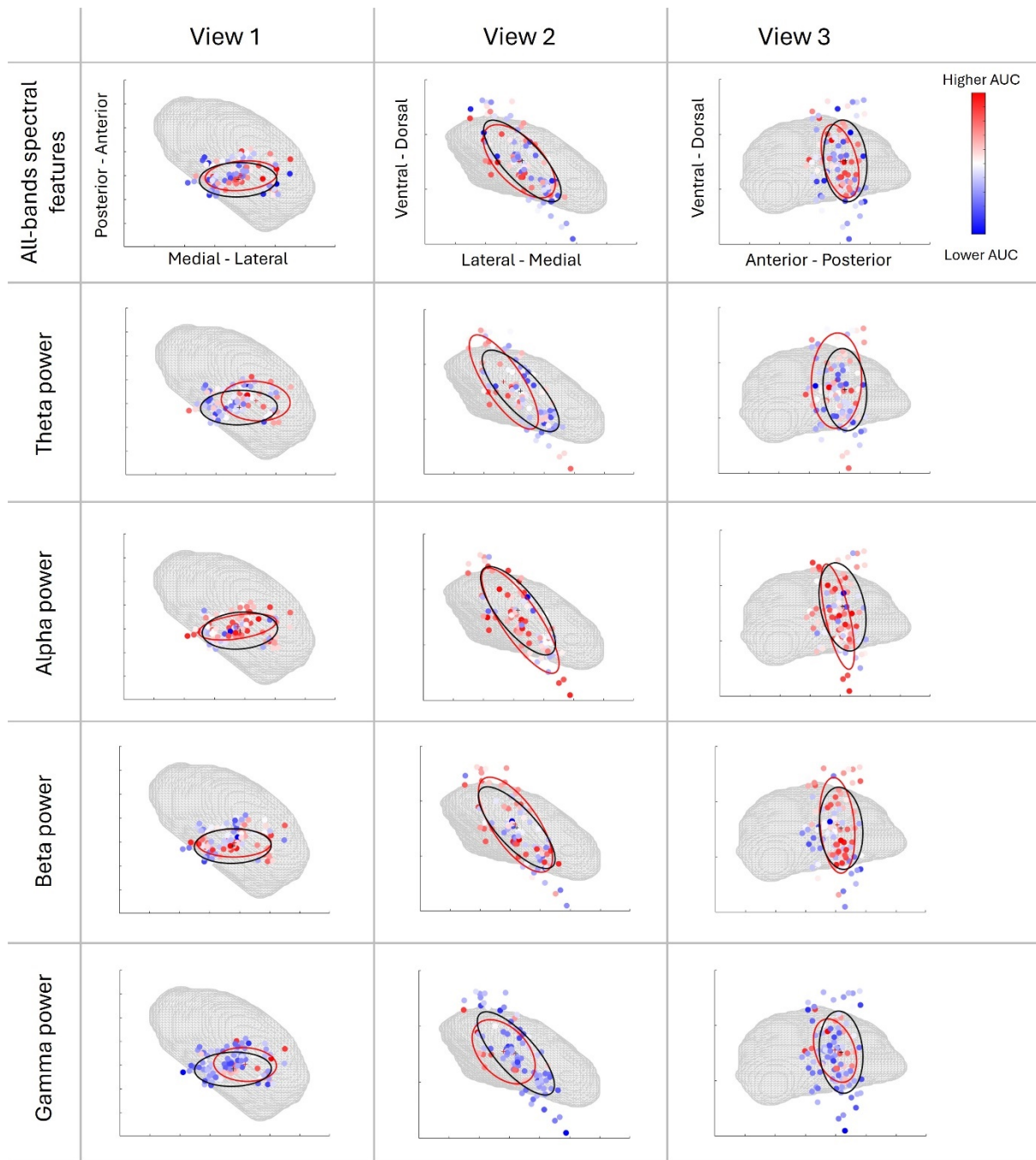
additional features (“all”, see Section 4.2.3 for further details on the feature extraction pipeline). The highest correlation can be observed between the “all” and “all_freqs” model.



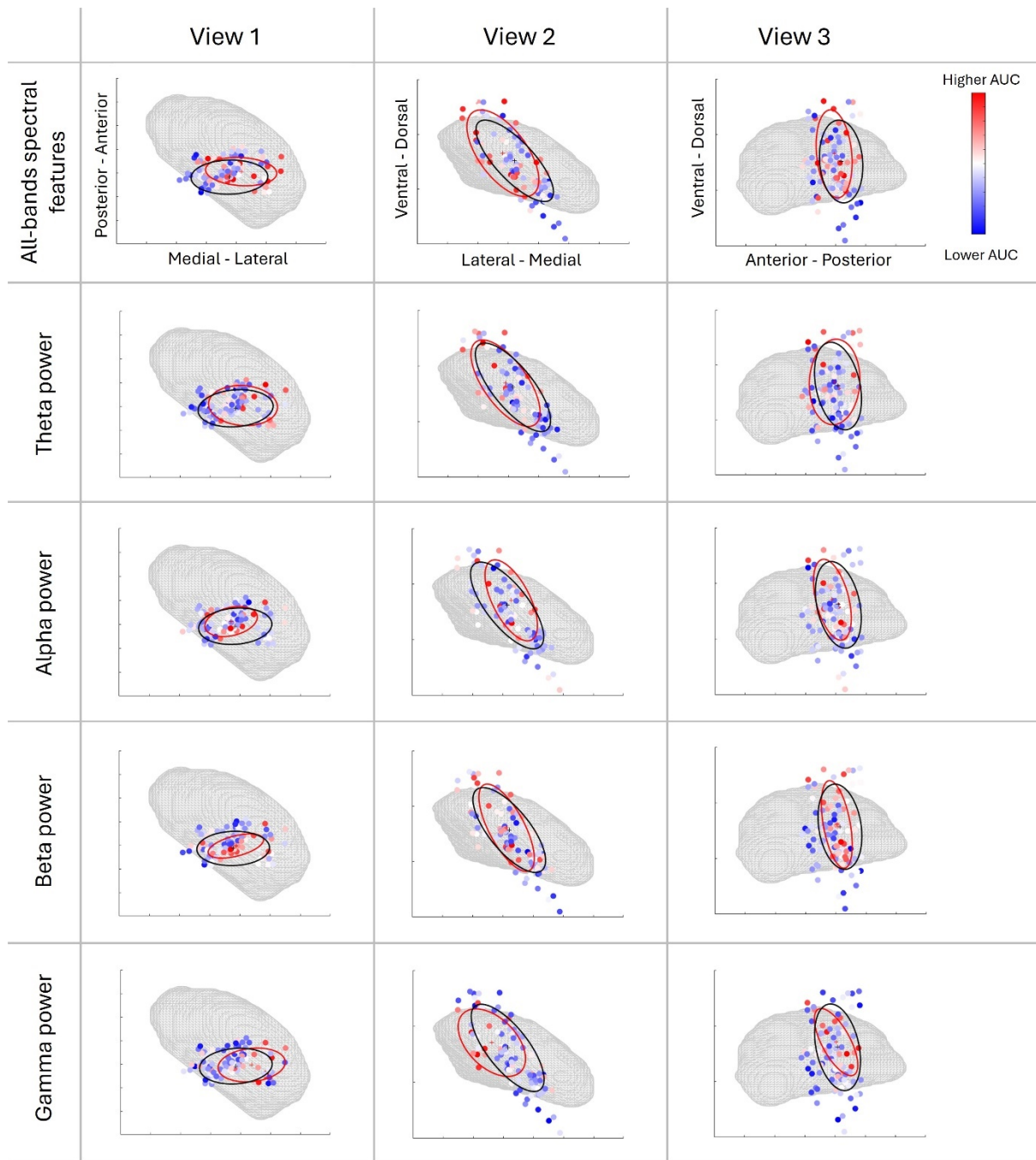
Supplementary Figure V: Movement-related Spectral and Functional Connectivity Modulation in the Motor Cortex. Grand average of z-scored spectral-power relative to baseline across all EEGs during (A) speech production ($n=34$ hemispheres) and (B) upper-limb movement ($n=32$ hemispheres). (C) Differences in movement-related averaged z-scored spectral changes relative to baseline per frequency bin between speech production task (blue) and upper-limb movement (orange) during 1s of maximum modulation. Clusters with significant power modulations (cluster-based permutation test, $p<0.05$) are marked with white contour. Gray bands marked the frequency range of significant differences in modulation between speech and upper-limb movement using cluster-based permutation analysis ($p<0.05$). Values in C are represented as mean $\pm$ SEM.



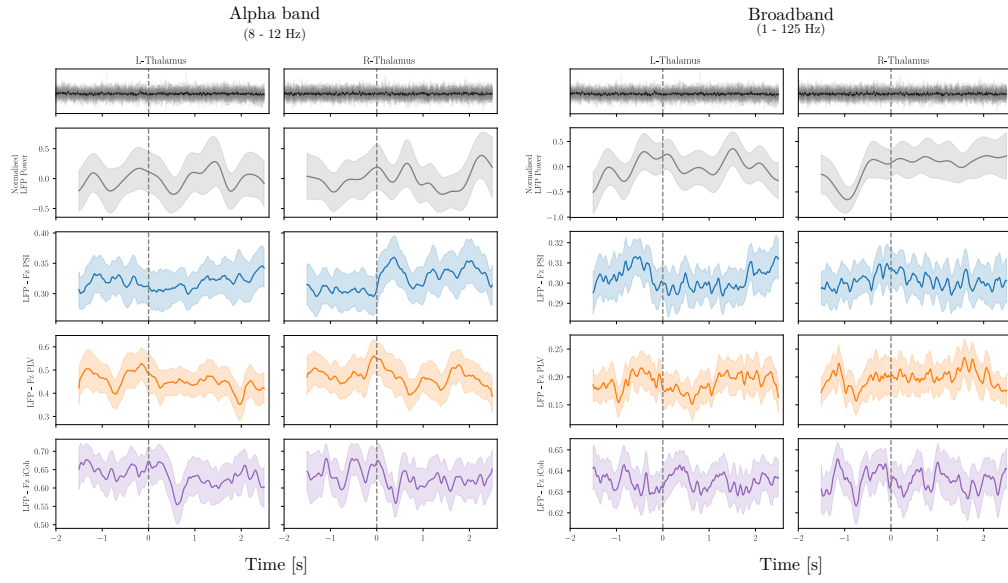
Supplementary Figure VI: Spatial distribution of AUC values for speech vs. rest using spectral feature in all or individual frequency bands. Clusters of bipolar contacts that showed highest AUC values are marked with red ellipse. Cluster of bipolar contacts that exhibited the highest beta power at rest is marked in black.



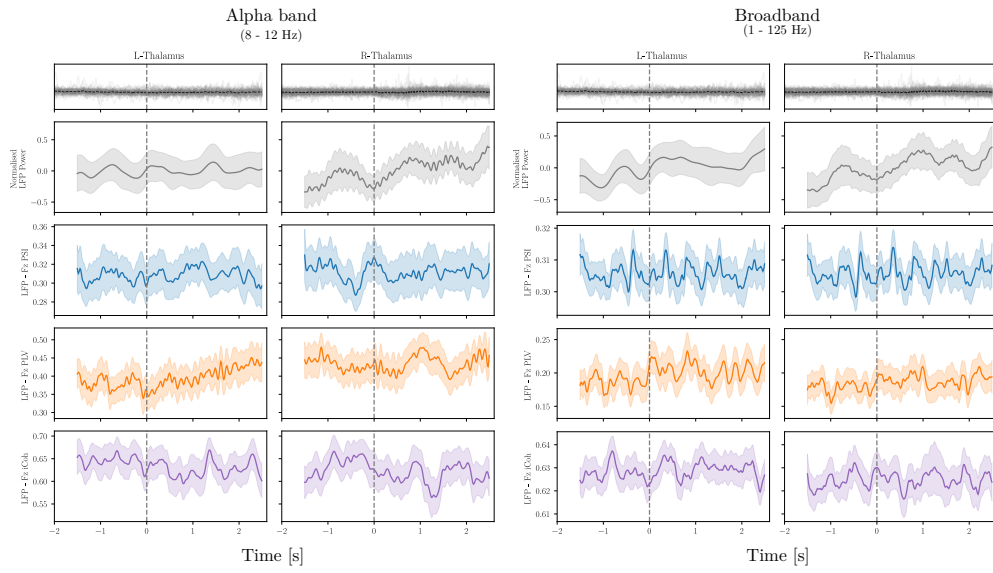
Supplementary Figure VII: Spatial distribution of AUC values for upper limb movement vs. rest using spectral feature in all or individual frequency bands. Clusters of bipolar contacts that showed highest AUC values are marked with red ellipse. Cluster of bipolar contacts that exhibited the highest beta power at rest is marked in black.



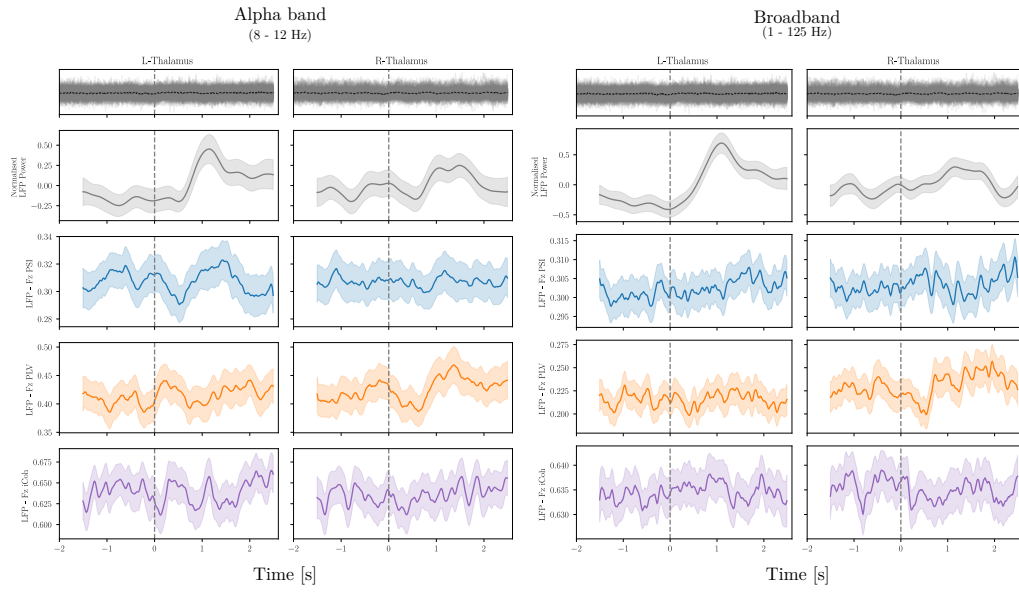
Supplementary Figure VIII: Spatial distribution of AUC values for speech vs. upper limb movement using spectral feature in all or individual frequency bands. Clusters of bipolar contacts that showed highest AUC values are marked with red ellipse. Cluster of bipolar contacts that exhibited the highest beta power at rest is marked in black.



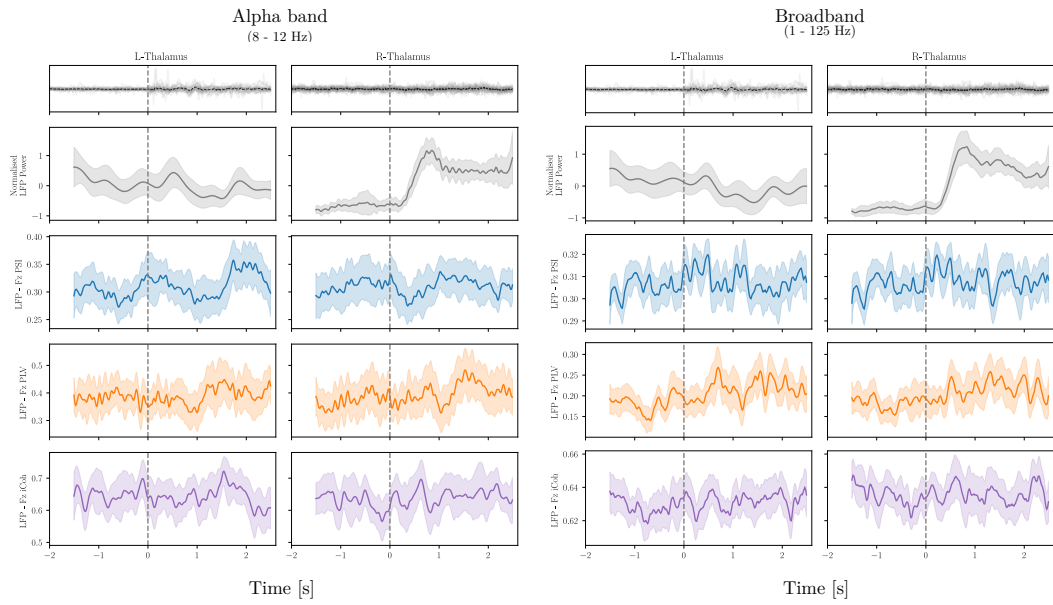
Supplementary Figure IX: Feature Progression (alpha band and broadband) for P1.



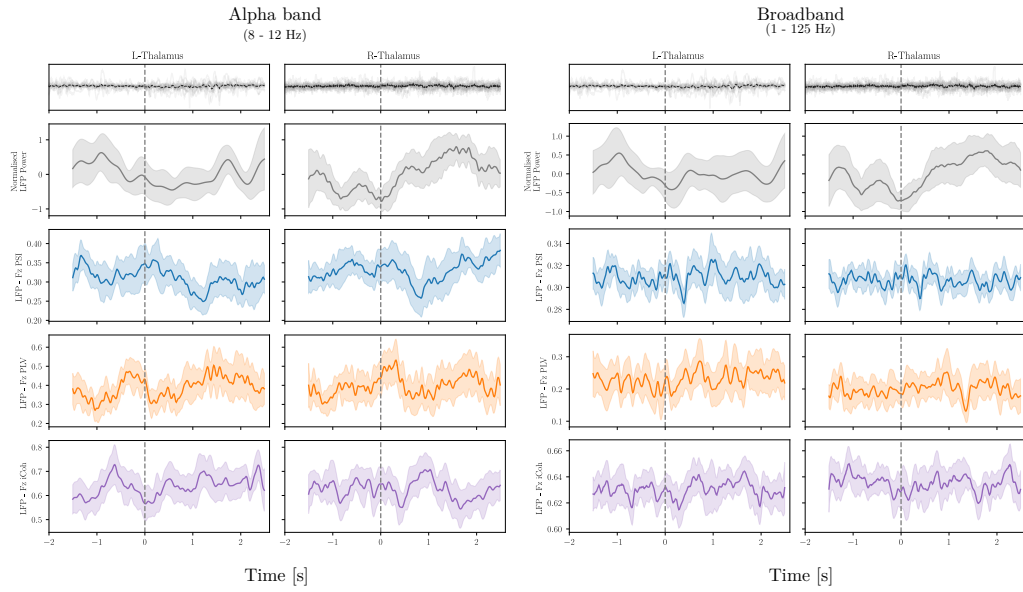
Supplementary Figure X: Feature Progression (alpha band and broadband) for P2.



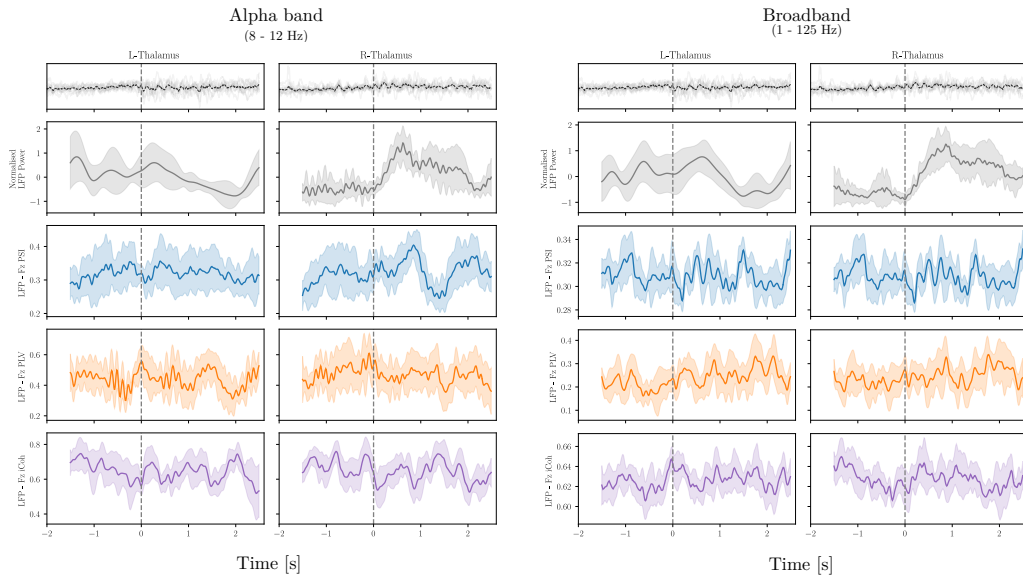
Supplementary Figure XI: Feature Progression (alpha band and broadband) for P3.



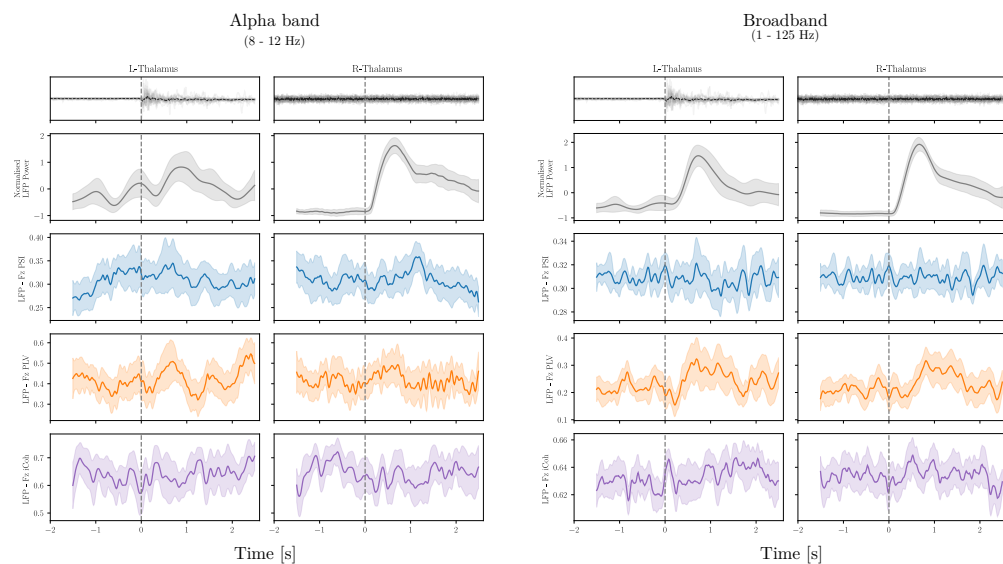
Supplementary Figure XII: Feature Progression (alpha band and broadband) for P4.



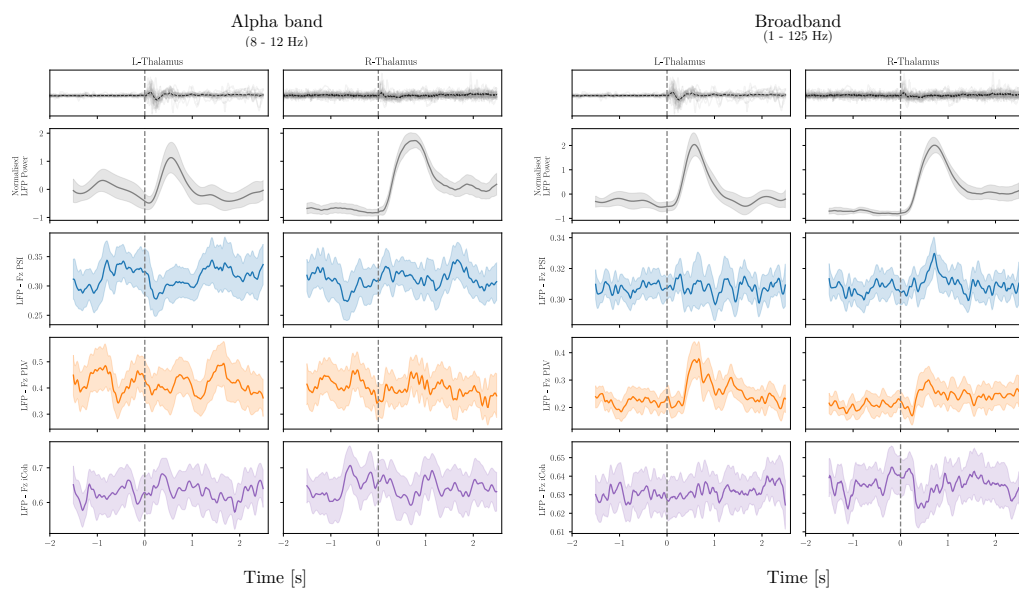
Supplementary Figure XIII: Feature Progression (alpha band and broadband) for P5.



Supplementary Figure XIV: Feature Progression (alpha band and broadband) for P6.



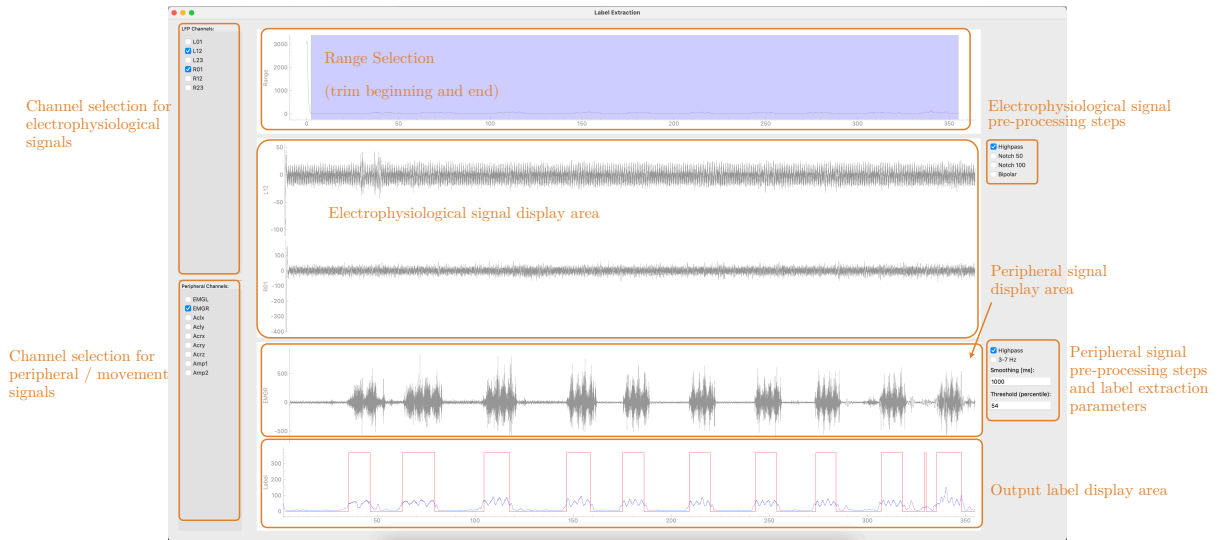
Supplementary Figure XV: Feature Progression (alpha band and broadband) for P7.



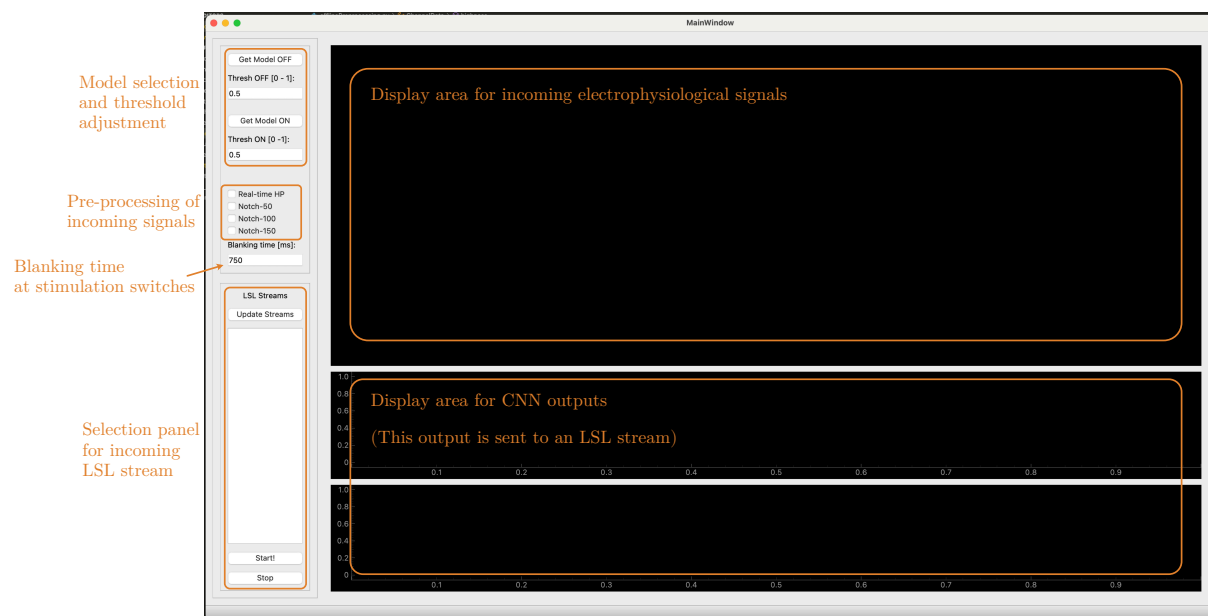
Supplementary Figure XVI: Feature Progression (alpha band and broadband) for P8.

	Delta/Theta		Alpha		Beta		Gamma		All bands	
	Connectivity	Connectivity, Power & Time-domain	Connectivity	Connectivity, Power & Time-domain	Connectivity	Connectivity, Power & Time-domain	Connectivity	Connectivity, Power & Time-domain	Connectivity	Connectivity, Power & Time-domain
P1	0.38 [0.39 - 0.59]	0.47 [0.37 - 0.60]	0.44 [0.36 - 0.61]	0.47 [0.37 - 0.65]	0.58 [0.34 - 0.61]	0.53 [0.34 - 0.63]	0.54 [0.37 - 0.61]	0.53 [0.37 - 0.63]	0.51 [0.36 - 0.62]	0.49 [0.38 - 0.62]
P2	0.49 [0.41 - 0.56]	0.50 [0.42 - 0.57]	0.45 [0.39 - 0.57]	0.47 [0.40 - 0.59]	0.51 [0.37 - 0.58]	0.51 [0.38 - 0.59]	0.53 [0.39 - 0.59]	0.52 [0.39 - 0.58]	0.44 [0.41 - 0.58]	0.49 [0.41 - 0.60]
P3	0.49 [0.43 - 0.55]	0.49 [0.44 - 0.54]	0.50 [0.45 - 0.55]	0.48 [0.43 - 0.56]	0.51 [0.44 - 0.55]	0.50 [0.43 - 0.57]	0.45 [0.43 - 0.56]	0.46 [0.44 - 0.56]	0.46 [0.44 - 0.55]	0.47 [0.44 - 0.55]
P4	0.47 [0.35 - 0.60]	0.43 [0.36 - 0.60]	0.48 [0.33 - 0.63]	0.39 [0.34 - 0.66]	0.39 [0.30 - 0.63]	0.36 [0.34 - 0.65]	0.63 [0.32 - 0.64]	0.51 [0.34 - 0.68]	0.52 [0.37 - 0.63]	0.44 [0.35 - 0.65]
P5	0.32 [0.34 - 0.64]	0.42 [0.35 - 0.63]	0.64 [0.29 - 0.65]	0.60 [0.30 - 0.68]	0.58 [0.30 - 0.65]	0.56 [0.29 - 0.68]	0.64 [0.30 - 0.69]	0.52 [0.32 - 0.71]	0.60 [0.31 - 0.66]	0.49 [0.31 - 0.69]
P6	0.49 [0.23 - 0.66]	0.44 [0.29 - 0.70]	0.53 [0.26 - 0.69]	0.38 [0.24 - 0.77]	0.38 [0.25 - 0.68]	0.41 [0.27 - 0.78]	0.71 [0.29 - 0.73]	0.55 [0.25 - 0.77]	0.45 [0.28 - 0.72]	0.50 [0.26 - 0.73]
P7	0.57 [0.34 - 0.62]	0.57 [0.32 - 0.64]	0.57 [0.34 - 0.63]	0.59 [0.32 - 0.67]	0.65 [0.31 - 0.66]	0.59 [0.33 - 0.68]	0.55 [0.33 - 0.67]	0.60 [0.30 - 0.69]	0.56 [0.32 - 0.64]	0.46 [0.32 - 0.68]
P8	0.45 [0.37 - 0.60]	0.47 [0.37 - 0.60]	0.47 [0.33 - 0.62]	0.47 [0.35 - 0.64]	0.47 [0.30 - 0.62]	0.49 [0.35 - 0.66]	0.43 [0.34 - 0.64]	0.47 [0.35 - 0.64]	0.48 [0.36 - 0.63]	0.42 [0.36 - 0.65]
Cross-patient Mean	0.46 [0.36 - 0.60]	0.47 [0.36 - 0.61]	0.51 [0.34 - 0.62]	0.48 [0.34 - 0.65]	0.51 [0.33 - 0.62]	0.49 [0.34 - 0.66]	0.58 [0.35 - 0.64]	0.53 [0.35 - 0.66]	0.51 [0.36 - 0.63]	0.48 [0.35 - 0.65]

Supplementary Table I: Feature Set and Frequency Band Sweep for the Cross-Validated Detection of Pre-Tic Periods. This table presents each patient's area under the ROC curve (AUC) for different feature sets (e.g., connectivity-only vs. combined features) and frequency bands (delta/theta, alpha, beta, gamma, and all bands combined). Values are shown in the format AUC [5th, 95th], where the bracketed numbers represent the 5th and 95th percentiles of the permutation-based null distribution. None of the observed AUC scores exceed the 95th percentile threshold, indicating that no feature set or frequency band provides above-chance classification of pre-tic intervals in this cross-validated analysis.



Supplementary Figure XVII: Custom Signal Labelling Software (written in Python). Allows to select and pre-process electrophysiological signals, and to extract a label from peripheral signals. The program incorporates functionality to send these data to a server and trigger the training of a ML system.



Supplementary Figure XVIII: Software Program for real-time CNN-driven aDBS (written in Python). Includes functionality to buffer signals from an incoming LSL stream, pre-process them, and process them through a neural network model in real time. The output can be used to guide a stimulator.