



UNIVERSITAT DE
BARCELONA

What's in a grasp? Describing motion-based brain states using reservoir computing

Master's Thesis

Master in Fundamental Principals of Data Science

Faculty de Mathematics and Informatics
University de Barcelona

2024-2025

Author: Luca Eric Di Croce

Supervisor: Dr. Ignasi Cos Aguilera

Barcelona, June 2025

“The human brain, then, is the most complicated organization of matter that we know.”

— Isaac Asimov

Contents

Abstract	4
Acknowledgements	5
1 Introduction	6
1.1 Problem Statement	6
1.2 Motivation	7
2 Objectives	8
3 Background	9
3.1 Neural Basis of Movement	9
3.1.1 Involved Brain Regions	9
3.1.2 Relevant Frequency Bands	9
3.2 Experimental Setup and Data Acquisition	10
3.2.1 Experiment Description	10
3.2.2 Local Field Potentials (LFPs)	10
3.2.3 Data Description	10
3.3 Reservoir Computing	11
3.3.1 Theoretical Background	11
3.3.2 Reservoir Dynamics	12
4 Methodology	14
4.1 Data Preprocessing	14
4.1.1 Bad Channel Removal	14
4.1.2 Powerline Noise Suppression	15
4.1.3 Outlier Smoothing	16
4.1.4 Frequency Band Partitioning	17
4.2 Data Processing	18
4.2.1 Spectrogram	18
4.2.2 Principal Component Analysis (PCA)	19
4.3 Modeling Approaches	20
4.3.1 Simple ESN	20
4.3.2 Region-Based ESN	20

4.3.3	Bandwidth-Based ESN	21
4.3.4	Region-Bandwidth Combined ESN	22
4.4	Model Parameters, Training and Evaluation	23
4.5	Analyzing Movement Dynamics	23
4.5.1	Principal Component Analysis	24
4.5.2	Classwise Reservoir Contributions	24
4.5.3	Timepoint Dynamics per Class	24
5	Results	25
5.1	Modeling Results	25
5.1.1	Simple ESN	25
5.1.2	Region-Based ESN	26
5.1.3	Bandwidth-Based ESN	27
5.1.4	Region-Bandwidth Combined ESN	28
5.2	Analyzing Movement Dynamics	29
5.2.1	Principal Component Analysis	29
5.2.2	Classwise Reservoir Contributions	29
5.2.3	Time Dynamics per Class	32
5.2.4	Distance Analysis On Full Reservoir Dynamics	33
5.3	Understanding Movement Dynamics	34
5.4	Comparison With Joshua Tapper’s Project	35
5.4.1	Accuracy	36
5.4.2	Manifold Expressiveness	36
5.4.3	Handedness Preference	36
5.4.4	Frequency Bands and Regions	36
6	Discussion	38
7	Conclusion	40
8	Limitations	42
9	Future Work	43
A	Code and GitHub Link	44
B	Author Contributions and Rights	45
C	Usage of LLMs throughout the project	46
	Bibliography	47

Abstract

Luca Eric DI CROCE

What's in a grasp? Describing motion-based brain states using reservoir computing

Keywords: Echo State Network, Neuroscience, Neural Signal Classification, Interpretability

This study explores the use of Echo State Networks (ESNs) to decode motion-related brain activity from Local Field Potential (LFP) time series recordings collected during a five-phase grasping task performed by a non-human primate. Each phase was executed with one of both hands, resulting in a 10-class classification problem (five trial phases times two hands).

We implemented a range of ESN architectures, from simple models to more structured variants incorporating frequency bands and brain regions. Leveraging the ESN's role as a dynamic feature generator, which transforms temporal patterns into a high-dimensional expressive state space, we achieved a strong performance: our simplest model reached 84% accuracy, while the most structured variant achieved up to 89%, surpassing most literature benchmarks for LFP-based motor classification.

In addition to accuracy, ESNs provided interpretable internal representations. Low-dimensional projections of the output weights revealed distinct, phase-specific neural signatures, and highlighted the contribution of both spectral, temporal and spatial features to motor decoding.

Acknowledgements

First and foremost, I would like to sincerely thank Dr. Ignasi Cos Aguilera for his invaluable guidance, encouragement and the inspiring ideas that significantly shaped this work. His expertise and insightful feedback have been once again instrumental in the completion of this project.

I am also sincerely grateful to the Universitat de Barcelona for offering an enriching academic environment and invaluable resources throughout both my Master's and Bachelor's studies. The commitment and expertise of the professors went far beyond what I had anticipated, providing exceptional guidance and support that greatly enhanced my learning experience.

I would like to acknowledge Joshua Tapper as a supportive collaborator throughout this process. Working on similar projects allowed us to compare results and benchmark our progress effectively, as well as gave us someone to call and share opinions.

I am also especially grateful to Michael DePass for providing the data and, more importantly, for his generous willingness to go above and beyond in offering extensive feedback and support throughout our discussions, which greatly contributed to enhancing this work.

To my friends, thank you for the moments of levity and motivation that kept me going during challenging times. Knowing I can always rely on you brings me an incredible sense of comfort.

I am deeply grateful to my family for their unwavering love, patience, and encouragement. Their constant support has been a pillar of strength throughout this journey.

Finally, I wish to thank Carlota for her unending support, companionship, encouragement and love, which have made this experience truly meaningful. I wouldn't have been able to achieve half of the things I have without her.

To all of you, thank you. This achievement is as much yours as it is mine.

Chapter 1

Introduction

The field of cognitive neuroscience is vast and complex, but one with an accelerating development and increasingly diverse applications. This in no small part is due to the more integrated usage of machine learning (ML) and AI, which has made it possible to extract structured meaning from complex and often noisy neural signals.

Among the growing array of ML techniques used, Echo State Networks (ESNs) have began emerging as particularly versatile and well-suited models for working with time series data. Based on reservoir computing, ESNs offer a compelling balance between complex and interpretable temporal modeling. In other words, their ability to process temporal information with low computational overhead has made them ideal candidates for decoding time-dependent signals from neural recordings.

In this study, we apply a range of ESN-based models to LFP recordings collected during a structured grasping task. The goal is to classify and understand the neural signatures that correspond to different movement phases, and to examine the impact of the signals from across brain regions and frequency bands. By doing so, we aim to evaluate the capacity of ESNs not only to perform classification tasks, but also to yield insights into the structure and dynamics of movement-related brain activity.

Beyond accuracy alone, we emphasize model interpretability and the insight our model gains from the data. Our long-term motivation is to contribute toward the creation of a formalized and interpretable vocabulary for LFP and EEG-based behavioral pattern recognition, as such a framework could streamline future experimentation, enable more systematic cross-subject and cross-task comparisons, and ultimately provide a more robust foundation for the study of motor cognition.

1.1 Problem Statement

In this study, we analyze Local Field Potential (LFP) recordings obtained during a five-phase grasping task performed by a trained monkey. The aim is to determine whether the neural activity patterns in each phase can be effectively distinguished and understood using ML models.

To do so, we have decided to use reasonably interpretable models that are still able to predict to a certain threshold of accuracy. At the end, we settled with Reservoir models, as they are not only very proficient in time series analysis, but are also very visually interpretable, due to their relatively straightforward and easy structure when compared to other deep learning algorithms.

We believe that the potential minor accuracy drop will be a worthy trade-off, as it will help us focus the problem from a different lens, and give us valuable insight we might have otherwise missed, as well as expand on the existing literature of LFP applications and utilities, and applied reservoir dynamics and Echo State Networks (ESNs).

1.2 Motivation

The execution of goal-directed movement, such as a grasp, engages complex patterns of neural activity across multiple brain regions and brainwave frequency bands. Decoding these patterns in an interpretable and potentially generalizable way remains a central challenge in neuroscience, particularly when attempting to link transient neuro-electrical signals with discrete behavioral states, and establish a methodology for behavioral pattern recording and analysis.

If these behavioral patterns could be systematically categorized and classified, it would not only enable the comparison of distinct brain states and the identification of their specific neural correlates, but also facilitate the development of a structured encoding vocabulary for local field potentials (LFPs), that, leveraging natural language processing techniques, such a vocabulary could support the real-time interpretation and visualization of LFP time-series recordings.

This capability would enhance experimental and clinical methodologies by providing immediate feedback during testing, allowing for quicker and more precise interventions in medical treatments or acute conditions, such as seizure detection and enabling personalized and adaptive neurological treatment strategies. Furthermore, the ability to decode and translate neural activity into interpretable outputs could help form the basis for communication with non-responsive or comatose patients, refinement of brain-computer interfaces (BCIs) through improved decoding of cognitive and motor intentions, the identification of deviations in neural signatures which could be used as biomarkers for disorders such as depression or ADHD, and even real-time cognitive state monitoring in high-stakes environments, such as aviation or surgery, where mental fatigue poses significant risks.

Chapter 2

Objectives

In this study, we aim at assessing whether reservoir models can reliably capture the specific neural code extracted from the brain via LFP time series for each of the motor brain states considered. Beyond classification accuracy, we intend to analyze the internal dynamics of the reservoir to characterize neural signatures or patterns associated with each phase, as well as the frequency bands and cortical regions that most contribute to the decoding process.

Through this dual goal approach, balancing performance and interpretability, we seek to advance both technical and theoretical understanding of motion-based brain states by theoretical means.

To this end, we have distributed the study around three distinct objectives:

- First, we aim to classify movement phases using LFP recordings with reservoir models, quantifying the specificity of the representation associated with each brain state.
- Once the model is trained, we propose to identify the contribution of each brain region and frequency band to each state, to understand the role of each bit throughout each phase, and to be able to compare our results with previous findings.
- Finally, we analyze dynamic changes in the reservoir structure across brain states, as characterized by the ESN. This will determine how specific are the neural signature patterns of each phase/motor state, as well as contributing to the understanding of the neural code for each phase.

Chapter 3

Background

3.1 Neural Basis of Movement

Understanding the neural basis of movement involves identifying the cortical regions responsible for motor control and how their coordinated activity enables complex, goal-oriented actions. Grasping, in particular, is a multi-phase task that involves both planning and execution components, each engaging different neurophysiological dynamics.

3.1.1 Involved Brain Regions

Grasping relies heavily on sensorimotor integration, which is coordinated across multiple regions including the primary motor cortex (M1), ventral and dorsal premotor cortices (PMv and PMd), and their interhemispheric counterparts. M1 is chiefly responsible for voluntary movement execution (Kakei, Hoffman, and Strick (1999)) while PMv and PMd are involved in planning, coordination, and adjustment of motor strategies based on sensory input (Davare et al. (2009)). By isolating channels from these regions, we aim to evaluate their individual and combined contributions to encoding grasp-related phases and to try to differentiate the hand used.

3.1.2 Relevant Frequency Bands

Neural oscillations across different frequency bands are known to reflect distinct cognitive and motor states. Low-frequency bands such as theta and delta are typically associated with high-level planning and error monitoring, while higher-frequency bands such as beta and gamma have been linked to sensorimotor coordination and muscular control.

In this study, due to the higher sensitivity of the LFP recoding, we are able to further divide the gamma range into standard gamma and high gamma, with ripple and fast ripple and multiunit capturing even more localized, high-speed dynamics. Evaluating each band's relative impact allows us to determine which frequencies carry task-relevant information during each movement phase.

3.2 Experimental Setup and Data Acquisition

3.2.1 Experiment Description

The data used was collected from a trained non-human primate performing a five-phase grasping task (DePass et al. (2022)). Each phase corresponds to a distinct component of motor behavior, ranging from initiation and reach to grip and hold. Recordings were made intracranially multiple 96-channel Utah arrays, enabling dense spatial sampling across multiple cortical regions.

This was one of the few experiments designed to capture and classify such a large amount of states, as the goal was not only to classify 5 states, but also to differentiate based on what hand was used and the rotation of said hand, with discrete values between 0° and 135°). While in our study we do not consider the angle, we do use the 5 phases as our classes for the 5-class approach, and add the hand dimensionality (right or left) for our 10-class task.

Please note that both the data collection and the experiments were not performed by us, but by a research group at the Université de Montréal (DePass et al. (2022)).

3.2.2 Local Field Potentials (LFPs)

Local Field Potentials are extracellular voltage recordings obtained through electrodes implanted in cortical tissue by specialized operations. They reflect the aggregate input (rather than output) of neuronal populations in the vicinity of the electrode, making them particularly sensitive to subthreshold events and synaptic potentials.

As such, they end up providing both superior spatial and frequency resolution when compared to EEGs, and are consequently increasingly favored in animal models for studying fine-grained neural dynamics, with roughly 30% of all published articles in google scholar about LFP in neuroscience having been published in the last 3 years.

3.2.3 Data Description

The dataset consisted of:

- Sampling rate: 2000 Hz
- Temporal duration: Each recording of each phase was of a quarter of a second, translating to 508 timepoints per sample
- Number of channels: 256 different channels
- Preprocessing: Segmentation by phase (epoching), with a sample per each phase but with no initial filtering or artifact rejection

- Brain region - channel equivalence table: The channels related to the 5 distinct cortical areas observed

This allowed us to conduct both full-signal and region/band-specific analyses, preserving as much variance as possible for initial model evaluations, as well as apply our own data preprocessing methodologies.

3.3 Reservoir Computing

Reservoir computing (RC) is a framework designed for processing temporal data using a high-dimensional dynamical system called a "reservoir," which projects input signals into a rich feature space. A simple, often linear readout layer is then trained to interpret the reservoir states.

This combination of high accuracy and low computational complexity has led to their increasing adoption across a wide range of fields, including biomedical signal analysis, machinery fault detection, data science, security systems, communication networks, chemistry modeling, environmental monitoring, and audio and speech processing (Zhang and Vargas (2023)).

In recent years, they have gained significant traction within the field of cognitive sciences as well, where they have been applied with great success to a variety of tasks, including the detection of attentional states (Jeong and Jeong (2020)), classification of emotional responses from EEG data (Tang et al. (2023)), and the diagnosis of neurological conditions such as epilepsy (Fourati et al. (2017)).

3.3.1 Theoretical Background

At its core, reservoir computing relies on the concept of a fixed, randomly connected recurrent network that transforms low-dimensional input signals into high-dimensional dynamic trajectories. These trajectories encode the temporal structure of the input in a nonlinear fashion, enabling effective classification or regression through a simple linear regression.

The reservoir per se consists of a collection of interconnected computational nodes, each maintaining an internal state that evolves over time. At each time step t , the state vector $\mathbf{x}(t)$ is updated as follows:

$$\mathbf{x}(t) = \tanh(\mathbf{W}_{in}\mathbf{u}(t) + \mathbf{W}_{res}\mathbf{x}(t-1))$$

where $\mathbf{u}(t)$ is the input vector, \mathbf{W}_{in} is the input-to-reservoir weight matrix, and \mathbf{W}_{res} is the recurrent reservoir weight matrix. The nonlinearity (in our case we used tanh, as it seems to be the most accepted choice) ensures the transformation of the input into a dynamic, high-dimensional space.

Notably, during training, the reservoir weights \mathbf{W}_{res} and \mathbf{W}_{in} remain fixed. The only component that is updated is the output layer, governed by the output weight matrix \mathbf{W}_{out} . This matrix is computed by collecting the reservoir states across time and applying ridge regression:

$$\mathbf{W}_{out} = \mathbf{Y}\mathbf{X}^\top (\mathbf{X}\mathbf{X}^\top + \lambda\mathbf{I})^{-1}$$

where \mathbf{X} is the matrix of collected reservoir states, \mathbf{Y} is the matrix of target outputs, λ is a regularization factor, and \mathbf{I} is the identity matrix.

This two phase training process, where we first defining the reservoir with input to collect internal states, and then solving a closed-form regression, offers both computational efficiency and analytical clarity, which is very advantageous in domains like neural decoding, as transparency is as valuable as performance.

3.3.2 Reservoir Dynamics

The most important feature of reservoir computing is the dynamics of the reservoir per se. These dynamics are governed by the randomly initialized weights \mathbf{W}_{res} , which define the recurrent connectivity between nodes. Through these connections, the reservoir is able to operate with short-term memory, retaining a decaying trace of historical inputs, which is a property crucial for modeling time-dependent phenomena.

To train the model, each reservoir node receives a weighted combination of the external input $\mathbf{u}(t)$, the previous reservoir state $\mathbf{x}(t - 1)$, and combines them using the non-linear response function. This results in the integration of both past and present information, giving rise to rich and informative internal dynamics. Three hyperparameters play a key role in shaping this behavior: the *spectral radius*, which controls the stability and temporal memory of the reservoir; the *input scaling factor* (α), which adjusts the sensitivity to new inputs; and the *sparsity*, which determines the proportion of active connections within the reservoir and influences its complexity and computational efficiency.

Once trained, the output of the reservoir is usually computed by linearly combining the current state of all nodes:

$$\mathbf{y}(t) = \mathbf{W}_{out}\mathbf{x}(t)$$

which implies that every node contributes to the final output based on its current activation.

In practical terms, this means that the reservoir acts as a dynamic feature generator, where the input sequence is fed into the network, and from these the internal states trace a trajectory in high-dimensional space. These trajectories are then interpreted by the output layer, which effectively separates them into relevant categories—in our

case, the distinct phases of a motor task. Because the reservoir is fixed, the overall computational cost is low, and the resulting system is highly suitable for real-time or embedded neural decoding tasks.

Chapter 4

Methodology

4.1 Data Preprocessing

During EEG or LFP recordings, due to the sensitive nature of the recording equipment and the variability of the recording source, one can always expect noise and unwanted interference to appear.

While it is not always possible to clean all channels, there are some specific instances of noise that are relatively straightforwards to clean, as we either know their distribution pattern (e.g powerline removal) or can easily identify their abnormal behavior.

To both remove the unwanted channels and clean the rest, the LFP recordings underwent a series of preprocessing steps, which lead to improved data quality and ensured the signals were suitable for time-series classification.

4.1.1 Bad Channel Removal

As explained previously, channels exhibiting abnormally high variance (exceeding two standard deviations above the mean) were identified as potentially faulty or noisy. These were deemed to be too irregular and not worth maintaining, so they were removed from all subsequent analysis in order to reduce signal artifacts and improve the reliability of the resulting models.

To do so, we applied the premade function (courtesy of Manuel Hernandez) called `automatic_bad_channel_detection`, modified from the Komosar, Fiedler, and Haueisen (2022) paper. This algorithm determines the median of the standard deviation throughout all the channels of one sample, and if it is determined to be above or below the acceptable threshold, the entire sample is dropped.

As each of our samples in raw data is divided into 5 samples of epoched data, to avoid class imbalance and present a more uniform measurement, we applied the bad channeled detection algorithm on the imported raw data, and dropped their equivalent epoched samples.

When we applied the function `automatic_bad_channel_detection` to our raw data, we found that a surprising large amount of samples were flagged as abnormal, with the resulting loss of 70 of our 315 raw samples, and their equivalent 350 out of 1575 epoched samples. Even so, removing these channels not only improved signal stability but also led to measurable improvements in classification accuracy across all model variants, with a noticed increase of 4 to 5 %.

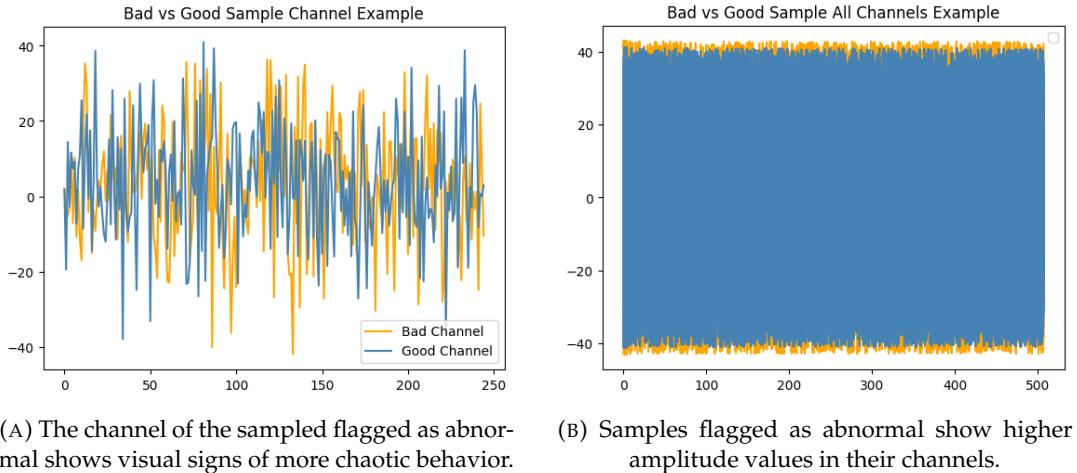


FIGURE 4.1: Comparison of bad channel detection using single- and multi-channel views.

4.1.2 Powerline Noise Suppression

Several methods were tested for removing powerline interference, which typically appears as narrow-band noise centered around 60 Hz (or 50 Hz depending on region) and its harmonics. Each considered method though offered a trade-off between effectiveness and preservation of surrounding signal content, so no ideal solution was established.

The first method tested early on with FFT-based subtraction, where the signal is transformed into the frequency domain using a Fast Fourier Transform (FFT). Harmonic peaks corresponding to powerline noise are then easily identified and suppressed (by either setting to 0 or some average value). While this method offers high precision and made sense when we were using FFT spectrograms as input for our model, it was eventually discarded as it could lead to spectral artifacts and unintentional attenuation of broader frequency regions due to windowing effects or imperfect peak isolation, as well as requiring unnecessarily expensive computation if not using the resulting spectrograms as input, which forced us to apply a inverse FFT.

We also tried implementing ZapLine filtering, a specialized algorithm designed to suppress periodic line noise using an adaptive filtering framework. It does so by

modeling line noise as a combination of sinusoids and removes it while maintaining the integrity of neighboring frequency components. Among the tested methods, ZapLine promised the best balance between interference removal and signal preservation, but we were sadly unable to use it due to compatibility errors.

Finally, we ended up using notch filtering, which applies a narrow band-stop filter centered on the interference frequency (e.g., 60 Hz) to eliminate the noise component. Although the application of notch filters tend to attenuate adjacent frequencies, which potentially distorted relevant signals that lie near the suppressed band, it was both computationally efficient and easy to implement. As such, using the right quality factor to minimize the unintended adjacent attenuation, we were able to effectively removed noise while preserving most of the adjacent spectral content, as seen in Fig. 4.2.

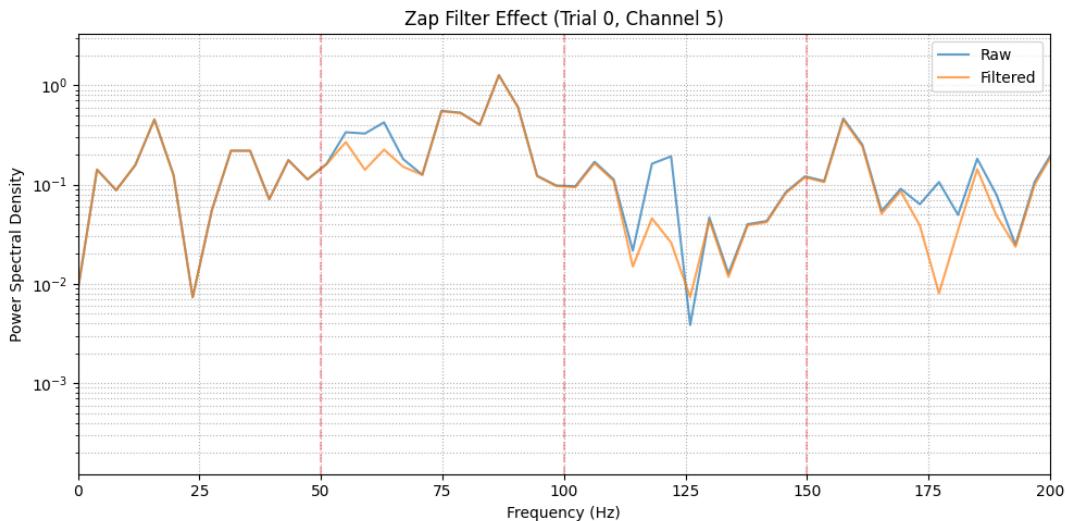


FIGURE 4.2: Visualization on the spectral dimension of the impact of notch filtering (Q 35).

As for the quality value of the notch filter, we empirically tested different possibilities, as seen in table 4.1, but ended up deciding on 35, as we found that it provided an optimal balance between attenuating the target frequencies while preserving the integrity of surrounding spectral components.

4.1.3 Outlier Smoothing

As mentioned at the beginning of this section, sometimes artifacts or abnormal behavior is introduced into our reading, be it due to equipment malfunction or neural irrelevant transient signals (e.g. eye blink artifacts).

While traditionally these could be eliminated by using a methodology such as ICA or some fine-tuned PCA, we decided to the much quicker and less data transformative outlier smoothing.

Quality Factor (Q)	Accuracy (%)
25	85.71
30	86.67
35	87.49
40	85.08
45	81.90
50	85.40
No Filter (Baseline)	83.49

TABLE 4.1: Classification accuracy at different notch filter quality factors (Q), compared to the unfiltered baseline, as tested on the simple ESN.

To do so, we applied a backward-looking rolling mean filter, considering per default the previous 5 values when possible or, if non are available (first 5 datapoints), setting the value to 0, for all datapoints with an amplitude surpassing two times the average standard deviations of all channels throughout all samples, which resulted in much more regular looking data, as seen in Fig. 4.3, as well as a slight increase in accuracy. Furthermore, this also allows us to be more confident in our generalization capabilities, as we ensure that our algorithms do not learn any non-realistic or dataset specific correlation.

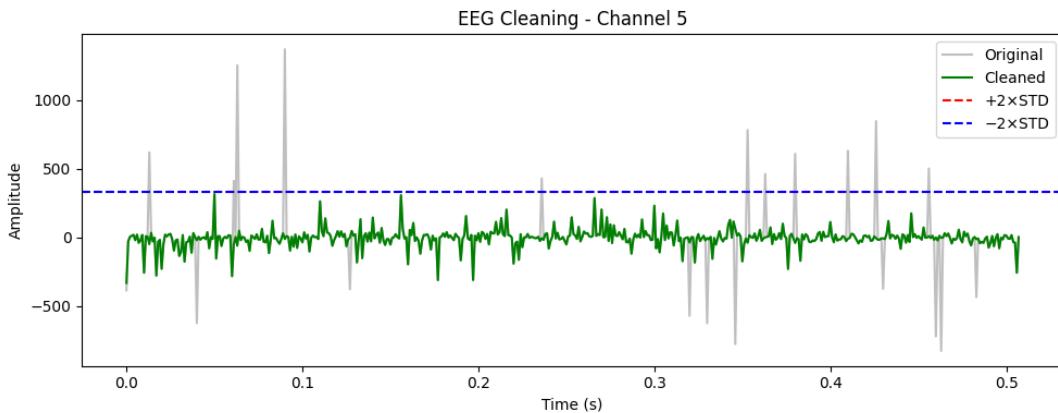


FIGURE 4.3: Visualization of original and cleaned timeseries, and threshold line. We can see that the threshold line is above the expected value, but it was considered beneficial to leave some leeway.

This process was done last as, being such a delicate operation, we wanted to make sure the average standard deviation captures was as close to the real one, which bad channels or powerline noise would heavily affect.

4.1.4 Frequency Band Partitioning

For some specific uses, we were interested in more than the full frequency range, and wanted to study how each of the canonical frequency bands behaved, considering as relevant canonical bands the theta, delta, alpha, beta, gamma, high gamma, ripple, and fast ripple bands.

In order to isolate the signal into these bands, we applied a standard bandpass filtering on their respective frequency ranges (as seen in Table 4.2), which allowed the construction of band-specific datasets for comparative model evaluation and more complex model building.

Frequency Band	Range (Hz)
Delta	0.4 – 4
Theta	4 – 8
Alpha	8 – 13
Beta	13 – 30
Gamma	30 – 70
High Gamma	70 – 100
Ripple	100 – 150
Fast Ripple	150 – 200
Multi-Unit	200 – 500

TABLE 4.2: Used neural frequency bands and their associated frequency ranges.

With this resulting data, as seen in Fig. 4.4, we could create targeted ESNs, or combine the information into our more complex hierarchical ESNs.

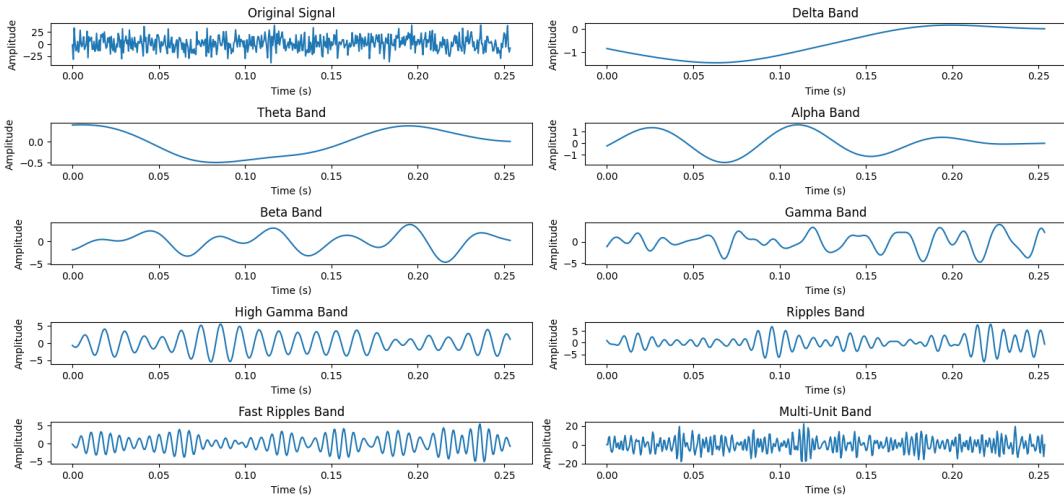


FIGURE 4.4: Resulting time series of each of the isolated frequency ranges.

4.2 Data Processing

4.2.1 Spectrogram

Initially in this project, our plan was to implement conventional deep learning (DL) models, specifically Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs, to classify the neural signals. However, following a shift

in focus from purely maximizing accuracy to prioritizing interpretability and neuro-scientific insight, we made the decision to restart the project and instead adopt Echo State Networks (ESNs) as our primary modeling approach.

During the initial deep learning phase, we sought to utilize the spatial sensitivity of CNN-based architectures by transforming our raw time series data into a time-frequency representation using the Short-Time Fourier Transform (STFT). This process yielded a spectrogram matrix, where frequency components, binned in intervals of 5 Hz, were tracked across four temporal windows. In addition to being well-suited for convolutional processing, this representation also facilitated some of our early preprocessing strategies, as outlined in Section 4.1.2.

Although this spectrogram-based approach has been shown to improve performance in deep learning models (Li et al. (2022); Ramos-Aguilar et al. (2020); Zabidi et al. (2012); Schirrmeyer et al. (2017); Wang et al. (2024)), we found that the benefit disappeared once we transitioned to ESNs. As a result, the spectrogram conversion was removed from the final pipeline in favor of preserving the raw temporal structure of the LFP signals, which is better suited for usage in ESNs.

4.2.2 Principal Component Analysis (PCA)

We also explored dimensionality reduction using Principal Component Analysis (PCA), for both increased computational efficiency and potential relevance of specific channels.

While PCA with 50 components preserved nearly all of the signal variance and had a high accuracy score as seen in Table 4.3, the reduction in computational cost was smaller than expected and, due to the transformations introduced by PCA, added an unnecessary layer of abstraction, making the outputs harder to interpret. As a result, this step was ultimately removed from the final pipeline in favor of maintaining clearer traceability between input features and model behavior.

PCA Components	Accuracy (%)
5	51.43
10	60.63
30	75.24
50	78.73
100	81.59

TABLE 4.3: Classification accuracy using different numbers of PCA components. Note that results are lower than expected due to limited preprocessing and number of training epochs.

4.3 Modeling Approaches

With our data preprocessed, we started constructing and fitting our models. For our purposes, we selected Echo State Networks (ESNs), as although deep learning models are often used on neural recording classification and analysis, specially if transformed previously into spectrograms as mentioned in Section 4.2.1, ESNs offer a balance between interpretability, accuracy and time computational cost which better suited our purposes, and are also widely underrepresented in current literature.

While recent ESN literature has proposed increasingly complex and stacked ensemble architectures to enhance performance (Gallicchio, Micheli, and Pedrelli (2017); Yang et al. (2022); Ma, Shen, and Cottrell (2017); Kawai, Park, and Asada (2019); Enel et al. (2016)), this study adopts a dual perspective, weighing both performance and interpretability. In doing so, we align more closely with the original approach of Jaeger (2001), adn offer a clearer and more simple ESN model.

4.3.1 Simple ESN

In the baseline model, the entire preprocessed LFP signals (without region or frequency band separation) for each trial were sent and fed channel per channel into our ESN. This simple version of our ESN served as the control against which all more complex variants were compared, as well as a template for all other models.

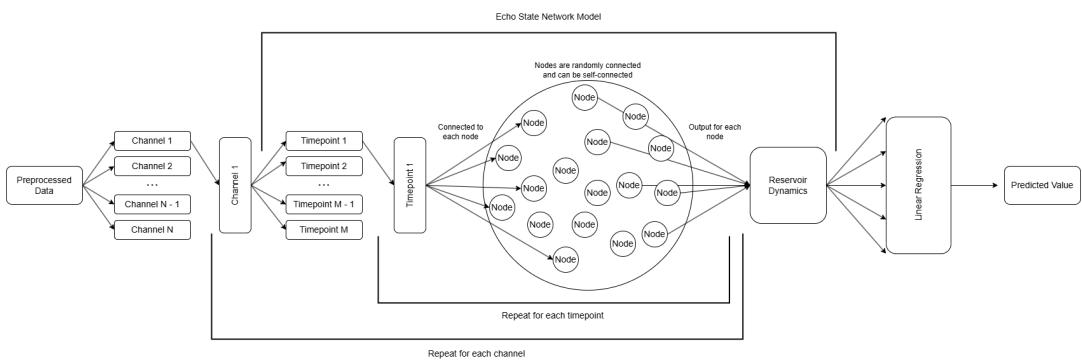


FIGURE 4.5: Simple Echo State Network Model

4.3.2 Region-Based ESN

As mentioned in Section 3.2, the LFPs were applied to a total of 5 different brain regions. Each of this brain regions, while playing an important role, present different patterns or importance based on each motion-state. To study this, we decided to isolate each brain region separately, and treat them as unique inputs, using the region-channel equivalence table given by the data, as seen in Table 4.4.

Region	Start Channel	End Channel
l_PMv	1	64
l_PMd	65	96
r_PMd	97	160
r_PMv	161	224
l_M1	225	256

TABLE 4.4: Channel index ranges for each brain region as indicated by DePass, 2022.

With our channels grouped according to their predefined brain regions (e.g., l_PMv, l_PMd, r_PMd, r_PMv, l_M1), we proceed to train and evaluate how our simple model would do if given only the information of a single channel, as seen in Fig. 4.6.

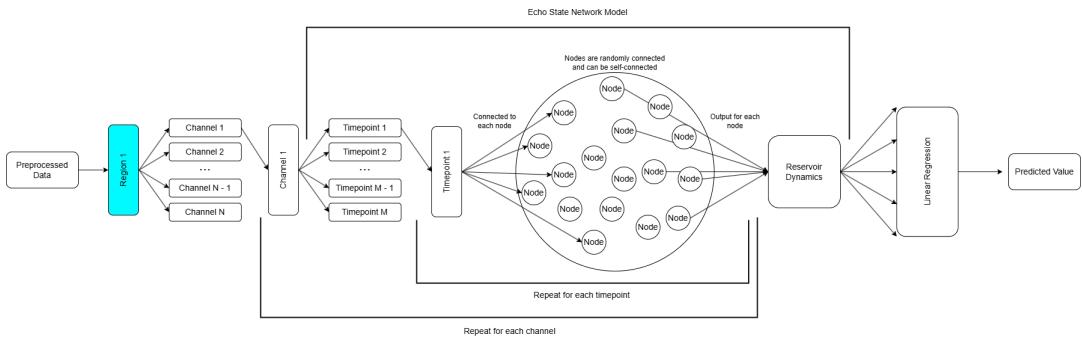


FIGURE 4.6: Single Region Echo State Network Model. Differences from simple ESN marked in blue.

In order to be able to also exploit and interpret the inter-region behavior and region role in each state, we also created a more complex hierarchical model, where we would apply the ESN to each region, and then combine all the inputs of each region into a single linear regression, as seen in Fig. 4.7.

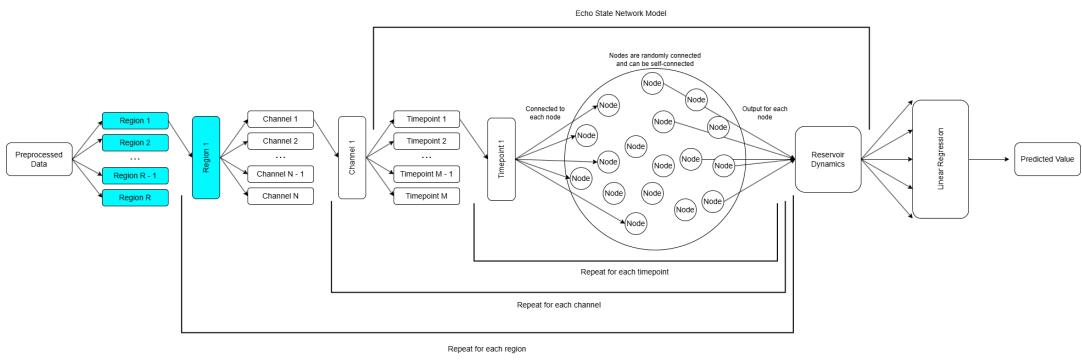


FIGURE 4.7: Multi-Region Echo State Network Model. Differences from simple ESN marked in blue.

4.3.3 Bandwidth-Based ESN

Similarly to each brain region, each frequency band also plays a different role in our brains, and to also study how each frequency band differs, we isolated each

frequency band as explained in Section 4.1.4 (see Table 4.2) and trained an ESN on the corresponding signal subset, as seen in Fig. 4.8, from which we studied the variation of accuracy as well as the resulting confusion matrices, in order to better interpret if any band specialized in any one state.

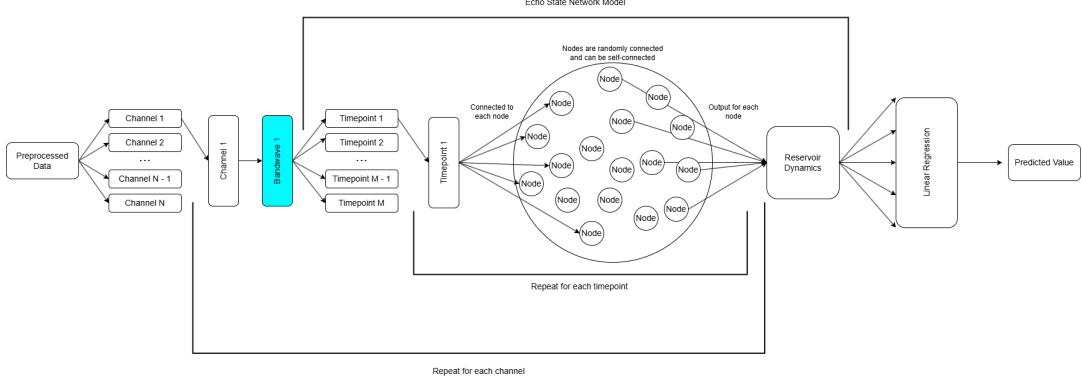


FIGURE 4.8: Single Bandwidth Echo State Network Model. Differences from simple ESN marked in blue.

Also mirroring the process done for the regions, we created the equivalent complex hierarchical model, which built a ESN that took as inputs the output reservoir states of the ESNs built on each region, as seen in Fig. 4.9.

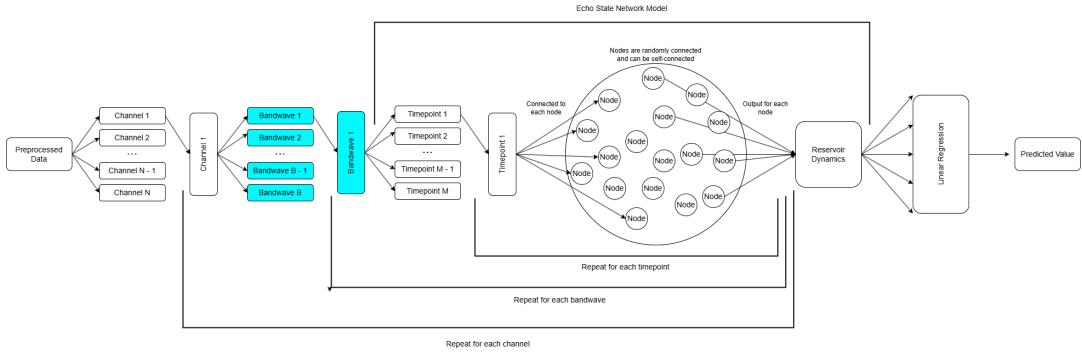


FIGURE 4.9: Multi-Bandwidth Echo State Network Model. Differences from simple ESN marked in blue.

4.3.4 Region-Bandwidth Combined ESN

Finally, in order to capture interactions between spatial and spectral features (frequency bands and regions), we created a complex hierarchical model with three different layers.

In said model, for each brain region, the signal was split into frequency bands, and an ESN was trained on each. The resulting reservoir states were hierarchically combined first within each region across bands and then across all regions, as seen in Fig. 4.10.

While this makes interpretation more complex, it is still possible with the right tools, and crucially, it allows a more extensive view on what role each frequency played in each region, as well as how they all interacted to define the neural signature.

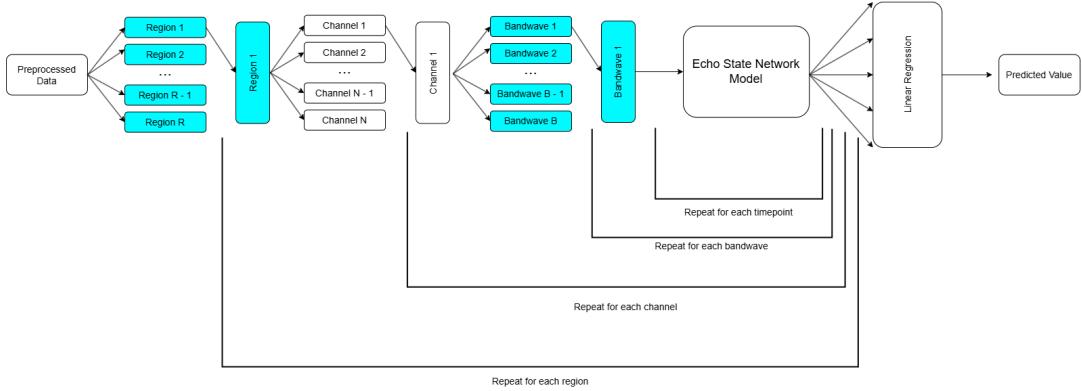


FIGURE 4.10: Region-Bandwidth Echo State Network Model. Differences from simple ESN marked in blue and ESN model compressed for clarity.

4.4 Model Parameters, Training and Evaluation

As mentioned in Section 3.3, when designing and training ESNs, there are a set of parameters one must decide, which affect the reservoir dynamics and behavior. We decided to use a spectral radius of 0.95, an input scaling factor $\alpha = 0.5$, and a sparsity of 10%, as they are generally considered to be values ideal for reliable reservoir behavior to ensure a light model not prone to overfitting.

We also implemented the output layer as a simple linear regression model, as it was the most interpretable model, which was trained for and evaluated our model performance on a stratified validation set, with all reported accuracy metrics based on this split. To interpret the results, we also examined the resulting confusion matrices, which provided insights into the separability of movement phases and highlighted the relative contributions of different input configurations.

4.5 Analyzing Movement Dynamics

To better understand what the ESNs are actually learning, we implemented a two-level feature analysis. The first, more straightforward approach involves applying Principal Component Analysis (PCA) to the output weight matrices \mathbf{W}_{out} , reducing them to three components. This allows us to represent the neural signatures learned by the model in a low-dimensional space that can be visualized and compared across classes.

For a more detailed and structural inspection, we analyzed how the reservoir dynamics vary across both classes and timepoints. Although we initially planned to

analyze the full reservoir-to-output space across all dimensions, this quickly became impractical. With 500 reservoir units, 508 timepoints, and 10 classes, the total structure spans over 2.5 million values, that both makes computation and meaningful direct visualization infeasible. To manage this, we narrowed the analysis to two more targeted strategies: one focused on how each reservoir unit contributes on average per class (across time), and one focused on how each timepoint is weighted across all classes (averaging across the reservoir units).

4.5.1 Principal Component Analysis

The Principal Component Analysis (PCA) is a non-linear dimensionality reduction methodology that is able to determine the data projection that maintains the highest possible data variance. In our case, we apply PCA to the full \mathbf{W}_{out} matrices, each shaped as 500 reservoir units by 508 timepoints, and reduce them to three components per class.

This gives us a way to visualize and compare the overall neural signature for each class in a 3D space. These low-dimensional projections allow for interpretable inspection of signature overlap and separation between classes, and could also be used to generate artificial signatures by perturbing coordinates within the known class clusters.

4.5.2 Classwise Reservoir Contributions

To investigate which reservoir nodes contribute most to each class and try to establish similar patterns between classes, we collapsed the temporal axis by averaging each \mathbf{W}_{out} matrix across timepoints. This gave us one 500-dimensional vector per class, indicating average reservoir influence, and were visually inspected using graphs and heatmaps. We also computed pairwise distances between them using both the Euclidean norm and cosine similarity to quantify how distinct or similar the reservoir contributions are between classes.

4.5.3 Timepoint Dynamics per Class

We also examined whether certain classes rely more on specific timepoints. For this, we collapsed the reservoir axis by taking the mean of \mathbf{W}_{out} values across all 500 reservoir units, producing a time series per class of shape (508,). However, since averaging alone may obscure class-specific dynamics, we also computed a broader set of statistics at each timepoint, including standard deviation, min, max, median, mode, and the 25th and 75th percentiles. These were plotted as heatmaps, and again compared using Euclidean and cosine distances to assess temporal alignment or divergence between classes.

Chapter 5

Results

5.1 Modeling Results

One of the main parameters we had to balance during execution was reservoir size. The larger the reservoir is, the more complex the representations it is able to make, but the less interpretable and more computationally costly it becomes.

To find a good trade-off, we empirically tested a set of possible values, as seen in Table 5.1, and determined that a reservoir size of 500 was the best fit for our needs.

Please note that the values seen are on fast preprocessed data with few training epoches. When applied to the fully preprocessed data, 50 nodes got an accuracy of 82% and 71% in 5 and 10 class predictions.

Reservoir Size	Accuracy (%)
50	60.32
100	60.12
150	63.12
200	64.07
500	65.03
750	66.82
1000	64.67
1500	66.54

TABLE 5.1: Classification accuracy obtained at varying ESN reservoir sizes. Low accuracy due to fast preprocessed data and few epochs.

5.1.1 Simple ESN

The base ESN achieved approximately 84% classification accuracy across 5 movement classes with no noticeable confusion between classes, as seen in Fig. 5.1.a. On the other hand, the accuracy for 10 classes was lowered to 79%, with an expected slight overlap or mirroring on the equivalent movement classes, as seen in Fig. 5.1.b.

This result confirmed that the reservoir model, despite its simplicity, could effectively distinguish among grasping phases based solely on raw LFP input, and even

what hand was used, which validated our preprocessing pipeline and confirmed the temporal structure of neural dynamics was preserved and properly learned to a surprisingly high degree.

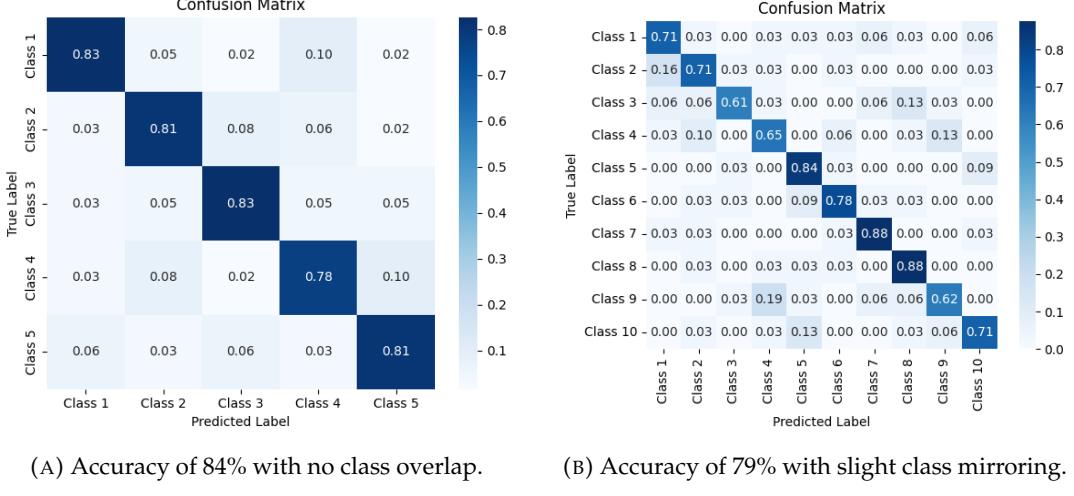


FIGURE 5.1: Confusion matrix for the simple ESN.

5.1.2 Region-Based ESN

Individually training ESNs on each brain region revealed accuracies ranging from 82% to 85% and from 73% and 79% on our 5-class and 10-class classification, as seen in Table 5.2. These findings seem to indicate a very distributed nature of motor representation across the cortex, to a considerably surprising degree, as all accuracies in the single region ESNs are almost identical, especially in the 5-class case.

Brain Region	5-Class Accuracy (%)	10-Class Accuracy (%)
l_PMv	84.44	79.37
l_PMD	84.13	73.65
r_PMv	82.54	76.83
r_PMD	83.13	77.46
l_M1	82.27	79.05

TABLE 5.2: Classification accuracy per brain region using region-specific ESNs for both 5-class and 10-class.

Surprisingly, the combined region model, which pooled reservoir states of each single region ESN into a linear regressor, reached up to 95% and 89% accuracy, an even higher accuracy than any individual region or the simple ESN, and the highest accuracy of the whole study, suggesting that each region encodes complementary aspects of the task, and combining them we can achieve equivalent or superior results than most simple deep learning techniques. As for the confusions, while the 5-class task didn't present any notable confusion, the 10-class shows some of the most defined class mirroring, as seen in Fig. 5.2, indicating the complexity of identifying the hand used for the task.

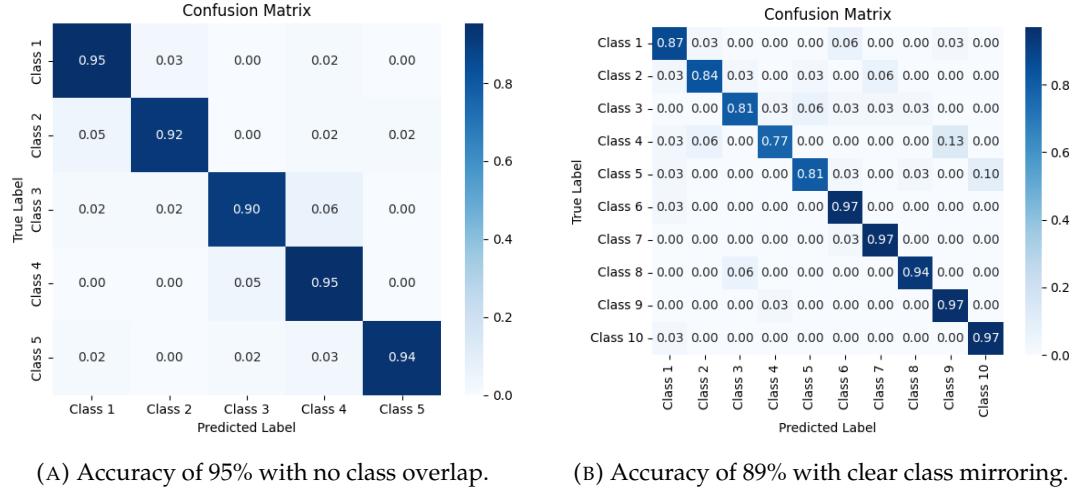


FIGURE 5.2: Confusion matrix for the multi-region ESN.

5.1.3 Bandwidth-Based ESN

On the other hand, when training on individual frequency ranges, the performance varied significantly by frequency band, with higher intensities seemingly being more representative, and consequentially, getting higher accuracy scores, as seen in Table 5.3.

Specifically, the Multi-Unit reaches accuracies of up to 84%+ and 74%+in the 5 and 10 class predictions, while theta, delta and alpha bands performed poorly (under 50% and 40% in 5 and 10 class prediction respectively), which supports the commonly established hypothesis that higher-frequency components are more closely tied to motor execution. We can also see a dip in accuracy in the High Gamma range, as reported by DePass et al. (2022) in their own analysis.

Frequency Band	5-Class Accuracy (%)	10-Class Accuracy (%)
Delta	38.41	28.57
Theta	45.40	36.19
Alpha	46.76	37.78
Beta	62.22	50.47
Gamma	70.15	52.69
High Gamma	66.66	49.84
Ripple	69.79	55.55
Fast Ripple	71.11	57.14
Multi-Unit	84.76	74.92

TABLE 5.3: Classification accuracy per frequency band using bandwidth-specific ESNs for both 5-class and 10-class tasks.

The combined bandwidth model, which pooled the reservoir states of the ESNs trained on each frequency band individually, reached a 81.2% accuracy with no notable class confusion on the 5-class task, and a 74% with minor class mirroring for the 10-class, as seen in Fig. 5.3.



FIGURE 5.3: Confusion matrix for the multi-bandwave ESN.

5.1.4 Region-Bandwidth Combined ESN

Finally, our most complex model, which pooled the reservoir states of each individually frequency band trained ESN model for each region and then pooled them once more globally, achieved one of the highest overall performance of almost 89% and 82% accuracy, which is surprisingly high for both 5-class and 10-class prediction, rivaling the combined region model. There was however a more pronounced class mirroring, specially for classes 1 and 5, but no significance class overlap for the 5-class as seen in Fig. 5.4. While the accuracy is highly impressive, due to its hierarchical complexity, direct interpretability was reduced. Nonetheless, if studied properly, it can provide valuable insight into the interaction between regional and spectral contributions and highlighted synergies that might not be visible in isolated models.

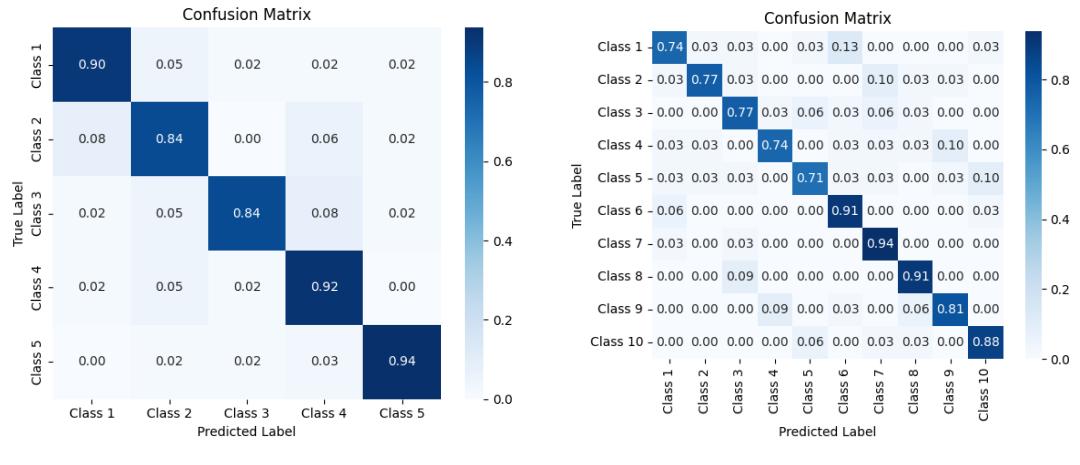


FIGURE 5.4: Confusion matrix for the multi-bandwave ESN.

5.2 Analyzing Movement Dynamics

5.2.1 Principal Component Analysis

To offer a simple interpretation of how our models are identifying each class, we can apply Principal Component Analysis (PCA) to the internal states of the trained ESN. This revealed separable clusters for each movement phase, indicating that the reservoir encodes phase-specific trajectories in its state space, as seen in Fig. 5.5, which we can establish the presence of stable, low-dimensional manifolds corresponding to the motor intentions of the nonhuman primate.

We can also note that in the 10-class PCA, the right and left hand phases have also been properly separated, even though to a smaller degree, indicating the expected similarity of neural signatures between these equivalent phases.

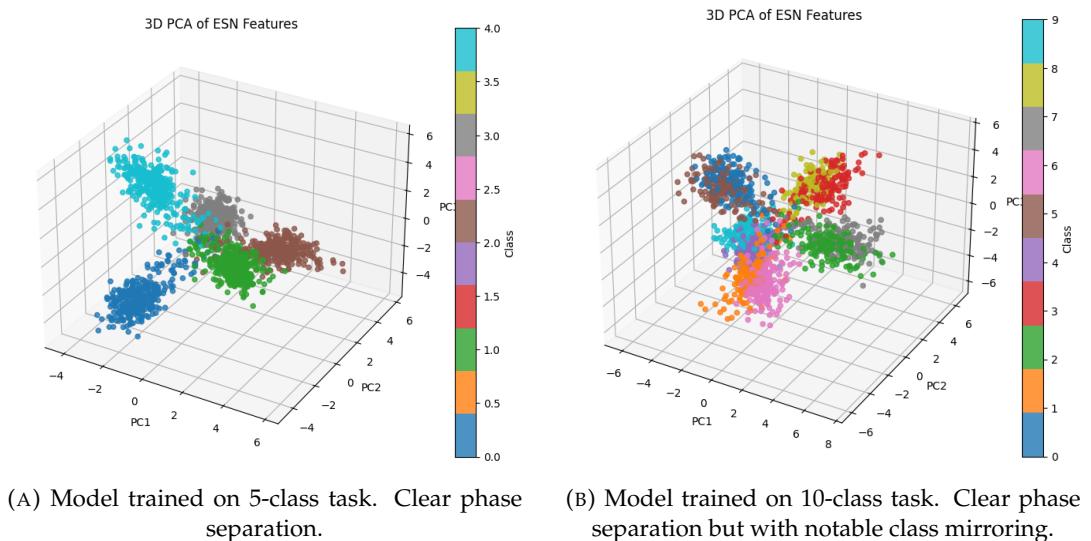


FIGURE 5.5: Human-interpretable visualization of the neurological signatures learned from the nonhuman primate testing by our region-bandwidth combined ESN.

These embeddings offer a first glimpse of how our model interprets the LFP data, the underlying biological behavior of our readings and how a neurological signature vocabulary could be created.

5.2.2 Classwise Reservoir Contributions

To attempt a fuller investigation on how the reservoir varied per class considering both the 5-class and 10-class predictions, we averaged the \mathbf{W}_{out} value of each node throughout the temporal dimension, for a resulting 500 values, using the same ESN for node coherence. In order to run some visual inspections, we then, for visual inspection and comparison, plotted those values on the reservoir network, with the color of each node representing their respective value, as seen in Fig. 5.6 and Fig.

5.7, and created a heatmap for clearer comprehension, as seen in Fig. 5.8, with each of the 500 values being distributed in a 50x10 matrix.

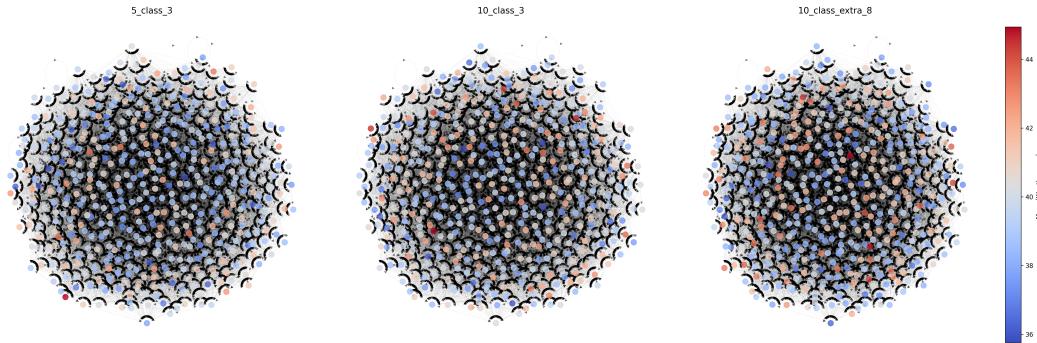


FIGURE 5.6: Comparison between reservoirs for class 3 and their respective counterparts in the 10-class prediction. No visual similarities can be observed when compared to other phases.

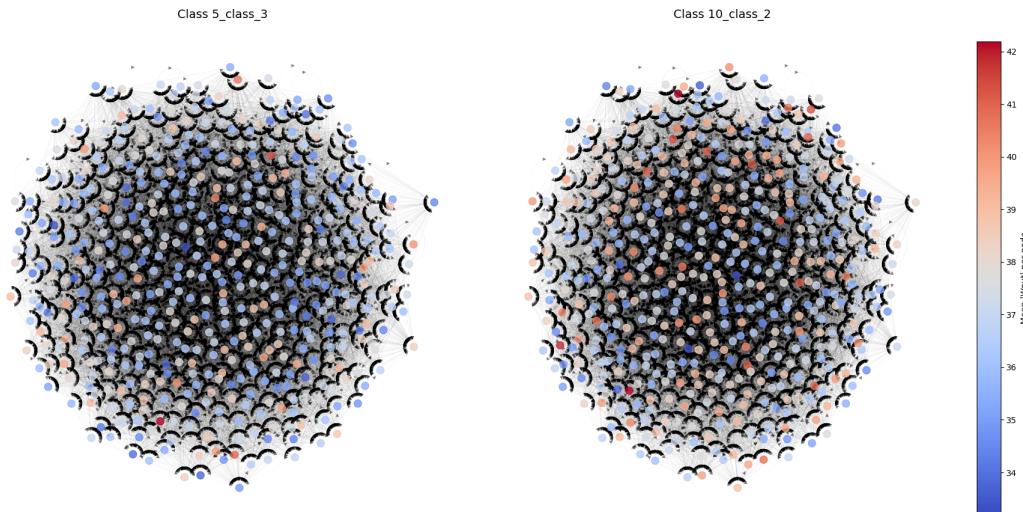


FIGURE 5.7: Comparison between reservoirs for class 3 in 5-class and class 2 in 10-class prediction. No notable difference from previous phase equivalent comparison.

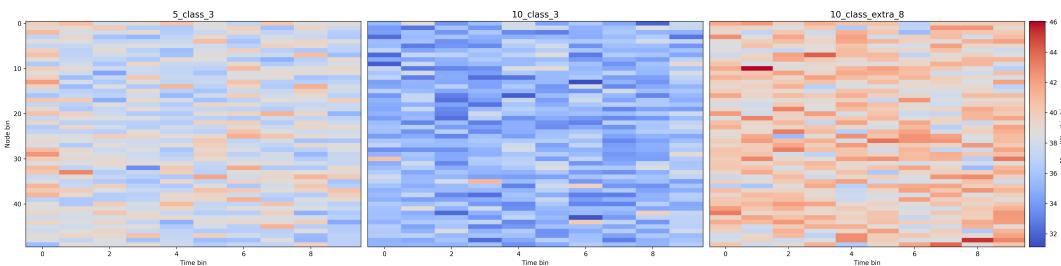


FIGURE 5.8: Heatmap comparison between reservoirs for class 3 in 5-class and class 2 in 10-class prediction. No emergent pattern can be distinguished.

Throughout all these comparisons, we could not find any visual equivalence between same phases for different classes, or any other type of emergent pattern. As

this methodology was overly complex to visualize, we attempted to train our ESN of 50 nodes, reaching similar results, see Fig. 5.9.

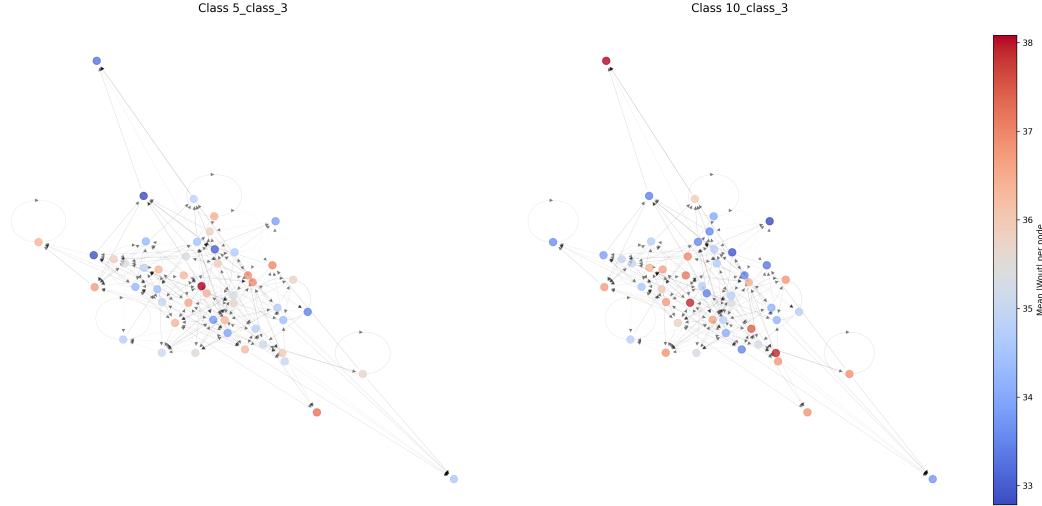


FIGURE 5.9: Network comparison of phase 3 in 5-class and 10-class prediction on a ESN with reservoir of 50. No conclusive patterns emerge.

Following this, we compared not the values themselves, but calculated metric distances between value sets, using both the Euclidean and cosine distance. These distances, obtained from the 500 node and the 50 node reservoir, are plotted in a matrix heatmap with each class to class distance for all 5 classes from the 5-class and the 10 classes from the 10-class prediction, with the sum of the total distances of the row written on the side, as seen in Fig. 5.10 and Fig. 5.11.

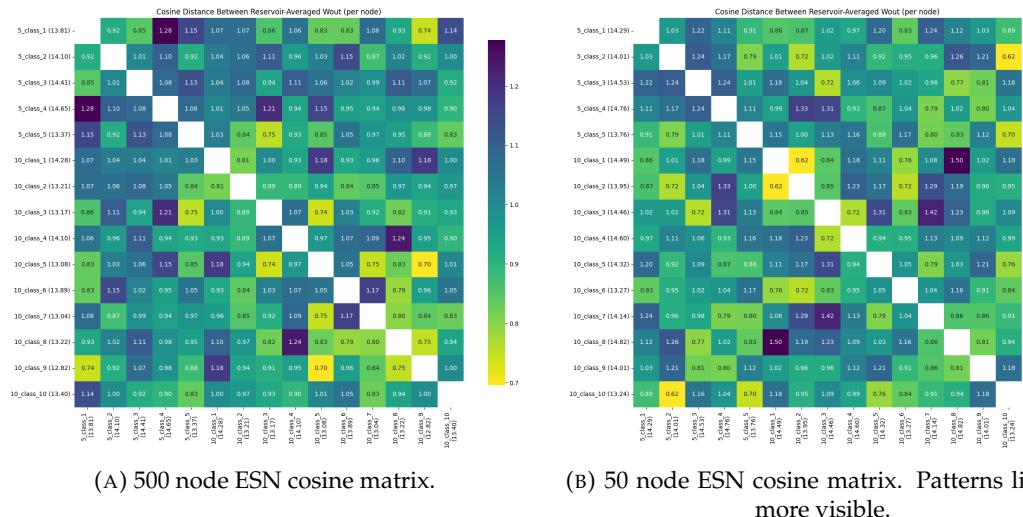


FIGURE 5.10: Cosine matrix heatmap for each time-collapsed vector. While some phase equivalences can be observed between the 5-class phases and their 10-class equivalents, no clear pattern can be seen.

When analyzing the interclass distances, we can observe a slight tendency for phases

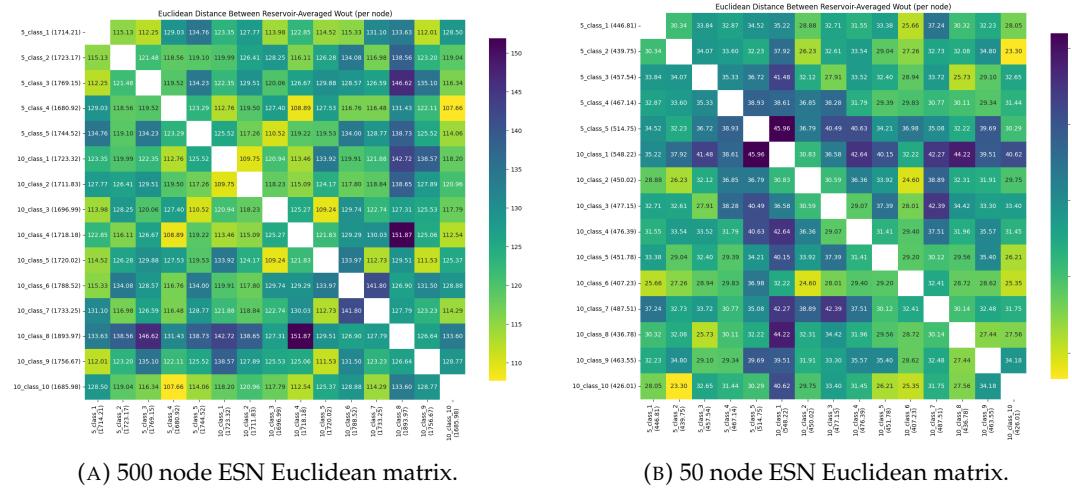


FIGURE 5.11: Euclidean matrix heatmap for each time-collapsed vector. Phase equivalences are more obvious, but no clear pattern can be seen.

from the 5-class to have a higher similarity to their 10-class equivalents, yet no clear patterns emerge once more.

5.2.3 Time Dynamics per Class

We similarly apply the previous logic, but collapsing the reservoir axis by taking statistical values, as explained in Section 4.5.2, of \mathbf{W}_{out} values across all 500 reservoir units, producing a time series per class of shape (508,), which are then represented in a stacked heatmap, as seen in Fig. 5.12.

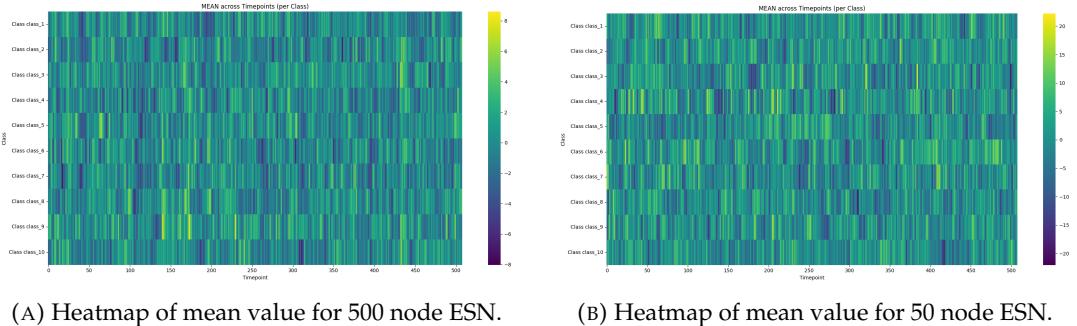


FIGURE 5.12: Heatmap of mean on 10-class prediction. No patterns or phase similarities were observed in mean or any other used value.

As no visual analysis was feasible, we calculated once more the Euclidean and cosine metric similarities, as described in the previous section. The results, seen in Fig. 5.13 and Fig. 5.14, show much clearer patterns, specially the euclidean distances and the comparisons ran on the 50 node ESN. These patters show that the 5 classes in the 5-class prediction are similar to their phase counterparts in the 10-class predictions, but the 10-class predictions themselves do not show the expected phase mirroring we've observed in the PCA.

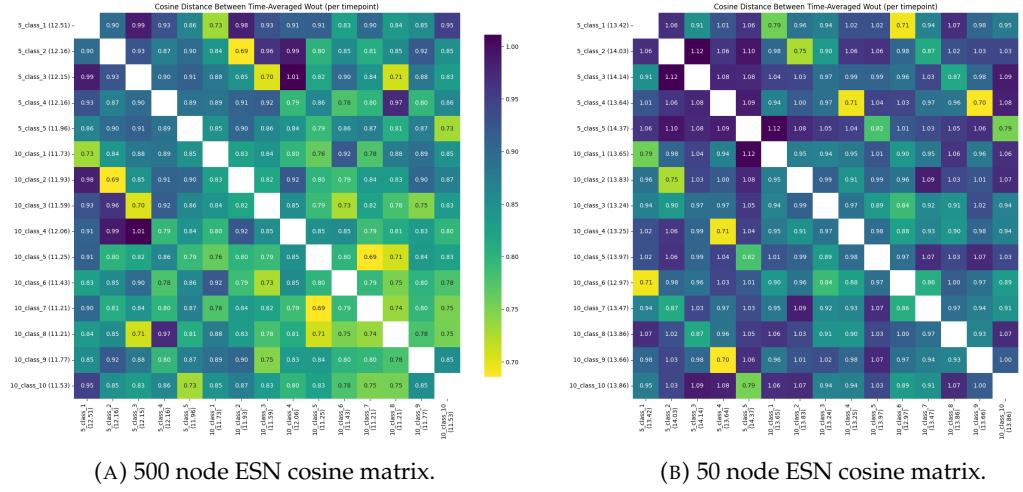


FIGURE 5.13: Cosine matrix heatmap for each reservoir-collapsed vector. While some phase equivalences can be observed between the 5-class phases and their 10-class equivalents, no clear pattern can be seen.

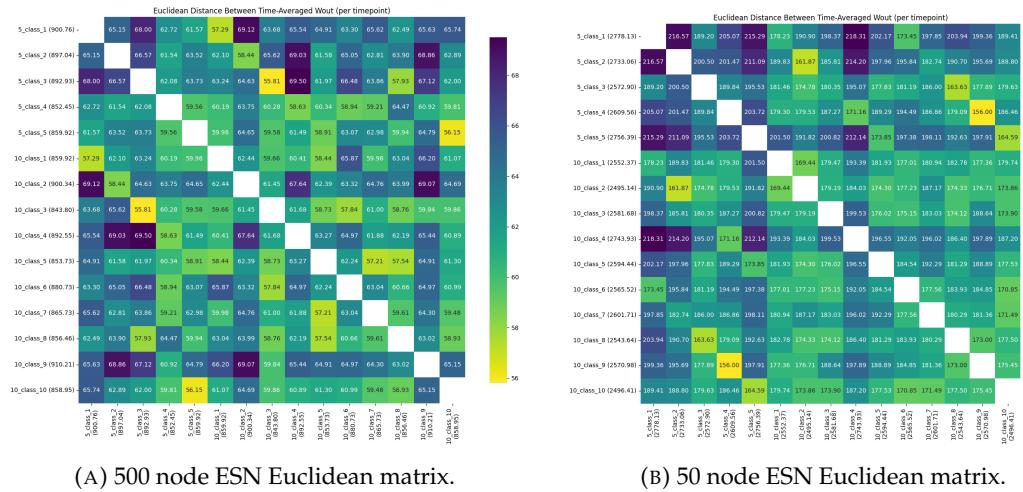


FIGURE 5.14: Euclidean matrix heatmap for each reservoir-collapsed vector. Phase equivalences are more obvious, but no clear pattern can be seen.

5.2.4 Distance Analysis On Full Reservoir Dynamics

The results obtained from the previous non-PCA analysis proved to be less informative than anticipated, which we suspect is due to the loss of a significant portion of information when collapsing either the temporal or reservoir dimensions. While, as mentioned previously in Section 4.5, visual analysis of the full reservoir state is unfeasible, both from a computational and interpretative point of view, applying our euclidean and cosine distances still results in a relative low number of calculations and no significant display burden. Following the same methodology used for the previous two class to class distance comparisons, we obtained Fig. 5.15 and Fig. 5.16, which showed starkly clearer similarities.

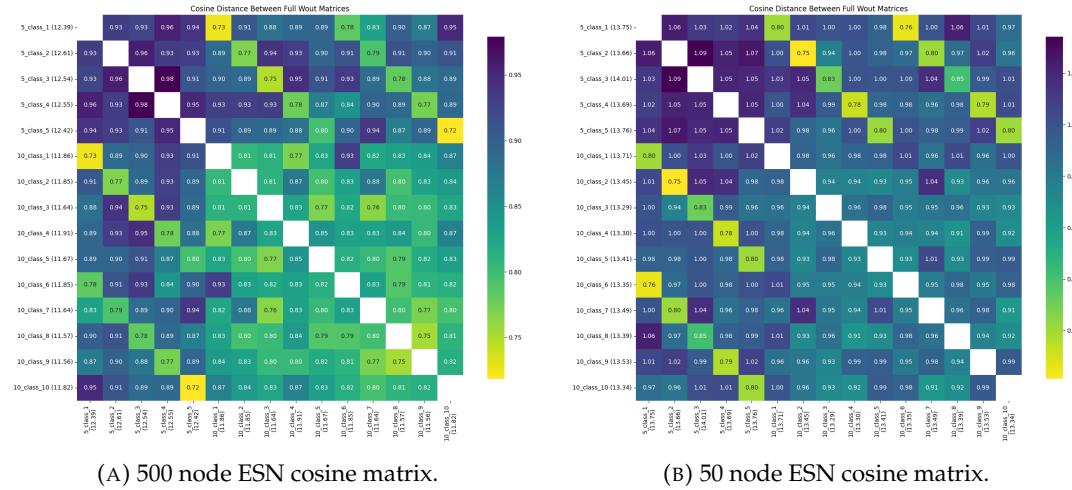


FIGURE 5.15: Cosine matrix heatmap for full flattened reservoir state. Remarkably clear patterns of phase similarities between phases in 5-class and 10-class prediction, but not phase mirroring in 10-class predictions.

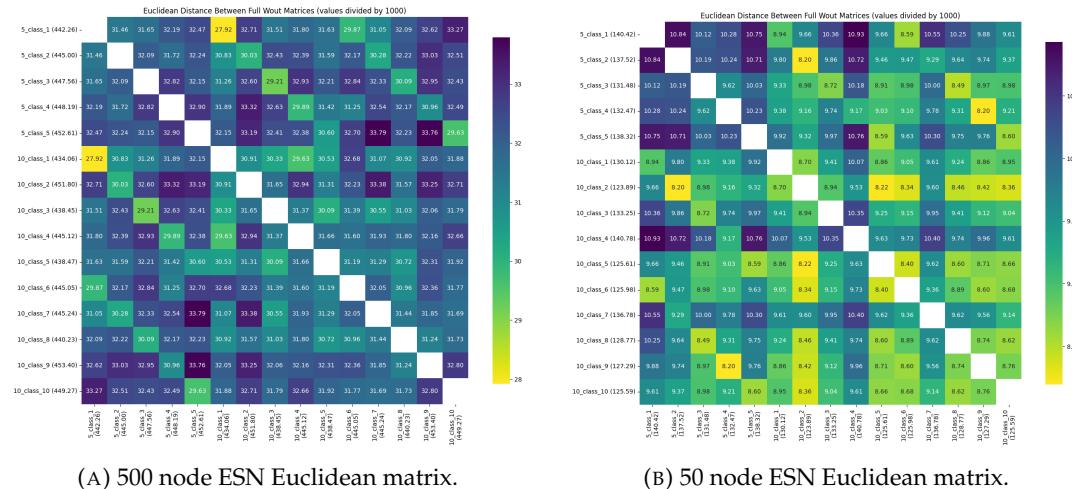


FIGURE 5.16: Euclidean matrix heatmap for full flattened reservoir state. Less clear patterns of phase similarities between phases in 5-class and 10-class prediction, but not phase mirroring in 10-class predictions.

5.3 Understanding Movement Dynamics

While our previous visualizations based on the collapsed temporal and reservoir axes offered limited interpretability, revealed only partial patterns of class-level correlation or distinctive neural signature, our PCA and distance-based analyses provided far more compelling results. These methods revealed, with remarkable clarity, consistent relationships between classes and meaningful structure within the neural signatures.

Through our PCA projections, we were able to demonstrate that each movement phase is associated with a unique, non-overlapping neural signature that occupies a

low-dimensional manifold. In the 10-class ESN, we observed a phenomenon we refer to as phase mirroring: classes corresponding to the same grasping phase but differing in handedness appear close together in PCA space, indicating that the model has captured a shared underlying structure despite their superficial distinction.

In contrast, the Euclidean and cosine distance matrices tell a slightly different story. While the 5-class ESN phases show strong similarity to their phase-matched counterparts in the 10-class ESN, the internal 10-class comparisons do not reflect the same phase mirroring observed in the PCA. This apparent contradiction can be explained by the linear nature of the distance metrics. Since the 10-class ESN must learn to discriminate between highly similar neural signatures, differentiated only by handedness, it is forced to focus on those subtle distinguishing features. This amplifies the perceived distance between mirrored classes in the linear metrics, even though their overall structure remains closely aligned, as shown in the nonlinear PCA projection.

In summary, while linear distance metrics emphasize differentiation for classification, the PCA reveals the broader structural similarity between mirrored phases, validating both the discriminative capacity of the ESN and the coherence of the encoded neural representations.

5.4 Comparison With Joshua Tapper’s Project

When comparing our findings to those presented in the work of Joshua Tapper, who analyzed the same dataset using a more conventional deep learning methodology, we observed a number of meaningful similarities as well as key differences. Tapper’s project focused primarily on maximizing classification accuracy, utilizing Convolutional Neural Networks (CNNs) and CNNs that feed into Long Short-Term Memory (LSTM) models as well as a parallel model that trains a CNN and LSTM separately and concatenates the results, whereas our study prioritized both interpretability and computational performance through the use of Echo State Networks. Despite this divergence in focus, the comparison between our results yielded additional insight into the behavior of both approaches.

It is worth noting that both projects began from the same preprocessing pipeline. As such, the dataset preparation and filtering procedures employed in Tapper’s study mirrored those I developed. This alignment ensures that observed differences in outcomes stem from the modeling strategies themselves rather than discrepancies in data handling.

5.4.1 Accuracy

One of the most notable outcomes of this comparison is the superior performance of the ESN-based models in terms of raw classification accuracy. While Echo State Networks were initially selected for their simplicity and reduced computational overhead, they consistently outperformed the more complex architectures used in Tapper’s study. Our 10-class ESN model achieved an accuracy of 89 percent, compared to the 82 percent reported in his 9-class framework. This performance gap is particularly striking given the increased granularity and notably lower complexity of our class structure, using 500 trainable parameters compared to the 250000 - 750000 parameters used in Tapper’s study (0.05% to 0.2% of the parameters).

5.4.2 Manifold Expressiveness

In addition to accuracy, the internal organization of model representations differed significantly. Our PCA-based visualizations revealed distinct phase clustering, with mirrored grasping phases, such as left and right hand movements of the same type, appearing closely aligned in low-dimensional space. This pattern was absent in Tapper’s visualizations, suggesting that the ESN’s reservoir structure captures a more coherent and discriminative representation of temporal neural dynamics. These results highlight the capacity of the ESN to encode functional similarities in a way that is more interpretable and visually separable, as well as being able to separate classes the deep learning models haven’t been able to, like the right- and left-handed baseline.

5.4.3 Handedness Preference

Both studies observed that tasks involving the right hand were consistently classified with higher accuracy than those involving the left. This finding could be attributed to several factors, including a possible innate handedness of the non-human primate, a slight imbalance in the dataset’s class distribution, or the presence of a higher number of recorded regions in the left hemisphere of the brain, which controls right-sided motor function.

5.4.4 Frequency Bands and Regions

With regard to the relevance of specific frequency bands and brain regions, our methodological paths diverged. Tapper employed SHAP values, channel ablation, and Integrated Gradients to estimate importance, and his results largely confirmed our findings for spectral contributions. Higher frequency bands were consistently found to carry more discriminative information across both studies, with the notable exception of the high gamma, as also noted by DePass et al. (2022).

However, regional relevance results varied considerably. Tapper’s findings showed fluctuating importance across different models and interpretability techniques, with some analyses indicating dominance of specific regions such as l_PMv, l_PMD, and r_PMD. Our results, in contrast, suggested a more evenly distributed contribution across all recorded regions, with each region independently supporting classification and yielding even greater performance when combined.

We believe this discrepancy may be explained by the ESN’s ability to maintain and process temporal dependencies through its reservoir structure. As described in Section 5.4.2, this structure allows the model to integrate information over time, leading to more stable and consistent estimations of each region’s relevance. In contrast, static attribution methods applied to deep learning architectures may be more susceptible to noise and redundancy, which could account for the variation in Tapper’s results.

Chapter 6

Discussion

The results presented above show the potential of Echo State Networks (ESNs) as effective and interpretable tools for decoding neural dynamics from Local Field Potential (LFP) data during a motion-based task.

The classification performance of the ESN models, particularly in the multi-region and the region-band configuration, demonstrate the ability to differentiate motion-based neurological signatures with equivalent accuracy to deep learning techniques, while maintaining their interpretability.

One of the most relevant findings is that, from the recorded regions, all contributed significantly yet similarly to classification accuracy. This seems to support the idea that all regions play an important role, with specialized tasks and distributed information throughout all of them, as seen by the multi-region model's higher accuracy when compared to single region models.

Similarly, the performance drop observed in the single-bandwidth models indicates that relevant information is distributed across frequencies, with more significant information in the higher frequency bands. The combination of multiple frequency bands into a unified model also proved to be an effective strategy, yet had an equal accuracy than our simple model, suggesting that our time-frequency decomposition is not useful from an accuracy standpoint, but still gains potential insight.

The region-bandwidth composite model yielded strong performance as well, yet its deeper architecture complicates interpretability. However, with the correct analysis tools, this complexity could be exploited to study the more nuanced interconnection between bands from different regions, leading to even further insights into the nature of each brain region and the role of the frequency bands.

In this context, our PCA- and distance-based analysis proved especially valuable. While representations based on mean or equivalent values across the reservoir or time dimensions showed little class-level structure and interpretability, nonlinear PCA projections revealed a clear separation between phases and a low-dimensional

organization of neural signatures. The mirrored phases in the 10-class model further suggest that the ESN was encoding shared features across grasping phases, even in the presence of handedness differences. These observations were further validated through cosine and Euclidean comparisons, although the linearity of the latter underrepresents the underlying neural similarity between mirrored classes in the 10-class model.

When compared to related work, our ESN approach performs robustly, achieving both high classification accuracy and strong interpretability. Although a direct comparison with studies such as Schirrmeister et al. (2017) may be misleading due to differences in task complexity and dataset scale, our results align well with the findings of DePass et al. (2022). In particular, our multi-region ESN is comparable to the highest accuracies in their study, while simultaneously offering greater interpretability and significantly reduced architectural complexity. Furthermore, our model perform equally as good as the framework developed by Joshua Tapper if not better, which used a combination of CNN and LSTM networks on the same dataset. Despite the deeper structure and focus on raw accuracy in his approach, our 10-class ESN achieved higher performance while maintaining a clearer and more structured internal representation of the neural dynamics, as visualized when exploring their respective ma.

Overall, these findings provide a promising foundation for the broader application of ESNs in cognitive neuroscience, particularly for the purpose of comprehension and identification of different neural behaivorl patterns, especially if interpretability and computational efficiency are prioritized over purely performance-driven goals, as well as helping to establish the utility of LFPs.

Chapter 7

Conclusion

In this work, we explored the utility of Echo State Networks (ESNs) for classifying motor-based brain states using Local Field Potential (LFP) recordings. By analyzing neural activity recorded during a five-phase grasping task in a nonhuman primate, we demonstrated that ESNs can reliably decode different behavioral phases with high accuracy while preserving interpretability.

Our layered approach, starting with a basic ESN and progressively incorporating region-specific and bandwidth-specific variations, enabled us to examine the contribution of both anatomical and spectral dimensions to classification. The results confirm that ESNs, despite their architectural simplicity, are capable of revealing meaningful structure in brain dynamics, and represent them in a human-interpretable format.

The bandwidth-specific models aligned well with prior assumptions and published literature, showing that higher frequency bands carried more relevant information for motor-task decoding. However, the results from the region-specific models were more unexpected. Each of the five recorded brain regions, when used independently, achieved over 82% accuracy on the 5-class task—despite existing literature often reporting unequal involvement across regions. Initially, we considered the possibility of signal bleed between channels as a confound, but the significant performance gain observed when combining regions in a hierarchical structure (achieving 96%+ accuracy) suggests otherwise, as these results point toward each region holding unique, non-redundant information relevant to the grasping task

Our visualization analyses, particularly those based on PCA and pairwise distance metrics, offered valuable insight into the internal structure of the ESN representations. While collapsing the temporal or reservoir dimensions yielded limited interpretability, PCA revealed a well-structured low-dimensional organization of the neural signatures—most prominently in the 10-class configuration. In this setting, we observed a consistent pattern of phase mirroring, where left- and right-handed variants of the same grasping phase clustered closely together, indicating that the

ESN encodes a shared underlying representation for each phase, modulated only slightly by handedness.

Complementary to this, the cosine and Euclidean distance matrices revealed a strong similarity between each 5-class phase and its corresponding pair in the 10-class model. This further reinforces the idea that each grasping phase carries a distinct and stable neural signature, which persists even when task granularity is increased.

Notably, the 10-class ESN achieved nearly the same accuracy as its 5-class counterpart, despite the added difficulty of distinguishing subtle within-phase differences. This consistency in performance, along with the structural clarity observed in PCA, highlights the ESN's ability to generalize across similar cognitive states while maintaining interpretable and phase-consistent internal representations.

These findings also serve as a strong validation of using LFP recordings in cognitive decoding tasks, as the richness and resolution of the LFP signals enabled our ESN models to distinguish between a larger set of motor states than typically achieved in similar experiments.

Overall, this work offers evidence not only for the viability of ESNs in neural signal classification, but also for their potential in producing interpretable, high-resolution behavioral decoding tools. By bridging accuracy and insight, we can advance both computational methodology and theoretical neuroscience, and take a step towards a more nuanced and accessible interpretations of complex neurophysiological data.

Chapter 8

Limitations

- **Single-subject dataset:** The analysis is based on data from a single nonhuman primate, limiting the generalizability of the results. Individual differences in brain anatomy and neural dynamics may affect model performance across subjects.
- **Single-task dataset:** Similarly, the data was recorded from a single motor task, which also limits the scope of our study. Different tasks or tasks of a different nature will inevitably show a different neural signature, which may affect the learning of our ESN model and force us to reconfigure it.
- **Limited interpretability tools:** While ESNs are more interpretable than many other ML models, our methods for extracting neural signatures, while effective, could be enhanced with advanced statistical or visualization techniques, specifically a fine grained analysis of the region-bandwidth combined ESN or other hierarchical ESNs could lead to a more complete understanding of the interconnection between brainwave frequencies and regions.

Chapter 9

Future Work

- **Generalization across subjects and tasks:** Future studies should include multiple subjects and diverse movement tasks to test the robustness and transferability of the ESN models.
- **Generalization across tasks of different nature:** It will also be necessary to test if this methodology can be implemented successfully as is to tasks of a different nature (non-motor), or modifications will be required.
- **Context-dependent modeling:** If both previous tasks can be completed, introducing contextual embeddings could allow ESNs to adaptively weigh different phases or conditions, improving both accuracy and interpretation and automatically adjusting the ESN to individuals and tasks, allowing increased performance and a comprehensible way of comprehending the impact tasks and individuals hold on the neural signatures.
- **Temporal vocabulary creation:** With sufficient labeled datasets, it may be possible to develop a structured vocabulary of neural patterns associated with specific motor actions, which could lead to a variety of improvements across multiple fields, such as real-time EEG and LFP translation, cognitive state monitoring, and the identification of biomarkers in the neural signature for medical disorders like ADHD, PTSD or depression.

Appendix A

Code and GitHub Link

The code used in this project can be found in the following GitHub repository:

Neural Signal Analysis

For any questions or inquiries about the repository, please feel free to reach out at
lucaeric.dicroce@gmail.com.

Appendix B

Author Contributions and Rights

The data and some algorithms, as motioned explicitly in the paper when referred to and attributed in the code, are taken from the DePass et al. (2022) paper. All appropriate rights are reserved to their respective authors.

The algorithm `automatic_bad_channel_detection` was taken from Manuel Hernandez, who improved on the existing code published in Komosar, Fiedler, and Haueisen (2022).

Appendix C

Usage of LLMs throughout the project

The use of ChatGPT (free online version) was used throughout this project for code structure, debugging and commenting. No code logic was written using ChatGPT, and all code and comments were manually checked and modified to ensure validity and consistency.

The use of ChatGPT (free online version) was also used in the writing of this report for paragraph clarity when deemed necessary. ChatGPT did not generate any of the original text or structure, and all text written by ChatGPT was rewritten by us to better suit the text flow and ensure validity and consistency.

Bibliography

- Davare, Marco et al. (2009). "Ventral premotor to primary motor cortical interactions during object-driven grasp in humans". In: *Cortex* 45.9, pp. 1050–1057.
- DePass, Michael et al. (2022). "A machine learning approach to characterize sequential movement-related states in premotor and motor cortices". In: *Journal of Neurophysiology* 127.5, pp. 1348–1362.
- Enel, Pierre et al. (2016). "Reservoir computing properties of neural dynamics in prefrontal cortex". In: *PLoS computational biology* 12.6, e1004967.
- Fourati, Rahma et al. (2017). "Optimized echo state network with intrinsic plasticity for EEG-based emotion recognition". In: *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part II* 24. Springer, pp. 718–727.
- Gallicchio, Claudio, Alessio Micheli, and Luca Pedrelli (2017). "Deep reservoir computing: A critical experimental analysis". In: *Neurocomputing* 268, pp. 87–99.
- Jaeger, Herbert (2001). "The "echo state" approach to analysing and training recurrent neural networks-with an erratum note". In: *Bonn, Germany: German national research center for information technology gmd technical report* 148.34, p. 13.
- Jeong, Dong-Hwa and Jaeseung Jeong (2020). "In-ear EEG based attention state classification using echo state network". In: *Brain sciences* 10.6, p. 321.
- Kakei, Shinji, Donna S Hoffman, and Peter L Strick (1999). "Muscle and movement representations in the primary motor cortex". In: *Science* 285.5436, pp. 2136–2139.
- Kawai, Yuji, Jihoon Park, and Minoru Asada (2019). "A small-world topology enhances the echo state property and signal propagation in reservoir computing". In: *Neural Networks* 112, pp. 15–23.
- Komosar, Milana, Patrique Fiedler, and Jens Haueisen (2022). "Bad channel detection in EEG recordings". In: *Current Directions in Biomedical Engineering*. Vol. 8. 2. De Gruyter, pp. 257–260.

- Li, Chengfan et al. (2022). "A deep learning method approach for sleep stage classification with EEG spectrogram". In: *International journal of environmental research and public health* 19.10, p. 6322.
- Ma, Qianli, Lifeng Shen, and Garrison W Cottrell (2017). "Deep-esn: A multiple projection-encoding hierarchical reservoir computing framework". In: *arXiv preprint arXiv:1711.05255*.
- Ramos-Aguilar, Ricardo et al. (2020). "Feature extraction from EEG spectrograms for epileptic seizure detection". In: *Pattern Recognition Letters* 133, pp. 202–209.
- Schirrmeister, Robin Tibor et al. (2017). "Deep learning with convolutional neural networks for EEG decoding and visualization". In: *Human brain mapping* 38.11, pp. 5391–5420.
- Tang, Xue-song et al. (2023). "A moth–flame optimized echo state network and triplet feature extractor for epilepsy electro-encephalography signals". In: *Mathematics* 11.6, p. 1438.
- Wang, Xianheng et al. (2024). "An in-depth survey on Deep Learning-based Motor Imagery Electroencephalogram (EEG) classification". In: *Artificial intelligence in medicine* 147, p. 102738.
- Yang, Xiaojian et al. (2022). "An improved deep echo state network inspired by tissue-like P system forecasting for non-stationary time series". In: *Journal of Membrane Computing* 4.3, pp. 222–231.
- Zabidi, A et al. (2012). "Short-time Fourier Transform analysis of EEG signal generated during imagined writing". In: *2012 international conference on system engineering and technology (ICSET)*. IEEE, pp. 1–4.
- Zhang, Heng and Danilo Vasconcellos Vargas (2023). "A survey on reservoir computing and its interdisciplinary applications beyond traditional machine learning". In: *IEEE Access* 11, pp. 81033–81070.