Decoding Intentions from Micro-Electrode Recordings of Neuronal Ensembles in Primates

Alireza Rouzitalab

THESIS SUBMITTED TO THE UNIVERSITY OF OTTAWA
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DOCTORATE IN PHILOSOPHY DEGREE
IN ELECTRICAL AND COMPUTER ENGINEERING

SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
FACULTY OF ENGINEERING
UNIVERSITY OF OTTAWA

Ph.D. in Electrical and Computer Engineering (2023) University of Ottawa

Electrical Engineering and Computer Science Ottawa, Ontario, Canada

TITLE: Decoding Intentions from Micro-Electrode Recordings of Neuronal Ensembles in Primates

AUTHOR: Alireza Rouzitalab

M.Sc., Electrical Engineering

Iran University of Science and Technology

SUPERVISORS: Dr. Jeongwon Park

Dr. Adam J. Sachs

NUMBER OF PAGES: 140

Abstract

Neuronal activities in the brain encode every decision, desire, or intention. Multiple brain regions are involved in translating intention into action. Detecting and decoding intentions directly from the brain could allow impaired individuals to communicate and interact with their environment despite central nervous system dysfunction.

Brain-computer interface (BCI) systems access neuronal activities and translate them into actions using a computer. BCIs are used in research studies to replace, restore, or replace neuromuscular functions. In addition, BCIs provide new insights into how the brain works, aiding in new treatments for neurological conditions.

BCI studies commonly target the primary motor cortex, the region of the brain most closely associated with volitional muscle control, with the expectation that signals from its neurons will be best suited for control of external effectors. Consequently, other brain regions are underrepresented in BCI studies.

This thesis focuses on two brain regions in primates with access to higher-order control over intention and movement: The prefrontal cortex and the basal ganglia system. These areas are vital for naturalistic movement and must be more widely explored for decoding intentions. We aim to find the movement information while the intentions have yet to transfer into planning.

One study in macaque monkeys explored eye movement intention, learning, and memory-related circuitry in the lateral prefrontal cortex (LPFC). In an eight-target saccade task, we could decode the target to which the monkeys would saccade before the eye movement began. Moreover, we decoded the abstract rule information acquired by the monkeys to find the correct target from the neuronal activities recorded from LPFC. In addition, the memory-related activities in LPFC were linked to monkeys' behaviour as evidence of the presence of working- and long-term memory circuitry in the prefrontal cortex.

In another study on Parkinson's disease (PD) patients, we explored the possibility of volitional control of brain activities, which can lead to a self-induced procedure to reduce the symptoms of PD. We recorded the local field potentials (LFP) of the subthalamic nucleus (STN) of nine PD patients performing a cognitive task during deep brain stimulation surgery. The patients could modulate their brain activities to change the colour of a central sphere to match the colour of a peripheral cue in a virtual reality task. They modulated the signal power in beta frequencies (13-30 Hz) and the rate of beta bursts (the fast episodes of changing amplitude in a short period in LFP's beta frequencies) based on the task conditions. Both beta power and beta bursts are associated with the pathological state in PD patients. A decodable volitional modulation of both presents the STN as a valuable region for BCI studies which could lead to self-regulation of PD symptoms.

The findings of this thesis contribute to the advancement of therapeutic systems used for various brain disorders like PD and Amyotrophic lateral sclerosis (ALS), as well as patients with disabilities that can benefit from assistive communicative technologies.

The study on the LPFC increased the decoding accuracy of saccade intentions compared to previous studies. Additionally, decoding associative rules is beyond the complexity of previous studies. We also showed the effects of previously learned associations on the learning rate of new rules and how this memory-retrieved information modulates neuronal activities.

Moreover, the study on the STN showed the volitional control of beta power and beta burst rates by PD patients, which can be used as therapeutic methods to improve the severity of the symptoms of PD.

Résumé

Les activités neuronales dans le cerveau codent chaque décision, désir ou intention. Plusieurs régions du cerveau interviennent dans la traduction de l'intention en action. La détection et le décodage des intentions directement à partir du cerveau pourraient permettre aux personnes atteintes de troubles du système nerveux central de communiquer et d'interagir avec leur environnement.

Les systèmes d'interface cerveau-ordinateur (ICO) accèdent aux activités neuronales et les traduisent en actions à l'aide d'un ordinateur. Les ICO sont utilisés dans des études de recherche pour remplacer, restaurer ou remplacer les fonctions neuromusculaires. De plus, les ICO fournissent de nouvelles informations sur le fonctionnement du cerveau, aidant ainsi à développer de nouveaux traitements pour les maladies neurologiques.

Les études sur les ICO ciblent généralement le cortex moteur primaire, la région du cerveau la plus étroitement associée au contrôle musculaire volontaire, en raison de l'attente que les signaux de ses neurones soient les mieux adaptés pour contrôler les effecteurs externes. Par conséquent, d'autres régions du cerveau sont sous-représentées dans les études sur les ICO.

Cette thèse se concentre sur deux régions cérébrales chez les primates qui sont censées avoir un contrôle de haut niveau sur l'intention et le mouvement: le cortex préfrontal et le système des ganglions de la base. Ces domaines sont essentiels pour le mouvement naturaliste et sont largement inexplorés en termes de décodage des intentions. Notre objectif est de trouver des informations sur le mouvement alors que les intentions ne se sont pas encore transformées en planification.

Une étude menée sur des singes macaques a exploré les circuits liés aux mouvements oculaires intentionnels, à l'apprentissage et à la mémoire dans le cortex préfrontal latéral (LPFC). Dans une tâche de saccades à huit cibles, nous avons pu décoder la cible vers laquelle les singes saccaderaient avant que le mouvement des yeux ne commence. De plus, nous avons décoder les informations de règle abstraite acquises par les singes pour trouver la cible correcte à partir des activités neuronales enregistrées dans le

LPFC. De plus, les activités liées à la mémoire dans le LPFC étaient liées au comportement des singes en tant que preuve de la présence de circuits de mémoire à court et à long terme dans le cortex préfrontal.

Dans une autre étude sur des patients atteints de la maladie de Parkinson (PD), nous avons exploré la possibilité de contrôle volontaire des activités cérébrales, qui peut conduire à une procédure auto-induite pour réduire les symptômes de la PD. Nous avons enregistré les potentiels de champ local (LFP) du noyau sous-thalamique (STN) de neuf patients atteints de la DP effectuant une tâche cognitive pendant une chirurgie de stimulation cérébrale profonde. Les patients ont été en mesure de moduler leurs activités cérébrales pour changer la couleur d'une sphère centrale pour correspondre à la couleur d'un indice périphérique dans une tâche de réalité virtuelle. Ils ont modulé la puissance du signal dans les fréquences bêta (13-30 Hz) et la fréquence des rafales bêta (les épisodes rapides de changement d'amplitude sur une courte période dans les fréquences bêta de LFP) en fonction des conditions de la tâche. La puissance bêta et les rafales bêta sont tous deux associées à un état pathologique chez les patients atteints de la DP. Une modulation volontaire décodable des deux présente le STN comme une région précieuse pour les études BCI qui pourraient conduire à l'autorégulation des symptômes de la DP.

Les conclusions de cette thèse contribuent à l'avancement des systèmes thérapeutiques utilisés pour divers troubles cérébraux tels que la maladie de Parkinson et la sclérose latérale amyotrophique (SLA), ainsi que pour les patients handicapés pouvant bénéficier de technologies communicatives d'assistance.

L'étude sur le cortex préfrontal latéral a augmenté la précision du décodage des intentions de saccades par rapport aux études précédentes. De plus, le décodage des règles associatives dépasse la complexité des études antérieures. Nous avons également montré les effets des associations préalablement apprises sur le taux d'apprentissage de nouvelles règles, ainsi que comment ces informations récupérées de la mémoire modulent également les activités neuronales.

De plus, l'étude sur le noyau sous-thalamique a révélé le contrôle volontaire de la puissance bêta et des taux d'éclatement bêta par les patients atteints de la maladie de Parkinson, ce qui pourrait potentiellement être

utilisé comme méthodes thérapeutiques pour améliorer la gravité des symptômes de la maladie de Parkinson.

Acknowledgments

I want to express my deepest gratitude to my supervisors, Dr. Jeongwon Park and Dr. Adam J. Sachs, for their support and guidance throughout my degree. Thank you for allowing me to study at such a renowned university surrounded by incredible researchers, colleagues, and students.

To Dr. Chadwick Boulay, please accept my deepest gratitude for supporting my research and guiding me towards valuable tools and methods, helping my Ph.D. research and future career.

I want to thank other members of the Sachs Lab for creating a positive and safe environment for research.

I would also like to thank the members of my thesis advisory committee, Dr. Martin Bouchard and Dr. Adrian Chan for their help, guidance, and support throughout my degree.

I would also like to thank the engineering graduate office for their support and guidance.

To my parents, I am grateful for always believing in and supporting me to achieve higher academic success.

Finally, I extend my heartfelt gratitude to my wife, Rana. Thank you for helping me and giving me the strength and hope to overcome the challenges I faced through my graduate studies.

Contributions

This study explored the possibility of decoding movement intentions before/without physical movement. We improved the decoding performance of saccadic eye movements using a custom deep neural network. The model's architecture can be used in other classification tasks with complex time series signals as inputs. In addition, we showed how learning modulates the neuronal activity decodable by our model. Moreover, we showed the learning rate of new rules based on previously learned rules and how the memory affects the LPFC neuronal activity.

Additionally, we showed that PD patients can modulate their STN activities volitionally. They performed a virtual reality task while controlling the STN beta power level and beta burst rates. Moreover, the signal processing pipeline cleaning the raw signal, filtering out the noise, and removing the outliers can be used prior to feature extraction and decoding.

List of Peer-Reviewed Journal Articles

A. Rouzitalab, C. B. Boulay, J. Park, and A. J. Sachs, "**Intracortical brain-computer interfaces in primates: a review and outlook**," Biomed. Eng. Lett., May 2023, doi: 10.1007/s13534-023-00286-8. **(Chapter 2)**

Author Contribution: Writing – original draft, A.R.; writing – review & editing, A.R., C.B.B., J.P., and A.J.S.; supervision, J.P. and A.J.S.

A. Rouzitalab, C. B. Boulay, J. Park, J. C. Martinez-Trujillo, and A. J. Sachs, "**Ensembles code for associative learning in the primate lateral prefrontal cortex**," Cell Reports, vol. 42, no. 5, May 2023, doi: 10.1016/j.celrep.2023.112449. **(Chapter 3)**

Author Contribution: Conceptualization, J.C.M.T.; methodology, A.J.S.; software, A.J.S.; validation, A.R.; formal analysis, A.R. and C.B.B.; investigation, C.B.B.; resources, J.C.M.T. and A.J.S.; data curation, A.R.; writing – original draft, A.R.; writing – review & editing, A.R., C.B.B., J.C.M.T., and A.J.S.; supervision, C.B.B., J.P., J.C.M.T., and A.J.S.; project administration, A.J.S; funding acquisition, J.C.M.T. and A.J.S.

A. Rouzitalab, C. B. Boulay, and A. J. Sachs, "**Volitional Control of Beta Activities in Parkinson's Disease Patients**," Brain Research, p. 148394, May 2023, doi: 10.1016/j.brainres.2023.148394. **(Chapter 4)**

Author Contribution: Alireza Rouzitalab: Software, Methodology, Validation, Formal analysis, Writing – original draft, Visualization. Chadwick B. Boulay: Methodology, Investigation, Data curation, Writing – review & editing. Adam J. Sachs: Conceptualization, Investigation, Writing – review & editing, Supervision, Funding acquisition.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALS	Amyotrophic lateral sclerosis
ARAT	Action Research Arm Test
BCI	Brain-Computer Interface
CNN	Convolutional Neural Network
CNS	Central Nervous System
DBS	Deep Brain Stimulation
DNN	Deep Neural Network
DOF	Degree of Freedom
EEG	Electroencephalography
fMRI	functional Magnetic Resonance Imaging
FOOOF	Fitting Oscillations and One Over F
GAM	Generalized Additive Model
GPi	Globus Pallidus Internus
GRU	Gated Recurrent Unit
HP	High-Performance
KF	Kalman Filter
LFP	Local Field Potential
LP	Low-Performance
LPFC	Lateral Prefrontal Cortex
LSTM	Long Short-Term Memory
M1	Primary Motor Cortex
MUA	Multi-Unit Acitivity
NHP	Non-Human Primate
NIRS	Near-Infrared Spectroscopy
NSLR-HMM	Naïve Segmented Linear Regression based on Hidden Markov Models
OCD	Obsessive Compulsive Disorder
PCA	Principal Component Analysis
PD	Parkinson's Disease
PFC	Prefrontal Cortex
PMDc	Premotor Cortex
PPC	Posterior Parietal Cortex

RBF	Radial Basis Function
RF	Random Forest
RLR	Regularized Logistic Regression
RNN	Recurrent Neural Network
RSS	Rule Similarity Score
SGD	Stochastic Gradient Descent
SMA	Supplementary Motor Area
STN	Subthalamus Nucleus
SUA	Single-Unit Activity
SVM	Support Vector Machine
t-SNE	t-Distributed Stochastic Neighbor Embedding
UPDRS	Unified Parkinson's Disease Rating Scale
VIM	Ventral Intermediate Nucleus
WM	Working Memory

Chapter 1:
Background

1.1. Overview

Since the first recording of the human brain signal in 1924, this bio-signal has been the subject of countless studies as it represents the activity of the most vital organ in the body. Almost 100 years later, the field of neuroscience, along with engineering and mathematical sciences, extracts unique information from brain activities that change our perspective on how this complicated organ works and simultaneously helps humans treat various disorders.

The brain works as a distributed system, i.e., various parts are essential and involved in the decision-to-action cycle. However, some regions and circuits are more known to be involved in the movement. This thesis focuses on two brain areas essential for movement: the prefrontal cortex (PFC) and the subthalamic nucleus (STN). Chapter 2 reviews the advances in brain-computer interfaces (BCIs), where neuronal activities are recorded from inside the brain. The *intracortical* signals carry valuable fast-changing information modulated by the brain's activities.

Chapter 3 explores the monkeys' lateral prefrontal cortex (LPFC) for saccade intentions, learning, and memory. This brain area holds essential information for movement planning, yet the LPFC needs to be thoroughly explored for BCIs. Our results show promising insights into what we can extract from LPFC in primates.

Chapter 4 explores the STN area of human individuals. The STN has been found as a primary target for Parkinson's disease (PD) research. Our results indicate that PD patients can be trained to control an essential indicator of PD with the help of an STN-driven BCI.

Chapter 5 discusses the results from the previous chapters and the outlook for further research that can benefit from our results.

As per the nature of this article-based thesis, some information might overlap across chapters, especially in this chapter and chapter 2. Nonetheless, in this chapter, we cover the essential topics to understand the following chapters sufficiently.

1.2. Brain Regions

1.2.1. Neuron Doctrine

The understanding of the nerve cells began by hypothesizing that nerves transmit fluid from the brain through the spinal cord to the body's periphery. However, this hypothesis was only confirmed once the microscope revealed the actual structure of cells in nervous tissue. A research study in the late 1800s produced detailed and accurate descriptions of nerve cells, developing a method of staining neurons with silver salts to reveal their entire cell structure under the microscope. This study discovered that nervous tissue is not a continuous web of elements but a network of discrete cells. This work led to the development of critical concepts and early evidence for the neuron doctrine. The neuron doctrine states that the basic unit of the nervous system is the neuron, an individual cell that processes and transmits information. Various studies supported the proof of neuron doctrine by showing the growth of dendrites and axons from isolated neurons and introducing electron microscopy [1]–[8].

The neuron doctrine is still considered a fundamental principle in modern neuroscience. However, recent research has expanded on this view by highlighting the importance of neural networks and the interactions between neurons. It is now understood that neurons do not function in isolation but rather communicate and interact with one another to form functional networks that underlie

behaviour and cognition [9]. Additionally, research has revealed the dynamic nature of neural connections, with the formation and elimination of synapses occurring throughout an organism's lifetime. This suggests the brain's plasticity - the ability to adapt and change in response to experience [10].

1.2.2. Structural Localization

The central nervous system (CNS) comprises the spinal cord and the brain. Subsequently, the brain is divided into functionally specific areas. Recent views in neuroscience oppose the modular distribution of functions into different divided areas. However, brain imaging during specific human tasks shows the behaviour-specific engagement of different brain regions. For example, the frontal lobe is heavily engaged during action planning, working memory [11], and the parietal lobe is engaged in processing sensory information [12]. The outermost layers of the brain are the gray matter or the cerebral cortex. The cerebral cortex is organized so that each hemisphere is primarily concerned with sensory and motor processes on the opposite side of the body. The areas of the cortex that were first pinpointed as essential for cognition were concerned with language. The study of aphasia, a language disorder that occurs when certain areas of brain tissue are destroyed, led to important discoveries in the neural science of human behaviour. The first area of the brain associated with a specific function was found in a language study [13]. Later, a neural processing model was proposed in another study that explains how the brain processes spoken and written language [14], [15].

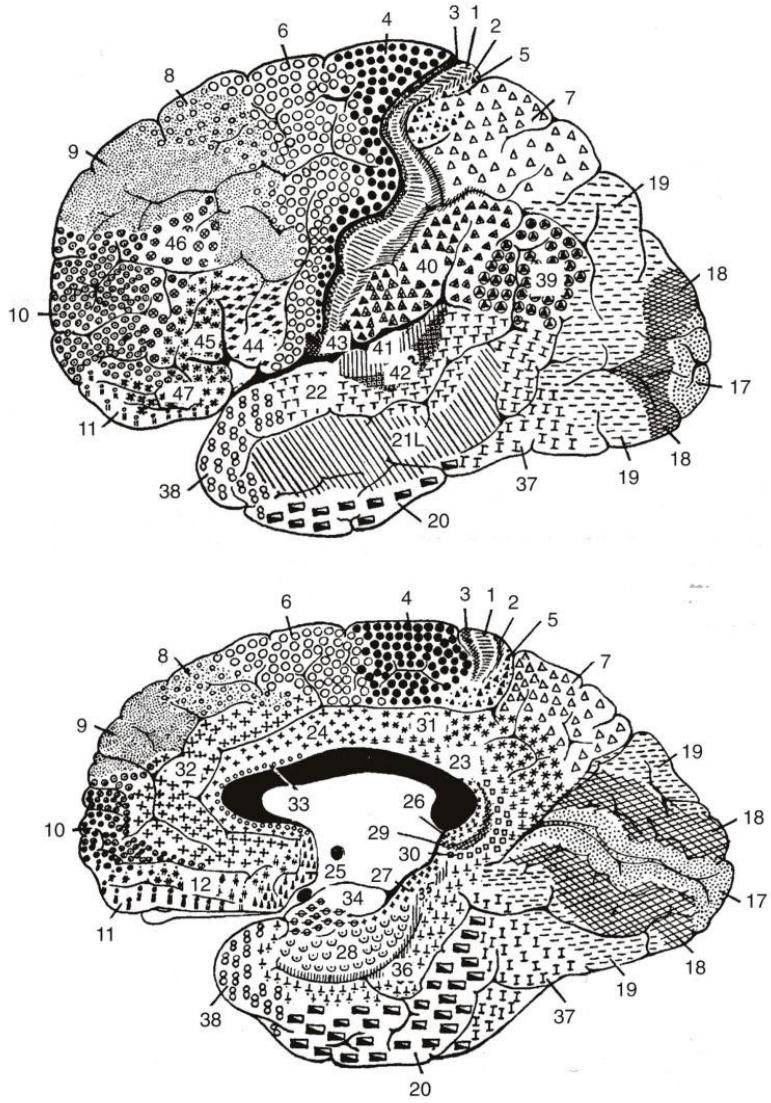

Figure 1.1. The Brodmann areas dividing the brain based on neuron cells' size and shape. Published in [16] and reproduced in [17]. (Top) The lateral view of the map. (Bottom) Medial view of the map.

A new point of view on cortical localization distinguished functional areas of the cortex based on the shapes of cells and variations in their layered arrangement [16]. This study analyzed the distribution of neurons based on their size and shape. This research led to identifying 52 other regions of the cerebral cortex, referred to as Brodmann's areas (Figure 1.1). Brodmann's framework is still widely employed in neuroscience. However, the map is refined and updated frequently.

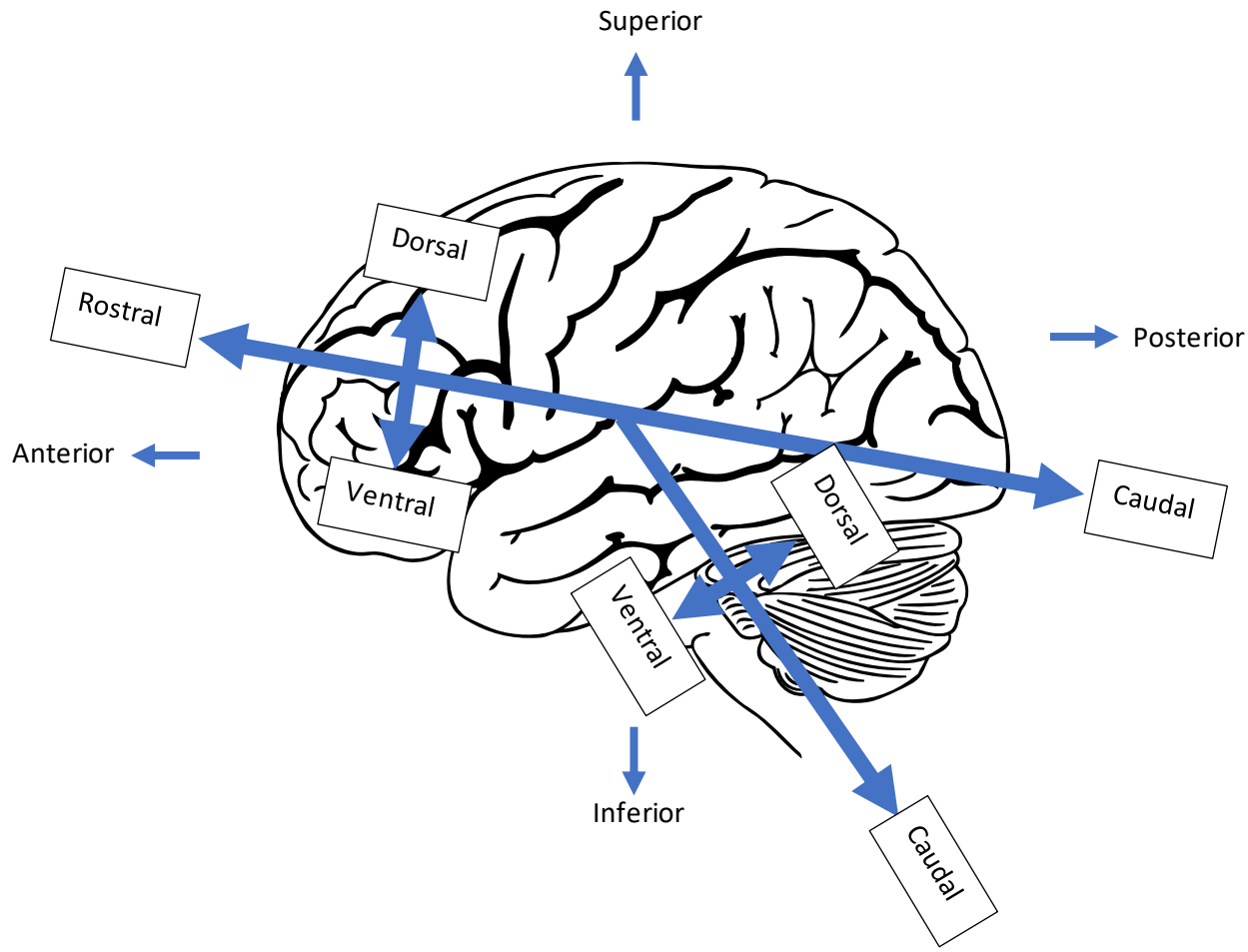

Figure 1.2. The axis of reference for the central nervous system. The ventral-caudal axis is split in two in primates as they walk on two legs. However, in animals walking on four legs, the brain and the spinal cord will have the same ventral-caudal axis [17].

1.2.3. Central Nervous System Organizations

The location and orientation of the CNS are described in a space with axes: the rostral-caudal and dorsal-ventral (Figure 1.2). These references, along with superior (top), inferior (bottom), anterior (front), and posterior (behind), are used in naming different regions and in addressing the position of regions.

The spinal cord, the most caudal part of the CNS, receives sensory information from various body parts in an ascending pathway to the brain. In addition, the spinal cord sends motor commands to muscles in a descending pathway from the brain [17], [18].

The brain comprises six regions: the medulla, pons, midbrain, cerebellum, diencephalon, and cerebral hemispheres. These divisions are found in both hemispheres of the brain with slight differences, and each is further subdivided into several anatomically and functionally distinct areas. The medulla, pons, and midbrain are collectively referred to as the brain stem. The midbrain, the smallest part of the brain stem, lies rostral to the pons. Neurons in the midbrain provide essential linkages between components of the motor system, particularly the cerebellum, basal ganglia, and cerebral hemispheres [19]–[23].

The substantia nigra provides input to a portion of the basal ganglia that regulate voluntary movements. The substantia nigra is the focus of interest as damage to its dopaminergic neurons is responsible for the pronounced motor disturbances characteristic of Parkinson's disease [24]–[26].

The cerebellum is a brain region located at the base of the skull, ventral to the cerebral hemispheres. It plays a crucial role in motor control, coordination, and regulating balance and posture. The cerebellum ensures the smooth and coordinated execution of movement and other motor-related functions [27]–[29].

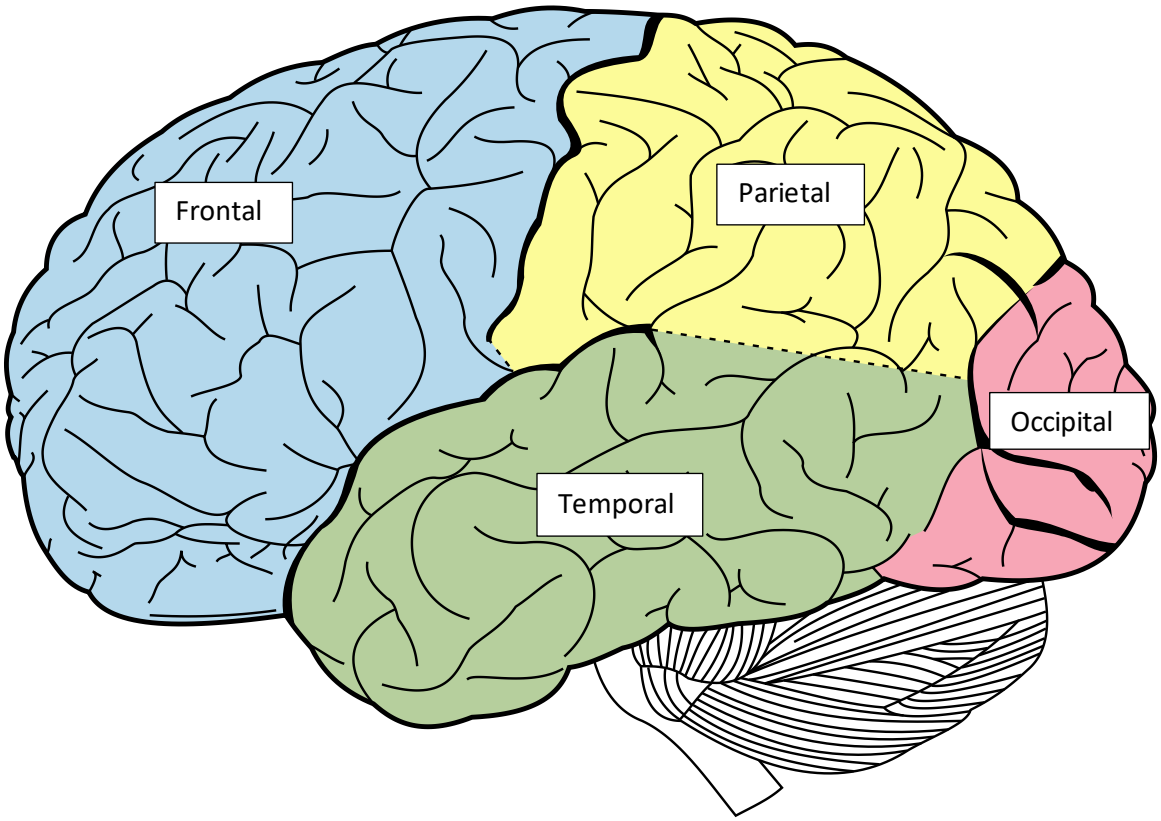

Figure 1.3. The four lobes of the cerebral cortex [30].

The diencephalon, located between the brainstem and the cerebral hemispheres, is essential in regulating various autonomic functions such as hunger, thirst, sleep, and wakefulness and controlling hormonal and cardiovascular systems. The diencephalon comprises several structures, including the thalamus, hypothalamus, and pineal gland. The thalamus relays sensory information, while the hypothalamus regulates various autonomic functions and homeostatic processes. In addition, the pineal gland produces the hormone melatonin, which helps regulate the sleep-wake cycle. Moreover, the diencephalon also involves various higher cognitive functions, such as attention and motivation [31], [32].

The brain's largest and most complex part is the cerebral hemisphere, also known as the telencephalon. The cerebral hemispheres consist of the cerebral cortex, the underlying white matter, and three deep-lying structures: the basal ganglia, amygdala, and hippocampal formation. The two hemispheres are interconnected by the corpus callosum, which is visible on the medial surface of the hemispheres [1].

While the spinal cord, brain stem, and diencephalon mediate many life-sustaining functions, the cerebral cortex is responsible for much of the planning and execution of actions in everyday life [1], [6]. The cerebral cortex includes four main lobes – frontal, parietal, temporal, and occipital – named after the overlying cranial bones (Figure 1.3). Each lobe includes many distinct functional subregions. The temporal lobe, for example, has distinct regions with auditory, visual, or memory functions.

Although the cerebral cortex on both sides of the brain is generally similar, some areas of the cortex on the two sides are functionally distinct. In right-handed people, for example, portions of the left cerebral cortex are specialized for language, whereas the right side of the brain is more related to visuospatial information processing.

1.2.4. The Cerebral Cortex

The primate cerebral cortex is the most elaborate species phylogenetically, with a highly complex shape formed by indentations or grooves known as sulci and bumps or elevations known as gyri. The sulci and gyri separate the cerebral cortex into four distinct lobes. The presence of the gyri and sulci is believed to be an evolutionary strategy to pack more neurons in the skull's limited space [33]. Although the cortex thickness does not vary substantially across different species, with a thickness of approximately 2 to 4 mm, the surface area is dramatically larger in higher primates,

particularly humans. This increase in surface area allows for more cortical neurons, providing a greater capacity for processing information.

The neocortex, the region of the cerebral cortex nearest to the brain's surface, is organized into layers and columns. This arrangement increases the computational efficiency of the cerebral cortex. The complex input-output system of the neocortex is efficiently organized in the orderly layering of cortical neurons. Each layer contains different inputs and outputs and has specific functions.

The neocortex, the surface of the cerebral cortex, comprises six layers [1], [34]:

1. Layer I (molecular layer) is the outermost layer and contains very few cells. It is composed mainly of neuroglia and fibers that support the cells in the other layers.

2. Layer II (external granular layer) contains small, densely packed neurons that receive input from the thalamus and transmit signals to the next layer.

3. Layer III (external pyramidal layer) contains large, triangular-shaped neurons called pyramidal cells. They are involved in transmitting signals to other regions of the brain and the spinal cord.

4. Layer IV (internal granular layer) contains small, densely packed neurons that receive input from the sensory organs. It is involved in the processing of sensory information.

5. Layer V (internal pyramidal layer) contains large pyramidal neurons that play a role in transmitting signals to other regions of the brain and the spinal cord.

6. Layer VI (multiform layer) contains a mixture of large and small neurons. It transmits signals to the brain's subcortical regions and regulates eye movements.

The study detailed in Chapter 3 recorded layers III and IV.

1.3. Cognition and Movement

1.3.1. Neural Maps

Neural maps are the brain's ability to represent information as a spatial map: The brain can use different locations within its neural network to encode and store information about different objects, features, or events. The concept of neural maps helps to explain how the brain processes and remembers sensory information.

For example, in the visual system, the brain creates a neural map of the visual world by using different brain regions to process different aspects of visual information, such as colour, motion, and form. Similarly, the brain creates a neural map of sound information in the auditory system by processing different aspects of sound, such as frequency and location.

These neural maps are created through experience-dependent plasticity, which is the ability of the brain to change and adapt in response to sensory inputs. Over time, the brain can modify the connections between neurons in response to the patterns of sensory information it receives, allowing it to build and refine neural maps representing our experience of the world.

Moreover, the brain creates a map of the body's surface. Early studies on animals recorded various brain regions while stimulating different body parts to find regions in which neuronal activities correlate with stimuli. These studies produced neuronal action potentials by stimulating specific body surfaces [35]. Subsequently, the human somatosensory cortex was mapped by stimulating

the cortex and asking the patients what they felt (sensory homunculus). Similarly, the primary motor cortex was mapped to various muscles (motor homunculus).

It is important to note that the bigger the associated area in the brain, the more precise and detailed the sensory and movement of the mapped area. For example, the index finger occupies more surface than the trunk in the somatosensory and the primary motor cortex.

In addition, these mappings can change with practice and experience, even in adults. For example, if a retired soccer player starts playing basketball, the associated motor area mapped to the foot will shrink while the arm will expand.

Many research studies have explored these maps in particular cases. For example, patients with a congenital fusion of the fingers (syndactyly) have one shared area mapped to both fingers. However, after surgically removing the fingers from each other, new mapping emerges in the same brain area to represent the two fingers separately [36].

Also, in case of losing a limb, the mapping reorganizes to represent the remaining body parts [37].

1.3.3. Cognition

The brain regions communicate hierarchically, and the information becomes more complex with each transition. The higher areas in this chain concern cognitive functions and are termed *the association cortex* [34].

Four lobes of the cortex comprise parts of the association cortex:

1. The parietal association cortex integrates sensory information from various body parts, such as the skin and muscles, to construct a coherent perception of the world.

2. The temporal association processes and integrates auditory, visual, and linguistic information. It is critical in language processing, particularly in spoken language comprehension. The temporal association cortex also helps us recognize objects, faces, and scenes and forms memories for events and episodes.

3. The limbic association cortex regulates emotions, motivation, and memory. This area of the brain plays a critical role in our emotional experiences. It is involved in processing emotionally-charged stimuli, including the perception of pain, pleasure, and fear. The limbic association cortex also regulates our motivation and drive and forms and retrieves memories, especially those emotionally charged. It is believed to play a vital role in processing memories associated with strong emotions, such as those related to survival.

4. The frontal association cortex is involved in a range of high-level cognitive processes. It is responsible for executive functions such as planning, decision-making, problem-solving, and regulating voluntary movements, speech, and other complex behaviours. In addition, it plays a role in working memory, the temporary storage and manipulation of information used in cognitive tasks. Moreover, it is involved in planning and initiating saccadic eye movements. It has been shown to directly connect to the brainstem and spinal cord regions that control eye movements.

Chapters 2 and 3 discuss the prefrontal cortex, part of the frontal association cortex, in more detail.

1.3.5. Movement

Movement is a complex process involving several brain regions' coordinated activity. The primary areas of the brain involved in the movement include the motor cortex, the basal ganglia, the cerebellum, the brainstem, and the thalamus. The motor cortex encodes the initiation and

coordination of the muscles, while the basal ganglia and the cerebellum are essential for the refinement of voluntary movement. On the other hand, the brainstem, which connects the spinal cord to the brain, is mainly concerned with reflexive movements as well as the control of posture. Finally, the thalamus is an information relay between the body's senses and the brain's cerebral cortex.

Chapter 2 discusses the motor cortex in more detail, especially the primary motor cortex. Finally, chapter 4 explores PD, one of the most widespread diseases concerning the basal ganglia, especially the STN.

1.4. Parkinson's Disease

PD is a neurodegenerative disorder characterized by cell loss in the substantia nigra. At PD symptom onset, this brain region can have lost 30-80% of its neurons compared to unaffected individuals [38]. The earliest documented pathological changes in PD, known as Braak stages 1 and 2, are pre-symptomatic. As the disease progresses to Braak stages 3 and 4, the substantia nigra, areas of the midbrain, and basal forebrain become involved. Finally, pathological changes appear in the neocortex.

The clinical diagnosis of PD is based on bradykinesia (slowness of movement or progressive halts), rigidity, and rest tremor. However, not all of these features may be present; thus, patients need to be referred to specialists in movement disorders for accurate diagnosis. Early diagnostic features may include a change of handwriting with micrographia (abnormally small handwriting), reduced facial expression, and a loss of arm swing on one side.

As the disease progresses, hypophonia (reduced speech intensity), drooling of saliva, and impairment of postural reflexes may develop. In addition, non-motor complications of the disease, such as depression, which occurs in 40% of PD patients, also become more prominent [39].

1.4.1. Unified Parkinson's Disease Rating Scale

The Unified Parkinson's Disease Rating Scale (UPDRS) is a widely used metric for evaluating the severity of PD in patients [40], [41]. It was developed to provide a comprehensive and standardized assessment of the symptoms and progression of the disease. The UPDRS is divided into four parts: Part I measures non-motor symptoms, Part II assesses motor symptoms related to activities of daily living, Part III evaluates motor examination, and Part IV assesses complications of PD therapy.

Part I of the UPDRS assesses non-motor symptoms, such as depression, anxiety, and sleep disturbance, which are common in PD patients and can significantly impact their quality of life.

Part II assesses motor symptoms related to activities of daily living, including activities like dressing, grooming, and eating. In addition, this section helps clinicians evaluate the impact of PD on a patient's daily functioning.

Part III evaluates motor examination, including tremors, rigidity, bradykinesia, and postural instability. This section provides essential information on the progression of the disease and helps clinicians monitor changes in the patient's motor symptoms over time.

Part IV assesses complications of PD therapy, including dyskinesia and wearing off, which are common side effects of PD medications. This section helps clinicians evaluate the effectiveness of current PD treatments and determine if a change in therapy is necessary.

In Chapter 4, analyzing the PD patients' brain activities, the motor-related symptoms are reported as the UPDRS III metric. Here is a general outline of the UPDRS III scoring system:

- 0-5: No or minimal symptoms
- 6-15: Mild symptoms
- 16-25: Moderate symptoms
- 26-35: Severe symptoms
- 36-50: Very severe symptoms

1.4.2. Deep Brain Stimulation

Deep brain stimulation (DBS) is a therapeutic method that uses electrical stimulation to target specific brain areas and has been found as an effective treatment for PD. The technique has been used for over 150 years, with the first demonstration of localized electrical excitability of the motor cortex in 1870 [42]. However, only in the development of human stereotaxic devices in the 1940s could neurosurgeons begin investigating the effects of stimulating deeper structures.

In the early 1950s, intraoperative stimulation was used to identify deep structures before lesioning the globus pallidus or thalamus. However, most reports in the 1950s focused on positive phenomena elicited by stimulation. The idea of treating neurologic disorders with chronic stimulation began to emerge in the 1960s, but stimulation was primarily used for targeting surgical lesions [43].

In the 1990s, implantable pacemaker technology was combined with chronically implanted deep brain electrodes for long-term chronic DBS. Since then, DBS has become increasingly used for

treating various disorders, including essential tremors and Parkinson's disease. For PD, different sites of stimulation provide different clinical effects. For example, thalamic stimulation in the ventral intermediate nucleus (VIM) region may reduce limb tremors but has little effect on other symptoms. Globus pallidus internus (GPi) stimulation may reduce all the significant motor manifestations of PD, including the reduction of dopa-induced dyskinesias and involuntary movements that can limit treatment efficacy. GPi stimulation also may reduce painful cramps and sensory symptoms that may occur when the benefit from individual doses of levodopa abates. However, GPi stimulation does not typically permit medication reduction, and this may be a severe limitation to those having drug-induced side effects such as orthostasis, psychosis, daytime lethargy, or cognitive impairment. In addition, STN DBS provides a similar reduction of motor symptoms, and several studies indicate that bilateral STN DBS improves gait, tremors, and bradykinesia [44].

Moreover, DBS has been studied as a potential treatment for refractory depression and obsessive-compulsive disorder (OCD). While current studies are limited, there is some evidence that DBS may be effective in improving symptoms in these conditions.

One study found that DBS of the subgenual cingulate white matter improved mood in four out of six people with treatment-resistant depression [45]. The targeted region was chosen based on previous research that had demonstrated increased fluorodeoxyglucose uptake in this area in people with depression. Another case report suggested that stimulation of the inferior thalamic peduncle may relieve depressive symptoms [46].

The neurophysiology of DBS has yet to be fully understood, but it is thought to influence various mechanisms involved in neuronal function and signaling. Different elements of the CNS have

different sensitivities to stimulation, which depend on factors such as the amplitude and temporal characteristics of the stimulation, the physiologic properties of individual cells, and the underlying pathophysiology of the disease. The relationship between stimulus amplitude and duration is essential in understanding DBS's effects. The duration must increase as the current amplitude decreases to produce a constant effect, and vice versa. For most neural elements, the form of this amplitude-duration curve is usually an exponential decay [47].

1.5. Brain-Computer Interfaces

BCIs are a rapidly growing research field aiming to provide the CNS with new outputs that are neither neuromuscular nor hormonal. Instead, the CNS produces outputs to respond to inputs from the outside world or the body. BCIs extract and convert CNS activity into artificial outputs.

The concept of BCIs was first introduced by discovering that electrical signals produced by the human brain could be recorded from the scalp [48]. After five years of further study, electroencephalography (EEG) was established as a primary tool for clinical diagnosis and brain research. Additional to EEG, various signals recorded by different levels of invasiveness are used in BCIs, such as MEG, functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (NIRS), firing rates, and local field potential (LFP) [49].

Regardless of the BCI input signals, the output commands control a device or a computer. BCIs can enhance the user's ability to perform tasks, e.g., controlling a powerful robotic arm or a healthy individual typing without moving a muscle. However, the main focus of this thesis and the majority of conventional BCI research studies are to replace or restore natural movement or communication for individuals suffering from trauma to their CNS. Our initial goal was to target different areas of the brain involved in motor processing and provide a unified algorithmic pipeline based on

engineering principles to achieve this. These principles and general decoding algorithms are reviewed in Chapter 2. However, we learned that the optimal algorithm is intricately related to the region of the brain (and thus stage in motor processing). For example, a multiclass discrete rule decoder from multi-electrode lateral prefrontal cortex input benefitted from an RNN-based decoder and dimensionality reduction (Chapter 3). While this same approach could plausibly be used for decoding the level of beta oscillation synchrony in STN, it isn't the most parsimonious or intuitive solution. Rather, a stage of processing that separates oscillatory activity, and parametrizes this on a subject-specific level provided the basis for this decoder (Chapter 4). This is intuitive because it leverages the oscillations that arise from the recurrent network of cortex-basal ganglia-thalamus, and also the oscillations that are known to be abnormal in Parkinson's disease.

BCIs can revolutionize how we interact with the world around us and improve the lives of people with disabilities. However, many challenges still need to be overcome before BCIs can be widely adopted. One challenge is an optimal method to extract meaningful features and translate them to intended movements or outputs. Additionally, a trade-off governs the recording of brain activities: As we get closer to neurons, the intentions modulate recorded information more precisely, but we limit the recorded brain area if we do not want to damage the brain. On the other hand, as we get further from the neurons, we can record larger areas. However, the recorded information is the vectorial summation of electric fields, which do not follow the intention modulations as closely. Recent advances in recording electrodes aim to address this challenge since no single cortical area is wholly responsible for natural movement [22], [49].

BCIs allow individuals to communicate with external devices through their brain activity by producing output commands for goal selection or process control [49]. For a BCI to select a goal, it must extract and decode a high-level intention. For example, if the BCI user is thirsty, a robotic

arm will pick up a glass of water or drives a motorized wheelchair to the refrigerator. On the other hand, a BCI processing the control signal must continuously extract and decode the intended arm movement and move a robotic arm in real-time to reach the glass of water. A BCI might also follow a synergistic combination of the methods.

Chapter 2 discusses and explores this idea in more detail.

1.6. Machine Learning and Decoding

1.6.1. Support Vector Machine

Support Vector Machines (SVMs) are a type of supervised machine learning algorithm that can be used for classification or regression tasks. The SVM finds the best boundary or decision surface that separates the different classes in the data. This boundary is called the maximum-margin hyperplane., the boundary farthest away from each class's closest data points, known as support vectors.

The SVM algorithm solves the following optimization problem that maximizes the distance between the hyperplane and the closest data points from each class:

$$\min \frac{1}{2} W^T W \; subject\ to\ y_i.(W^T.x_i + b); i \in \{1, \dots, n\} \qquad \text{Equation 1.1}$$

where W is the normal vector of the hyperplane, x_i is a data point, y_i is the class label of x_i, and b is the bias term. The constraint ensures that the hyperplane correctly classifies the data points, and the objective function tries to maximize the margin.

In the case of non-linearly separable data, the SVM algorithm uses the kernel trick technique to transform the data into a higher-dimensional space where a linear boundary can separate the

classes. The kernels used in SVMs include the polynomial and the radial basis function (RBF) kernel.

SVMs are particularly useful in cases where the number of features is greater than the number of samples, such as in image recognition or natural language processing. Additionally, SVMs have a built-in regularization parameter, which helps to avoid overfitting.

The SVMs are binary classifiers. However, using methods such as one vs. one or one vs. all they were used for multi-class decoding. Note that this adds computational load if the number of classes are high.

1.6.2. Logistic Regression

Logistic regression is a type of supervised machine learning algorithm that is used for classification tasks. It is a probabilistic model used to parameterize the relationship between a binary dependent variable and one or more independent variables by fitting a logistic curve to the data. The logistic curve is a sigmoid function that maps the input features to a probability between 0 and 1, which can then be thresholded to make binary predictions.

$$P(y = 1|x) = 1/(1 + e^{(-W^T.x-b)}) \quad \text{Equation 2.2}$$

where x is the input feature vector, W is the weight vector, and b is the bias term. The goal of logistic regression is to learn the values of W and b that maximize the likelihood of the observed data.

Regularized logistic regression adds a regularization term to the objective function to prevent overfitting. Different regularization techniques can be used in logistic regression, but the most common are L1 (also known as Lasso) and L2 (also known as Ridge).

L1 regularization adds a term to the objective function that is proportional to the absolute value of the weight coefficients, and it tends to shrink the weights of less important features to zero.

L2 regularization adds a term to the objective function that is proportional to the square of the weight coefficients. It tends to shrink the weights of all features toward zero but less aggressively than L1 regularization:

$$P(y = 1|x) = \frac{1}{1 + e^{(-W^T.x - b)}} + \lambda \|W\| \qquad \text{Equation 3.3}$$

where λ is the regularization parameter and controls the strength of the regularization. For example, a larger value of λ will result in smaller weight coefficients and a simpler model, while a smaller value of λ will allow for larger weight coefficients and a more complex model.

1.6.3. Random Forest

Random Forest is an ensemble machine learning algorithm for classification and regression tasks. It combines multiple decision trees, where each tree is built using a different random subset of the data and features. The final output of a Random Forest is the average or majority vote of the individual trees.

The decision trees used in a Random Forest are built using bootstrap aggregating, or bagging, where a random sample of the data is replaced to create multiple training sets. The decision trees are then built independently on each training set, and their outputs are combined to form the final output of the Random Forest.

In the case of classification, each tree in the forest casts a vote for the most popular class among the data points in the terminal nodes. The final decision of the random forest is then defined as the

mode of the class votes. Assuming the number of trees in the forest is T, and the number of classes is C, then the final decision can be represented as:

$$y = \underset{1 \le c \le C}{\operatorname{argmax}} \sum_{t=1}^{T} y_t = c \qquad \text{Equation 4.4}$$

where y_t is the class label predicted by the *t*-th tree.

In regression, the final output of the Random Forest is the average of the predictions of all the decision trees:

$$y = \frac{1}{T} \sum_{t=1}^{T} y_t \qquad \text{Equation 5.5}$$

where y_t is the prediction made by the *t*-th tree.

One of the main advantages of Random Forest is that it reduces overfitting by averaging the outputs of multiple decision trees. Additionally, Random Forest can handle large numbers of features and interactions between them, which makes it well-suited for high-dimensional data.

Another advantage of the Random Forest algorithm is that it measures each feature's importance, which can be used for feature selection. The feature importance measure is based on the average impurity decrease overall in forest trees. The impurity can be calculated using Gini impurity or information gain. The feature importance can be represented as:

$$f_i = \sum_{t=1}^{T} \Delta_i \qquad \text{Equation 6.6}$$

where f_i is the importance of feature *i*, and Δ_i is the decrease in impurity due to feature *i*.

1.6.4. Convolutional Neural Network

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm well-suited for classification tasks. They are designed to learn and recognize patterns in data, such as the trends in neuronal firing rates.

The main building block of a CNN is the convolutional layer, which applies filters to the input data. The filters are typically small, typically 3x3 or 5x5, and are designed to detect specific features in the input data. The convolution operation is performed by sliding the filter over the input data and computing the dot product between the filter and the input data at each location. The output of the convolutional layer is a set of feature maps, one for each filter:

$$y(i,j,k) = \sum x(p,q) \times W(i-p, j-q, k) \qquad \text{Equation 7.7}$$

where $y(i,j,k)$ is the output feature map for the k-th filter, $x(p,q)$ is the input data, and $W(i-p, j-q, k)$ is the weight of the k-th filter at position $(i-p, j-q, k)$.

The feature maps are then passed through a non-linear activation function, such as *ReLU*, to introduce non-linearity into the model:

$$f(x) = \max(0, x) \qquad \text{Equation 8.8}$$

where $f(x)$ is the output of the activation function, and x is the input.

After the convolutional layer, a pooling layer is typically applied to reduce the spatial dimensions of the feature maps while retaining the important features. The most common types of pooling are max and average pooling. Max pooling selects the maximum value from each window, while average pooling computes the average value:

$$y(i,j,k) = \max\left(x(p,q)\right)$$

Equation 9.9

where $y(i,j,k)$ is the output feature map for the k-th filter after pooling, and $x(p,q)$ is the input feature map in the pooling window.

1.6.5. Recurrent Neural Network

Recurrent Neural Networks (RNNs) are a type of deep learning algorithm that is particularly well-suited for tasks that involve sequential data, such as neuronal activities recorded through time. They are designed to learn and recognize patterns in data present in sequences, such as time series or text.

One of the most popular types of RNNs is the Long Short-Term Memory (LSTM) network. LSTMs are designed to overcome the limitations of traditional RNNs by introducing a "memory cell" mechanism that can retain information for an extended period allowing LSTMs to effectively process long sequences and maintain a "memory" of the data that has been seen so far.

An LSTM network comprises a series of LSTM cells, each of which takes the current input and the previous hidden state as input. The LSTM cell has three gates: the forget gate, the input gate, and the output gate. The forget gate controls the information that is forgotten from the previous cell state, and the input gate controls the information added to the cell state. The output gate controls the information passed to the next hidden state:

Equation 10.10

$$f_t = sigmoid\left(W_f \times [h_{t-1}, x_t] + b_f\right)$$

Equation 11.11

$$I_t = sigmoid(W_i \times [h_{t-1}, x_t] + b_i)$$

Equation 12.12

$$O_t = sigmoid(W_o \times [h_{t-1}, x_t] + b_o)$$

where f_t, I_t, O_t are the forget, input, and output gates, W_f, W_i, W_o are the weight matrices, h_{t-1} is the previous hidden state, x_t is the current input, and b_f, b_i, b_o are the biases. The sigmoid function ensures that the gates are between 0 and 1.

The cell state is then updated by combining the forget gate, the input gate, and a new candidate cell state:

$$C_t = f_t \times c_{t-1} + I_t \times \tanh(W_C \times [h_{t-1}, x_t] + b_c) \quad \text{Equation 13.13}$$

where C_t is the current cell state, c_{t-1} is the previous cell state, tanh is the hyperbolic tangent function, W_C is the weight matrix, h_{t-1} is the previous hidden state, x_t is the current input, and b_c is the bias.

Finally, the current hidden state, h_t, is computed by combining the output gate and the cell state:

$$h_t = O_t \times \tanh(C_t) \quad \text{Equation 14.14}$$

The output of the LSTM network can be obtained by applying a fully connected layer with a *softmax* activation function to the final hidden state:

$$y_t = softmax\ (W_y \times h_t + b_y) \quad \text{Equation 15.15}$$

where y_t is the output, W_y is the weight matrix, and b_y is the bias.

After several convolutional, pooling, and RNN layers, the feature maps are flattened and passed through a fully connected layer known as a dense layer to make the final prediction. The dense layer comprises a set of neurons, each connected to all the neurons in the previous layer. The output

of the dense layer is a set of probabilities for each class in the classification task or a set of continuous values for a regression task:

$$y = W.x + b$$

Equation 16.16

where y is the output of the dense layer, W is the weight matrix, x is the input, and b is the bias.

To minimize the loss function, CNNs can be trained using various optimization algorithms, such as stochastic gradient descent (SGD), Adam, and RMSprop. Standard loss functions include cross-entropy for classification and mean squared error for regression.

The Gated Recurrent Unit (GRU) [50] is another type of RNN architecture that helps address the vanishing gradient problem and improves the ability of the network to capture long-term dependencies in sequential data.

GRU is similar to the more well-known LSTM architecture but has a simplified structure with fewer gates. GRUs have two main components: a hidden state and an update gate.

The hidden state in a GRU is similar to the cell state in an LSTM. It represents the memory of the network and captures the relevant information from previous time steps. The hidden state is updated at each time step based on the input and the previous hidden state.

The update gate controls the flow of information in the GRU. It determines how much of the previous hidden state should be passed to the current time step and how much of the new input should be integrated into the hidden state. The update gate considers the current input and the previous hidden state, and its output is a value between 0 and 1 that determines the trade-off between old and new information.

GRUs also have a reset gate, which determines how much of the previous hidden state should be ignored. The reset gate is used to discard irrelevant information from the previous hidden state.

Compared to LSTMs, GRUs have a simpler structure and require fewer parameters. This makes them computationally more efficient and easier to train in some cases. However, LSTMs tend to perform better in tasks that require modeling long-term dependencies.

1.7. Conclusion

The brain comprises functionally distinct regions, while none of the body's functions is handled in only one region. Understanding the aspects concerning these regions will help us better comprehend disorders and explore new ways to tackle them.

This chapter explored the background information essential to find brain regions suitable for our research. Additionally, the disorders benefiting from this research are reviewed. Moreover, the machine learning tools and methods used in our research are studied.

1.8. References

[1] E. R. Kandel, *Principles of neural science*. New York; Toronto: McGraw-Hill Medical, 2013. Accessed: Nov. 23, 2018. [Online]. Available: http://www.myilibrary.com?id=396874

[2] T. Dickinson *et al.*, *ON SCIENCE 10*, 1st edition. McGraw Hill School Indige, 2009.

[3] P. A. Young, P. H. Young, and D. L. Tolbert, *Basic Clinical Neuroscience*, 3rd edition. Philadelphia: Wolters Kluwer, 2015.

[4] D. Purves *et al.*, *Neuroscience*, 6th edition. New York: Oxford University Press, 2017.

[5] J. Ward, *The Student's Guide to Cognitive Neuroscience*, 4th edition. London : New York: Routledge, 2019.

[6] D. T. Lisa Dufraimont Don Stuart, *Evidence: Principles And Problems 13th Edition*. Carswell, 2021.

[7] L. Squire, D. Berg, F. E. Bloom, S. du Lac, A. Ghosh, and N. C. Spitzer, Eds., *Fundamental Neuroscience*, 4th edition. Amsterdam ; Boston: Academic Press, 2012.

[8] M. Bear, B. Connors, and M. Paradiso, *Neuroscience: Exploring the Brain*, Fourth, International Edition edition. Philadelphia: Lippincott Williams and Wilkins, 2015.

[9] S. M. Kosslyn, G. Ganis, and W. L. Thompson, "Neural foundations of imagery," *Nat Rev Neurosci*, vol. 2, no. 9, Art. no. 9, Sep. 2001, doi: 10.1038/35090055.

[10] C. S. Green and D. Bavelier, "Exercising your brain: A review of human brain plasticity and training-induced learning," *Psychology and Aging*, vol. 23, pp. 692–701, 2008, doi: 10.1037/a0014345.

[11] A. M. Owen, J. J. Downes, B. J. Sahakian, C. E. Polkey, and T. W. Robbins, "Planning and spatial working memory following frontal lobe lesions in man," *Neuropsychologia*, vol. 28, no. 10, pp. 1021–1034, 1990, doi: 10.1016/0028-3932(90)90137-d.

[12] M. E. Goldberg, "Parietal Lobe," in *International Encyclopedia of the Social & Behavioral Sciences*, N. J. Smelser and P. B. Baltes, Eds. Oxford: Pergamon, 2001, pp. 11051–11054. doi: 10.1016/B0-08-043076-7/03471-9.

[13] P. Broca, "Sur le siège de la faculté du langage articulé," *Bulletins et Mémoires de la Société d'Anthropologie de Paris*, vol. 6, no. 1, pp. 377–393, 1865, doi: 10.3406/bmsap.1865.9495.

[14] C. Wernicke, "THE APHASIC SYMPTOM-COMPLEX: A Psychological Study on an Anatomical Basis*," *Archives of Neurology*, vol. 22, no. 3, pp. 280–282, Mar. 1970, doi: 10.1001/archneur.1970.00480210090013.

[15] C. Wernicke, "Der aphasische symptomencomplex...," 1874, Accessed: Jan. 20, 2023. [Online]. Available: http://cesimadigital.pucsp.br/handle/bcd/2246

[16] K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig : Barth, 1909. Accessed: Feb. 26, 2021. [Online]. Available: http://archive.org/details/b28062449

[17] H. G. J. M. Kuypers, "The Descending Pathways to the Spinal Cord, their Anatomy and Function," in *Progress in Brain Research*, vol. 11, J. C. Eccles and J. P. Schadé, Eds. Elsevier, 1964, pp. 178–202. doi: 10.1016/S0079-6123(08)64048-0.

[18] C. A. M. Wheeler-Kingshott *et al.*, "Investigating Cervical Spinal Cord Structure Using Axial Diffusion Tensor Imaging," *NeuroImage*, vol. 16, no. 1, pp. 93–102, May 2002, doi: 10.1006/nimg.2001.1022.

[19] D. Ferrier, *The Functions of the Brain*. Smith, Elder, 1886.

[20] U. Hasson, H. C. Nusbaum, and S. L. Small, "Task-dependent organization of brain regions active during rest," *Proceedings of the National Academy of Sciences*, vol. 106, no. 26, pp. 10841–10846, Jun. 2009, doi: 10.1073/pnas.0903253106.

[21] P. Molenberghs, R. Cunnington, and J. B. Mattingley, "Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies," *Neuroscience & Biobehavioral Reviews*, vol. 36, no. 1, pp. 341–349, Jan. 2012, doi: 10.1016/j.neubiorev.2011.07.004.

[22] S. Genon, A. Reid, R. Langner, K. Amunts, and S. B. Eickhoff, "How to Characterize the Function of a Brain Region," *Trends in Cognitive Sciences*, vol. 22, no. 4, pp. 350–364, Apr. 2018, doi: 10.1016/j.tics.2018.01.010.

[23] L. Pessoa, "Understanding brain networks and brain organization," *Physics of Life Reviews*, vol. 11, no. 3, pp. 400–435, Sep. 2014, doi: 10.1016/j.plrev.2014.03.005.

[24] C. R. G. Jones and M. Jahanshahi, "The Substantia Nigra, the Basal Ganglia, Dopamine and Temporal Processing," in *Birth, Life and Death of Dopaminergic Neurons in the Substantia Nigra*, G. Giovanni, V. Di Matteo, and E. Esposito, Eds. Vienna: Springer, 2009, pp. 161–171. doi: 10.1007/978-3-211-92660-4_13.

[25] R. M. Norman, "CEREBELLAR ATROPHY ASSOCIATED WITH ÉTAT MARBRÉ OF THE BASAL GANGLIA," *J Neurol Psychiatry*, vol. 3, no. 4, pp. 311–318, Oct. 1940.

[26] A. M. Graybiel, "The basal ganglia," *Current Biology*, vol. 10, no. 14, pp. R509–R511, Jul. 2000, doi: 10.1016/S0960-9822(00)00593-5.

[27] P. L. Strick, R. P. Dum, and J. A. Fiez, "Cerebellum and Nonmotor Function," *Annual Review of Neuroscience*, vol. 32, no. 1, pp. 413–434, 2009, doi: 10.1146/annurev.neuro.31.060407.125606.

[28] J. D. Schmahmann, J. B. Weilburg, and J. C. Sherman, "The neuropsychiatry of the cerebellum — insights from the clinic," *Cerebellum*, vol. 6, no. 3, pp. 254–267, Sep. 2007, doi: 10.1080/14734220701490995.

[29] D. M. Wolpert, R. C. Miall, and M. Kawato, "Internal models in the cerebellum," *Trends in Cognitive Sciences*, vol. 2, no. 9, pp. 338–347, Sep. 1998, doi: 10.1016/S1364-6613(98)01221-2.

[30] H. Gray, *Anatomy of the Human Body*. Lea & Febiger, 1918.

[31] W. E. L. G. Clark, "Functional Localization in the Thalamus and Hypothalamus.∗," *Journal of Mental Science*, vol. 82, no. 337, pp. 99–118, 1936, doi: 10.1192/bjp.82.337.99.

[32] M.-T. Herrero, C. Barcia, and J. Navarro, "Functional anatomy of thalamus and basal ganglia," *Childs Nerv Syst*, vol. 18, no. 8, pp. 386–404, Aug. 2002, doi: 10.1007/s00381-002-0604-1.

[33] T. Sun and R. F. Hevner, "Growth and folding of the mammalian cerebral cortex: from molecules to malformations," *Nat Rev Neurosci*, vol. 15, no. 4, pp. 217–232, Apr. 2014, doi: 10.1038/nrn3707.

[34] D. Purves *et al.*, "An Overview of Cortical Structure," *Neuroscience. 2nd edition*, 2001, [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK10870/

[35] W. Penfield and E. Boldrey, "Somatic Motor And Sensory Representation In The Cerebral Cortex Of Man As Studied By Electrical Stimulation," *Brain*, vol. 60, no. 4, pp. 389–443, Dec. 1937, doi: 10.1093/brain/60.4.389.

[36] A. Mogilner *et al.*, "Somatosensory cortical plasticity in adult humans revealed by magnetoencephalography.," *Proc Natl Acad Sci U S A*, vol. 90, no. 8, pp. 3593–3597, Apr. 1993.

[37] D. E. Feldman and M. Brecht, "Map plasticity in somatosensory cortex," *Science*, vol. 310, no. 5749, pp. 810–815, Nov. 2005, doi: 10.1126/science.1115807.

[38] H.-C. Cheng, C. M. Ulane, and R. E. Burke, "Clinical Progression in Parkinson's Disease and the Neurobiology of Axons," *Ann Neurol*, vol. 67, no. 6, pp. 715–725, Jun. 2010, doi: 10.1002/ana.21995.

[39] C. A. Davie, "A review of Parkinson's disease," *British Medical Bulletin*, vol. 86, no. 1, pp. 109–127, Jun. 2008, doi: 10.1093/bmb/ldn013.

[40] Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, "The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations," *Movement Disorders*, vol. 18, no. 7, pp. 738–750, 2003, doi: 10.1002/mds.10473.

[41] C. G. Goetz *et al.*, "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disorders*, vol. 23, no. 15, pp. 2129–2170, 2008, doi: 10.1002/mds.22340.

[42] G. Fritsch and E. Hitzig, "Uber die elektrische Erregbarkeit des Grosshirns," *Arch Anat Physiol Wiss Med, pp. 300–332; 1960. Reprinted in: G. von Bonin (transl). Some Papers on the Cerebral Cortex, pp. 73–96. Spring3eld, IL: Thomas.*, vol. 37, pp. 300–332, 1870.

[43] J. S. Perlmutter and J. W. Mink, "Deep Brain Stimulation," *Annual Review of Neuroscience*, vol. 29, no. 1, pp. 229–257, 2006, doi: 10.1146/annurev.neuro.29.051605.112824.

[44] A. Horn, W.-J. Neumann, K. Degen, G.-H. Schneider, and A. A. Kühn, "Toward an electrophysiological 'sweet spot' for deep brain stimulation in the subthalamic nucleus," *Hum Brain Mapp*, vol. 38, no. 7, pp. 3377–3390, Jul. 2017, doi: 10.1002/hbm.23594.

[45] H. S. Mayberg *et al.*, "Deep brain stimulation for treatment-resistant depression," *Neuron*, vol. 45, no. 5, pp. 651–660, Mar. 2005, doi: 10.1016/j.neuron.2005.02.014.

[46] F. Jiménez *et al.*, "A patient with a resistant major depression disorder treated with deep brain stimulation in the inferior thalamic peduncle," *Neurosurgery*, vol. 57, no. 3, pp. 585–593; discussion 585-593, Sep. 2005, doi: 10.1227/01.neu.0000170434.44335.19.

[47] R. Ramasubbu, S. Lang, and Z. H. T. Kiss, "Dosing of Electrical Parameters in Deep Brain Stimulation (DBS) for Intractable Depression: A Review of Clinical Studies," *Front Psychiatry*, vol. 9, p. 302, Jul. 2018, doi: 10.3389/fpsyt.2018.00302.

[48] H. Berger, "Über das Elektrenkephalogramm des Menschen," *Archiv f. Psychiatrie*, vol. 87, no. 1, pp. 527–570, Dec. 1929, doi: 10.1007/BF01797193.

[49] J. Wolpaw and E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, USA, 2012.

[50] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "*Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*." arXiv, Dec. 11, 2014. doi: 10.48550/arXiv.1412.3555.

Chapter 2:
Intracortical Brain-Computer Interfaces in Primates: A Review and Outlook

2.1. Abstract

Brain-computer interfaces (BCI) translate brain signals into artificial output to restore or replace natural central nervous system (CNS) functions. Multiple processes, including sensorimotor integration, decision-making, motor planning, execution, and updating, are involved in any movement. For example, a BCI may be better able to restore naturalistic motor behaviors if it uses signals from multiple brain areas and decodes natural behaviors' cognitive and motor aspects. This review provides an overview of the preliminary information necessary to plan a BCI project focusing on intracortical implants in primates. Since the brain structure and areas of non-human primates (NHP) are similar to humans, exploring the result of NHP studies will eventually benefit human BCI studies. The different types of BCI systems based on the target cortical area, types of signals, and decoding methods will be discussed. In addition, various successful state-of-the-art cases will be reviewed in more detail, focusing on the general algorithm followed in the real-time system. Finally, an outlook for improving the current BCI research studies will be debated.

2.2. Brain-Computer Interface

2.2.1. Intracortical Brain-Computer Interface

Intracortical BCI is an invasive branch of BCI that records brain activity within the brain tissue, usually from the cerebral cortex [1]. In intracortical BCIs, recording electrodes are closer to the neurons. Therefore, we can study the neural activity of a single neuron or population of neurons in the desired brain area. The target area of the brain can diversely change the mechanism

of a BCI research study and is usually decided upon in the early stages of preparation. In addition, various cortical areas are involved in movements, which are critical in decision-to-motor transformations [2–5]. Many BCIs aim to replace motor function lost due to CNS trauma or disease by decoding movement intentions from the motor cortex and translating them to movements of effectors like robotic arms or computer cursors. However, naturalistic motor behaviors are only encoded partially in the motor cortex. Thus, BCIs may benefit from various brain areas activated during a specific task movement. As a result, one (or multiple) brain area(s) will be the target of BCI studies. Subsequently, the study enters imaging and mapping stages to precisely target the best sub-area in the brain. Combining EEG recordings, CT scans, MRI imaging, fMRI recording, skull, and cortex 3D modeling will localize the most active sub-area of the targeted cortex for recording array implantation.

2.2.2. Brain Areas

Decision-making is choosing a course of action toward a desired outcome in the future. However, various events in the world cannot be predicted entirely and may alter the outcome. Cortical and subcortical areas of the brain process the limited information provided to make the best decision with the most acceptable consequences. Although various brain areas are involved in this process, the motor and pre-motor areas of the brain have been the primary targets of BCI studies.

Motor Areas

Motor information in the brain is encoded by various populations of neurons. These populations in different brain areas work in series and parallel during a movement. This hierarchical and parallel processing pattern is reproduced in all motor systems inside the brain.

The motor areas are constructed of multiple functional sub-areas forming hierarchical systems in the brain. The primary area (M1) includes an ensemble of neurons, and the population activity is mapped to different peripheral movements. The secondary area is the premotor cortex, concerned chiefly with planning and the global aspects of motor behavior, such as the muscles involved in a movement or the sequence in which they should contract. The tertiary area is located in the prefrontal cortex, concerned with executive behavior control.

The motor behavior of the brain during voluntary movements comprises two steps: plan and execute. Studies show that neurons in the parietal, premotor, prefrontal, and motor regions are involved in these steps [7–14]. Motor planning starts before motor execution and could continue during the movement. For example, planning is the neural process that first locates the desired object in the space in a reach-and-grasp movement. Subsequently, the object's physical properties, such as size and shape, are transformed into a grip. Lastly, the planning information translates into moving towards the object until the object is grasped.

Recordings of hand and arm kinematics show that even the execution of reach and grasp are not in sequence but essentially simultaneous [15–19]. The hand shapes the grip while the arm makes the reaching movement. Different body parts involved in a single motor task may execute the move simultaneously. Furthermore, they do not use a single unified mapping of the world. Instead, the brain encodes the space we experience using multiple maps with reference frames related to different motor circuits controlling the eyes, arms, hands, and other effectors. These maps have changed and adapted to the motor effector's specific needs [20–24].

The motor cortex neurons modulate activity related to kinematic and kinetic parameters. Task-related parameters such as instruction cues and movement observation also modulate the ensemble activities in the motor cortex [25–32]. Recordings in the motor area show that adjacent neurons

may have highly different activities, relating these neurons to completely different parameters. As a result, it is possible to encounter various signals related to voluntary movement recorded from a small, implanted electrode array [33].

The motor cortex has a critical role in controlling voluntary movement; it contains ensembles of neurons that convert a plan of action into motor commands to execute the plan. This execution will not be achieved by static mapping of muscles and moving body parts. Instead, the internal organization of the primary motor cortex and spinal connections will dynamically convert the current state's motor intentions and sensory feedback into motor output commands [34–36].

Decision-Making

The prefrontal cortex (PFC) is involved in various stages of the perception-to-action cycle. Many types of information are encoded in PFC, yet unlike the motor or sensory cortex, no precise mapping exists except for the eye movement direction map [38, 39]. Studies show that sensory inputs, planning and executing various actions, memory-related tasks, and attention activate the neurons in PFC [40–45]. The PFC uses this variety of information to have a 'higher order' control over behaviors. For example, patients with damage to PFC may look ordinary but cannot perform daily activities.

Every intended action depends on the ability to remember the intention, in some cases complicated like a chosen lifestyle. These intentions and the ability to remember them until we carry them out originate in PFC [46–49]. An intention-directed behavior is associated with a PFC subarea called orbital-ventromedial. This area connects to the subcortical systems, the hypothalamus, and the amygdala. These areas are involved when we feel hungry, thirsty, or more complicated feelings such as fear, aggression, and arousal. As a result, the orbital-ventromedial

PFC has access to object features like shape, color, taste, and textures. The orbital-ventromedial PFC projects the appropriate emotional response to the dorsolateral PFC and then the premotor cortex to trigger appropriate action. For example, the intention of drinking could be triggered when a glass of cold water with desired visual features is nearby, and we feel thirsty [44, 47, 50–52].

PFC is not exclusively concerned with movement planning. Much of the research in PFC includes memory and, more specifically, working memory (WM) [41, 53]. Nevertheless, executive motor control and WM might be related since they depend on holding information over time. WM is the ability to keep the information for a short time, usually on the scale of seconds, and manipulate it mentally. A good example is holding two numbers in mind and trying to multiply them mentally. WM also retrieves and uses information from long-term memory, making WM an input/output pathway for the long-term storage sites from a modeling point of view [54–60].

Despite recent studies, we still need to understand the population-scale organization of WM. The neural systems that drive and support WM are broadly distributed through the brain, specifically PFC. However, whether the systems supporting the WM are distributed or form multiple specialized networks is still being studied. Understanding the PFC becomes crucial in research studies investigating and treating a cognition disorder. Also, how to approach PFC neural activity recorded by WM tasks is dependent on understanding how these networks are organized [61]. By large-scale recording across lateral PFC of macaque monkeys, WM activity has been categorized into three anatomically specific modes. The first two modes encode early and late forms of memory storage, and the third mode predicts behavioral variability after delay, consistent with the response preparation [61].

2.3. Signal Recording

2.3.1. Electrodes

Various types of electrodes can be used to monitor and record brain activity. In addition, the research study might utilize electrodes to stimulate the neurons with electrical waves sent to a particular brain area. Intracortical recording of brain activities includes electrodes entering the brain's gray matter. These sensors could be individual electrodes penetrating the brain or an array of electrodes implanted on the surface of the cortex. Thus, intracortical sensors are classified into three main groups: Microwires, Microwire bundles, and Microelectrode arrays [62].

Microwires are the simplest form of these sensors. Initially, a glass electrode filled with electrolytes to monitor intracellular activities of neurons [63] evolved into conductive metal microwires which, except for the tip in contact with tissue, are insulated [64]. However, these electrodes are unsuitable for long time recording, and the chronic reactions of brain tissues should be considered in their long-term usability [65]. Therefore, a different design of wire recording was proposed and developed, which could record single and multi-unit activity [66]. This electrode has cone-shaped glass insulating gold wire(s) inside. The long-term stability of recording is the main improvement over standard wire recording techniques.

Neural activities of single neurons are usually separated by assuming that the waveform shape of one neuron will only differ by amplitude through time, and different neurons' activities have different waveform shapes [67–70]. These assumptions may differentiate one neuron activity into two or more units recorded by the same microwire [62]. Electrode bundles of twisted microwires have been developed to overcome these uncertainties [62, 71]. These microwire bundles consist of platinum-iridium alloy microwires insulated by Teflon.

Additionally, multielectrode arrays contain multiple electrode shanks on a plate base. The first array developed in 1967 had five tungsten microwires [72]. In 1992, the Utah array, a glass/silicon composite array of needle-like electrodes, was developed, which recorded from 100 electrodes simultaneously [73]. Later in 1994, a three-dimensional silicon-based microelectrode array called the Michigan probe was developed [74], which had multiple recording sites along the length of each shank and various electrodes on a base. Although ceramic-based, polymer-based, and multielectrode arrays have been developed, the Utah array remains the most popular in human studies [75-76, 62].

2.3.2. Signals

Three general types of neuronal signals can be obtained from the intracortical recording: single-unit activity (SUA), multi-unit activity (MUA), and local field potential (LFP) [77]. LFPs are neurons' low-frequency (<300 Hz) analog activity close to the electrode. Any electrical current in the brain leads to an extracellular field that fluctuates based on the contributing currents. The amplitude and frequency of LFPs will change based on the activity of neurons close to the electrode. A higher amplitude LFP is associated with the inverse of the distance to a source neuron related to the task being executed [78].

Moreover, the beta-band (13-30 Hz) and the gamma-band (30-100 Hz) peaks in LFPs have valuable information, which has led to much research specific to these brain signals. For example, beta activities have been associated with concentration and problem-solving and are a critical source of studies on Parkinson's disease. In addition, gamma activities have been associated with perception, attention, working memory, and disorders like schizophrenia and autism [79].

The SUA is the high-frequency activity of a single neuron. By detecting the spikes in this high-frequency (>300 Hz) signal and assigning each spike to its corresponding neuron in a complicated process called spike sorting, we derive a signal which shows the spiking of a single neuron through time (i.e., spike trains). Furthermore, the number of spikes in specific time bins measures the firing rates of a neuron through time. Spike trains and firing rates are the variable features containing motor or cognitive activity information. The MUA is derived similarly to the SUA, with the difference that the resulting spikes are not sorted and are the simultaneous activity of multiple neurons [80].

2.3.3. Single Neuron vs. Population Activities

Single-Neuron Representation

The various methods of studying the brain usually fall into one of the categories of single-neuron or population representation of activities. Regardless of how we record neural activities, we could always investigate single-neuron involvements in movement and learning.

Prefrontal Cortex

Lesion studies show that PFC involves learning a principle or rule from experience and applying it to new and similar situations [81]. The neural basis of this fact is studied by recording single neurons in the PFC of monkeys [45]. They indicated whether two consecutive shown objects were a match or a non-match, releasing a lever based on the changing rule. Some neurons in PFC showed more significant activity during match trials, and their firing rate changed greatly based on the rules. This preliminary study concluded that single neurons in the prefrontal cortex encode abstract rules. However, it has been found that abstract rules are reflected more strongly and earlier in the premotor cortex than PFC in well-practiced monkeys [82]. Understanding the rules is an

integral part of decision-making. Another research in a similar task with the binary decision of holding or releasing a lever has studied the single neuron responses of 708 randomly selected neurons from the PFC of monkeys [83]. The task comprises five phases: fixation, stimulus, delay, rule, and response. The monkeys are prompted to hold or release a lever based on the rule in each trial. The study found that 8% of neurons in the stimulus phase of the task and 18% during the delay phase significantly code the subjective decisions of monkeys. This research shows the possibility of decoding simple decisions from single neurons in the PFC. However, another study states that the single neuron activities in PFC during WM-related tasks are heterogeneous and intensely dynamic, making them unstable for working memory representations [84].

Motor Cortex

Neurons in the motor cortex generate different activations for various movements based on different features, including the direction of the movement or velocity. By looking at each recorded neuron during a particular movement, we might find correlations between how a single neuron fires and the direction of the movement. One of the essential features in single neuron representations is the tuned activity of the neuron to a specific movement direction [85, 86].

In a study on monkeys' motor cortex, the activities of 606 cells related to proximal arm movement were studied while making two-dimensional arm movements in eight directions. As a result, the frequency and intensity of discharge in 75% of 323 active neurons were related to a preferred direction [27]. The shape of the resulting tuning curve has been a subject of debate. One study finds that the frequency of discharge in 75% of the 241 directionally tuned cells is a sinusoidal function (usually cosine) of the direction of movement [27]. However, another study shows evidence of various exhibited tuning profiles from cell to cell [87].

Regardless of the shape of tuning curves, the directional preference of motor cortex neurons has been a valuable source of information for decoding motor cortex activity. Although there are limitations to single-neuron decoding of movement and information representation [88], averaging the activity through time has produced a more accurate decoding output in several studies [89–92].

Population/Ensemble representations

Prefrontal Cortex

The PFC plays an essential role in WM, and studies suggest a relatively stable population-level representation of remembered stimulus features in the PFC [84]. In addition, another study has found that ensemble activities in rats' prefrontal cortex changed during new task learning, allowing them to decode previous and future goal choices [38]. Moreover, another study on monkeys shows the dynamic coding of stimulus flashing frequency in the population activities of PFC neurons [93]. In this study, monkeys indicated whether the flashing frequency of one stimulus was higher or lower than the previous stimulus. At the same time, the population activity of 899 neurons was recorded during this working memory. Other studies found that the LFP gamma (30-100 Hz) oscillation dynamics and the spike rates from the PFC correlate with the sensory information held in working memory. Moreover, the gamma power correlated with the number of objects stored in the WM [94-99].

Motor Cortex

Studies suggest decoding motor cortex activities is more precise when combined actions of a large population of widely tuned neurons are studied [100-101]. For example, single-neuron studies show the directional tuning among neurons in three-dimensional arm movements [32].

However, since many cells are active for any movement, generating action in a specific direction depends upon the activity in the neural ensemble as a population coding [102]. Population vector in the motor cortex is considered a vectorial summation of individual neurons' votes to a particular movement where active neurons make a weighted contribution along the axis of their preferred direction [102]. In many brain regions, a large population of neurons' response to stimuli or a motor command leads to a single outcome interpreted as a distributed representation of the information [103].

One study on rhesus monkeys has found the population vector to be a robust predictor of the direction of arm movement in space and relatively immune to cell loss [100]. In that study, the discharge frequency of 224 active neurons from 282 total recorded cells was used to form a population vector that could predict arm movement's direction before the movement onset, where the monkeys reach toward a peripheral target from a central point.

In population recording of neurons, although targeting a fraction of neurons involved in a movement, it is essential to know that only some cells add valuable information to the set. For example, some cells might not be active during a specific task; even if they are, there might be a significant correlation of information between neighbor cells. Thus, dimensionality reduction methods are applied to the recordings before other translations or decoding are performed [104, 105]. If there are D measured variables, but there are K explanatory variables (where $K < D$), dimensionality reduction methods extract these K explanatory (or latent) variables. The latent variables are a shared input or a collective role of unobserved neurons in a network we are recording. They form a K-dimensional space representing the important shared activity of a population response [106]. The lower-dimensional surfaces within the entire space are called neural manifolds. In a manifold, the activity of each neuron is presented as a linear combination

of latent variables plus some additive noise [107]. The neural manifolds, like cortical activity, could change or emerge through time with long-term learning [108, 109]. If a movement requires the acquisition of new neural modes outside of the manifolds, it takes longer for the subject to learn that motor activity [107, 110].

Most motor actions follow the kinematic principles, yet neurons act based on practical solutions shaped by evolution. For example, it is unclear whether the neural activity in the motor cortex relates to muscles or abstract movement features [26, 111–114]. The motor cortex encodes the kinematics and kinetics of each movement. Kinematics refers to parameters that describe the movement, such as position, velocity, acceleration, joint angles, and length of muscles. Kinetics involves causal forces and muscle activity. Findings show that some neurons in the primary motor cortex encode the kinematics of movement and some kinetics. Of course, this relation is not linear, and some neurons are found to discharge firmly during weak activities and be relatively quiet during a strong movement of a finger [115–119].

There are different views regarding movement parameters encoded by neuron populations. Traditional theories state that motor cortex neurons are tuned for parameters such as direction. Still, some neural populations are directionally tuned to velocity and non-directionally sensitive to speed [120]. Neuronal activities governing desired kinematics and required movement kinetics are generated simultaneously in different yet overlapping neural populations. Thus, the role of the motor cortex may be the transformation between what movement to make (kinematics) and how to make it (kinetics) [115, 118, 121].

2.4. Decoding Methods

An essential characteristic of a BCI system is the estimation of intentions. To control a computer or machine with brain signals, we need to translate the neural activities into intention and then into commands; This challenging process is called decoding. Various decoders and decoding methods exist in machine learning, each performing better than others in specific situations. For any research, usually, multiple decoders are tested to find the best decoding performance. The best decoding methods are particular to the research task. For example, decoding a movement intention is continuous (i.e., based on the previous kinematic state), while decoding an intended target in a saccade task is discrete. However, the methods suitable for a given task might require a mixture of regression and classification or an ensemble of multiple methods to vote on the decoded outcome.

BCIs, one popular decoding method to translate motor activities into movement is the Kalman filter (KF) [122]. KF uses previous states of the hand (e.g., kinematic features like position, velocity, and acceleration) plus neural activity (e.g., firing rates of current and previous time bins) and estimates the next state of the hand as follows:

$$x_k = function\ (x_{k-1}, z_k) \qquad \text{Equation 2.17}$$

where x is the state vector, and z is the neural activity (i.e., firing rate). In each step, the KF model is updated so the predictions are subsequently corrected towards the true state whenever a new batch of data arrives. There are two main steps in the KF model: prediction and updating. If we assume there is no external influence (e.g., force) that will alter our states, the prediction step will go as follow:

$$\hat{x}_k = F_k \hat{x}_{k-1}$$
$$P_k = F_k P_{k-1} F_k^T + Q_k$$

Equation 2.18

where F is the state transition obtained from X, P is the estimate covariance matrix, and Q is the covariance of additional uncertainty from the environment called process noise. In this step, a 'best estimate' and a 'new uncertainty' are predicted from the 'previous best estimate' and 'old uncertainty'. Then we will update the model in the next step:

$$\hat{x}'_k = \hat{x}_k + K_k(z_k - H_k \hat{x}_k)$$
$$P'_k = P_k - K_k H_k P_k$$
$$K_k = P_k H_k^T (H_k P_k H_k^T + R_k)^{-1}$$

Equation 2.19

where H is the observation function obtained from Z and X; R is the covariance of the observation noise. Subsequently, the new 'best estimate' $\hat{x}'_k$, and the updated uncertainty E'_k will be fed to the next iteration of prediction and updating.

Multiple versions of KF have been developed that could perform better in nonlinear systems. The extended KF is developed to address nonlinear systems by simply linearizing all nonlinear models to apply the traditional KF [123]. Unscented KF is another extension of KF, which uses a set of appropriately chosen weighted points to address a system's nonlinearity [124]. In this version, the state-to-state and state-to-measurement are not simply two matrices; two functions make these transitions. Recent advances in machine learning sciences, especially deep-learning algorithms, have provided new tools for BCIs. For example, a BCI study on monkeys used long short-term memory (LSTM), an extension of the recurrent neural network (RNN) method, to decode kinematic features during a center-out, bimanual reaching, and bipedal walking

on a treadmill task [126]. They report a significant improvement in performance using this decoding method over state-of-the-art unscented and standard KF.

For classification tasks (e.g., decoding the intended target in a saccade task), various methods can be utilized based on the complexity of the signal and the recorded brain area. If the translation of neural activities to intentions proves linear simple and fast methods might be a good fit (e.g., the linear discriminant analyses or regularized logistic regression). Support vector machines [127] can be a practical method for more complex transitions.

2.5. Successful Brain-Computer Interface

2.5.1. Motor BCI

Many BCI research studies have explored the functionality of the motor cortex of primates. Though far from perfect, some successful BCI cases developed and tested on humans are some of the most sophisticated BCIs developed.

BrainGate

BrainGate was developed to translate intended hand motion into a neural cursor with which the patient could perform activities like opening emails or operating a television [129]. In this study, a male 25-year-old patient with spinal cord injury resulting in complete tetraplegia was implanted with a 100-microelectrode array in the arm area knob of M1. A linear filter translated the patient's intended hand movements into cursor movements. To train the model, since the patient cannot physically move the cursor, he was asked to imagine following the cursor movements made by a physician. The filter was then used to decode activity and drive a neural cursor. For example, during the center-out task, the patient has acquired between 73-95% of targets.

Algorithm

1. A technician moved a 2D cursor on a screen through random targets.
2. The patient was asked to imagine moving his hand like the cursor.
3. The single and multi-unit activity of M1 (arm area) was recorded.
4. A linear filter regressed neural response matrix (containing neuron's firing rates) onto technician-controlled cursor position.
5. The filter was updated with each one-minute block of data.
6. The neural activity of the patient was then translated into cursor position using the following closed-form solution comprised of least-squares formulation:

$$u = R.f = R(R^T R)^{-1} R^T k \qquad \text{Equation 2.20}$$

where u is the reconstructed 2D position, R is the neural response matrix, f is the linear filter, and k is the 2D position.

A neural-controlled robotic arm was developed in one of the following projects based on BrainGate (BrainGate2) [130]. The neural activity of the arm and the hand area in M1 are decoded using a Kalman filter to continuously update an estimate of the participant's intended hand movement and control a robotic arm to perform three-dimensional reach and grasp movements. Similar to BrainGate, to initialize and train the model, participants were asked to imagine controlling the robotic arm as they watched it move. In this study, one participant performed self-paced reaches to successfully drink from a mug of coffee 11/12 times within a single session [131].

Recent research from the same group has combined a recurrent neural network (RNN) decoding model with a language-predictive model to estimate the handwriting intentions of their participant

[132]. In this study, the motor activities of neurons in the hand knob in a patient with spinal cord injury have been associated with attempted handwriting letters to train the models that achieved the speed of 90 characters per minute with 94.1% real-time accuracy.

University of Pittsburgh Medical Center

In this project, the motor cortex signals of a participant with tetraplegia were translated into movement dynamics of a seven-degree-of-freedom robotic arm to reach and grasp a target [133]. A velocity-feature vector (3D translation of the arm, 3D wrist orientation, and grasp) was related to the acquired firing rates in 30ms time bins. Optimal linear estimation was used with ridge regression to find the coefficients of the linear model. Different tasks of reaching and grasping have been tested and reported in this research. The first task is moving the robot arm towards one of the six targets on a board indicated by an LED and performing an orientation of grasp announced by a computer-generated voice. The second task is grabbing, moving, and different objects of different shapes on a target surface. Other tasks were also selected from the action research arm test (ARAT) [146] for the participant to perform under complete brain control of the robotic arm. This project was the first human study to achieve natural control in complex movements of a prosthetic arm.

Algorithm

1. A seven-degree-of-freedom robotic limb was moved through space toward an indicated target on a board.
2. The patient was asked to watch the movement of a robotic limb carefully.
3. Neural data were recorded during this observation phase.
4. The activity of each unit was linearly modeled by the seven velocities described above.

$$f = b_0 + b_x v_x + b_y v_y + b_z v_z + b_{\theta x} v_{\theta x} + b_{\theta y} v_{\theta y} + b_{\theta z} v_{\theta z} + b_g v_g \quad \text{Equation 2.21}$$

f is the firing rate of a unit, b is the coefficient from B, coefficient matrix, and v is velocity from V, velocity matrix.

5. Indirect optimal linear estimation with ridge regression was used to fit the model, and a unit that did not fit the model was excluded.
6. 80 trials (6 minutes) of observation phase data were used to train the model so the participant could control the limb in the next phase.
7. 80 trials of participant-controlling phase data were also used to optimize the final model.
8. Training and optimizing the model were performed daily.

The Ohio State University

Many people who have paralysis have disruption of the signal pathway between the brain and the muscles. This research study implements a new information path from the brain to the muscles [134, 135]. This project uses the intracortical motor activity of a participant with quadriplegia to activate his forearm muscles to eventually give them the ability to grasp, manipulate, and release objects. In this study, the hand area of the motor cortex was implanted with a Utah array, and the acquired signal decoded the participant's movement intention using wavelet decomposition. This intention was then encoded to evoke the desired response from the muscles in the forearm. A graphical hand on a monitor showed a movement or posture to train the decoder, and the participant imagined/attempted the cued movement. For example, in one of the tasks completed by the participant, they were able to grasp a bottle, pour its contents into a jar, release the bottle, pick up a stir stick, and use it to stir the jar's contents. The participant completed this task three out of five times in 10 minutes, each taking less than 52 seconds.

Algorithm

1. The stimulation patterns for desired movements of the thumb and the wrist were first calibrated by trial and error to find stimulation levels corresponding to low, medium, and high amounts of joint deflection.
2. The participant was prompted by a small graphical hand on a monitor to attempt the gesture or movement.
3. Four-scale wavelet decomposition was applied to each 100ms block of data.
4. One-second-wide boxcar filter was applied to the mean of wavelet coefficients for each channel to smooth the data.
5. Four scales were averaged and standardized to be used as a feature (called MWP or mean wavelet power) for a block of data from each channel (96 features for every 100ms)
6. Multiple decoders were used simultaneously, each for a trained movement.
7. The features and the cues were fed to a support vector machine for each motion using a nonlinear Gaussian radial basis function kernel.
8. The decoder with the highest output score (r) was used to determine the stimulation intensity level.

$$f(r) = \begin{cases} 0, & r \leq 0 \\ r(m-l)/0.2 + l, & 0 < r \leq 0.2 \\ r(h-m)/0.2 + 2m - h, & 0.2 < r \leq 0.4 \\ h, & r > 0.4 \end{cases} \qquad \text{Equation 2.22}$$

   m, l, and h are the three stimulation levels corresponding to the medium, low and high amounts of joint deflection, respectively.
9. The decoders were updated with every 100ms block of data.

This research is the first implementation of intracortical motor cortex BCI to stimulate the muscles in real-time to regain functional hand movement.

John Hopkins University

In this research, the participant patient suffering from spinal cord injury controlled two robotic arms with the neuronal activities of the motor and somatosensory cortices of both their brain hemispheres [136]. In addition, the study adopted a shared control strategy to control two semi-autonomous twelve-degree-of-freedom (12 DOF) robotic arms.

Algorithm

1. In a gesture-based 2D control strategy, a mapping familiar to the patient was adapted as follows: up (open hand), down (2-finger pinch), toward the midline (wrist flex), away from the midline (wrist extend), and no movement (hand rest).
2. Neural signals were recorded from six multielectrode arrays from 384 channels, and the normalized firing rates were calculated as features.
3. Based on linear discriminant analyses, the decoding model was trained by the patient imagining different gesture combinations.
4. The shared-control strategy was applied to limit the DOFs controlled by neural signals. In complex tasks, up to 4 DOFs were controlled by the patient (2 DOFs for each brain hemisphere).
5. Since the robots are performing the majority of the control, a moving average filter on the input monitors the active engagement of the user. The robots will halt the autonomous movement controls if this control signal falls below a threshold.

2.5.2. Cognitive BCI

The term "cognitive" broadly defines many aspects of brain activity during thinking and understanding. Higher cognitive functions of the brain are independent of the details of how the actions are executed [137] and are more concerned about the final achievement and evaluating the course of activities. For example, imagine a bowl of fruits; If we prefer one fruit over another, signals in our brain indicate this preference which eventually could decide what fruit we will reach to pick [138]. Using a cognitive-related signal from the brain to control a BCI is less explored than execution-related BCIs. The idea is to decide whether we could decode the higher-level intention of a user and translate it to a computer or machine output. Cognitive neural prosthetics are BCI systems that use cognitive signals to control an external device [148].

The difference between a motor and a cognitive BCI is the approach to understanding the user's intention. For example, in a motor BCI, the user thinks about moving their arm to open a window, and the machine will translate the moving intentions until it reaches the window and opens it. However, in a cognitive BCI, the user intuitively imagines an open window, and the machine will open the window [137]. Therefore, extracting the user's cognitive state can produce a more meaningful output. For example, suppose a mute person wants to use the BCI for communication. In that case, the cognitive BCI can extract the appropriate signals directly from the speech-related area of the brain rather than using a letter board and a moving cursor based on motor signals [138]. To this end, we need to know which part of the brain generates the appropriate signals. Here we also examine the implementation of a cognitive BCI and how the intentions are decoded.

Compared to motor BCIs, the number of developed cognitive BCIs is very limited. This number is much more limited when the subject is human. For example, one research study in the

posterior parietal cortex (PPC) decoded motor imagery of a tetraplegic human [139]. They indicate that motor imagery recorded from a population of neurons in PPC can be translated into cursor movements or robotic limb control. In this study, a 32-year-old patient with paralyzed limbs due to spinal cord injury was implanted with two microelectrodes in the reaching-related area on the superior parietal lobule and a grasping-related area at the junction of the intraparietal and postcentral sulci. The recorded spiking activity was used to control a 17-degree-of-freedom robotic arm and a 3D cursor. The patient could control these external devices by imagining movements like moving their hands to their mouth, rotating their shoulder, or touching their nose. An important finding was that the patient's neurons encoded both the goal and imagined trajectory of movements. To test the spatial goal and trajectory, they asked the patient to imagine moving the cursor to a cued target on a screen, sometimes masking it from view. From the 124 units recorded during this task, 19% coded only the movement's goal, 54% coded only the movement trajectory, and 27% coded both. Classification accuracy of more than 90% was achieved using the goal-oriented units. This finding is the first high-level decoding of motor intentions in humans, which shows an up-and-coming tool yet to be fully utilized in BCIs.

2.6. Outlook and Discussion

To decode the motor impulses until the task is finished, we require prior knowledge of the motivation for the maneuver. Studies suggest that due to the brain's distributed nature of motor planning, multiple cortical areas should be employed to operate the BCI systems better [140, 141]. However, even some developments using only one cortical area try to use other sources of information to accompany the motor signals to get better results. For example, in a reach-and-grasp task using a robotic arm, one study has decreased the dependency on motor signals by using a vision-guided shared control of the robotic arm [142]. As soon as the mounted camera on the

robotic arm detects a possible target in the direction of movement, it will form an appropriate hand positioning to grasp the object later. This autonomous movement will eliminate the need to decode grasp patterns from motor signals.

Another study recorded simultaneously from multiple cortical areas alongside M1, like dorsal premotor cortex (PMd), supplementary motor area (SMA), PPC, and primary somatosensory cortex (S1), to make the maximum use of available information from brain signals [143]. The study suggests that even though a significant sample of M1 neurons consistently provides the best estimation of all motor parameters (e.g., position and velocity), the other areas also contribute to BCI performance. Also, using more available neurons from the beginning will compensate for the loss of neurons later, and the performance stays at an acceptable level for a more extended period.

Decision-making BCIs are another example of using higher-level and motor signals to control a system. For example, one study has obtained information about the targets, goals, and motor preparation before movement onset, indicating the significant advantage of decision-making BCI over a traditional motor BCI [144]. This decision estimation will help the BCI correct errors and minimize processing delays. As a result, the controlled device will produce desired outputs (or perform the desired action) without heavy processing load and execution delays of continuous motor commands. Motor BCIs can significantly benefit from additional information like an abstract representation of actions or target objects [145]. For example, in the reach-and-grasp task, the motor cortex continuously encodes the kinematics to reach the target. Suppose the model can decode the user's cognitive state, for example, a visual feature of the intended target, like color or shape. Many potential targets can be eliminated, resulting in fast-responding BCI with minimal

delays and errors. Therefore, all available information should be combined synergistically to improve a traditional BCI.

2.7. Conclusions

Brain-computer interfaces are primarily developed to assist or restore motor-related shortcomings and to help the user communicate or control external devices. To this goal, going straight to motor signals from the brain is the right solution. However, other brain regions are involved in the decision-to-movement process. For example, the prefrontal cortex (PFC), or more precisely the lateral PFC, has been a valuable source of information [147], [149-154], which has yet to be explored for human BCIs. This area encodes high-level control and intention in the brain for decision-making. It also controls the eye movements used to decode saccade intentions in monkeys [147]. Additionally, the PFC is heavily involved in memory-related tasks, which could help the decoding models access the intended target through the information on shapes and colors. Finally, the PFC is an excellent candidate to be utilized along with the motor cortex in multi-area BCIs to improve error detection, correction, and decision-making intention estimation.

2.8. References

[1] Wolpaw J, Wolpaw EW (2012) Brain-Computer Interfaces: Principles and Practice. Oxford University Press, USA

[2] Lotze M, Montoya P, Erb M, et al. (1999) Activation of Cortical and Cerebellar Motor Areas during Executed and Imagined Hand Movements: An fMRI Study. J Cogn Neurosci 11:491–501. https://doi.org/10.1162/089892999563553

[3] Picard N, Strick PL (1996) Motor Areas of the Medial Wall: A Review of Their Location and Functional Activation. Cereb Cortex 6:342–353. https://doi.org/10.1093/cercor/6.3.342

[4] Roland PE, Larsen B, Lassen NA, Skinhoj E (1980) Supplementary motor area and other cortical areas in organization of voluntary movements in man. J Neurophysiol 43:118–136. https://doi.org/10.1152/jn.1980.43.1.118

[5] Schnitzler A, Salenius S, Salmelin R, et al. (1997) Involvement of Primary Motor Cortex in Motor Imagery: A Neuromagnetic Study. NeuroImage 6:201–208. https://doi.org/10.1006/nimg.1997.0286

[6] Mirabella G, Lebedev MA (2017) Interfacing to the brain's motor decisions. J Neurophysiol 117:1305–1319. https://doi.org/10.1152/jn.00051.2016

[7] Berti A, Bottini G, Gandola M, et al. (2005) Shared Cortical Anatomy for Motor Awareness and Motor Control. Science 309:488–491. https://doi.org/10.1126/science.1110625

[8] Desmurget M, Reilly KT, Richard N, et al. (2009) Movement Intention After Parietal Cortex Stimulation in Humans. Science 324:811–813. https://doi.org/10.1126/science.1169896

[9] Desmurget M, Sirigu A (2009) A parietal-premotor network for movement intention and motor awareness. Trends Cogn Sci 13:411–419. https://doi.org/10.1016/j.tics.2009.08.001

[10] Fried I, Katz A, McCarthy G, et al. (1991) Functional organization of human supplementary motor cortex studied by electrical stimulation. J Neurosci 11:3656–3666. https://doi.org/10.1523/JNEUROSCI.11-11-03656.1991

[11] Haggard P, Clark S, Kalogeras J (2002) Voluntary action and conscious awareness. Nat Neurosci 5:382. https://doi.org/10.1038/nn827

[12] Lau HC, Rogers RD, Haggard P, Passingham RE (2004) Attention to Intention. Science 303:1208–1210. https://doi.org/10.1126/science.1090973

[13] Singer W (2001) Consciousness and the Binding Problem. Ann N Y Acad Sci 929:123–146. https://doi.org/10.1111/j.1749-6632.2001.tb05712.x

[14] Sirigu A, Daprati E, Pradat-Diehl P, et al. (1999) Perception of self-generated movement following left parietal lesion. Brain 122:1867–1874. https://doi.org/10.1093/brain/122.10.1867

[15] Begliomini C, De Sanctis T, Marangon M, et al. (2014) An investigation of the neural circuits underlying reaching and reach-to-grasp movements: from planning to execution. Front Hum Neurosci 8:. https://doi.org/10.3389/fnhum.2014.00676

[16] Castiello U, Dadda M (2018) A review and consideration on the kinematics of reach-to-grasp movements in macaque monkeys. J Neurophysiol 121:188–204. https://doi.org/10.1152/jn.00598.2018

[17] Hoff B, Arbib MA (1993) Models of Trajectory Formation and Temporal Interaction of Reach and Grasp. J Mot Behav 25:175–192. https://doi.org/10.1080/00222895.1993.9942048

[18] Kukke SN, Curatalo LA, de Campos AC, et al. (2016) Coordination of reach-to-grasp kinematics in individuals with childhood-onset dystonia due to hemiplegic cerebral palsy.

IEEE Trans Neural Syst Rehabil Eng Publ IEEE Eng Med Biol Soc 24:582–590. https://doi.org/10.1109/TNSRE.2015.2458293

[19] Vaidya M, Kording K, Saleh M, et al. (2015) Neural coordination during reach-to-grasp. J Neurophysiol 114:1827–1836. https://doi.org/10.1152/jn.00349.2015

[20] Andersen RA, Buneo CA (2002) Intentional Maps in Posterior Parietal Cortex. Annu Rev Neurosci 25:189–220. https://doi.org/10.1146/annurev.neuro.25.112701.142922

[21] Buneo CA, Andersen RA (2006) The posterior parietal cortex: Sensorimotor interface for the planning and online control of visually guided movements. Neuropsychologia 44:2594–2606. https://doi.org/10.1016/j.neuropsychologia.2005.10.011

[22] Cisek P, Kalaska JF (2010) Neural Mechanisms for Interacting with a World Full of Action Choices. Annu Rev Neurosci 33:269–298. https://doi.org/10.1146/annurev.neuro.051508.135409

[23] Colby CL, Duhamel J-R (1996) Spatial representations for action in parietal cortex. Cogn Brain Res 5:105–115. https://doi.org/10.1016/S0926-6410(96)00046-8

[24] Colby CL, Goldberg ME (1999) Space and Attention in Parietal Cortex. Annu Rev Neurosci 22:319–349. https://doi.org/10.1146/annurev.neuro.22.1.319

[25] Bensmaia SJ, Miller LE (2014) Restoring sensorimotor function through intracortical interfaces: progress and looming challenges. Nat Rev Neurosci 15:313–325. https://doi.org/10.1038/nrn3724

[26] Churchland MM, Cunningham JP, Kaufman MT, et al. (2012) Neural population dynamics during reaching. Nature 487:51–56. https://doi.org/10.1038/nature11129

[27] Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT (1982) On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. J Neurosci 2:1527–1537. https://doi.org/10.1523/JNEUROSCI.02-11-01527.1982

[28] Humphrey DR, Schmidt EM, Thompson WD (1970) Predicting Measures of Motor Performance from Multiple Cortical Spike Trains. Science 170:758–762. https://doi.org/10.1126/science.170.3959.758

[29] Paninski L, Fellows MR, Hatsopoulos NG, Donoghue JP (2004) Spatiotemporal Tuning of Motor Cortical Neurons for Hand Position and Velocity. J Neurophysiol 91:515–532. https://doi.org/10.1152/jn.00587.2002

[30] Pohlmeyer EA, Solla SA, Perreault EJ, Miller LE (2007) Prediction of upper limb muscle activity from motor cortical discharge during reaching. J Neural Eng 4:369–379. https://doi.org/10.1088/1741-2560/4/4/003

[31] Schwartz AB (1994) Direct cortical representation of drawing. Science 265:540–542. https://doi.org/10.1126/science.8036499

[32] Schwartz AB, Kettner RE, Georgopoulos AP (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. I. Relations between single cell discharge and direction of movement. J Neurosci Off J Soc Neurosci 8:2913–2927

[33] Brandman DM, Cash SS, Hochberg LR (2017) Review: Human Intracortical Recording and Neural Decoding for Brain–Computer Interfaces. IEEE Trans Neural Syst Rehabil Eng 25:1687–1696. https://doi.org/10.1109/TNSRE.2017.2677443

[34] Bizzi E, Mussa-Ivaldi FA, Giszter S (1991) Computations underlying the execution of movement: a biological perspective. Science 253:287–291. https://doi.org/10.1126/science.1857964

[35] Pew RW (1984) A Distributed Processing View of Human Motor Control. In: Prinz W, Sanders AF (eds) Cognition and Motor Processes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 19–27

[36] Stelmach GE, Diggles VA (1982) Control theories in motor behavior. Acta Psychol (Amst) 50:83–105. https://doi.org/10.1016/0001-6918(82)90053-1

[37] Shanechi MM, Orsborn AL, Moorman HG, et al. (2017) Rapid control and feedback rates enhance neuroprosthetic control. Nat Commun 8:13825–13825. https://doi.org/10.1038/ncomms13825

[38] Baeg EH, Kim YB, Huh K, et al. (2003) Dynamics of Population Code for Working Memory in the Prefrontal Cortex. Neuron 40:177–188. https://doi.org/10.1016/S0896-6273(03)00597-X

[39] Bruce CJ, Goldberg ME (1985) Primate frontal eye fields. I. Single neurons discharging before saccades. J Neurophysiol 53:603–635. https://doi.org/10.1152/jn.1985.53.3.603

[40] Asaad WF, Rainer G, Miller EK (1998) Neural Activity in the Primate Prefrontal Cortex during Associative Learning. Neuron 21:1399–1407. https://doi.org/10.1016/S0896-6273(00)80658-3

[41] Fuster J (1997) The Prefrontal Cortex Anatomy, Physiology and Neuropsychology of the Frontal Lobe

[42] Fuster JM (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. J Neurophysiol 36:61–78. https://doi.org/10.1152/jn.1973.36.1.61

[43] Rainer G, Rao SC, Miller EK (1999) Prospective Coding for Objects in Primate Prefrontal Cortex. J Neurosci 19:5493–5505. https://doi.org/10.1523/JNEUROSCI.19-13-05493.1999

[44] Rao SC (1997) Integration of What and Where in the Primate Prefrontal Cortex. Science 276:821–824. https://doi.org/10.1126/science.276.5313.821

[45] Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953–956. https://doi.org/10.1038/35082081

[46] Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50:7–15. https://doi.org/10.1016/0010-0277(94)90018-3

[47] Damasio Antonio R., Everitt Barry John, Bishop Dorothy, et al. (1996) The somatic marker hypothesis and the possible functions of the prefrontal cortex. Philos Trans R Soc Lond B Biol Sci 351:1413–1420. https://doi.org/10.1098/rstb.1996.0125

[48] Drevets WC, Price JL, Simpson JR, et al. (1997) Subgenual prefrontal cortex abnormalities in mood disorders. Nature 386:824. https://doi.org/10.1038/386824a0

[49] Weinberger DR, Berman KF, Zec RF (1986) Physiologic Dysfunction of Dorsolateral Prefrontal Cortex in Schizophrenia: I. Regional Cerebral Blood Flow Evidence. Arch Gen Psychiatry 43:114–124. https://doi.org/10.1001/archpsyc.1986.01800020020004

[50] Kobatake E, Wang G, Tanaka K (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. J Neurophysiol 80:324–330. https://doi.org/10.1152/jn.1998.80.1.324

[51] Murata A, Gallese V, Luppino G, et al. (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. J Neurophysiol 83:2580–2601. https://doi.org/10.1152/jn.2000.83.5.2580

[52] Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398:704. https://doi.org/10.1038/19525

[53] Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci Off J Soc Neurosci 16:5154–67

[54] Baddeley A (1992) Working memory. Science 255:556–559. https://doi.org/10.1126/science.1736359

[55] Bechara A, Damasio H, Tranel D, Anderson SW (1998) Dissociation Of Working Memory from Decision Making within the Human Prefrontal Cortex. J Neurosci 18:428–437. https://doi.org/10.1523/JNEUROSCI.18-01-00428.1998

[56] Braver TS, Cohen JD, Nystrom LE, et al. (1997) A Parametric Study of Prefrontal Cortex Involvement in Human Working Memory. NeuroImage 5:49–62. https://doi.org/10.1006/nimg.1996.0247

[57] Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7:415–423. https://doi.org/10.1016/S1364-6613(03)00197-9

[58] Ericsson KA, Kintsch W (1995) Long-term working memory. Psychol Rev 102:211–245. https://doi.org/10.1037/0033-295X.102.2.211

[59] Goldman-Rakic PS (1995) Cellular basis of working memory. Neuron 14:477–485. https://doi.org/10.1016/0896-6273(95)90304-6

[60] Kane MJ, Engle RW (2002) The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. Psychon Bull Rev 9:637–671. https://doi.org/10.3758/BF03196323

[61] Markowitz DA, Curtis CE, Pesaran B (2015) Multiple component networks support working memory in prefrontal cortex. Proc Natl Acad Sci U S A 112:11084–9. https://doi.org/10.1073/pnas.1504172112

[62] Im C, Seo J-M (2016) A review of electrodes for the electrical brain signal recording. Biomed Eng Lett 6:104–112. https://doi.org/10.1007/s13534-016-0235-1

[63] Renshaw B, Forbes A, Morison BR (1940) Activity of isocortex and hippocampus: electrical studies with micro-electrodes. J Neurophysiol 3:74–105. https://doi.org/10.1152/jn.1940.3.1.74

[64] Wolbarsht ML, MacNichol EF, Wagner HG (1960) Glass Insulated Platinum Microelectrode. Science 132:1309–1310. https://doi.org/10.1126/science.132.3436.1309

[65] Geddes LA, Roeder R (2003) Criteria for the selection of materials for implanted electrodes. Ann Biomed Eng 31:879–890

[66] Kennedy PR (1989) The cone electrode: a long-term electrode that records from neurites grown onto its recording surface. J Neurosci Methods 29:181–193. https://doi.org/10.1016/0165-0270(89)90142-8

[67] Jansen RF, Ter Maat A (1992) Automatic wave form classification of extracellular multineuron recordings. J Neurosci Methods 41:123–132. https://doi.org/10.1016/0165-0270(92)90055-I

[68] Kreiter AK, Aertsen AMHJ, Gerstein GL (1989) A low-cost single-board solution for real-time, unsupervised waveform classification of multineuron recordings. J Neurosci Methods 30:59–69. https://doi.org/10.1016/0165-0270(89)90075-7

[69] Salganicoff M, Sarna M, Sax L, Gerstein GL (1988) Unsupervised waveform classification for multi-neuron recordings: a real-time, software-based system. I. Algorithms and implementation. J Neurosci Methods 25:181–187. https://doi.org/10.1016/0165-0270(88)90132-X

[70] Wheeler BC, Heetderks WJ (1982) A Comparison of Techniques for Classification of Multiple Neural Signals. IEEE Trans Biomed Eng BME-29:752–759. https://doi.org/10.1109/TBME.1982.324870

[71] McNaughton BL, O'Keefe J, Barnes CA (1983) The stereotrode: A new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. J Neurosci Methods 8:391–397. https://doi.org/10.1016/0165-0270(83)90097-3

[72] Marg E, Adams JE (1967) Indwelling multiple micro-electrodes in the brain. Electroencephalogr Clin Neurophysiol 23:277–280

[73] Jones KE, Campbell PK, Normann RA (1992) A glass/silicon composite intracortical electrode array. Ann Biomed Eng 20:423–437. https://doi.org/10.1007/BF02368134

[74] Hoogerwerf AC, Wise KD (1994) A three-dimensional microelectrode array for chronic neural recording. IEEE Trans Biomed Eng 41:1136–1146. https://doi.org/10.1109/10.335862

[75] Moxon KA, Leiser SC, Gerhardt GA, et al. (2004) Ceramic-based multisite electrode arrays for chronic single-neuron recording. IEEE Trans Biomed Eng 51:647–656. https://doi.org/10.1109/TBME.2003.821037

[76] Wester BA, Lee RH, LaPlaca MC (2009) Development and characterization of in vivo flexible electrodes compatible with large tissue displacements. J Neural Eng 6:024002. https://doi.org/10.1088/1741-2560/6/2/024002

[77] Nicolas-Alonso LF, Gomez-Gil J (2012) Brain Computer Interfaces, a Review. Sensors 12:1211–1279. https://doi.org/10.3390/s120201211

[78] Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407–420. https://doi.org/10.1038/nrn3241

[79] Burns SP, Xing D, Shapley RM (2011) Is Gamma-Band Activity in the Local Field Potential of V1 Cortex a `Clock' or Filtered Noise? J Neurosci Off J Soc Neurosci 31:9658–9664. https://doi.org/10.1523/JNEUROSCI.0660-11.2011

[80] Waldert S, Pistohl T, Braun C, et al. (2009) A review on directional information in neural signals for brain-machine interfaces. J Physiol-Paris 103:244–254. https://doi.org/10.1016/j.jphysparis.2009.08.007

[81] Milner B (1963) Effects of Different Brain Lesions on Card Sorting: The Role of the Frontal Lobes. Arch Neurol 9:90–100. https://doi.org/10.1001/archneur.1963.00460070100010

[82] Muhammad R, Wallis JD, Miller EK (2006) A Comparison of Abstract Rules in the Prefrontal Cortex, Premotor Cortex, Inferior Temporal Cortex, and Striatum. J Cogn Neurosci 18:974–989. https://doi.org/10.1162/jocn.2006.18.6.974

[83] Merten K, Nieder A (2012) Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. Proc Natl Acad Sci 109:6289–6294. https://doi.org/10.1073/pnas.1121084109

[84] Murray JD, Bernacchia A, Roy NA, et al. (2017) Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. Proc Natl Acad Sci 114:394–399. https://doi.org/10.1073/pnas.1619449114

[85] Gerstner W, Kistler WM (2002) Spiking Neuron Models: Single Neurons, Populations, Plasticity. Cambridge University Press

[86] Mukamel R, Ekstrom AD, Kaplan J, et al (2010) Single-Neuron Responses in Humans during Execution and Observation of Actions. Curr Biol 20:750–756. https://doi.org/10.1016/j.cub.2010.02.045

[87] Amirikian B, Georgopulos AP (2000) Directional tuning profiles of motor cortical cells. Neurosci Res 36:73–79. https://doi.org/10.1016/S0168-0102(99)00112-1

[88] Rolls ET, Changeux J -p, Konishi M, et al. (1987) Information Representation, Processing, and Storage in the Brain: Analysis at the Single Neuron Level

[89] Chestek CA, Batista AP, Santhanam G, et al. (2007) Single-Neuron Stability during Repeated Reaching in Macaque Premotor Cortex. J Neurosci 27:10742–10750. https://doi.org/10.1523/JNEUROSCI.0959-07.2007

[90] Frost WN, Katz PS (1996) Single neuron control over a complex motor program. Proc Natl Acad Sci 93:422–426. https://doi.org/10.1073/pnas.93.1.422

[91] Schmidt EM (1980) Single neuron recording from motor cortex as a possible source of signals for control of external devices. Ann Biomed Eng 8:339–349. https://doi.org/10.1007/BF02363437

[92] Truccolo W, Hochberg LR, Donoghue JP (2010) Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. Nat Neurosci 13:105–111. https://doi.org/10.1038/nn.2455

[93] Barak O, Tsodyks M, Romo R (2010) Neuronal Population Coding of Parametric Working Memory. J Neurosci 30:9424–9430. https://doi.org/10.1523/JNEUROSCI.1875-10.2010

[94] Bastos AM, Loonis R, Kornblith S, et al. (2018) Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. Proc Natl Acad Sci 115:1117–1122. https://doi.org/10.1073/pnas.1710323115

[95] Honkanen R, Rouhinen S, Wang SH, et al (2015) Gamma Oscillations Underlie the Maintenance of Feature-Specific Information and the Contents of Visual Working Memory. Cereb Cortex 25:3788–3801. https://doi.org/10.1093/cercor/bhu263

[96] Howard MW (2003) Gamma Oscillations Correlate with Working Memory Load in Humans. Cereb Cortex 13:1369–1374. https://doi.org/10.1093/cercor/bhg084

[97] Miller EK, Lundqvist M, Bastos AM (2018) Working Memory 2.0. Neuron

[98] Roux F, Wibral M, Mohr HM, et al. (2012) Gamma-Band Activity in Human Prefrontal Cortex Codes for the Number of Relevant Items Maintained in Working Memory. J Neurosci 32:12411–12420. https://doi.org/10.1523/JNEUROSCI.0421-12.2012

[99] Lundqvist M, Rose J, Herman P, et al (2016) Gamma and Beta Bursts Underlie Working Memory. Neuron 90:152–164. https://doi.org/10.1016/j.neuron.2016.02.028

[100] Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science 233:1416–1419. https://doi.org/10.1126/science.3749885

[101] Chapin JK, Moxon KA, Markowitz RS, Nicolelis MAL (1999) Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. Nat Neurosci 2:664–670. https://doi.org/10.1038/10223

[102] Georgopoulos AP, Kettner RE, Schwartz AB (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. J Neurosci 8:2928–2937. https://doi.org/10.1523/JNEUROSCI.08-08-02928.1988

[103] Shamir M (2014) Emerging principles of population coding: in search for the neural code. Curr Opin Neurobiol 25:140–148. https://doi.org/10.1016/j.conb.2014.01.002

[104] Ethier C, Oby ER, Bauman MJ, Miller LE (2012) Restoration of grasp following paralysis through brain-controlled stimulation of muscles. Nature 485:368–371. https://doi.org/10.1038/nature10987

[105]Serruya MD, Hatsopoulos NG, Paninski L, et al. (2002) Brain-machine interface: Instant neural control of a movement signal. Nature 416:141–142. https://doi.org/10.1038/416141a

[106]Cunningham JP, Yu BM (2014) Dimensionality reduction for large-scale neural recordings. Nat Neurosci 17:1500–1509. https://doi.org/10.1038/nn.3776

[107]Gallego JA, Perich MG, Miller LE, Solla SA (2017) Neural Manifolds for the Control of Movement. Neuron 94:978–984. https://doi.org/10.1016/j.neuron.2017.05.025

[108]Fu M, Yu X, Lu J, Zuo Y (2012) Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. Nature 483:92–95. https://doi.org/10.1038/nature10844

[109]Rioult-Pedotti M-S, Friedman D, Hess G, Donoghue JP (1998) Strengthening of horizontal cortical connections following skill learning. Nat Neurosci 1:230–234. https://doi.org/10.1038/678

[110]Sadtler PT, Quick KM, Golub MD, et al. (2014) Neural constraints on learning. Nature 512:423–426. https://doi.org/10.1038/nature13665

[111]Aflalo TN, Graziano MSA (2007) Relationship between Unconstrained Arm Movements and Single-Neuron Firing in the Macaque Motor Cortex. J Neurosci 27:2760–2780. https://doi.org/10.1523/JNEUROSCI.3147-06.2007

[112]Mussa-Ivaldi FA (1988) Do neurons in the motor cortex encode movement direction? An alternative hypothesis. Neurosci Lett 91:106–111. https://doi.org/10.1016/0304-3940(88)90257-1

[113]Scott SH (2008) Inconvenient Truths about neural processing in primary motor cortex. J Physiol 586:1217–1224. https://doi.org/10.1113/jphysiol.2007.146068

[114]Todorov E (2000) Direct cortical control of muscle activation in voluntary arm movements: a model. Nat Neurosci 3:391–398. https://doi.org/10.1038/73964

[115]Ashe J, Georgopoulos AP (1994) Movement parameters and neural activity in motor cortex and area 5. Cereb Cortex N Y N 1991 4:590–600. https://doi.org/10.1093/cercor/4.6.590

[116]Evarts EV, Fromm C, Kroller J, Jennings VA (1983) Motor Cortex control of finely graded forces. J Neurophysiol 49:1199–1215. https://doi.org/10.1152/jn.1983.49.5.1199

[117]Fetz EE, Finocchio DV (1975) Correlations between activity of motor cortex cells and arm muscles during operantly conditioned response patterns. Exp Brain Res 23:217–240. https://doi.org/10.1007/BF00239736

[118]Georgopoulos AP, Ashe J, Smyrnis N, Taira M (1992) The motor cortex and the coding of force. Science 256:1692–1695. https://doi.org/10.1126/science.256.5064.1692

[119]Georgopoulos AP, Caminiti R, Kalaska JF, Massey JT (1983) Spatial coding of movement: A hypothesis concerning the coding of movement direction by motor cortical populations. Exp Brain Res 49:327–336

[120]Moran DW, Schwartz AB (1999) Motor cortical representation of speed and direction during reaching. J Neurophysiol 82:2676–2692. https://doi.org/10.1152/jn.1999.82.5.2676

[121]Ashe J (1997) Erratum to “Force and the motor cortex”: [Behavioural Brain Research 86 (1997) 1–15]1PII of original article: S0166-4328(96)00145-31. Behav Brain Res 87:255–269. https://doi.org/10.1016/S0166-4328(97)00752-3

[122]Kalman RE (1960) A New Approach to Linear Filtering and Prediction Problems. Trans ASME–Journal Basic Eng 35–45

[123]Jazwinski AH (1970) Stochastic processes and filtering theory. Academic Press, New York, N.Y.

[124]Julier SJ, Uhlmann JK (1997) New extension of the Kalman filter to nonlinear systems. In: Signal Processing, Sensor Fusion, and Target Recognition VI. International Society for Optics and Photonics, pp 182–194

[125]Gilja V, Nuyujukian P, Chestek CA, et al. (2012) A brain machine interface control algorithm designed from a feedback control perspective. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp 1318–1322

[126]Tseng P-H, Urpi NA, Lebedev M, Nicolelis M (2019) Decoding Movements from Cortical Ensemble Activity Using a Long Short-Term Memory Recurrent Network. Neural Comput 1–29. https://doi.org/10.1162/neco_a_01189

[127]Hearst MA, Dumais ST, Osuna E, et al. (1998) Support vector machines. IEEE Intell Syst Their Appl 13:18–28. https://doi.org/10.1109/5254.708428

[128]Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). pp 1–6

[129]Hochberg LR, Serruya MD, Friehs GM, et al. (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442:164–171. https://doi.org/10.1038/nature04970

[130]Hochberg LR, Bacher D, Jarosiewicz B, et al. (2012) Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature 485:372–375. https://doi.org/10.1038/nature11076

[131]Ajiboye AB, Willett FR, Young DR, et al. (2017) Restoration of reaching and grasping in a person with tetraplegia through brain-controlled muscle stimulation: a proof-of-concept demonstration. Lancet Lond Engl 389:1821–1830. https://doi.org/10.1016/S0140-6736(17)30601-3

[132]Willett FR, Avansino DT, Hochberg LR, et al. (2021) High-performance brain-to-text communication via handwriting. Nature 593:249–254. https://doi.org/10.1038/s41586-021-03506-2

[133]Collinger JL, Wodlinger B, Downey JE, et al. (2013) High-performance neuroprosthetic control by an individual with tetraplegia. The Lancet 381:557–564. https://doi.org/10.1016/S0140-6736(12)61816-9

[134]Bouton CE, Shaikhouni A, Annetta NV, et al. (2016) Restoring cortical control of functional movement in a human with quadriplegia. Nature 533:247–250. https://doi.org/10.1038/nature17435

[135]Sharma G, Friedenberg DA, Annetta N, et al. (2016) Using an Artificial Neural Bypass to Restore Cortical Control of Rhythmic Movements in a Human with Quadriplegia. Sci Rep 6: https://doi.org/10.1038/srep33807

[136]Handelman DA, Osborn LE, Thomas TM, et al. (2022) Shared Control of Bimanual Robotic Limbs With a Brain-Machine Interface for Self-Feeding. Front Neurorobotics 16:

[137]Min B-K, Chavarriaga R, Millán JDR (2017) Harnessing Prefrontal Cognitive Signals for Brain-Machine Interfaces. Trends Biotechnol 35:585–597. https://doi.org/10.1016/j.tibtech.2017.03.008

[138]Musallam S, Corneil BD, Greger B, et al. (2004) Cognitive Control Signals for Neural Prosthetics. Science 305:258–262. https://doi.org/10.1126/science.1097938

[139]Aflalo T, Kellis S, Klaes C, et al. (2015) Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. Science 348:906–910. https://doi.org/10.1126/science.aaa5417

[140]Taylor DM, Tillery SIH, Schwartz AB (2002) Direct Cortical Control of 3D Neuroprosthetic Devices. Science 296:1829–1832. https://doi.org/10.1126/science.1070291

[141] Wessberg J, Stambaugh CR, Kralik JD, et al. (2000) Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature 408:361–365. https://doi.org/10.1038/35042582

[142] Downey JE, Weiss JM, Muelling K, et al. (2016) Blending of brain-machine interface and vision-guided autonomous robotics improves neuroprosthetic arm performance during grasping. J NeuroEngineering Rehabil 13:28–28. https://doi.org/10.1186/s12984-016-0134-9

[143] Carmena JM, Lebedev MA, Crist RE, et al (2003) Learning to Control a Brain–Machine Interface for Reaching and Grasping by Primates. PLoS Biol 1:. https://doi.org/10.1371/journal.pbio.0000042

[144] Mirabella G, Lebedev MA (2017) Interfacing to the brain's motor decisions. J Neurophysiol 117:1305–1319. https://doi.org/10.1152/jn.00051.2016

[145] Murata A, Fadiga L, Fogassi L, et al (1997) Object Representation in the Ventral Premotor Cortex (Area F5) of the Monkey. J Neurophysiol 78:2226–2230. https://doi.org/10.1152/jn.1997.78.4.2226

[146] Lyle RC (1981) A performance test for assessment of upper limb function in physical rehabilitation treatment and research. Int J Rehabil Res 4:483–492

[147] Boulay CB, Pieper F, Leavitt M, et al. (2016) Single-trial decoding of intended eye movement goals from lateral prefrontal cortex neural ensembles. J Neurophysiol 115:486–499. https://doi.org/10.1152/jn.00788.2015.-Neu

[148] Andersen RA, Hwang EJ, Mulliken GH (2010) Cognitive Neural Prosthetics. Annu Rev Psychol 61:169–190. https://doi.org/10.1146/annurev.psych.093008.100503

[149]Mendoza-Halliday D, Torres S, Martinez-Trujillo JC (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. Nat Neurosci 17:1255–1262. https://doi.org/10.1038/nn.3785

[150]Leavitt ML, Pieper F, Sachs AJ, Martinez-Trujillo JC (2018) A quadrantic bias in prefrontal representation of visual-mnemonic space. Cerebral Cortex 28:2405–2421

[151]Backen T, Treue S, Martinez-Trujillo JC (2018) Encoding of Spatial Attention by Primate Prefrontal Cortex Neuronal Ensembles. eNeuro 5:ENEURO.0372-16.2017. https://doi.org/10.1523/ENEURO.0372-16.2017

[152]Johnston R, Doucet G, Boulay C, et al (2021) Decoding Saccade Intention From Primate Prefrontal Cortical Local Field Potentials Using Spectral, Spatial, and Temporal Dimensionality Reduction. Int J Neur Syst 31:2150023. https://doi.org/10.1142/S0129065721500234

[153]Johnston R, Abbass M, Corrigan B, et al (2023) Decoding spatial locations from primate lateral prefrontal cortex neural activity during virtual navigation. J Neural Eng. https://doi.org/10.1088/1741-2552/acb5c2

[154]Tremblay S, Doucet G, Pieper F, et al (2015) Single-Trial Decoding of Visual Attention from Local Field Potentials in the Primate Lateral Prefrontal Cortex Is Frequency-Dependent. J Neurosci 35:9038–9049. https://doi.org/10.1523/JNEUROSCI.1041-15.2015

Chapter 3:
Ensembles codes for associative learning in the primate lateral prefrontal cortex

3.1. Summary

The lateral prefrontal cortex (LPFC) of primates is thought to play a role in associative learning. However, it remains unclear how neuronal ensembles in this area dynamically encode and store memories for arbitrary stimulus-response associations. We recorded the activity of neurons in LPFC of two macaques during an associative learning task using multielectrode arrays. During task trials, the colour of a symbolic cue indicated the location of one of two possible targets for a saccade. During a trial block, multiple randomly chosen associations/rules were learned by the subjects. We conducted a state-space analysis and found that LPFC neuronal ensembles rapidly learn new stimulus-response associations mirroring the animals' learning rates. Multiple associations acquired during a training session can be stored in a neuronal subspace. They can be retrieved hours or days after learning. Finally, ensemble knowledge of old associations can facilitate learning of new, similar associations. These results indicate that neuronal ensembles in the primate LPFC provide a flexible and dynamic substrate for associative learning.

3.2. Introduction

Learning to associate visual information with a deeper symbolic meaning is a trademark of primate cognition and behaviour. For example, prosocial behaviour in chimpanzees has been attributed, in large part, to the associative learning [1]. In humans, associative learning has long been recognized as a key cognitive domain necessary for successful environmental and social interactions [2]–[4]. The prefrontal cortex has been implicated in associative learning. There is a demonstrated

relationship between the increasing size of the prefrontal cortex and phylogenetically and cognitively more advanced species [5]. In its simplest form, associative learning involves mapping (and remapping) arbitrary stimulus-response associations [6], [7]. In humans, Milner [8] demonstrated that lesions of the dorsolateral frontal lobe result in impaired performance in learning new arbitrary stimulus-cue associations (rules), resulting in perseverating old rules. These findings have been replicated in macaques with lateral prefrontal cortex (LPFC) lesions performing a similar task [9].

The LPFC receives input from different sensory systems and sends projections to multiple cortical areas [10]. Neurons in LPFC encode sensory and motor signals, movement planning and execution, working memory, attention, reward expectation, and task rules [11]–[22]. Moreover, the LPFC is one of the brain regions in which changes in the selectivity of neurons during learning of new task rules have been documented [18], [23], [24]. This suggests circuits within the LPFC possess a degree of plasticity and connectivity with the rest of the brain, allowing them to store arbitrary stimulus-response associations (referred to as rules hereafter) and adjust to the contingencies of changing environments. For example, the decision to reach for one out of two objects may quickly change depending on the context (e.g., if we have wine, then we may reach for cheese, but if we are having coffee or tea, then we may reach for a pastry at the same table).

A previous study [13] has shown that during an associative learning task, single neurons in the LPFC of macaque monkeys can represent associations between visual cues (e.g., the colour of a fixation point) and motor responses (e.g., a saccade in a particular direction). The study showed that while representations of saccade direction seem to remain stable during learning, the representation of the cue seems to change as animals learn the task. This selectivity is similar to the one observed in areas involved in the coding of associative memories, such as the hippocampus

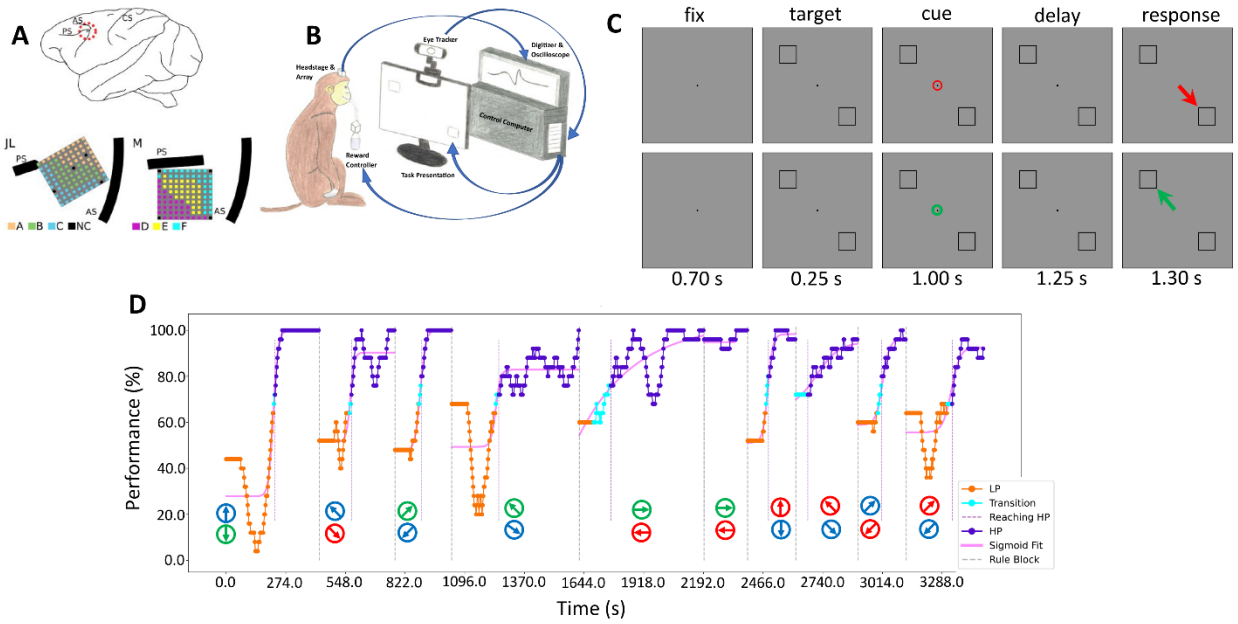

Figure 3.1. Electrophysiological Recordings and Behavioural Task

(A) Anatomical location of chronic implant for monkey JL and monkey M. A, B, C, D, E, and F are different channel groups recorded during different sessions. NCs are the not-connected channels.

(B) The task and reward setup in which the LPFC activity is recorded along with the eye movements.

(C) The different periods of the task in which the monkey should make a saccade to one of the two presented targets based on the colour of the cue.

(D) Each session contains blocks of trials with different colour rules. The vertical gray dashed lines separates different blocks. The trials before the sigmoid fit (pink line) reaches 65% are considered LP (orange), the trials after the sigmoid fit crosses 75% are considered HP (purple), and the trials in between are the transitioning period (cyan). The vertical purple dashed line indicates the start of the HP trials.

[25] and the parietal cortex [26]. Over the last few years, studies in the LPFC have shown that single neuron's performance, evaluated neuro-metrically, explained less behavioural variability than the performance of neuronal populations [27], [28]. It is currently unclear how ensembles of LPFC neurons acquire selectivity for new stimulus-response associations during associative

learning tasks and how they encode and store such associations in the distributed activity of neuronal populations.

We address these issues by recording the activity of neuronal ensembles in the LPFC of behaving monkeys using chronically implanted multielectrode arrays (Figure 3.1a) during an associative learning task (Figure 3.1b). During the task, animals made a saccade to one of two targets presented on a computer screen depending on the colour of a central fixation point to obtain a fluid reward (e.g., if green, saccade to target 1, if red, saccade to target 2) (Figure 3.1c). Once the animal learned the rule and performed above the 75% accuracy level for some random number of trials, we changed the meaning of the colours either to a new colour combination or to a combination that included one or two of the colours but changed the colour-target mapping contingency. We found that neuronal ensembles in the LPFC acquire selectivity for the novel component of the rule (colour) and its association with a saccade direction as the animals learn the task. Furthermore, we trained a deep neural network on the recorded data and demonstrated neuronal ensembles in LPFC rapidly adapt during a session to encode the new rules. We conducted state-space analyses, showing that the rules are encoded in separate neural subspaces. The distance between the attributes of the encoded rules reflects the similarity between rules and the time it took the animals to acquire the rule.

3.3. Results

3.3.1. Behaviour

The animals performed the task in blocks (vertical gray dashed lines in Figure 3.1d separates the blocks). As a block begins, the monkey learns the rules for that block based on receiving the reward for correct trials only, resulting in them choosing the correct target more often i.e., better

performance from the monkey (Figure 3.1d). When a new block with new rules starts, there is a drop in performance. However, as the monkey progresses in a block, the performance improves. In blocks with repeated rules (e.g., red colour saccade to the left target) relative to previous blocks, the monkeys tend to perform well from the beginning, which suggests that the monkeys store the learned rule in long-term memory and retrieve it when needed. Notice that in our paradigm, the rules for a block could be new rules or the combination of a new and an old rule. We fit a sigmoid function to parameterize the monkeys' performance curve. We consider the performance above 75% as indicating the animals had acquired the task rules [29]. With a gap of one standard deviation of the performance distribution, performance below 65% was considered to indicate that the animals had not acquired the task rules yet (horizontal dashed lines in Figure 3.1d). During the trials with performance between 65% and 75%, the animals had some knowledge of the rule but had not fully acquired them. The trials in which the monkey had not acquired the rule yet are used in our analyses as the low-performance trials (LP), and the trials in which the monkey has acquired the rule as the high-performance (HP) trials.

3.3.2. Electrophysiological data modeling

As a confirmatory step, we explored the recordings to observe the emergence/increase of the rule information coding in HP compared to LP trials. For this section we analyzed the neuronal activities in single channel and single unit levels. The single unit activities were obtained by thresholding the multiunit activities of single channel responses filtered at >250Hz, followed by automatic spike sorting using MKSort (Matthew Kaufman and Ripple, Inc., https://github.com/ripple-neuro/mksort).

To confirm the existence and persistence of rule information in single channels, we used a rule decoding SVM-based model to find the single channel performances (Figure 3.2A). In one session

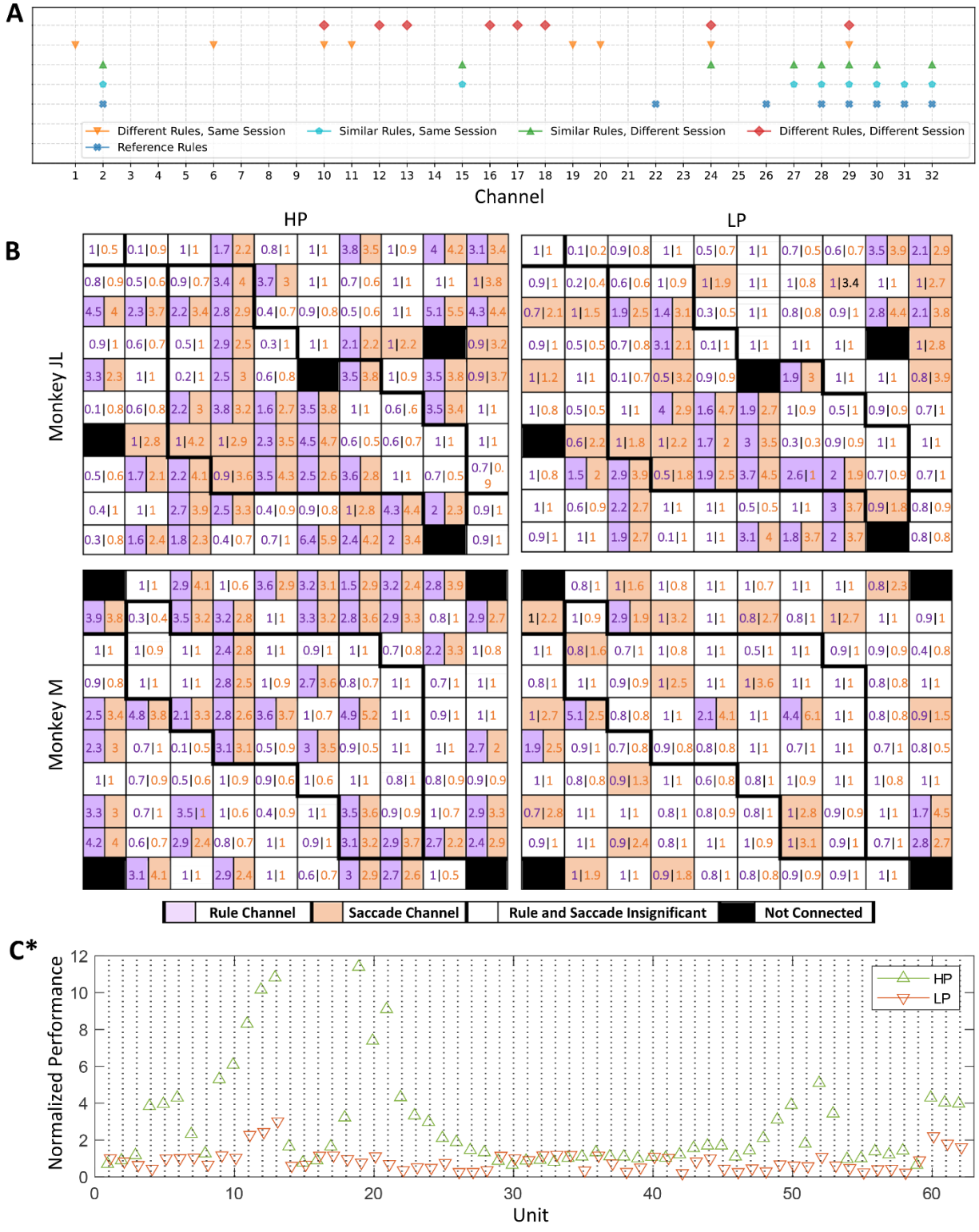
A
Different Rules, Same Session
Similar Rules, Same Session
Similar Rules, Different Session
Different Rules, Different Session
Reference Rules
Channel
B
HP
LP
Monkey JL
Monkey M
Rule Channel
Saccade Channel
Rule and Saccade Insignificant
Not Connected
C*
Normalized Performance
Unit
HP
LP

Figure 3.2. Single Channel and Single Unit Analysis

(A) Analysis of best performing channels for rule decoding. A high number of the best performing channels decoding a given rule recorded from the same electrode bank remains the same in the same session and different block (75%), and different session (62.5%). However, 12.5% of the best performing channels decoding different rules from the same and different session remain the same.

(B) The normalized rule and saccade log-likelihood of single channels. White channels have lower than chance log-likelihood encoding rule (purple numbers) and saccade (orange numbers). Channels with significant saccade log-likelihood are "saccade channels" (orange). Channels with significant rule log-likelihood are "rule channels" (purple). A channel can have significant log-likelihood for saccade and rule (split channels). The black electrodes are not connected to wires. Comparing the LP and HP trials, the number of saccade channels have increased by 26% (38 to 48) in monkey JL and 60% (25 to 40) in monkey M. On the other hand, comparing the LP to HP trials, the number of rule channels have increased by 56% (25 to 39) in monkey JL and 485% (7 to 41) in monkey M.

(C) Increase in single unit rule decoding performance from LP (red rectangular) to HP (green rectangular) trials. *p=4.34e-11, N=62 for each distribution.

(X) (see Figure A1) with repeating rules (two blocks with the same rules: B1 and B2), we trained the model on all trials except B2 and tested on B1 in a 3-fold cross-validation scheme and recorded the top 25% performing channels (blue x in Figure 3.2A). We repeated the analysis in session X, training on all trials except B1, testing B2 in a 3-fold cross-validation scheme, and recording the top 25% performing channels (cyan octagons in Figure 3.2A). Moreover, we trained the model on session X and tested it on a block containing different rules (B3) in a 3-fold cross-validation scheme and recorded the top 25% performing channels (orange rectangles in Figure 3.2A). In addition, in another session (Y) recorded from the same electrode bank, which contained the same rules as B1 and B2, we repeated the analysis on a similar block, B4. We trained the model on

session Y, tested it on B4 in a 3-fold cross-validation scheme, and recorded the top 25% performing channels (green rectangles in Figure 3.2A). Lastly, we trained the model on session Y, tested it on a block with different rules (B5) in a 3-fold cross-validation scheme, and recorded the top 25% performing channels (red diamonds in Figure 3.2A). In Session X, 75% of the top-performing channels in decoding the B2 rules remained the same as B1, while 12.5% of the top-performing channels decoding the B3 rules were the same as B1. Also, in Session Y, 62.5% of the top-performing channels in decoding B4 rules remained the same as B1, while 12.5% of the top-performing channels decoding the B5 rules were the same as B1.

In another single-channel analysis, we used a generalized additive model [30] to explore the relationship between the single-channel neuronal responses and different task components (see Methods). The saccade direction and the rule were tested against the single-channel neuronal responses in HP and LP trials (Figure 3.2B). The model finds the log-likelihood of neuronal responses given a certain condition (saccade and rule) with the chance likelihood. Normalizing the given likelihoods with their chance level, we recorded each channel's saccade (orange numbers) and rule (purple numbers) normalized likelihood. Channels with a lower than the chance likelihood for saccade and rule are white. The channels with higher than chance likelihoods for saccade are the 'saccade channel' (orange channels). The 'rule channel' (purple channels) have higher than chance likelihoods for the rule. A channel can be a 'rule channel' and a 'saccade channel' simultaneously if both likelihoods are higher than chance (split channels). The black electrodes are not connected to any channel. Comparing the LP and HP trials, the number of saccade channels has increased by 26% (38 to 48) in monkey JL and 60% (25 to 40) in monkey M. On the other hand, comparing the LP to HP trials, the number of rule channels has increased by 56% (25 to 39) in monkey JL and 485% (7 to 41) in monkey M (Figure 3.2B) (also see Figure A2).

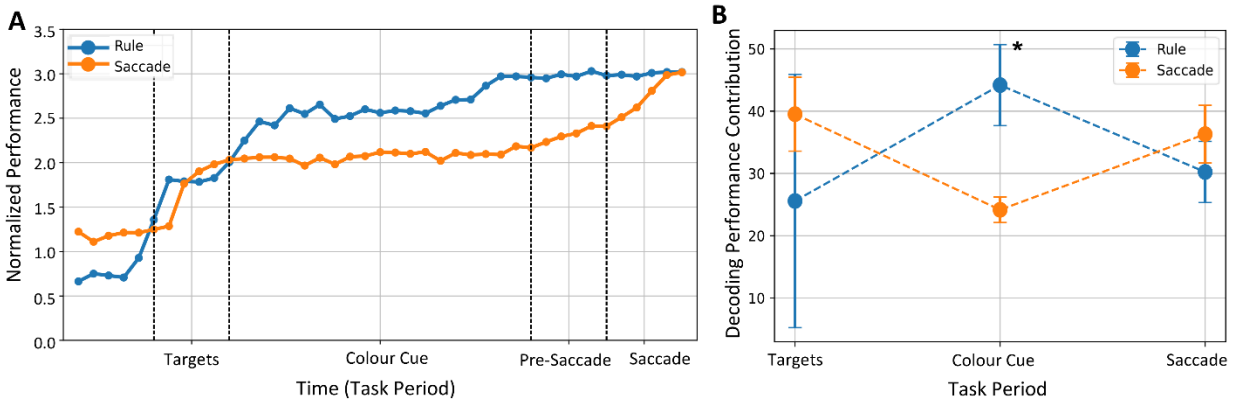

Figure 3.3. Temporal Analysis of Rule and Saccade Decoding

(A) The performance of the rule decoder and the saccade decoder given the neuronal activity recorded through different periods of the task (increasing window of input information) normalized by the decoders' chance level.

(B) The contribution of the information coded in different periods of the task to the decoding model's performance across all sessions. *p=0.019 for monkey JL, p=0.029 for monkey M comparing the performance contribution distribution of the colour cue period in rule decoder (blue) versus saccade decoder (orange). N=100 for each distribution.

Moreover, as an additional confirmatory step, we sorted the single channel recordings of one session into single unit activities to explore the increase in coding of rule information in HP trials compared to LP. We used an SVM-based model to decode the rules from the single unit spike rates (Figure 3.2C) normalized by each unit's chance level rule decoding. The normalized single unit rule decoding increased significantly from LP to HP trials in the 62 sorted units (p=4.34e-11 Kruskal-Wallis [31], and Dunn [32] post-hoc) (Table 3.1).

3.3.3. Decoding Analysis

Using a support vector machine classifier, we decoded the associated rule and the monkeys' intended saccade direction in each trial from all of the channel activity at multiple intervals after

the target onset. Compared to the chance decoding level, the rule decoder achieves a high performance level faster than the saccade decoder. This is likely due to utilizing the colour cue period more than the saccade information. It is important to note that the rule model is decoding 13 different rules in Figure 3.3A while the saccade model decodes 8 different directions. As a result, decoding performance normalized by the chance level (performance after shuffling trial labels) is used to compare both models (Figure 3.3A). Also, across multiple decoding repetitions, the errors are too small to be visible in the normalized scale, therefore, are not shown in Figure 3.3A.

We separated these intervals into three task periods: target presentation, colour cue, and saccade planning/movement. We analyzed the saccade and rule decoding models' performance increase across all sessions.

The three task periods have different relative contributions of rule decoding and saccade decoding. On average, saccade models showed higher performance during the target presentation and the saccade period. However, during the colour cue period, the rule decoding models performed better ($p=0.019$ for monkey JL, $p=0.029$ for monkey M, Kruskal-Wallis, and Dunn post-hoc) (Table 3.1, Methods). The statistics from all sessions show that decodable information, important to distinguish between different rules, is encoded in neuronal activities during the colour cue period (Figure 3.3B). Thus, the colour rule information is encoded early during the trial (cue period) to produce the correct motor response (saccade) later during the trial.

3.3.4. Latent Space Analysis of Rule Coding

In this task, the colour cue and the target location may be confounded. In order to dissociate the colour cue from the target location/saccade direction, we analyzed trials with the same saccade

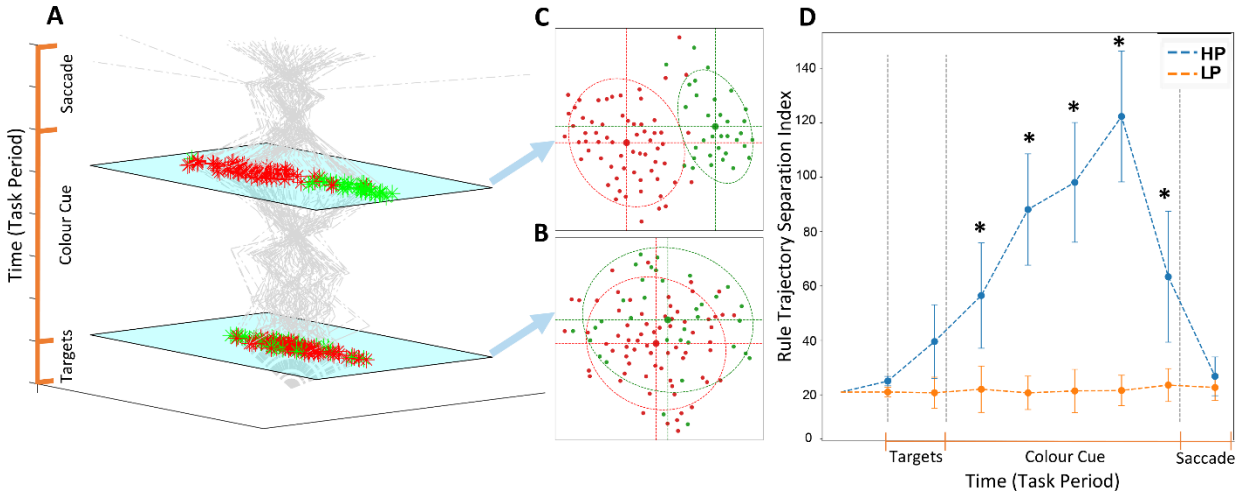

Figure 3.4. Latent Space Analysis of Rule Coding

(A) Two subsets of HP trials with the same target location and different cue colours are transformed into a low dimensional latent space using PCA and t-SNE. The trajectories are plotted on a 3-dimensional temporal-latent space. As we progress through the task periods, the trajectories pass through two separate attractors and move back towards each other as the pre-saccade planning to the same target location begins.

(B) During the target presentation period, the two rules occupy the same cluster in the latent area and the separation of two rules is not obvious.

(C) During the colour cue presentation period, the two rules occupy two different clusters in the latent space.

(D) Analyzing the separation of the two rules through task periods in the latent space using the Calinski-Harabasz index, gives us a separation score that increases during the colour cue presentation to a maximum, and decreases as the monkeys start to plan the saccade. Moreover, the same analysis on the LP trials with the same saccade direction and different cue colours yields no separation between the two rules since the monkey has not acquired the rules and the resulting saccade is the same for all. *p=8.9e-4 for monkey JL, p=1.7e-7 for monkey M comparing the distribution of the separation index of the rules during the colour cue period of HP (blue) versus LP (orange) trials. N=100 for each distribution.

direction but different cue colours. Thus, in this comparison, the monkey saccades in the same direction towards the same target location, but the information prompting the movement arises from different cue colours. Using two dimensionality reduction methods, we transformed the trials into a latent space to investigate the separation of trials with different cue colours in that latent space. We transformed the firing rates of HP trials with the same saccade direction and different cue colours into a 50-dimensional space using principal component analysis (PCA). Subsequently, the principal components are transformed into a 2-dimensional space using the t-distributed stochastic neighbor embedding (t-SNE) [33] method for visualization.

The firing rates obtained in these trials are transformed into a three-dimensional space, two dimensions of the latent space and time (Figure 3.4A). The latent space trajectories through time show the separation of the two colours through different stages of the task. During the target presentation period, before the colour cue is presented, the trajectories of the two different rules occupy the same space with no distinct separation (Figure 3.4B). However, during the colour cue period, the trajectories move toward two separate regions of the latent space, distinguishing between the two rules (Figure 3.4C). The separation of the two rules analyzing multiple pairs of rules with the same saccade direction is quantified by the Calinski-Harabasz index [34] (Figure 3.4D). As the monkeys receive the cue colour information, the separation increases, and as they make a saccade toward the same target, the separation decreases. Importantly, applying the same index on the LP trials shows negligible separation during the stages of the task since the presence of the coded task rules in neuronal activities in these trials is not as strong as in the HP trials ($p=8.9e-4$ for monkey JL, $p=1.7e-7$ for monkey M, Kruskal-Wallis, Dunn post-hoc) (Table 3.1).

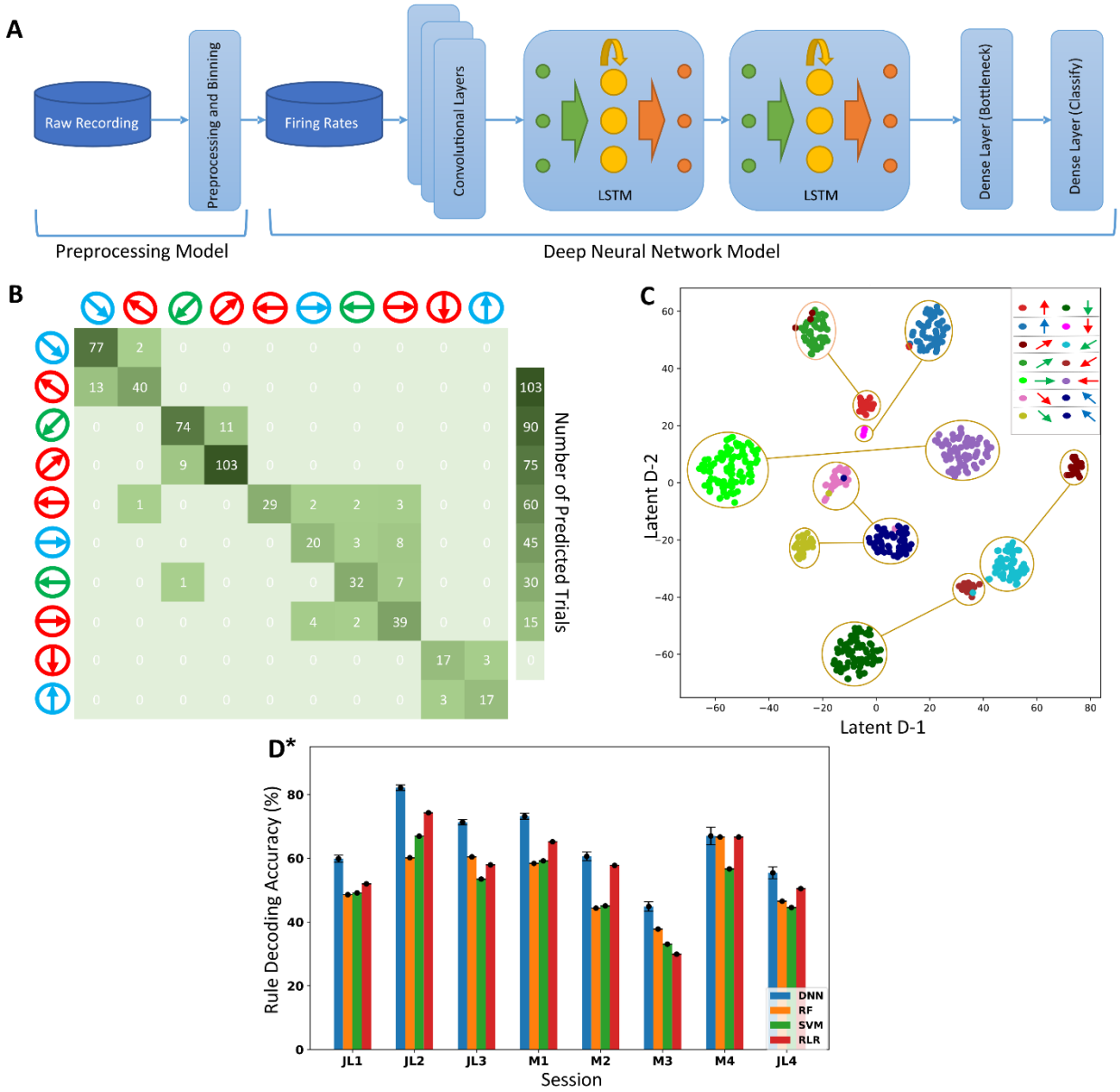

Figure 3.5. Multi-Unit Decoding of the Rules

(A) The processing pipeline and the DNN architecture. The raw recordings are processed, and the calculated firing rates are passed to a DNN. Multiple convolutional layers decode the neuronal activity through multi-units and time. Two LSTM layers process the time-series signals before passing through two dense layers for prediction.

Figure 3.5. Continued.

(B) The confusion matrix for a 10-rule session of monkey JL shows the predicted labels by the DNN model along the x-, and the correct labels along the y-axis.

(C) To visualize the rule classification by DNN in one session, the output of the second LSTM layer of the DNN model in a 13-rule session of monkey JL has been transformed into a 2-dimensional latent space using PCA and t-SNE. Different rules occupy different clusters in this space. The gold loops combine the two opposing targets in each block of trials. The figure legend shows the rule associated with each cluster.

(D) Comparing the performance of DNN, random forest (RF), support vector machine (SVM), and regularized logistic regression (RLR) models, decoding the rules from the population activity of LPFC neurons in all recorded sessions of monkey JL and M. *p=0.02 for monkey JL, p=0.02 for monkey M comparing rule decoding accuracy distribution of the DNN (blue) versus the shuffled label chance level. N=1000 for each distribution.

3.3.5. Decoding of multiple rules

Multiple methods were applied to decode the trial-associated rules. The input firing rates are the threshold-crossing spikes recorded before the response period, counted in 50*ms* time bins. Four models performed better than the rest across all sessions: deep neural network (DNN), random forest (RF), support vector machine (SVM), and regularized logistic regression (RLR). We compared the performance of these four methods decoding the rule from the population activities of LPFC neurons recorded during the task cue period. The DNN (Figure 3.5A) outperformed all other methods across all sessions with maximum decoding accuracy of 80% in a twelve-rule session (chance level decoding of 10.3%) (Figure 3.5D). Across all sessions, the decoding accuracy using the DNN was significantly higher than the chance level (p=0.02 for monkey JL, p=0.02 for monkey M, Kruskal-Wallis, Dunn post-hoc) (Table 3.1). Based on the number of

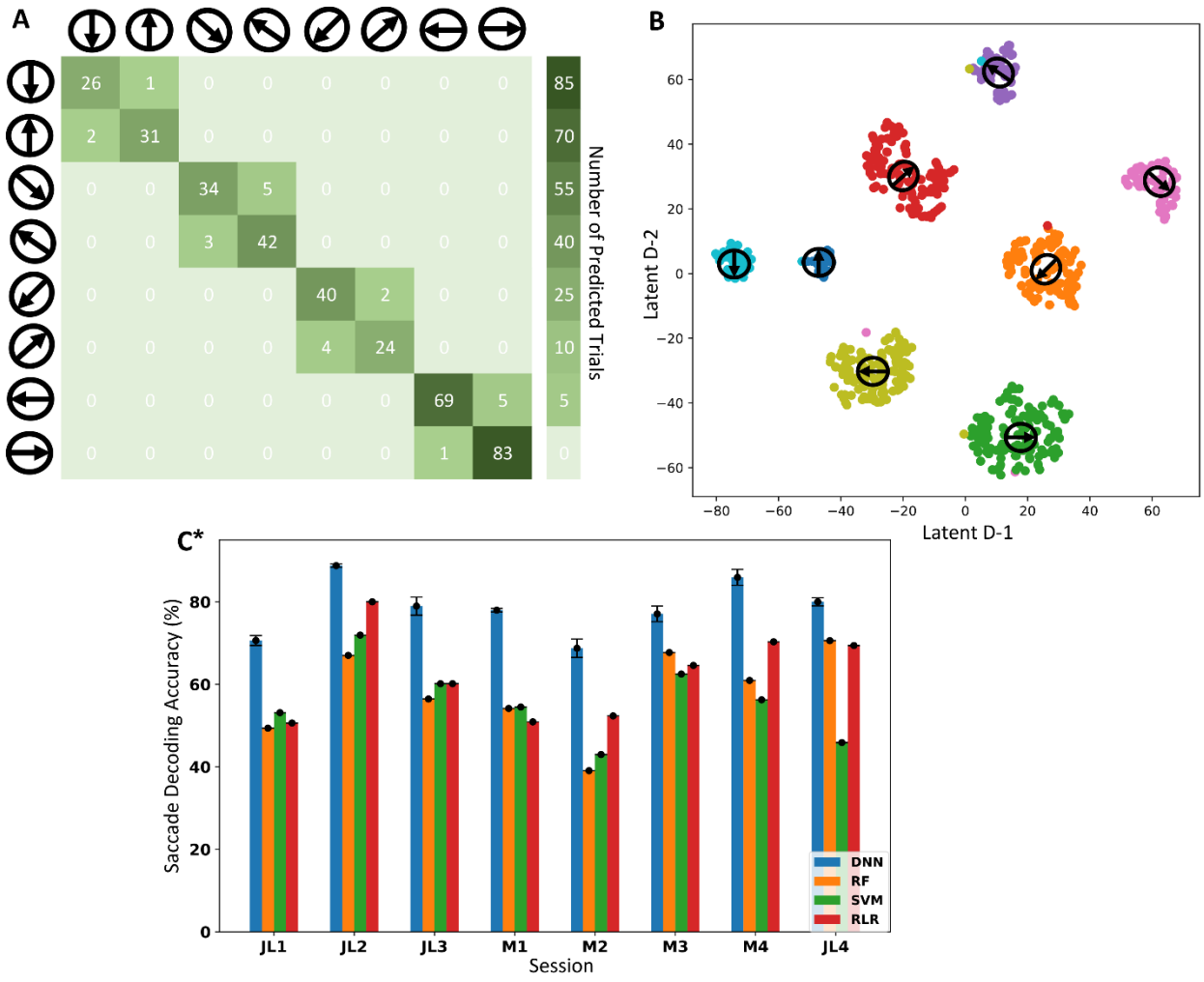

Figure 3.6. Multi-Unit Decoding of Saccades

(A) The confusion matrix for the predictions made by the deep neural network (DNN) model in a session of monkey M. The predicted labels are along the x-axis, and the correct labels are along the y-axis. (B) To visualize the saccade classification by DNN in one session, the output of the second LSTM layer of the DNN model in a session of monkey M is transformed to a 2-dimensional latent space using PCA and t-SNE. Different saccade directions, highlighted by black arrows, occupy the same space in eight separate clusters.

Figure 3.6. Continued.

(C) Comparing the performance of DNN, random forest (RF), support vector machine (SVM), and regularized logistic regression (RLR) models, decoding the intended saccade from the population activity of LPFC neurons in all recorded sessions of monkey JL and M. *p=0.01 for monkey JL, p=0.02 for monkey M comparing the saccade decoding accuracy distribution of the DNN (blue) versus the shuffled label chance level. N=1000 for each distribution.

randomly chosen rules in each session, there are different numbers of categories into which a trial can be classified. The confusion matrix in Figure 3.5B quantifies the decoding performance of a 10-class rule session of monkey JL. In line with the latent space analysis of rule coding, the majority of the rules with the same saccade direction and different cue colours were predicted correctly. Thus, one can distinguish between population activities that arose from multiple different rules with similar or different saccade directions. Importantly, the model decodes information beyond directionally tuned activities corresponding to the different colour rules.

In addition, we transformed the output of the second long short-term memory (LSTM) layer of the DNN, before being passed to the prediction layers, into a 2-dimensional latent space using PCA and t-SNE to visualize the separation of different rules in a 13-class session of monkey JL (Figure 3.5C). There are clear segregations of different colour-saccade direction rules in the latent space. Similar saccade directions corresponding to different colours (e.g., green down left saccade and red down left saccade) are easily discriminable. However, they tend to cluster close to each other in the latent space. The lines link clusters of rules that were run within the same block. There is no apparent segregation of rules as a function of the block, suggesting that the subspace encodes rules based on their similarity and not the time they were learned.

3.3.6. Decoding of Saccades

We tested multiple methods to decode the intended saccade direction from the firing rates recorded during the colour-cued saccade task. We compared the performance of the DNN, RF, SVM, and RLR methods decoding the intended saccade direction from the population activities of LPFC neurons recorded before the saccade (Figure 3.6C). The DNN has a similar architecture to the rule decoding DNN with exceptions in the input and classifier layers (Figure 3.5A). The DNN outperformed the other three methods across all sessions, obtaining significantly higher performance than the chance level ($p=0.01$ for monkey JL, $p=0.02$ for monkey M, Kruskal-Wallis, Dunn post-hoc) (Table 3.1) and maximum average decoding of 90% in both monkeys.

The DNN's performance in one session of monkey M is quantified with a confusion matrix (Figure 3.6A). In addition, we transformed the output of the second LSTM layer to a low-dimensional latent space using the PCA and t-SNE methods to see the clustering of different saccade directions, where ~90% of the trials are categorized correctly in this session of monkey M, occupying separate clusters in the latent space (Figure 3.6B).

3.3.7. Rule Similarity Analyses

The monkeys performed a colour-cued saccade task in which the rule governing cue-target association was changed several times during each experimented session. A new set of colour cues corresponding to two opposing targets governs each trial block. One (or both) of the new rules might be repeated or similar to a previous rule that monkeys learned in the same session. The previously learned rules affect the monkeys' performance in future session trials. If the monkey has learned that the colour red means to saccade toward the up-right target, and if the same rule (red to up-right target) appears in the same session, it will not take long for the monkey to perform above the "rule acquired" threshold. Once the rule has been learned, it will remain in memory for

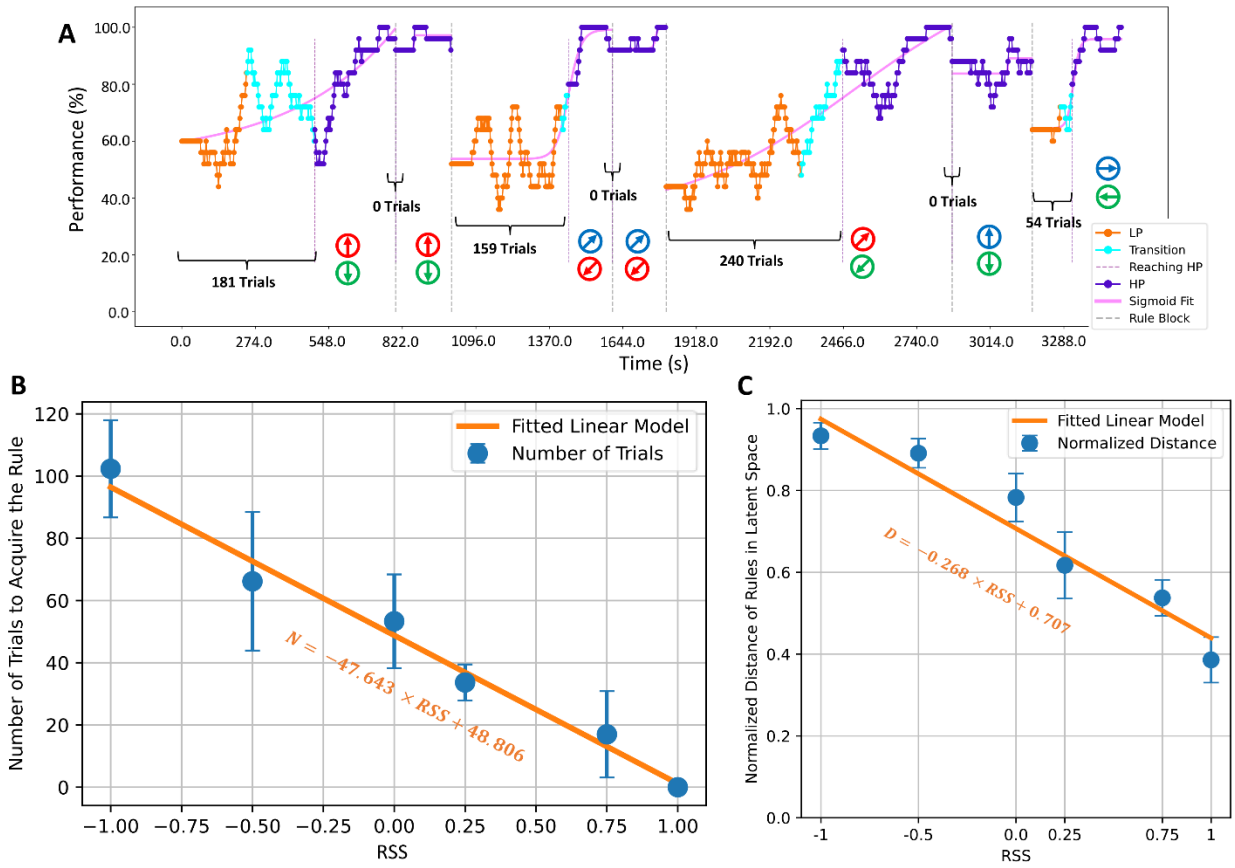

Figure 3.7. Impact of Rule Similarity on Learning

(A) An example session of monkey M performance. Similar to Figure 1D with addition of black brackets indicating the number of trials required for the monkey to acquire the new rules.

(B) We quantified the changes of rules between blocks of trials into a rule similarity score (RSS). The score of 1 means that the monkey has seen a very similar rule recently, like the second block in (A); While the score of -1 means that the monkey has seen an opposing rule recently, like the fifth block in (A). Any other change in rules is scored between -1 and 1. Across all sessions, as the rule similarity increases, the monkey needs fewer number of trials to adjust to the new rules. The orange equation is the fitted linear model to show the decreasing trend. N (Number of trials to acquire the rule) is the y-axis and RSS is the x-axis.

a while to be reused. We quantified this pattern using a rule similarity score (RSS). This score shows whether and how much the current rule resembles the previous rules in the same session. Two extremes are assumed for the RSS, containing other scenarios in between. If one of the two

Figure 3.7. Continued.

(C) The HP trials are transformed into a 100-dimensional latent space and the distance between trials of different rules in different blocks are calculated. By normalizing the distances, we were able to combine the results in different sessions. By analyzing the normalized distances based on their RSS we found a trend in which the blocks with higher RSS are closer (have lower distance value) in the latent space, compared to the blocks with lower RSS. The orange equation is the fitted linear model to show the trend. D is the y-axis and RSS is the x-axis.

current rules is repeated exactly as it is (at least one colour signaling the same target location as before), the RSS equals 1. On the other hand, if one of the two current rules is repeated exactly as opposite (at least one colour signaling the opposite target location as before), the RSS is equal to -1. Other scenarios follow an algorithm (see Methods) that assigns an RSS to the new block of trials based on previous rules learned by the monkey in the same session.

The monkeys' behavioural performance in a session with rules changing in different blocks of trials is shown in Figure 3.7A. Analyzing all sessions, we obtained the RSS for each transitioning block of trials to a new set of rules. Then, counting the number of trials the monkey needs to acquire the new rule, we found that as the RSS increases from -1 to 1, the monkeys need fewer trials to learn the new rule (Figure 3.7B). For example, the monkeys needed zero trials to learn the new rules in the second block in Figure 3.7A, where the RSS is 1 (It is important to note that since the assignment of new rules for the new block is random, sometimes the rules are repeated in two adjacent blocks like block 1 to 2 and 3 to 4). In other words, when the monkeys learn the task rules, they can easily access and utilize them later in the same session. The task rules remain in their memory for at least one session. We also analyzed this trend by fitting a linear model to the

RSS distribution (Figure 3.7B). The slope of the fitted model is -47.643 (confidence interval between -9.57 and -56.52) with the r-squared value of 2.3e-6, which shows a trend in which the number of trials needed by the monkeys to acquire the new rule decreased significantly as the RSS increased.

Furthermore, we transformed the colour cue period of the HP trials into a 100-dimensional latent space using PCA to assess the latent state of blocks of trials based on their RSS. We defined a normalized distance variable between different blocks of trials in all sessions and categorized the calculated distances based on the RSS between blocks (see Methods). We found an inverse correlation between the defined distances and the RSSs where the distances between blocks with higher RSS are lower than those between blocks with lower RSS (Figure 3.7C). We fit a linear model on the distance and RSS distribution, resulting in a slope of -0.268 (confidence interval between -0.301 and -0.010). It is important to note that the slope is calculated with a normalized distance along the y-axis such that $0 \leq D \leq 1$.

3.4. Discussion

This study assessed LPFC coding of associative memory for stimulus-response associations (rules). First, we demonstrate that neuronal populations and single neurons show selectivity for rules once acquired. We then used decoders applied to subsets of the data and found that the rule could be successfully decoded from the neural activity independently of movement parameters. Moreover, we showed that rules are encoded in low dimensional subspaces of neural activity, and the distance between rules in neural space correlates with the similarity between rules assessed behaviourally. The distance in state space between the representations of the rules correlates with the similarity between the rules and behavioural performance.

3.4.1. Encoding of task rules in LPFC

Previous studies have shown that single neurons in LPFC encode stimulus-response associations during an associative learning task. Asaad and coworkers [13] used a similar task to ours and demonstrated that LPFC neurons are selective for a stimulus-response association (object-saccade). In that study, they used one rule combination per session and recorded the activity of one or few neurons at a time. Our study examined the learning of several rules during the same session and recorded from many neurons simultaneously to document the responses of populations of rule-selective neurons. Thus, by analyzing channel recordings, which are representative of subpopulations of neurons within a small cortex region covered by the electrode (~300μm), we found that the activities in some channels change to encode the rule information. As a proxy for rule learning within the region, we found that the number of rule channels increases in HP trials compared to LP trials. The fact that we found mainly learning for the association and not for the saccade direction indicates that what is learned is the abstract mapping of the cue colour into the saccade direction rather than a motor plan for making the saccade to a location. The latter is supported by the fact that we found populations of neurons that can discriminate between two rules with different colours and the same saccade direction.

The coding of task rules (i.e., the number of rule channels) in LP trials is not zero. There are several reasons for this finding. First, in these experimental sessions, the monkeys acquired the rule gradually over several trials. We create thresholds that classify trials as HP, LP, and intermediate to aid analysis. Most of the analyses focused on HP and LP trials. It should be noted that even in the LP trials, the monkeys' performance was typically greater than chance. Thus, the monkeys may have somewhat learned the rules but have not yet reached the performance threshold to categorize some trials as HP. In addition, in many cases, the monkeys learned parts of the rule or

an opposite rule during previous trials. So, it is possible that neuronal activity reflects the similarity between an acquired and a new rule. Indeed, we found a similarity between patterns of neuronal activity that correlate with rule similarity.

3.4.2. Encoding of task rules and saccades during different task periods

Given the activities recorded during different trial periods, we decoded the rule associated with each trial. We found that the rule decoder approaches the maximum performance during the colour cue presentation period. On the other hand, the saccade decoder does not perform as well as the rule decoder until the pre-saccade period. These results can be explained by rule information becoming available during the cue period before saccade information. It is also possible that saccade information becomes more accurate closer to the response period by recruiting neurons that are selective for the location of the intended saccade. Indeed, we have previously shown that decoding spatial attention from LPFC neurons becomes more accurate immediately before a saccade to the attended location is made [35]. Moreover, it is possible to accurately decode the direction of visually guided saccades from populations of neurons with LPFC [19]. It is difficult to determine whether the decoded information during the period before the saccade is related to motor preparation or to spatial attention. However, what seems clear is that the rule and saccade information can be dissociated at the level of the LPFC circuitry (Figure 3.4). The latter point is consistent with the proposal that the LPFC encodes abstract information that is further passed to premotor areas where movement parameters are specified [18], [36], [37].

3.4.3. Separation of rule and behaviour representation

We analyzed HP trials with the same saccade direction to explore the independent coding of the rule from the saccade preparation and movement, following two different colour cues. We showed

that populations of neurons in LPFC can discriminate between two rules that map into a similar motor command suggesting that these two aspects of the rules are orthogonally encoded within the area [36], [37]. This independence of rules attributes is paramount for cognitive flexibility, a main function of the LPFC [38].

The separation between trajectories was negligible in LP trials, likely because the rule was not learned, and information is insufficient to separate the rules clearly. Thus, as monkeys learn the rule, the neural trajectories of different cue colours separate during the cue presentation. This finding indicates that changes in the activity of the population that occurred during the learning of new rules correlate with the behavioural performance of the animals. The latter agrees with studies showing an impaired ability to acquire new rules after lesions of the LPFC in humans and monkeys [8], [9].

3.4.4. LPFC enables flexible storage of associations in long-term memory.

A main finding of our study was the storage of rules in activity subspaces in the LPFC. The analysis of rule similarity summarized by the RSS between revisited rules in one session suggests that the learned information is accessible to the monkey for at least the duration of the session. The persistence of up to 14 different rules is beyond the threshold of what is typically considered a short-term memory [39]. Thus, our results rule out that the animals were using short-term memory alone to solve the task. Furthermore, the latent space analysis of LPFC activity during the cue presentation but plotted as a function of RSS (Figure 3.7C) demonstrates that proximity in PCA space correlates with rule similarity and the ability of the animals to learn the task in a few trials. The strong similarity between the behavioural finding that ease of learning correlates with rule similarity (Figure 3.7B) and the PCA state proximity implies that LPFC plays a critical role in learning the arbitrary stimulus-response associations [13], [18]. Moreover, although we tested a

limited number of rules (up to 14), our results indicate that the LPFC has a storage capacity that allows learning an arbitrary number of stimulus-response associations and new associations from previous experience (rule similarity analyses). We further propose that ensembles of neurons tuned for different rules provide a high dimensional space for flexibly encoding arbitrary numbers of associative experiences, a fundamental aspect of primate intelligence.

3.4.5. Limitations of the study

The main aim of this study was to investigate the neural correlates of associative learning and memory in the LPFC of macaques. Although we achieve animals to learn multiple rules, there are limited trials within a session until animals get satiated. Thus the number of rules we tested is limited. Moreover, the array isolation decays over time, so our study had a time window to collect the data. One must also acknowledge that although microelectrode arrays allow recording from neuronal ensembles of dozens or more neurons, this number is still small compared to the millions of neurons recruited during the task within a few square millimeters of the LPFC. Increasing the number of neurons recorded may produce more accurate results. Using technologies such as Calcium imaging may also improve this issue, although their temporal resolution and practicality are not optimal for primate studies. Such technologies could also improve our ability to reliably identify cell types during the recordings and their specific role in the neural computation underlying ensemble codes.

3.6. STAR METHODS

3.6.1. KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Python 3.8	Python Software Foundation	https://www.python.org/
NumPy 1.19	Community project	https://numpy.org/
SciPy 1.6	Community library project	https://scipy.org/
Tensorflow 2.2	Google Brain Team	https://tensorflow.org/
Keras 2.4	François Chollet	https://keras.io/
MATLAB R2021a	Mathworks	https://www.mathworks.com/

3.6.2. RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources should be directed to the lead contact, Alireza Rouzitalab.

Data and Code Availability

Data reported in this paper will be shared by the lead contact upon request.

Original code used in this paper will be shared by the lead contact upon request.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

3.7. Method Details

3.7.1. Experimental animals

Two adult male monkeys (Macaca fascicularis: monkey “JL” and monkey “M”) were used in this study. All procedures complied with the Canadian Council of Animal Care guidelines and were

approved by the McGill University Animal Care Committee. Animals were pair-housed in enclosures, and interactive environmental stimuli were provided for enrichment. During experimental days, fluid was controlled (minimum of 35 mL kg−1 day−1) in such a manner that the animals could earn any additional amount to the minimum through the successful performance of the task. Fluid intake was supplemented to reach this quantity if it was not achieved during the task, and restriction was lifted during nonexperimental days.

3.7.2. Electrophysiological recordings

Surgical procedures were carried out under general anesthesia with endotracheal intubation. During the initial procedure, the animals were implanted with three titanium head posts – one was placed posterior to the supraorbital ridge in the midline and the other two on the petrosal bones, superior to the occipital protuberance behind each ear. During the experimental tasks and recordings, the head posts interfaced with a head holder to maintain a fixed head position.

In a subsequent procedure, we chronically implanted a 10x10 multi-electrode array (Blackrock Microsystems) in each monkey's left LPFC anterior to the knee of the arcuate sulcus and caudal to the posterior end of the principal sulcus (Brodmann area 8a) (Figure 3.1a). The scalp was incised, and the incision was electrocauterized. The scalp was retracted and the pericranium excised to limit biological reaction around the implant. A high-power drill (Anspach) was used to create a craniotomy over the left LPFC. Wet gelfoam was applied to the epidural edges for hemostasis. The electrode array connector was fixed to the skull with cranial screws, and the array wires were bent in anticipation of the electrode positioning. The array wires were secured with Silastic sealant (silicone polymer, World Precision Instruments). The dura was opened with a no. 11 blade and Reynold scissors. The electrode was placed over the target to lay with the plane of the array parallel to and just touching the cortex with no tension in the wire and inserted 1 mm into

the cortex using a pneumatic inserter (Blackrock Microsystems). The dural flap was laid atop the array, and a synthetic dura was placed over (Durepair, Medtronic, Minneapolis, MN). The bone flap was replaced and secured using cranial fixation plates and screws (Synthes). Gaps in the bone were filled with Silastic and the scalp was released from retraction. A small incision was made in the scalp to make the connector percutaneous. The scalp was closed in layers with buried vicryl suture in the galea and staples to the skin. The animal fully recovered from the surgery within 1 week.

Data from 32 electrodes per session were recorded using a Cerebus Neural Signal Processor (Blackrock Microsystems). The signals were buffered at the head stage (1× amplification; ICS-96), then band-pass filtered (0.3 Hz/1-pole, 7.5 kHz/3-pole, analog) and digitized (16 bit, 1 μV/bit) at 30 kHz. The occurrence of a spike was detected whenever the digitally high-pass filtered (250 Hz/4-pole) raw data passed a threshold (manually adjusted between -4 and -4.5 × noise standard deviation). Spike times and waveforms (48 samples at 30 kHz, 1.6 ms) were saved to disk for later offline analysis.

Raw broadband data was linearly transformed with a sphering matrix to yield independent signal channels. The sphering matrix linearly transforms a vector of random variables with a known covariance matrix into a set of new variables whose covariance is the identity matrix, meaning that they are uncorrelated, and each has variance 1. We then applied a threshold to the resulting signal. To detect events in the recordings, we keep the signals crossing the $-0.5 \times \sigma$ threshold, where σ is the standard deviation of the signal and remained below that threshold for at least 60 μs.

After detecting the threshold-crossing events, we apply a sanitizing method on the event train. We keep only the largest event where they occur within 300 μs of each other. Also, if an event appeared simultaneously on more than 30 % of channels, it was discarded as noise.

We first apply principal component (PC) analysis to the train to sort these events to convert the waveforms into PC features. We then applied isosplit5, a clustering method, to sort the events based on the PC features. Then for each cluster, we re-applied principal component analysis recursively to branch deeper into more clusters up to a maximum depth.

While recording the monkeys' eye movements, the gaze tracker reports a confidence value for each recorded data point. To sanitize this data, we clipped out recorded values with a confidence level of less than 60%.

We detected saccade events using the Naïve Segmented Linear Regression based on Hidden Markov Models (NSLR-HMM) method[40] on our gazing signals.

We segmented processed data into uniform length trials. We sliced each trial between –0.15*s* to +0.2*s* around the Target onset and –0.05*s* to +1.35*s* around the Cue resulting in 1.75 *s* long analysis window. This segmentation intentionally excludes data from 0.05*s* before saccade onset to solely focus on intended saccade decoding during premovement planning. Anecdotally, if peri-saccade data are included then decoding is trivial, with 100% decoding accuracy for most recorded sessions.

To obtain the firing rates, we binned the spike trains in 50ms windows and convolved them with a rectangular kernel.

3.7.3. Behavioural Task

The monkeys performed a modified colour-cued center-out delayed saccade task. During this task a custom computer program controlled the stimulus presentation and monitored eye position signals (SR Research, Ottawa, Canada). Visual stimuli were back-projected on a screen using a video projector (NEC WT610, 1024×768 pixels, 85*Hz*). The screen was positioned 100*cm* from

the animal's eyes. The animal initiated a trial by maintaining fixation on a central fixation point—a square of length 0.2 *deg* located in the middle of the screen—for 700*ms*. Two target boxes of length 5.1 *deg*, symmetric with reference to the fixation point, then appeared in two of eight locations. The potential target locations were the 45 *deg* steps arranged radially 12.7 *deg* away from the middle of the screen. A colour cue in the shape of an annulus of radius 1.3 *deg* then appeared around the fixation point 250*ms* after target onset and remained on the screen for another 1,000*ms*. The cue had three possible colours: red, green, and blue. The colours dictate to which target (top, top-right, right, or bottom-right vs. bottom, bottom-left, left, or top-left), the monkey should make a saccade in order to receive its reward. After the annulus disappeared, the monkey maintained fixation on the central fixation point for another 250*ms* until the fixation point disappeared, when the monkey had to make a saccade to the correct target box (Figure 3.1c). The monkey was required to fixate on the target location for 250*ms* to receive the reward. The rules associated with the target location and the colour cue changed randomly after the monkey learned the rule. For each task event, the stimulus presentation software sent digital signals to both the neurophysiology and eye tracker systems for offline synchronization (Figure 3.1b).

3.7.4. Statistical Analysis

The normality of the distribution of each result was tested by D'Agostino and Pearson's method [41]. Based on this test, the non-parametric or parametric tests were used to confirm the hypothesis. A post-hoc test was carried when we found statistically significant results. Highly significant results are marked with * in the figures. A summary of the tests for different analyses and figures can be found in Table 3.1.

Table 3.1. Summary of Statistical Analyses

Figure	Normal Distribution Test	Hypothesis Test	Post-Hoc Test	Monkey JL				Monkey M				N = (DoF+2)/2
				p	H/T	F	DoF	p	H/T	F	DoF	
2C	Not Gaussian	Kruskal-Wallis	Dunn	4.3e-11	43.45	479.7	122	n/a	n/a	n/a	n/a	Number of single units
3B	Not Gaussian	Kruskal-Wallis	Dunn	0.019	5.46	9.95	6	0.029	4.74	7.50	6	Number of sessions
4D	Not Gaussian	Kruskal-Wallis	Dunn	8.9e-4	11.04	32.08	38	1.7e-7	27.27	195.74	38	Number of trials
5D	Not Gaussian	Kruskal-Wallis	Dunn	0.02	5.33	9.48	6	0.02	5.39	9.71	6	Number of sessions
6C	Not Gaussian	Kruskal-Wallis	Dunn	0.01	5.47	9.96	6	0.02	5.46	9.94	6	Number of sessions

3.7.5. Generalized additive model

A generalized additive model (GAM) is a technique to simultaneously estimate the relationship between a response variable and multiple predictor variables. Unlike generalized linear models, GAMs can account for arbitrary nonlinear relationships between predictors and the response. Since the relationship between stimulus variables and spikes is often nonlinear and nonmonotonic, GAMs are an attractive option for modeling neural responses. In neuroscience experiments, the neural response may be influenced by external inputs – which may either be continuous-valued signals (e.g., velocity, position) or discrete (e.g., target onset, colour cue, reward) – as well as neuron's response in the past (spike history). In its most general form, the model can be written as: $g\big(E(y)\big) = \sum_{i=1}^{N_C} f_i(x_i) + \sum_{k=1}^{N_D} (u_k * z_k) + h * y$, where x_i is the *i*-th continuous-valued input variable, $f_i(.)$ is any generic nonlinear function operation on x_i, z_k is the *k*-th discrete value, u_k is the temporal filter operating on z_k, h is the spike-history kernel, $g(.)$ is the link function, $E(.)$ is the expectation value of the response, and N_C, N_D denote the total number of continuous- or discrete-valued inputs respectively.

Given the response y and inputs x_i, z_k, the goal is to recover f_i, u_k, and h under some assumed link function g. This can be solved by computing the maximum a posteriori estimates as $\{\hat{f}, \hat{u}, \hat{h}\} =$

$argmax(L(f,u,h\,|y,x)\ P(f,u,h))$, where $L(f,u,h\,|y,x) = P(y|x;f,u,h)$ is the model likelihood and $P(f,u,h)$ is the prior over the model parameters [30].

Analyzing the changing rules, saccades, and colours in our task, given the single channel neuronal spiking activity, shows which inputs can explain the activities better. The inputs with the highest likelihood, which explain the highest fraction of the spike occurrence, are chosen as the best model by the algorithm [42].

3.7.6. Decoding methods

We compared the performance of our deep RNN-based model against multiple popular approaches using scikit-learn [43]. The first method was regularized logistic regression [44], [45]. For time-series signals, we may have many features relative to the number of training examples which could lead to overfitting. We used L2 regularization to help mitigate overfitting.

The second method that we tested is the support vector machine (SVM) algorithm. SVM translates the input data into a higher dimension space and then tries to find the hyperplanes that separate the data to multiple classes with the least possible error rate [46]–[48]. Furthermore, by applying the appropriate kernel to SVM, we can turn a linear model into a non-linear model to better fit our data. In the SVM model, we used radial basis function (RBF) as the kernel with the gamma coefficient of $1/(n_{features} \times Variance(X_{input}))$. Here, the $n_{features}$ is the number of time samples in each trial.

The third approach tested is the random forest method. Random forests comprise multiple relatively uncorrelated trees that predict the desired outcome for inputs based on various features and, ultimately, vote for the most popular result [49]. This method has proved to perform well due

to the ensemble nature of the approach, which takes multiple aspects of input data into account [50]–[52].

We trained, evaluated, and tested all methods in a 10-fold scheme, recording the validation accuracy as the model's performance.

3.7.7. Deep neural network parameters

Model definition, training, and evaluation were implemented in Tensorflow (2.2) using the Keras API. To read the data for our model, we start with a sub-model inspired by EEGNet [53]. The data are spatially filtered and then temporally filtered using a sequence of convolutional layers, yielding a lower-dimensional transformation of the input data. The output from this sub-model is then fed to a recurrent neural network (RNN). RNNs are frequently used to decode time-series signals, and we tested three different RNNs: simple RNN, gated recurrent unit (GRU) [54], and long short-term memory (LSTM) [55]. GRU and LSTM yielded the same results for our purposes, and we chose to work with LSTMs. The LSTM outputs were fed to two sequential non-linear layers to produce a prediction of the intended saccade target (Figure 3.5a).

The model weights were trained using backpropagation with the NAdam optimizer to minimize the loss. The cost function is defined to calculate the multi-class loss between model output and signal labels. To this end, we used the cross-entropy loss: $CELossFunction\ (Y,S) = -\sum_i^c y_i \log s_i$, where C is the number of classes, y_i is the ground truth for the *i*-th class and s_i is the model's score for the *i*-th class.

The LSTM layer in our network uses L2 regularization for both input and recurrent kernels. This method adds the square magnitude of coefficients to the cost function to avoid overfitting. The ridge regression changes the cost function into $L = CELossFunction(Y,S) + \lambda\sum_{i=1}^{N} |w_i|$, where

Y is the true label, S the score vector for classes, $CELossFunction(Y,S)$ the original cross-entropy loss function, λ the L2 regularization weight and w_i the i-th coefficient.

To avoid both overfitting and overconfidence in our deep network, we used another regularization technique called label smoothing [56]. This technique mixes the labels from binary values of one-hot encoded vectors with a uniform distribution to make the gap between correct and incorrect classes less obvious: $Y_{LS} = (1-\alpha) \times Y_{one-hot} + \alpha/C$, where α determines the amount of smoothing.

Theoretically, there is not much limit on how deep a multi-layer network can get. Successful models have many different numbers of layers, usually higher than three [57]–[60]. We tested multiple architectures different in the number of convolutional and RNN layers. This is an additional step towards generalizing our approach, which differs from the hyperparameter optimization that usually occurs when the model architecture is decided. After choosing the best-performing architecture, several methods were used to ensure the network did not produce unreal and non-deterministic results. The chance level was calculated using the label shuffling technique. The model is trained and validated using a 10-fold cross-validation scheme. During each run, we save the model with the highest validation accuracy. Our model hyperparameters were the number of units in the RNN layer, the learning rate, the L2 regularization weight, the label smoothing ratio, the number of RNN layers, and the number of units in the fully connected layers. We optimized the hyperparameters using a hold-out data set. The final hyperparameters are listed in Table 3.2.

Deep neural networks tend to perform differently with different weight initializations. Due to the process of loss minimization and the search for the global minima in the loss landscape, it is important where we start. We always initialize the random generators to a fixed seed in our program to avoid this. We ran the training and evaluation process for 100 run to verify whether

our model performance was deterministic. If the results do not show a large variance, the initialization is not playing a major role in the model's performance. The decoding results showed a maximum of ~5% variance.

Table 3.2. Optimized Hyperparameters

Hyperparameter	Optimized Value
Number of Units in LSTM layers	64
Convolution Kernel Length	25
Learning Rate	0.001
Dropout Rate	0.4
L2 Regularization Weight	0.001
Label Smoothing Ratio	0.2
Number of Units in the Dense layer (Bottleneck)	64

3.7.8. Rule similarity score algorithm

The following pseudo-code shows the principle governing the assignment of the rule similarity score to a new block of trials:

```
Target Group 1 = {Right, Up Right, Up, Up Left};
Target Group 2 = {Down Right, Down, Down Left, Left};
N = number of blocks in current session;
RSS = 0;

for colour in block colours:
        if colour is repeated:
                if repetition occurs more than N/2 blocks before:
                        r = 2;
                else:
                        r = 1;
                if target is 0 deg apart: #same target
                        RSS = 1;
                        break;
                else if target is 180 deg apart: #opposite target
                        RSS = -1;
                        break;
                else if target is 45 deg and in same group:
                        RSS += 0.5/r;
                else if target is 45 deg apart and in other group:
                        RSS -= 0.5/r;
                else:
                        RSS += 0;
```

3.7.9. Latent space distance

In each session, the dimensions of the colour cue period of the learned trials are reduced to 100 points using the PCA method. First, 20 principal components are extracted through time, then 5 principal components are extracted from different channels, i.e., space. Thus, each trial is represented by 100 principal components to capture all the necessary dynamics of the cue period.

As each block of trials consists of two rules, we defined the following variable as the distance between trials of two blocks:

$D = \left\| \frac{{A_1}^T A_1}{N_{A_1}} - \frac{{B_1}^T B_1}{N_{B_1}} \right\| + \left\| \frac{{A_2}^T A_2}{N_{A_2}} - \frac{{B_2}^T B_2}{N_{B_2}} \right\|$, where A_1 and B_1 are the transformed trials with the same colour cue in the two blocks (since we only have 3 different colour cues, at least one colour cue is repeated between blocks), A_2 and B_2 are the transformed trials with the second rule in the two blocks (with the same colour cue or different) with number trials noted with N_{A_1}, N_{B_1}, N_{A_2}, and N_{B_2} respectively, and $\|.\|$ is the matrix norm.

The trial matrices of A and B have dimensions of $N_A \times 100$ and $N_B \times 100$, which makes $A^T A$ and $B^T B$ 100×100 matrices. Also, to eliminate the effect of the number of trials with each rule, the $A^T A$ and $B^T B$ matrices are divided by N_A and N_B respectively.

In each session, the normalized distances are divided by the maximum value of calculated D to have values between 0 and 1. Then, the distances are categorized based on the RSS between the two blocks and pooled across all sessions to analyze the correlation between the distances and the RSSs.

3.8. References

[1] C. Heyes, "Simple minds: a qualified defence of associative learning," *Philos Trans R Soc Lond B Biol Sci*, vol. 367, no. 1603, pp. 2695–2703, Oct. 2012, doi: 10.1098/rstb.2012.0217.

[2] J. M. Barnes and B. J. Underwood, "'Fate' of first-list associations in transfer theory," *Journal of Experimental Psychology*, vol. 58, no. 2, pp. 97–105, 1959, doi: 10.1037/h0047507.

[3] D. Hume, *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning Into Moral Subjects; and Dialogues Concerning Natural Religion*. Oxford University Press, London, 1739.

[4] K. Koffka, *Principles of Gestalt Psychology*. Psychology Press, 1935.

[5] J. B. Smaers, A. Gómez-Robles, A. N. Parks, and C. C. Sherwood, "Exceptional Evolutionary Expansion of Prefrontal Cortex in Great Apes and Humans," *Current Biology*, vol. 27, no. 5, pp. 714–720, Mar. 2017, doi: 10.1016/j.cub.2017.01.020.

[6] E. R. Guthrie, *Psychology of Learning*. Oxford, England: Harper, 1935.

[7] E. L. Thorndike, "Animal intelligence: An experimental study of the associative processes in animals," *The Psychological Review: Monograph Supplements*, vol. 2, no. 4, pp. i–109, 1898, doi: 10.1037/h0092987.

[8] B. Milner, "Effects of Different Brain Lesions on Card Sorting: The Role of the Frontal Lobes," *Arch Neurol*, vol. 9, no. 1, pp. 90–100, Jul. 1963, doi: 10.1001/archneur.1963.00460070100010.

[9] T. L. Moore, S. P. Schettler, R. J. Killiany, D. L. Rosene, and M. B. Moss, "Effects on executive function following damage to the prefrontal cortex in the rhesus monkey (Macaca mulatta)," *Behav Neurosci*, vol. 123, no. 2, pp. 231–241, Apr. 2009, doi: 10.1037/a0014723.

[10] J. D. Wallis and E. K. Miller, "From Rule to Response: Neuronal Processes in the Premotor and Prefrontal Cortex," *Journal of Neurophysiology*, vol. 90, no. 3, pp. 1790–1806, Sep. 2003, doi: 10.1152/jn.00086.2003.

[11] R. A. Andersen, S. Kellis, C. Klaes, and T. Aflalo, "Toward more versatile and intuitive cortical brain-machine interfaces," *Curr. Biol.*, vol. 24, no. 18, pp. R885–R897, Sep. 2014, doi: 10.1016/j.cub.2014.07.068.

[12] S. Musallam, B. D. Corneil, B. Greger, H. Scherberger, and R. A. Andersen, "Cognitive Control Signals for Neural Prosthetics," *Science*, vol. 305, no. 5681, pp. 258–262, Jul. 2004, doi: 10.1126/science.1097938.

[13] W. F. Asaad, G. Rainer, and E. K. Miller, "Neural Activity in the Primate Prefrontal Cortex during Associative Learning," *Neuron*, vol. 21, no. 6, pp. 1399–1407, Dec. 1998, doi: 10.1016/S0896-6273(00)80658-3.

[14] J. Fuster, *The Prefrontal Cortex Anatomy, Physiology and Neuropsychology of the Frontal Lobe*. 1997.

[15] J. M. Fuster, "Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory.," *Journal of Neurophysiology*, vol. 36, no. 1, pp. 61–78, Jan. 1973, doi: 10.1152/jn.1973.36.1.61.

[16] G. Rainer, S. C. Rao, and E. K. Miller, "Prospective Coding for Objects in Primate Prefrontal Cortex," *J. Neurosci.*, vol. 19, no. 13, pp. 5493–5505, Jul. 1999, doi: 10.1523/JNEUROSCI.19-13-05493.1999.

[17] S. C. Rao, "Integration of What and Where in the Primate Prefrontal Cortex," *Science*, vol. 276, no. 5313, pp. 821–824, May 1997, doi: 10.1126/science.276.5313.821.

[18] J. D. Wallis, K. C. Anderson, and E. K. Miller, "Single neurons in prefrontal cortex encode abstract rules," *Nature*, vol. 411, no. 6840, pp. 953–956, Jun. 2001, doi: 10.1038/35082081.

[19] C. B. Boulay, F. Pieper, M. Leavitt, J. Martinez-Trujillo, and A. J. Sachs, "Single-trial decoding of intended eye movement goals from lateral prefrontal cortex neural ensembles," *J Neurophysiol*, vol. 115, pp. 486–499, 2016, doi: 10.1152/jn.00788.2015.-Neu.

[20] K. R. Bullock, F. Pieper, A. J. Sachs, and J. C. Martinez-Trujillo, "Visual and presaccadic activity in area 8Ar of the macaque monkey lateral prefrontal cortex," *Journal of Neurophysiology*, vol. 118, no. 1, pp. 15–28, Jul. 2017, doi: 10.1152/jn.00278.2016.

[21] D. J. Freedman, M. Riesenhuber, T. Poggio, and E. K. Miller, "Categorical Representation of Visual Stimuli in the Primate Prefrontal Cortex," *Science*, vol. 291, no. 5502, pp. 312–316, Jan. 2001, doi: 10.1126/science.291.5502.312.

[22] R. F. Squire, B. Noudoost, R. J. Schafer, and T. Moore, "Prefrontal contributions to visual selective attention," *Annual Review of Neuroscience*, vol. 36, pp. 451–466, 2013, doi: 10.1146/annurev-neuro-062111-150439.

[23] D. J. Freedman and J. A. Assad, "Experience-dependent representation of visual categories in parietal cortex," *Nature*, vol. 443, no. 7107, pp. 85–88, Sep. 2006, doi: 10.1038/nature05078.

[24] P. Schwedhelm, D. Baldauf, and S. Treue, "The lateral prefrontal cortex of primates encodes stimulus colors and their behavioral relevance during a match-to-sample task," *Scientific Reports*, vol. 10, no. 1, Art. no. 1, Mar. 2020, doi: 10.1038/s41598-020-61171-3.

[25] R. A. Gulli *et al.*, "Context-dependent representations of objects and space in the primate hippocampus during virtual navigation," *Nat Neurosci*, vol. 23, no. 1, pp. 103–112, Jan. 2020, doi: 10.1038/s41593-019-0548-3.

[26] L. J. Toth and J. A. Assad, "Dynamic coding of behaviourally relevant stimuli in parietal cortex," *Nature*, vol. 415, no. 6868, Art. no. 6868, Jan. 2002, doi: 10.1038/415165a.

[27] T. Backen, S. Treue, and J. C. Martinez-Trujillo, "Encoding of Spatial Attention by Primate Prefrontal Cortex Neuronal Ensembles," *eNeuro*, vol. 5, no. 1, p. ENEURO.0372-16.2017, Mar. 2018, doi: 10.1523/ENEURO.0372-16.2017.

[28] S. Tremblay, F. Pieper, A. Sachs, and J. Martinez-Trujillo, "Attentional Filtering of Visual Information by Neuronal Ensembles in the Primate Lateral Prefrontal Cortex," *Neuron*, vol. 85, no. 1, pp. 202–215, Jan. 2015, doi: 10.1016/j.neuron.2014.11.021.

[29] S. P. McKee, S. A. Klein, and D. Y. Teller, "Statistical properties of forced-choice psychometric functions: Implications of probit analysis," *Perception & Psychophysics*, vol. 37, no. 4, pp. 286–298, Jul. 1985, doi: 10.3758/BF03211350.

[30] T. Hastie and R. Tibshirani, "Generalized Additive Models," *Statistical Science*, vol. 1, no. 3, pp. 297–310, Aug. 1986, doi: 10.1214/ss/1177013604.

[31] W. H. Kruskal and W. A. Wallis, "Use of Ranks in One-Criterion Variance Analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, 1952, doi: 10.2307/2280779.

[32] O. J. Dunn, "Multiple Comparisons Using Rank Sums," *Technometrics*, vol. 6, no. 3, pp. 241–252, Aug. 1964, doi: 10.1080/00401706.1964.10490181.

[33] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008.

[34] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, Jan. 1974, doi: 10.1080/03610927408827101.

[35] S. Tremblay, G. Doucet, F. Pieper, A. Sachs, and J. Martinez-Trujillo, "Single-Trial Decoding of Visual Attention from Local Field Potentials in the Primate Lateral Prefrontal Cortex Is Frequency-Dependent," *J. Neurosci.*, vol. 35, no. 24, pp. 9038–9049, Jun. 2015, doi: 10.1523/JNEUROSCI.1041-15.2015.

[36] N. Ramnani, I. Toni, O. Josephs, J. Ashburner, and R. E. Passingham, "Learning- and Expectation-Related Changes in the Human Brain During Motor Learning," *Journal of Neurophysiology*, vol. 84, no. 6, pp. 3026–3035, Dec. 2000, doi: 10.1152/jn.2000.84.6.3026.

[37] J. B. Rowe, I. Toni, O. Josephs, R. S. J. Frackowiak, and R. E. Passingham, "The Prefrontal Cortex: Response Selection or Maintenance Within Working Memory?," *Science*, vol. 288, no. 5471, pp. 1656–1660, Jun. 2000, doi: 10.1126/science.288.5471.1656.

[38] N. P. Rougier, D. C. Noelle, T. S. Braver, J. D. Cohen, and R. C. O'Reilly, "Prefrontal cortex and flexible cognitive control: Rules without symbols," *Proceedings of the National Academy of Sciences*, vol. 102, no. 20, pp. 7338–7343, May 2005, doi: 10.1073/pnas.0502455102.

[39] P. M. Bays and M. Husain, "Dynamic Shifts of Limited Working Memory Resources in Human Vision," *Science*, vol. 321, no. 5890, pp. 851–854, Aug. 2008, doi: 10.1126/science.1158023.

[40] J. Pekkanen and O. Lappi, "A new and general approach to signal denoising and eye movement classification based on segmented linear regression," *Scientific Reports*, vol. 7, no. 1, Art. no. 1, Dec. 2017, doi: 10.1038/s41598-017-17983-x.

[41] R. D'Agostino and E. S. Pearson, "Tests for Departure from Normality. Empirical Results for the Distributions of b2 and √ b1," *Biometrika*, vol. 60, no. 3, pp. 613–622, 1973, doi: 10.2307/2335012.

[42] K. Lakshminarasimhan, "neuroGAM." Apr. 30, 2021. Accessed: Oct. 04, 2021. [Online]. Available: https://github.com/kaushik-l/neuroGAM

[43] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[44] S.-I. Lee, "Efficient L1 Regularized Logistic Regression," *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)*, p. 8, 2006.

[45] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," *J Stat Softw*, vol. 33, no. 1, pp. 1–22, 2010.

[46] C. Cortes and V. Vapnik, "Support-vector networks," *Mach Learn*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[47] J. A. K. Suykens and J. Vandewalle, "Least Squares Support Vector Machine Classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, Jun. 1999, doi: 10.1023/A:1018628609742.

[48] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer Science & Business Media, 2013.

[49] Tin Kam Ho, "Random decision forests," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Aug. 1995, vol. 1, pp. 278–282 vol.1. doi: 10.1109/ICDAR.1995.598994.

[50] C. Nguyen, Y. Wang, and H. N. Nguyen, "Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic," *JBiSE*, vol. 06, no. 05, pp. 551–560, 2013, doi: 10.4236/jbise.2013.65070.

[51] M. Pal, "Random forest classifier for remote sensing classification," *International Journal of Remote Sensing*, vol. 26, no. 1, pp. 217–222, Jan. 2005, doi: 10.1080/01431160412331269698.

[52] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 93–104, Jan. 2012, doi: 10.1016/j.isprsjprs.2011.11.002.

[53] M. Hassan, M. Shamas, M. Khalil, W. El Falou, and F. Wendling, "EEGNET: An Open Source Tool for Analyzing and Visualizing M/EEG Connectome," *PLoS ONE*, vol. 10, no. 9, p. e0138297, Sep. 2015, doi: 10.1371/journal.pone.0138297.

[54] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches," *arXiv:1409.1259 [cs, stat]*, Oct. 2014, Accessed: Aug. 31, 2020. [Online]. Available: http://arxiv.org/abs/1409.1259

[55] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[56] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *arXiv:1512.00567 [cs]*, Dec. 2015, Accessed: Sep. 30, 2021. [Online]. Available: http://arxiv.org/abs/1512.00567

[57] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv:1704.04861 [cs]*, Apr. 2017, Accessed: Sep. 24, 2020. [Online]. Available: http://arxiv.org/abs/1704.04861

[58] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *arXiv:1703.06870 [cs]*, Jan. 2018, Accessed: Sep. 24, 2020. [Online]. Available: http://arxiv.org/abs/1703.06870

[59] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces," *J. Neural Eng.*, vol. 15, no. 5, p. 056013, Oct. 2018, doi: 10.1088/1741-2552/aace8c.

[60] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *arXiv:1612.08242 [cs]*, Dec. 2016, Accessed: Sep. 24, 2020. [Online]. Available: http://arxiv.org/abs/1612.08242

Chapter 4:
Volitional Control of Beta Activities in Parkinson's Disease Patients

4.1. Abstract

Patients diagnosed with Parkinson's disease (PD) have difficulty initiating and executing movements due to an acquired imbalance of the basal ganglia thalamocortical circuit secondary to loss of dopaminergic input into the striatum. The unbalanced circuit is hyper-synchronized, presenting as larger and longer bursts of beta-band (13-30 Hz) oscillations in the subthalamic nucleus (STN). As a first step toward a novel PD therapy to improve symptoms through beta desynchronization, we sought to determine if individuals with PD could acquire volitional control of STN beta power in a neurofeedback task. We found a significant difference in STN beta power between task conditions, and relevant brain signal features could be detected and decoded in real-time. This demonstration of volitional control of STN beta motivates the development of a neurofeedback therapy to modulate PD symptom severity.

4.2. Introduction

The basal ganglia are brain regions clustered deep inside the brain, and they are important for motor control [1], especially in goal-oriented movements [2], and in gating movement onset [3].

In PD, the loss of dopaminergic input to the striatum unbalances the basal ganglia–thalamocortical circuit. Subsequently, the unbalanced circuit results in pathological movement inhibition. Microelectrode recordings from the STN of the basal ganglia are abnormal in PD, and the magnitude of abnormality correlates with motor impairment severity [4]–[6]. The pathological state is characterized by abnormally large amplitude and long-duration bursts of the beta band (13–30 Hz) synchronized activity in the STN [7] and the motor cortex [8]. Beta activity in PD

desynchronizes in response to L-Dopa [9], or deep brain stimulation (DBS) therapy and the desynchronization magnitude is proportional to symptom improvement [10]. Furthermore, beta-triggered DBS provides equivalent therapy to always-on DBS with fewer side effects [11].

Given the association between STN beta-band synchronization and PD symptom severity, we propose that targeted STN beta-band desynchronization should improve PD symptoms, and this could be a novel therapy for PD patients. However, it is unknown if pathological STN beta can be modulated volitionally.

Real-time neurofeedback of brain signals enables participants to modulate their brain activity over weeks and months with concomitant changes in ADHD severity [12], [13] or epilepsy seizure frequency [14], [15], and over seconds as a form of alternative communication in a BCI [16]. Moreover, studies have shown that BCI-based volitional control of motor cortical beta oscillations affects reaction time in healthy individuals [17]. In addition, others have demonstrated volitional control of motor cortical beta oscillations in PD patients [18]. In this study, we sought to demonstrate the rapid acquisition of STN beta control in PD patients.

Table 4.1. Information on recorded sessions, and patients' PD conditions

Patient/Session	Number of Trials		Age	Sex	UPDRS 3	Improvement on Meds (%)	Calculated Beta Band
	Orange	Blue					
P008	13	15	59	M	48	67	15-30 Hz
P009	14	14	47	F	31	74	15-40 Hz
P010	10	13	57	M	40	45	25-40 Hz
P014	33	24	63	F	56	63	20-35 Hz
P015	9	14	69	M	N/A	N/A	25-45 Hz
P017	9	11	54	M	44	43	20-40 Hz
P018	8	14	73	M	45	53	20-40 Hz
P019	13	13	67	F	51	65	25-45 Hz
P020	13	13	50	F	48	85	15-30 Hz

We tested the ability of nine PD patients undergoing DBS implantation surgery (Table 4.1) to acquire volitional control of STN beta-band oscillations. The post-hoc reconstructed recording sites used to drive the feedback (Figure 4.1A, red dots) were mostly in or around the motor part of the STN (Figure 4.1A, orange part of nucleus). The other recorded sites not used to drive the feedback (Figure 4.1A, yellow dots) were also in the STN. Significant modulation of beta power was not observed in the other two recorded channels that were not used to drive feedback. Participants observed a sphere in a virtual reality scene, and the colour of the sphere was controlled by a transformation of the instantaneous power of ongoing beta oscillations recorded from a microelectrode in the STN (Figure 4.1B). At the beginning of each trial, participants were cued to change the colour of the ball toward either orange or blue. Participants were instructed to make the colour orange by imagining stiff and slow Parkinsonian-like movements and to change it to blue by imagining smooth and easy movements. Still, they were free to explore and use whatever mental strategy worked best at controlling the ball's color.

The experiment aims to determine whether individual participants can acquire volitional control of their brain signals, discriminable between target classes.

4.3. Results

4.3.1. Change in Beta Power

The beta band wave is acquired by bandpass filtering the recorded LFP between 13-30 Hz. Figure 4.1C illustrates one of the beta waves recorded during one session (solid line). The beta wave amplitude envelope (dashed line) was acquired by applying a Hilbert transform [26] on the beta wave time signal to encapsulate the beta bursts. Moreover, analyzing the LFP in the frequency domain provides valuable information for this study. Periodograms provide the signal's

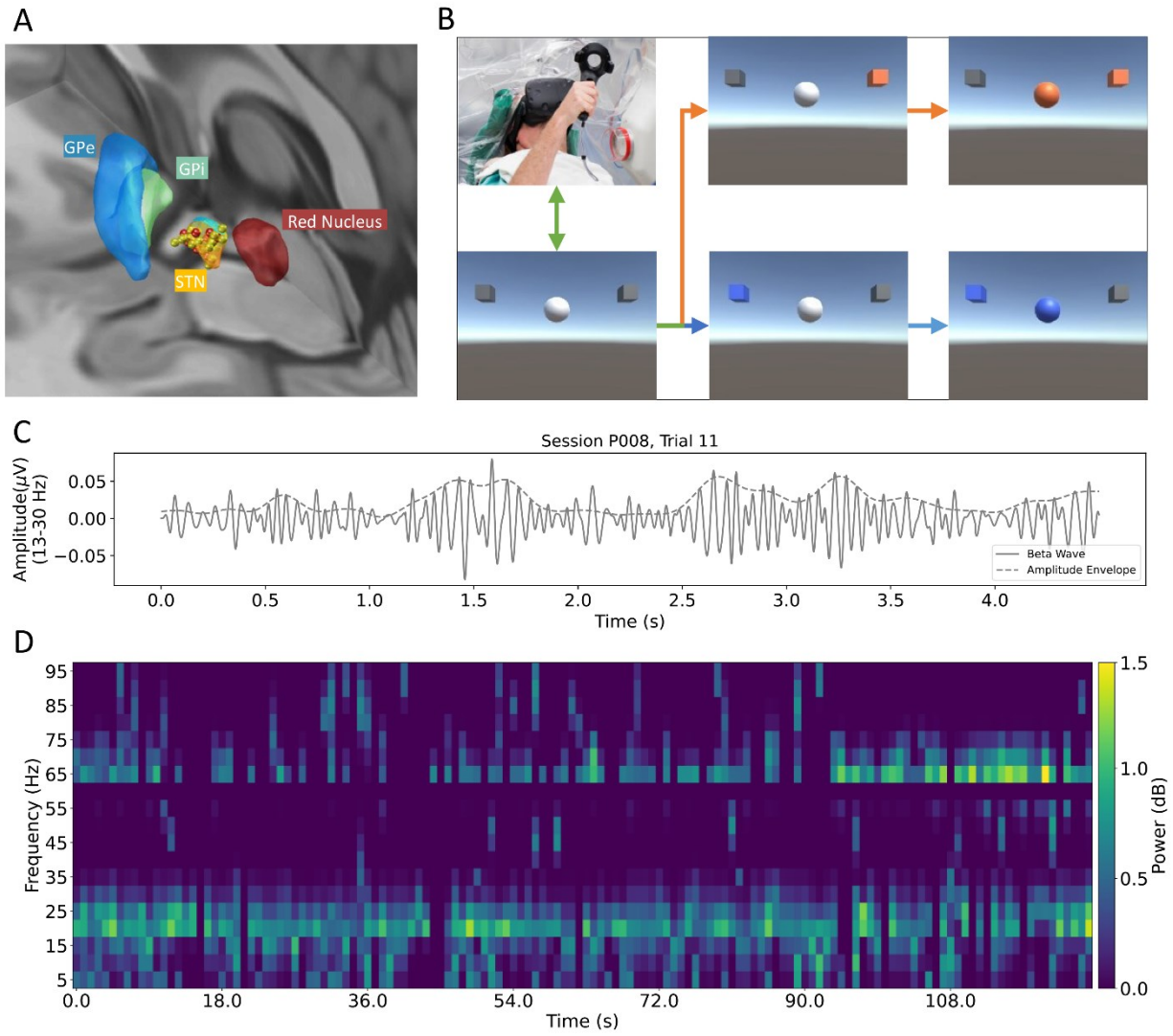

Figure. 4.1. The procedure to record modulated LFPs in PD patients.

(A) Microelectrode recording sites from nine participants reconstructed from surgical notes and co-registration of postoperative CT with preoperative MRI then transformed to template space. The recording sites used to drive the feedback are indicated in red, and recorded sites that were note used for feedback are indicated in yellow. The anatomical structures from an atlas are the pallidum (external, blue; internal, green), the red nucleus (red), and the STN (motor, orange; associative, blue; limbic is not visible)

power in different frequencies through time. Investigating the periodogram of each session, the frequency band with more varying power through time may yield better information on the

Fig 4.1. Continued.

(B) During the DBS surgery, the patient donned a VR headset and performed a task to change the color of a central sphere. One of the two targets eccentric to the sphere, changed colors indicated the target color. The patient then imagined one of two opposing scenarios to change the color of the sphere to match the cued target color. (* The patient's image is taken during another task which required the VR controller).

(C) LFP recorded during one sample trial is filtered to the beta band (solid line). The beta waves are enveloped using the Hilbert transform (dashed line)

(D) The spectrogram of a session showing the changing LFP power in different frequency bands. The beta band (13-30 Hz) shows significant change in modulation through time.

changing condition across trials. Figure 4.1D illustrates the periodogram for one session where the power values are persistently high and vary in the beta band. It is important to note that the LFP powers are filtered at 60 Hz to avoid the powerline noise affecting the results. The power spectra of the filtered LFP are calculated using the Welch method [19]. This method estimates the power spectral density (PSD) of time signals by dividing the signal into successive blocks, forming the periodogram for each block, and then averaging. Separating the PSD into different trials, we sorted the trials into two categories based on the colour of the cue presented to the patients. The number of trials for each condition is not always equal and varies across sessions (Table 4.1). To compare the PSD levels based on the trials' conditions, we averaged the trials of the two conditions to obtain two separate PSD distributions for the orange and blue trials. Across all sessions, the orange trials had significantly higher signal powers, specifically in beta frequencies, compared to the blue trials (Figure 4.2A) ($p<0.02$ Kruskal-Wallis [20], and Dunn [21] post-hoc) (Table 4.2). Significant results (statistically tested and resulting $p < 0.05$) are marked with * in Fig. 2. Moreover, the

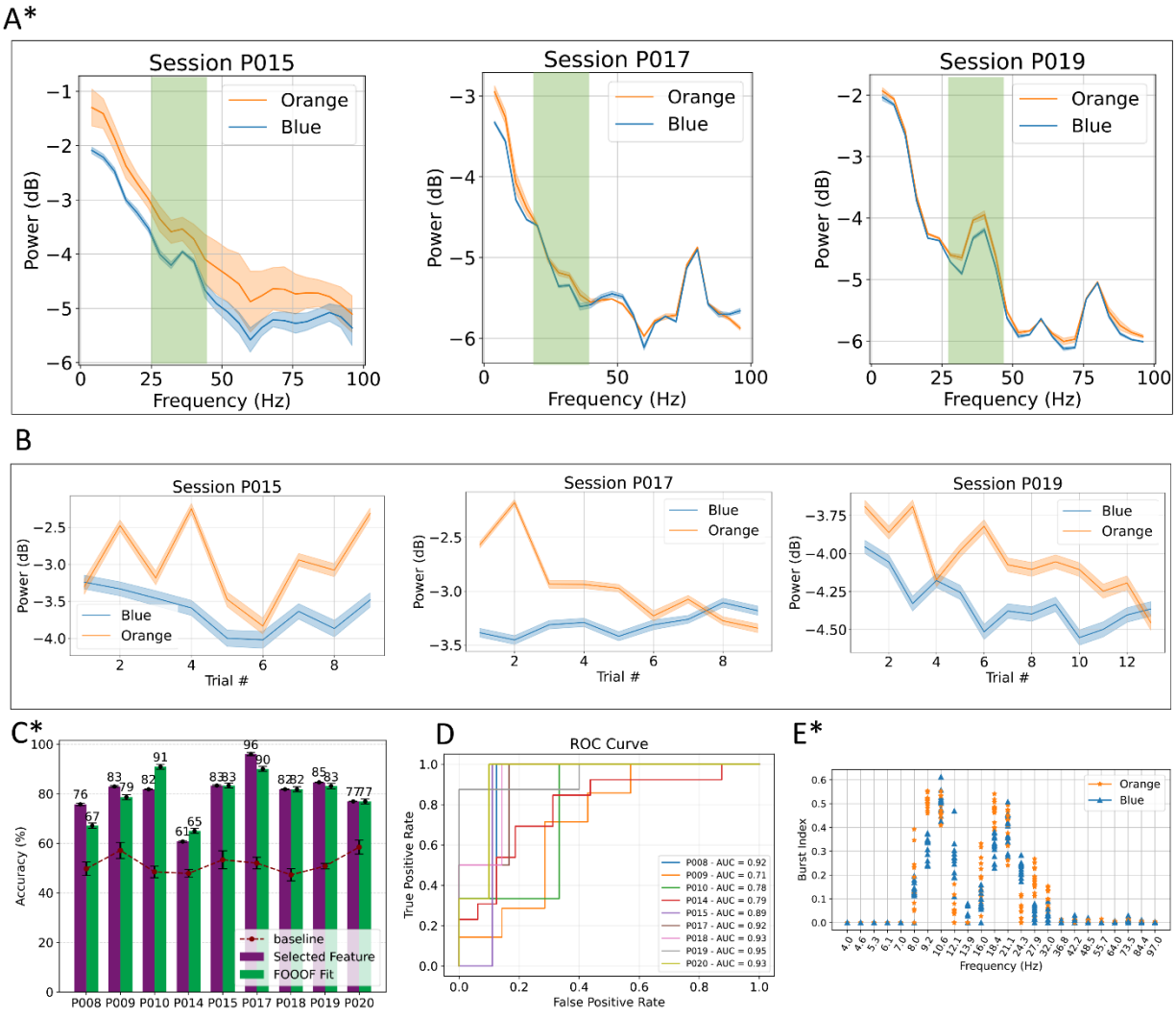

Figure 4.2. The varying modulation of beta between the two task conditions.

(A) The average power of the two conditions is significantly different in certain frequency bands. However, the beta power peaks appeared in different windows (The green rectangular).

(B) The beta power peaks are detected using a FOOOF model. Grouping the beta powers of the two conditions results in two sets of separated features across trials. The features can be used as the input to a decoding model, separating the orange and blue conditions.

(C) Two SVM models decode the orange and blue conditions using manually selected features (purple) vs. automatic features selected by the FOOOF model (green). Both decoders achieved classification accuracies (Y-axis) significantly higher than chance level in all patients (X-axis).

Figure 4.2. Continued.

(D) The ROC curve depicting true positive rate vs. false positive rate for the binary decoding (orange vs. blue) in Fig. 2C for the nine patients. The area under the curve (AUC) for each patient is printed in the caption.

(E) Quantifying the burstiness of the LFP with an index (Y-axis) in different frequencies (X-axis) shows that the beta and alpha (8-32 Hz) bursts are significantly higher than other frequencies. Moreover, the beta bursts in the two conditions are significantly different in the ~9.2 Hz, 16 Hz, 18.4 Hz, 24.3 Hz, and 27.9 Hz. The frequencies (X-axis) are selected in a way to give higher resolution to lower frequencies.

distribution analysis of the PSD of orange vs. blue trials shows significance separation pooled across patients ($p=0.004$, two sample t-test) (Table 4.2). Thus, the PD patients were able to control and modulate their brain activities, manifested by the change in LFP signal power in the frequency domain. Subsequently, the significant difference between orange and blue trials enables the classification of two conditions given the beta peak powers.

4.3.2. Explanatory Model

Across various patients, the beta peaks occurred in different frequencies and sometimes outside the 13-30 Hz range. Moreover, in some patients, the gamma peaks appeared to indicate the trial's condition. As a result, we modeled each trial with the " fitting oscillations and one over f " (FOOOF) method [22] to extract the occurring peaks' powers and center frequencies as decoding features to classify the trials into two conditions. This method divides the signal's power into periodic and aperiodic (1/f-like) components that can be separately analyzed as periodic parameters (center frequency, power, bandwidth) and aperiodic parameters (offset, exponent). Afterward, we used the periodic parameters as input features to a discriminatory method to classify

the trials based on their conditions (orange vs. blue). By using the explanatory method, we eliminate the need to modify the frequency bands to find the power peaks' center frequencies.

4.3.3. Decoding

We trained two SVM-based classifier models to decode the condition's colour given the manually selected power peak values and the selected power peak values by the FOOOF method. The significantly higher-than-chance decoding of manually selected features indicates brain activity modulation by PD patients, albeit unsuitable for BCI purposes. However, by modeling the power spectra for each trial, we were able to decode the conditions with an automatic feature-selection step serving as an appropriate pipeline for a BCI system. We were able to decode the trials' conditions with significantly higher than chance levels (Figure 4.2C) (p=0.0003 Kruskal-Wallis, and Dunn post-hoc using manual and FOOOF features) (Table 4.2). Moreover, analyzing the normalized powers pooled across all patients, we found a significant separation of the two conditions using mixed linear model (MLM) method [23] (p=0.016, MLM) (Table 4.2). The performance of the decoding analyses (Figure 4.2C) are quantified using ROC curves (Figure 4.3C) and confusion matrix pooled across all patients (Figure A3).

4.3.4. Change in Beta Burst Rate

We calculated a burst index to quantify the presence of episodic bursts in various frequency bands. We separated the frequency bands as fractional exponents of 4 to have higher resolution for lower frequencies ($f = 4^i; i \in \{1, 1.1, 1.2, \dots, 3.4\}$). Comparing the burst index in different band frequencies, the burst rates are significantly higher in 8-32 Hz compared to other frequency bands (Figure 4.2D) (p=4.8e-11 Kruskal-Wallis, and Dunn post-hoc) (Table 4.2). Moreover, separating the orange and blue trials, the burst index for the conditions were significantly different (p<0.0006

Kruskal-Wallis, and Dunn post-hoc) (Table 4.2) for multiple alpha and beta frequencies (9.2 Hz, 16 Hz, 18.4 Hz, 24.3 Hz, and 27.9 Hz in one session) (Figure 4.2E).

4.4. Discussion

There is increasing acceptance in the research and clinical communities that pathological synchronization in the basal ganglia–thalamocortical circuit is an important biomarker for the parkinsonian state. Indeed, clinical researchers are currently investigating the utility of subthalamic beta as a trigger for adaptive DBS with promising results [11]. Other researchers are already beginning to investigate chronic feedback training of motor cortical beta oscillations as a therapeutic tool to treat PD. However, they have not yet been tested for effects on motor symptoms [18]. However, an fMRI neurofeedback study suggests neurofeedback training can affect motor symptoms [23].

We demonstrated that neurofeedback training of pathological synchronization in the STN is possible. The PD patients in this study modulated their brain activities, affecting beta bursts' power and rate. These modulations are statistically different in patients' two cognitive states and are detectable using fast decoder models for BCIs.

The current study's outcomes provide impetus toward developing and trialing a clinical therapeutic device that is technically already feasible but not yet indicated. Moreover, the post hoc analyses of the study's data provided insight into how such a device and trial should be designed to estimate optimal brain signal features to drive the feedback.

4.5. Experimental Procedure

4.5.1. Recruitment

Participants are recruited from a pool of patients undergoing the surgical implantation of deep brain stimulation electrodes to treat Parkinson's disease. During the clinical evaluation for DBS implantation surgery, some patients were determined good candidates for participation in the research study. The potential participants were informed about the research study and, if interested, met the team for more details. Lastly, we performed the informed consent process with candidates interested in participating in the research study.

4.4.2. Inclusion/Exclusion Criteria

The inclusion criteria for the standard-of-care procedure [24] include any inclusion criteria for the research. Potential participants may be excluded from the study if their Parkinsonian symptoms do not include bradykinesia of either upper extremity.

4.4.3. Procedure

We recruited nine participants from a pool of patients who will undergo a clinical procedure to implant DBS electrodes bilaterally into the STN as part of the standard of care treatment for Parkinson's disease. (At our center, this is done using a frameless approach (REF)). Each participant will be analyzed independently; no treatment group assignment or blinding is necessary.

Participants were awake and cogent throughout the entire surgical procedure. Immediately prior to the clinical procedure, a researcher reminded the participant about the experimental details, including the research equipment, what the research team was measuring, and what was asked of

the participant. The researcher then fitted and calibrated the research equipment to the participant, including a virtual reality headset.

During the clinical procedure, the neurosurgery team exposed the brain and inserted the microelectrodes through the prefrontal cortex down to the STN using a skull-mounted Microdrive (Nexdrive). The neurosurgery team paused the insertion at various sites to assess the neurophysiology of the structures encountered by the electrodes. The electrodes were connected to a medical device designed to monitor electrophysiological activity from human brain tissue during DBS electrode implantation. The signals captured by the medical device were available to the research computer in real-time. The research computer was always electrically and functionally isolated from the medical device. The neurosurgery team paused at several different sites so the research team could conduct its experiments using the research computer.

The standard-of-care clinical procedure duration is 6-8 hours; the research added 40-60 minutes.

Table 4.2. Summary of statistical analyses and the results.

Non-Parametric Statistics (Figure #)	Hypothesis Test	Post-Hoc Test	P	H/T	F	DoF
2A	Kruskal-Wallis	Dunn	0.02	5.3	9.5	6
2C-Purple	Kruskal-Wallis	Dunn	0.0003	12.8	46.1	16
2C-Green	Kruskal-Wallis	Dunn	0.0003	12.7	46.0	16
2D-(8-32 Hz vs. rest)	Kruskal-Wallis	Dunn	4.8e-11	43.23	484.5	54
2D-(9.2 Hz orange vs. blue)	Kruskal-Wallis	Dunn	7.1e-6	20.17	109.0	26
2D-(16 Hz orange vs. blue)	Kruskal-Wallis	Dunn	0.006	7.52	15.13	26
2D-(18.4 Hz orange vs. blue)	Kruskal-Wallis	Dunn	7.0e-6	20.17	108.9	26
2D-(24.3 Hz orange vs. blue)	Kruskal-Wallis	Dunn	6.0e-6	20.5	112.4	26
2D-(27.9 Hz orange vs. blue)	Kruskal-Wallis	Dunn	3.7e-5	17.0	77.6	26
Distribution Statistics (Figure 2A, Pooled Beta Peak Power(dB) for All Patients)	Mean	Variance	Pooled STD	T-test	T-critical	p
Orange	-2.24	0.58	0.75	2.9	1.96	0.004
Blue	-2.53	0.54				
MLM Statistics for Orange vs. Blue Pooled Across All Patients	Coefficient	STD ERR	Z	P>|Z|	[0.025	0.975]
	0.426	0.178	2.402	0.016	0.078	0.774

4.5. References

[1] M. R. DeLong, G. E. Alexander, S. J. Mitchell, and R. T. Richardson, "The contribution of basal ganglia to limb control," in *Progress in Brain Research*, vol. 64, H.-J. Freund, U. Büttner, B. Cohen, and J. Noth, Eds. Elsevier, 1986, pp. 161–174. doi: 10.1016/S0079-6123(08)63411-1.

[2] Y. Darbaky, C. Baunez, P. Arecchi, E. Legallet, and P. Apicella, "Reward-related neuronal activity in the subthalamic nucleus of the monkey," *Neuroreport*, vol. 16, no. 11, pp. 1241–1244, Aug. 2005, doi: 10.1097/00001756-200508010-00022.

[3] C. Loukas and P. Brown, "Online prediction of self-paced hand-movements from subthalamic activity using neural networks in Parkinson's disease," *J Neurosci Methods*, vol. 137, no. 2, pp. 193–205, Aug. 2004, doi: 10.1016/j.jneumeth.2004.02.017.

[4] A. A. Kühn *et al.*, "Pathological synchronisation in the subthalamic nucleus of patients with Parkinson's disease relates to both bradykinesia and rigidity," *Exp Neurol*, vol. 215, no. 2, pp. 380–387, Feb. 2009, doi: 10.1016/j.expneurol.2008.11.008.

[5] A. Pogosyan *et al.*, "Parkinsonian impairment correlates with spatially extensive subthalamic oscillatory synchronization," *Neuroscience*, vol. 171, no. 1, pp. 245–257, Nov. 2010, doi: 10.1016/j.neuroscience.2010.08.068.

[6] H. Tan *et al.*, "Frequency specific activity in subthalamic nucleus correlates with hand bradykinesia in Parkinson's disease," *Exp Neurol*, vol. 240, pp. 122–129, Feb. 2013, doi: 10.1016/j.expneurol.2012.11.011.

[7] G. Tinkhauser, A. Pogosyan, H. Tan, D. M. Herz, A. A. Kühn, and P. Brown, "Beta burst dynamics in Parkinson's disease OFF and ON dopaminergic medication," *Brain*, vol. 140, no. 11, pp. 2968–2981, Nov. 2017, doi: 10.1093/brain/awx252.

[8] A. Oswal, P. Brown, and V. Litvak, "Synchronized neural oscillations and the pathophysiology of Parkinson's disease," *Current Opinion in Neurology*, vol. 26, no. 6, pp. 662–670, Dec. 2013, doi: 10.1097/WCO.0000000000000034.

[9] S. T. Howard, M. B. Hursthouse, C. W. Lehmann, and E. A. Poyner, "Experimental and theoretical determination of electronic properties in l-dopa," *Acta Cryst B*, vol. 51, no. 3, Art. no. 3, Jun. 1995, doi: 10.1107/S0108768194011407.

[10] N. Jenkinson and P. Brown, "New insights into the relationship between dopamine, beta oscillations and motor function.," *Trends in neurosciences*, vol. 34, no. 12, pp. 611–8, Dec. 2011, doi: 10.1016/j.tins.2011.09.003.

[11] S. Little *et al.*, "Adaptive deep brain stimulation in advanced Parkinson disease.," *Annals of neurology*, Jul. 2013, doi: 10.1002/ana.23951.

[12] A. A. Alegria *et al.*, "Real-time fMRI neurofeedback in adolescents with attention deficit hyperactivity disorder," *Hum Brain Mapp*, vol. 38, no. 6, pp. 3190–3209, Mar. 2017, doi: 10.1002/hbm.23584.

[13] M. M. Lansbergen, M. van Dongen-Boomsma, J. K. Buitelaar, and D. Slaats-Willemse, "ADHD and EEG-neurofeedback: a double-blind randomized placebo-controlled feasibility study," *J Neural Transm (Vienna)*, vol. 118, no. 2, pp. 275–284, 2011, doi: 10.1007/s00702-010-0524-2.

[14] M. B. Sterman and T. Egner, "Foundation and Practice of Neurofeedback for the Treatment of Epilepsy," *Appl Psychophysiol Biofeedback*, vol. 31, no. 1, pp. 21–35, Mar. 2006, doi: 10.1007/s10484-006-9002-x.

[15] J. E. Walker and G. P. Kozlowski, "Neurofeedback treatment of epilepsy," *Child and Adolescent Psychiatric Clinics*, vol. 14, no. 1, pp. 163–176, Jan. 2005, doi: 10.1016/j.chc.2004.07.009.

[16] J. Wolpaw and E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, USA, 2012.

[17] C. B. Boulay, W. A. Sarnacki, J. R. Wolpaw, and D. J. McFarland, "Trained modulation of sensorimotor rhythms can affect reaction time.," *Clinical neurophysiology*, vol. 122, no. 9, pp. 1820–1826, Mar. 2011, doi: 10.1016/j.clinph.2011.02.016.

[18] P. Khanna *et al.*, "Neurofeedback Control in Parkinsonian Patients Using Electrocorticography Signals Accessed Wirelessly With a Chronic, Fully Implanted Device," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1715–1724, Oct. 2017, doi: 10.1109/TNSRE.2016.2597243.

[19] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, Jun. 1967, doi: 10.1109/TAU.1967.1161901.

[20] W. H. Kruskal and W. A. Wallis, "Use of Ranks in One-Criterion Variance Analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, 1952, doi: 10.2307/2280779.

[21] O. J. Dunn, "Multiple Comparisons Using Rank Sums," *Technometrics*, vol. 6, no. 3, pp. 241–252, Aug. 1964, doi: 10.1080/00401706.1964.10490181.

[22] T. Donoghue *et al.*, "Parameterizing neural power spectra into periodic and aperiodic components," *Nat Neurosci*, vol. 23, no. 12, Art. no. 12, Dec. 2020, doi: 10.1038/s41593-020-00744-x.

[23] Lindstrom, M.J., Bates, D.M., 1988. Newton-Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data. Journal of the American Statistical Association 83, 1014–1022. https://doi.org/10.2307/2290128

[24] L. Subramanian, M. B. Morris, M. Brosnan, D. L. Turner, H. R. Morris, and D. E. J. Linden, "Functional Magnetic Resonance Imaging Neurofeedback-guided Motor Imagery Training and Motor Training for Parkinson's Disease: Randomized Trial," *Frontiers in Behavioral Neuroscience*, vol. 10, p. 111, Jun. 2016, doi: 10.3389/fnbeh.2016.00111.

[25] M. S. Okun, H. H. Fernandez, R. L. Rodriguez, and K. D. Foote, "Identifying candidates for deep brain stimulation in Parkinson's disease: the role of the primary care physician.," *Geriatrics*, vol. 62, no. 5, pp. 18–24, May 2007.

[26] E. Bedrosian, "A product theorem for Hilbert transforms," Proceedings of the IEEE, vol. 51, pp. 868-869,1963.

Chapter 5:
Conclusion and Discussion

5.1. Conclusion

Looking inside a black box system controlling and initiating our every move led to remarkable findings in human history, which answered many questions yet left many more unanswered new ones. Some behaviours and attributes thought to be caused by demons possessing the human body were found to be neurological disorders. We have learned much about the brain, but the body's most complex organ still has unresolved mysteries. This thesis explored two brain regions essential to movement and modeled said regions as systems with known inputs, unknown uncertainties, and measured outputs. Our studies attempted to find the best-describing model, explaining the outputs modulations.

Chapter 1 surveyed the background information and history, which can improve the reader's experience in the following chapters.

Chapter 2 explored intracortical BCIs in more detail as another essential piece of information prior to Chapters 3 and 4. The intracortical BCIs address various CNS disorders and benefit millions of people once they become more efficient and less costly.

Chapter 3 studied the PFC of macaque monkeys to explore multiple hypotheses. The saccade movements are controlled in the LPFC; additionally, the PFC is involved in memory circuitry, making it essential for movement planning and learning. Recording from the LPFC was an attempt to simultaneously access activities encoding saccade intentions, learning, and memory. In an 8-target saccade task, we were able to decode saccade intentions. In addition, we studied the effects of learning task rules on LPFC ensemble activities, using which we also classified the different

rules acquired by the monkeys. Finally, we demonstrated that monkeys could learn new rules in fewer trials if they had previously learned a similar rule. In addition, this finding correlates with the distance of ensemble activities in a latent lower-dimensional space.

Chapter 4 studied the STN of PD patients to explore their ability to control an indicator of PD. Nine patients participated in a real-time task attempting to control the beta powers recorded from their STN while receiving feedback. In the offline analyses of the recorded activities, we demonstrated that their attempts modulated the beta powers and bursts.

5.2. Discussion

A freely moving arm would require the real-time processing of motor cortex neurons from the shoulder, arm, hand, and fingers, possibly from other cortical areas like the cerebellum and brainstem. Capturing this activity within the brain is impossible for many reasons (health risk considerations being the most important). Since our brain is never still, another problem is the difficulty of controlling the motionless state. For example, suppose the patient's intention to move a wheelchair to the window can be decoded based on activity from the PFC. In that case, it may then be possible to execute this intention using a motorized BCI wheelchair without the need to decode the required body motion or in conjunction with decoding movement intent from the motor cortex.

Given the vast amount of information encoded in this region, the PFC in humans has yet to be explored. Following the study presented in chapter 3, we believe that LPFC can be a significantly beneficial source of information added to the primary motor cortex in intracortical BCI studies.

Additionally, the research study presented in chapter 4 can be extended to a more prolonged, continuous use of the BCI by PD patients to monitor its effects on PD symptoms using the UDPRS III metrics.

Appendix

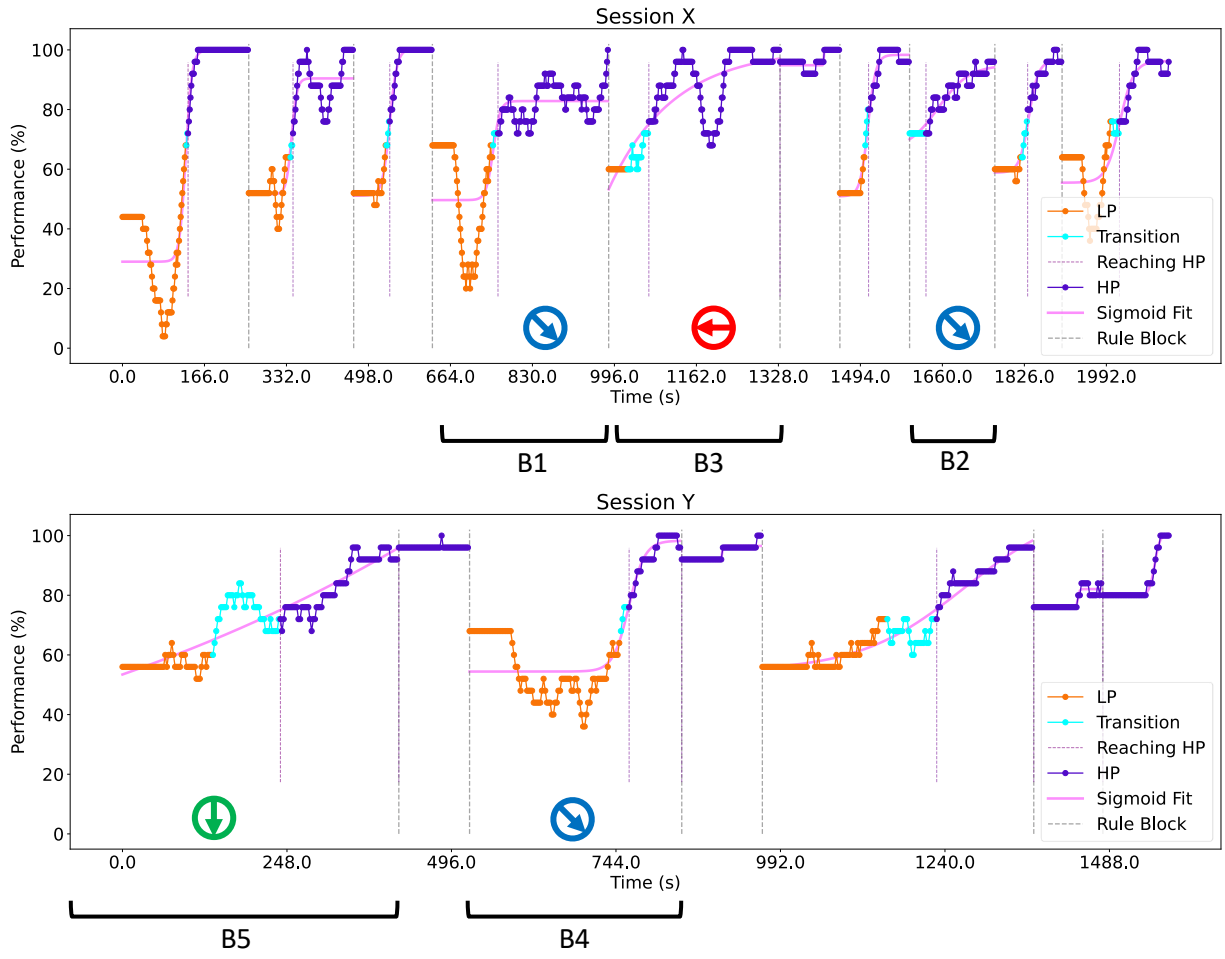

Figure A1. Two Sessions Recorded from the Same Electrodes, Related to Figure 3.2A.

To test the persistence of the best decoding channels two sessions (X and Y) were chosen which are recorded through the same electrodes from the same monkey. The single channel performance of a rule decoder is compared in Figure 3.2A tested on blocks B1, B2, B3, B4, and B5.

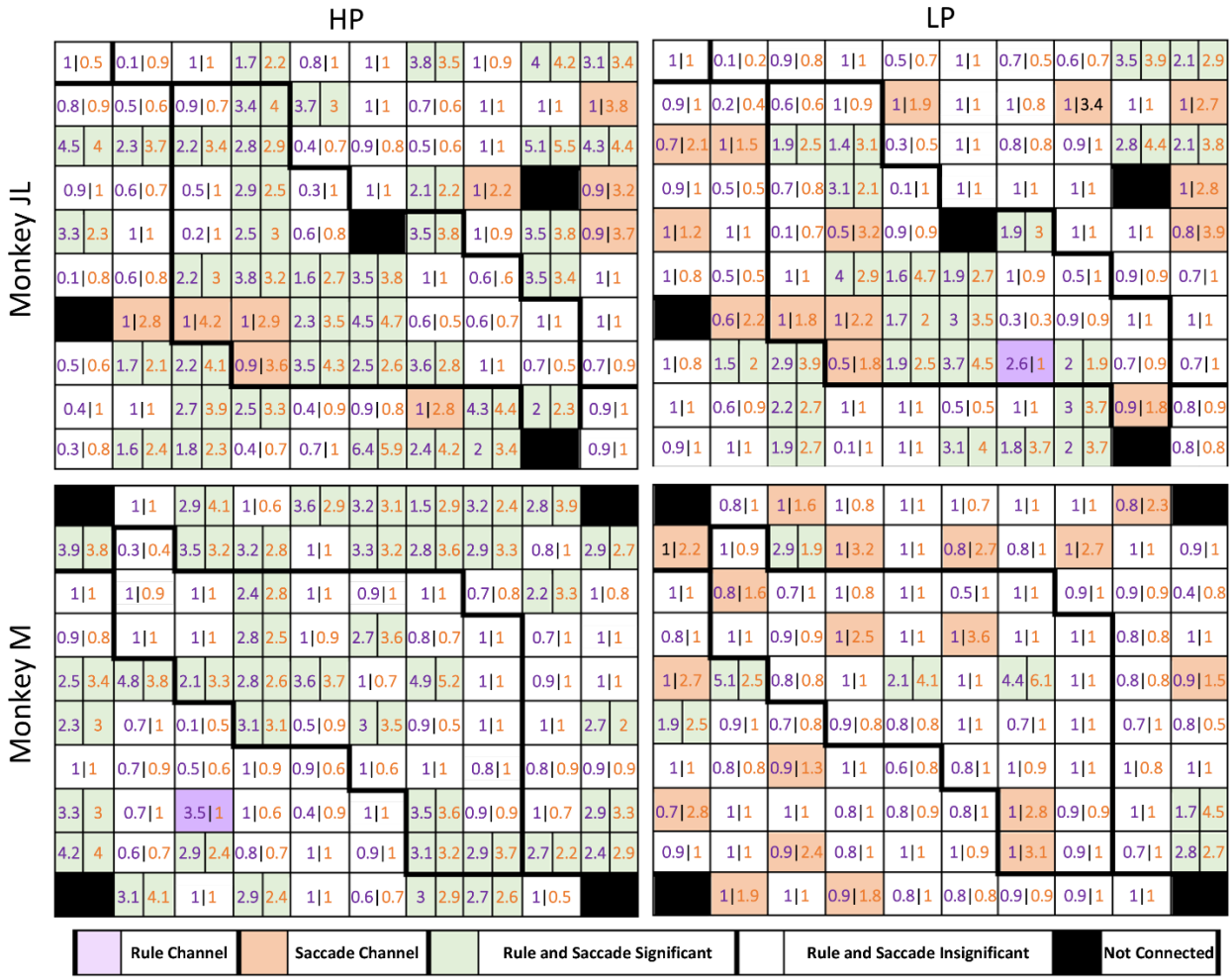

Figure A2. Single Channel Analysis, Related to Figure 3.2B.

The normalized rule and saccade log-likelihood of single channels were calculated. White channels have lower than chance log-likelihood encoding rule (purple numbers) and saccade (orange numbers). Channels with significant saccade log-likelihood are "saccade channels" (orange). Channels with significant rule log-likelihood are "rule channels" (purple). A channel can have significant log-likelihood for saccade and rule (green channels). The black electrodes are not connected to wires. Comparing the LP and HP trials, the number of saccade channels (orange and green) have increased by 26% (38 to 48) in monkey JL and 60% (25 to 40) in monkey M. On the other hand, comparing the LP to HP trials, the number of rule channels (purple or green) have increased by 56% (25 to 39) in monkey JL and 485% (7 to 41) in monkey M.

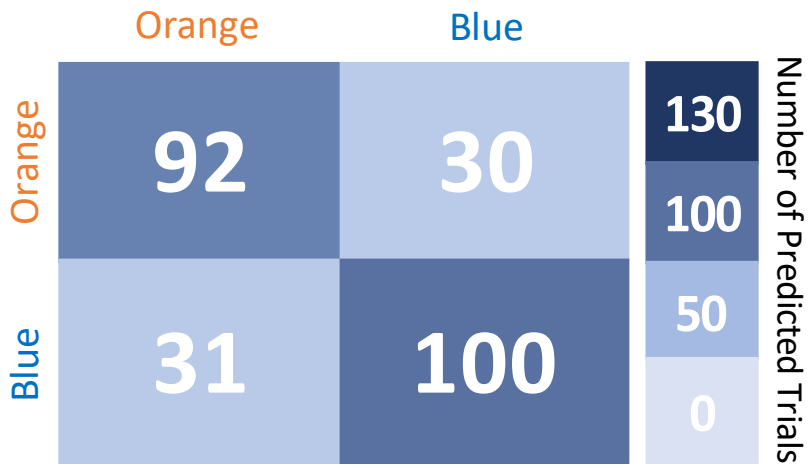

Figure A3. The Binary Classification's Confusion Matrix, Related to Figure 4.2C.

Pooling performance of the SVM decoder across all patients, decoding the orange vs. blue trials using the beta peak powers.