PERSONALIZED BRAIN-COMPUTER INTERFACES FOR AMYOTROPHIC LATERAL SCLEROSIS

A Dissertation in Engineering Science and Mechanics
by
Andrew Geronimo

© 2015 Andrew Geronimo

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

May 2015
The dissertation of Andrew Geronimo was reviewed and approved* by the following:

Steven J. Schiff  
Professor of Engineering Science and Mechanics, Neurosurgery, and Physics  
Dissertation Advisor, Chair of Committee

Zachary Simmons  
Professor of Neurology and Humanities

Patrick Drew  
Assistant Professor of Engineering Science and Mechanics and Neurosurgery

Rick Gilmore  
Associate Professor of Psychology

Bruce J. Gluckman  
Associate Professor of Engineering Science and Mechanics, Neurosurgery, and Biomedical Engineering

Judith A. Todd  
Department Head, P. B. Breneman Chair, Professor of Engineering Science and Mechanics

*Signatures are on file in the Graduate School.
Brain-computer interfaces (BCIs) are a potential last line of communication for those in the late stages of amyotrophic lateral sclerosis (ALS). Following the precedent seen in the field of personalized medicine, this thesis proposal focuses on tailoring BCI devices to personal factors which influence disease outcomes. These factors include the physical, cognitive, and behavioral presentations of the patient as well as the contributing genetic factors. It is with this type of patient-centered personalization that I aim to establish and improve communication in a larger portion of BCI users. I show that previously unused features reflecting task vigilance can be used to increase BCI speed in certain individuals. I also show that psychological changes associated with ALS, rather than physical symptoms, can affect the desire and ability to use a BCI communication system. Analysis of BCI data indicates a frontal shift and delayed timing of discriminable features during a P300 task for ALS patients. Furthermore, patients with cognitive impairment uniquely benefit from BCI features capturing functional connectivity compared to the traditional power features used in a motor-imagery task. In light of the methods for personalization defined in this work, I provide outlook on possible avenues for future BCI development, along with some thoughts on the ethical guidelines for implementation of these systems as assistive communication tools.
CONTENTS

List of Figures vii

List of Tables ix

Acknowledgments x

Chapter 1
Introduction 1

Chapter 2
Literature Review 4
  2.1 Brain-Computer Interfaces 4
  2.1.1 Significant Early Contributions to the Field 4
  2.1.2 Recording methods 8
  2.1.3 BCI paradigms and associated brain phenomena 11
  2.1.4 The BCI pipeline 16
  2.2 Amyotrophic Lateral Sclerosis 24
  2.2.1 Signs and Symptoms of ALS 25
  2.2.2 Brain Imaging in ALS 27
  2.2.3 Assistive andAugmentative Communication in ALS 29
  2.3 Unsolved issues with BCIs 30
  2.3.1 Asynchronous Use 30
  2.3.2 BCI Inefficiency and Predictors of Success 32
  2.3.3 Limitations of the locked-in 33
### Chapter 3

**Gating mechanisms for BCI**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Evoked Responses to Asymmetric Cueing [1]</td>
<td>36</td>
</tr>
<tr>
<td>3.2</td>
<td>Use of single trial gating signals to optimize motor-imagery BCI [2]</td>
<td>39</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Introduction</td>
<td>39</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Methods</td>
<td>40</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Results</td>
<td>50</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Discussion</td>
<td>54</td>
</tr>
</tbody>
</table>

### Chapter 4

**Personalization - User to System**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Desire for BCI use among ALS participants is affected by behavioral health [3]</td>
<td>59</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Introduction</td>
<td>59</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Methods</td>
<td>60</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Results</td>
<td>63</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Discussion</td>
<td>68</td>
</tr>
<tr>
<td>4.2</td>
<td>Cognitive impairment negatively impacts BCI use [4]</td>
<td>71</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Introduction</td>
<td>71</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Methods</td>
<td>72</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Results</td>
<td>75</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Discussion</td>
<td>80</td>
</tr>
<tr>
<td>4.3</td>
<td>Repeats of hexanucleotide G\textsubscript{4}C\textsubscript{2} in \textit{C9ORF72} correlate with quality of BCI performance [5]</td>
<td>84</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Introduction</td>
<td>84</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Methods</td>
<td>86</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Results</td>
<td>87</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Discussion</td>
<td>92</td>
</tr>
</tbody>
</table>

### Chapter 5

**Personalization - System to User**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Feature Optimization</td>
<td>94</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Introduction</td>
<td>94</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Methods</td>
<td>97</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Results</td>
<td>99</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Discussion</td>
<td>106</td>
</tr>
<tr>
<td>5.2</td>
<td>Features describing neural connectivity for BCI</td>
<td>107</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Event-related coherence changes</td>
<td>107</td>
</tr>
<tr>
<td>5.2.2</td>
<td>State estimation based on neural field modeling</td>
<td>114</td>
</tr>
</tbody>
</table>
Chapter 6
Conclusions, Future Directions, and Ethical Considerations

6.1 Personalized deployment and design of BCI devices in ALS
   6.1.1 Prospects for gating
   6.1.2 System Targeting
   6.1.3 System Personalization
   6.1.4 Alternative features

6.2 Ethical Considerations
   6.2.1 Making judgments about BCI use
   6.2.2 Informed Consent
   6.2.3 Issues arising from continuous BCI use

Appendix A
Supplementary Materials

Bibliography
LIST OF FIGURES

2.1 Berger’s EEG .......................... 5
2.2 The Farwell and Donchin speller ....................... 6
2.3 Directional tuning of neurons in the motor cortex ........ 7
2.4 Motor-imagery paradigm .................................. 15
2.5 P300 paradigm ........................................ 17
2.6 EEG artifact examples .................................. 19
2.7 EEG reference types .................................... 20
2.8 Time- and frequency-based features ...................... 22
2.9 Heterogeneity of ALS ................................... 26
2.10 Brain changes in the ALS-FTD continuum ............... 28

3.1 Event-related lateralization ............................. 38
3.2 Phase and amplitude extraction via Hilbert Transform .... 43
3.3 Construction of the VEP template ........................ 45
3.4 VEP template for all subjects ............................ 46
3.5 Flowchart of the gating analysis .......................... 49
3.6 Results of gating by right hemisphere variables .......... 51
3.7 Accuracy and bit rates from the gating procedure .......... 54

4.1 ALS clinic study procedures ............................. 62
4.2 Initial patient survey responses ........................... 65
4.3 Patient BCI requirements ................................. 66
4.4 Interest in BCI functions is affected by cognition and behavior 66
4.5 Logistic regression model for participation in the BCI pilot study ...... 67
4.6 Success of BCI use influences patient opinion on the follow-up survey . 68
4.7 BCI accuracy among patients and controls .................. 77
4.8 P300 quality of ALS patients ............................. 78
4.9 Regression of BCI quality on patient factors ................ 79
4.10 Motor-imagery quality of ALS patients .................... 80
4.11 Elevated high-frequency power occurs in cognitively impaired patients

4.12 Older patients perform better with P300

4.13 Histogram of GGGGCC repeat lengths within C9ORF72

4.14 Disease factors do not correlate with sub-threshold expansions

4.15 Repeat length is inversely related to P300 quality

4.16 Repeat length is inversely related to motor-imagery quality

5.1 P300 error among patient participants by classifier complexity

5.2 P300 error by electrodes used in classification

5.3 P300 error generated from reduced electrode sets

5.4 P300 error by time points used in classification

5.5 Motor-imagery error by electrode, averaged over participant groups

5.6 Motor-imagery error generated from reduced electrode sets

5.7 Differential mu power between right and left trials

5.8 Differential coherence in the mu band between left and right trials

5.9 Classification results from power and coherence features

5.10 2D representation of field-, state-, and sensor-space

5.11 Modifications to the estimation framework

5.12 Replication of kernel estimates in the Freestone framework

5.13 Reconstruction of the simulated field

5.14 Spectral analysis of field reconstruction

5.15 Estimation fidelity of the EEG fields

5.16 Spatial frequencies in the sensor and basis space for the EEG fields

5.17 Estimated kernels from motor-imagery EEG data
### LIST OF TABLES

3.1 Results of the permutation test with the three gating variables . . . . . . 52  
3.2 Simulated gating results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53  
4.1 ALS patient characteristics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64  
4.2 Characteristics of ALS patients and control participants in the BCI study 76  
5.1 Parameters used in the neural field estimation framework . . . . . . . . 120
I would first like to thank my advisor, Dr. Steven Schiff for his constant encouragement and guidance. I am also indebted to Dr. Zachary Simmons, without whom my past and future work in the ALS Clinic would not be possible. For first recognizing and nurturing my potential, I thank Dr. Mst Kamrunnahar. Thanks to my committee members, Dr. Patrick Drew, Dr. Rick Gilmore, and Dr. Bruce Gluckman. You challenged me to think beyond my own little world and have elevated the quality of this work. Thanks to the nurses, specialists, and staff who welcomed me as a part of the team, especially Travis Haines, Judy Lyter, Heidi Runk, Beth Stephens, and Susan Walsh. To my fellow graduate students who have lent me their heads, both literally and figuratively, you have succeeded in helping me keep my own. This work would not be possible without the following funding sources: NIH grant K25NS061001(Kamrunnahar), the endowment funds of Harvey F. Brush, the Paul and Harriet Campbell Fund for ALS Research, the ALS Association Greater Philadelphia Chapter, and many other private donations to the Penn State Hershey ALS Center.

A special thanks to all the patients who gave of their precious time, your excitement brought meaning and joy to my work.

Last, to my parents and family, even though your eyes glazed over at even the briefest of research descriptions, I owe you all my success.
In the winter of 1995, French editor Jean-Dominique Bauby suffered a stroke which left him nearly completely paralyzed. Following months of adjustment to the reality of his physical condition, he painstakingly wrote *The Diving Bell and the Butterfly* with the help of a translator, using only the blink of an eye. Twenty years later, the field of brain-computer interfacing is on the cusp of an answer to the question posed in the final pages of his memoir.

As an emerging field of scientific study, brain-computer interface (BCI) research has experienced a divergence of nomenclature during its development, with systems labeled brain-response interfaces, thought-translation devices, brain-machine interfaces, and direct-neural interfaces, to name a few. Although there is no academic society or governing entity which formally organizes BCI researchers, the community set out at the 4th International Workshop on BCI to establish a definition, and collectively agreed on four main points [6]. A BCI:

- Must directly record activity from the brain
Utilize at least one intentionally-modulated brain signal

- Processing of this signal must occur in real-time
- Feedback must be given to the user to communicate success in the task.

Aside from their important uses as tools of scientific inquiry, the potential applications of BCIs are widespread, encompassing novel forms of biofeedback for purposes ranging from immersive entertainment to vigilance monitoring. Much of the research into BCI technology has been with the goal of enhancing the quality of life for individuals with conditions causing neurological dysfunction by aiding in motor control [7], communication [8] and rehabilitation [9]. The focus of this dissertation will center on the use of BCI for augmentative and alternative (AAC) communication for those with neurological disorder, specifically those with amyotrophic lateral sclerosis (ALS). This disease causes progressive degeneration of motor neurons, leading to muscle weakening, and in some cases, to locked-in syndrome similar to what was experienced by Bauby. The physical limitations of this patient population necessitate the development of a practical tool which can augment communication and mobility.

Until very recently, BCI-AAC systems studied in healthy individuals and ALS patients were created with the assumption of homogeneous, strictly motor limitations across users. As one group of researchers note, this assumption couldn’t be further from the truth. “Disease heterogeneity in ALS has multiple dimensions, including genetic origin, site of onset, rate of decline, the presence of cognitive impairment and the relative degree of upper and lower motor neuron involvement” [10]. In light of this, it may be of little surprise that ALS patients continue to under-perform in BCI compared to neurologically healthy individuals, and in critical situations often achieve no control whatsoever.

As will be described in this thesis, nothing short of a patient-specific approach to BCI design will be suitable for achieving robust control in the majority of patients. In the following chapters, I will describe how the heterogeneity observed in ALS patients describes a pattern of device utility. Even more critically, I outline methods to personalize the BCI system to account for specific motor and non-motor limitations of the disease. I do this because I hope to see BCI technology become a valuable and effective means of communication for those living with ALS, and hope to implement personalized BCI methods into device design in the future.
The following work builds upon the successes of scalp-based BCI communication discussed in Chapter 2, while addressing the unique challenges to deployment faced by the ALS population. Early work in my thesis focused on BCI device personalization applied to healthy, young individuals (Chapter 3), while my later work focused on the challenges of patient-to-device and device-to-patient personalization for ALS users (Chapters 4 & 5). The findings emerging from these studies are presented in short below.

- Previously unused brain signatures can be employed as gating mechanisms as part of a hybrid BCI, in order to increase the accuracy and speed of the BCI system by identifying periods of task vigilance and cortical excitability.

- Behavioral abnormalities occurring in some ALS patients reduce their interest in BCI adoption.

- Cognitive deficiencies associated with ALS can lead to poor performance in standard BCI protocols. This poses a clinical challenge, as patients with cognitive deficits are just as eager to use a BCI for communication.

- Preliminary evidence exists that genetic screening of the ALS-linked C9ORF72 genetic expansion could form a predictor of BCI system success.

- Patterns of feature extraction differ between control and patient participants on two BCI tasks. Furthermore, we show how patterns of control signals diverge for ALS patients who experience cognitive impairment.

The final chapter addresses some of the future questions and design implementations that make up the next step in this line of work. Furthermore, a portion of the discussion is dedicated to ethical matters pertaining to my present and future work in this field.
2.1 Brain-Computer Interfaces

2.1.1 Significant Early Contributions to the Field

Although the field of BCI is relatively new, the neuroscientific and engineering discoveries that preceded its definition have a much longer history. The contributions to the field of BCI, which have grown at an increasing rate since the 1990’s, rise from a foundation built in electroencephalography, neurophysiology, and computer science. Comprehensive reviews on the discoveries which preceded the technical and conceptual foundation of BCIs are available [8, 11]. Here I overview some of the work that has gone into the development of tools for studying the electrical phenomena of the brain at the level of the scalp, the definition and refinement of paradigms used to infer information about a mental state from neural activity, and the progress in computational and statistical methods for making these inferences.

By the 19th century, there was a growing body of knowledge about the electrical impulses associated with living tissue, and electrical studies of the brain had been conducted in a number of animal species. Hans Berger, a German neuropsychiatrist, is credited as the first to document the nature of the human electroencephalogram (EEG) in 1929 [12]. Berger used a galvanometer in his early studies, more advanced than those used by his predecessors, which allowed him more sensitivity and bandwidth in
his recordings. Although he is most famously credited for showing the occipital alpha rhythm (Figure 2.1) and blocking of this rhythm, he would also describe the EEG characteristics of sleep, consciousness, hypoxia, and various brain disorders in human subjects [12, 13]. A substantial review is given on the developments of EEG that occurred throughout the world after the popularization of the technique by Berger [12]. By the 1950s, EEG was a concept well known in medicine and the academic world; nearly all university hospitals had at least one EEG machine [12]. During this time, EEG flourished as a tool used to study epileptic foci, the neurological correlates of sleep, analysis of evoked potentials, as well as leading to the development of more invasive bioelectrical sensors. In the last 60 years, the clinical use of EEG has diminished due to the availability of superior non-invasive techniques for structural and functional brain imaging, but researchers still employ EEG for its applicability to human subjects and desirable temporal resolution, in order to related the electrical activity of the brain to the complex human behavior.

The birth of proper BCI research likely rose out of the study of operant conditioning, where the voluntary regulation of EEG characteristics in the presence of feedback had been explored by researchers since the 1960’s. This type of biofeedback was subsequently realized as a potential BCI by Birbaumer and colleagues [11]. His group showed that slow cortical potentials (SCPs) were able to be trained by subjects to be increased or reduced at will, providing at least rudimentary brain control of a computer system [14]. Around a similar time, a separate class BCIs emerged, which were based on the user’s innate response to rare visual stimuli. Farwell and Donchin created the now famous and perhaps most widely successful system, the P300 speller, (Figure 2.2) [15] which utilized the electrical potential evoked by rare target stimuli to identify the user’s letter of intent. This system achieved an increased range of outputs for the user at a faster speed. It represented a dramatic increase in the usability of such a system for a broader

![Figure 2.1: Berger’s recording of the electroencephalogram, along with a 10 Hz sine wave reference [12].](image)
spectrum of users, without the need for a lengthy conditioning process. A third system based on the imagination of motor movements was pioneered primarily in the work of Pfurtscheller and Aranibar, who identified the suppression of cortical rhythms that accompanied self paced movements [16]. From these observations, BCI systems based on the imagination of movement were also developed, functioning on a similar principal of cortical desynchronization. Unlike the SCP systems before it, motor-imagery BCIs have been shown to be capable of producing multidimensional control through the use of different mental imageries.

Other advances in recording technologies and computation have contributed significantly to the field. Berger performed his recordings using a single pair of electrodes connected to a galvanometer, recording on photographic paper [12]. The sophistication of our tools today allow us to process enormous amounts of data from sensors that provide an image of the brain in far greater detail. Indeed, some of the most fantastic results come from intracranial BCI systems. In 1982, Georgopoulos et al. published a landmark study in the field of single unit recordings, showing that primary motor cortex neurons are tuned preferentially to movement in one direction (Figure 2.3, [17]). Their recordings from the motor cortex of the monkey showed that direction of movement was coded in these individual cells, and movement trajectories could be reconstructed from their firing rates alone. They later showed that the population coding of many neurons in the motor cortex were responsible for the range of movements in 3-D space [18]. The work of Nicolelis et al. expanded on this work by designing a robotic appendage that could be controlled purely from recording direction sensitive neurons in the motor cortex, and that this control existed with and without the presence of overt movement [19]. This phenomenon, which was found to be true in humans as well, along with the field of neural spike encoding/decoding, contributed making invasive BCI techniques at the state-of-the-art for brain-controlled prosthetic development.

![MESSAGE BRAIN](image)

Figure 2.2: The speller matrix used to evoke P300 potentials for communication in the first of its type BCI [15].
Figure 2.3: Georgopoulos showed that a single neuron in the motor cortex of the monkey have a preferred firing direction [17]. In the case of the neuron shown here, the highest firing rates are when the monkey moves its arm between 90 and 210 degrees.

Besides some notable exceptions, such as the work at the BrainGate project [20] where microelectrode arrays are implanted into the brain, most invasive studies in humans use subdural grids which record the electrocorticogram (ECoG). This was first achieved in a closed-loop paradigm by Leuthardt et al. [21]. ECoG currently holds much of the potential of the BCI field as it delivers improved spatial resolution and signal quality of underlying cortical populations compared to EEG while also being relatively safer to implement than more invasive methods. However, at this point the majority of subjects participating in ECoG studies are those for which a grid is being implanted for epilepsy localization purposes.
2.1.2 Recording methods

Electroencephalography

BCIs use a number of methods to record the activity of the brain. The earliest described and most commonly used still is the EEG, which is a measure of electrical potential of the brain sensed by electrodes placed on the scalp. The macroscale electrical potentials measured by EEG are generated at the cellular level by the transfer of ions in and out of neurons and glia. Specifically, synchronous dendritic current flows of millions of cells having similar columnar organization are shown to contribute to macroscale electrical phenomena [22]. Additionally, potassium-mediated current flows associated with glial depolarization may serve to amplify this effect [12]. The currents associated with these membrane depolarizations that flow through the extracellular space generate field potentials, and the time course on which they occur dictate the frequencies observed in EEG. For groups of directional neurons, for example in layer V of the cortex, the extracellular currents along the main axes of the neurons sums to the generation of a dipole layer. These dipoles are mesosource generators of electrical potentials, which undergo some low-pass transformation by a conductive medium [22]. These types of large summations of synchronous potentials are the type that can be measured at a distance of an EEG scalp electrode [12].

Various factors contribute to the generation of EEG functional dipoles in the cortex. Brain tissue is inhomogeneous, with conductance varying by location. It is also anisotropic, especially white matter tracts. The skull, which possesses both of these properties, has a major effect on the volume conduction of the EEG. An inner layer of conductive cancellous bone shunts current due to the greater resistance of the outer layer of cortical bone [22]. As a result, a single scalp electrode measures the neural activity from tissue masses containing $10^8 - 10^9$ neurons [22].

The EEG is a measure of voltage, determined as the potential difference between an active electrode and a reference. This means that all EEG recordings are bipolar, however, EEG can be recorded using a number of choices for the reference electrode. These references can be determined by physical location, or virtually by subtraction. Linked-ear or linked-mastoid references are popular choices for physical references, as they are as electrically distant as can be achieved on the head, but suffer a critical drawback. The actual reference provided by these linked electrodes varies depending
on the differential resistance at their interfaces [22]. If the resistances are equal, the reference is average potential between the ears, but if the resistance at one ear is much higher, the potential will be heavily weighted towards the other ear. Additionally, this linkage can cause current flows between ears, leading to electrochemical potentials at the electrode interface that are detrimental to recordings. An ideal reference is dependent on the task being studied, and it should be electrically distant from the generators involved in the task. The test for a good reference is if the EEG dynamics of interest remain unchanged if the reference is moved. This means that the source of the dynamics of interest are sufficiently far from the reference to not be affected by its exact placement [22].

**Magnetoencephalography**

The magnetoencephalogram (MEG) is the electromagnetic complement to EEG, and also possesses a substantial history as a tool for neuroscience research. Superconducting quantum inference devices (SQUIDs) were developed as magnetic sensors in the 1960’s. Later that same decade MEG recording with SQUID technology was demonstrated by Cohen et al., and is currently is the most widely used application of this technology for biosensing, although magnetocardiography and low-field magnetic resonance imaging are active areas of research [23]. As part of Cohen’s characterization of MEG, he and his colleagues performed both theoretical and experimental demonstrations of the differences between magnetic and electrical recordings of the brain. [24, 25]. Using both theory and data from evoked potential studies, they showed three fundamental properties of MEG: the MEG lead field pattern is rotated 90 degrees with respect to EEG, MEG displays slightly better localization of sources due to insensitivity of volume current conduction by the skull, and there is asymmetry in the EEG field due to the radial component of the dipole from which MEG is unaffected [25].

Somewhat of a paradox has existed in MEG research concerning the usefulness of simultaneous MEG and EEG. Malmivuo describes the concept in terms of Helmholtz’s Theorem [26], which describes a general electric field as the sum of two vector fields: the flow source and the vortex source. For this type of recording, we measure these sources as the electric and magnetic fields from the brain. For each of these fields, there are three orthogonal lead field components (sensitivity to electromagnetic fields) which are independent of each other. It is this independence which has led researchers
to believe that the addition of magnetic recordings could provide new information not found in EEG. While each component is independent, the signal recorded by them is not since they originate from the same volume source. Therefore, the independence of sensitivity distributions of the two methods does not imply the independence of recorded signals.

Even in light of this, a common finding that has been echoed with a number of researchers is that the information provided by simultaneous EEG and MEG recordings is complementary for source localization [27, 28, 29, 30]. One reason for the complementary nature of the two methods is that the sensitivities of the lead fields are different for the two recording methods [26]. This can arise from heterogeneities in the volume conductor [28]. MEG selectively identifies tangential sources, or those arising from cortical infolds [22, 26], which can be used to identify source topology of, for example, photic alpha entrainment as having significant tangential generators [31]. EEG is more sensitive to deeper sources within the brain, while MEG is better at localizing sources closer to the surface [32], and is less susceptible to volume conduction effects or skull deformities than EEG [22, 33].

There have been a few BCI studies which employ MEG. While useful as a tool to probe brain activity, MEG is mainly hindered by its large size and cost, and special considerations that accompany the used of super-cooled magnets. MEG has been used in a motor-imagery paradigm to communicate binary decisions using modulation of sensorimotor rhythms [34]. MEG was used in the same way to decode hand movement in four directions [35]. Using simultaneously recorded EEG, it was shown that MEG was superior at distinguishing between the four states. MEG has been combined with source localization techniques during a visual evoked potential-based BCI to demonstrate how cortical generator locations change with stimulation frequency [36].

Other Methods

In addition to macroscale measurements of electromagnetic brain phenomena, BCIs have also been implemented with fine-scale electromagnetic measurements, as well as metabolic correlates of neural activity. Invasive electrical measures include EC0G [37] arrays and implanted depth electrodes [38]. The former measures the local potentials of a large population of neurons, while the latter measures the multi-unit spiking activity of neurons in the proximity of the electrode. The potential for technologies which achieve
the recording quality of surface electrodes without the risk of invasive brain surgery are obviously needed. A potential candidate is the epidural screw electrode, which is implanted within the skull, a source of major information loss in electrical brain recordings, but without penetrating the dura mater, and thus avoiding the complications of infection and replacement [39, 40].

Metabolic BCIs allow for the indirect assessment of neural activity, usually through the measurement of blood flow. Functional magnetic resonance imaging (fMRI) [41] records a blood oxygenation level-dependent (BOLD) response which is a measure of oxygen saturation in brain tissue that can be correlated with neural activity. Near infrared spectroscopy (NIRS) [42] uses an optical signal to measure the concentration of oxygenated hemoglobin in the blood, also signaling the metabolic demands of active tissue.

Although these technologies will not be covered in further detail here, their contributions to the field have been substantial in many ways. Invasive recordings have enabled the longest-running applications of continuous BCI use [38], and allowed for control at the finest levels of detail, down to individual finger movement [43]. Metabolic BCIs have allowed for communication to be established in advanced stages of neuromuscular paralysis, where electrical BCIs have so far failed [44]. These technologies certainly help augment our understanding brain phenomena, however EEG remains the most widely used modality because of its high temporal resolution, its ease of use, and its knowledge base encompassing over 90 years of theoretical and experimental work.

2.1.3 BCI paradigms and associated brain phenomena

Types of BCIs

Over the relatively short history of brain-computer interaction, a number of paradigms have been shown to be successful in extracting information of intent directly from neural activity. Here an overview of the paradigms is given, with focus on two such mechanisms which are used in this research, the P300 and motor-imagery BCI paradigms.

BCI paradigms can most generally be distinguished by the type of interaction the user has with the system. Three general categories of BCI systems are those which rely on active, reactive, or passive mechanisms for control [45]. Active BCI paradigms are those which rely on the self-modulation of brain signal in a process that is completely
endogenous; the user is capable of generating the control signals without external cuing or stimulation by the computer. Active approaches include all forms of mental imagery, of which motor imagery, spatial navigation, mental arithmetic, and verbal imagery represent some which have been used successfully. Slow cortical potentials fall into this category as well [14], in which users are trained through feedback conditioning to generate long bouts of cortical depression to facilitate binary communication.

Reactive BCIs rely on the user to respond to a stimulus from the computer. The control signals generated as part of a reactive BCI would not be possible without the stimulus provided by the computer, which can be auditory, tactile, or visual. Perhaps the most successful BCI device to date is the P300 speller, an example of a reactive system because user delivers their intent by reacting to one of many target stimuli, differentiated by the presence of an evoked brain potential. Another visual BCI in this group is the steady-state visually evoked potential (SSVEP) BCI, which presents the user with a number of stimuli flickering at different frequencies [46]. The user delivers intent by reacting selectively, in this case focusing their attention, on the stimulus which corresponds to that intent. Reactive BCIs can also be facilitated with other sensory modalities such as somatosensation and audition [47, 48].

The last category is the passive BCI, which is completely absent of voluntary human control. Although there is some controversy around ‘diluting’ the definition of BCI with the inclusion of devices that serve as more or less cognitive state monitors [49], the inclusion of passive monitoring, especially as components of a ‘hybrid’ BCI system, has gained momentum [6].

**Motor-imagery**

Imagery is a widely used tool in BCI because it requires no overt movements, and generates a distinct pattern of cortical activation which can be discriminated as focal deviations in the EEG. Various types of mental imagery have been used in BCI control, including spatial navigation, mental rotation, mental arithmetic, word imagery, auditory imagery, and motor imagery [50, 51, 52, 53]. Although these studies point to a number of imagery techniques, motor imagery has received the most attention as a mental strategy for achieving communication of voluntary intent through a BCI. This may have been because this type of imagery provides high inter-session stability [51], or simply because of the straightforward implementation of imagery as a possible aid for motor
system dysfunction.

The performance of motor imagery is not an intuitive task; in most individuals, it is a skill that requires substantial practice. Learning how to control these signals is similar to other physical skills. By using a combination of practice and feedback, users are able to modulate brain activity effectively. Kinesthetic, or first-person imagery of motor action, rather than visual, or third-person imagery, has been shown to be more effective at enhancing cortical excitability [54] and better for generating desynchronization over motor regions of the brain [55]. The imagination of motor action is not only useful in BCI research, but also helps to determine efficacy of rehabilitation of upper extremities. Various tests and revisions on these tests have been developed to assess the ability to perform both visual and kinesthetic imagery [56, 57].

There is a substantial literature unifying the phenomena of motor execution, motor imagery, and motor observation through anatomical, functional, and lesional studies. Jeannerod proposes a “unifying mechanism for motor cognition”, which he states imagery and action observation provide a neural simulation for motor action [58]. Bolstered mainly by fMRI evidence, there is large overlap in the activated motor brain regions involved in all three motor states, although, understandably, there are differences the magnitude and topography of those tasks, as well as patterns of functional connectivity. [58, 59]. Those who are trained and skilled performers of motor imagery, for example, show greater activation of pre-motor areas, a locus of motor planning [60, 61]. A typical estimate is that imagery produces about 30% of the intensity of overt action as recorded by fMRI [62].

The brain circuitry involved in coordinated motor actions and imagery originate at the level of the neuron, although the large scale fronto-parietal circuitry that contributes to the generation of oscillations are more relevant to EEG-based recording. These brain oscillations that are related topographically and functionally to motor representation are termed the sensorimotor rhythms (SMR). Although they encompass multiple frequency bands, the first discovered in 1952 by Gastaut [63] was central or Rolandic mu rhythm, which oscillates at 8-13 Hz and decreases in the presence of movement. This rhythm is related in to the occipital alpha because of its similar rate (8-13 Hz) and its classical interpretation as an idling rhythm [64]. However, the mu rhythm has been shown to be anatomically [64] and functionally [65] distinct from the occipital alpha rhythm. Thalamo-cortical loops are involved in the generation of mu rhythms in cats [66, 67],...
and at least for its occipital counterpart, the alpha rhythm, this functional circuit exists in humans as well [68]. The loop receives additional inputs from secondary motor areas, including the basal ganglia and cerebellum [69]. In the case of motor imagery, there is an inhibitory influence from the supplementary motor area in modulating mu power changes [69, 70]. Mu rhythms are popularly employed in BCI because they are of high amplitude and easy to resolve with EEG recordings. However, the activity in higher frequency bands such as beta (18-25 Hz) and gamma (30-100 Hz) have been implicated in EEG, ECoG, and intracranial recordings to involve localized processing [71, 72] and are causal for widespread changes in mu activity [73].

An example of a typical center-out motor imagery task is shown in Figure 2.4. The mechanism for control in this task involves the depression of mu and beta rhythms when motor-imagery is performed, a phenomenon termed event-related desynchronization (ERD) [74]. This term was chosen because the oscillations of EEG represent postsynaptic firing in a large group of neurons [75]. ERD is interpreted as the electrophysiological result of activated cortical areas becoming involved in the initiation of a motor behavior. Factors such as task complexity, attention, and performance efficiency contribute to the amplitude and spread of desynchronization [76]. This blocking begins to occur during the planning stages of motor action, and manifests shortly before the action is performed [64, 76]. The patterns of desynchronization are somatotopically localized to the representation areas of the imagined limb, especially in the upper frequency bands known to reflect the activations of smaller scale networks. These power changes at the scalp are in line with intracranial recordings [72]. The signals of interest in an imagery task are the topographical differences in SMR desynchronization, which are calculated commonly through a frequency-based method. The differences in SMR power over the pre-motor and motor cortices enable classification of covert motor intent into a command signal [74].

The P300

The P300 is a positive EEG evoked potential that occurs around 300 ms in response to a novel, infrequent, or unexpected stimulus, commonly called the ‘oddball’ response. The P300 is typically recorded over the medial centro-parietal cortex, peaking anywhere from 300-900 ms after the stimulus, with amplitude proportional to the rarity of the stimulus [77]. A theory for P300 generation centers on context updating [78]. According to
Figure 2.4: (Top) A two-class center out motor imagery paradigm, in which the user controls the cursor (ball) to either a left or right target (box) on the screen. (Middle) Spectrogram of channel C4 during a motor imagery run, with the timing of left and right imagery trials marked on the side in blue and red. Each marker corresponds to a three-second period of imagery. (Bottom) The average of the spectra over all left and right time windows. Discrimination between the two tasks is evident in the mu and beta bands.

This theory, a change to stimulus attributes act on attentional and memory processes to update the context of the stimulus while producing the P300 evoked potential. Without this detection of change, only exogenous sensory evoked potentials are generated as a
result of the stimulus. There are diverse structures implicated in the generation of the subcomponents of the P300 [77]. However, there are some commonalities found between intracranial recordings, combined fMRI/EEG, and lesion studies. Involvement of the temporo-parietal junction implicates a circuit pathway between frontal and temporo-parietal regions of the brain for generation of this response. [78]. Distinct varieties of the P300 have been demonstrated, with the component most relevant to the oddball response inherent in the P300 spelling task, being the P3b potential. This specific potential is characterized by its maximal response in parietal cortex and occurrence that has slightly greater latency than the anterior P3a evoked potential [78]. For the P3b, the actions of the posterior parietal cortex also factor in as both an integrator of visuomotor inputs as well as processing goal-directed attention [77].

The P300 speller, which was first developed in the late 1980s employs this covert marker for selective attention to establish a communication channel directly between brain and machine. The system relies on the user to react to a specific letter of intent out of a larger group of randomly flashing letters in a way that generates a P300 only for the intended letter [79]. Usually, these devices are implemented by arranging the alphabet in a grid on a computer screen, and flashing the rows and columns of the grid in a random pattern. With this type of system, a letter is decoded as the one at the intersection of the row and column which elicit the largest P300 response. Other stimulation methods include flashing the letters in randomized groups in order to limit adjacency errors (Figure 2.5), [80]. The process of P300 classification generally requires the averaging of multiple trials, and there is a trade off between device accuracy and speed. Of current BCIs available, these devices perform the best in terms of accuracy and information transfer, being able to deliver nearly 100% accuracy and 20 bits of information transferred per minute [80].

2.1.4 The BCI pipeline

The signals generated by active or reactive brain modulation comprise only half of a functional brain-computer interface. The computation performed on these signals, namely the pre-processing, feature extraction, classification, and feedback are critical to the function of the system. Perhaps most of the effort that has gone into improving BCI technology has centered on methods of computation for brain signal translation. Of
Figure 2.5: Top: A checkerboard P300 speller, with four letters highlighted at once. The user is focusing on the letter ‘A’. Bottom: Representative EEG signals from one individual. The blue line is the average EEG following the target letters, and the red line is the average EEG following the non-target letters. The P300 is best observed in Cz (with the axis markers), although it can be seen in many central and parietal channels.

course, different feature extraction methods and classification techniques are used for each type of biosignal and each type of paradigm. Extensive reviews are given on these
elsewhere [81, 82, 83], but I would be remiss to fail to give at least an overview of the major trends in signal processing that have accompanied BCI research.

**EEG preprocessing**

Some recording methods are more sensitive to artifacts than others, but all experience at least some infiltration by signals not relevant to the task. These can include muscular, ocular, respiratory, and cardiac artifacts, as well as line noise from nearby electronic devices running on 60 Hz alternating current power. Two common examples of jaw and ocular artifact are displayed in Figure 2.6. Often the first step in feature extraction is to remove or reduce these artifacts from the raw data before proceeding. Line noise is often removed during recording through use of digital bandpass filters implemented within the amplifier. Most brain signals useful for EEG-based BCI are below the 60 Hz range of line noise, so the filtering of these signals yields minimal loss of information.

For addressing muscular and ocular signals that may appear in the recording, additional electromyogram (EMG) or electrooculogram (EOG) electrodes may be integrated into the system. This way, trials having these artifacts can simply be removed from further steps, or can be used to create spatial filters which reduce artifact contamination [84, 85]. Although respiratory and cardiac signals may be picked up by certain EEG leads, these artifacts are much smaller if present at all, predictable, and can also be removed using spatial filters.

As mentioned earlier in the introduction, each EEG channel is the bipolar voltage difference between the active and reference electrode. Often, the reference electrode is chosen to be in an inactive region, but the reference may be recalculated later in order to isolate signals of interest more effectively. These virtual, as opposed to physical, references are generated to create reference-free or more localized recordings of neural activity (Figure 2.7). The common average reference (CAR) theoretically provides a reference-free basis by subtracting the average of all electrode potentials. Justification for this relies on the assumption that the recording electrodes comprise a surface containing all brain currents within a volume, which, for all but the densest electrode arrays, is an assumption that is generally not met. Other references, such as the bipolar or Laplacian reference, utilize pairs of closely separated electrodes to estimate with higher resolution the local potential gradient in the direction between the electrodes [22]. The surface Laplacian, or second spatial derivative of scalp potential, estimates the current
Figure 2.6: An example EEG trace corrupted by two types of muscular artifacts. The jaw clench occurs between second 124 and 125, and is best isolated in temporal channels, although it can be seen in most channels. Eye blinks, on the other hand, are best viewed in frontal channels (blue) and can be isolated with proper referencing.

density exiting through a portion of the skull. This band-pass filter is used to emphasize local sources and suppress others. Whichever referencing scheme is utilized, care should be taken in the interpretation of data, as the sensitivity to source location and orientation is affected by this choice [86]

Additional preprocessing methods are often used in bioelectric recordings to isolate data of interest before performing feature extraction and classification. Such algorithms include spatial filters, component maps, and dimensionality reduction algorithms [82]. Such preprocessing can serve a few purposes. The Independent Component Analysis (ICA) spatial filter, for example, serves to separate multichannel EEG into statistically independent components [87]. ICA is commonly used to perform artifact rejection; one could separate out an eye blink signal, discard it, and remix the components back together to achieve eye blink-free data [88]. Another useful feature of a spatial filter is for dimensionality reduction, especially for systems intended for online operation. Principal Component Analysis (PCA) is a common tool used in signal processing, as it uses the covariance structure of a multivariate data set to find a set of orthogonal components projected on a principal subspace [89]. Using PCA, one can extract a few principal components which captures majority of the variance in the data, thereby reducing the required processing with minimal loss of information. Related to PCA is the Common
Figure 2.7: (a) An example EEG trace of channel P3 colored by the type of reference used to generate it. Black – Unipolar (ear referenced). Blue – Common average reference. Red – Laplacian reference. (b) Examples of power spectra in channel C4 using the three referencing schemes, and (c), the average visually evoked potential in P3 over one run. Visual cueing stimulus occurs at time zero.

Spatial Patterns (CSP) algorithm, which is used to find the maximally different proportions of the combined variances of two sets of EEG data, e.g. left and right motor imageries [90].

**Feature Extraction**

A feature is the piece of information that is extracted from the brain signal to be interpreted by the classifier. The process can be very straightforward or highly technical, however the purpose of feature extraction is to isolate a relevant signal of interest to the task, so that a lower dimensional signal can be handed off to the classifier. The dimensionality reduction accomplished by feature selection serves two purposes: it makes the classifier more generalizable by not over fitting the highly dimensional training data, and it allows for faster online computations, especially when using complex classification schemes.

The nature of features used in BCI control are highly dependent on the recording modality and paradigm; features extracted in an EEG-based motor imagery task bear little resemblance to the neuronal firing rate codes that are used for control in the same task using intracranial recordings. For EEG-based BCI implementations, where signals
of interest are transient changes in ongoing oscillations, both time- and frequency-based features may be extracted (Figure 2.8). Substantial reviews have been compiled that pick apart the intricacies of feature calculation and selection methods [82, 83, 91].

Many of the methods, especially those related to motor imagery, rely on spectral analysis methods. These are used when the information of interest is contained in non-phase locked oscillatory changes, which can be lost by averaging trials in the time domain. Spectral analysis requires signal transformation into the frequency domain, using methods such as the Fourier and Wavelet transforms [92]. Rapid versions of these algorithms allow for time-frequency decomposition of online data, to allow for extraction of frequency and power in real time.

Features in the time domain are also useful for many BCI applications. A natural example is the extraction of evoked potential features using a system such as a P300 speller, in which case the signals of interest are timed to a specific stimulus. Simple averaging in the time domain allows for noise to be suppressed, increasing the quality of the control signal. Following averaging, features utilized have been as simple as peak amplitude or area under the curve during a period of the evoked potential. Another useful time-domain method is the extraction of band powers from motor imagery data. Band power calculations first involve filtering in the relevant frequency band and then estimating the envelope of the square of that signal to determine the approximate power of that frequency component over time [76]. This method allows for visualization of the changes in power in the sensorimotor rhythms due to movement imagination, and underlies the analysis for calculations of ERD, which is defined as deviation from baseline band power [76].

In addition to features computed directly from the time and frequency representations of individual channels, other methods compute statistics about the interactions between channels, either through cross-correlations or coherences [93, 94]. Additionally, random process models have been applied to EEG data in order estimate parameters of best fit that can be used to in classification [95].

Classification

Classification is at the heart of the brain-computer interface. This process allows the computer to make decisions about the intent of the user based on the available data. As with feature extraction, classification techniques are widely varying and data-dependent,
Figure 2.8: (a) Time-based features used in a P300 spelling task. Black line is a portion of EEG from channel Cz, red and black patches are search windows for maximum EEG amplitude for target and non-target trials. Asterisks mark the maximum amplitudes in those windows. (b) Band power features for motor imagery data during left (blue patch) and right (red patch) imagery periods. The band power is calculated by filtering the EEG signal in channel C4, squaring it, and then calculating the envelope of this signal. This provides an estimate of the time-varying power in that frequency. (c&d) Frequency-based features for the P300 and motor-imagery tasks in the same set of channels. Discriminable features are identified by differential power levels or frequency bands between classes.

with both simple and complex implementations [82, 83]. Broad categories of classifiers emerge from the literature: those based on thresholds, generation of linear and nonlinear decision boundaries, neural networks, clustering algorithms, and eclectic variations on Bayesian inference.

Linear discrimination methods are by far the most popular and have received attention due to their success in a field-wide BCI algorithm competition [96]. These include linear discriminant analysis, and linear support vector machine algorithms. Both of these algorithms utilize the covariance structure of a set of labeled training data in order to build a decision boundary that maximizes the between-to-within group variance or
the maximum margin of separation between classes, respectively.

Classifiers can operate continuously, as in the case of a motor-imagery system for driving limb prostheses, or they can operate discretely, making a decision every few seconds after sufficient data has accumulated, as in the case of the P300 speller. In the case of the former, the classifier has to choose from a small set of outcomes, in the simplest case to drive a prosthetic to the left or right. On the other hand, a P300 classifier is usually making a decision from thirty or more options. Classifiers can be built from a dedicated set of training data recorded on the same day or even from a previous session. Alternatively, adaptive methods allow for immediate classification as trials become available [97, 98]

**Feedback**

A final and necessary component of the BCI system is the generation of feedback. Feedback allows the user to be aware of their success using the system and, if necessary, how to adapt to improve their performance. The most common form of feedback is a visual output of the decision of the classifier on the computer screen, although other forms of feedback have been used. Feedback modality is especially important for users with compromised sensation. For example, users unable to focus on a computer screen may be better served receiving cuing signals and feedback by sound [99, 100]. Alternatively, those with spinal cord injury, who have motor and sensory limitations of the limbs, may not be able to use a tactile feedback system as well as someone with a primarily motor disorder.

These methods of interfacing are suited for different types of applications. P300 are more suited for spelling and environmental manipulation though some form of augmented reality interface, and they are more appropriate for discrete multi-class decision making. This type of device has proven its usefulness in clinical applications, for both patients with paralysis, degenerative disease, and disorders of consciousness [101, 102]. These systems are implemented communication platforms, which use the classification of the user’s brain signals to communicate words or actions to an effector, which can be a text document for composing an email [101], or a controller for a wheelchair [103].

On the other hand, the active modulation of an SMR BCI has more appropriate applications for continuous, graded control of low dimensional systems. With sufficient training, motor imagery offers a level of control that is most similar to motor action.
Although the spatial resolution of an EEG-BCI is not nearly as fine as invasive methods, amazing levels of control have been achieved in those who train to become proficient. For example, robust three-dimensional control of a cursor on a computer screen [104] and even a quadcopter to targets in a real-world environment [105] have been demonstrated using EEG-based motor imagery. Immersive feedback using virtual reality environments has been used to explore how control signals are modulated in life-life operating scenarios [106, 107, 108].

Other forms of feedback, primarily employed with intracranial BCI platforms, rely on real-time control of prosthetics. Users are able to gauge the progress of their training and generate rewards using a physical effector on their environment. These often employ sophisticated control systems to achieve robust control. Both non-human primates and human volunteers have utilized motor imagery along with invasive recording of multi-unit neural responses to control prosthesis to achieve remarkably natural movement [20, 109]. Invasive recordings enable recording from multiple neurons with preferred firing tuned to a specific motor modality and direction of action, and thus allow for control of multiple degrees of freedom. Motor imagery has also been applied to systems using ECoG [43]. With a grid implanted on the cortex, the imagination of individual finger action was able to be decoded using this type of system.

2.2 Amyotrophic Lateral Sclerosis

ALS is a progressive neuromuscular disease with roughly 6,000 new diagnoses in the US each year [110], with a mean onset age of 55. ALS results in progressive muscle weakening, leading to eventual loss of voluntary limb movement, inability to speak or eat, and respiratory failure. The mean duration of survival after diagnosis is three to five years [111]. ALS was first described in 1869, and more than a century later there exists a very basic and incomplete understanding of its pathologic mechanisms [112]. 5-10% of presenting ALS cases appear to have a genetic component [112], although this number may be underestimated [113]. For the remaining patients who experience sporadic ALS, the etiology is unknown. The resulting motor neuron loss appears to be a result of oxidative damage, mitochondrial dysfunction, defects in axonal transport, growth factor deficiency, and glutamate excitotoxicity. Similar changes in neighboring astrocytes may also disrupt glutamate transmission, exacerbating excitotoxicity and hastening cell death.
Even though the exact disease pathway is an area of active study, the overt manifestations encompass gross motor changes that warrant BCI intervention. However, other manifestations of the disease could also be relevant to BCI design: non-motor changes and genotype-phenotype associations.

2.2.1 Signs and Symptoms of ALS

The formal El Escorial criteria used to diagnose ALS is based on presence of progressive upper motor neuron and lower motor neuron abnormalities in multiple regions, while also ruling out other motor disorders [114]. Beyond these common criteria, the variations in disease onset, aggressiveness, and extra-motor effects are substantial. ALS onset can be bulbar, spinal, or mixed in nature, with poorer prognosis for those with bulbar onset [115]. Some patients experience primarily upper motor neuron symptoms of spasticity and brisk reflexes, while some experience primarily lower motor neuron symptoms of muscular atrophy and paralysis [116] (Figure 2.9). These variations are assessed in the clinic through the use of a standardized scale, the ALS Functional Rating Scale - Revised (ALSFRS-R) [117], neurologist evaluation, and electrophysiological recordings. The ALSFRS-R is a widely-used, 12-item, ALS-specific questionnaire assessing physical function in the bulbar, upper limb, lower limb, and respiratory domains. Each item is scored from 0 (poorest function) to 4 (normal function), and the scores are added to produce a total score from 0 to 48.

In some patients, ALS results in degeneration outside the traditional motor regions, leading to cognitive and behavioral abnormalities. Although not a defining component of the disease, behavioral and cognitive symptoms can appear together or separately in disease manifestation [118]. In rare cases, patients can meet the criteria for frontotemporal lobe degeneration (FTLD), which can result in frontotemporal dementia (FTD). The evidence for a continuum of disease that links ALS and FTD is backed by clinical, pathological, and genetic evidence [119]. Physical symptoms are a diagnostic criteria for the disease, but some behavioral or cognitive symptoms occur in 10-75% of patients, with dementia occurring in 15-41% [114].

Behavioral symptoms are coincident with cognitive symptoms in about half of patients, and with depression in a smaller portion of patients [118]. The connection between ALS and the behavioral variant of FTD (bvFTD) is evident given similarities in
grey matter changes of the anterior cingulate, temporal pole, and prefrontal cortex, as well as the underlying white matter in these regions [119]. Mental rigidity is the most common behavioral change associated with ALS. Patients can display decline in social interpersonal conduct, self-centeredness, impairment in regulation of personal conduct, apathy, irritability, emotional blunting and loss of insight [123, 114].

Many of the cognitive functions that have been shown to be compromised in ALS are associated with frontal regions that are degraded in FTD (Figure 2.9). The most common cognitive deficit is in executive functioning, relating to the ability to organize information, shift attention, and inhibit behavior [114]. A common test of executive functioning and working memory is verbal fluency, measured by intrinsic response generation and widely agreed to be affected [114, 121, 124, 125, 126, 127, 128, 129]. Other studies have also pointed to language deficits [121, 128, 130], highlighting possible limi-
itations in linguistic or semantic knowledge. This finding is controversial, as other studies have pointed to normal abilities on semantic knowledge, measured by naming and grammar tests [127, 131, 132]. Recently, another trend in research has appeared, with scientists classifying ALS patients with cognitive limitations into two main categories, those with executive dysfunction, and those with language dysfunction [128, 130]. In their analyses, some patients show reduction in performance on tasks in one or both of these realms of cognition. Deficits in language, executive functioning, and physical ability lead to changes in discourse, the content and productivity of which is altered in ALS patients. Other commonly observed cognitive deficits are attention/mental control [114, 125, 126, 127], visual and verbal memory impairments [121, 127, 129], and free recall [114, 125, 133].

Genomic risk factors have been found in both the familial and sporadic types of the disease [113, 134, 135], and provide further evidence in support of an ALS-FTD spectrum [136] (Figure 2.9). Such findings include the numerous mutations of SOD1, FUS, TARDBP, of which there are evidence for greatly different phenotypes [113]. A recently discovered genetic marker is the repeat of hexanucleotide GGGGCC in the C9ORF72 gene, which in healthy people is rarely repeated more than five times, but in 40% of familial ALS patients and 7% of sporadic cases can be repeated hundreds and even thousands of times [134]. The expansion of this gene results in approximately equal proportions of individuals with ALS and FTD [135]. The forming consensus is that a multi-genetic contribution underlies diverse pathological mechanisms which converge on the ALS phenotype.

### 2.2.2 Brain Imaging in ALS

Today, diagnosis of ALS is done independent of brain imaging methods such as EEG and MRI. However, structural magnetic resonance imaging can be used as an exclusion criteria for other neurodegenerative diseases, lesions, or myelopathies [137]. MRI has been used with much success in ALS research in realms other than diagnostic testing. There are many results which are specific to particular regions of the brain. I will not attempt to give all of the findings here, but rather an idea of the heterogeneity of structural and functional changes that occur within ALS.

Brain imaging in ALS confirms atrophy in primary motor regions, although MRI
Figure 2.10: (Left) The brain of a patient with sporadic ALS, showing atrophy of the frontal lobe [138]. (Right) Graphical representation of the changes in grey matter (Motor Cortex, MPFC = medial prefrontal cortex, Temporal pole) and their underlying white matter tracts (CST = corticospinal tract, CC = corpus callosum, ILF = inferior longitudinal fasciculus) in the ALS-FTD continuum [119].

studies produce inconsistent results [137]. An imaging finding that is consistent across MRI studies is degeneration in the corticospinal tract, measured using diffusion tensor imaging and reflecting disease severity [137, 119]. A recent study [10] confirmed pathological and physiological findings that the presentation of bulbar or limb symptoms is associated with atrophy in the corresponding brain regions, and that the level of degeneration is linear with the functional score in these areas.

There are non-motor anomalies in brain structure as well. The frontotemporal and parietal regions are additionally affected, as both grey matter and underlying white matter in frontotemporal and parietal regions display signs of atrophy [119, 137, 139]. Further evidence for an ALS-FTD continuum comes from imaging studies which show a distinct trend in degeneration in motor, frontal, and temporal areas (Figure 2.10). Those with bvFTD showed degeneration in the anterior cingulate, motor and premotor cortices, similar to the pattern observed in ALS patients. Damage to the anterior cingulate has been shown to produce the hallmark increase in apathy common to both groups [119]. Those with FTD produced greater overall atrophy in grey matter regions, including the prefrontal and temporal cortex, as well as in the striatum. ALS patients overall showed more white matter changes, especially in the corticospinal tract. A distinguishing feature in the continuum of structural changes was the spread of pathology into the anterior temporal lobe that exists in ALS patients co-diagnosed with FTD, but does not exist in patients with classic ALS [119].
The use of EEG for describing electrophysiological changes in ALS have been inconclusive, leading one to believe that gross brain activity is essentially unaffected by the disease [140]. Currently, the use of EEG for diagnostic purposes does not meet the sensitivity and specificity requirements needed to justify it as a clinical tool, and is not one that is regularly used for that purpose. While resting state changes in gross EEG do not provide a reliable marker for disease progression, event-related potentials (ERPs) measured by EEG under certain conditions has been shown to be sensitive to the cognitive status of ALS patients. Both the N200, a negative wave associated with target identification [12], and the P300 potentials arising from an oddball task of auditory stimulation have been shown to be of significantly longer latency in ALS patients [141]. These findings highlight the changes that occur in the ALS brain which affect the endogenous components of the ERP, specifically the N200 and the P300, while leaving the exogenous early components unaffected.

2.2.3 Assistive and Augmentative Communication in ALS

A product of the disease is the eventual loss of muscular coordination responsible for speech production. At some point, 80-95% of ALS patients are unable to use their own speech to communicate [142]. As the disease progresses, many people with ALS develop locked-in syndrome (LIS), which describes individuals who are awake and conscious but in a physical state of almost complete immobility and loss of verbal communication [143]. The last muscular control available to an ALS patient with LIS are the muscles of the anal sphincter and the eye [144] both of which are lost as the patient transitions to completely locked-in syndrome (CLIS). In CLIS, even eye movement has been compromised, eventually leaving a portion of the 2% of ALS patients who receive invasive ventilation support, the majority of which are unplanned, without any means of communication [145]. For this reason, patients with ALS are one of the most cited clinical target populations for BCI use [146], as these systems would remain the only plausible alternative for interpretation of intent while the patient is in a state of complete voluntary muscular insufficiency.

BCI has been researched for use as an augmentative and assistive communication (AAC) technology for those living with advanced ALS. Common low-technology AAC devices include communication boards and notepads. High-technology systems may
include computer-based speech synthesizers, as well as eye- and head-tracking systems. AAC acceptance overall in the ALS population has increased from 73% prior to 1996 to 96% by 2004 [147]. Motor-disabled patients consistently rate mobility, communication, activities of daily living, and employment as the main reasons to improve independence through the use of AAC technology [148, 149]. Those who use the devices for the longest are those who opt for mechanical ventilation [150] and continue to use them while receiving breathing support.

The majority of BCI systems have been tested in healthy individuals. However, a few studies have explored the applicability of BCI-AAC devices in patients afflicted with ALS. The first were SCP-based BCI systems explored by Birbaumer et al [11], who achieved communication with ALS patients and continued to do so after they developed LIS. Using language-based P300 systems, ALS patients have shown accuracy rates similar to or slightly below the accuracy of healthy control subjects [151, 152]. Another system which has achieved success in the ALS community is the imagery-based BCI [153, 154]. Recent results indicate that the type of successful imagery and location of disease onset are not correlated [153, 155]. However, the presence of significant bulbar involvement was found to be negatively correlated with the ability to modulate motor rhythms [155]. This could signal that behavioral alteration, which is more prevalent in the bulbar subtype [156], is partially responsible for the decrease in performance.

A caveat remains, however, as ALS patients having CLIS at the time of implementation were unable to gain control of the device [157], similar to the results found in SCP studies [11]. There is evidence that P300 speller systems may be the most reliable and rapid BCI-AAC communication devices for these patients [158], although the usefulness of this paradigm may be limited primarily to the auditory modality in CLIS [144].

### 2.3 Unsolved issues with BCIs

#### 2.3.1 Asynchronous Use

One of the biggest challenges facing deployment of BCIs for communication purposes are the cue-based paradigms by which the systems operate. The tasks the user performs, cued by the computer, are synchronized to the system that performs the decoding. This
allows for little flexibility regarding the pacing of the system, including when to send user intent to the computer. The ideal BCI would be one that is self-paced, the user controls when and how often to communicate their intent. A system that translates the signals of unknown timing into commands is called asynchronous. In practice, this means that a user has the ability to initiate a “no-control” option, during which the computer does not perform classification or generate feedback from the brain signals.

Classifying a state of unknown timing is what makes asynchronous BCI such a difficult prospect. Different approaches have been taken which dictate the timing of the system. Signals that exceed threshold [159, 160], have been used to indicate an action state, while others require the control signal be sustained, accumulating a sufficient amount of information, [161, 162]. Additionally, refractory periods on sequential classification have been used to give structure to self-paced decision making [161]. Other mediators of asynchronous control include low frequency oscillations in the 1-4 Hz range [163], motor rhythm idling states [164], as well as steady state evoked potentials [160]. Of different types of mental imageries, motor-imagery has been shown to be one of the most effective tasks at discriminating between active and idle states, while auditory imagery proved to be least separable [52]. Beta rebound found after the imagination of foot movement has been used to initiate control over a BCI [159], as well as indicate the system for intent to begin a second phase of classification in a hybrid-type BCI [161].

Perhaps the sub-field of BCI most interested in asynchronous use are the BCI-controlled wheelchairs. Often these devices are paired with a semi-autonomous guidance systems that generally integrate information about the environment to weight decisions from brain signals accordingly [103, 162, 165]. In this way, a user is able to make self-paced decisions about the direction of the wheelchair without having to worry about the avoiding obstacles or minor trajectory changes.

Work in Chapter 3 addresses a novel solution for asynchronous BCI use by utilizing supplementary brain information in the assessment of mental state. By taking advantage of previously-unused EEG signatures of attentiveness and cortical excitability, we describe an additional potentially useful mechanism for creating self-paced BCI control.
2.3.2 BCI Inefficiency and Predictors of Success

For some BCI tasks, particularly the SCP and motor-imagery paradigms, a substantial percentage of individuals are unable to gain control of the BCI even after a typical amount of training with the device. These particular paradigms are the most user intensive in terms of training, and require sustained attention to complete. Despite training, it is estimated that roughly 20% of healthy individuals fail to achieve an effective control strategy with a motor-imagery paradigm [166, 167], although estimates have also been higher [168]. The process of training on a BCI can be very time consuming and may result in frustration if the user is unable to establish control over the system. For this reason, different electrophysiological markers have been assessed for predicting performance on a BCI system over a short training period, before much time is spent training on the system. With such predictors, we hopefully are able to begin the process of personalized engineering an effective BCI system.

Different variables have been found to be predictive of BCI success, ranging from electrophysiological tests to psychological evaluations. Here we focus on predictors of performance for the SMR-BCI in particular. Decent reviews of the field are given in [169, 170]. In the latter, the authors discuss the identification of useful performance predictors, as well as the time scale of the predictor, or whether long or short term performance can be associated with the factor.

Blankertz et al. found that a combination of resting sensorimotor rhythm amplitudes were predictive of subsequent trainability on the motor-imagery task, and were able to establish a fairly strong correlation between online accuracy and their performance predictor [166]. Halder et al. found that there were initially substantial differences in fMRI activation of the supplementary motor area in high vs. low performers [171]. Structural MRI has also been able to distinguish high performing and low performing SMR-BCI users [168]. Central white matter changes, rather than gray matter changes were the relevant features for making this distinction.

Other research has focused on how performance varies within subjects over time on a trial-by-trial basis. High-gamma power in the pretrial period has been shown to correlate with SMR-BCI performance [172]; the fronto-parietal network that generates these rhythms is hypothesized to be associated with attentional processes. Training strategies for maintaining the user in a state receptive to BCI control have been proposed using classic meditation practices [173] as well as neurofeedback [170].
Earlier lines of research focused on psychological factors. “Locus of control of reinforcement”, a measure of whether perception of the result of a person’s attempt at BCI control is a result of their own performance/personality, or a result of luck, coincidence, or destiny was used to study BCI ability [174]. The authors’ conclusion was that those comfortable with technology and their power to control it generated higher accuracy with the motor-imagery task. Motivation has also been associated with device utility, both positively (feeling of challenge and confidence in mastering the device) as well as negatively (fear of failure) [158]. Both ability to concentrate on a task, and measures of visuo-motor coordination have been associated with high SMR-BCI performance [167]. Furthermore, self-rating of kinesthetic imagery quality via imagery questionnaire has correlated with performance on the SMR BCI task [175]. Interestingly, users who are experienced in meditation are also able to modulate brain rhythms more effectively [176].

Work in both Chapters 3 & 4 directly address the issues associated with BCI performance predictors and device literacy. In Chapter 3, alternate EEG predictors are assessed as a potential gating mechanisms for improving BCI performance by allowing classification to occur only during times which are most favorable. In Chapter 4, we elaborate on a possible source of device inefficiency in ALS patients, the high prevalence of mild to moderate cognitive dysfunction found in this population.

2.3.3 Limitations of the locked-in

A major caveat concerns the goal of personalizing a communication device for the severely disabled. Currently, the state-of the art of assistive communication in ALS is the eye-tracking device. Arguably, if eye function is intact, this type of system is a superior tool for intuitive, rapid communication. However, as oculomotor function declines, or gaze holding becomes too fatiguing, the efficacy and desire to use such a device among patents deteriorates [177]. The unique utility of a BCI-AAC in ALS is for those who have lost residual eye movement, which gradually occurs during the transition to CLIS.

Although the capacity to control a BCI is relatively conserved in ALS patients with LIS, no attempts at establishing communication using a non-invasive electrophysiological recording have been successful [178]. This is true even when intensive conditioning
was performed in patients who possess positive responses to passive tests of cognition [179]. Furthermore, there is no indication that invasive recording techniques produce imagery signals with greater discriminability [144], although for technological and ethical reasons, this a relatively untested BCI paradigm. Recently, marginal communication has been established in an ALS patient after having been completely locked in for two years [44]. This BCI system, based on NIRS technology, measured metabolic brain activity on a 25-second time scale to infer ‘yes/no’ responses to aurally-delivered question prompts.

These results, especially the 100% accuracy achieved during certain sessions, are highly encouraging. Lack of success with other CLIS paradigms have been attributed to technology incompatibilities [44], limitations of sensory input due to loss of ocular control [99], cycles of vigilance [178] as well as the extinction of goal-directed behavior [180]. The latter phenomenon, which is hypothesized to occur as a result of the complete abolition of motor control and feedback, represents a significant cognitive change that is hypothesized to occur in CLIS.

In the latter half of Chapter 4 we detail certain clinical factors which correlate with a decline in BCI function. In opposition to physical health, which may decline while BCI performance is maintained, psychological health impairment correlates with decreases in performance in two types of BCI systems. Although no completely locked-in patients were evaluated in our study, we elaborate on the cognitive factors that may contribute to low performance, and in Chapter 5 describe potential strategies to increase performance through personalization.
CHAPTER 3

GATING MECHANISMS FOR BCI

One of the biggest challenges for sustained BCI use is the issue of asynchronous control. BCI users are often presented with systems governed by rigid timing; the cue is presented during a fixed period and the user is given a set time to complete the task, after which classification occurs and feedback is presented. What happens when the user is focused elsewhere, unable to complete the task, or ready to stop? The need for asynchronous control is apparent, especially in cases where the user may not be able to manually input a command or turn off the device.

The following two studies were performed during the first two years of my thesis work. Although they hardly present a final solution to the issue of asynchronous timing within the BCI environment, they offer a possible mechanism for offloading more control to the user. These studies were performed in healthy, college-aged participants, and consider alternate features found in the recording, which, although ignored in the primary classification task, could serve an entirely different purpose as a gating feature. As a gating mechanism, secondary neural signals could serve as a switch to activate or inactivate the process of primary task classification, effectively setting the start and endpoints of system use, as well as the timing of trials within the operational period. In these studies we looked for gating signals reflecting states of vigilance and task readiness that are able to be extracted from the EEG as information supplementary to the primary task. The work done in processing and validating these measures in offline analysis are described in the remainder of the chapter.
3.1 Evoked Responses to Asymmetric Cueing [1]

In the case of a motor-imagery paradigm, the evoked responses to synchronization cues are often overlooked or treated as artifact. Typical stimuli used in motor-imagery paradigms are unilateral in nature [87, 95, 181]; they take the form of arrows or boxes presented to the user in a single visual hemifield, which can lead to asymmetric visually evoked potentials (VEPs) in the recorded EEG. Substantial literature on VEPs describes the hemispheric asymmetries in evoked potentials recorded from the scalp that occur as a result of imbalanced visual stimuli. Subjects exposed to a unilateral stimulus exhibit shorter latencies for the characteristic positive and negative evoked potentials over the visual cortex contralateral to the stimulus [182]. These potentials, denoted as P1-P3 and N1-N3, occur in the 30-300 ms following flash stimulus presentation to the foveal retina. The increase in transmission time is explained anatomically by signal delay across the corpus callosum [183]. The implications of this phenomenon have yet to be utilized in a BCI paradigm. In this study, we assessed whether the VEPs due to cue asymmetry be used as part of a hybrid BCI as a priming system for attention detection in an imagery task.

Data were recorded from 5 subjects (ages 18-28, two female) in this study. Nineteen EEG channels were arranged on the scalp according to the standard 10-20 system, referenced to linked earlobes. Additionally, signals were recorded from four EOG channels for the purpose of artifact correction. Signals recorded from the scalp were amplified and bandpass filtered between 0.1 and 30 Hz using a commercial EEG recording system (Guger technologies, www.gtec.at). Signals were sampled at 256 Hz. The study was conducted in accordance with guidelines approved by the Institutional Review Board of Penn State University.

In order to assess the effect of unbalanced vs. balanced stimuli with minimal additional factors, we used three different types of arrow cues. The unbalanced cue extended from the center of a fixation cross to the periphery. In contrast, the balanced cues were centered in the middle of the cross, using only the directionality of the arrowheads to indicate cue type. For these long and short balanced arrows, their ends continued to the periphery or halfway to the periphery of the screen. These visual stimuli were the cues used to direct subject intention. Each subject performed four motor imagery sessions lasting approximately one hour each. Each session comprised four to five runs. Each
run consisted of 50 trials, and the presentation of the cues in each run was randomized. In each trial, the subject performed five seconds of imagery of hand movement on the side indicated by the cue.

Raw data were first run through a regressive EOG correction algorithm [184] to remove artifact associated with the activity recorded from the four EOG channels. Following this, trails were rejected if the amplitude of the EEG in any of the channels exceeded 60 mV. Artifact control resulted in the removal of 0%, 4.8% 25.5%, 13.8%, and 0% of total trials for the five subjects. Data corrected for artifact were used in subsequent analysis.

Hemispheric evoked potential asymmetry can be measured with the event-related lateralization (ERL) statistic, as in (3.1) [185].

\[
\text{ERL}(O_1,O_2) = \frac{VEP_{O_2}(L) - VEP_{O_1}(L) + VEP_{O_1}(R) - VEP_{O_2}(R)}{2}
\]  

This equation gives the event-related lateralization across channels O1 and O2 in the visual cortex in response to the mean VEP evoked by left (L) and right (R) cued trials. Significant lateralization was found in four of the five subjects who participated in the study. The period of lateralization was highly stereotyped, occurring with a negative peak at 200 ms and a positive peak around 300 ms (Figure 3.1). The interval of significant lateralization due to cue type was found to be 203-305 ms after cue presentation, as determined by an ANOVA \((p < 0.01)\). Post hoc paired significance tests confirmed that the ERL evoked by the unbalanced cue was found to be significantly different from both short and long balanced cues. The common finding across subjects was that ERL occurs due to latency discrepancies in two major evoked potential peaks. Because the contralateral signals consistently lead the ipsilateral ones, the ERL during the negative portion of the evoked potential has a negative sign, and during the positive component yields a positive lateralization value.

The expected consequence of this hemispheric lateralization is an increase in the accuracy of the classifier before the sensorimotor modulations produced by motor imagery are expected to occur. Indeed, this increase was observed in all subjects who demonstrated cortical lateralization, and was found roughly 200-700 ms following cue presentation (Figure 3.1). This classification increase occurs uniquely for the unbal-
balanced cue because the evoked potentials due to each of these cues contain hemispheric timing differences which are not present in the case of balanced cues. This finding motivates a few things for future motor-imagery BCI work. In order to classify intent of the motor actions of the user and not an evoked artifact of the stimulus, a balanced cue should be used. Alternatively, the data from the initial period following cue presentation should be discarded in classification of the imagery-task.

This alone is not a new result for the VEP or BCI communities. However, this result served as an impetus to rethink the purpose of the cuing mechanism. The arrow stimulus in a BCI training paradigm serves to direct the intent of the user, not create spurious classification features independent of intent. In a self-generated feedback task with no cue, the user is in control of the direction of the imagery. In this case of free form control, the stimulus could be re-purposed as a probe for assessing user attention. Given the user is attentive to the task, the positive classification of an evoked potential may serve to gate the continuation of the remainder of the trial. Otherwise, the trial may be aborted or the system inactivated. With this goal in mind for the cue, the stimulus must be chosen so that it either evokes a strong and consistent VEP for an individual trial, or possibly even a strong lateralization of the VEP, so that real time assessment of task vigilance may be made.

Although this work did not produce any groundbreaking results in asynchronous BCI, it focused our goal of using visual cues for purposes other than just directing the user. In the following study, we consider additional gating variables for their predictive value for success in an imagery task, so that we may use these features to enable user-regulated timing of a motor imagery interface as part of a two-stage BCI system.

Figure 3.1: (Top) Aggregate ERL curves for all five subjects due to unbalanced as well as short and long balanced cues that were presented at time 0. (Bottom) Classification accuracies for detecting left vs. right trials using evoked potential traces for each type of cue.
3.2 Use of single trial gating signals to optimize motor-imagery BCI [2]

3.2.1 Introduction

As was described in Section 2.1, the imagination of motor action is a common mental task used for delivering intent without overt action, and is reflected as changes in the SMRs over motor and premotor cortical areas. Success in controlling an SMR-based BCI system depends on the user’s ability modulate these motor rhythms from a baseline state. Little is known about why some individuals are able to do this well and why others are not, but one such predictor, the level of baseline SMR amplitude, has been shown to be positively correlated with the success of the individual in completing the task [166]. BCI performance has also been demonstrated to be associated with pre-trial gamma band (70-80 Hz) amplitude [172].

EEG oscillations such as the SMRs have been given roles as modulatory gating mechanisms for information transfer between the cortex and subcortical structures such as the thalamus [12], and may explain the dependence of imagery performance on these rhythms. Using transcranial magnetic stimulation (TMS), low amplitude oscillations were shown to be representative of an excitable state of the motor cortex [186]. These findings indicate that the regulation of SMRs, specifically the mu rhythm, is likely to condition the cortex for forthcoming perception or action [187]. A BCI that relies on changes in mu rhythm may be affected by pre-trial mu amplitude simply because high mu before the trial offers opportunity for larger decreases from baseline level. However, the ability to form motor imagery may also be modulated by oscillations representing motor network excitability.

Early work by Bishop demonstrated that the phase of ongoing cortical oscillations partially predicted the seemingly random variations in evoked potentials due to visual stimulation [188]. This demonstration led to the hypothesis that these oscillations represented cyclical cortical excitability. Later studies showed how these periods of excitability could be demonstrated by consistent changes in perceptual or motor thresholds. Phase and amplitude of the occipital alpha at the time of visual stimulus presentation has been shown to be predictive of stimulus detection [12, 187]. Somatosensory cortical evoked responses to painful stimuli are also facilitated by the amplitude of 8-
13 Hz mu motor rhythms [189]. Further evidence comes from the work of Kruglikov et al. [190], who showed that auditory evoked potential morphology is a function of broadband EEG phase.

In addition changes in excitability resulting from oscillatory activity, evoked responses to varying stimuli have been shown to be modulated by region of fixation [191, 192], and serve as a marker of visual attention. Visual attention is controlled by a distributed network of cortical and subcortical areas which act to provide “bias signals” that enhance or suppress the responses to visual stimuli [192]. The distinction between fixation and attention in the context of a BCI is an important one. In a cued BCI paradigm, attention is required for the subject to understand the cue and react to it, but fixation on that cue is not absolutely necessary [191]. Of the components of the VEP, the N1, N2c, and P3b peaks have been shown to be influenced by visual attention [12, 191, 193]. Increases in amplitude of the first two components and a decrease in the latency of the P3b have been shown to be correlated with visual attention.

The aim of this study was to identify intra-individual predictors of single trial BCI success. By finding brain signatures which correlate with improved performance, we can effectively increase the accuracy and bit rate achieved with the interface. One possible way to do this is by gating trials with low predicted performance. We evaluated the utility of three of these brain signatures as gating variables: the amplitude and phase of ongoing motor rhythms, and the visually evoked potential produced in response to the cuing mechanism. These were chosen because of the evidence for these features to associate with motor-readiness and task vigilance, respectively. We show that these peri-stimulus features of the EEG, which are generally discarded during motor-imagery feature selection, are correlated with performance in a subset of subjects and can be used to boost the information transfer rate of a BCI communication system.

### 3.2.2 Methods

**Experimental Setup**

A commercial EEG recording system (Guger technologies, www.gtec.at) was used to acquire data from subjects. Data were sampled at 256 Hz and band pass filtered at 0.1-30 Hz. This bandpass range was chosen to preserve VEP components and oscillations in the mu band. Data were recorded within the Simulink environment in MATLAB and
stored on a notebook computer (Dell Latitude E6400) running Windows XP. Subjects were seated in a chair facing an LCD monitor which displayed cuing and feedback information. The experimental protocol was approved by the Institutional Review Board of Penn State University.

**BCI Paradigm**

Nine right-handed volunteers, all male with ages ranging from 18-37, participated in a cue-paced, one-dimensional center-out motor imagery task. Channels FC3, FC4, C5, C3, C1, C2, C4, C6, CP3, CP4, P5, P3, P1, P2, P4, P6, PO3, and PO4 were recorded, in addition to three EOG electrodes placed on the lateral canthi as well as just above the nasion. All channels were referenced to linked earlobes, and ground was placed at Fpz. Each subject performed four sessions over a two week period. Each session lasted approximately 1.5 hours.

During each session, the subject performed five runs of 60 trials each, divided equally between left, right, and no-target cues that were presented in a randomized sequence. The first run of each session, the training run, was used to train the classifier so the remaining four testing runs could be used to give feedback to the subject as they performed the task. In each trial, the subject was cued by an arrow pointing in the left or right direction. Arrows were displayed symmetrically in the visual field to minimize asymmetry in the evoked potential due to cue type, which in Section 3.1 was shown to falsely contribute to BCI classification accuracy independent of user-driven modulation [1, 194]. Subjects were instructed to perform imagery of an object-oriented grasping action for the hand corresponding to the direction of the arrow being displayed. If no arrow appeared, the subjects were informed to relax and were given no visual feedback.

In each trial of the training run, a fixation cross would first appear at which time the subject would relax. If this were followed by an arrow cue one second later, the subjects were instructed to perform imagery while the arrow remained on screen for two seconds. Each trial was followed by a random inter-trial period of 1-2 seconds. During the testing runs, a target appeared on the far side of the screen in the direction of the arrow, and during the two seconds in which they performed imagery, a cursor moved on the screen to provide the subject with feedback.

Feedback served as an indicator to the subject how closely their EEG signals of intent matched templates for left and right cues developed in the training run. This
was done by summing the squared distance $D$ of the band power in trial $n$ from the $m \in \{\text{left, right}\}$ template band power over all time points $t$ and features $f$ (3.2). Each template $\mu$ consisted of mu and beta band powers from channels C3, C4, P3, and P4 during the period of motor imagery, making up $f = 8$ total band power features.

\[
D_{n,m}(t) = \sum_{f=1}^{8} (x_{n,f}(t) - \mu_{m,f}(t))^2.
\] (3.2)

Here, $x_{n,f}$ represents a band power feature from a single trial. The feedback at each time point was calculated as the log ratio of the left distance over the right distance (3.3).

\[
Feedback_n(t) = Feedback_{n}(t - 1) + \log \left( \frac{D_{n,\text{left}}(t)}{D_{n,\text{right}}(t)} \right).
\] (3.3)

This control algorithm drove the cursor to the left when this ratio was less than one, and to the right when the ratio was greater than one. Feedback was accumulated over the length of the trial and then reset to the center of the screen at the beginning of the next trial. $Z_n$, the success of the subject at performing imagery during trial $n$, was defined according to (3.4),

\[
Z_n = \begin{cases} 
Feedback_n(T), & \text{if } \text{Cue} = \text{right} \\
-Feedback_n(T), & \text{if } \text{Cue} = \text{left}, 
\end{cases}
\] (3.4)

where the time index $T$ marks the end of the trial. Trials were classified as successful if at the end of the imagery period the cursor was on the same side of the screen as the target. Over the four sessions, each subject completed 16 test runs consisting of 960 trials, with 320 trials each for left, right, and no-target cues. The data belonging to the 320 trials of left and right cues were analyzed offline using the methods described below.

**Preprocessing**

**Artifact correction**

The first step in offline data preprocessing was removal of eye-related artifacts that resulted from blinking or eye movement. This was a two-step process including artifact reduction and trial rejection. Artifact reduction was accomplished by linear regression
This least-squares method assumes the linear superposition of neural and artifact sources to produce the measured signal. Assuming the independence of the artifact sources and the neural sources, data can be used to find a weight matrix, which represents the projection of noise sources onto neural sources. We used the 18 EEG channels as our recorded signal $Y$, and the three EOG channels as the noise sources $U$, to find the weight matrix $W$, and solve for decontaminated neural sources $S$ (3.5).

$$Y_{18 \times T} = S_{18 \times T} + W_{18 \times 3} \cdot U_{3 \times T}.$$ (3.5)

Here, $T$ is the length of the data segment from which the weight matrix is calculated. In practice, this method may be suboptimal if there is significant leakage of the task-relevant EEG data into the designated noise channels, rendering the assumption of independence void. To minimize this effect, we solved for $W$ using data which was sampled from when the EOG channel exceeded 75 $\mu$V, as in the case during a blinking event.

Following artifact reduction, trials were rejected if channel FC3 displayed absolute amplitude of greater than 50 $\mu$V. To control for artifact in unintentional movements, bipolar electrodes were placed on the forearms of a subset of the subjects to record EMG muscle activity during the recording session. EOG and EMG were both analyzed offline to rule out possible contamination from overt eye and arm movements.

**Extraction of gating variables**

*Amplitude and phase of peri-stimulus mu*

Data in channels C3 and C4 were filtered with a Laplacian spatial filter to localize the mu rhythms specific to the motor cortex [195]. Peak mu frequency in these channels was found for each subject.

Figure 3.2: Phase and amplitude extraction of a single EEG trial. (a) Trial $n$ of EEG data. (b) Signal was filtered in the subject-specific mu band. (c) The phase of the band-limited signal at $\tau_n$, the time of cue presentation, is marked by the vertical dashed line. (d) Amplitude of the band-limited signal.
using the multitaper spectral analysis method [196]. The EEG data in these channels were filtered using a zero-delay filter in a 2 Hz range around the peak mu frequency. A Hilbert transform was applied to this band-limited signal to generate the analytic signal \( S_a(t) \), comprising a real part \( S(t) \) made up of the original data, and an imaginary part \( H(t) \), its Hilbert transform (3.6),

\[
S_a(t) = S(t) + iH(t). \tag{3.6}
\]

The amplitude of the signal \( A(t) \) was calculated as the Euclidean norm of the real and imaginary parts of the analytic signal (3.7).

\[
A(t) = \sqrt{S^2(t) + H^2(t)}. \tag{3.7}
\]

The instantaneous phase \( \phi(t) \) was calculated as the four quadrant inverse tangent of the ratio of the imaginary part of the analytic signal to the real part (3.8). Phase ranged from \(-\pi\) to \(\pi\).

\[
\phi(t) = \arctan\frac{H(t)}{S(t)}. \tag{3.8}
\]

For each trial \( n \), the amplitude and phase of mu in C3 and C4 at the time of cue presentation were extracted (Figure 3.2). Mu amplitude was the first gating variable. Linear regression between trial success and amplitude was used to determine association. We assumed that the success at the end of the trial was a linear function of the natural-log-transformed peri-stimulus mu amplitude (3.9). Log transformation was performed on the spectral features to enforce the normality of the data as well as reduce outlier effects [197].

\[
Z_n = \alpha \times \ln(A(\tau_n)) + \beta. \tag{3.9}
\]

In this equation, \( \alpha \) and \( \beta \) are the parameters of the regression, and \( A(\tau_n) \) is the magnitude of the mu rhythm at the time of the cue of trial \( n \). We used the slope of the regression \( \alpha \), as the measure of correlation between mu amplitude and trial success.

Mu phase was the second gating variable. We assumed trial success was a cosine function of the phase of the mu rhythm.

\[
Z_n = \gamma \times \cos(\phi(\tau_n) - \delta) + \varepsilon. \tag{3.10}
\]
Figure 3.3: Construction of the VEP template from phase-corrected left cue trials, channel PO4-FC4, subject 7. (a) A single trial of VEP data. (b) Phase-matched correction by subtraction of a non-cued trial with similar phase (black line) results in the phase-corrected trial (red line). (c) Uncorrected trials of the same phase group and their average. (d) Corrected trials of the same phase group and their average. (e) All eight phase groups and the average resulting VEP (black trace). Regions with thick horizontal lines indicate times during which the data in the phase bins were significantly different as determined by ANOVA between the 8 phase groups (p<.05, Bonferroni corrected). (f) The final template for this subject is the average of the phase corrected trial groups (black line).

Here, $\gamma$, $\delta$, and $\epsilon$ are regression parameters specifying the amplitude, phase, and offset of the fitted cosine, and $\phi(\tau_n)$ represents the phase of the mu rhythm at the time of the cue of trial $n$. We specified the frequency of the cosine to be one cycle per 360°of phase. To find the parameters of this cosine model, we performed a non-linear regression using the MATLAB function nlinfit.m. The correlation measure between mu phase and trial success was the amplitude of this fitted curve, $\gamma$.

**Match to VEP template**

The final gating variable was related to the quality of the VEP. In order to define
Figure 3.4: Uncorrected (left) and phase match corrected (right) VEPs for all subjects. This represents for all subjects what are panels (e) and (f) for subject 7 in Figure 3.3. Colored lines are the eight phase groups, and the black trace is the resulting average VEP. Regions with thick horizontal lines indicate times during which the data in the phase bins were significantly different as determined by ANOVA between the 8 phase groups (p<.05, Bonferroni corrected).
whether the subject produced an evoked response during a single trial that represented good fixation, we needed to define a template that represented a typical VEP waveform for that subject.

Following 0.1-30 Hz bandpass filtering, we first performed phase-matched control trial correction, following the methodology defined by Kruglikov [190]. This was done because the phase of EEG at the time of the cue presentation introduces an averaging bias. Individual trials were corrected using phase-matched control trials, or trials in which a similar alpha phase was evident but in which there was no stimulus presented. By subtracting out a signal with similar phase, we remove the predominant alpha rhythm and are left with a trace that better reflects the underlying neural correlate of the VEP (Figure 3.3). Such a procedure removes peri-stimulus bias by phase of the evoked potential in all subjects (Figure 3.4).

Once corrected for phase, templates for left and right cues were defined by averaging EEG data over trials corresponding to each cue type for bipolar channels PO3-FC3 and PO4-FC4. This referencing scheme was the closest approximation to the standard Oz-Fz [198] we could accomplish with our electrode montage. The timing of the template ranged from 100 to 400 ms after cue presentation.

A metric describing the match of individual trials to the template was created using a matched filter approach. This technique comes from radar communication as a method for maximizing the probability of detecting a target waveform in the presence of Gaussian noise [199]. Although a derivation can be found elsewhere [200], the optimal filter for maximizing the signal to noise ratio of a signal generated by a linear, time invariant system with added Gaussian noise is the time-reversed version of the transmitted signal. This time-reversed template is called the matched filter. We applied the template as the transmitted signal \( s(t) \), and the phase-corrected single trial data as \( r_n(t) \). The output of the matched filter \( y_n(t) \) was the convolution of the trial data with the time-reversed template (3.11). The output of the matched filter that we used was the central value of this convolution, normalized to the magnitude of the template (3.12). This value, which we called the matched filter value (MFV), was related to the signal to noise ratio of the VEP, and served as the third gating variable. Trials with large MFV were interpreted as being trials which the subject produced a robust VEP.

\[
y_n(t) = r_n(t) * s(-t). \tag{3.11}
\]
Here, $*$ represents the convolution function, and $y_n(0)$ is the central value of the output of the convolution. The magnitude of the template $s$, in this case the Euclidean norm, is denoted by vertical brackets.

Similarly to the previous gating variables, we determined the relationship between the MFV and success on the BCI task. Again, we assumed that the trial success was a linear function of the MFV (3.13).

$$Z_n = \zeta(MFV_n) + \eta.$$  

(3.13)

The slope of this regression, $\zeta$, reflects the correlation between of the outcome of the trial and the match to a VEP template of good fixation.

**Permutation testing for significance**

The three gating variables were tested for significant correlation with trial success in each subject. Because the dependent variable, the trial success, was not normally distributed, we chose to perform a non-parametric permutation test to determine the statistical significance of the regression parameters. We first computed the test statistic $Q_{obs}$, the slope of the linear regression ($\alpha, \zeta$) or the amplitude of the fitted cosine ($\gamma$). Then we shuffled the $Z$ success outcomes between the trials and used the same fitting procedures to calculate the permuted statistics, $Q(k)$ for each permutation $k = 1\ldots K$, where $K = 1000$.

This permutation was performed with each gating variable from EEG features in two channels. Correction for multiple comparisons was done for each hypothesis test in the following manner, as described in more detail in [201]. For each permutation $k$, both the maximum $Q_{max}(k)$ and minimum $Q_{min}(k)$ values were chosen from the two $Q(k)$ statistics evaluated for both channels. This resulted in $K$ values making up each of the maximum and minimum empirical distributions of the randomized data set. The calculation of p-statistics, the probability of $Q_{obs}$ belonging to the null distribution of the permuted set, is given in (3.14).

$$p_{high} = \frac{\sum_{k=1}^{1000} (H(Q_{max}(k) - Q_{obs})) + 1}{1000 + 1}.$$  

(3.14)
Figure 3.5: This flowchart shows a generic example of the permutation test for determining significant relationships between gating variables and trial success. 

1. Test data from left and right trials are sorted using the rankings determined by each gating method. Shown here, the trials are ranked according to mu amplitude at the time of the cue, $A(\tau_n)$. Trials with high $A(\tau_n)$ are shown on the bottom of the block. The feedback at the end of these trials is also shown, with left trials in blue and right trials in red. Feedback was subsequently simplified to trial success, $Z$. 

2. The observed test statistic, $Q_{obs}$ is the slope of the regression that relates mu amplitude to trial success. The permuted test statistic $Q(k)$ was computed $K=1000$ times after permuting the success scores of the trials to $Z_{k,n}$. 

3. The $k^{th}$ value for the maximum and minimum null distributions $Q_{max}$ and $Q_{min}$ were found from $Q(k)$ across both channels tested. Finally the p-value of the test statistic distribution is computed as the percentage of values in this null distribution exceeding that of the observed test statistic.

Here, $H$ is the Heaviside function. Ones are added to the numerator and denominator to include the $Q_{obs}$ in our null distribution. $p_{low}$ was also calculated similarly using $Q_{min}$ to test for significance from both tails of the null distribution. Findings of either $p_{high}$ or $p_{low}$ less than 0.025 are significant deviations from chance. The generic process of trial ranking and permutation testing for significance is shown Figure 3.5.

**Trial gating simulation**

To simulate the effect of utilizing the potential gating variables online, we computed how the accuracy and bit rates would change for each subject if a portion of the trials having
the least predictive value for task completion were removed. This predictive value was determined from the output of a linear discriminant analysis (LDA) classifier, which had been trained with the gating variables as the input and trial success as the outcome variable. Trials in the test set which were classified as having the lowest predictive value were gated, meaning they were stopped before the imagery period began. The simulation was repeated at different thresholds, from gating no trials up to gating 70% of the total trials. Gated trials were skipped 500 ms after stimulus presentation and were interpreted as no-decision while taking 1.5 seconds (1 second before cue and 0.5 seconds after) to complete. Allowed trials resulted in a decision while taking four seconds to complete. Accuracy was defined as the number of allowed trials performed correctly over the total number of allowed trials. 5-fold cross validation was used to find an average of the accuracy of the gating procedure at each threshold. The complete cross validation procedure was completed 50 times for each gating scenario and the accuracy was taken as the average. The bits per trial, $B$ was calculated for an $N = 2$ choice task having accuracy $P$ following the form of Wolpaw et al. (3.15) [8].

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{N - 1}.$$  \hfill (3.15)

The bit rate $BR$ in bits/sec was calculated using (3.16).

$$BR = B / T_\mu,$$  \hfill (3.16)

where the average trial time, $T_\mu$, was increased to represent a penalty for disposing trials. $T_\mu$ was calculated from the number of allowed trials, $N_a$, and the gated trials, $N_g$, using (3.17).

$$T_\mu = \frac{(4N_a + 1.5N_g)}{N_a}.$$  \hfill (3.17)

### 3.2.3 Results

**BCI performance**

The operation of the BCI was successful for eight of the nine subjects, who achieved classification accuracies for a left vs. right imagery task of 64.1-97.8%. Subject 2, with an accuracy of 49.5%, indicative of random BCI control using this method, was not
Figure 3.6: Three EEG features were considered as potential gating variables (shown here are the right hemisphere features): mu amplitude in C4 at stimulus presentation (left column), mu phase in C4 at stimulus presentation (middle column), and match to PO4-FC4 VEP template (right column). For each feature, all recorded trials were divided into eight groups, from low feature value to high, represented by the eight colored lines/points. The first row is the average of the EEG in these groups. The second row is the average mu suppression from baseline in C4 in these groups. The last row is the mean and standard error of the success across all trials by group.

considered successful at operating the BCI system and was omitted from analysis in this study. LDA with 10-fold cross validation was performed for features derived from EOG and EMG channels. Mean classification accuracy for all subjects using EOG features was 50.6 \pm 3.8\% and mean classification accuracy for the four out of 8 subjects with recorded EMG features was 48.0 \pm 2.1\%. Bounds are standard deviations.
Table 3.1: Results of the permutation test with the three gating variables

Values in the six columns represent value of $Q_{obs}$, or the slope/amplitude parameters of the three regressions in two channels. Bolded values indicate a significant $Q_{obs}$ at the 0.025 level.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Amplitude ($\alpha$)</th>
<th>Phase ($\gamma$)</th>
<th>MFV ($\zeta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel</td>
<td>C3</td>
<td>C4</td>
<td>C3</td>
</tr>
<tr>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.012</td>
<td>0.013</td>
<td>0.010</td>
</tr>
<tr>
<td>2</td>
<td>0.003</td>
<td>0.008</td>
<td>-0.016</td>
</tr>
<tr>
<td>3</td>
<td>0.016</td>
<td><strong>0.044</strong></td>
<td>0.006</td>
</tr>
<tr>
<td>4</td>
<td>-0.017</td>
<td>0.003</td>
<td>0.013</td>
</tr>
<tr>
<td>5</td>
<td><strong>0.029</strong></td>
<td><strong>0.055</strong></td>
<td>-0.012</td>
</tr>
<tr>
<td>6</td>
<td>-0.001</td>
<td>0.011</td>
<td>-0.007</td>
</tr>
<tr>
<td>7</td>
<td>0.013</td>
<td><strong>0.027</strong></td>
<td>0.021</td>
</tr>
<tr>
<td>8</td>
<td><strong>0.018</strong></td>
<td>0.012</td>
<td>0.015</td>
</tr>
</tbody>
</table>

Mu amplitude is predictive of trial success

Trials were ranked according to the three gating variables of mu amplitude, mu phase, and MFV. Figure 3.6 shows the grand average results for the study, in which trials are subdivided into eight groups based on ranking by each gating variable. Average EEG traces for each group are shown in subplots (a), (d), and (g), while mu suppression from baseline is given in the second row in subplots (b), (e), and (h). For groups ranked by baseline mu amplitude, there is a significant difference in the level of suppression during imagery (Figure 3.6b, black bar indicates ANOVA with $p < .05$, Bonferroni corrected for 768 time points). MFV shows positive association between gating variable group and trial success, mu phase shows no association, and the positive correlation between trial success and mu amplitude is significant (Figure 3.6c, $r^2 = .86$, $p < .001$). Positive correlations between trial success and both mu amplitude and MFV are also evident in the left hemisphere channel C3.

On an individual basis, the permutation test procedure determined that subjects 3, 5, 7, and 8 produced in at least one channel a slope of regression, $\alpha$, which was significantly greater than that produced by the random null distribution, indicating a significant positive correlation between mu amplitude and motor-imagery task success (Table 3.1). Non-linear regression produced no significant cosine fits between mu phase and trial performance. Finally, only subject 8 displayed a significant correlation between the
Table 3.2: Simulated gating results

Simulated gating using the three gating variables. Columns Acc<sub>o</sub> and BR<sub>o</sub> give the accuracy and bit rate for the online operation of the device. BR<sub>max</sub> gives the 95% confidence interval of the highest bit rate achieved when a percentage of trials up to 70% were gated 500 ms after cue presentation over the 50 iterations of the simulation. Bolded ranges mark where a significant improvement in bit rate was achieved by the simulation.

<table>
<thead>
<tr>
<th>Subject</th>
<th>SMR amplitude</th>
<th>SMR phase</th>
<th>MFV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acc&lt;sub&gt;o&lt;/sub&gt; (%)</td>
<td>BR&lt;sub&gt;o&lt;/sub&gt; (bits/min)</td>
<td>BR&lt;sub&gt;max&lt;/sub&gt; (bits/min)</td>
</tr>
<tr>
<td>1</td>
<td>80.9</td>
<td>4.45</td>
<td>(4.49-4.57)</td>
</tr>
<tr>
<td>2</td>
<td>76.1</td>
<td>3.10</td>
<td>(3.06-3.07)</td>
</tr>
<tr>
<td>3</td>
<td>86.7</td>
<td>6.51</td>
<td>(6.52-6.59)</td>
</tr>
<tr>
<td>4</td>
<td>69.4</td>
<td>1.67</td>
<td>(1.65-1.67)</td>
</tr>
<tr>
<td>5</td>
<td>64.1</td>
<td>0.87</td>
<td>(1.40-1.46)</td>
</tr>
<tr>
<td>6</td>
<td>67.5</td>
<td>1.36</td>
<td>(1.40-1.43)</td>
</tr>
<tr>
<td>7</td>
<td>76.9</td>
<td>3.31</td>
<td>(3.24-3.30)</td>
</tr>
<tr>
<td>8</td>
<td>97.8</td>
<td>12.70</td>
<td>(12.77-12.83)</td>
</tr>
</tbody>
</table>

MFV gating variable, with high values of the gating variable in both channels PO3-FC3 and PO4-FC4 corresponding to greater trial success.

**Trial Gating Simulation**

Table 3.2 gives the 95% confidence interval of the bit rates achieved by each subject in the offline simulation of gating with each of the three potential gating features. In most instances where the permutation test indicated a subject would benefit from gating trials of low value, the simulation of offline gating yielded an increased bit rate. Bolded values show where the mean bit rate achieved by the 50 simulation repetitions was significantly larger than the original bit rate (t-test, p < .05). Again, amplitude of the SMRs appears to be the variable most suited for gating the motor-imagery system. Figure 3.7 gives an example of how the accuracy and bit rate change for each subject as a function of the number of trials that are gated based on mu amplitude. On an individual basis, even in this small subject population, there was one subject for which phase gating could have improved the communication rate, and one in which MFV gating was beneficial.
Figure 3.7: Average accuracy and bit rates from 50 repetitions of the gating simulation.

a) Average accuracy achieved for each subject after random gating of trials up to 70%.
b) Resulting bit rate. c) Average accuracy achieved for each subject after gating of trials up to 70% based on mu rhythm amplitude at the time of the cue. d) Resulting bit rate.

3.2.4 Discussion

Prospects for gating features

In BCI, single trial identification of user intent is a difficult task; nevertheless, we have shown that ongoing oscillatory activity of the motor cortex contains additional information that can be used to produce a modest but significant increase in motor-imagery task performance. Of the three features explored, the technique of selecting for trials with high mu amplitude was the most consistently useful to limit decision making to times when users are most likely to perform higher quality imagery.

The relationship between oscillatory behavior influencing motor excitability is controversial [202], but the evidence is stronger for cortical excitability changes to be correlated with SMR amplitude than with phase. The lack of correlation between the phase of the mu rhythm and subsequent suppression efficacy is in agreement with other studies which found that phase had little effect on mu suppression [202]. Whereas the phase at the time of cue presentation did have an effect on early evoked potential morphology (Figure 3.3e), it is not surprising that the effects do not propagate 1-3 seconds after the cue, during which mu suppression is evident. On the other hand, all subjects displaying significant correlation between mu amplitude and BCI task success performed better when the amplitude of mu was high over channels C3 and C4 at the beginning of the
This could be due to two factors. Our classifier was based on reduction of mu from a baseline level to an attenuated level, and as a result, trials with larger baseline mu amplitude have a greater potential for mu power reduction. In addition, the amplitude of the mu rhythm reflects the state of excitability of its generating network [202]. While our finding is in disagreement with studies that point to a desynchronized cortex as one primed for motor output [186, 202], Blankertz et al. showed that a successful predictor of imagery performance across subjects was an increased resting state mu rhythm [166]. In a follow up study, they showed that this performance increase was associated with larger recruitment of motor and premotor regions. They concluded that recruiting more synchronized neurons for motor imagery led to higher resting mu amplitude as well as higher performance [203]. In our study we show that not only is this relationship between resting mu amplitude and imagery performance valid between subjects, but also on a single-trial basis within subjects.

The third tested gating variable did not produce enough evidence to warrant its use as a potential metric for predicting motor-imagery success. Risner et al. showed that, following phase-matched correction of VEP data using unstimulated controls, average VEPs in different phase groups display the same VEP morphology [204]. We also found this to be the case for early evoked potentials, although after phase-matched control trial correction, there was some phase dependence in later evoked potentials in certain individuals (Figure 3.4, subjects 3 & 7). Because the MFV was calculated from phase-corrected VEP data, the value of this metric can be interpreted to be free of oscillatory alpha bias and instead can be associated with attention to the visual task. However, using a subject-specific VEP template to define the MFV, we were unable to find a consistent correlation between this marker of attention and the subsequent modulation of motor rhythms. Although subject 8 did display a relationship between high MFV and good performance, overall there was no group-wide trend to support this result. This has two important implications for BCI research. The first is that this type gating may provide performance improvements on an individual level. Notably, this gating mechanism significantly improved the already high bit rate of the highest performing BCI user in this study, possibly by gating the rare trials in which the user was not attentive to the task. The second is that fixation to the visual cue is not critical for task success in most users. For these users, control over the BCI communication device should not be limited by
their ability to directly fixate with cues on the screen, a critical allowance in the case of severe oculomotor impairment.

**Trial Gating Simulation**

The results of the gating simulation closely matched the results of the permutation test. Those subjects who performed better during high mu amplitude trials also achieved the greatest benefit from gating trials with low mu amplitude, with the exception of subject 7. However, more subjects benefited from gating than indicated by the permutation test. This may be due to the fact that a multivariate classifier was used to rank the trials as opposed to the permutation test run on each individual channel. The strong improvements seen in subjects 5 and 6 are also partially due to the low initial accuracy and bit rate of online BCI operation. This type of gating holds the biggest potential for improvement for subjects who do not regularly perform at a high level. For high performing users, the increase in accuracy needs to be substantial for this method to be useful.

**Study Limitations**

Although the oscillation of cortical rhythms is associated with the depolarization and hyperpolarization of large groups of neurons, at the level of EEG the linkage with motor cortical excitability may be tenuous; at this level of measurement from scalp, both excitatory and inhibitory neurons contribute to oscillatory behavior [205]. Consequently, broadly recorded signatures of rhythmic activity may be unable to describe excitability in a local group of motor neurons responsible for hand imagery. This is especially true when we attempt to use the excitability of the motor region as a predictor of SMR modulation that occurs on the order of 500-1000 ms later.

The lack of findings for the MFV gating variable could be due to poor characterization of individuals’ VEPs. The recording parameters included a 0.1-30 Hz bandpass filter, which is narrower than the clinical recommendation of 1-100 Hz for identifying individual VEP peaks [198]. As a result, these peaks may have undergone some attenuation and blurring, and may not have been the ideal template for assessing attention with small variations in peak amplitude.

A substantial limitation comes from the use of feedback in our study. Because users
were experiencing feedback, we cannot assure that good and bad trials were unlinked. Bad feedback on a trial may have led to subsequent good performance on the following trial. This is a limitation of the retrospective offline analysis performed in this study and these findings need to be replicated in a real-time adaptive scenario in a future study.

Lastly, we seek personalized algorithms to improve performance in individual participants. Although across a population of individuals mu suppression was the most consistently helpful feature to base gating on, we did observe that phase as well as MFV were both capable boosting the information transfer rate of the system for certain users. For particular individuals, creating multivariate gating system based upon features that contribute to that person’s physiology may be more accurate than using single features. Such an exploration of combining features may be useful to explore in larger populations.
The work in this chapter stemmed from a clinical rotation course I took within the ALS clinic at the Hershey Medical Center in the spring of 2012. My work up to that point had focused on the implementation of BCI systems in college-aged individuals, but I had a goal of designing these communications systems for patients in need of them. The knowledge I gained by this rotation, by interactions with Dr. Simmons, the nurses, staff, and most importantly the patients, led me to formulate a goal of study in this clinic, one of patient-centered BCI device personalization.

The necessity for individualized BCI was apparent upon observing the heterogeneity of symptoms experienced by patients. Most surprising was the presence of cognitive deficits that were identifiable even to my untrained eye, which, as mentioned in Chapter 2, affect a significant proportion of ALS patients. The lack of research studying the effects of cognitive decline on BCI use was the motivation for pursuing this line of study. However, as the work in the clinic progressed, additional clinical variables were studied for their influence in BCI utility.

This chapter is split into three parts which focus on the larger goal of BCI personalization in the face of disease heterogeneity. In it, physical, psychological, and even genetic factors are assessed for their influence on a patient’s willingness to accept a BCI-AAC, and the utility of such a system to establish an alternative channel of communication.
4.1 Desire for BCI use among ALS participants is affected by behavioral health [3]

4.1.1 Introduction

Successful use of BCI devices has been documented in individuals with ALS [153, 206], but adoption of this technology as an AAC tool has been largely absent outside the research setting. This may be due to the fact that relatively few researchers assess the suitability of the technology in this population, particularly with respect to consideration of the extra-motor limitations of ALS patients. Widespread dissemination of this technology will require an understanding of, and accommodation for, the heterogeneity of ALS, which extends beyond the motor deficits hallmark to the disease.

As described in Section 2.2, cognitive deficits accompanying ALS can affect verbal fluency, attention, language, visual and verbal memory and learning. A substantial portion of patients exhibit signs of the behavioral variant of FTD, which is characterized by altered regulation of interpersonal conduct and emotional blunting, which can be described by caregivers as uncharacteristic irritability, selfishness, or disinterest [114]. Certain behavioral abnormalities, such as inflexibility of thought and resistance to new ideas, which can occur prior to motor manifestations [207], have been observed to be associated with AAC device rejection [149]. Huggins et al., [208], showed that BCI technology was well received, but that the devices fell short of meeting patients’ requirements for acceptable levels of accuracy and speed. However, determination of cognitive and behavioral function were not performed in those patients, nor were they given a chance to use a BCI systems before giving their assessments.

We conducted an initial survey on BCI acceptance, a BCI training protocol, and a follow-up survey to identify the factors that contribute to device acceptance. We hypothesized that patients exhibiting behavioral impairment would be less likely to have a favorable opinion of these technologies on the initial survey, as a consequence of increases in apathy and mental rigidity [114, 156]. We also hypothesized that the success a user experienced while using a BCI would influence their opinion on the follow-up survey of the utility and practicality of a BCI-AAC for long-term use.
4.1.2 Methods

Study Procedure

All patients attending a single multidisciplinary ALS clinic were informed of the study, and those who met inclusion and exclusion criteria were offered the opportunity to participate. Inclusion criteria were: 1) Age 18 years or older; 2) Diagnosis of definite, probable, probable laboratory-supported, or possible ALS by revised El Escorial research criteria [209]. Those with clinically significant dementia, as determined by the ALS clinic neurologist, were excluded. The caregivers were one of the individuals who accompanied the patient to their appointment at the clinic, often a spouse, relative, or full-time caregiver. Patients and their caregivers both consented to the study which was approved by the Institutional Review Board of the Penn State Hershey College of Medicine. The initial visit was followed by a three-month interim period, during which a subset of patients volunteered for a pilot BCI protocol. After completing this protocol, or at the time of their next clinic appointment three months later, participant pairs attended the follow-up session.

During each session, patients were seen by the attending neurologist, and their ALS Functional Rating Scale - Revised (ALSFRS-R) was recorded. Demographic data, including age, education level, gender, time since symptom onset, phone use, computer use, and travel distance to the clinic were also recorded. The ALS Cognitive Behavioral Screen (ALS-CBS) was administered to patients and caregivers. This five-minute screen has two components which assess behavior and cognition in ALS. The cognitive portion of the exam is quartered into sections assessing attention, concentration, tracking, and initiation, while the behavioral questionnaire asks caregivers about alterations in personality and behavior in the affected individual. These items are particularly important indicators of frontotemporal dysfunction in ALS [114, 121]. This instrument has been validated against a full neuropsychological battery to detect cognitive or behavioral impairments in ALS patients [210]. Patients were categorized as cognitively impaired if they scored below 17 (cognitive screen, range 0-20), and behaviorally impaired with a score below 37 (behavioral questionnaire, range 0-45). FTD was defined by a score of less than 11 on the cognitive portion and/or less than 33 on the behavioral portion. Scores for each portion of the ALS-CBS were averaged across the initial and follow-up sessions.
Following the screening procedure, patients and caregivers were introduced to BCI technology through a ten-minute demonstration by the investigator. This approach was used to produce more informed responses on the subsequent survey. The introduction was standardized through use of a prepared presentation given by the same investigator. The screening investigator answered participants’ questions to the best of his knowledge, and sought to standardize responses across participants. Topics covered were: mechanism of control, potential applications, current communication efficacy, instructions for daily use, and successes in ALS, with focus on P300 and motor-imagery-based spelling systems.

Patient and caregiver participants then completed a survey concerning their opinion of BCI-AAC technology (patient survey in Appendix A). Patients rated how often they had used a phone and computer in the month before their diagnosis. They reported if they had ever used assistive devices for verbal or written communication, mobility, feeding, or respiration, and also reported the improvement in their quality of life with use of these devices. They were asked to rate their interest in using a P300- or motor-imagery-based BCI system for communication use on a scale from 1 to 5. On the same scale, they rated their interest in using a BCI for ten potential functions, which ranged from television control to robotic arm use, as well as the importance of the features of the BCI system, such as appearance, accuracy, and invasiveness of electrode. Finally, patients were asked about their requirements of the system, such as level of accuracy and speed. Questions on the survey concerning system features, desired functions, and system requirements were modeled from a previous survey [208]. Caregivers completed a shorter survey with similar questions, covering their level of technology use as well as their interest in having their loved one use a BCI as an assistive tool.

The patients who took part in the pilot BCI study were seen for four sessions over the course of 1-2 months, and then the survey was re-administered. During the sessions of the pilot BCI protocol, patients underwent application of an EEG cap, and completed two 30-minute tasks: a P300 speller, and two-class motor-imagery cursor control. For the P300 speller, participants were instructed to mentally count when a target letter was flashed. For the motor-imagery task, participants were instructed to perform kinesthetic imagery of their left or right hand when cued. Both protocols involved a training run, in which a classifier was built using stepwise linear discriminant analysis based on the features extracted from the EEG data [206]. These features were the time averages of target
Figure 4.1: Forty-two patient and caregiver pairs enrolled in the study. During the initial session, the researcher administered the ALS-CBS screen, gave an introduction to BCI technology, and had the patient and caregiver complete surveys. In the three month period between initial and follow-up sessions, 22 patients participated in a BCI pilot study, 10 opted not to participate in the BCI study, and 10 dropped out of the study. During the follow-up, the screen and survey were re-administered.

and non-target trials for the P300 task, while for the motor-imagery task, the frequency-band powers of the SMR for left and right trials were used. In the remaining testing runs, participants received feedback as output from the classifier; they were shown the selected letter in the case of the P300 task, and observed the movement of the cursor on the screen for the motor-imagery task. For the purposes of this study, the main outcome of the pilot BCI protocol was the accuracy generated online by each participant for each task over the testing runs. A diagram of the study procedure is given in Figure 4.1.

Survey Analysis

We employed the Wilcoxon rank sum test to determine whether the occurrence of cognitive or behavioral dysfunction responses altered patients initial opinions of the desired
functions, features, and requirements of a BCI. This was done for all patients who were deemed either cognitively normal or impaired, and behaviorally normal or impaired based on the results of the ALS-CBS. No distinction was made for patients who scored low enough to be indicated as having possible FTD. Tests were corrected for multiple comparisons by a conservative Bonferroni adjustment.

Secondly, a logistic regression model was used to determine which independent variables contributed to the main outcome of acceptance, the decision to participate in the BCI pilot study. Data from all patient participants was used in the logistic regression model. Confounding variables which were highly correlated were eliminated before applying the data to the model. The statistical package R was used to compute the logistic regression, and the log-odds ratios are used to describe the relative importance of each of the factors for influencing the acceptance outcome. Forward and backward selection techniques were used to evaluate model fit based on the Akaike Information Criterion (AIC). For clarity, continuous variables were normalized before performing the regression.

Finally, we tested whether user opinions changed based on their perceived success at using the device. Participants of each BCI task were deemed “high performers” if the accuracy they produced online was above the median level. The Wilcoxon rank sum test was used to test for whether changes in participant’s interest level on the follow-up survey was significantly different between high and low performers. Only data from participants who completed the BCI pilot study were included in this analysis.

4.1.3 Results

Patient Characteristics

Forty-two patients participated in the initial session of the study, all but one with a caregiver. The demographics of the patient sample are given in Table 4.1. The fraction of patients presenting with bulbar, limb, and respiratory symptoms was 23.8%, 71.4%, and 2.4%, respectively.

Twenty-six (61.9%) patients exhibited cognitive impairment, while 15/41 (36.6%) exhibited behavioral impairment. Ten patients (23.8%) were classified as having both cognitive and behavioral impairment, and seven patients (16.7%) were identified as having possible FTD. Abnormalities in cognition and behavior were not correlated with one
Table 4.1: Patient characteristics. Median and range for 42 patient participants.

<table>
<thead>
<tr>
<th></th>
<th>Median (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient age, years</td>
<td>59 (42-81)</td>
</tr>
<tr>
<td>Time since symptom onset, mo.</td>
<td>30 (9.5-143)</td>
</tr>
<tr>
<td>ALSFRS-R</td>
<td>30.3 (7-46)</td>
</tr>
<tr>
<td>Gender, % male</td>
<td>59.5%</td>
</tr>
<tr>
<td>Education, years</td>
<td>14 (11.5-24)</td>
</tr>
<tr>
<td>Distance to clinic, miles</td>
<td>29.2 (0 -147)</td>
</tr>
<tr>
<td>ALS-CBS cognitive score</td>
<td>16 (4-19)</td>
</tr>
<tr>
<td>ALS-CBS behavioral score</td>
<td>38 (18-45)</td>
</tr>
</tbody>
</table>

another, the average scores on each test component did not change significantly between testing sessions, and there was no correlation between the results of the ALS-CBS and education or age.

**Survey Results**

The majority of patients (78.6%) used some form of assistive technology in their daily living, with mobility devices being the most prevalent (69.0%) and highly rated for improving quality of life (2.7/3 improvement rating). Fewer people used assistive technologies for verbal communication (26.2%), and ratings of quality of life improvement were lower for this type of technology (2.3/3 improvement rating).

Patients’ rankings of important system features identified accuracy, variety of functions, and standby reliability (robustness against false-positive decision making) as items with the highest median response (Figure 4.2, left). Patients rated computer and wheelchair control as desirable functions, but rated robotic arm use and temperature control as the functions of least interest (Figure 4.2, right). Caregiver-reported ratings of BCI functions mirrored patient’s responses. Most participants (59.5%) required that the BCI system be at least 80% accurate, with 90.5% of respondents satisfied with a system that could produce 90% accuracy. Nearly three-fourths of respondents would be satisfied with a speed between 15 and 19 letters per minute. Fifty percent were willing to undergo 21 minutes or more of setup time. Many responders (38.1%) opted for an ideal training time that was short (between 2 and 5 sessions), while few (16.7%) were willing to invest more than 20 sessions to become trained. Fifty percent of respondents indicated they would tolerate the system incorrectly leaving standby mode (standby un-reliability) once every five hours or more (Figure 4.3).
Figure 4.2: Patient survey responses for the initial session. Boxplots are shown with a diamond indicating the mean, the vertical line at the center of the box as the median, with the box extending to the 25th and 75th quartiles. Outliers are shown as + signs. (Left) Patient rating of the importance of features of a BCI system. Ratings ranged from 1 being “Not important to me” to 5 being “Very important”. (Right) Patient rating of their interest in the potential functions for a BCI assistive technology. Ratings ranged from 1 being “Not interesting to me” to 5 being “Very interesting”

**Impact of cognitive and behavioral impairments**

On an individual item basis, ratings of system features, functions, and requirements were not significantly altered by the presence of either cognitive or behavioral deficits in patients. However, aggregate ratings of BCI functions, or the sum of reported interest in all functions, were affected (Figure 4.4). Of all patients surveyed, those with behavioral abnormalities were likely to have lower interest in the BCI for its suggested functions ($\chi^2 = 24.56, p < .05$). Cognitive impairment significantly affected these ratings in the opposite direction ($\chi^2 = 14.87, p < .05$), with impaired individuals rating their interest in BCI functions as higher than those with normal cognitive function.

**Factors predicting BCI acceptance**

Examination of the data led us to combine the 12 sub-scores of the ALSFRS-R into three aggregate scores covering bulbar, motor, and breathing function. Phone use was eliminated as a factor due to its low variance among responders. Of the remaining factors tested, Figure 4.5 shows those that were retained after model selection. High computer use and bulbar function positively contributed to participation in the second study, while
Figure 4.3: Histograms of requirements from BCI devices, as rated by patients from the initial session.

Figure 4.4: Aggregate responses for interest in all the potential BCI functions as a percentage of total responses. Distributions of responses are separated by those who were determined to be normal or impaired in the behavioral (left plot) and cognitive (right plot) domains. Patient ratings of desired functions were significantly affected by abnormal cognition and behavior, with impairment in these domains producing higher and lower reports of interest, respectively.
maleness, behavioral score, driving distance, and education all negatively contributed. Of these, the factors which significantly contributed to pilot study participation were bulbar sub-score \( p = 0.013 \), driving distance \( p = 0.015 \), and education \( p = 0.005 \).

**BCI use affects opinion**

For those who did not participate in the pilot BCI study, there was no change at follow-up in interest in the P300 or motor-imagery BCI systems for home use, the potential functions of the device, or the features of the system.

BCI study participants were likely to change their rating of interest based on their perceived performance (Figure 4.6). This is clearest for the motor imagery task, where poor performers decreased their stated interest in the system more than high performers at a level near significance \( p = 0.063 \) after correction for two comparisons. Although this trend was not significant for the P300 task, high performers increased their rating overall in the follow-up session, while low performers decreased their average rating.

For those who did not participate in the pilot BCI study, there was no change at follow-up in interest in the P300 or motor-imagery BCI systems for home use, the potential functions of the device, or the features of the system.
Figure 4.6: Patient-reported change in interest in the two BCI systems after use in a pilot protocol. Response groups were separated by performance as well as pre- and post-BCI pilot study. Changes in opinion from pre- to post-session reached significance between the high and low performers on the motor-imagery task, with low performers reporting significantly lower interest than high performers at follow-up.

4.1.4 Discussion

Despite the relatively small size of this study, our patient population appears to be representative of the ALS patient population as a whole. The median age, the proportion of men, the fraction presenting with bulbar symptoms, and the proportion demonstrating cognitive and behavioral deficits were consistent with published data [114, 211, 212]. Our main findings demonstrate that factors of cognitive and behavioral health contribute to a patient’s interest in pursuing BCI technology for assistive communication. After using a trial BCI device, changes in interest reflected the patient’s perceived level of performance.

Patients require that a BCI system be multifunctional, accurate, and robust against false positives. System appearance was generally unimportant, whereas other factors, such as preference in electrode type, showed great variability in responses. Some rated electrode type as a very important feature in their decision to use a BCI (desired minimal invasiveness), while others did not find this feature of great importance (were open to more invasive electrode types). The desired accuracy levels of 80-90% are achieved currently using both types of BCI systems, and the cap setup of 10-20 minutes that is
current standard practice would be acceptable to many of our patients. Although a training period of 2-5 sessions is sufficient for the P300 speller, a motor-imagery training program generally takes upwards of 20 sessions [153]. An aspect identified by a majority of users for improvement is a desired communication speed of 15 letters per minute. Although this speed is elusive for a scalp-based motor-imagery system, for evoked potential BCIs, speeds of at least 15 letters per minute have been accomplished [213]. The results of this survey were largely in agreement with the literature [208]. A few of the discrepancies between studies are likely due to differences in item description by the research teams. For example, the importance of standby reliability and the utility of a robotic arm were descriptions that likely varied between investigators.

The relative lack of interest in those with behavioral impairment in using a BCI is consistent with the apathy and mental rigidity seen commonly as signs of behavioral dysfunction associated with ALS [114, 156]. Those with cognitive impairment were more interested than those with normal cognition, driven by the large percentage of respondents in the impaired group that indicated high interest (score 5). Patients with FTD have notably poor insight and judgment [214, 215], and often fail to acknowledge physical and cognitive deficits. Cognitive deficits associated with FTD lead to reduced capability for abstraction [214]. These two factors may have led to unrealistically positive assessments. As anecdotal support, we have found that patients with cognitive impairment generally rate all aspects of their quality of life as being the highest possible, and often describe their physical deficits as being minimal. The finding that psychological impairment parallels BCI acceptance is of high clinical relevance, as these clinical features influence the course of technological intervention decided on by the patient, caregiver, and physician.

In our multivariate analysis, education and driving distance were negative predictors of pilot study participation, while high bulbar function was a positive contributor (Figure 4.5). Long travel distances were an understandable deterrent to attending the four study sessions, while highly educated individuals could find the devices insufficient for their needs and frustrating to use. Alternatively, highly educated individuals may still be in the workforce and unable to participate in the pilot study, or they may have other activities that occupy their time. High bulbar function as a positive factor was somewhat surprising, given that we expected patients with speech difficulties to be more eager to participate. However, bulbar involvement has also been associated with behavioral
impairment [156]. It should be mentioned that certain aspects of patient life, such as work and family commitments, were not assessed. Although they may have had an impact on the patient’s decision to participate, these factors remain unresolved in this study.

This study has limitations of course. The relatively small sample size of the survey group limits the power of the significant findings, and the method of sampling patients from the clinic on a voluntary basis introduces a bias towards opinions which are favorable to BCI technology. The use of a full neuropsychological screen would have been more accurate in gaging the cognitive and behavioral health of the participants, however the ALS-CBS was used mainly due to limited time during the clinical visit. Finally, the training period of four sessions for the BCI protocol was chosen to provide participants with a practical demonstration of the utility of the communication interface. While this may have been long enough for the P300 speller system, an extended training period may have been required for the users to reach their optimal level of performance with the motor-imagery system. Ratings of the motor-imagery system may have suffered due to this insufficient training period.

Critically, this study showed how perception of success caused a change in patient opinion of the devices. Participants who performed well using the systems tended to maintain or increase their interest in pursuing one of the devices, while those who did not achieve satisfactory performance saw a marked decrease in their desire to use the device. This finding indicates the need for trials of these expensive BCI-AAC devices early in the design stage of device development. Such trials, with a patient’s perception of performance, will likely determine whether the patient accepts and will continue using the device. The observed changes in opinion also underline the importance of BCI-informed respondents in the collection of survey data relating to perceived usefulness of such devices.
4.2 Cognitive impairment negatively impacts BCI use [4]

4.2.1 Introduction

People with ALS have for some time been considered primary candidates for BCI communication systems. Although the capacity to control a BCI has been shown to be relatively conserved in these patients, only marginal communication has yet been established for those experiencing CLIS of late-stage ALS [44]. One explanation for unsuccessful BCI use in CLIS is based on the loss of goal-directed thinking behavior by the user. This phenomenon, proposed as a type of cognitive impairment associated with the ALS/FTD syndrome, results from the complete lack of motor control and subsequent feedback [180]. In this study, we explored the effect of cognitive and behavioral impairment, as well as other clinical variables, as possible indicators for losses in performance in two BCI tasks: the P300 speller and motor-imagery cursor control.

The amplitude of the P300 decreases and latency increases with normal aging [216, 217, 218]. ALS has been shown to further prolong the latencies of the P300 response for visual and audio paradigms [133, 141, 219]. Some [133, 220], but not all [141] groups report lower amplitude responses in patients compared to controls. The degree of functional motor impairment, as measured using the ALSFRS-R, was not found to be correlated with P300 abnormality [141, 206, 220, 221]. The altered morphology of the oddball evoked response in ALS has been suggested to result from reduced focus and attention [219] as well as abnormal memory processing [222], and possibly even a failure of cognitive association [141], although there are conflicting results on whether cognitive impairment associated with ALS leads to P300 abnormality [220].

Persons with ALS have also been shown to be able to suppress their SMRs for the purpose of operating a motor-imagery BCI [153], although baseline SMR levels are of lower amplitude in patients with ALS [223] and FTD [224]. Efficacy of SMR modulation is not associated with physical functional disability [180], nor is the locus of motor weakness correlated with the ability to form useful imagery of that limb [155]. Bulbar involvement, however, has been shown to negatively influence imagery ability and therefore BCI performance [155, 223].

The goal of this study was to identify ALS-related factors that contribute to success
using P300 and motor-imagery BCI systems. We expected that cognitive and behavioral impairments would impact successful BCI operation more than physical factors. For the P300 speller, we anticipated that altered cognition would weaken BCI performance due to reduced attention, consistent with the literature linking both ALS and FTD to a reduced P300 response. For the motor-imagery BCI, we expected baseline SMR spectral power levels would be the main predictor of BCI performance, as reported previously [166], but also expected that cognitive impairment would be detrimental to achieving the SMR suppression required for device control.

4.2.2 Methods

Study Procedure

A subset of twenty-five patients from the survey study described in Section 4.1 were enrolled in this study. All met the criteria for definite, probable, probable laboratory-supported, or possible ALS, and were absent of overt dementia. Fifteen control participants were recruited from the nearby area by means of a community bulletin. In addition to being neurologically healthy, control participants were age and gender matched to the patient group. The study protocol was approved by the Institutional Review Board of the Penn State Hershey College of Medicine and all participants provided informed consent.

Patients had demographic and clinical variables recorded, and were administered the ALS-CBS, as described in Section 4.1. Control participants were also administered the cognitive portion of the ALS-CBS. Participants were categorized as cognitively impaired if they achieved a score below 17 on the cognitive portion (range 0-20), and were behaviorally impaired if they scored below 37 (range 0-45) on the behavioral assessment. The ALS-CBS was administered before and after the BCI study procedures, and the two scores were averaged.

Patient participants completed four sessions of BCI recordings over the course of 1-2 months, with each hour-long session split between two BCI paradigms: a P300 spelling system, and a two-class motor-imagery center-out task. Control participants completed two sessions. Each run of the P300 spelling task consisted of copy spelling a four letter word, where each trial culminated in the selection of one letter after flashing each grid icon 20 times. A checkerboard-type speller with 32 targets was used to evoke the P300 signal from the user [80]. Targets were highlighted for 187.5 ms, and the inter-
stimulus interval was 62.5 ms. Targets were randomly assigned into groups of four at the beginning of each letter presentation, and the order of groups was randomized as well. Group order was constrained so that no target flashed consecutively, i.e. a specific letter flashed no more than once every half second. During each session the participants also performed one calibration run in the ‘covert’ spelling setting; they were instructed to focus their eyes on the center of the letter matrix, which contained a fixation cross, and to only direct their attention to the target letter in the grid periphery. P300 responses were defined by amplitude and latency in channels Fz, Cz, and Pz. P300 amplitude was the maximum magnitude over a baseline of the average EEG response in these channels 250-500 ms after the stimulus [78]. Latency was the time after the stimulus at which this maximum occurred.

Each run of the motor-imagery task consisted of ten left and right trials, in which the subject was instructed to perform kinesthetic imagery of their left and right hands, as well as ten no-go trials, during which they were instructed to relax. The first run of each session entailed calibration of the classifier without feedback. In the remaining feedback runs, real-time classification of brain features enabled the cursor to move freely from the center of the screen under the control of the user.

During the recordings, EEG electrodes were affixed in an electrode cap at nineteen locations in the 10-20 system, with ground at Fpz, and referenced to linked earlobes. Additionally, for purpose of artifact reduction, electrodes were placed on the forehead and lateral canthi to record the electrical activity of eye movement. Signals were amplified and digitized with two g.USBamp amplifiers. Data acquisition, signal processing, and feedback generation were performed by a customized program in BCI2000. Aside from the specifics of the feature extraction step, the processing pipelines for the two systems were very similar, and included artifact reduction and rejection, feature extraction, and classification. Ocular artifact reduction was automated though a regression procedure as described in Section 3.2, followed by rejection of epochs which exceeded an amplitude of $\pm 75\mu V$. Features used for the P300 spelling system were the stimulus time-locked average EEG signals, downsampled to 20 Hz. The features extracted during the motor-imagery task were the spectra in 2 Hz bins from 5-30 Hz from each Laplacian-referenced EEG channel. Classifier generation was implemented using the stepwise$\textit{fit.m}$ function in MATLAB, which utilized stepwise selection of regression coefficients on training data to generate a classifier for predicting the two target classes.
online [80]. This classifier was used to generate real-time feedback for the user, and was updated after each feedback run.

**Analysis**

The online classification accuracy for each task was reported as the percentage of trials finished correctly in the last two runs for each session. For most individuals, this meant that the P300 accuracy statistic was calculated from the percent letters out of 32 (4 sessions with 2 runs spelling a 4 letter word) that were spelled correctly in the online session. The motor-imagery task online accuracy was calculated from the percentage of trials out of 160 (4 sessions with 2 runs of 10 left and 10 right trials) in which the cursor finished on the correct side of the screen. For patients and control participants who completed fewer than four sessions, accuracy statistics were calculated from a reduced number of trials.

We also wanted to define the differences in EEG features for the control and patient groups. For the P300 task, we report latency and amplitude of the evoked potential under overt and covert spelling conditions. For the motor-imagery task, we calculated the BCI performance predictor established by Blankertz et al. [166]. This predictor was calculated as the difference of the amplitude of the power spectrum and the fit of a \( 1/f \) noise spectrum in the SMR range. Statistical comparisons of sample means between groups were handled with the Wilcoxon rank-sum test.

In addition to reporting accuracy and feature differences between patients and controls, we defined a classifier-independent approach to determining the robustness of the signals recorded from the participants, which we call ‘quality’. The rationale for considering feature quality was to distinguish the potential for BCI aptitude from a performance metric based on an individually-trained classifier. The quality of our two tasks was calculated as the standard distance between data belonging to the two classes of interest: left and right trials in the case of motor-imagery, and target and non-target trials for the P300 speller.

As an example, \( Q_{C3} \), the quality of motor-imagery in channel C3, was calculated as the standard distance between left and right average power spectra,

\[
Q_{C3} = \frac{\mu_{C3,L} - \mu_{C3,R}}{\sqrt{\sigma^2_{C3,L} + \sigma^2_{C3,R}}},
\]  

(4.1)
where $\mu_{C3,L}$ was the average power spectrum for left trials in the channel C3, and $\sigma_{C3,L}$ was the standard deviation over these trials. Motor-imagery quality was defined for each individual and Laplacian-referenced channel for frequencies 5-30 Hz. Quality was defined similarly for P300 as between the target and non-target averages 250-500 ms after the stimulus.

To determine whether quality of the P300 and motor-imagery signals were dependent on physical or psychological factors, we performed multiple linear regression, with the quality scores as the dependent variables. The independent variables of interest were the motor, bulbar, and respiration sub-scores of the ALSFRS-R, as well as the behavioral and cognitive scores of the ALS-CBS. Each of these factors were normalized before performing the regression. Model selection was performed through minimization of the Akaike Information Criterion (AIC), a measure that reflects the goodness of fit of the model as well as its complexity. The model selection procedure was performed in R [225] using both forward and backward feature selection algorithms. The factors which contributed to a minimized AIC through both forward and backward feature selection were retained in the final reduced model.

### 4.2.3 Results

Demographics for patients and controls are given in Table 4.2. Controls were well matched for age and gender, but had more years of formal education. Of the 19 patients that provided information on region of onset, 2 presented with bulbar onset, 13 with limb onset, 3 with simultaneous bulbar and limb onset, and 1 with simultaneous limb and respiratory onset.

One participant did not complete the ALS-CBS because of severe communication difficulties, and another participant did not complete the behavioral portion of the screen because of the lack of an available caregiver. Among patients, the median cognitive score on the ALS-CBS was 16 (range 9-19), with 14/24 (58%) defined as cognitively impaired. The median behavioral score was 37.5 (range 18-45), with 8/23 (35%) defined as behaviorally impaired. The median cognitive score for the control participants was 17.5 (range 15.5-19), with 3/15 (20%) defined as cognitively impaired. The distribution of cognitive scores in the two groups was quite different. Abnormal controls were just below the cutoff, consisting of 2 scores of 16 and one of 15.5. In contrast, patient
Table 4.2: Patient and control participant demographics. Statistics are given as median and range (min-max), if applicable, and p-values compare control and patient samples.

<table>
<thead>
<tr>
<th></th>
<th>Patients (n=25)</th>
<th>Controls (n=15)</th>
<th>p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>58 (45.5-74)</td>
<td>55 (45.73.5)</td>
<td>0.364</td>
</tr>
<tr>
<td>Education, years</td>
<td>14 (11.5-24)</td>
<td>18 (12-24)</td>
<td>0.006</td>
</tr>
<tr>
<td>Gender, % male</td>
<td>68</td>
<td>60</td>
<td>0.624</td>
</tr>
<tr>
<td>ALSFRS-R</td>
<td>30 (0-46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSSO, months</td>
<td>32 (12-113)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ALSFRS-R = ALS Functional Rating Scale - Revised, TSSO = Time since symptom onset.

scores were skewed towards the lower end of the range, with nine patients scoring lower than the poorest performing control subject. Overall cognitive scores were significantly higher for the control group ($p = .013$).

Exceptions to the recording procedure described in the methods included two patients (P06, P28) who completed only one session of the study and one patient (P23) who completed three sessions. One control participant (S14) also only completed one session. Electrode impedances often exceeded the recommended 5 kΩ level [12] for individual participants. Of the 19 EEG electrodes attached during each session, 81% registered impedance values less than 10 kΩ, and 90% were less than 30 kΩ. There was no correlation found between electrode impedance and the EEG features relevant to the BCI tasks, nor the performances achieved on those tasks.

**P300 Speller**

Although P300-BCI performance varied within patient and control participant groups, control participants on average performed at a higher level ($p = 0.033$, Figure 4.7). The average accuracy achieved by the patients on the P300 task was 68% (range 0-100%). For control participants, the mean accuracy was 86% (range 38-100%). The P300 quality for patients was generally highest in channels along the midline, and strongest in the central electrodes (Figure 4.8b). Timing of the target VEP varied among individuals as well, with most participants having a large positive deflection around 200 ms which was sustained through 400 ms (Figure 4.8c). Many individuals also had a large negative deflection sometime after 400 ms. Control participants had similar topology and timing of P300 quality.

The average amplitude of the patients’ P300, that is the largest positive deflection
of the VEP in channel Cz during the 250-500 ms window after the target stimulus, was 3.0 µV (±1.5 µV), with latency occurring 334.2 ms (±77.9 ms) after the stimulus. Control participants had a slightly larger mean P300 amplitude than patients of 3.7 µV (±1.5 µV, \( p = 0.105 \)), although their mean latency of 330.7 ms (±77.7 ms, \( p = 0.723 \)) was not different. P300 latency for patients was similar to a previous report, but the amplitude was significantly lower (4.06 µV from [206], \( p = 0.003 \) for a one-sample t-test). Compared to the overt spelling task, the mean latency of the P300 during the covert spelling condition was longer for both patients (399.3 ±72.7 ms, \( p = 0.004 \)) and controls (426.3 ±81.7 ms, \( p = 0.002 \)), but the amplitude was unchanged for both groups.

Regression was performed to determine whether the mean absolute value of P300 quality in channels Fz, Cz, and Pz occurring 250-500 ms after the stimulus was correlated with any patient characteristics. Of the factors evaluated, cognitive score of the ALS-CBS was the only variable to positively correlate with P300 quality (Figure 4.9,
Figure 4.8: Quality of the evoked response between target and non-target trials. (a) Top: Target (red) and non-target (black) averages in channel Cz for participant P27. Bottom: Quality is defined as the sum of the absolute value of the standard difference between the target and non-target trials in a specified time window, shaded here for the 250 – 500 ms period. (b) Qualities in each channel at this time period are given for each participant. (c) Qualities at each time point are also shown for each participant, as the (d) channel average of qualities in Fz, Cz, and Pz (shown here for participant P12).

\( p < 0.05 \). However, the stepwise model selection procedure, which added and removed factors to minimize the AIC, retained both cognitive and behavioral scores in the final model. Both of these scores were positively associated with P300 quality, indicating that participants who scored high on the psychological screen were more likely to perform well on the spelling task. A break down of sub-scores on the cognitive portion of the ALS-CBS indicated that deficiencies in tasks of attention and tracking were correlated with poor P300 feature quality (\( r^2 = 0.178, p = 0.04 \) and \( r^2 = 0.173, p = 0.04 \)).

**Motor-imagery task**

Group-wide performance was lower on the motor-imagery task (Figure 4.7), with an average accuracy achieved by the patients of 60% (range 45-97%). Five out of the 25 patient participants were able to control the motor-imagery center out task at the 70% level suggested for productive use. For control participants, the mean accuracy was 62% (range 46-90%), with 3 of the 15 control participants achieving 70% accuracy.

For the five patients who were successful in motor-imagery control of the center-out task (P09, P21, P31, P34, P39), channels C3 and C4 produced the highest quality
Figure 4.9: Cognition positively predicts BCI signal quality. (Left) Regression of P300 quality on patient characteristics. Bars indicate the mean regression coefficients for each factor, along with 95% confidence interval of this estimate. (Right) Regression with motor-imagery quality as the dependent variable.

motor-imagery responses (Figure 4.10b). In these individuals, both mu and beta band sensorimotor rhythms were modulated to control the cursor (Figure 4.10c). In other individuals, there was little differentiation in SMR power between trial types (P16, P27), or there was an absence of defined SMRs altogether (P02, P29). The BCI performance predictor, which reflects the prominence of these resting state rhythms, was positively correlated with the quality of the motor imagery signal ($r^2 = .717$, after removal of outlier P16). We found no difference in the BCI performance predictor between patients and controls.

The difference in motor-imagery quality between channels C4 and C3, averaged over the 5-30 Hz range served as the dependent variable for the regression on the disease factors. Again, cognitive score of the ALS-CBS was the only variable to positively correlate with motor-imagery quality (Figure 4.9, $p < 0.05$). The model reduction procedure retained only cognitive score as the factor which minimized the AIC. A breakdown of the sub-scores of the cognitive screen revealed that tasks involving attention and word initiation were significantly correlated with the quality of the motor-imagery features ($r^2 = 0.26$, $p = 0.01$ for both sub-scores).

In some patient participants, strong power in gamma-range frequency bands consistently appeared in frontal and temporal electrodes. The ratio of spectral power in the 40-50 Hz range to task-relevant power in the 8-24 Hz range was calculated in channels Fp1, Tz, and Cz. An ANOVA determined there to be a significant difference in this ratio between the cognitively-impaired patients, cognitively-normal patients, and controls ($p < 0.05$ in each channel). Subsequent pairwise comparisons revealed that patients
Figure 4.10: Quality of SMR discrimination between right and left tasks. (a) Quality is defined as the mean standard distance between left (blue) and right (red) spectra in the $5 - 30$ Hz range. (b) Qualities in each channel are given for each participant. (c) Qualities at each frequency are also shown for each participant. (d) The frequency plots in (c) are calculated from the mean of the difference of the qualities in channels C3 and C4 (shown here for participant P34).

with cognitive impairment possessed an elevated power ratio compared to the other two groups. (Figure 4.11).

Task dominance, or differential success using one of the systems, was observed in many participants. A metric was generated which described the dominance produced by each participant as the residual of the regression of motor-imagery task accuracy on P300 task accuracy. The main factor which predicted task dominance was the participant’s age, with older subjects performing relatively better using the P300 system than the motor-imagery system (Figure 4.12).

4.2.4 Discussion

Of the factors considered for this study, those related to cognitive function, rather than physical function, were determined to be a major predictor of successful BCI operation. Patients with ALS who scored higher on the cognitive portion of the ALS-CBS were more likely to achieve higher accuracies using the P300 and motor-imagery systems. A possible mechanism for this loss in performance in patients with cognitive impairment
Figure 4.11: Cognitively-impaired patients display elevated ratios of power in the 40 – 50 Hz band over the average power in the 8 – 24 Hz band during no-go trials of a motor-imagery task. ANOVA identified significant differences between cognitively-impaired patients, cognitively-normal patients, and control participants in channels Fp1, T7, and Cz (p<0.05). Significant pairwise differences between groups at the p<0.05 level are indicated by an asterisk.

was a decrease in the signal-to-noise ratio within task-relevant EEG frequencies. There was also evidence that behavioral dysfunction negatively affects P300 speller performance. Finally, there was a tendency for older participants to achieve relatively better performance with the P300 system.

Riccio et al. [226] asserted that attention is an important factor in predicting P300 BCI performance. Specifically, those who are better at detecting a target letter (self-reported) in a serial visual stream are also better at the P300 task. They attribute this to the ability to update an attentional filter. In our paradigm, we found the quality of the P300 consistent throughout the length of the trial and from the first run of the session to the last, even for patients who scored poorly on measures of attention, indicating that the length of the trial or the session was not a determinant of the quality of the P300 control signal.

Poor performance in those with impaired cognition may have been due to deficits in ability to update their attentional filter to recognize the target stimulus. This would be consistent with the executive function deficits seen in FTD, which are typically demonstrated through tests which measure shifting of cognitive set, initiation of behavior, inhibition of inappropriate responses, and control of attention [227]. The prefrontal damage that occurs in ALS can often lead to the presence of reduced focus and attention, hindering the production of a stable P300 [222]. However, we cannot rule out the possibility that visual dysfunction may have played a role in the low performance experienced by some patients, as reported recently [99].
The morphology of the evoked potentials between patients and controls was remarkably similar in both overt and covert conditions, albeit with a non-significant reduction in the amplitude of the P300 in the ALS sample as a whole. Among patients, P300 amplitude and latency were not significantly affected by physical function, measured by the ALSFRS-R, nor did we observe an increase in P300 latency with age. This may have been partly due to limited age range compared to studies which reported this effect [216, 217, 218], as well as the smaller sample size. The mean latency of the P300 found in this study was similar to the 360 ms previously reported [206], although the amplitude was lower. Both of these metrics are difficult to compare because of the differences in trial design and stimulation protocol between studies. In addition, the lower amplitude might be due in part to the higher electrode impedances in this study. The observed increase in latency of the P300 from the overt to the covert attention condition reflects the loss of the 200 ms peak during the covert condition, with results very similar to Treder et al. [191]. The extinction of the fixation-related portion of the evoked potential during the covert trial results in a subsequent decrease in BCI system performance [191, 228].

Cognitive impairment was the major disease-related factor that impaired the effective use of the motor-imagery system. Motor-imagery is dependent on the user’s ability to self-generate imagery of a detailed motor action from memory. The user needs to repeat this imagery when cued, as well as switch imageries quickly, a task heavily dependent on executive function. As in the case of the P300, the sub-score of attention was a particularly important indicator of motor-imagery quality, while the initiation sub-score...
was also positively correlated with signal quality. Both findings implicate executive dysfunction as detrimental to success on the motor-imagery task. Motor dysfunction, as measured by the motor sub-score of the ALSFRS-R, was not correlated with motor-imagery performance, consistent with previous reports [155, 180]. Although age was not found to be a significant factor in predicting BCI performance overall, we did find that there was task dominance for the P300 system in older individuals. This finding implicates P300 systems as potentially more useful for older users, while the motor-imagery systems may be better utilized by younger users.

In contrast to previous findings [223], we did not observe a significant difference in the resting state power of the mu rhythms in ALS patients compared to controls, although we did observe an increase in the ratio of task-irrelevant to task-relevant power in patients with cognitive impairment (Figure 4.11), which is in agreement with other studies that showed a reduction of motor rhythm amplitudes in FTD [224]. These findings lead us to believe that the changes in the brain specific to ALS-associated cognitive impairment may reduce the utility of a motor-imagery BCI because of the decrease in signal to noise ratio of the sensorimotor rhythms. On the other hand, with sufficient training, patients without cognitive impairment may be able to modulate SMR well enough to achieve accuracy levels comparable to healthy individuals. Our short training period of four sessions may have been inadequate for the participants to achieve optimal results. The equally poor performance of the control participant group on this task leads us to believe that, given a longer period of training, a greater number of patients could achieve motor-imagery control.

The main findings of our study lead us to make to some recommendations for BCI communication devices in ALS. First, the younger individuals in our study demonstrate a dominance for the motor-imagery BCI paradigm, while the older individuals perform better with the P300 system compared to the motor-imagery system. Although performance is only one factor for determining BCI device adoption, this age-paradigm interaction may be a useful tool for recommending systems to patients of different ages. Second, we have shown that both of the standard BCI paradigms tested fall short of supporting useful communication for cognitively impaired patients. However, patients with cognitive impairments are just as likely as those without such impairments to desire a BCI for assistive communication use [3]. This presents a meaningful clinical problem if we are to offer BCI systems to patients with ALS, because up to half of individuals...
with ALS have cognitive or behavioral impairment. For these users, the design of the BCI will have to accommodate for reduced cognitive ability, or they may prove inoperable. When designing assistive devices, cognitive load can be alleviated by reducing memory loads and distractions and providing information in more intuitive forms [229]. Modifications to BCI systems have been attempted using centralized, serial stimulus presentation (RSVP), and language prediction models [230], as well as using symbols representing common daily activities [231]. Similar modifications will need to be tested for the potential to improve communication for ALS patients with reduced cognitive capacity.

4.3 Repeats of hexanucleotide G_4C_2 in C9ORF72 correlate with quality of BCI performance [5]

4.3.1 Introduction

In modern health care, the term “personalized medicine” is often associated with treatments and therapeutics which are targeted towards an individual, often on the basis of one’s genetic diversity. The treatment of ALS is no different, and the genetic anomalies that shape the specific course of illness may help to define subsets of patients who will be receptive to new pharmaceuticals, or in the case of the following work, assistive communication devices. As mentioned in Chapter 2, mutations in a number of genes have been implicated in a growing fraction of the total ALS cases. Of these, some have been shown to also contribute to the development of frontotemporal dementia.

The repeat expansion in the gene C9ORF72, found in open reading frame 72 on chromosome 9, is one of the most recent discoveries that has provided another biological link between ALS and FTLD. Abnormal expansion of the G_4C_2 (GGGGCC) hexanucleotide in this sequence occurs in about 40% of familial ALS and 7% of sporadic ALS cases, while being rare in healthy individuals. C9ORF72 repeat expansion is also prevalent at rates <2% in Huntington’s, Alzheimer’s, and Parkinson’s diseases [232, 233], similar to the trinucleotide repeats responsible for an array of additional neurogenetic disorders [234]. ALS patients possessing the pathologic repeat expansion exhibit earlier onset [232, 235, 236, 237, 238, 239], more rapid progression, and earlier death [236, 237, 239, 240]. Penetrence of the mutation increases with age, from
a state of non-penetrance at age 35 to full penetrance by age 80, with nearly 100% of those with the expansion developing symptoms by this time [241, 242, 243]. Among ALS pedigrees, there is evidence for genetic anticipation, with earlier age of onset in the descendants of expansion carriers. [241]

The question arises as to what defines a pathological repeat length of the C9ORF72 gene? In the pair of papers that originally established the link between the repeat and ALS-FTD, the cutoff for a pathologic expansion was determined to be greater than 30 hexanucleotide repeats [244, 245]. Polymerase chain reaction (PCR) is capable of determining whether a patient has the repeat length of greater than 30, but is not ideal for determining the repeat length that could be in the thousands. For this, Southern Blotting is typically employed. [232, 233, 238, 246]. To complicate matters, significant somatic heterogeneity has been observed, with different expansion lengths found in different tissues, even between regions of the brain [233, 239, 247].

Still uncertainties remain about the phenotypic differences occurring as a result of long expansions and those of intermediate length (7-30 repeats). One study showed that five FTD patients with repeat lengths of 20-22 were similarly cognitively impaired as the four patients with longer repeats (>30) [248], while another study showed that four patients with repeats of 20-22 were phenotypically similar (age of onset, higher prevalence of dementia) to those with >30 [249]. Also, intermediate numbers were shown to have a risk effect in familial FTLD, with lengths of 12-21 associated with earlier onset and shorter survival [241]. On the other hand, others have found no such correlations between sub-threshold repeat lengths and phenotype in ALS [250].

There has been limited work done to determine the structural and functional brain changes that co-occur with the C9ORF72 repeat expansion. MRI studies point to structural changes common among pathologic repeat carriers [236, 237, 251, 252, 253]. For one, there are gray matter changes that extend beyond the degeneration of the corticomotor pathways typically involved in ALS. Transcranial magnetic stimulation was used to assess cortical excitability in ALS patients [243]. Both those with sporadic ALS and those possessing the repeat expansion displayed increased motor evoked potentials compared to controls, indicating similar pathological hyper-excitability of the cortex. In an EEG study of seven ALS/FTLD patients with the repeat expansion, two showed generalized slowing of the background activity, while another two showed intermittent abnormal delta-theta activity temporally [253].
The question of whether genotype, particularly in this case as it applies to expansion of the \textit{C9ORF72}, produces measurable changes in the performance of an EEG-based BCI task remains to be seen. In this section, we explore the potential of this genetic marker as a screening mechanism for brain-computer interface utility, and describe how the intermediate repeat expansions found in a sample of ALS patients interplay with cognition and performance on two BCI tasks.

### 4.3.2 Methods

The majority of this protocol was described in the previous section. Twenty-five patients that were being treated at a multidisciplinary ALS clinic were enrolled in the study. Those with clinically significant dementia, as determined by the ALS clinic neurologist, were excluded. Fifteen age and gender matched control participants were also included in the study. In addition to being neurologically healthy, control participants were age and gender matched to the patient group.

The first goal of the analysis was to determine whether the repeat lengths, of intermediate size or otherwise, correlated with any of the physical or psychological symptoms exhibited by the patients in the study. These include correlations between repeat length and bulbar vs. limb onset, impaired cognition and behavior, as well as age at the onset of symptoms.

In order to determine whether a genetic screen may contain predictive value for BCI utility, we also evaluated the interaction of repeat length with two measures of BCI performance in each task: online accuracy and signal quality. These outcome variables of accuracy and quality were the same as those described in Section 4.2. Motor-imagery quality was defined for each Laplacian-referenced channel. A single measure of ‘classic’ motor-imagery quality for each participant was determined by averaging the quality differences in the paired channels F3-F4, C3-C4, and P3-P4 over the frequency range 5-30 Hz. Quality was defined similarly for P300 as between the target and non-target averages. The timing of the ‘classic’ quality signal was averaged over the 250-500 ms period following the stimulus, in channels Fz, Cz, and Pz. The quality measures were regressed with repeat length to assess for interaction. Additionally, we defined a cutoff level at the median repeat length in order to determine how groups of users possessing a high repeat length compared to those with a low repeat length. For these tests, we used a
Figure 4.13: Histogram of GGGGCC repeat lengths within C9ORF72 in 24 ALS patients

Figure 4.14: Neither (a) cognitive impairment, (b) behavioral impairment, (c) bulbar onset, nor (d) age at symptom onset were correlated with sub-threshold expansions of C9ORF72.

the Wilcoxon rank-sum test, or in the case of comparing these two groups with controls, a one-way ANOVA.

4.3.3 Results

Of the patients enrolled in the study, one participant (P38) did not complete the ALS-CBS because of severe communication difficulties, and another participant (P43) did not complete the behavioral portion of the screen because of the lack of an available caregiver. All twenty-five patients and fourteen of fifteen controls provided samples for assessment of the C9ORF72 repeat expansion, although one DNA sample from a patient
was of insufficient quality to perform analysis. Therefore, analysis is performed on the twenty-four patients and fourteen controls with complete genetic results.

Overall cognitive scores on the ALS-CBS were significantly higher for the control group \((p = .019)\), with a median score for patients of 16 and a median score for controls of 17.5. The median behavioral score for patients was 38.25 (range 18-45). All patients in the study possessed sub-threshold expansion lengths of 15 hexanucleotide repeats or less (Figure 4.13). These sub-threshold repeat lengths in patients were not found to be significantly associated with cognitive impairment, behavioral impairment, bulbar onset, or age at symptom onset (Figure 4.14).

Patients were divided into ‘low-repeat’ and ‘high-repeat’ groups based on a repeat length above or below the median length of eight. Figure 4.15 displays the differences in the P300 evoked potential between patients in the two repeat groups and controls. Controls demonstrated consistently better quality in the evoked potential over a wide range, including early potentials at 200 ms, the classic P300 (250-500 ms after the cue), and even late rebound potentials. We focused on the quality of the classic P300, as it is not as affected by fixation of eye gaze as the 200 ms potential. Classic P300 quality, averaged over the channels Fz, Cz, and Pz was found to be most different between the control group and the ALS group with high repeat lengths. Furthermore, the negative correlation between the repeat length and classic P300 quality averaged over the 250-500 ms range was significant (Figure 4.15c, \(R^2 = 0.21, p = 0.024\)). Control participants, who achieved the highest performance overall on the P300 system, displayed a significantly higher accuracy rate compared to the high-repeat group as determined by ANOVA and subsequent Tukey post-hoc tests (Figure 4.15d). Although there was a tendency for patients in the low repeat group to achieve higher accuracies on the online P300 task than those in the high repeat group, this difference was not significant \((p > 0.05)\).

Figure 4.16 displays the same analysis for the motor-imagery task. Here, the three groups have similar quality, as there was slightly better than random performance for all but a few of the best patient and control participants. The variability of imagery ability among participants yielded little differentiation between the qualities in specific frequencies across groups (Figure 4.16a). When contralateral electrode pairs were averaged to generate a measure of classic motor-imagery quality, controls displayed clear peaks around the typical mu and beta ranges (Figure 4.16b). Patients produced a more distributed band of frequency desynchronization, which led to a broader, al-
though smaller amplitude quality of motor-imagery. Averaged over the 5-30 Hz range, the classic motor-imagery quality exhibited a significant negative correlation with gene repeat expansion length (Figure 4.16c, $R^2 = 0.19, p = 0.033$). Although the patients with high repeat lengths averaged lower accuracies on the online motor-imagery task (Figure 4.16d), this difference between groups was not significant ($p > 0.05$).
Figure 4.15: Repeat length is inversely related to P300 quality. (a) P300 quality in channels Fz, Cz, and Pz in high-repeat patients (red), low-repeat patients (blue), and controls (black). Greater values indicate better discriminability of the P300 control signal. Horizontal bars of magenta, dark red, and dark blue indicate significant ANOVA pairwise differences between high-low, high-control, and low-control groups at \( p < .05 \). Thicker horizontal bars indicate significance after Bonferroni correction for all time windows tested. (b) The classic P300 quality, or the average over the three channels. The range of a typical P300 signal (250-500 ms) is shaded in gray. (c) A significant negative correlation exists between the average of the classic P300 quality in the 250-500 ms range and the length of the repeat expansion among patients. (d) Accuracy on the P300 task by repeat group. The asterisk indicates a pairwise difference between the high-repeat and control groups following a significant ANOVA result.
Figure 4.16: Repeat length is inversely related to motor-imagery quality. (a) Motor-imagery quality in channels F3, F4, C3, C4, P3, and P4 for the three participant groups. Refer to Figure 4.15 for information about horizontal lines. (b) Classic motor-imagery quality, or the average over the paired differences in quality between F3-F4, C3-C4, and P3-P4. The range of a typical SMR modulation (5-30 Hz) is shaded in gray. (c) A significant negative correlation exists between the average of the classic motor-imagery quality in the 5-30 Hz range and the length of the repeat expansion among patients. (d) Accuracy on the motor-imagery task by repeat group.
4.3.4 Discussion

All of the patients included in the study possessed a sub-threshold repeat expansion, indicating that none had C9ORF72-linked ALS [245]. Furthermore, these patients did not display any of the signs and symptoms typically associated with this linkage. Patients were no more likely to have cognitive or behavioral impairments, as opposed to those with the pathological expansion who often have co-occurring frontotemporal dementia. Studies of FTD patients possessing intermediate length repeats (~12-22) were associated with similar levels of cognitive impairment, age of onset, and prevalence of carrying high-risk alleles as individuals with longer repeats [241, 248, 249, 254], although this finding was not observed in ALS [250]. Many of the patients in our study fell outside of this intermediate range, possessing repeat lengths of eight or less.

Certain findings from this study warrant a further look in a larger group of patients, specifically, the association between BCI task proficiency and C9ORF72 repeat expansion length. Both the quality of the control signals for the P300 and motor-imagery tasks were negatively associated with gene expansion length. The online accuracies produced in these individuals also trended in this direction, although the difference between groups was not clear enough to make claims about performance effects of longer repeat lengths. Given that impaired cognition, a factor shown to impede BCI performance, was not associated with these longer repeat lengths, additional assessment is required to determine the size of this intermediate length effect, as well as a possible causal mechanism.
The personalization addressed in the previous chapter focused on identifying patient characteristics that predicted BCI performance, so that a user could be paired with a system which was most likely to result in a compatible interface. As the complement to the previous chapter, this chapter explores personalization of the BCI system from the other direction. Here, we identify ways in which we could personalize the operation of the system to optimally utilize the brain signals generated by the specific user. Personalization of system operation can be achieved through a number of means. In the data collection stage, selective recording of subsets of electrodes [255, 256] can minimize setup time and discomfort. At the level of feature selection, only the most discriminable time and frequency features could be forwarded to the classifier [185]. Personalization can also occur at the level of the interface, whether by cuing modality [48, 257], speed [258], or number of classes [258, 259], to name a few.

The need for personalization in the most general sense was made apparent within the first few sessions of recording in the ALS clinic. Things such as screen adjustment and positioning, system pacing, and electrode subsets were small modifications that were sometimes made so that the device was better able to be used. These changes were often made due to physical restrictions, usually as a result of bulky wheelchairs with on-board monitors, but timing changes to the visual stimuli were also made to accommodate for longer trials or slower flash times. Although these are anecdotal examples, I was able to experience firsthand what difference could be made in BCI performance by small
alterations in system parameters. The remainder of the chapter focuses on engineering solutions for device personalization, so that adaptation to individual brain patterns may allow for maximum performance. Additionally, we report trends in personalization that occur across neurologically healthy individuals and ALS patients with and without associated cognitive impairment.

In this chapter, two projects are described which were accomplished with offline analysis of BCI data. First, optimization was carried out on features generated from both P300 and motor-imagery tasks. This was done to find how the optimal subsets of electrodes and spatio-temporal locations of class-discriminable brain data changed on an individual and group level. Second, alternative task features related to network connectivity patterns were assessed for classification utility. A simple measure of coherence was compared to a novel data assimilation technique that uses a biophysical neural field model to estimate cortical connectivity. The results of these analyses provide a unique perspective on the opportunities and challenges for device deployment in ALS, and presage future work employing these techniques online to boost BCI-AAC performance.

5.1 Feature Optimization

5.1.1 Introduction

Feature selection is a central component in the BCI system as a process that converts large quantities of EEG data into concise chunks of information to be fed to the classifier. The calculation of features allows for the classifier to work with a lower dimensional training set. This is one way to prevent the classifier from over fitting the training data, and it also speeds up the process of training the classifier. Feature selection can be fixed, by performing one set of transformations on the data to extract similar features for all individuals, or it can be optimized in such a way that the smallest set of the most discriminable features are used. Our lab has explored feature optimization techniques using exhaustive search methods to determine optimal electrode configurations [255], and eigenvalue methods for selecting robust frequency features in a motor-imagery task [185]. In this section, we perform feature optimization offline on data recorded from ALS patients and controls during the P300 and motor-imagery BCI tasks described in
Section 4.2.

As mentioned in Chapter 2, ALS and FTD produce overlapping patterns of degeneration that affect multiple regions within the central nervous system. Patients with primarily ALS show large white matter changes, while those with FTD exhibit substantial gray matter changes. However, little work has been done to describe the spatio-temporal changes to task-specific patterns of brain activity that occur as a result of disease state. In the following paragraphs, I describe the anticipated differences in feature optimization to be found between ALS patients and controls, as well those occurring as the result of cognitive impairment.

**P300 in ALS**

The work in Chapter 4 reviews some of the changes to the morphology of the P300 that occur due to normal aging as well as ALS. The most common changes involve reduced amplitude and longer latency of the P300 that occurs with normal aging, with similar morphological changes occurring in ALS and FTD. Although the topographical changes in P300 scalp distribution have not been clearly studied, Hanagasi et al., demonstrated that both patients and controls display a parietal maximum for the P3b component. Severens et al. also showed similar topography of the P300 between early-stage ALS patients and controls [47]. These studies describe a topographical stationarity to the P3b oddball potential despite significant neurological change. It should be noted that a different expectation occurs for the novelty P300, or P3a potential, which is elicited by a slightly modified novelty detection task. ALS patients display smaller amplitude of P3a over frontal sites, with no effect over central and parietal sites [222]. For this reason, Raggi et al. suggest that novelty P300 paradigms may be better suited to study the disruption in frontal networks accompanying cognitive decline in ALS [222].

When heterogeneity of the ALS population is accounted for, differences in the P300 become apparent. Ogawa et al. showed that patients with bulbar disease onset had significantly delayed P300 peaks compared to those with limb onset [260]. The cognitive decline sometimes co-occurring with ALS may also lead to alterations in the P300. According to Vierrege et al., dorsolateral prefrontal and anterior cingulate cortices have a role in the direction of attention. Some of the cognitive deficits found in ALS patients include reduced attention and inhibition, and PET studies have shown that regional blood flow in these brain areas are reduced [261]. FTD can cause delay in the P300 signal,
although this delay is not as large as in dementia of the Alzheimer’s type [262]. Others have noted that the delay in the P300 may be substantial in some demented patients, but non-existent in others, therefore ruling out the P300 as a reliable diagnostic tool [263]. However, in the case of BCI feature personalization, identifying longer P300 latencies on an individual basis could make the difference between a functional and non-functional BCI device. Therefore, in addition to the weaker amplitude of the P300 signal, we may expect the timing of the positive peak to be delayed and for compensatory reorganization to shift the location of the maximum in certain individuals with cognitive impairment.

Motor-imagery in ALS

The hallmark of ALS is degeneration of upper and lower motor neurons, but it is less clear how the large-scale, oscillatory activity of primary motor and associative areas of the brain are affected. Most studies have been performed with neuroimaging and blood flow analysis to describe the changes in brain function due to movement action and imagery. fMRI has shown increased activity during motor execution in regions primarily anterior but also posterior to the motor cortex, as well as in areas associated with motor learning, as in the basal ganglia and cerebellum [137, 264]. These changes have been confirmed with positron emission tomography (PET), showing increased activation in auxiliary motor areas, premotor, and parietal association areas during a motor task [265]. These studies point to an expansion of the output sensorimotor zone, which may represent cortical plasticity, or ongoing changes that occur to compensate for a loss of pyramidal cells in the motor cortex [264, 265], an effect that has also been observed in stroke patients [264]. Shifts to ipsilateral sensorimotor cortex that occur in these patients are another example of cortical reorganization [264, 266]. Conversely, lower regional cerebral blood flow was observed in prefrontal areas during both free and stereotyped motor tasks, which may be reflected as a correlate of neuropsychological deficits often found in the disease [265].

Again, disease heterogeneity adds nuance to this narrative. Lule et al. point out that this type of functional reorganization is not as evident in cases of purely lower motor neuron degeneration, and is a result of degeneration of both upper and lower motor neurons [264]. Kollewe and colleagues documented differential fMRI BOLD responses to hand and tongue movements in ALS patients with and without bulbar involvement. In
both groups of patients, there was an increase in the number of activated voxels during hand movement, but during tongue movement, only patients with bulbar signs showed reduced activation [267]. This study shows that compensatory mechanisms exist for the neurodegeneration of hand motor areas which do not exist in bulbar regions, possibly explaining the faster disease progression in bulbar-onset ALS [267].

Evidence for a different mechanism is at work in motor imagination. While some studies have shown stronger recruitment of premotor cortical areas, as well as cognitive areas related to motor planning, [268], others have demonstrated reduced BOLD activity in the left anterior parietal lobule, the anterior cingulate, and prefrontal cortex [269]. The reduction in activity seen in imagery may reflect the disruption of networks responsible for imagery outside of the primary motor cortex.

Expected Findings

We performed an exhaustive search for the optimal spatio-temporal features for a P300 and motor-imagery task in our patients. We were interested in finding if there was a difference in the optimal feature set between controls and patients with and without cognitive deficits. We expected to observe an overall decrease in feature robustness in the ALS disease state, but also anticipated a general shift in typical activations (central electrodes for motor-imagery and 300 ms delay for P300) to alternate locations and times, providing evidence that more personalized configurations may be optimal for users with neurodegeneration.

5.1.2 Methods

Exhaustive optimization was performed to determine the most effective feature vectors for classification in 25 ALS patients and 15 control participants. This optimization procedure was carried out for P300 and motor-imagery data separately. To facilitate the optimization procedure, all jobs were run on the Penn State computing cluster, with 12 processors dedicated to each person’s feature set.

The full feature set consisted of data from 19 (P300, linked-ears reference) or 6 (motor-imagery, Laplacian referenced) channels. For each trial, there were 5 time-binned EEG amplitude features for P300 or 10 frequency-binned EEG power features for motor-imagery, both of which had been down-sampled to reflect averages within the
range of 280-500 ms or 5-50 Hz, respectively.

The details of the optimization procedure are described for the analysis of P300 data, although the same method was used for the motor-imagery data. The only difference was the size of the spatio-temporal feature vector for each trial. Data from all four sessions was concatenated into a single matrix. Each row within this matrix contained the average EEG of a single stimulus code in one trial of P300 data. This row was associated with a class label, indicating whether it was a target code or a non-target code, and contained 19 channels with 5 time features evenly spaced over the 280-500 ms period, resulting in 95 features per stimulus code.

Two sets of optimization were performed for each task, due to computational restrictions. First, the set of time features were kept constant within each feature vector, while the full search of electrode combinations was performed. For the P300 task, this amounted to \((2^{19} - 1)\) 524,286 subsets of electrodes containing all time features to be processed through the classifier. As an example, on the first classification iteration only the first channel was used. The data matrix contained \(~1000\) rows, each being an EEG average over 10 flashes of a single stimulus code within a trial, and 5 columns, for the five time features in the first channel. This was repeated for all combination of electrode channels. Similarly, a second optimization was performed with the electrodes held constant and the space of the time features searched, for an additional \((2^5 - 1)\) 31 classifications. This method of separate electrode and time search avoided the need to do a full search on all possible combinations of 95 electrode and time features with \(~4e^{28}\) classifications.

The data matrix sent to the classifier contained a 1:8 ratio of target to non-target flashes. Data were parcellated into ten groups in order to calculate and average test set error from the classifier using 10-fold cross validation. In each of the ten iterations of the classifier, a random 90% of the stimulus codes (rows) were used for training the LDA classifier which used a 1:8 prior, while the remaining rows were used for classification. The test set error was defined as the fraction of trials in the test group which were misclassified. This error was averaged over the ten folds of the procedure to determine the average accuracy for that electrode combination.

For analysis, the number and type of features, specifically electrodes, time points, and frequencies, which yielded optimal classification were identified. To identify the spatio-temporal differences in feature space among participants, the classification error
from the iteration using only a single electrode, time point, or frequency was used. Other methods were attempted which used all subsets up to a certain dimension that included a specific feature, but the resulting trends were simply spatially averaged versions of the single feature scheme. We considered how these optimal feature sets changed across participant groups. Finally, we asked how the optimal feature set could be reduced in each group, while maintaining 80% of the optimal error. As an example, for an individual achieving an optimal error of 5%, the allowable error at a level 80% optimal was determined as (5.1).

\[
\text{Error}_{80\%} = 12.5 - (12.5 - \text{Error}_{opt}) \times 0.8
\]

\[
= 6.5\%
\]

where the theoretical random classifier error was 12.5%, due to the 1:8 prior. The configuration with the smallest number of electrodes reaching this level of accuracy was chosen as the reduced electrode set.

5.1.3 Results

The exhaustive search of feature space was expedited by the resource of the Penn State computing cluster. A maximum of four jobs were able to be run at a time, with each job allocated 12 cores. These 12 cores were utilized within MATLAB through the use of the \textit{parfor} parallelized computation loop. The bulk of the processing power went into performing classification on hundreds of thousands of feature vectors, and when parallelized in this way, each job experienced a roughly 6-fold increase in speed over the local computer. For the largest P300 datasets, a single subject’s data took roughly 1.5 hours to analyze, and as a whole, the time of total computation was reduced from roughly 384 hours to 16 hours.

P300

The optimization scheme used 10-fold cross validation of an LDA classifier, of which the version in MATLAB defaults to an equal prior on classes when not specified. Because the occurrences of target and non-target classes were unequal, a classifier with
equal expectation of each class resulted in a relatively high false positive rate (FPR = 1 - true negative rate). To bias the classifier to produce fewer false positives, the classifiers used in this analysis employed a 1:8 target:non-target prior. As can be seen in Figure 5.1, the classifier with this empirically-calculated prior produced lower overall error, due to the significant improvement in the true negative rate (TNR).

The first observation, apparent in Figure 5.1, was that the performance of the classifier improved with increasing complexity, or number of electrodes and time points included in the feature vector. The exception to this trend were those individuals who had a limited number of trials in their feature matrix due to fewer testing sessions. The relatively smaller number of trials made the classifier over fit the data when greater numbers of features were used, resulting in lower test set performance.

The critical task for this optimization exercise was to find whether the optimal electrode locations were different depending on the study group. The cross validation error for each subject that resulted from using only a specific electrode is shown in the top of Figure 5.2. The differences in classification across electrodes within most individuals was relatively small, although in some (S14, P39) the differential in test set error due to electrode location was quite large. When visualized across groups, some differences appeared between controls and patients (Figure 5.2), although ANOVA statistics for each electrode revealed no significant differences in electrode errors between groups, when Bonferroni corrected for 19 comparisons. In the control participants, the locus of the P300 signal had a distinct posterior maximum, with little discriminative information contained in the frontal channels, typical of a P3b oddball potential. The central, parietal, and
occipital regions defined the optimal regions for classification in these 15 individuals. On the other hand, in the ALS disease state, a more distributed response resulted from individuals benefiting from both frontal channels (P16, P34) and parieto-occipital channels (P31, P39). This distributed response occurred in the cognitively impaired group as well, although the general quality of the signals was lower and the resulting error higher in nearly all channels.

Control and patient participants displayed similar levels of optimal accuracy (Figure 5.3, left). Data were examined to determine what subset of the optimal configuration would be required to produce at minimum 80% of the optimal accuracy. The majority of those in the ALS group could achieve the accuracy at 80% their max using the same number of electrodes as control participants (Figure 5.3, center), with high performers in both groups generating 80% optimal performance using anywhere from 50-80% fewer electrodes (Figure 5.3, right). The reduction in required electrodes appears to scale with performance, indicating that there is high information content over few electrodes in the high performers, and little information content over many electrodes in the low performers.

Finally, optimal times for classification within the P300 task were examined by considering features from five time bins. Errors achieved by the classifier using only features from one time point are given for each participant in the top of Figure 5.4. Error was minimized for the control group during the time bins centered on 325 and 375 ms, and the lowest errors achieved by ALS patients were in the 475 ms window. The accuracy at each time point was not significantly different between groups, due to the high variability of accuracy within groups (Figure 5.4).
Figure 5.2: (Top) Error in each electrode for each participant. Errors are on the same scale, which shows relative performance between subjects. (Bottom) Absolute error in each electrode, averaged over participant groups. Red colors indicate high error rates, blue colors indicate low error rates.
Figure 5.3: P300 error from reduced electrode sets in patients (∗) and controls (o). (Left) Optimal error achieved over all electrode combinations plotted against the number of channels used in that optimal combination. (Center) The optimal error vs. the minimum number of channels needed to maintain 80% of the optimal error level. (Right) Optimal error vs. the percentage reduction in channels of the reduced set.

Figure 5.4: (Top) Error at each time point for each participant. (Bottom) Error at each time point, averaged over participant groups.
Figure 5.5: Electrode errors in each participant group from features from individual channels (top) and frequencies (bottom), on the same scale. For the electrode maps, red colors indicate high error, and blue low error.

**Motor-imagery**

Analysis for the motor-imagery task was performed in the same way, except feature vectors were constructed from Laplacian-referenced EEG data that had been transformed into frequency space and the spectrum from each trial parcellated into 10-5 Hz bins from 5-50 Hz. Similar to the results of the P300 optimization, increasing the complexity/channels fed into the classifier resulted in a lower error rate, although the error rates as a whole were higher in this task, near random level for many individuals. Due to the inconsistent performance of users on the motor-imagery task, these results should be interpreted cautiously.

Comparing participant groups, the cognitively impaired patients produced the highest error rates, while the cognitively normal patients produced the lowest error rates. The control participants, possibly because of the shorter training period of two sessions, performed within the range of the two patient groups. The relatively poor accuracies seen across groups masked the good performance of a minority of participants in the control
and cognitively normal groups. Averaging electrode performance over groups shows the relative difference in electrode importance between the three groups (Figure 5.5, top row). In cognitively normal individuals, the source of motor-imagery signals was in the C3/C4 pair of electrodes. These are the electrodes which putatively sample the activity of the primary motor cortex, which, along with the pre-motor cortices, are the primary generators of imagery signals. An ANOVA indicated that there was a significant difference in the accuracy of the three groups due to classification based on features extracted from channel C4, after Bonferroni correction for the six channels tested. A post-hoc analysis revealed that the controls and cognitively normal ALS patients achieved higher accuracy from the features extracted from this channel than those with cognitive impairment. The cognitively impaired patients, while performing overall at a random level, seemed to achieve little discriminability of motor-imagery signals in any channel tested.

Optimization was also run to evaluate which frequency features were most useful for discrimination of classes. The results are less clear in this case, but useful frequency features tended to occupy the mu and beta ranges (7.5 Hz, 12.5 Hz, 17.5 Hz, 22.5 Hz, and 27.5 Hz). Again, increasing the number of frequency features generally improved classification, but as in the case of the P300, this effect plateaued. Averaging over groups, a large frequency band over in the SMR range facilitated the classification of motor-imagery (Figure 5.5, bottom row). No groups appeared to use frequencies above 35 Hz to control the motor-imagery BCI, and an ANOVA indicated that there was a significant difference in the accuracy of the three groups due to classification based on features in the frequency bin centered at 22.5 Hz, Bonferroni corrected for the ten frequency bins tested. A post-hoc analysis revealed that the cognitively-normal ALS patients achieved higher accuracy from the features extracted at this frequency.

There were no obvious patterns for electrode retention at 80% optimal accuracy, but a large percentage of controls and patients saw no electrode reduction at the 80% level, due to the fact that one or two channels made up the optimal set (Figure 5.6). On the other hand, there were quite a few individuals who would be able to reduce the number of channels by half in order to maintain 80% of their optimal error level. These individuals were of a wide range of proficiencies, and included control as well as patient participants.
Figure 5.6: Motor-imagery accuracy from reduced electrode sets in patients (•) and controls (○). (Left) Optimal error achieved over all electrode combinations plotted against the number of channels used in that optimal combination. (Center) The optimal error vs. the minimum number of channels needed to maintain 80% of the optimal error level. (Right) Optimal error vs. the percentage reduction in channels of the reduced set.

5.1.4 Discussion

The reason for performing feature optimization for a BCI task is two fold. First, it allows the classifier to account for diversity among individuals by utilizing unique features, which are sometimes outside the realm of ‘normal’ ranges expected for these tasks. With this optimal feature set, the user can perform at a higher level than if the feature extraction were a fixed process. Additionally, the dimension reduction accomplished by feature selection makes simple linear classifiers more robust to over-fitting, decreases online computation time, and requires the use of fewer sensors.

The optimization performed in this section is relatively inelegant compared to other feature reduction schemes, although the method used guarantees finding the feature set producing the lowest error. The obvious drawback is the increase in computational demands, which was mitigated by parallelizing the optimization procedure. As is, this technique is reasonable to do in a multi-session BCI environment, as the current procedure took at most three hours per participant. However, if immediate feature reduction is required, stepwise algorithms for feature selection are regularly used, as well as dimensionality reduction techniques, both of which have their downsides.

The optimization performed over electrodes and time points in the P300 task indicate that there may be subtle differences in the optimal feature set produced by ALS patients. As expected from the literature, the times closer to the end of the 500 ms window were most useful for classification in patients, while times in the 300-400 ms window yielded the lowest error in controls. Additionally, ALS patients exhibited a wider distribution
of P300 features in frontal electrodes, while in controls, discrimination between classes was exclusively found in centro-parietal regions. For all performers, the relative benefit to feature reduction scales with performance. For high performers, the use of many channels produces redundancy. These users can maintain high levels of performance even after eliminating up to 80% of the electrodes, while low performing users need an increased feature space to maximize the P300 averaging effect.

Not much can be conclusively said about feature differences across groups in the motor-imagery task. This is due to the majority of individuals unable to effectively modulate their motor rhythms in response to hand imagery. The differences between groups represent the biases introduced by one or two successful motor-imagery users, averaged over the larger cohort of near-random performers. Among high performers in control and patient groups, the channels with greatest discriminability were the central electrodes C3 and C4. Although there was no observed shift in frequencies or electrode locations of discriminable data in the motor-imagery task due to ALS, there is substantial reduction of the feature space possible by utilizing only these two channels for classification.

5.2 Features describing neural connectivity for BCI

5.2.1 Event-related coherence changes

Introduction

To define the neural features that accompany a mental task, the majority of current BCIs use individual sensor activations; firing rates from local populations of neurons, regional metabolic changes, or the transient changes in the brain’s electromagnetic field recorded at the level of the cortical surface and scalp. By limiting features to the activity of individual sensors, information about neural interaction is lost. With that in mind, diverse methods have been used to define regional coupling, causality, and synchrony in EEG data. Reviews on the nature and applicability of these methods, which range from coherence [94], Granger causality [270], phase synchronization [271], as well as nonlinear descriptors of synchrony [272] are given elsewhere [273, 274, 275, 276].

Evaluation of these tools in the analysis of nonlinear EEG data has revealed potential pitfalls in their interpretation [276], stemming from the inherent variability in
the resting-state network compared to changes due to functional coupling. However, when combined with traditional BCI metrics, connectivity features have been shown to provide an additional boost in performance [277, 278]. In this analysis, we consider a simple measure of coherence, in comparison to traditionally used power features utilized in a motor-imagery task. We determine whether there are significant differences between features across participants and whether this translates to improved classification of imagery state.

Methods

Based on the findings of the online BCI study described in Section 4.2, we were interested in determining the utility of coherence features across three groups: control participants, cognitively normal patients, and cognitively impaired patients. Of interest was whether features describing the interaction between electrode sites in the mu and beta bands provided additional information for classification, and how this information varied across group type.

Spectral features were calculated using Welch’s method, implemented in MATLAB. Power spectral density features were calculated from the same six Laplacian-referenced channels used online, while coherence was calculated for each of the 19 electrodes referenced to linked earlobes. For each method, the time period from 0-3 seconds after the presentation of the cue was used, encompassing the assumed period of imagery. Windows of one second length were passed through a hanning window, and the periodogram calculated. The window was shifted by half a second and the process repeated, with the output being the power spectral density, or average of these periodograms normalized by the sampling frequency. The magnitude squared coherence was calculated as the ratio of the cross spectral density magnitude to the product of the autospectral densities.

\[
C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f))}
\]  

(5.2)

In addition to describing changes in power and coherence among participants, we compared the accuracy achieved in offline classification using these features in a 10-fold cross validation scheme. The averages of the test set errors over the ten folds using typical power features vs. coherence features were compared. There were twelve features in the case of band power, accounting for the average band power in the six
channels over two frequency bands. Although there were 30 total coherence features between the 6 channels F3/4, C3/4, and P3/4, this was reduced to 9 features, retaining all connections to and from channels C3 and C4. Accounting for both frequency bands, this was 18 features total for the coherence data matrix. This reduction in the size of the coherence feature vector was done to minimize error discrepancies between feature types due to the size of the feature vector. For each individual, cross validation error was determined for each of the feature sets, and an error differential was calculated as the difference between the two errors. This error differential was compared across patient groups using a one-way ANOVA to determine whether there was an advantage of using coherence or power features depending on the user’s neurological state.

**Results**

The power differences across left and right trials followed the pattern expected from SMR desynchronization, with the best performers achieving differential power changes in multiple channels per hemisphere (Figure 5.7). Significant differences in mean power between left and right trials at the $p < 0.05$ level, as determined by repeated t-tests, are shown as colored circles at the electrode, with larger circles indicating where this interaction was significant after the test statistic had been Bonferroni corrected for six inferences. In general, channels in the left hemisphere (F3, C3, P3) displayed increased power during left trials compared to right trials as a result of mu (and beta) suppression in these channels during right hand imagery. The opposite was true for right hemisphere channels. Not shown is the plot for beta power differences, in which participants such as S16, P36, and P39, who achieved relatively high online motor-imagery accuracies, showed more consistent patterns of hemispheric desynchronization. As expected, those with poor performance in the motor-imagery task failed to display significant mu power modulation depending on trial type.

There are similarities and differences apparent when comparing power to coherence changes in these individuals. In Figure 5.8, coherence differences between left and right trials in the mu band can be seen for patients and controls. Significant differences in mean coherence between left and right trials are shown as lines between two of the 19 unipolar EEG channels. Thicker lines indicate where this interaction was significant after the test statistic had been Bonferroni corrected for 171 ($[N_{elec} \times (N_{elec} - 1)]/2$) electrode pair inferences. Of note are the large hemispheric differences in coherence seen
Figure 5.7: Power spectral density in the mu frequency band in each Laplacian-referenced channel, compared across left and right motor-imagery trials in controls and patients. Significant power differences across trial type, as determined by repeated 2-sample t-tests, are indicated by colored circles at each electrode. Larger circles indicate significance after Bonferroni correction.

for the high-performing individuals. The pattern of these differences indicated increased coherence for right hemisphere channels during left trials and increased coherence for left hemisphere channels during right trials. As opposed to the clear hemispheric differentiation of coherence observed in high-performing subjects, there were large scale, whole brain coherence changes between left and right trials in other participants, e.g. P02, P05, P29, and P43. Interestingly, these users all achieved relatively low online performance, and did not demonstrate strong hemispheric power differences in either the mu or beta bands.

To determine whether coherence features could be useful in motor-imagery clas-
Figure 5.8: Coherence in the mu frequency band in each channel, compared across left and right motor-imagery trials in controls and patients. Significant coherence differences across trial type, as determined by repeated 2-sample t-tests, are indicated by lines between electrodes. Thicker lines indicate significance after Bonferroni correction.

Classification, an offline classifier was implemented. The classifier compared the errors achieved when using typical power features vs. coherence features. The results for each individual, shown in Figure 5.9, indicated high variability both in overall performance and relative utility of coherence features. The error differential between classifiers using power and coherence features was not different across groups of control participants, cognitively normal patients, and cognitively impaired patients, although certain patients in the cognitively impaired group did achieve substantial differential benefit from using the classifier based on coherence features. Notably, patients P06 and P43 saw improvement in their test set errors from 46.25% to 34.5% and from 47.9% to 37.1% when switching from power to coherence features for classification.
Figure 5.9: Classification errors from offline cross validation using power and coherence features separately. (Bottom right) Differences in these errors averaged over participant groups. Error differentials are not different between groups.

Discussion

The spectral measure of coherence has been previously used to describe the interaction between electrode sites during motor planning and execution. Rappelsberger et al. described a contralateral coherence increase in the alpha band in premotor and motor areas just prior to movement onset [94]. Leocani et al. described event related coherence increases due to self-paced movements that were localized spatio-temporally with event related desynchronization in frontotemporal areas [93]. The contralateral increase in coherence described in these studies was similar to what was observed in the high-performing motor imagery users in this analysis. Furthermore, coherence changes in response to imagery overlapped spatially and temporally with suppression of SMR power.

For most individuals there was no appreciable difference in intensity of coherence between imagery states. Of the subjects which did show consistent coherence changes,
the high-performing users demonstrated clear hemispheric lateralization of this effect. However, there were also individuals who demonstrated whole brain changes to coherence values which were hemisphere invariant. Although the functional interpretation of whole-brain coherence changes is less clear, for both of these groups, these features should theoretically allow for classification into left and right states. Furthermore, coherence in higher frequency bands has been demonstrated to produce complementary and somatotopically specific changes as a result of motor action [279]. Functional connectivity changes at the gamma (30-100 Hz) level were not assessed in this analysis, although they may prove useful in this type of imagery task.

Overall, power features proved to be more useful for the classification task. This does not mean that for certain individuals (e.g. P06, P43) there existed a greater level of discriminable information in features of coherence compared to traditional power changes. It should be noted that both of these individuals were indicated as cognitively impaired by the ALS-CBS, thus generating the non-significant trend that coherence features were more useful for those with cognitive impairment compared to those with intact cognition. Despite this trend, the majority of individuals in the cognitively impaired group performed at a random level, and it would be unfounded to make any statements about the optimal features to use in this group of individuals. That being said, on an individual level, feature selection including coherence measures could prove beneficial in classification. In the case of P06, discrimination between classes was due to hemispheric lateralization of coherence found in the beta band, while in P43, there existed a spatially diffuse increase in coherence during right trials over a large range of frequencies. Measures of coherence, although overlooked in healthy individuals, may be a more robust mechanism for control in those with diffuse neurodegeneration, although further confirmation of this required. Most likely, a unique combination of both coherence and power features will result in the optimal classifier.
5.2.2 State estimation based on neural field modeling

Introduction

In line with our goal for device personalization, we sought to use additional features as input to the classification system. In this section, we move beyond empirical measures of functional coupling discussed earlier in the chapter, in order to use the accumulated knowledge about brain anatomical networks that produce large scale electrical activity. In doing so, we constrain the estimates of population functional connectivity to a plausible subset of realistic interactions, thus reducing the influence of irrelevant features. Neural models are effective tools used in neuroscience to accomplish such data assimilation, and can provide a physiologically-plausible structural basis for estimates of brain connectivity.

Lumped-parameter neural models are able to emulate different aspects of brain activity, from thalamo-cortical generators of occipital alpha rhythms [280, 281, 282], to high frequency rhythms and epileptic activity [283], sensorimotor generators [284], and oddball responses [285]. Although the dynamics of these models are not nearly as complex as those defined using populations of individual neurons [283, 286], they have the benefit of being parameterized by a low number of variables. The original ‘neural mass’ models [280, 287], consisting of discrete interacting cortical structures, have recently been generalized to occupy a continuum, resulting in ‘neural field’ models [288, 289]. Continuum models such as these are adept at defining activity found between cortical columns, such as traveling waves.

Neural mass/field models integrate the postsynaptic potentials of a neural subpopulation, and transform this membrane potential into a spike train that influences neighboring populations. These time-evolving networks are then forward mapped to sensor-space and fit to bioelectric data. Our proposed neural field model follows the framework given in Freestone et al. [290]. Their procedure improves on other models by converting the continuous field potential into a discretized state-space by Galerkin projection onto a set of basis functions, followed by tracking of states and parameters using a variation on the Kalman Filter. This allows for assimilation of neural data in order to track parameters which represent connectivity strengths between populations, as well as time constants. With this model, we attempted to infer differential parameter distributions that exist due to the intentional states of a BCI paradigm.
Methods

The estimation framework of Freestone et al. [290]

To develop a set of BCI features representing connectivity between brain regions during a motor-imagery task, we looked towards the work of Freestone et al. [290]. This group applied a neural field model for the purpose of estimating connectivity kernels between neural populations recorded at the cortical surface. In this section, their framework, which was originally suited for microelectrode recordings on the cortex, was modified to enable estimation of connectivity strengths between larger scale populations measured by EEG.

The estimation framework is based on knowledge of the generative model that is assumed to underly electrical phenomena in the brain. Their discretized integro-difference equation (IDE) describes the evolution of a neural field on the surface of the cortex at locations $\mathbf{r}$ as

$$
\mathbf{v}_{t+T_s}(\mathbf{r}) = \xi \mathbf{v}_t(\mathbf{r}) + T_s \int_{\Omega} w(\mathbf{r}, \mathbf{r}') f(\mathbf{v}_t(\mathbf{r}')) d\mathbf{r}' + e_t(\mathbf{r})
$$

(5.3)

where $\mathbf{v}_t$ is the membrane potential of the spatially distributed neural field over domain $\Omega$ at time $t$, $T_s$ is the time step of the field, and $\xi$ represents the memory in the field, and is related to the time step and the synaptic time constant of the underlying dynamics. These dynamics are represented by a sigmoidal function $f$ that transforms the membrane potential of a population into a firing rate as in

$$
f(\mathbf{v}(\mathbf{r}', t)) = \frac{1}{1 + \exp(\varsigma(\nu_0 - \mathbf{v}(\mathbf{r}', t)))}.
$$

(5.4)

This function is defined by the firing threshold $\nu_0$ and the slope $\varsigma$. To update the membrane potential from one state to the next, the firing rate is convolved with the spatial connectivity kernel $w$, which is composed of three Guassian functions $\psi(\mathbf{r}, \mathbf{r}')$, scaled by amplitudes $\theta$, following

$$
w(\mathbf{r}, \mathbf{r}') = \psi^\top(\mathbf{r}, \mathbf{r}') \theta.
$$

(5.5)
Here, $\psi$ is a 3x1 vector of Gaussian basis functions, $\theta$ is a 3x1 vector of scaling parameters, and $^\top$ represents the matrix transpose. It is assumed that the width of the connectivity basis functions are known, and the scaling parameters are to be estimated from the data. For the simulated field, by which the authors assess the efficacy of their framework, $\theta = [100 -80 5]^\top$, generating a connectivity kernel between populations that facilitates center excitation, surround inhibition, and weak long-range excitation, sometimes referred to as a Mexican Hat function (Figure 5.10). Finally, the IDE includes a disturbance term $e_t$ defined by spatial covariance $\gamma(r - r')$.

In the paper of Freestone et al., the field of membrane voltage is propagated forward into electrode space containing sensors $y$ that exist at locations $r_n$, according to

$$y_t(r_n) = \int_\Omega m(r_n - r') v_t(r') d r' + e_t(r_n). \quad (5.6)$$

This equation describes the mapping between the membrane potential and electrode space, with sensor noise $e_t$ defined by a multivariate normal distribution, and using a Gaussian observation kernel $m$, which defines the falloff of sensitivity of the electrode with width $\sigma^2_m$.

$$m(r_n - r') = \exp \left( - \frac{(r_n - r')'(r_n - r')}{\sigma^2_m} \right). \quad (5.7)$$

Thus far described has been the underlying generative model for field potential and sensor activation. In the application of this framework to EEG data, we had available sensor activations $y(r_n)$, and no information about the underlying field. However, from this model we were able to perform a projection into state space so that estimation of the underlying field could be performed. In order to perform this estimation, Freestone et al. decompose the neural field $v_t$ into a set of states $x_t$ and associated basis functions $\phi(r)$ of the form

$$v_t(r) \approx \phi^\top(r) x_t \quad (5.8)$$
\[
\phi(r_b - r') = \exp\left(-\frac{(r_b - r')^\top (r_b - r')}{\sigma_\phi^2}\right). \tag{5.9}
\]

Here, the basis functions with centers at nodes \( r_b \) are Gaussian shaped with width \( \sigma_\phi^2 \). To project the field into state space, the basis decomposition of \( \nu_t \) is substituted into the IDE model, followed by pre-multiplication of the equations by the field basis functions \( \phi \). Following simplification provided in [290], the state space model takes the following form:

\[
x_t + 1 = \int_\Omega \Psi(r') f(\phi^\top(r')x_t) d\theta + \xi x_t + e_t \tag{5.10}
\]

\[
y_t = Cx_t + \varepsilon_t \tag{5.11}
\]

With relevant state-space matrices taking the form:

\[
[\Psi(r')]_{ij} \overset{\Delta}{=} T_s \Gamma^{-1} \int_\Omega \phi(r) \psi_i(2c_j + r' - r) d\mathbf{r} \tag{5.12}
\]

\[
\Gamma \overset{\Delta}{=} \int_\Omega \phi(r) \phi^\top(r) d\mathbf{r} \tag{5.13}
\]

\[
e_t \overset{\Delta}{=} \Gamma^{-1} \int_\Omega \phi(r) e_t(r) d\mathbf{r} \tag{5.14}
\]

\[
C_{ij} \overset{\Delta}{=} \int_\Omega m(r_i - r') \phi_j(r') d\mathbf{r}' \tag{5.15}
\]

These four entities, which are employed in the estimation procedure, are predefined analytically before entering the iterative estimation loop. The job of the estimator is to assimilate the data from the sensors with knowledge about the underlying connectivity among populations of the unobserved field in order to find the states \( x_t \), the connectivity kernel parameters \( \theta \), and the synaptic dynamics \( \xi \).

This is accomplished with a two-part iterative algorithm, in which state estimation is performed by the unscented Rauch-Tung-Striebel smoother (URTSS), and parameter estimation is done through minimization of the squared error of the state prediction. For each trial of data, the URTSS is run forward and backward to generate the best estimate of the states given the current estimate of the connectivity kernel. Then the parameter vector is updated with a least squares estimator which minimizes the sum of the
Figure 5.10: 2D representation of field space (black), state space (blue), and sensor space (red). The points at each level represent the relative spacing of nodes. Each node of the field (black points) has assumed dynamics which transform the membrane voltage into a firing rate. The black line shows the connectivity kernel, or the influence the central node has on the surrounding nodes in the field. The blue line represents the basis function used for the Galerkin projection; it shows the relative contribution of the field nodes to the central basis node. The red line shows the state space representation of a single row of the observation matrix, showing the influence of each basis function on the amplitude of the sensor (filled red circle).

Squared errors of the predicted state update. Explicit formulations of the URTSS and least squares algorithms are given in [290]. The estimation procedure stops when the parameters change < 5% from their previous estimate. For replication of the simulations performed by the original authors, 50 instances of the neural field were generated, resulting in 50 parameter estimates.

**Modifications to the estimation framework**

The transition from cortical measurements to the scalp EEG required significant reformulation of the model in order to accommodate for incongruities with scalp data. The domain of EEG space was an order of magnitude larger, scaling from tens of millimeters to tens of centimeters. Many variables were able to scale with the increasing dimensions, like the discretization of the field, and the spacing and width of the basis functions. Others, such as the connectivity kernel width, sensor and field disturbance variance were
kept in proportion. Two data sets were available for testing, one with a high-density 129 electrode montage, and another with a low density 19 electrode montage. Modifications are summarized in Table 5.1 and Figure 5.11, and describe primarily accommodations for the increased spatial domain of the field, although there were changes to the sampling frequency as well the formulation of the connectivity kernel $w$. All changes are detailed further here:

- **Dimensions of field, bases, and sensors.** All dimensions related to the neural field were increased by ten-fold. Dimensions originally $\pm 10$ mm in the $x$ and $y$ direction have been modified to $\pm 10$ cm.

- **Node spacing.** Spacing of the node locations ($r$) increased, along with the spacing of the basis functions ($r_b$) and sensors ($r_n$). The locations of the sensors were constrained to a grid governed by the actual locations of the electrodes in their respective geodesic net and 10-20 system configurations. From the high-density recording, 46 of the central electrodes were retained in the estimation procedure, while all 19 of the electrodes were used in the low-density estimate.

- **Parameterization of observation and basis functions.** Width of observation kernels and field basis functions increased in proportion to the spacing of nodes in order to achieve similar falloff with distance. Due to the increased distance between sensors, the observation kernel width was increased so that each electrode sampled from a larger proportion of field space. This is consistent with the greater spread of the field to sensor space in EEG compared to microgrid measurements. The new spacing and width of the basis functions limited the maximum spatial frequency able to be reconstructed by the field to 0.06 cycles/cm. In the original framework, this maximum frequency was 0.12 cycles/mm (this has not been scaled by a factor of 10, so in effect, the spatial resolution of the original field was double the new field, due to increased basis function spacing).

- **Parameterization of the connectivity kernel.** The variance of the connectivity kernel was doubled in the case of the high-density EEG field, and tripled in the case of the low-density EEG field. These increases were based on the spatial extent of the resulting connectivity kernel in terms of the influence of nearby basis functions.
Table 5.1: Parameters used in the neural field estimation framework. The left column replicates the values of Freestone et al. [290], while the values in the remaining columns indicate changes to accommodate for high- and low-density EEG data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Freestone et al. [290]</th>
<th>High-Density EEG</th>
<th>Low-Density EEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial domain, $\Omega$</td>
<td>$\pm 10$ mm</td>
<td>$\pm 10$ cm</td>
<td>$\pm 10$ cm</td>
</tr>
<tr>
<td>Spatial discretization, $\Delta$</td>
<td>$.5$ mm</td>
<td>$1$ cm</td>
<td>$1$ cm</td>
</tr>
<tr>
<td>Time step, $T_s$</td>
<td>$.001$ s</td>
<td>$.004$ s</td>
<td>$.004$ s</td>
</tr>
<tr>
<td>Synaptic time constant, $\tau$</td>
<td>$0.01$ s$^{-1}$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Firing threshold, $v_0$</td>
<td>$1.8$ mV</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Activation function slope, $\zeta$</td>
<td>$.056$ mV$^{-1}$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Connectivity kernel widths, $\sigma_p$</td>
<td>$[1.8, 2.4, 3.6]$ mm</td>
<td>$[3.6, 4.8, 12]$ cm</td>
<td>$[5.4, 7.2, 18]$ cm</td>
</tr>
<tr>
<td>Number of sensors, $n_y$</td>
<td>196</td>
<td>46</td>
<td>19</td>
</tr>
<tr>
<td>Distance between sensors, $\Delta_y$</td>
<td>$1.5$ mm</td>
<td>$2.66$ cm</td>
<td>$5$ cm</td>
</tr>
<tr>
<td>Observation kernel width, $\sigma_m$</td>
<td>$0.9$ mm</td>
<td>$1.6$ cm</td>
<td>$3$ cm</td>
</tr>
<tr>
<td>Observation noise variance, $\Sigma_e$</td>
<td>$.1 \times I_{n_y}$ mm$^2$</td>
<td>$.5 \times I_{n_y}$ cm$^2$</td>
<td>$1.67 \times I_{n_y}$ cm$^2$</td>
</tr>
<tr>
<td>Disturbance spatial cov. width, $\sigma_f$</td>
<td>$1.3$ mm</td>
<td>$1.3$ cm</td>
<td>$1.3$ cm</td>
</tr>
<tr>
<td>Disturbance variance, $\sigma_d$</td>
<td>$0.1$ mV</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of basis functions, $n_x$</td>
<td>81</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>Dist. between basis functions, $\Delta_\phi$</td>
<td>$2.5$ mm</td>
<td>$5$ cm</td>
<td>$7.5$ cm</td>
</tr>
<tr>
<td>Width of basis functions, $\sigma_\phi$</td>
<td>$1.58$ mm</td>
<td>$3.12$ cm</td>
<td>$4.68$ cm</td>
</tr>
</tbody>
</table>

- **Synaptic dynamics.** We assumed the same model for field generation as the original model. The membrane potential of the field is transformed via a parameterized sigmoidal function to a firing rate. This firing rate influences neighboring neural populations through Mexican hat connectivity. Amplitude of the EEG data was scaled to the same size as the simulated data, so that the sigmoid parameters acted on the data in a similar manner.

- **Sampling frequency.** The simulations were carried out at 1000 samples/second. The EEG recordings were acquired at 250 samples/second, which prompted us to change the sampling parameter in the model. This quadrupling of the time step of the field affected both the time constant parameter $\xi$ and the state space connectivity matrix $\Psi$.

To assess the validity of the estimation procedure, we first looked at the reproduction of the sensor activations $y$ from the reconstructed states $x$. This allowed us to determine whether the iterative state tracking procedure effectively assimilated the data fed in by the sensors or if errors accumulated, resulting in unreasonable estimates for the parameters.
Figure 5.11: Modifications to the estimation framework are given for the high-density EEG recordings (center column) and the low-density recordings (right column), with the original simulation framework on the left. The top row shows the extent and spacing between the nodes in the field, in mm. The second row displays the locations of the field states in blue, as well as the associated basis function for a single state. The third row displays the locations of the sensors in red, as well as the observation kernel for the center-most sensor. The bottom row once again shows the locations of the states, along with the connectivity kernel of the state at the center of the field. All rows use the same color scale except the last row, for which the amplitude of the EEG connectivity kernels are four times larger due to the increased time step.
The main goal was to find the distributions of kernel amplitude parameters, $\theta$, across imagery states. A major issue for assessing the validity of the kernel estimates was that there was no information about the actual nature of the underlying connectivity kernels. The expectation was that the estimated kernels would display a similar center excitation, surround inhibition pattern, as this was not expected to change with scalp recordings. From a standpoint of utility as a feature for classifying individual BCI trials, the values of the parameters were not as critical as the consistency of estimates across trials. We performed classification on the first two kernel parameters of $\theta$, describing the amplitude of center and surround influence, in order to determine if the active state of mental imagery could be distinguished from the state of rest. This required the classifier to be trained on left and right imagery data in a single class, with no-go data belonging to the other class. Ten-fold cross validation of the LDA classifier was performed on the roughly 480 (160 per cue) trials for each individual, and the test set error was reported.

**Results**

**Replication of the simulated estimation framework**

The first goal of the neural field estimation procedure was to replicate the findings of Freestone et al., which was accomplished through generation of a custom MATLAB code. As can be seen in Figure 5.12, the estimated kernels for three different field simulations, representing three cases of field connectivity, matched those estimates presented in the original paper. Between the original and reproduced model, there were no significant differences in the distribution of any of the $[\theta; \xi]$ parameters yielded by the estimation scheme.

The influence of the synaptic dynamics on kernel reconstruction were also replicated as in [290]. The ideal values for $\zeta$ and $v_0$ were close to the .56 and 1.8, with these parameters producing the lowest root mean squared error between the simulated kernel and the estimated kernel. Critically, increasing the field size does not substantially alter these dynamics, which was determined by repeating this simulation on a field which had been enlarged by a factor of ten.

Reconstruction fidelity could be directly assessed in the case of the simulated field by comparing the estimated states, transformed by the field basis functions ($\phi^\top x$) with the underlying true field, $v$. As seen in Figure 5.13, field reconstruction captured the
dynamics at large spatial scales, but missed higher frequency components of the field. This was an inherent limitation of the reconstructed estimate due to the greater spread of basis functions compared to the nodes of the field. The field and its reconstruction were analyzed to determine their spectral distributions. Power spectra in two dimensions were computed using the 2D Fourier transform (Figure 5.14). In the simulated case, these power spectra were able to be calculated for the field $\nu$, as shown in equation 5.17.

$$\hat{\nu}(k, l) = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} e^{-2\pi i \left(\frac{k}{X} + \frac{l}{Y}\right) x} \nu(x, y)$$ \hspace{1cm} (5.16)$$

$$P_{\nu} = |\hat{\nu}(k, l)|^2$$ \hspace{1cm} (5.17)$$

Evaluation of the power spectral density demonstrated the spatial spectral falloff for the observations from the simulated electrode grid ($P_\nu$) and the reconstructed subspace ($P_{\phi x}^\top$). These two should roughly match because the configuration of the basis functions are chosen to account for the full spatial bandwidth of the observations. As can be seen from the figure, the relatively greater spacing of the basis functions generated a loss of

Figure 5.12: (Top) Selected simulations (A-C) using three different kernel amplitudes, from [290]. (Bottom) Replication of this framework produces nearly identical kernel estimates. The solid line is the actual kernel used to generate the data, the dashed line is the mean of the trial estimates, and the shaded region is the confidence interval for the mean.
Figure 5.13: Field reconstruction can be directly assessed with a simulated field. (Top) The underlying field \( \nu \) is subtracted from the reconstruction in state space, transformed by basis functions \( \phi \). (Bottom) Electrode activations and the estimated states projected into sensor space, and the resulting error.

Figure 5.14: Plots of spectral falloff for each field from left to right: The simulated neural field, the electrode space, the reconstructed field, the error between the field and the reconstruction. The reconstruction captures the dynamics at large spatial scales, but misses higher frequency components of the field.

high frequency information in the reconstruction. This analysis provided an example of the spatial frequencies that could be expected to be lost in the estimation procedure. Of course, these high frequency losses could be mitigated by choosing a higher density of basis functions, at the cost of significantly higher computational complexity. In the case of estimation using EEG data, the reason a greater number of basis functions was not used was due to the problems encountered with over-fitting sensor data to a higher-dimensional basis space.

**Results with EEG data**

The modifications to the estimation framework involved altering the spatial dimensions of the field, the number of node, basis, and sensor elements, and the extent of the con-
Figure 5.15: Observed sensor activity at one time point (left column), estimated states projected into sensor space (middle column), and the resulting error for high density (top) and low density (bottom) EEG data. The state-space formulation tracks the gross features of the field, but cannot resolve observation dynamics on smaller spatial scales.

nectivities between these elements (Figure 5.11). With the EEG data, there was no underlying field to assess the reconstruction fidelity of the estimation, since there is no information about the cortical generators that lead to the measured signal. We could, however, compare the sensor activations to a projection of the estimated states into sensor space (Figure 5.15). This provides a check that the states and parameters of the model were reconstructing the observed dynamics, and that there were no systematic errors accumulating in the estimation procedure. The relative ratio of the error in the observed field was on a similar order for the high density recordings, although the error increased as the number of basis functions decreased, as in the case of the low density field.

We could also assess if the basis functions were sensitive to the frequency content of the observations. As can be seen in Figure 5.16, the spatial frequency content of the observation kernel was still above the -3 dB point at the Nyquist frequency. However, the falloff pattern was similar to the falloff of the basis functions, which both scaled by a factor of two when going from the high-density EEG field to the low-density field.

As can be seen from Figure 5.17, there was a tendency for the Mexican Hat connectivity kernel estimates to exhibit a negative central peak. This was the case for all 5 subjects with the low-density EEG recordings. The exception to this was high-density EEG from S002, whose kernel distribution displayed a positive central peak. Interestingly, this individual also achieved the greatest consistency among kernel estimates and
exhibited the greatest differentiability of kernels between mental states.

For both S001 and S002, the magnitude of the kernel estimate is largest for the no-go trials. For individuals with low-density EEG recordings, this trend appeared to be reversed, although in these individuals there were no significant differences in kernels between the three mental states. A reasonable question to ask, especially of the kernel estimates of participant S002, is whether these features could be used in a classifier to predict state of mental imagery in an online BCI. For the 10-fold cross validation of a classifier trained on imagery and non-imagery trials, the test set classification errors aligned with the variability of the kernel estimates. The lowest error of 19.4% resulted from the relatively high discriminability between no-go and imagery trials in participant S002. On the other hand, all of the participants with low density recordings generated errors >40%, or near random.

**Discussion**

Freestone et al. provide a novel framework for tracking the states of a neural field, while simultaneously estimating the parameters of spatial connectivity within a simulated cortical field. We set out to modify this framework to estimate relative connectivity strengths between states of the neural field underlying motor-imagery EEG recordings. With our modified framework, we achieved limited success in differentiating between
the active states of imagery and rest periods using filtered EEG data. Although there were not as many subjects with high-density EEG recordings available, the use of a greater number of sensors appears to be better suited for distinguishing between mental states. We anticipate the estimation of the low-density EEG scheme suffers because of the focal nature of the signal of interest. While a large array of electrodes was used in the space of state estimation, only a small portion of these electrodes are expected to sample from the region of scalp exhibiting task-relevant changes in spectral power and coherence. Certainly, this framework would be more appropriately suited for a high-density EEG array or even an ECoG grid, both of which pack a higher density of sensors potentially along a region of brain that is engaged in the motor-imagery task.

There are other aspects of the modified estimation scheme which have the potential for improvement. Many of the modifications made to the model were straightforward, such as alteration of the size and spacing of elements in estimation space. However, the model was originally designed for estimating field connectivity from the measurements provided by a microelectrode grid, and the difference in the spatial scale of the model also requires additional experimental confirmation of parameters, and re-imagining of the assumed generative dynamics.

The first aspect of the model to be addressed in the future is the generative model...
for the dynamics of the unobserved field. As opposed to a microgrid array, which may reasonably record from neuronal ensembles that experience the influence of center excitation and surround inhibition from its neighbors, the large scale, synchronous postsynaptic activity recorded by EEG may require a different implementation. A Gaussian or Laplacian connectivity kernel, rather than the Mexican Hat type, may better describe the influence of nearby cortical activity. The fixed width of the connectivity kernel bases may pose additional limitations by not allowing interactions to occur on multiple scales.

In a follow-up paper, the authors of the original model utilized a family of B-spline wavelets and associated scaling functions to create a multi-resolution neural field estimation framework [291]. This modification allows for inference of connection strengths on multiple scales, which would be of great use with EEG data, as often the precise scale of interaction is unknown.

The convolution of this kernel with the sigmoidal firing rate function may also be inappropriate for use when the characteristic scale is an order of magnitude larger. Simplification of the connectivity kernel may be less desirable than a more realistic neural mass model for SMR rhythm generation that includes multiple neuron types [285]. Other groups have used discretized damped linear wave equations to describe current source density [292] fit to EEG, or estimated parameters using the spectra of the EEG rather than the time series representation [284].

In addition to the changes in the proposed model for field generation, there are some additional aspects of the current model that are lacking. Specifically, the omission of time delays becomes more critical at larger spatial scales. Delays in signaling among distant neural populations may be included through implementation of a defined propagation velocity. Finally, the observation kernel in the modification remained of the Gaussian form. A more appropriate observation kernel for EEG applications might take the form of a lead field matrix, including information about the location of the source, the electrode, and the intervening media. Ideally, the observation kernel would be created on an individual basis to represent personal differences in geometry and distribution of skull tissue.

One use of this type of data assimilation is to track the single trial parameter differences for the purpose of classification, as in the case of a BCI. This will always be a more difficult task than finding differences in the average connectivity over a number of trials, and will require reformulation of the estimation algorithm to make it operate
in an online scenario. While this is a worthwhile goal, this model could also be used to assess the functional changes that occur as a result of the neurodegeneration inherent to ALS. Using the tools of regional connectivity elaborated in the previous section, this estimation framework may be employed to describe motor and non-motor related network changes in the ALS brain.
6.1 Personalized deployment and design of BCI devices in ALS

6.1.1 Prospects for gating

In Chapter 3, trial gating is demonstrated to improve accuracy and speed of the BCI interface for certain individuals. All three of tested gating mechanisms are derived from EEG data that is supplementary to the primary task, and therefore may be considered a form of hybrid BCI that restricts the timing of the primary classification to periods of high task potential. This allows for asynchronous timing of the system controlled by the user’s state of vigilance or desire to use the device. The study on BCI gating was performed offline using motor imagery data, and indicated that most individuals would benefit from gating trials with low baseline mu amplitude, while gating by the robustness of the VEP was also beneficial for some. These procedures require validation in an online implementation. A future study will employ similar analysis for defining EEG features which predict single trial motor imagery performance. Something as simple as a thresholding procedure for mu amplitude may be used to determine whether motor-imagery trials are allowed or gated. Alternatively, the hybrid system could em-
ploy multivariate classification on available gating signals to make this decision. Online gating of low-predictive trials will allow us to assess whether this type of procedure improves performance and more importantly makes device use more enjoyable. What remains to be seen is how the feedback of the gating procedure affects the stationarity of these signals, whether the user is able to adapt to repeated gating by modulating EEG correlates of vigilance, and whether user-defined asynchrony is effective using this method. This is a valuable avenue of study, as the potential for improvement is greatest for low performing users, and could possibly offer a solution to BCI inefficiency seen in a significant proportion of users.

6.1.2 System Targeting

Chapter 4 focused on how the unique capabilities and limitations presented by ALS patients affect the utility of standard BCI systems. These results of this study are of high clinical relevance, as disease heterogeneity can influence the course of technological intervention decided on by the patient, as well as the level of achievement possible with different devices. It was found that behavioral impairment, which can affect up to half of ALS patients, was associated with a relative lack of interest for pursuing BCI as an assistive communication technology. This finding is consistent with the behavioral signs of apathy and mental rigidity associated with the disease [114, 156]. At the same time, those with cognitive impairment were more interested in pursuing BCI than those with normal cognition, possibly as a result of loss of insight for cognitive deficits.

As my research with BCI-AAC devices in ALS continues, additional studies will help confirm these findings and probe further the reasons for device rejection. Assessment of patients’ neuropsychological states in greater detail is required to achieve a better understanding of their cognitive and behavioral limitations. The ALS-CBS was used primarily due to time restrictions, but more comprehensive tests for cognition and behavior in ALS may serve to acquire a better understanding of changes in these domains specific to the disease [293]. Of course, permitting more time for BCI training would be a goal of future work, as it would allow more users to reach their optimal performance level, so that they could better make decisions about long-term BCI use. This is especially true of the motor-imagery system, for which lack of overall control likely contributed to poor post-training evaluations of this system.
Of foremost importance in future study is to identify the reasoning behind disinterest in BCI-AACs. Depression may affect motivation for exploring new technologies, and its impact on BCI interest will need to be quantified. Self-reported motivation has been shown to affect the quality of P300 signals in a spelling task [158], while increasing motivation using monetary incentives or immersion in virtual feedback environments significantly improve performance on BCI tasks, most notably in poor performers [294, 108]. Nijboer et al. [158] present a more comprehensive assessment of motivation that encompasses four factors: mastery confidence, incompetence fear, interest, and challenge. It should be interesting to determine which of these factors play a role in device rejection among behaviorally impaired patients, and which associate with higher device acceptance among those who are cognitively impaired.

With regard to system targeting based on factors of the disease, the only variable which correlated with differential performance on the BCI tasks was age. The correlation was not particularly strong, but does indicate that older patients perform relatively better using a P300 system. This finding needs to be reproduced in an additional sample, taking into account motivational and cognitive factors. It has been shown that both P300 amplitude and motor imagery ability decline with age [217, 295], so individual assessment of P300 generation and imagery ability will be needed to make decisions about optimal BCI paradigms.

6.1.3 System Personalization

Of the disease factors measured in this study, those related to cognitive function, rather than physical function, were determined to predict successful BCI operation. Specifically, the cognitive sub-score of attention was associated with BCI proficiency in both P300 and motor-imagery tasks. This is consistent with the executive function deficits seen in FTD, in which prefrontal damage can often lead to the presence of reduced focus and attention. Although the studies in this dissertation did not directly assess how performance and cognition are related in locked-in syndrome, the loss of BCI capability in CLIS may in part be considered the extrapolation of worsening performance with increasing cognitive limitations found here.

The morphology of the P300 and resulting task performance is reduced with age and cognitive dysfunction. Modifications to P300 systems have attempted to overcome
these limitations by reducing cognitive load. Some teams have used centralized spellers and language prediction models, by which distractions and memory load are reduced by removing the need to search for letters in a grid. Symbolic spellers allow for rapid communication of the common needs of daily living, and have been shown to be invariant to the cognitive status of the patient. Finally, entirely new paradigms eliciting alternative evoked potentials such as the N400 in response to familiar faces have been shown to be robust against cognitive decline compared to the P300. Modifications such as these may make evoked potential-based tasks more viable for patients with subtle cognitive impairment, and will need to be tested for improved communication in ALS patients with reduced cognitive capacity.

Modifications to motor-imagery tasks that address cognitive decline are less studied, and will require more intensive participant training to evaluate. Certainly, the low number of training sessions factored into the poor performance of the ALS participants on the motor-imagery task. Future work will involve at least ten sessions of training. The use of a sufficient training interval may reveal the differential timing of training effects (or lack of such effects) that occur in individuals with cognitive limitations. As was the case with the P300 speller, the cognitive sub-score of attention was a particularly important indicator of motor-imagery quality. In addition to the results of the cognitive screen, the presence of elevated EEG power in frequencies unrelated to the task was also common in low-performing individuals. Further work will study the linkage between these two findings, along with methods to reduce this task-irrelevant signal. This may involve utilizing alternative forms of mental imagery that originate from brain networks which are less affected by the neurodegenerative processes occurring in ALS, such as auditory imagery and spatial navigation.

Additional study is required to determine whether sub-threshold genetic markers for C9ORF72-linked ALS/FTD correlate with cognitive impairment, a trend which was not indicated in the study performed in this dissertation. This linkage would provide evidence for graded cognitive decline with repeat length, a finding whose impact would extend far beyond the field of BCI. Without this, the relationship between repeat length and performance on both of the tasks needs to be assessed for alternative causal influences.

Finally, the limitations of patients with CLIS were not directly assessed in this dissertation, as all ALS patients retained at least residual eye movement. Up to this point,
there has been very little success of achieving reliable communication in these patients, although a number of reasons for this are becoming evident. First, the cuing and feedback system may be inappropriate for use in these patients who often develop visual impairments. Second, scalp-based electromagnetic measures of brain activity may not be appropriate, given that the only minor success for these patients was achieved with near infrared spectroscopy, which measures blood oxygenation over long time scales. Last, the work presented here points to gradual loss in cognition, specifically attention, as a possible mechanism for the extinction of BCI use in CLIS. For any of these modes of failure, we have a better idea of how to personalize systems to CLIS users by using modified inputs, recording methods, and training techniques to facilitate device success.

### 6.1.4 Alternative features

Substantial improvements to BCI performance can be made by optimization of neural features extracted from ALS patients, the beginnings of which are detailed in Chapters 3 & 5. System gating, as described above, employs unused features recorded during the primary BCI task, in order to make decisions about task readiness. As was demonstrated offline, this type of gating could significantly improve the performance and speed of the interface.

Personalized feature optimization allows the classifier to account for diversity among individuals by utilizing unique brain signatures, which are sometimes outside of the ‘normal’ ranges expected for these tasks. With this optimal feature set, the user can perform at a higher level than if feature extraction were a fixed process. Additionally, the dimension reduction accomplished by feature selection makes simple linear classifiers more robust to over-fitting, decreases online computation time, and requires the use of fewer sensors. Future studies could employ feature reduction algorithms based on exhaustive search, and measure the stationarity of these optimal signals both inter- and intra-session.

From the results of our P300 study, we expect high performing users to achieve the greatest reduction in sensor number, while ALS patients would benefit from classification periods around 500 ms after the presentation of the stimulus. Assessment of the changes to the optimal feature set in motor-imagery will rely on a more extensive training period. Results from the high-performing users indicate imagery in controls and
ALS patients could both be classified using a low number of channels, specifically C3 and C4 over the primary motor regions, leading to a substantial reduction in electrodes and feature space for classification.

The use of coherence as a measure of interaction between electrode sites during motor imagery proved to be relevant for classification in a subset of individuals. Of the subjects which did show consistent coherence changes resulting from motor imagery, the high-performing users demonstrated clear hemispheric lateralization, displaying a contralateral increase in coherence within the SMR frequency bands. There were also individuals who demonstrated whole-brain changes in coherence values which were hemisphere invariant. Measures of coherence, although overlooked in healthy individuals, were found to be somewhat more effective with ALS patients demonstrating cognitive impairment, possibly due to compensation of functional networks through large-scale brain recruitment. Future work will address whether features of coherence used in conjunction with SMR power features can be utilized in online classification more effectively than either alone, and whether this improvement applies to cognitively impaired users as well.

We used a novel method for data assimilation using a neural field as the underlying generative model and a modified Kalman filter to track the states of this field. This model was a modified version of that developed by Freestone et al. [290]. Such a model was used to generate estimates of functional connectivity of the underlying neural field that exist during the intentional states of a BCI paradigm. With our modified framework, we achieved limited success in differentiating between the active states of imagery and rest periods using filtered EEG data, although the high-density sensor system produced less variable results, and appears to be the most promising avenue for future work.

The next step for such an estimation scheme is to develop a scalp-based grid with closely packed sensors localized to the region of interest. This will allow for assessment of kernel validity using high-density recordings in a larger group of subjects. Before exploring this, three modifications to the estimation procedure are apparent. These include the development of an observation function based on lead field models, the addition of signaling delays, and connectivity kernels which allow for multi-scale level interactions between field nodes. If kernel estimates arising from these recordings were to prove viable, the next step would be to implement the single trail tracking of kernel parameters for use in a BCI classification scheme. Additionally, this model could be used to assess
the functional changes that occur as a result of the neurodegeneration that accompanies ALS.

6.2 Ethical Considerations

A small section has been devoted to the ethical discussion that has accompanied the development of BCI. For simplicity, this overview is limited to matters surrounding research and application of brain-computer interfaces for communicative purposes in locked-in patient populations. Other discussions arising from the topics of BCIs in the classroom, the courtroom, and the military are given elsewhere [296, 297, 298]. The ethical basis for brain-computer interfacing emerges from the relatively new field of neuroethics, itself a recent extension of bioethics. It is from two major accomplishments in bioethics that we build the foundation of ethical discourse on the topic: The Declaration of Helsinki and the Belmont Report.

The Declaration of Helsinki is a series of 37 ethical principles established by the World Medical Association in 1964 for physicians and those performing medical research involving human subjects and data [299]. Last modified in 2013, it establishes the life, health, dignity, self-determination, and confidentiality of research subjects as paramount importance in medical research. It dictates that the benefits associated with research should outweigh the risks, and for researchers to continuously monitor this relationship and modify study procedures accordingly. The declaration outlines the process of study oversight by research ethics committees, which are set up to approve and monitor research on human participants. It documents the procedures of acquiring informed consent, for placebo use, vulnerable populations, testing unproven clinical interventions, and the dissemination of study results to the public.

The Belmont report was published by the U.S. government in 1979 as a set of guidelines for biomedical research on human subjects [300]. Within the report, the Belmont commission put forward three general principals which serve as a framework for investigators and review boards, which are respect for persons, beneficence, and justice. Respect for persons stipulates that individuals are treated with appropriate autonomy over their course of treatment, and protections given to those with diminished autonomy. Beneficence is an obligation to minimize harm while at the same time maximizing the benefits delivered through the intervention. Both the judgment of patient autonomy and
risks and benefits should be assessed regularly. The mandate of justice ensures that consideration is given towards the equal distribution of burden and reward that accompanies scientific inquiry. This is to insure that those who are financially or socially impoverished do not bear the majority of the burden (undergoing scientific procedures) without bearing any of the reward (access to resulting treatments). In practice, the framework of the Belmont report serves as a regulatory model for modern institutional review boards. The applications of such a framework pertain to the process of informed consent and voluntariness, subject selection, and the evaluation of risks and benefits.

6.2.1 Making judgments about BCI use

As an experimental technology, the main barrier to BCI deployment is the achievement of sufficient communication throughput to merit use in therapeutic and rehabilitative intervention. The level at which experimental technology should be implemented comes down to a judgment relating the risks associated with device use to the benefit achieved by the system. This assessment must account for the many avenues of failure that could occur, whether from implanted electrode rejection, improper classifier calibration, or effector failure in the case of wheelchair or prosthesis. For devices that carry significant risks, only those patients who stand to benefit the most from the system, like a locked-in patient with little residual communication, would be considered good candidates. This application of invasive technology to an already compromised patient population requires consideration of the responsibilities we have to these patients, and our obligations as researchers to improve the quality of life through low risk, high benefit communication systems while also respecting their autonomy. Less rigorous selection criteria may be used if the device carries a lower risk, and there appears to be a clear divide among researchers about current risk/benefit ratios when comparing non-invasive to invasive BCI systems [301].

The difference between research and treatment should be made clear to prospective users of BCI. These systems currently occupy a space between research and treatment, often in an experimental stage, but also capable of providing benefit. If the intent of a BCI intervention is mainly as research, patients may experience “therapeutic misconception”, which can lead to decisions that might not be in their best interest [302]. As an example, the prospect of BCI communication can influence decisions about life support,
specifically the choice to receive invasive mechanical ventilation. This decision should
not be made without full understanding of the expected outcomes of the BCI protocol.
As BCI systems advance towards long-term clinical implementation, some suggest that
they be offered only after the individual has chosen to be ventilated [303]. On the other
hand, options for BCI communication should be discussed with the patient early in their
treatment course. Even though these topics may be upsetting to patients and family, it
has been shown that such discussions allow for an enhanced sense of control over the
disease [304].

Lastly, the choice a patient makes about communication technologies is not indepen-
dent of the financial cost to themselves and their family. At this point, BCI technologies
remain quite expensive and pose an additional burden to patients and their families. This
creates a problem for equal distribution of the technology, which may in part be allevi-
ated by reductions in system prices as commercial systems enter the market, as well as
future coverage by insurance policies.

6.2.2 Informed Consent

The procedure of informed consent should instill in the participant an accurate depiction
of what to expect from the research or therapeutic intervention, with special care given
to the operation of the system and expected performance [305]. In progressive diseases
such as ALS, the process of informed consent should begin early, so that the user and
family have enough time reach an informed decision [305]. BCI experts agree that the
research team should have a common interaction with the patient about the potential
risks and benefits [301]. The practice of integrated team care is the standard within ALS
clinics, and BCI experts should be working closely with the clinical staff to deliver the
same message relating to informed consent and decision making [305]. The goal of
the device is to maximize the quality of life for the patient and their family; conveying
disjointed expectations about potential benefits could lead to alterations in end-of-life
decisions, which could result in deleterious effects on the user and their loved ones
[301, 302].

The process of obtaining informed consent when the user is in a non-communicative
state relies on the consent of legal representative and possibly the assent of the user. A
legal representative may be a family member or a neutral party, depending on the laws
of the country. In either case, major issues with representative consent come from the evidence that patients may rate their quality of life as satisfactory and worth preserving, while others may ascribe them to have a poor quality of life [178]. There is general agreement among experts that an attempt should be made to establish BCI communication with a completely locked-in patient if consent of their legal representative is given. However, this does not apply to systems that involve invasive recordings [301].

6.2.3 Issues arising from continuous BCI use

Once a user consents to BCI intervention, special concerns arise from regular and continuous use of the system. One such concern is privacy of bioelectric data generated by the user. With improving sensors and algorithms, the researcher is able to tell more and more about a person’s state of mind or intention [302]. Regulation on the use of that data, which is generally controlled in a research setting, will be less well controlled with commercially-developed products [306]. Also, the issue of handling incidental findings, although not unique to BCI development, could be problematic for all users, but especially for recreational users of the devices. If we are able to record from the brain such detailed information to be able to make assessments about voluntary intention, then it is not implausible to regularly encounter gross abnormalities in related bioelectric phenomena. The question of reporting these incidental findings then becomes highly relevant, as a plan of action needs to be in place for situations affecting otherwise healthy users.

A major issue in BCI prosthetic and mobility applications is determining liability in the case of failure or misuse. For BCI applications with high risk, such as wheelchair navigation, at what point is the user responsible for the actions of an imperfect machine? Jens Clausen argues that “...humans are often in control of dangerous and unpredictable tools such as cars and guns. Brain-machine interfaces represent a highly sophisticated case of tool use, but they are still just that. In the eyes of the law, responsibility should not be much harder to disentangle” [307]. Similarly, others note that there are legal structures in place which deal with ascription of liability in our use of machines [308], and many BCI experts anticipate that BCI users will maintain responsibility for the actions of the BCI device [301]. Grüber elaborates on this viewpoint, speculating that proper regulation of high-risk BCI systems will require minimum requirements for de-
vice reliability, documentation of user training, regular licensure, and acquisition of liability insurance [309].

Continuation of treatment directives becomes a concern when a patient who has been using the BCI for communication is no longer deemed competent to make their own decisions. Right-to-die decisions made by patients with disorders of consciousness have come under scrutiny due to controversial cases related to brain injuries resulting in disorders of consciousness [310]. Like traumatic brain injury, ALS can eventually lead to a brain state in which communication is no longer feasible, and the prospect of recovery is without precedent. Locked-in syndrome is not considered a disorder of consciousness, but it is necessary for patients to communicate end-of-life directives at regular intervals, in the event further communication becomes infeasible. Treatment directives stated by patients before becoming non-communicative should not be overwritten easily, and should incorporate the legal representative. Ideally, the patient should make decisions about treatment directives in the case of reduced autonomy due to dementia [304].

Other ethical questions for the future-thinking BCI researcher may involve societal impacts [306], unintended changes to the brain and personality [303, 307, 311], especially in BCI-enabled neurorehabilitation regimes [302], the actions of sub-personal mental agency [308], and deleterious side effects of a hyper-efficient mental connection in a physical world [306]. These topics are given a passing glance in this discussion because I, like other researchers in the field [301], feel an inadequate body of research exists to make judgments about these potential consequences of BCI use.
APPENDIX A

SUPPLEMENTARY MATERIALS
Brain-computer interface technology patient survey

Thank you for taking part in this survey.
The survey will take approximately 10 minutes to complete.

Section 1: Patient Information

1. Education (years):

2. Technology Use

   Computer/Tablet Use:
   - Last month
   - Month before diagnosis

   Phone Use:
   - Last month
   - Month before diagnosis

Section 2: Current assistive technology usage

3a. Have you ever used an assistive device for verbal communication?  Yes  No

   Answer questions 3b&c if you answered ‘Yes’ to the previous question.

3b. What forms of assistive verbal communication do you currently use or have used in the past?
   - Text-to-speech synthesizer
   - Eye/forehead tracking
   - Other

3c. How has your quality of life improved with the addition of your assistive verbal communication device?
   - Not At All
   - Minor Improvement
   - Major Improvement

4a. Have you ever used an assistive device for written communication?  Yes  No

   Answer questions 4b&c if you answered ‘Yes’ to the previous question.

4b. What forms of assistive written communication do you currently use or have you used in the past?
   - Symbolic communication board
   - Speech-to-text transcriber
   - Eye/forehead tracking
   - Other

4c. How has your quality of life improved with the addition of your assistive written communication device?
   - Not At All
   - Minor Improvement
   - Major Improvement

5a. Do you currently have a feeding tube for feeding assistance?  Yes  No

   Answer question 5b if you answered ‘Yes’ to the previous question.

5b. How has your quality of life improved with the addition of your feeding tube?
   - Not At All
   - Minor Improvement
   - Major Improvement
6a. Have you ever used an assistive device to aid in your mobility during your time with ALS?
   Yes  No

Answer questions 6b & c if you answered 'Yes' to the previous question.

6b. Which assistive devices have you used now or in the past course of your ALS treatment to aid in your mobility?
   - Cane
   - Walker
   - Manual/Power Wheelchair
   - Brace
   - Other

6c. How has your quality of life improved with the addition of your assistive mobility device?
   - Not At All
   - Minor Improvement
   - Major Improvement

7a. Do you currently use a BiPAP/breathing device?  Yes  No

Answer questions 7b, c, d & e if you answered 'Yes' to the previous question.

7b. How often do you use the device per day while not sleeping?
   - Never
   - Less than 2 hours
   - 2-6 hours
   - 6-10 hours
   - More than 10 hours

7c. Does it ease your daily activities?  Yes  No

7d. Does it help you sleep?  Yes  No

7e. How has your quality of life improved with the addition of your BiPAP/breathing device?
   - Not At All
   - Minor Improvement
   - Major Improvement

8. What are your reasons for not adopting assistive technologies to aid with Communication, Feeding, Mobility, and Breathing?
   - Financial
   - Time to Learn
   - Not Effective Enough
   - Aesthetics
   - Will make me weaker
   - Other

9. Please explain your answers to the previous question in a little more detail.

---

Section 3: Opinions toward BCI technology

10. Knowing the prognosis for ALS patients includes decreased verbal and written communication ability, how interested would you be in pursuing these BCI assistive communication technologies at some point in the future? Rate your interest in each device on the scale below, with (1) indicating no interest, and (5) indicating great interest.

   Not Interested ——— Very Interested
   (1)  (2)  (3)  (4)  (5)

   P300 Speller
   Dasher
11. If at some point in the future you were unable to perform these tasks, in which other BCI functions for assistive technology would you be interested? Rate your preference for each function on the scale below, with (1) indicating no interest, and (5) indicating great interest.

<table>
<thead>
<tr>
<th>Function</th>
<th>Not Interested</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bed Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lift</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light Switch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheelchair Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robotic Arm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker Phone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12. If you were to use a BCI system as an assistive technology, specify the importance of each of the following system features. Rate the importance of each item on the scale below, with (1) meaning the feature is not important to you, and (5) indicating great importance.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Not Important to me</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appearance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCI Functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setup Simplicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Setup Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standby Reliability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Electrodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

13. What would be the minimum accuracy requirement for you to use a BCI device?

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Less than 60%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>95%</th>
<th>100%</th>
</tr>
</thead>
</table>

14. What would be your required speed of a BCI device used for communication (letters per minute)?

<table>
<thead>
<tr>
<th>Speed</th>
<th>Less than 5</th>
<th>5-9</th>
<th>10-14</th>
<th>15-19</th>
<th>20-24</th>
<th>25+</th>
</tr>
</thead>
</table>

15. What is the maximum setup time you would tolerate (in minutes)?

<table>
<thead>
<tr>
<th>Setup Time</th>
<th>Less than 10</th>
<th>10-20</th>
<th>21-30</th>
<th>31-45</th>
<th>46-60</th>
<th>More than 60</th>
</tr>
</thead>
</table>

16. What is the maximum number of training sessions you would desire to reach optimal performance?

<table>
<thead>
<tr>
<th>Training Sessions</th>
<th>1 only</th>
<th>2-5</th>
<th>6-10</th>
<th>11-15</th>
<th>16-20</th>
<th>20+</th>
</tr>
</thead>
</table>

17. What is the minimum time you would tolerate for the system to incorrectly leave standby mode?

<table>
<thead>
<tr>
<th>Time</th>
<th>Less than 15 min</th>
<th>30 min</th>
<th>1 hour</th>
<th>2-5 hours</th>
<th>5+ hours</th>
</tr>
</thead>
</table>

Thank you for participating in this survey!
Non-technical abstract

A brain-computer interface (BCI) is defined by four criteria. It must (1) record activity from the brain that is (2) intentionally modulated by the user. (3) The processing and classification of neural activity must occur in real time, after which (4) the user receives feedback of the result. The functions of a brain-computer interface can range from communication interfaces and rehabilitation tools, to biofeedback, gaming, training, and vigilance monitoring. In this thesis, I focus on the personalization of BCI systems applied as assistive communication tools for patients with amyotrophic lateral sclerosis (ALS). Also known as Lou Gehrig’s disease, ALS is a neurodegenerative disorder with approximately 6000 new cases in the United States each year. The disease produces degeneration of motor neurons, leading to eventual cessation of voluntary muscular activity, and the need for alternative forms of communication.

BCI communication in advanced ALS has achieved limited success, with users often performing at a lower level than young, healthy individuals in whom the devices are regularly tested. Furthermore, only marginal communication has been established in patients whom are completely locked-in, or without any means of communication with the outside world. In this dissertation, personalization of the BCI system is used to overcome some of the disadvantages faced by ALS patients, as well as provide potential solutions to the BCI inefficiency found in late-stage ALS.

Three projects were undertaken, the first of which focused on novel predictors of BCI performance in healthy individuals. We employed previously unused brain signatures to assess vigilance before the BCI task, in order to selectively gate the operation of the system. By utilizing one of these features, the amplitude of ongoing motor-related brain oscillations, the communication speed of the interface was able to be substantially increased for certain individuals. The second study focused on the user-to-system personalization of a BCI system for ALS. We determined that psychological factors, more than physical factors, contributed to the acceptance of BCI communication systems among these individuals, and that cognition was also a major determinant of device success. Finally, we performed offline analysis on the BCI data to explore possible avenues of system-to-user personalization. Specifically, feature optimization and novel tools for defining brain connectivity were assessed for applicability within this population. Significant improvements in system performance were made possible byoptimiz-
ing the feature set, indicating that functional brain changes occurring in ALS warrant an individualized approach to feature extraction.

For patients with ALS, the need for assistive BCI devices occurs when other methods for communication fail. The results of this study inform both the clinical guidelines for device prescription and the optimal methods of feature extraction. These were often found to be modulated by disease heterogeneity, specifically psychological changes that often co-occur in the disorder. Future work involving online confirmation of these findings will further validate these tools as essential components of BCI systems.


Andrew Geronimo

CONTACT INFORMATION
W301 Millennium Science Complex
The Pennsylvania State University
University Park, PA 16802

Phone: (201) 961-2289
E-mail: geronimo.andrew@gmail.com

EDUCATION
The Pennsylvania State University, University Park, PA
PhD, Dept. Engineering Science and Mechanics, GPA: 3.94 2009 - Present
  • Thesis: “Personalized brain-computer interfaces for amyotrophic lateral sclerosis”.
The College of New Jersey, Ewing, NJ
BS, Engineering Science, GPA: 3.83 2005 - 2009

RESEARCH EXPERIENCE
The Pennsylvania State University, University Park, PA
Research Assistant, Center for Neural Engineering 2009 - Present

REFEREED JOURNAL PUBLICATIONS


REFEREE SERVICE
• IEEE Transactions on Biomedical Engineering

TEACHING EXPERIENCE
The Pennsylvania State University, University Park, PA
Teaching Assistant August 2012 to December 2012
  • Teaching assistant for EMCH 211: Engineering Statics

PROFESSIONAL MEMBERSHIPS
Institute for Electrical and Electronics Engineers (IEEE), Member, 2011 – Present
IEEE Engineering in Medicine and Biology Society (EMBS), Member, 2011 – Present
Society for Neuroscience, Member 2012 – Present

AWARDS
The Pennsylvania State University, University Park, PA
  • Dr. Richard E. Llorens Graduate Award, 2013-2014.
  • Sabih and Guler Hayek Graduate Scholarship, 2013.

Other Awards
  • IEEE EMBS Excellence in Neural Engineering Travel Award. $700 for travel to the 5th International Conference on Neural Engineering, April 2011.
  • BCI 2000 Student Scholarship. $1000 for travel to the Fourth International Brain-Computer Interface Meeting, May 2010.

TECHNICAL SKILLS
Electroencephalographic recording techniques
Brain-computer interface design
Programming Languages: MATLAB/Simulink, LabVIEW, Introductory C++, LaTeX