Probabilistic Methods for Improving the Performance of the P300 Speller Brain Computer Interface

by

Dmitry Kalika

Department of Electrical and Computer Engineering
Duke University

Date:_____

Approved:

_____
Leslie M. Collins, Supervisor

_____
Galen Reeves

_____
Stacy L. Tantum

_____
Loren W. Nolte

_____
John A. Board Jr.

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Electrical and Computer Engineering in the Graduate School of
of Duke University

2018

ABSTRACT

Probabilistic Methods for Improving the Performance of the P300 Speller Brain Computer Interface

by

Dmitry Kalika

Department of Electrical and Computer Engineering
Duke University

Date:_____

Approved:

_____
Leslie M. Collins, Supervisor

_____
Galen Reeves

_____
Stacy L. Tantum

_____
Loren W. Nolte

_____
John A. Board Jr.

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical and Computer Engineering in the Graduate School of Duke University

2018

Abstract

Many augmentative and alternative communication (AAC) devices have been developed to aid individuals who have some form of severe neuromuscular disorder to communicate with the outside world, such as Amyotrophic Lateral Sclerosis (ALS). Eye-trackers are used as a primary communication device for people with ALS until they lose the ability to use their eyes. As an alternative to eye-trackers, the P300 speller brain-computer interface (BCI) is a non-invasive mode of communication that utilizes electroencephalography(EEG) data.

The P300 speller relies on eliciting and detecting event related potentials (ERPs) that occur in the EEG data when a rare or unexpected visual or auditory stimulus is presented to the user; however, only visual stimuli are used in this work. The P300 speller displays characters or symbols in a grid on a computer screen and presents (i.e. flashes) a subset of the characters simultaneously; these presentations are known as the stimuli. The presentation of the target (i.e., desired) character should elicit an ERP. After many repetitions of presentations, the speller attempts to estimate the desired character using the EEG data. This thesis is composed of two primary methods to improve the P300 speller: A novel data-driven adaptive stimulus selection paradigm based on maximizing the expected discrimination gain (EDG) metric; and fusion of an eye-gaze data stream to develop a hybrid P300 and eye-gaze speller.

Many pseudo-random stimulus presentation paradigms (i.e., patterns) have been developed to improve the accuracy and decrease the time required to communicate via the P300 speller. Few data-driven, adaptive, stimulus presentation paradigms have been developed, however, they are computationally expensive, thus have limited flexibility in the groups of characters that can be presented simultaneously. In this thesis, a novel data-driven, adaptive, stimulus selection approach based on maximizing the expected discrimination gain (EDG) is introduced. Various restrictions are set on the characters that can be presented based on system and physiological constraints. Simulations show that even with various restrictions on the proposed adaptive paradigm, the adaptive paradigm yields a higher accuracy and a decrease in time required to spell compared to the most commonly used row/column random paradigm. Online results show that the proposed paradigm decreases the time required to spell, however, there is a slight decrease in speller accuracy.

In addition to setting restrictions based on physiological effects, this thesis presents work done on explicitly modeling refractory effects, probabilistically, on a subject-specific basis. Refractory effects occur when the time between target stimulus presentations is not sufficiently long, resulting in decreased SNRs of the ERP. By modeling the refectory effects, the adaptive stimulus selection paradigm can automatically choose characters to present that minimize refractory effects, without having to explicitly set ad-

hoc restrictions. Offline simulations showed that modeling refractory effects explicitly has the potential to further increase the accuracy, and decrease the time required to spell.

Beyond improving the independent P300 speller, there has been recent interest in developing a hybrid (or "fused") BCI system. In this thesis, a probabilistic hybrid P300 and eye-tracker is developed and its effectiveness is explored. The hybrid speller collects both eye-tracking and EEG data in parallel, and the user spells the characters in the same way that they would spell them using the traditional P300 speller. Both online and offline experiments are performed to analyze the hybrid speller. Online results showed that for the fifteen non-disabled participants, the hybrid speller improved accuracy and reduced the time required to spell a character. Offline simulations showed that the system is more robust to eye-gaze abnormalities than a stand-alone eye-gaze system.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Lately I've been joking about how I accidentally stumbled onto a Ph.D. Though this statement isn't completely accurate, it's not far from it. If it weren't for my brilliant advisor, mentors, peers, and my loving family and friends, I would not have made it this far.

Thank you to all of the teachers at Neshaminy High School and professors at Pitt that motivated me to pursue a bachelors and then a Ph.D. in electrical engineering. Specifically, I would like to thank Hank Oppenheimer, David Bender, Mahmoud El Nokali, Steven Jacobs, Luis Chapparo, and Ervin Sejdic.

Thank you to all of the people that participated in my data collections; your participation allowed me to verify whether or not my simulations would actually translate to a usable, online system. Thank you to the organizations that funded my research throughout graduate school: the Duke University ECE department; U.S. Army RDECOM CERDEC Night Vision and Electronic Sensors Directorate; Duke Wisenet NSF; the NIH; and discretionary funding from Leslie Collins' applied machine learning lab (AMLL). Without the financial support none of my work would have been possible.

Leslie Collins, thank you for inviting me to join the best lab at Duke. Since I've been here, the AMLL has consistently been filled with students, post-docs, and research scientists that are extremely bright, hard-working, and sociable. You have created a lab atmosphere that encourages everyone to work hard, yet still have fun – both in and

outside of the lab. You have always promoted a strong work-life balance, which has ensured that I (and others in the lab) stayed at least somewhat sane. Thank you for legitimately caring about my well-being – the bi-weekly emails I received from you as my defense was approaching, urging me that everything would be alright, significantly reduced my stress levels and encouraged me to keep pushing towards graduation. I can't thank you enough for everything you've done for me.

Thank you to all the mentors I've had over the years: Peter Torrione, for significantly improving my MATLAB programming skills. Kenny Morton, for instilling in me the Bayesian mindset and for sharing with me the misery that is Pitt sports. Mary Knox, for your guidance over the short time we worked together. Sandy Throckmorton, for teaching me that every good presentation has a flow-chart. Boyla Mainsah, for brainstorming with me almost daily, which allowed me to make research progress at a significantly faster pace than I would have otherwise. I'm also deeply appreciative for the time you spent on editing my latest conference papers, this thesis, and my doctoral defense presentation.

Thank you to all of the current and former graduate students in the AMLL for keeping me sane, as well as motivating me to be the best researcher I can be. Nick Czarnek, I'm really glad that we were lab mates at both Pitt and Duke; having somebody to study with and learn new concepts with made the transition between undergrad and grad school palatable. Joe Camilo, the last few months you've been telling me, "hurry up and

graduate already." I'm not sure I would have ever graduated without this incredible advice. Kedar Prabhudesai, your constant happiness is contagious and managed to brighten my mood after every one of our interactions. Jillian Clements, I'm glad that we were able to become close friends throughout my time here; I enjoyed every one of our conversations about boring adult finance stuff like ETFs and 401k's. Jordan Malof, Daniel Reichman, and Patrick Wang, I greatly appreciate all of the brainstorming sessions I've had with each of you over the years. Lidea Shahidi, Bohao Huang, and Kevin Chu, thank you all so much for your insightful feedback during various lab meetings and exam practices. Rayn Sakaguchi and Ken Colwell, you were both major contributors in making the AMLL the most fun lab at Duke. Ken, those googly eyes that you put on objects throughout the lab are still making me laugh at least once a week. Achut Manandhar, Sara Duran, and Jill Desmond, thank you for welcoming me into the lab with open arms. I truly felt like I was part of the lab family after my first day here.

I was also fortunate enough to also be surrounded by many great people at Duke outside of the AMLL. Thank you to all of of the people in the Duke ECE department who encouraged me to be less of a recluse: Guy Lipworth, Alex Mrozack, David Carlson, Laura Carlson, John Marcus, Kyle Ulrich, Devin Ulrich, Jon Soli, Lorelei Soli, Michael Boyarsky, Greg Spell, Christian Nadell, Stacey Kuo, Andrew Traverso, Andrew Boyce, Ugonna Ohiri, Korine Ohiri, Michael Martinez, Michaela Kane, Tim Sleasman, and Trisha Dupnock. Thank you to the people who helped me put together the best intramural

ultimate Frisbee team Duke had ever seen: Chris Jackman, Jake Coleman, Kevin Liang, Kristen Collar, Dean Culver, Danton Noriega-Goodwin, Michael Reinsvold, and Bradley Feiger. Lastly, thank you to all the members and coaches of the Duke club ultimate Frisbee team from 2012-2014 for giving me the opportunity to enjoy playing the sport I love most.

I've been blessed with a kind, loving, and supportive family. Thank you to my parents, Inna and Gena for pushing me to work hard, yet giving me the freedom to choose my own path in life. Thank you to my younger brother Michael for always keeping me in the loop about what slang words are still cool to use. Thank you to my grandparents Valya, Vasya, Ira, and Shuryik, for showering me with love every time I come visit, albeit infrequently. Now that I'm done with school and moving closer to home, I promise to visit and call more often.

I'm not even sure words can describe how thankful I am for my girlfriend of 4 years, Rebecca Delk. She probably deserves this Ph. D. more than I do. She's had to put up with my constant, grad-school induced, whining and neediness for pretty much the entirety of our relationship. During extra stressful periods of grad school (such as exams and conference deadlines) she would drop everything to drive 1.5 hours to be with me; often she would even surprise me with assortments of gifts and snacks. I've enjoyed just about every moment we've spent together, and I'm excited for many more. Now hurry up and finish your school year so you can finally come live with me in Maryland!

1 Introduction

Various augmentative and alternative communication (AAC) devices have been developed in order to aid individuals who have some form of severe neuromuscular disorder to communicate with the outside world. One disease that results in the need for AAC systems is Amyotrophic Lateral Sclerosis (ALS), also known as Lou Gehrig's disease. ALS is a progressive neurodegenerative disease characterized by loss of voluntary muscle control. Over time, people diagnosed with ALS lose their ability to move and communicate through speech or gestures [1]. In some cases, as the disease progresses, eye movement becomes the only voluntary muscle control that an individual with ALS can perform [2]. There are more than 12,000 definite diagnoses of ALS in the US, and approximately 5000 individuals are diagnosed each year [3]. The target population for the AAC devices discussed in this work are those with ALS [4].

Studies have shown that eye-trackers can be used as one mode of communication for people with ALS who retain much of their oculomotor control (e.g., [2], [5]–[9]), allowing them to spell words by gazing at characters on a screen, and having the direction of the gaze automatically detected. Most modern eye-trackers emit infrared (IR) light and the direction of eye gaze is determined based on the reflection of the IR light off of the pupil and cornea ([10], [11]). Eye gaze is used for detecting the spatial position that is being looked at on the screen, and object selection on the screen is typically achieved by maintaining gaze for a specified dwell time, typically 500-1000ms for novice spellers [12].

An eye-tracker typically requires a very short calibration period (1-5 minutes) (e.g., [13]) and is easy to set up. Although eye-tracking is commonly used as a mode of communication for people with ALS, there are several limitations with gaze-control systems. People with ALS are known to have oculomotor impairment ([9], [14]–[17]), such as saccadic intrusions during fixation (e.g., [18]), horizontal and vertical gaze limitations (e.g.,[19]), slower saccades (e.g.,[20]), and nystagmus (e.g.,[19]), a condition that causes a person to make repetitive, uncontrolled eye movements. These abnormalities will sometimes render the user unable to efficiently use an eye-tracker as their primary method of communication ([8], [21], [22]). Eye-trackers can also become inaccurate due to excessive sunlight (e.g., [2]) or with the use of corrective lenses [2].

As an alternative to eye-trackers, the P300 speller brain-computer interface (BCI) [23] is a non-invasive mode of communication that utilizes electroencephalography(EEG) data, and this system has also been shown to allow people with ALS to communicate [22], [24]–[26]. P300 spellers rely on eliciting and detecting event-related potentials (ERPs) that occur in EEG data when an uncommon stimulus is presented to the user [27]. P300 spellers are named after an ERP that occurs approximately 300 ms after an uncommon event occurs, although other ERP responses in the EEG data can also be used for classification (e.g. [28], [29]). In a typical visual P300 speller, characters or symbols are displayed on a computer screen and groups of characters are sequentially illuminated (i.e., flashed or presented) on the screen. While the flashing is occurring, the subject attends to the

character or symbol in the grid that they intend to spell. The flash of the "desired" or "target" character is the uncommon event that should elicit an ERP. Typically, the EEG data collected after each flash has a low signal to noise ratio (SNR), so many repetitions of flashes are performed until there is enough information that has been gathered to determine which character is being spelled.

The P300 speller can be used as a communication device when the eye-tracker is unusable, as the ERP response remains mostly stable over the progression of ALS [21], [22], [30]–[32]. The primary drawback associated with the P300 speller is that the time it takes to spell a character can be lengthy (~5-20 seconds per character), particularly in comparison to spelling times associated with an eye-tracker (~1 second per character). The P300 speller also requires a significantly longer set up time and calibration or training phase compared to an eye-tracking communication device.

Work has been done towards improving the accuracy and decreasing the time required to communicate using the P300 speller. Some of the improvements involve cosmetic changes to the interface. In [33], [34], the speller presents characters using various colors so that the user is less distracted by nearby character presentations. In [35], [36], faces were presented instead of characters to increase the amplitude of the ERP response. Traditionally, the subsets of characters presented to the user are either rows and columns of characters arranged in a grid layout, and these subsets are randomly selected to be presented to the user. Instead of randomly created flash groups, other methods have

exploited error-correcting codes from coding theory to optimize the composition and presentation order of flash groups, as is considered in [37], [38]. These approaches, however, do not rely on previously observed EEG data that was collected in the sequence of spelling in order to optimize the stimulus selection.

Some work has been done on the development of an adaptive, data-driven, stimulus presentation paradigm for the P300 speller [39], [40]. A data-driven approach selects stimuli based on previously observed stimulus responses; this allows the speller to select stimuli that can provide more information about the target character than a random stimulus selection. In Park et al. [39], a partially observed Markov decision process (POMDP) was used to adaptively decide which row or column to flash. The row and column decisions were made independently and the intersection of the selected row and column was then selected as the target character estimate. The POMPD approach, however, becomes intractable for a real-time system when considering a search space considerably larger than row and column flash groups. Ma et al. [40] implemented a hierarchy of variable-sized flash groups based on a language model. The hierarchical approach, however, is easily susceptible to error propagation, especially when considering the amount of incorrect selections in users with low accuracy levels [40].

The first aim of this work was to develop an adaptive stimulus selection framework that aims to maximize the amount of information gained after every stimulus presentation. The basis of the framework is the expected discrimination gain (EDG). The

EDG is a metric introduced by Kastella [41] as a sensor sampling strategy. The proposed approach explicitly models the expected information gain for the upcoming stimulus presentation, as opposed to the POMDP which tries to maximize a reward function over a measurement horizon. Since the proposed approach is myopic, it is significantly less computationally complex than the PODMP approach, and therefore allows the flexibility to dynamically create flash groups using a greedy search.

One of the primary concerns that should be considered when developing a stimulus selection paradigm for a speller-based BCI are physiological effects that can render the automatic discrimination between target and non-target stimulus presentations difficult. Refractory effects, for example, occur when a target character is presented in rapid succession; this can reduce the amplitude and increase the latency of the ERP response [42], [43]. When refractory effects occur, target stimulus presentations can be incorrectly classified as non-target presentations. Conversely, non-target presentations can be classified as target presentations if an adjacent character presentation distracts the user [25], [44]; this is known as an adjacency distraction. The adaptive stimulus presentation paradigm developed in this work enforces constraints on the stimulus presentation groups to attempt to reduce the effects of these physiological effects.

In addition to setting constraints based on physiological effects, work was done on explicitly modeling refractory effects, probabilistically, on a subject-specific basis. By

modeling the refectory effects, the adaptive stimulus selection paradigm can automatically choose characters to present that minimize refractory effects, without having to explicitly set *ad-hoc* constraints. By modelling the refractory effects on a subject-specific basis, the adaptive stimulus selection paradigm can increase the time between character presentations if subjects are more prone to refractory effects, and vice versa.

Another physiological phenomenon that can provide information about what the user of a P300 speller is trying to spell is eye-movement. There has been interest in combining EEG data and eye-gaze data with the goal of developing a hybrid (or fused) BCI (hBCI) communication system (e.g. [45], [46]). A system that can successfully combine eye gaze and EEG has the potential to increase robustness to oculomotor impairment through the additional information provided by EEG while taking advantage of the increased communication speed provided by eye gaze. Hybridizations have taken several forms in the literature, with the majority using the EEG and eye-gaze inputs for separate control purposes (e.g. [46]–[53]).

Dong et al. [54] demonstrated the potential for improving BCI performance through parallel fusion of eye gaze and EEG by incorporating eye gaze into a motor imagery BCI for cursor control. Gaze was classified in only two directions (left, right). Choi et al. [55] developed a hybrid P300 and eye-gaze speller by limiting the characters that can be selected to a 3 x 3 grid centered at the eye-gaze location. In Dong et al. and Choi et al., the eye-gaze was used to limit the characters that could be selected. In these

studies, hard-thresholds on eye-gaze position are used to prune the set of selectable characters. These thresholds are fixed and do not adapt to the user's oculomotor ability.

The second aim of this work is to develop an algorithm that uses eye-gaze and EEG data in parallel to estimate the probability of a character being the target character after each collected data point. The probabilistic algorithm will allow the system to adapt to each individual use; for example, if it is known that eye-gaze data is typically inaccurate, more EEG and eye-gaze data will be collected before making a character selection. The full two-dimensional spatial position will be used, thereby increasing the potential information that can be used by the BCI system to improve performance.

This document is organized as follows: In Chapter 2, the background information for this work is introduced. In Chapter 3, the proposed adaptive stimulus presentation paradigm based on the EDG metric is introduced. In Chapter 4, the adaptive stimulus presentation paradigm is improved by incorporating a model for refractory effects in the EDG computation. In Chapter 5, an algorithm is introduced for a hybrid P300 and eye-gaze speller. Chapter 6 will briefly describe the primary contributions of this work as well as suggest avenues for future work.

2 Background

One of the most common AAC devices for individuals with ALS is the eye-tracker (e.g., [2], [5]–[9]). However, if the person with ALS is unable to use an eye-tracker due to oculomotor abnormalities, a BCI can be used. The P300 speller BCI has been one of the most widely researched AAC's for individuals with ALS who do not have consistent eye gaze (e.g., [21], [22]). The P300 speller relies on the presence of an event related potential (ERP) that exists in the measured EEG signal after an unexpected auditory or visual stimulus event (oddball). The presence of the ERP can be used to select characters on a virtual keyboard [23].

There are two main contributions in this work: developing a probabilistic hybrid eye-gaze and P300 speller, and developing a data-driven, adaptive, stimulus presentation paradigm for the P300 speller. In section 2.1, communication via eye-tracking is described. The eye-gaze data used in this work will also be introduced and examples of the eye-gaze data that can be collected with our system is shown. In section 2.2, the P300 speller is introduced and methods that have been used to improve the P300 speller are discussed. Next, the physiological phenomena that can affect the accuracy and spelling speed of the P300 speller are discussed in section 2.3. Finally, the metrics used in this work to measure the performance of both independent and hybrid spellers are discussed in section 2.4.

2.1 Eye-tracking for Communication

Most modern eye-trackers emit infrared (IR) light and learn the direction of eye-gaze based on the reflection of the IR light from the pupil and cornea ([10], [11]). An eye-tracker typically requires a very short calibration period (~1 minute) (e.g., [13]), is easy to set up, and is relatively inexpensive compared to an EEG-based system. The eye-tracker can be used to communicate via "eye-typing" (e.g., [56]–[62]). To eye-type, a virtual keyboard is presented to the user, and the user selects characters by focusing on a character that they intend to spell. During spelling, when the detected eye-gaze is closest to a particular character for a predetermined amount of time, known as the dwell time, the system selects that character, and spelling proceeds. Dwell times are typically set to 500-1000 ms for novice spellers [12].

Unfortunately, dwell based eye-gaze spelling systems fail to perform well if the user has significant oculomotor impairment. Individuals with ALS are known to have oculomotor impairments ([9], [14]–[17]) such as saccadic intrusions during fixation (e.g., [18]), horizontal and vertical gaze limitations (e.g.,[19]), slower saccades (e.g.,[20]), and nystagmus (e.g.,[19]). Saccadic intrusions and nystagmus can increase the variance of the measured eye-gaze data significantly, rendering the user of the system unable to dwell on the target character throughout the dwell period. Gaze limitations can result in the user being unable to gaze at the character that they are intending to spell for the required dwell time.

2.1.1 Eye Gaze Data

The Tobii Technology X2-30 eye-tracker [55] was used to collect eye-gaze data in this work. The X2-30 is a binocular, infrared eye-tracker that samples at a frequency of 30 Hz. The eye tracker estimates the user's gaze, or point of focus, on the monitor in both the vertical and horizontal axis. Prior to collecting eye gaze data in a spelling task, the system is calibrated for each subject with the Tobii eyeX software. The calibration procedure asks the user to follow a marker with their eyes as it moves around the screen for approximately 30 seconds, and then uses that data to learn a physiological 3D eye model [63]. The initial calibration performed using Tobii's calibration software normalizes the gazes such that all gazes on the monitor are $\in [0,1]$ on both the horizontal and vertical axes, regardless of the size of the monitor. The eye-gaze estimates are not restricted to the length of the monitor, therefore gaze estimates may exist outside of [0,1] on either the horizontal or vertical axis. The monitor used in this work has a resolution of 1920 x 1080 (width x height) pixels, and is 22.9" along the diagonal.

A single observation of eye-gaze data is denoted as $Y_t \in \Re^{1x2}$, and t observations are denoted as $\boldsymbol{Y}_t \in \Re^{tx2}$, where the two dimensions are the horizontal and vertical eye-gaze. Figure 2.1 provides an example of eye-gaze data that was collected from a study participant with ALS while they were attempting to eye-type on a virtual keyboard.

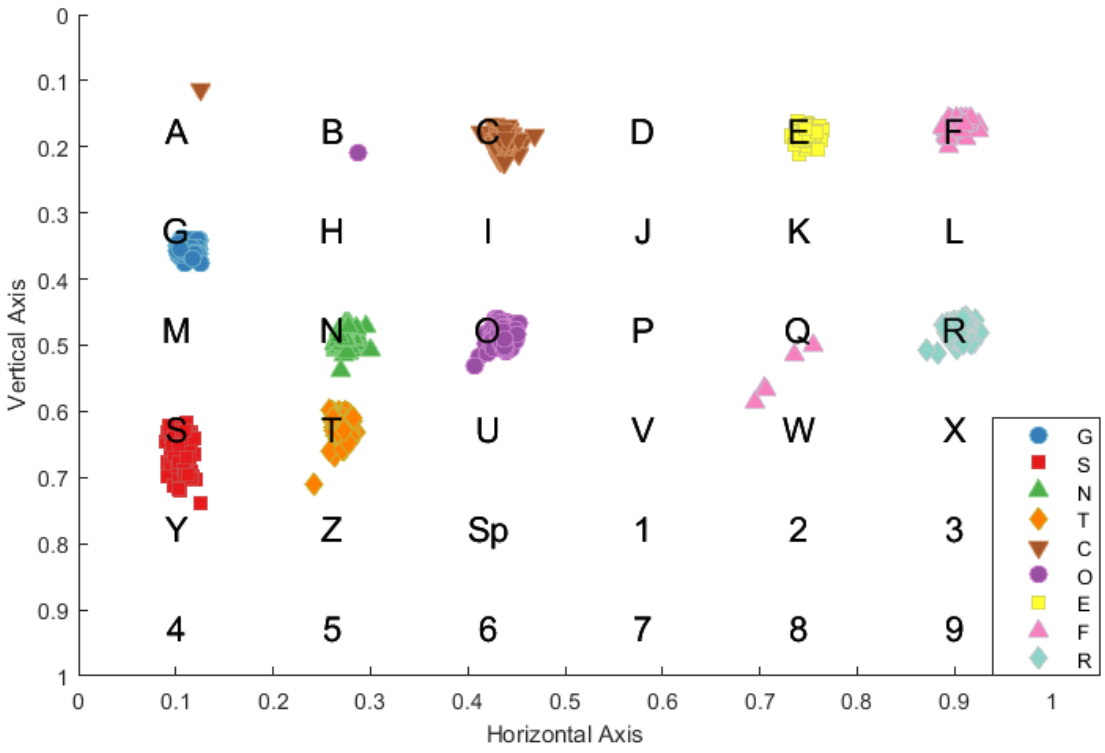

Figure 2.1 – This figure provides an example of eye-gaze data collected on a 6 x 6 character grid from a study participant with ALS while they were attempting to spell various characters. The color/character combination corresponds to the character that the participant was instructed to focus on. There are 168 eye-gaze data points for each character being spelled, and during this data collection there were 12 characters being spelled. Some characters are spelled more than once, and thus there are only 9 unique characters spelled.

2.2 P300 Speller

When a person is unable to use an eye-tracker due to oculomotor impairment, a BCI such as the P300 speller can be used as an alternative communication device [22]–[26]. The P300 speller leverages ERPs that are elicited when an infrequent auditory or visual stimulus is presented to the user [27]. This infrequent stimulus presentation is known as

the oddball or target stimulus response. The ERP response can be seen in the EEG data as a positive peak that occurs approximately 300ms after the auditory or visual stimulus is presented; this positive peak is termed the P300. Figure 2.2 shows averaged EEG responses that were measured when a target and background (non-target) stimulus event occurs.

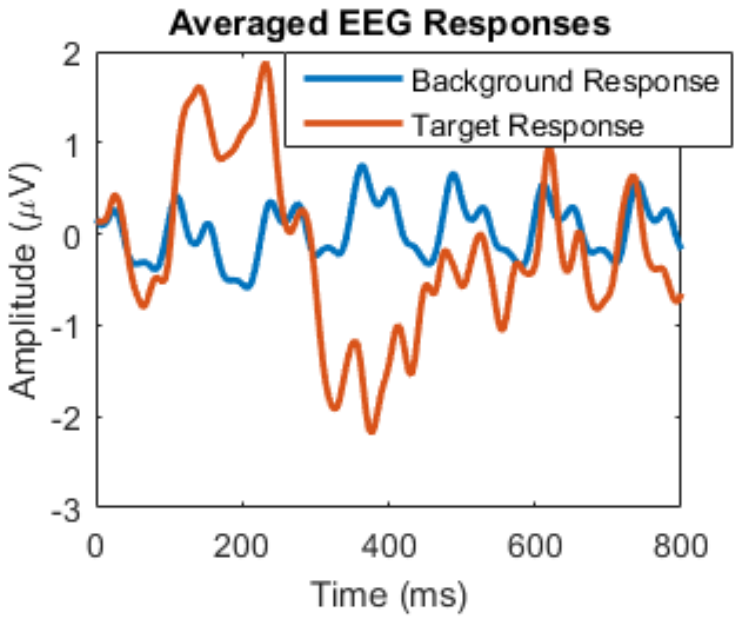

Figure 2.2 – Averaged EEG response measured from a subject when a target (red) background (blue) stimulus event occurs.

In a visual P300 speller, characters or symbols are displayed on a computer screen and groups of characters are sequentially illuminated (i.e., flashed or presented) on the screen. While the presentations are occurring, the user attends to the character or symbol in the grid that they intend to spell. The presentation or flash of the "desired" or "target"

will typically elicit an ERP. The P300 speller signal processing algorithm attempts to estimate the character that the user is intending to spell by determining the character presentations that elicited the ERP responses.

A flow chart describing the components of the visual P300 speller system is provided in Figure 2.3. The user is presented with a grid of characters and/or symbols on a computer screen. The user then chooses a single character on the P300 speller grid and focuses on that character (i.e., the target character). The grid illuminates, or flashes, a subset of the total characters on the grid (section 2.2.1); the subset of the characters being presented is termed a flash group. The flash group presentations on the screen are synchronized with the EEG signal acquisition. EEG data is collected (section 2.2.2), beginning at the onset of the illumination of a flash group, for some preset amount of time. The collected window of EEG data is then pre-processed and features are extracted (section 2.2.3). The features are automatically analyzed with a trained, user-specific, algorithm, or classifier, to generate a classifier score. This classifier score is then used to update a character scoring function in order to estimate the character that the user is intending to spell (section 2.2.4).

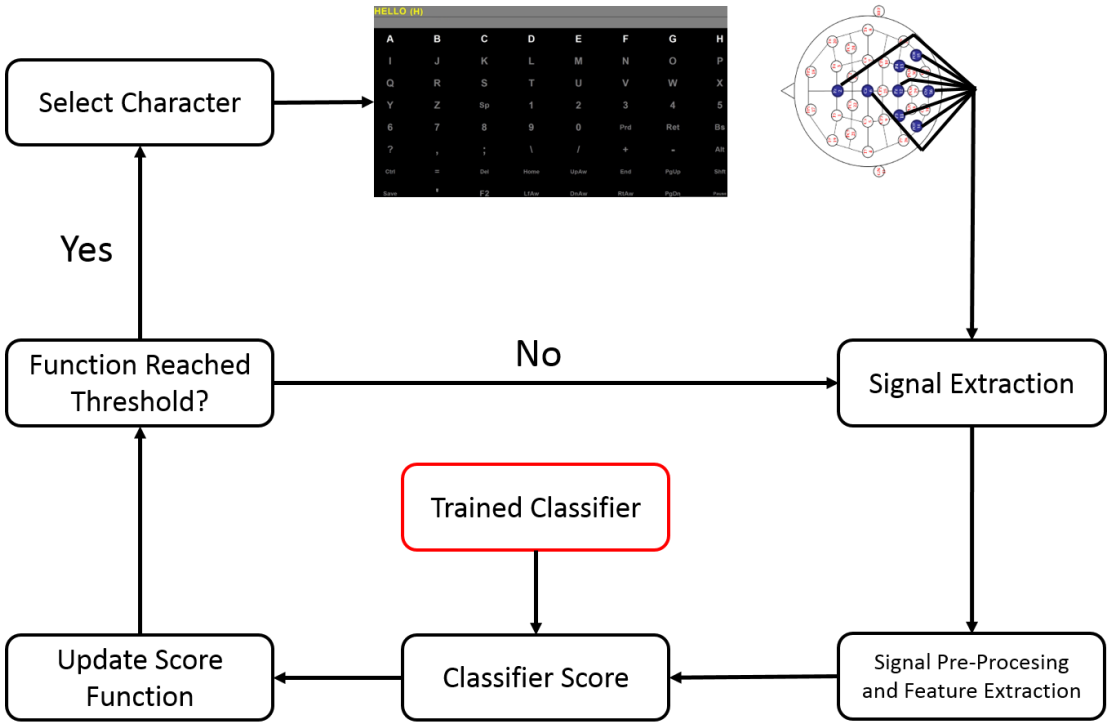

Figure 2.3 – P300 Speller Flow Chart. The user of the P300 speller (upper right) focuses on the character that they are intending to spell. The grid illuminates a subset of the characters on the grid. After each subset illumination, a window of EEG data is collected and preprocessed. A classifier score, obtained from the features extracted from the window of EEG data, is used to update the scoring function. After the scoring function meets some threshold, or the maximum number of presentations (i.e., illuminations) is exceeded, the character that maximizes the scoring function is selected.

2.2.1 P300 Speller Interface

A P300 speller grid displays all the symbols or characters that are available for the user to spell. At each flash index, t, a flash group (i.e., subset of characters) is illuminated. Let $\mathcal{F}_t \in \{0,1\}^{1\,x\,M} = [\mathbb{f}_{t,1}, \mathbb{f}_{t,2}, \ldots, \mathbb{f}_{t,M}]$ denote a binary vector that represents a flash group at time index t. $\mathbb{f}_{t,m} \in \{0,1\}$ indicates whether character m at t is presented, where a '1'

indicates a character that is presented and a '0' denotes a character that is not presented. M is the total number of characters in the grid, and m is an indexing value for the characters. $\boldsymbol{\mathcal{F}}_t \in \{0,1\}^{t,M} = [\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_t]$ is the matrix of all flash groups after t presentations. An example of a P300 speller grid is shown in Figure 2.4. In the Figure, the characters in the top row are being presented, or flashed, to the user.

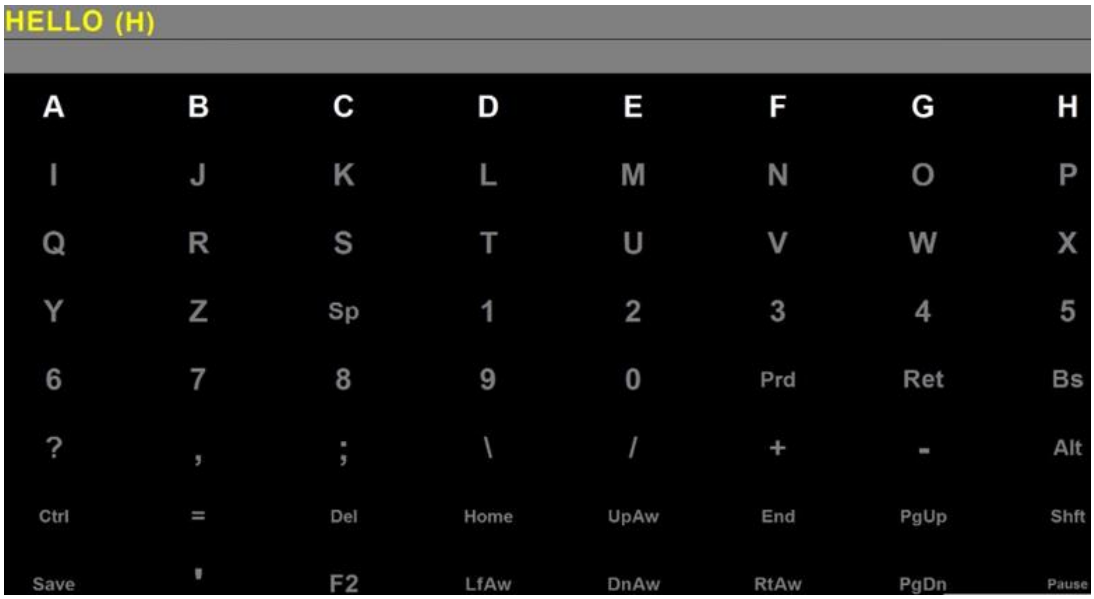

Figure 2.4 - A 9 x 8 (9 rows and 8 columns) P300 speller grid with a subset of characters illuminated in the top row. The grid displays all of the characters (characters, numbers, punctuation, commands, etc.) available for the user to spell. In this figure, the top row of characters are being presented to the user by being illuminated.

2.2.2 EEG Signal acquisition and processing

Electrical signals are collected externally on the scalp using EEG. In this work, EEG data was collected using 32-channel caps from Electro-Cap International Inc. connected to 16-channel GugerTec g.USBAMP Biosignal amplifier. An electrode map of the EEG cap

is shown in Figure 2.5. Data from the Fz, Cz, P3, Pz, P4, P07, P08, and Oz electrodes are collected [64], [65] as this set of electrodes was shown to generally yield high performance in the P300 speller [65]. The reference and ground are set to the right and left mastoid electrodes, respectively. The raw EEG data were collected at 256 Hz. Raw EEG data collected from the scalp is then pre-processed (i.e., filtered) to remove artifacts and amplify the EEG signal [1]. After the signal is pre-processed, features to be used for classification are extracted (see section 2.2.3).

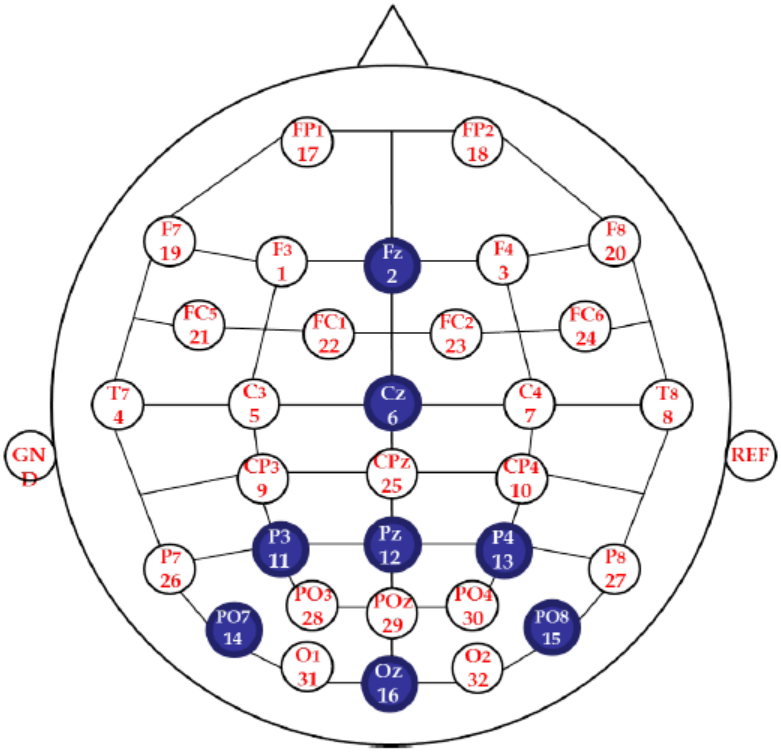

Figure 2.5 – Locations on the scalp where EEG electrodes were placed. The blue electrodes denote the electrodes used in this work (Fz, Cz, P3, Pz, P4, P07, P08, and Oz).

2.2.3 Feature Extraction and Classification

The feature extraction process from Krusienski et al. [65] is used in this work. A single feature observation at flash index, t, will be denoted $X_t \in \mathcal{R}^{1\,x\,D}$, where D is the dimensionality of an EEG feature observation. An 800ms time window is extracted from eight channels. The pre-processed EEG data is then decimated, by averaging, yielding 15 EEG data samples across 8 channels, or 120 total samples when vectorized; the 120 vectorized samples constitute a single feature observation.

Prior to the to the testing phase, a classifier is trained using labeled observations of features generated during a calibration, or training, phase of the experiment. A label, $\ell_t \in \{0,1\}$, is assigned at each flash index, where 0 corresponds to an observation where the target character is not presented, and 1 corresponds to an observation where the target character is presented. The entire training dataset consists of all observations of features, $\boldsymbol{X}_T = [X_1\ , X_2, \dots, X_T]$, and the binary labels, $\mathcal{L}_T = \{l_1, l_2, \dots, l_T\}$ collected during a training phase that consists of T presentations. Training the classifier requires learning a function that transforms a feature observation, X_t, into a classifier score, $z_t \in \Re^1$, such that the separation of classifier scores between target and non-target observations is maximized. A vector of classifier scores is denoted as $Z_t \in \Re^t = [z_1, z_2, \dots, z_t]$.

Krusienski et al. [65] shows that linear and non-linear classifiers yield similar classification accuracy for the EEG data features that were used; therefore, a linear classifier will be used for this work . In training a linear classifier, a classifier weight vector,

$W_{EEG} \in \mathcal{R}^{1\,x\,D}$, is learned such that $z_t = X_t {W_{EEG}}^{\mathrm{T}}$, where T denotes a matrix transpose. The stepwise-linear discriminant analysis (SWLDA) classifier is commonly used for processing EEG data (e.g., [11], [12], [13]), and will therefore be used to learn W_{EEG} in this work. After W_{EEG} is learned with the training data, classifier scores, z_t can be estimated for testing data (i.e., data whose label is unknown).

2.2.4 Target Character Estimation

During the testing phase, the goal of the P300 speller is to select the character that the user is attempting to spell. Determining if a presentation is a target or non-target presentation is difficult due to the low SNR of EEG data [67]. Therefore, character selection is typically made after many stimulus presentations [68], [69]. The total process of spelling for each character is termed a *trial*. The number of stimulus presentations for a given trial can either be fixed or determined dynamically, termed static and dynamic stopping respectively. Static stopping was used in the original P300 speller implementation [23] and is still used in some recent P300 literature (e.g., [25], [33], [34]). Dynamic stopping methods, however, have been shown to decrease the amount of time required to spell with the P300 speller compared to static stopping methods [70] and have been more commonly used in recent literature (e.g., [71]–[76]).

2.2.4.1 Static Stopping

The typical method to select characters via static stopping is the cumulative moving average (CMA) [23]. In the CMA character selection algorithm, a cumulative score

is computed after a fixed number of stimuli have been presented to the user during a trial; the cumulative score for each character is the average of all classifier scores, in a trial, for all observations in which that character was presented. The cumulative score is defined as the following:

$$\psi_{m,T} = \frac{\sum_{t=1}^{T} z_t \mathbb{1}_{t,m}}{\sum_{t=1}^{T} \mathbb{1}_{t,m}} \tag{2.1}$$

where $\psi_{m,T}$ is the cumulative score for character index m, after completing T presentations in a trial. The speller then selects the character that yields the maximum cumulative score; this is formally written as $\hat{c}_* = \arg\max_m \psi_{m,T}$.

The static stopping configuration, however, can result in low character estimation accuracy if there are too few stimulus presentations or a slow communication rate if there are too many stimulus presentations, as discussed in [72], [77]–[80]. To alleviate the dependency on choosing the optimal number of stimulus presentations, a dynamic stopping configuration can be used instead of a static stopping configuration.

2.2.4.2 Dynamic Stopping

Previous work has shown that using dynamic stopping, can increase the user communication rates [70]. Throckmorton et al. [72] developed a Bayesian probabilistic dynamic stopping algorithm, where the data collection can be terminated if any of the

characters' probability of being the target character reaches a predefined threshold. The character probabilities in Throckmorton et al. are updated as follows:

$$p(c_m = c_*|Z_t, \boldsymbol{\mathcal{F}}_t) = \frac{p(z_t|c_m = c_*, \mathcal{F}_t)p(c_m = c_*|Z_{t-1}, \boldsymbol{\mathcal{F}}_{t-1})}{\sum_{m=1}^{M} p(z_t|c_m = c_*, \mathcal{F}_t)p(c_m = c_*|Z_{t-1}, \boldsymbol{\mathcal{F}}_{t-1})} \tag{2.2}$$

$$p(z_t|c_m = c_*, \mathcal{F}_t) = \begin{cases} p(z_t|H_1), \mathcal{F}_{t,m} = 1 \\ p(z_t|H_0), \mathcal{F}_{t,m} = 0 \end{cases} \tag{2.3}$$

where z_t is the classifier score at the current flash index t; c_m is a character indexed by m; c_* is the target character for a trial; $p(c_m = c_*|Z_t, \boldsymbol{\mathcal{F}}_t)$ is the posterior probability for c_m; $p(z_t|c_m = c_*, \mathcal{F}_t)$ is the classifier score likelihood (i.e., conditional classifier score probability density) at presentation t; and $p(c_m = c_*|Z_{t-1}, \boldsymbol{\mathcal{F}}_{t-1})$ is the prior character probability, equivalent to the posterior probability after the previous presentation, $t-1$, conditioned on all previous classifier scores, Z_{t-1}, and previous flash groups $\boldsymbol{\mathcal{F}}_{t-1}$. The conditional classifier score density for $c_m = c_*$ is $p(z_t|H_1)$ if c_m is presented or $p(z_t|H_0)$ otherwise. After every presentation, t, the posterior probability is updated for c_m for all m using the Bayesian update equation (2.2). When $p(c_m = c_*|Z_t, \boldsymbol{\mathcal{F}}_t)$ reaches some

threshold for any m, data collection is terminated and the character c_m is selected as the estimated target character, $\widehat{c_*}$.

The conditional data probability density functions (PDFs), $p(z|H_1)$ and $p(z|H_0)$, can be estimated by first transforming all EEG training data features $\boldsymbol{X}_T$ to classifier scores, Z_T using the trained weights $\boldsymbol{W}_{EEG}$ learned from $\boldsymbol{X}_T$ and $\mathcal{L}_T$. Next, Z_T are grouped by target and non-target presentations and the conditional probability densities for target and non-target classifier scores be estimated independently. In Throckmorton et al. [72], the kernel density estimator (KDE) is used to estimate the conditional probability densities $p(z|H_1)$ and $p(z|H_0)$. Figure 2.6 shows an example of the target and non-target classifier PDFs that were estimated for a subject, with the KDE, using the collected training data. Figure 2.7 shows the evolution of the posterior probability for a trial using the Bayesian update equation (2.2). After the posterior probability of a target character exceeds the threshold (set to 0.9 in the example), data collection is terminated and the character exceeding the threshold is selected.

Throughout the remainder of this work, the dynamic stopping configuration as proposed by Throckmorton et al. [72] will be used for character selection and determining when to terminate the data collection. For simplicity, any reference to *Bayesian algorithm* in this work will refer to the Throckmorton et al. Bayesian character selection algorithm.

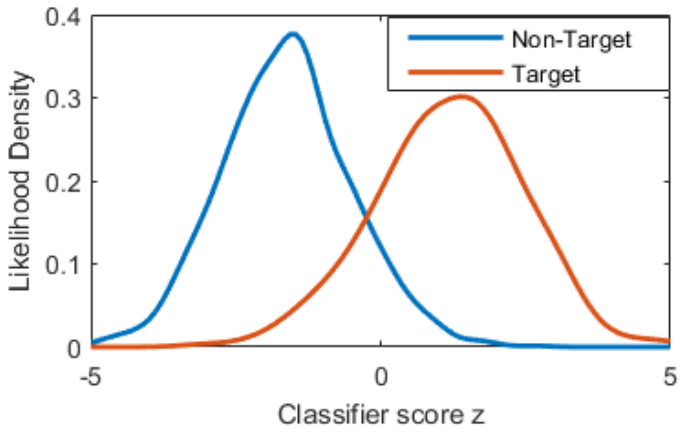

Figure 2.6 – An example of the target and non-target conditional classifier score PDFs

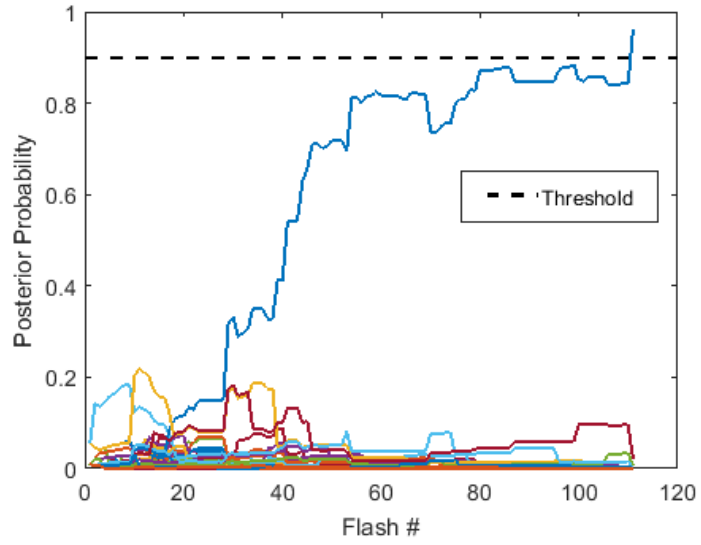

Figure 2.7 – An example of posterior probability updates as a function of flash index, using collected EEG data. The Bayesian algorithm was used to update the posterior probability for all the characters. The dashed-black line indicates the probability threshold that terminates the epoch. Each solid line shows the posterior probability for a character.

2.2.5 Stimulus Presentation Paradigms

An important aspect of the P300 speller is the stimulus presentation (i.e., flashing) paradigm (i.e., pattern). P300 spellers that present one character at each flash index are referred to as rapid serial visual presentation (RSVP) paradigms [81]. The RSVP speller, however, will usually require many character presentations before selecting the estimated target character. To increase the character presentation rate, most P300 speller systems flash multiple characters simultaneously on a speller grid.

The most common presentation paradigm used for P300 spellers is the row column (RC) paradigm [23], where the flash groups consist of only row and column flashes. Figure 2.8 shows all of the possible flash groups in sequence for a RC stimulus presentation paradigm given a 9 x 8 grid. The set of all possible flash groups in a presentation paradigm is sometimes referred to as a codebook in the literature (e.g., [82], [83]). Though the RC paradigm requires fewer presentations than the RSVP paradigm, it can be negatively affected by physiological phenomena such as refractory effects and adjacency distraction errors (see section 2.3). Refractory effects occur when the time between target presentations is insufficient, and adjacency distractions occur when characters adjacent to the target character are presented.

The checkerboard (CB) paradigm [25] was created as a way to alleviate the physiological phenomena that negatively impact the accuracy of the RC paradigm. The

CB paradigm creates two virtual matrices from a checkerboard overlay on a grid. Only characters in the same virtual matrix can be presented within the same flash group.

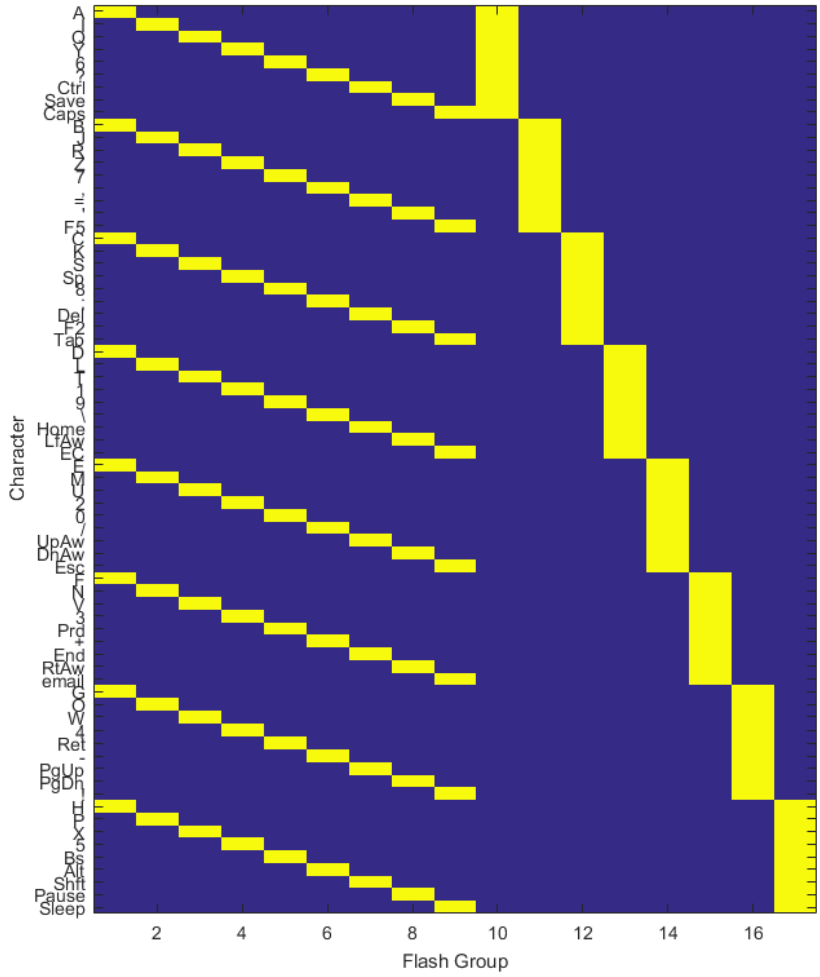

Figure 2.8 – All possible flash groups (i.e., the codebook) for a row-column (RC) stimulus presentation paradigm given a 9 x 8 character P300 speller. A yellow entry denotes that a presentation for the character shown on the y-axis occurs during the flash group on the x-axis. For example, if flash group 1 is presented, the characters A, B, C, D, E, F, G, and H are presented, constituting a presentation/flash of a row of characters.

2.2.5.1 Adaptive Stimulus Selection

A significant amount of work has been done towards designing stimulus presentation paradigms that aim to improve the accuracy and decrease the time it takes to spell a character using the P300 speller (e.g., [37], [38], [84]–[86]). Most of these improvements, however, do not rely on previously observed EEG data that was measured during the spelling sequence to optimize the stimulus selection (i.e., they are not data driven). These non-adaptive paradigms can present a subset of characters such that little to no new information is actually gained. For example, if all of the characters in the first row have a target character probability of ≈ 0 or a low cumulative score, presenting that first row of characters again provides essentially no new information to the speller.

Mainsah et al, [37], developed a codebook generation paradigm that derives a codebook that maximizes the predicted accuracy of the P300 speller as a function of the user's EEG data discriminability; this paradigm was termed the performance-based paradigm (PBP). The data collected in the training phase (i.e., training data) is first used to determine the EEG data discriminability, and subsequently the PBP codebook is generated. However, during the testing phase of the experiment, random flash groups are selected from the codebook generated after the training phase. Though the PBP paradigm adapts to the user's ability to use the P300 speller, it is not adaptively choosing stimuli to present to the subject. In this work, adaptive paradigms refer only to paradigms where stimuli are adaptively selected during the testing phase.

Some work has been done on the development of an adaptive, data-driven, stimulus presentation paradigm for the P300 speller. In Park et al. [39], a partially observed Markov decision process (POMDP) was used in conjunction with a RC paradigm to adaptively decide which row or column to flash. The row and column decisions were made independently, and the intersection of the selected row and column was then selected as the target character estimate. Ma et al. [40] implemented a hierarchy of variable-sized flash groups based on a language model. Both adaptive flashing methods showed improvement over a RC speller, showing the potential of an adaptive flashing paradigm.

2.3 Physiological Phenomena

In the previous subsection, stimulus presentation paradigms were introduced and methods used to improve stimulus presentation paradigms were discussed. In this section, physiological phenomena that can affect the accuracy and spelling speed of the P300 speller will be discussed. It is vital to consider these physiological phenomena when designing a stimulus presentation paradigm as will be done in Chapter 4.

2.3.1 Refractory Effects

Refractory effects occur when the time between target stimulus presentations is not sufficiently long, resulting in decreased SNRs of the ERP [43], [87]. Consequently, refractory effects make it more difficult to distinguish between nontarget and target responses, negatively impacting the character selection abilities of the classifier. The

degree of refractory effects can be quantified by amount of time between two consecutive target character presentations, also refer to as the target-to-target interval (TTI). In this study, the TTI is defined based on the number of stimulus presentations between two consecutive target character presentations. For example, if 'T' and 'N' are denoted as target and non-target stimulus presentations, respectively, a flash pattern denoted TNNT represents a TTI of 3. Previous work has shown that the amplitude and latency of a P300 response can be affected by the TTI [42], [43], [82], [88], [89].

2.3.1.1 Signal Detectability

To analytically assess how distinguishable a P300 signal is from a non-P300 signal, the detectability index [90] can be used. The Detectability index [90] is a metric that measures the separation between the target and non-target distributions of two Gaussian distributions. The higher the detectability index, the easier it is to distinguish between target and non-target observations. Formally, the detectability index is defined by equation (2.4).

$$d' = \frac{\mu_S - \mu_n}{\sqrt{\frac{1}{2}(\sigma_s^2 + \sigma_n^2)}} \tag{2.4}$$

where μ_s and μ_n are the means for the target and non-target distributions respectively, and σ_s^2 and σ_N^2 are the variances for the target and non-target distributions, respectively.

To compare the detectability of a P300 signal at different TTIs, training and testing data from Throckmorton et al. [72] were analyzed. For each subject, the classifier scores, $\boldsymbol{z}$, were grouped by TTI, and the detectability at each TTI was calculated. Figure 2.9 shows the detectability index of the classifier scores, z, as a function of TTI between target and non-target flashes averaged across subjects using data recorded in the Throckmorton et al. [72] static stopping data.

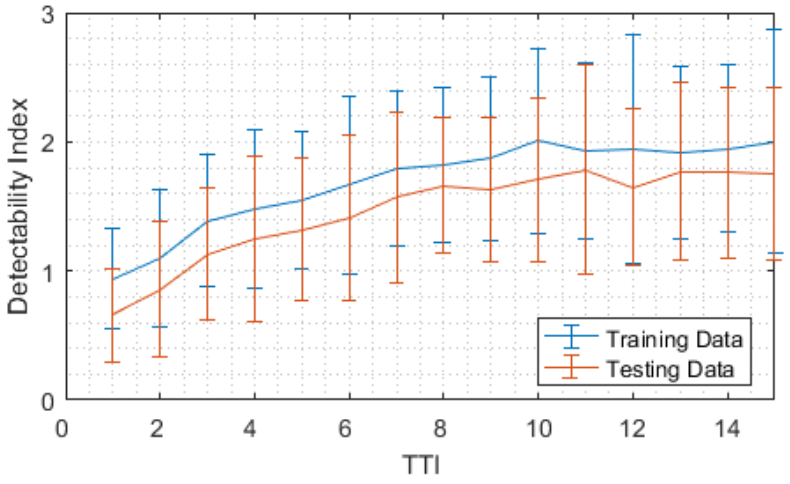

Figure 2.9 – Detectability index of the SWLDA classifier scores, z, as a function of TTI between target and non-target flashes averaged across subjects in the Throckmorton et al. [72] static stopping experiment. The red line corresponds to the testing data and the blue line corresponds to the training data. The solid lines represent the mean detectability index as a function of TTI and the bounds represent $\pm$ 1 standard deviation. As the TTI increases, the detectability increases as well. However, as the TTI reaches approximately 10, the detectability begins to plateau.

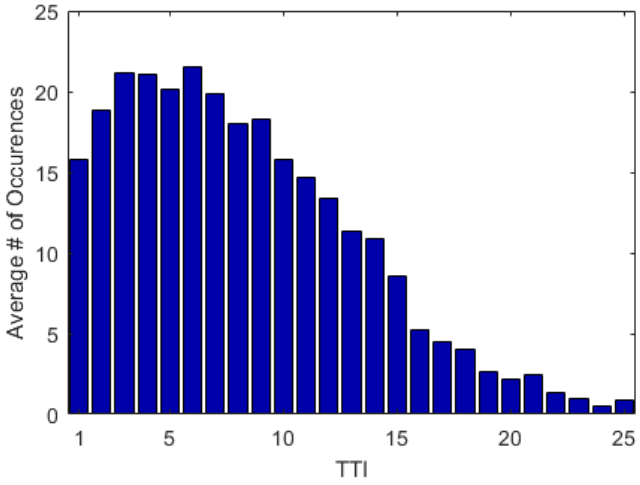

Figure 2.10 - The distribution of average TTI occurrences in the Throckmorton et al. [66] static stopping testing data.

The high variance across participants in higher TTIs occurs because there are many more low-TTI target flashes than high TTI target flashes in the RC paradigm. This can be seen in Figure 2.10, which shows the distribution of average TTI occurrences in the Throckmorton et al. [72] data. As Figure 2.9 shows, the longer TTI target flashes typically lead to a more discriminable P300 response than lower TTI target flashes. These results are consistent with prior studies (e.g., [43], [82], [87]).

Refractory effects can be alleviated by using a presentation paradigm that inherently or explicitly restricts characters from being presented with a low TTI. The CB paradigm, for example, has a minimum TTI of 7 for a 6 x 6 character grid. Other paradigms that have been studied (e.g., [37], [39]) explicitly set a minimum TTI restriction of 3.

2.3.2 Adjacency Distraction

Typically, the most commonly confused characters are adjacent to the target character [25], [44]. This type of error is referred to as an adjacency distraction error. In the RC paradigm, the most common errors occur in the same column or row as the target character, with a higher likelihood of errors for characters directly adjacent to the target character. Adjacency distraction errors are difficult to remove when using a speller grid because removing adjacent character presentations requires the knowledge of the target character. One way to mitigate the effects of these errors is to ensure that adjacent characters cannot be presented in the same flash group. The CB paradigm, for example, can only present characters in one of the two checkerboards, therefore adjacent characters are never presented. This paradigm yielded an error map such that the most common errors were no longer adjacent to the target character [25].

2.3.3 Fatigue

Several researchers have shown that the amplitude and latency of the P300 response can be affected by the user's fatigue level [91]. In [91], participants were instructed to use an auditory P300 speller while under medium and high mental work-loads. It was shown that under high work-loads, the users became more fatigued and performed the auditory P300 task worse than under medium-workloads. While it is difficult to prevent fatigue, it is possible to monitor the level of fatigue by monitoring alpha band power. Several studies have shown that the alpha band (7.5–12.5 Hz) power decreases as the level of mental

workload or fatigue increases [92]–[94]. In [95], the alpha band power was used as a threshold to determine whether or not to select a character spelled by the P300 speller. Using the alpha-band power threshold improved the accuracy of the speller.

2.4 Performance Assessment

In the P300 literature, accuracy, expected stopping time (EST), and bit-rate are the most commonly used metrics to determine the performance of the speller (e.g., [66], [72], [82], [96]). Accuracy, $\mathcal{A}$, is the ratio of characters spelled correctly to the total number of characters spelled. While the accuracy is an important metric to determine how well the speller works, it should be presented along-side with the time required to spell. For example, the user may prefer a speller that is 85% accurate, and requires 5 seconds to spell a character to a speller that is 90% accurate and requires one minute to spell a character. Therefore, the average number of presentations, or EST, is another important metric when evaluating the BCI performance. The trade-off between accuracy and expected stopping time can be assessed with the bit-rate metric. Bit-rate (BR) is used in information theory to determine the amount of information that is being transmitted over a channel [97]. In the context of the P300 speller, the bit-rate, BR, is defined as the following:

$$BR = \log(M) + \mathcal{A}\log(\mathcal{A}) + (1-\mathcal{A})\log\left(\frac{1-\mathcal{A}}{M-1}\right) \tag{2.5}$$

where M is the number of possible selections on the P300 speller grid. A higher bit-rate indicates that more information is being transferred. Given that the amount of characters on a speller grid is constant, BR increases when $\mathcal{A}$ increases, or EST decreases. Inversely, BR decreases when $\mathcal{A}$ decreases, or EST increases.

These metrics can be evaluated online or offline. The online evaluation uses the spelled character during the online phase to measure the metrics. The offline evaluation uses EEG data collected during the online phase in offline simulations.

2.4.1 Online Evaluation

Online evaluation of accuracy, average number of presentations, and bit-rate are measured by the output of the P300 speller. The online evaluation is able to take into account physiological factors (e.g., refractory effects, fatigue, and distractions) that offline evaluation may ignore. Online evaluations are typically considered the 'gold-standard' when evaluating a communication device [82]. Online metrics, however, typically have a low number of spelled characters, and therefore have high variances and granularity in their performance estimates.

2.4.2 Offline Evaluation

Offline simulations can be used a way to increase the precision of the performance metrics by using the collected EEG data or synthetic data drawn from the online data collection. Offline estimates compute the projected accuracy, projected bit-rate, and

projected EST. These projected metrics attempt to estimate the respective online metrics if infinitely many characters were spelled during the online data collection.

One method that can be used to estimate the projected metrics is the Monte Carlo (MC) simulation, a random sampling method used to estimate the expected value of a distribution [98]. The MC simulation can be performed by using the collected EEG data or an estimated distribution of the classifier scores of the collected EEG data [37]. In this work, all offline simulations are performed using the MC approach. All of the MC simulations in this work are performed using the Throckmorton et al. dynamic stopping framework [72]. The following steps are taken to complete a single iteration of MC simulation in this work:

1. Initialization:
    a. Uniformly randomly select the target character, c_*.
    b. Initialize target character probabilities, $p(c_m)$
    c. Initialize flash groups (i.e., $\mathcal{F}_1, \mathcal{F}_2, \dots,)$
    d. Set $t = 1$
2. Do while stopping criteria has not been met
    a. Draw z_t, the EEG classifier scores at t. Draw an H_1 example from the testing data if $\not{f}_{t,m} = 1$, or draw an H_0 example from the testing data otherwise
    b. Update target character probabilities
    c. Update flash groups if using adaptive stimulus selection paradigm

d. Set $t = t + 1$

3. Estimate target character

Figure 2.11 below shows the projected accuracy, average number of presentations, and bit-rate as function of the number of MC simulation iterations. As can be seen in the figure, the projected accuracy plateaus at ~50 iterations and the projected number of presentations and bit-rate plateau even earlier.

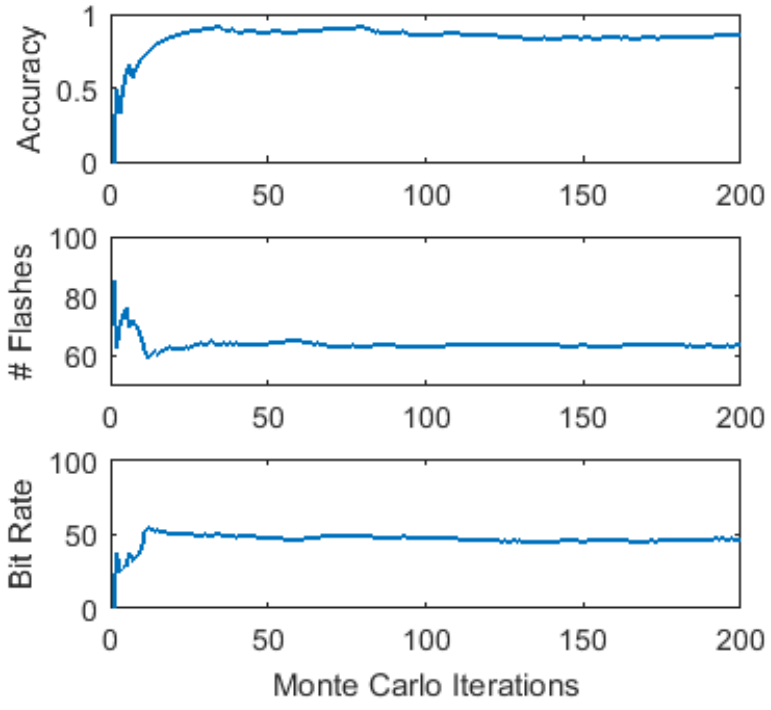

Figure 2.11 – The projected accuracy, projected presentations, and projected bit-rate as a function of the number of MC simulation iterations.

While the MC simulations increase the precision of the performance estimates, current simulation models ignore physiological effects during real-time BCI, such as refractory effects, fatigue, and user behavior modification due to feedback. Therefore, when assessing the performance of a P300 speller configuration, it is beneficial to perform both online and offline evaluations of the data.

2.5 Summary

The background material required for the remainder of this work is discussed in this chapter. First, the way in which an eye-tracker is used for communication is discussed; furthermore, the eye-gaze hardware and data format used in this work, particularly in chapter 5, is described. Next, the P300 speller was introduced and various advances that were made to improve the accuracy and spelling speed of the P300 speller were discussed. Physiological phenomena that can affect the accuracy and spelling speed of the P300 speller were also described. Finally, the evaluation metrics that are to assess the performance of the communication devices in this work were introduced.

3 Adaptive Stimulus Selection for the P300 Speller

As discussed in Chapter 2, various stimulus presentation paradigms (e.g., RC, CB) have been developed for the P300 speller. However, most of these paradigms are not data driven [23], [25], [37], [81] in that they do not rely on previously observed EEG data to optimize stimulus presentation, and as such are not data adaptive. Some previous work has been done on the development of a data-driven adaptive paradigm [39], [40]. The POMDP approach proposed in [39] is intractable for a real-time system when considering a search space considerably larger than row and column flash groups and the hierarchical approach [40] is easily susceptible to error propagation.

Of the data-driven approaches described above, the data-driven adaptive stimulus presentation proposed in this work is most closely related to the POMDP approach as it does not use a hierarchy structure. Instead of maximizing a reward function over a time horizon, the proposed approach attempts to maximize the information gained, using the EDG metric, for every stimulus presentation, or flash. This myopic approach to stimulus presentation allows for dynamically created flash groups that are not restricted to a dictionary of flash groups such as rows and columns.

In this chapter, the Kastella sampling framework based on the EDG is introduced and then modified to adapt it to the P300 speller framework. Physiological phenomena that might negatively affect the accuracy or speed of the proposed EDG adaptive paradigm are discussed and constraints on the proposed stimulus presentation

framework are introduced to mitigate these effects. Numerical simulations using synthetic data are carried out in order to measure the efficacy of the proposed paradigm. Finally, the EDG adaptive stimulus presentation paradigm is tested in an online data study to evaluate its performance in a real-time P300 speller.

3.1 Kastella Sampling Framework

The Kullback-Leibler (KL) divergence was used by Kastella as the measure of discrimination between two distributions [99]. The KL divergence is a non-symmetric similarity measure between two probability distributions; the higher the KL divergence, the more discriminable the distributions are. The continuous and discrete forms of the KL divergence between distributions p and q (denoted as vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ if discrete), parametrized by x, are shown in equations (3.1) and (3.2), respectively.

$$D_{KL}(p\|q) = \int p(x) \log\left(\frac{p(x)}{g(x)}\right) dx \tag{3.1}$$

$$D_{KL}(\boldsymbol{p}\|\boldsymbol{q}) \quad = \sum_{x\in S} \boldsymbol{p}_x \log\left(\frac{\boldsymbol{p}_x}{\boldsymbol{g}_x}\right) \tag{3.2}$$

where S is the set of all values that x can take in the discrete case. KL divergence is non-negative and is only equal to 0 when distributions p and q are equivalent.

Kastella presents the problem of optimizing a sensor data sampling strategy that maximizes the correct identification ('target' or 'non-target'), where targets are spaced within a grid of cells where data can be sampled. Translating this idea to the P300 speller, the 'target' is the character that the user is intending to spell that is located within the grid of character cells. The overall goal of Kastella's work is to maximize the expected discrimination gain, ΔD_{KL}, that is achieved by a collected sample. The discrimination gain is defined as:

$$\Delta D_{KL}(c) = E[D_{KL}(\mathcal{p}_{c,t+1} \| \mathcal{p}_{c,0}) | c] - D_{KL}(\mathcal{p}_{c,t} \| \mathcal{p}_{c,0}) \tag{3.3}$$

where $\mathcal{p}_{c,t+1}$ is the probability that cell c is a 'target' at time index $t+1$, given that c is sampled. Intuitively, the discrimination gain is the information that is expected to be gained after cell c is sampled at $t+1$ relative to the information that we have at the current flash t. The expectation in (3.3) can be expanded to:

$$E[D_{KL}(\mathcal{p}_{c,t+1} \| \mathcal{p}_{c,0}) | c] = \int_{-\infty}^{\infty} [D_{KL}(\mathcal{p}_{c,t+1} \| \mathcal{p}_{c,0}) | z^*_{c,t+1} = z] p(z^*_{c,t+1} = z | \mathbf{z}_{c,t}) \, dz \tag{3.4}$$

where $z^*_{c,t+1}$is the observation of cell c at $t+1$; $z^*_{c,t+1}$ is unknown, therefore it is integrated out over all possible observations, z, to obtain the expected discrimination between probabilities obtained at time index $t+1$ and the initial probabilities. The term $p(z^*_{c,t+1} =$

$z|\boldsymbol{z}_{c,t}$) is the data probability of the classifier score given all previously observed observations of cell c.

After calculating the expected discrimination gain for all cells c, a decision is made to sample the cell with the maximum expected discrimination gain. Kastella showed that using this method of adaptive sampling yielded a higher accuracy, given the same number of collected samples, than randomly choosing cells to sample.

3.2 Expected Discrimination Gain Optimization Function

The proposed data-driven adaptive stimulus presentation paradigm couples the Kastella stimulus selection paradigm [41] and the Bayesian dynamic stopping algorithm [72]. As discussed in section 2.2.4.2, the Bayesian dynamic stopping algorithm can increase the spelling speed and accuracy of the P300 speller by estimating the target probability of a character after every stimulus presentation.

To simplify the notation, the Bayesian update equation can be rewritten as follows:

$$\mathcal{p}_{t,m} = \frac{\ell_{t,m}(z_t)\mathcal{p}_{t-1,m}}{p(z_t|z_{T-1})} = \frac{\ell_{t,m}(z_t)\mathcal{p}_{t-1,m}}{\sum_{c=1}^{M}\ell_{t,c}(z_t)\mathcal{p}_{t-1,c}} \tag{3.5}$$

$$\ell_{t,m}(z_t) = \begin{cases} \ell 0(z_t), & \mathcal{f}_{t,m} = 0 \\ \ell 1(z_t), & \mathcal{f}_{t,m} = 1 \end{cases} \tag{3.6}$$

It is important to note that (3.5) and (3.6) are identical to (2.2) and (2.3), but with simplified notation. $\wp_{t,m}$ and $\wp_{t-1,m}$ are the prior and posterior probabilities for character c_m. $\wp_{t,m}$ and $\wp_{t-1,m}$ are synonymous to $p(c_m = c_*|Z_t, \boldsymbol{\mathcal{F}}_t)$ and $p(c_m = c_*|Z_{t-1}, \boldsymbol{\mathcal{F}}_{t-1})$, respectively. $\ell 0(z_t)$ and $\ell 1(z_t)$ are the target and non-target classifier score likelihoods, equivalent to $p(z_t|H_0)$ and $p(z_t|H_1)$.

With the simplified notation for the Bayesian update equation, the ΔD_{KL} from equation (3.3) can be rewritten as follows:

$$\Delta D_{KL}(\mathcal{F}^*_{t+1}) = E_{z^*_{t+1}|Z_t}[D_{KL}(\mathcal{P}_{t+1}\|\mathcal{P}_0)|\mathcal{F}^*_{t+1}] - D_{KL}(\mathcal{P}_t\|\mathcal{P}_0) \tag{3.7}$$

where ΔD_{KL} is the EDG at t given a flash subset $\mathcal{F}^*_{t+1}$. $\mathcal{P}^*_{t+1}$ is the hypothetical posterior distribution at $t+1$ over all possible future observations, conditioned on the current set of observations $z^*_{t+1}|Z_t$. The exact value of $\mathcal{P}^*_{t+1}$ is unknown, hence the expectation is over $z^*_{t+1}|Z_t$. $\mathcal{P}_0$ is the character probability vector prior to data collection (i.e., the prior).

The following optimization goal is then used to select a flash group to present at $t+1$:

$$\mathcal{F}^{\ddagger}_{t+1} = \underset{\boldsymbol{\mathcal{F}}^*_{t+1}}{\operatorname{argmax}} \Delta D_{KL}(\mathcal{F}^*_{t+1}), \forall \mathcal{F}^*_{t+1} \in \boldsymbol{\Omega} \tag{3.8}$$

where the goal is to choose a subset of characters to present, $\mathcal{F}_{t+1}^{*}$, from a search space of flash groups, $\boldsymbol{\Omega}$, such that ΔD_{KL} is maximized. The resulting flash group presented at $t+1$ is $\mathcal{F}_{t+1}^{\ddagger}$.

The work presented in Kolba [100] shows that if the sensor samples are conditionally independent, $\Delta D_{KL}(\mathcal{F}_{t+1}^{*})$ can be rewritten as the expected KL divergence between the current and hypothetical posterior probabilities. Though this assumption is not entirely accurate due to various physiological effects, it is the assumption made in the Bayesian update equation (3.5). Consequently, $\Delta D_{KL}(\mathcal{F}_{t+1}^{*})$ can be rewritten as follows:

$$\Delta D_{KL}(\mathcal{F}_{t+1}^{*}) = E_{z_{t+1}^{*}|Z_t}[D_{KL}(\mathcal{P}_{t+1}^{*} \| \mathcal{P}_t) | \mathcal{F}_{t+1}^{*}] \tag{3.9}$$

$$\Delta D_{KL}(\mathcal{F}_{t+1}^{*}) = \int_{-\infty}^{\infty} \sum_{m=1}^{M} \wp_{t+1,m}^{*} \log\left(\frac{\wp_{t+1,m}^{*}}{\wp_{t,m}}\right) p(z_{t+1}^{*}|Z_t) dz_{t+1}^{*} \tag{3.10}$$

Replacing $\wp_{t+1,m}^{*}$ in (3.10) with the update equation (3.5) yields the following:

$$\Delta D_{KL}(\mathcal{F}_{t+1}^{*}) = \int_{-\infty}^{\infty} \sum_{m=1}^{M} \frac{\wp_{t,m}\ell_{t+1,m}(z_{t+1}^{*})}{\cancel{p(z_{t+1}^{*}|Z_t)}} \log\left(\frac{\cancel{\wp_{t,m}}\ell_{t+1,m}(z_{t+1}^{*})}{p(z_{t+1}^{*}|Z_t)\cancel{\wp_{t,m}}}\right) \cancel{p(z_{t+1}^{*}|Z_t)} dz_{t+1}^{*} \tag{3.11}$$

$$= \int_{-\infty}^{\infty} \sum_{m=1}^{M} \wp_{t,m}\ell_{t+1,m}(z_{t+1}^{*}) \log\left(\frac{\ell_{t+1,m}(z_{t+1}^{*})}{p(z_{t+1}^{*}|Z_t)}\right) dz_{t+1}^{*} \tag{3.12}$$

The binary choice in the character likelihood assignments in (3.6) during the Bayesian update process (3.5) allows evaluation of the integral in (3.12) to be simplified to (3.13). More specifically, the discrete summation can be split into two summations; the group of presented characters and non-presented characters. By doing this separation, $\ell_{t+1,m}$ becomes $\ell 1$ if the m is presented (i.e., $\mathbb{f}_{t+,m} = 1$) or $\ell 0$ otherwise. This yields the following for ΔD_{KL}:

$$\Delta D_{KL}(\boldsymbol{\mathcal{F}}^*_{t+1}) = \int_{-\infty}^{\infty} \sum_{\forall m: \mathbb{f}^*_{t+1,m}=1} \wp_{t,m} \ell 1(z^*_{t+1}) \log\left(\frac{\ell 1(z^*_{t+1})}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1} + \int_{-\infty}^{\infty} \sum_{\forall m: \mathbb{f}^*_{t+1}=0} \wp_{t,m} \ell 0(z^*_{t+1}) \log\left(\frac{\ell 0(z^*_{t+1})}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1} \tag{3.13}$$

$$= P1_t \int_{-\infty}^{\infty} \ell 1(z^*_{t+1}) \log\left(\frac{\ell 1(z^*_{t+1})}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1} + P0_t \int_{-\infty}^{\infty} \ell 0(z^*_{t+1}) \log\left(\frac{\ell 0(z^*_{t+1})}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1} \tag{3.14}$$

where $P1_t$ and $P0_t$ are equivalent to:

$$P1_t = \sum_{\forall c:\, \mathbb{f}^*_{t+1,c}=1} \wp_{t,c} \tag{3.15}$$

$$P0_t = \sum_{\forall c:\, \ell^*_{t+1,c}=0} \mathscr{p}_{t,c} \tag{3.16}$$

$$P0_t + P1_t = \sum_{c=1}^{M} \mathscr{p}_{t,c} = 1 \tag{3.17}$$

The conditional data probability can be expressed as $p(z^*_{t+1}|Z_t) = \sum_{c=1}^{M} \ell_{t,c}(z_t)\mathscr{p}_{t,c}$. Similar to the decomposition in (3.13), the conditional data probability in (3.14) can be decomposed by grouping characters that are and grouping characters that are not presented at $t+1$ as shown in (3.18).

$$p(z^*_{t+1}|Z_t) = \sum_{c=1}^{M} \ell_{t,c}(z_t)\mathscr{p}_{t,c} = \sum_{\forall c:\, \ell^*_{t+1,c}=1} \mathscr{p}_{t,c}\ell 1(z^*_{t+1}) + \sum_{\forall c:\, \ell^*_{t+1,c}=0} \mathscr{p}_{t,c}\ell 0(z^*_{t+1}) \tag{3.18}$$

This can be further simplified to the following:

$$p(z^*_{t+1}|Z_t) = P1_t\ell 1(z^*_{t+1}) + P0_t\ell 0(z^*_{t+1}) \tag{3.19}$$

Using (3.19) as the conditional data probability, (3.14) can be rewritten as a function of z^*_t, $\ell 0$, $\ell 1$, $P1_t$ as shown in (3.20). All $P0_t$ can be replaced with 1-$P1_t$ by exploiting the sum of probabilities as shown in (3.17).

$$\Delta D_{KL}(\mathcal{F}^*_{t+1}) \tag{3.20}$$
$$= P1_t \int_{-\infty}^{\infty} \ell 1(z^*_{t+1}) \log\left(\frac{\ell 1(z^*_{t+1})}{P1_t \ell_1(z^*_{t+1}) + (1-P1_t)\ell_0(z^*_{t+1})}\right) dz^*_{t+1}$$
$$+ (1-P1_t) \int_{-\infty}^{\infty} \ell 0(z^*_{t+1}) \log\left(\frac{\ell 0(z^*_{t+1})}{P1_t \ell_1(z^*_{t+1}) + (1-P1_t)\ell_0(z^*_{t+1})}\right) dz^*_{t+1}$$

It is important to note that when evaluating the integral in (3.20) to determine ΔD_{KL}, z^*_{t+1} is being marginalized out over $\ell 0$ and $\ell 1$. Consequently, if $\ell 0$ and $\ell 1$ are only dependent on the next observation z^*_{t+1}, ΔKL_t depends only on $P1_t$. This simplification is particularly useful because the integral in (3.20) does not need to be directly computed to determine ΔKL_t for a proposed subset of characters presented, $\mathcal{F}^*_{t+1}$. Instead, a ΔD_{KL} curve as a function of $P1_t$ can be precomputed immediately following the training phase. Examples of the ΔD_{KL} curve over three $\ell 0$ and $\ell 1$ distributions are shown in Figure 3.1.

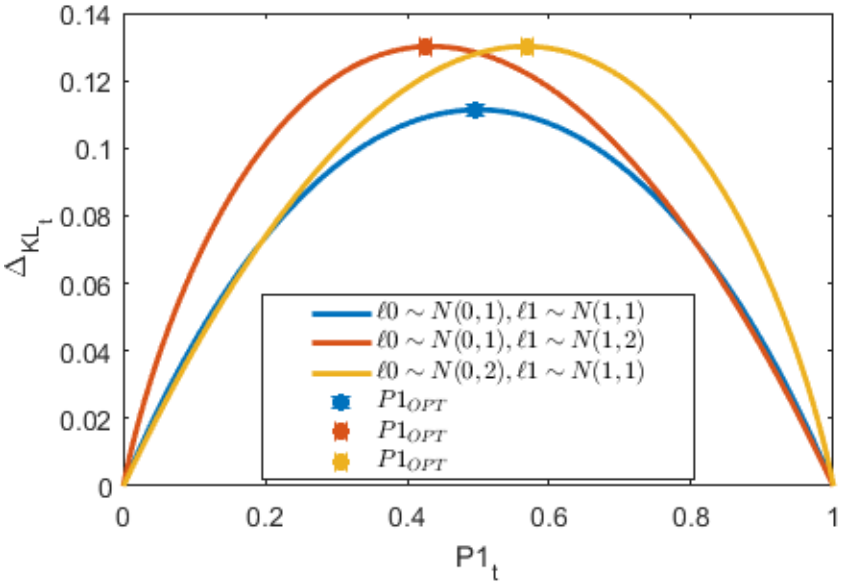

Figure 3.1 - Three examples of pre-computed expected discrimination gain curves, ΔKL_t as a function of $P1_t$. $P1_{OPT}$ is the $P1_t$ that yields the maximum ΔKL_t for a given function. $\ell 1$ and $\ell 0$ are distributed by a normal distribution $\mathcal{N}(\mu, \sigma^2)$ where μ and σ^2 are shown in the legend.

Given a proposed $\mathcal{F}_{t+1}^*$, $P1_t$ is computed per (3.15), and then ΔD_{KL} is indexed from the pre-computed curve.

As an aside, it is interesting to note that if the conditional independence assumption between sensor samples is made, the EDG metric is equivalent to the conditional mutual information (MI) [99] between c_* and z_{t+1}^* given Z_t: $I(c_*, z_{t+1}^* | Z_t)$. The proof for this equivalence is shown in appendix A. This equivalency was also noted in [101].

3.2.1 Greedy Expected Discrimination Gain Combinatorial Optimization

Given a search space of flash groups, $\boldsymbol{\Omega}$, the EDG can be used to determine the next flash group presentation, $\mathcal{F}_{t+1}^{\ddagger}$. For a fixed set of flash groups, such as row and column or checkerboard flash groups, $\mathcal{F}_{t+1}^{\ddagger}$ can be determined by selecting the flash group whose $P1_t$ maximizes ΔD_{KL}. However, stimulus selection from a small search space is unlikely to give the most optimal flash group. The flexibility to dynamically create flash groups is desirable in order to better achieve the objective to maximize ΔD_{KL}. Given the exponentially large space of 2^M possibilities, however, an exhaustive search is impractical.

To expand the search space, $\boldsymbol{\Omega}$, a greedy approach for stimulus selection is adopted in this work. Given the current character probabilities, $\mathcal{P}_t$, the character that maximizes ΔD_{KL} is iteratively chosen to be added to the flash group until no remaining character increases ΔD_{KL}. This method of stimulus selection is termed the greedy EDG adaptive paradigm; the greedy combinatorial optimization will be used to maximize the EDG for the remainder of this work.

3.3 Numerical Simulations for Adaptive Stimulus Presentation Paradigm

Following the framework described in section 2.4.2, numerical MC simulations of the P300 speller character selection process are performed to compare the proposed greedy EDG adaptive stimulus paradigm to the RC and CB paradigms. Initial simulations are performed to analyze the ideal or upper bound performance of the greedy EDG

adaptive stimulus presentation paradigm (section 3.3.1). Next, restrictions are placed on the greedy EDG adaptive paradigm to account for physiological and system constraints; simulations are performed for the greedy EDG adaptive stimulus presentation paradigm with the various restrictions (section 3.3.2).

For the greedy EDG adaptive paradigm with ideal conditions (section 3.3.1), the EDG optimization function shown in equation (3.8) is used. The ΔD_{KL} curve as a function of $P1_t$, generated by equation (3.20) is first used to compute the EDG. Next, the greedy combinatorial optimization described in section 3.2.1 is used to choose the subset of characters to present, $\mathcal{F}_{t+1}^{\ddagger}$.

Based on the flash groups defined by a stimulus paradigm condition, for non-target and target flash groups, the classifier scores are drawn from normal distributions for $\ell 0$ and $\ell 0$ respectively, defined accordingly:

$$\ell 1 \sim \mathcal{N}(\mu_1, \sigma^2) \qquad (3.21)$$
$$\ell 0 \sim \mathcal{N}(\mu_0, \sigma^2)$$

where μ_0 and $\mu_1$are the mean parameters for $\ell 0$ and $\ell 1$, respectively, and σ is the common standard deviation. For the simulations, σ^2 is set to 1, and μ_0 is set to 0. μ_1 is determined based on d' as defined in equation (2.4); given that $\sigma^2 \coloneqq 1$ and $\mu_0 \coloneqq 0$, $\mu_1 = d'$.

The Bayesian dynamic stopping update equation (3.5) is used to compute the target character probabilities. The target character probabilities are initialized to be

uniformly distributed. $\ell 0$ and $\ell 1$ are defined by (3.21) where parameters $\sigma^2 = 1, \mu_0 = 0, and\ \mu_1 = d'$. The stopping threshold, P_{th}, is set to 0.9, and the maximum flashes prior to selecting a character, t_{max}, is set to 120 flashes for all paradigms. A 72-character grid size is used for all numerical simulations. Selection accuracy and the average number of flashes prior to character selection, denoted as the expected stopping time (EST), are estimated as a function of d', with performance results averaged over 1500 iterations.

3.3.1 Numerical Simulation with Ideal Conditions

Assuming no physiological or speller constraints, a simulation using synthetic data is performed to compare the performance of three stimulus presentation paradigms: RC random, CB random, and the proposed greedy EDG adaptive paradigm. The RC and CB random paradigms select flash groups from RC and CB codebooks using a uniformly random draw. A comparison between flash groups using the ideal greedy EDG adaptive paradigm and RC random paradigms is shown in Figure 3.2. The adaptive paradigm presents large subsets of characters initially, and then will prune the presented characters until the target character is determined. The RC paradigm, however, will randomly present characters regardless of the target character probabilities.

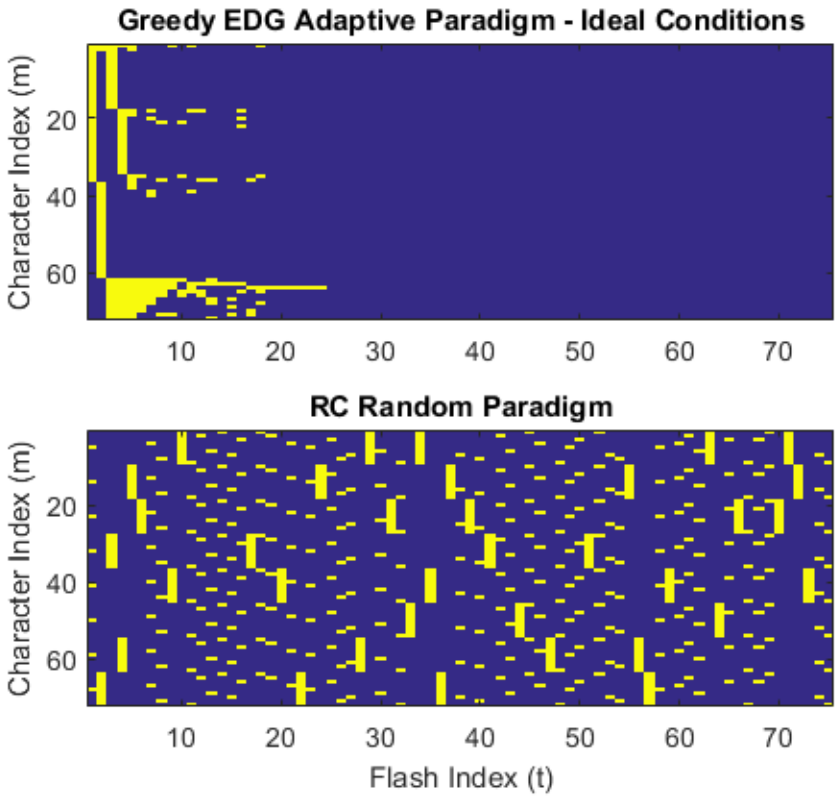

Figure 3.2 – Example of stimulus presentations for a single iteration of the numerical simulations for the greedy EDG adaptive paradigm (top) and the RC random paradigm (bottom). Yellow indicates that a character, m, at a flash index, t, is presented. No stimuli are presented after $t = 25$ when using the EDG adaptive paradigm as the character is already selected in the simulation.

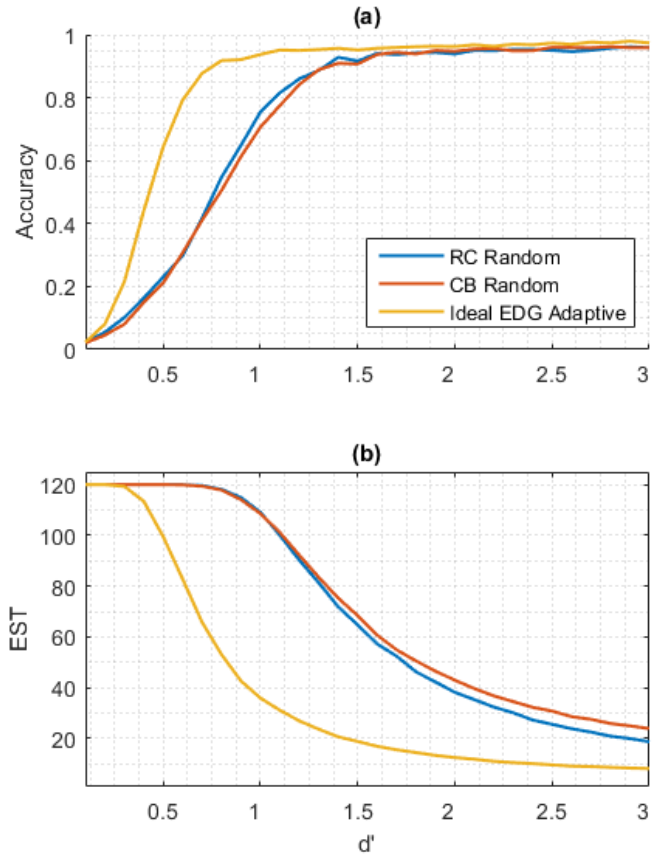

Figure 3.3 – Simulation results to compare performance of RC, CB, and greedy EDG adaptive paradigm as a function of d' in ideal conditions. (a) shows the accuracy as a function of d', and (b) shows the EST as a function of d'. Across all d', the ideal greedy EDG adaptive paradigm has higher accuracy and lower EST than both the RC and CB paradigms.

The accuracy and EST of the simulation results as a function of d' are shown in Figure 3.3. Figure 3.3 (a) and (b) shows the accuracy and EST, respectively, of the RC, CB, and ideal greedy EDG adaptive paradigms as a function of d'. Across all d', the greedy EDG adaptive paradigm has a significantly higher accuracy and lower EST than both RC

and CB paradigms. For example, at a $d' = 0.5$, the ideal greedy EDG adaptive has an accuracy ~45% higher than both RC and CB paradigms; at a $d' = 1$, the ideal greedy EDG requires 75 fewer flashes than both CB and RC paradigms. This suggests that the adaptive paradigm based on the EDG metric can in theory yield a higher accuracy and lower EST than the random RC and CB paradigms.

3.3.2 Numerical Simulations with Realistic Conditions

The results for the ideal greedy EDG adaptive paradigm shown in section 3.3.1 are promising, however, they ignore various physiological constraints (see section 2.3) as well as online system constraints. In this section, simulation results are shown for three different independent restrictions; observation delay (OD) 3.3.2.1, maximum characters presented in a flash group 3.3.2.2, and minimum TTI 3.3.2.3. Furthermore, results are shown for a simulation that jointly considers the OD, character presentation, and TTI restrictions in section 3.3.2.4.

3.3.2.1 Observation Delay Constraints

In ideal simulations, it was assumed that the classifier observation, z_t, was available prior to determining $\mathcal{F}_{t+1}^{\ddagger}$. However, this is typically not the case during online BCI implementation. Following each stimulus presentation, a time window of EEG data is analyzed to yield a classifier score. Consequently, there is an observation delay between the presentation of $\mathcal{F}_{t+1}$ and its associated classifier score, z_t, as illustrated in Figure 3.4. The window for which the EEG data is being collected over includes multiple stimulus

presentations. For example, in Figure 3.4 it can be observed that $\mathcal{F}_{t+1}, \mathcal{F}_{t+2}, \ldots, \mathcal{F}_{t+6}$ will have already been presented prior to observing z_t. OD is defined as the observation delay, and δ is the number of additional stimulus presentations prior to being able to compute z_t.

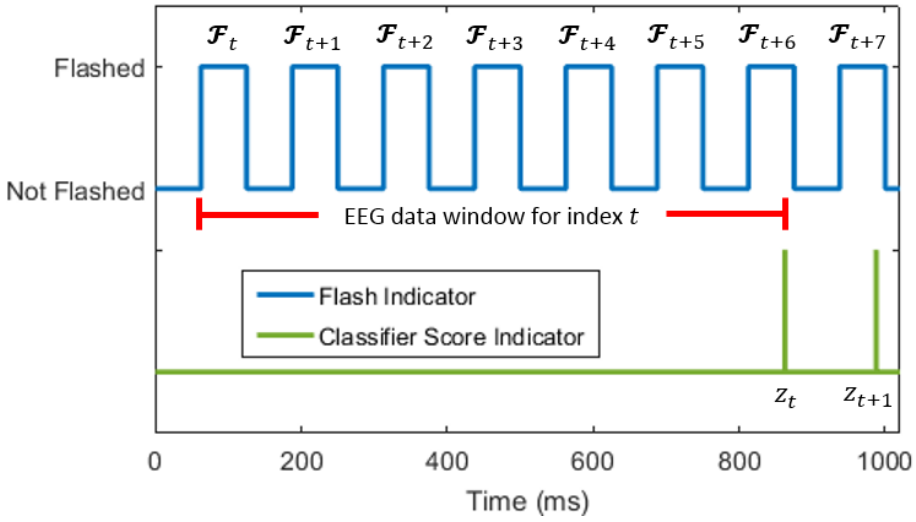

Figure 3.4 – Diagram showing an example of the time between a stimulus presentation and an EEG observation for the stimulus presentation. In this example, $\mathcal{F}_{t+1}, \mathcal{F}_{t+2}, \ldots, \mathcal{F}_{t+6}$ will have already been presented prior to observing observation z_t because z_t requires EEG data over a window longer than the inter-stimulus interval.

Incorporating the observation delay modifies the optimization function for choosing a stimulus presentation to the following:

$$\mathcal{F}^{\ddagger}_{t+1+\delta} = \arg\max_{\mathcal{F}^{*}_{t+1+\delta}} \Delta D_{KL}(\mathcal{F}^{*}_{t+1+\delta}), \forall \mathcal{F}^{*}_{t+1+\delta} \in \boldsymbol{\Omega} \tag{3.22}$$

This new optimization function requires calculating $\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta})$. Equation (3.9) can be rewritten to calculate $\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta})$ instead of $\Delta D_{KL}(\mathcal{F}^*_{t+1})$ as shown below:

$$\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta}) = E_{z^*_{t+1}, z^*_{t+\delta}, \dots, z^*_{t+1+\delta}|Z_t}[D_{KL}(\mathcal{P}^*_{t+1+\delta} \| \mathcal{P}_t)|\mathcal{F}^*_{t+1}] \tag{3.23}$$

Equation (3.23) can then be expanded to:

$$\begin{aligned} &\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta}) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}\sum_{m=1}^{M} \wp^*_{t+1+\delta,m} \log\left(\frac{\wp^*_{t+1+\delta,m}}{\wp_{t,m}}\right) p(z^*_{t+1}, z^*_{t+2}, \dots, z^*_{t+1+\delta}|Z_t) dz^*_{t+1} dz^*_{t+2} \dots dz^*_{t+1+\delta} \end{aligned} \tag{3.24}$$

The multi-variate integral in (3.24) used to compute $\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta})$ is computationally expensive as it requires $\delta + 1$ integrals with infinite bounds. Therefore, a simplification is applied such that:

$$\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta}) \approx E_{z^*_{t+1+\delta}|Z_t}[D_{KL}(\mathcal{P}^*_{t+1+\delta} \| \widehat{\mathcal{P}_{t+\delta}}) | \mathcal{F}^*_{t+1+\delta}] \tag{3.25}$$

where $\widehat{\mathcal{P}_{t+\delta}}$ is an estimate of $\mathcal{P}^*_{t+\delta}$. $\widehat{\mathcal{P}_{t+\delta}}$ is the expectation of $\mathcal{P}^*_{t+\delta}$ as shown in the equation below:

$$\widehat{\mathcal{P}_{t+\delta}} = E_{z^*_{t+1}, z^*_{t+2}, \ldots, z^*_{t+\delta} | Z_t}[\mathcal{P}^*_{t+\delta}] = \mathcal{P}_t \tag{3.26}$$

With this simplification, the pre-computed ΔD_{KL} curve described in section 3.1 can be used to compute $\Delta D_{KL}(\mathcal{F}^*_{t+1+\delta})$, and the optimization function, (3.25), can be used to choose $\mathcal{F}^*_{t+1+\delta}$ with a greedy search.

The same simulation is performed as with the greedy EDG adaptive paradigm in section 3.3.1, however, (3.25) is used to update flash group $\mathcal{F}^*_{t+1+\delta}$ instead of updating $\mathcal{F}^*_{t+1}$. Simulations are performed with OD set to 3, 5, and 7. The accuracy and EST of the simulation, with various OD restrictions, as a function of d' are shown in Figure 3.6.

The results in Figure 3.6 show that OD has little to no effect on the accuracy of the speller using the adaptive paradigm; at OD = 3, 5, and 7, the accuracy is approximately equal to that of the ideal adaptive paradigm. As OD increases, the EST of the speller increases across all but low d' (< 0.5).

3.3.2.2 Character Presentation Limit

Presenting too many characters simultaneously can lead to more adjacency distractions, and to an increase in cognitive load. None of these physiological effects are desirable as they can all lead to a decrease in speller accuracy. However, a decrease in the number of characters presented, much like any other restriction on $\boldsymbol{\Omega}$, can decrease the EDG in $\mathcal{F}^{\ddagger}_{t+1}$ as the search space shrinks.

The goal of the simulation is to analyze the effect of various character presentation restrictions on the performance of the greedy EDG adaptive paradigm. A simulation is performed such that there is a limit set on the number of characters that can be presented during a single flash. Formally, the restriction on $\mathbf{\Omega}$ results in $\mathbf{\Omega}_\alpha$ as shown below:

$$\forall \mathcal{F}_{t+1}^{*} \in \mathbf{\Omega}_{\alpha}, \left[\sum_{\forall m: \phi_{t+1,m}^{*}=1} 1 \right] \leq \alpha_{\max} \tag{3.27}$$

where $\alpha_{\max}$ is the maximum number of characters that can be in a candidate flash group, $\mathcal{F}_{t+1}^{*}$. $\mathbf{\Omega}_{\alpha}$ is flash group search space with the character presentation limit.

The same simulation is performed as in section 3.3.1, except the flash group search space is limited to $\mathbf{\Omega}_\alpha$ as shown in (3.27). Simulations are performed with $\alpha_{\max}$ set to 1, 5, 9, and 13. During stimulus selection, the greedy search is terminated after $\alpha_{\max}$ has been reached. The accuracy and EST over a range of d' with a flash group search space of $\mathbf{\Omega}_{\alpha}$ (with varying $\alpha_{\max}$) is shown in Figure 3.5; the accuracy and EST of the RC and CB paradigms are also shown.

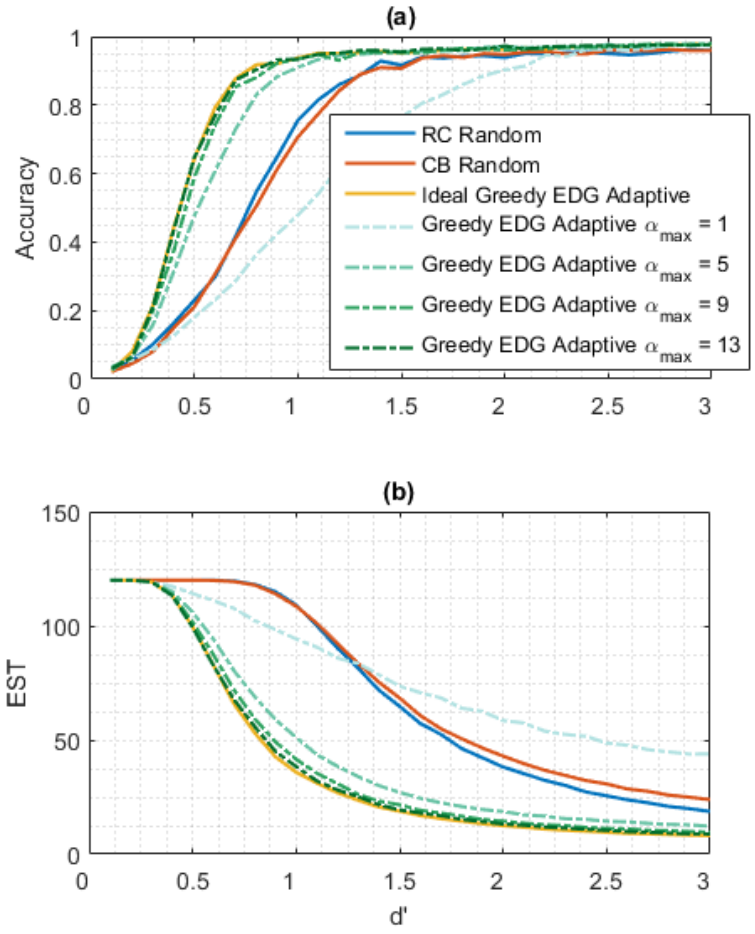

Figure 3.5 - Simulation results that provide a comparison between RC, CB, and various greedy EDG adaptive paradigms with character presentation limits. (a) shows the accuracy and (b) shows the EST of a stimulus presentation paradigm as a function of d'. When α_{max}, the maximum number of characters presented concurrently, is set to 1 (i.e., only 1 character can be presented at a time), RC and CB yield a higher accuracy across all d'. Greedy EDG adaptive with $\alpha_{max} = 1$ requires fewer presentations up to a $d' \approx 1.3$. When $d' > 1.3$, both RC and CB paradigms require fewer presentations than the greedy EDG adaptive paradigm with $\alpha_{max} = 1$. As α_{max} increases, the accuracy increases and EST decreases.

The results provided in Figure 3.5 show that only presenting a single character at a flash index for a 72-character grid is insufficient as the accuracy of greedy EDG adaptive paradigm with $\alpha_{max} = 1$ is consistently lower than RC and CB paradigms at all d'. When α_{max} is increased to 5, the accuracy and EST are higher and lower, respectively, for the greedy EDG adaptive paradigm than the both RC and CB paradigms. As α_{max} increases, the performance of the adaptive paradigm improves, however, only a slight increase in accuracy and a slight decrease in EST is shown between greedy EDG adaptive paradigms with $\alpha_{max} = 9$ and $\alpha_{max} = 13$.

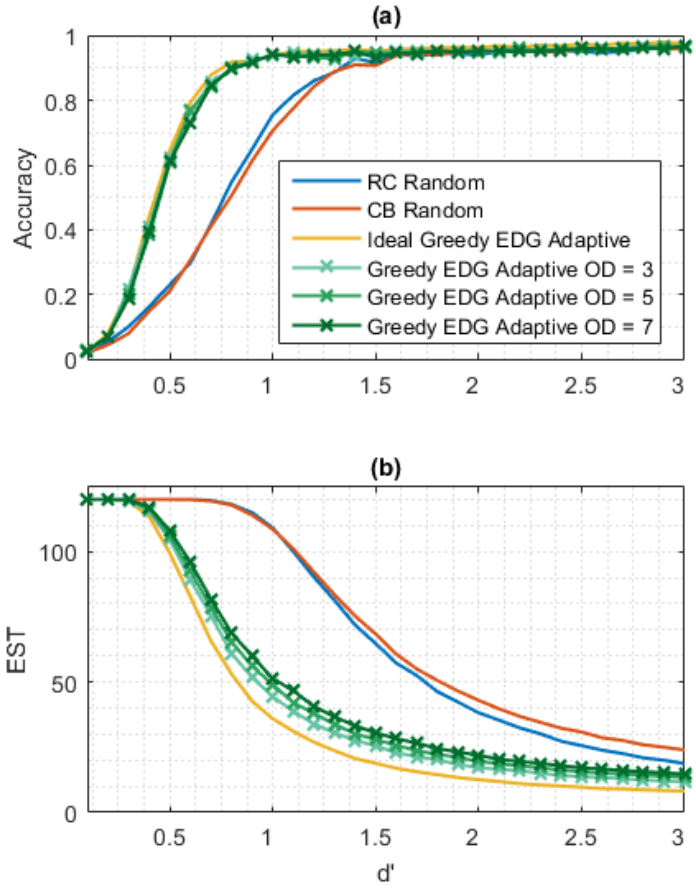

Figure 3.6 - Simulation results that provide a comparison between RC, CB, and various greedy EDG adaptive paradigms with various observation delays (ODs). (a) shows the accuracy and (b) shows the EST of a stimulus presentation paradigm as a function of d'. An increase of OD has little impact on the accuracy, however, it does increase the EST of the speller slightly. At an OD of 3,5, and 7, the greedy EDG adaptive paradigm has a higher accuracy and lower EST than the RC and CB random paradigms.

3.3.2.3 Refractory Effects Constraints

Refractory effects are important to consider when designing a stimulus presentation paradigm (section 2.3.1). The discriminability of a target observation

depends on the time interval between target stimulus presentations as shown in Figure 2.9. In this section, a minimum TTI is imposed in between a character's presentation. The optimization function for the flash group selection remains the same, however, the flash group space is restricted to $\boldsymbol{\Omega}_{TTI}$, defined in (3.28).

$$\forall \mathcal{F}_{t+1}^{*} \in \boldsymbol{\Omega}_{TTI}: \left[\forall m: \left[\mathbb{f}_{t+1,m}^{*} + \mathbb{f}_{t,m} + \cdots + \mathbb{f}_{t+2-TTI_{min},m} \leq 1\right]\right] \tag{3.28}$$

where TTI_{min} is the minimum TTI imposed in between a character's presentation. Equation (3.28) states that if any character, m, is presented between t and $t + 1 - TTI_{min}$, that character cannot be elicited at $t + 1$. Note that setting $TTI_{min} = 1$ yields $\boldsymbol{\Omega}_{\text{TTI}} = \boldsymbol{\Omega}$ (i.e., no TTI restrictions).

The same simulation is performed as with ideal conditions in section 3.3.1, however, the flash group search space in the simulation is restricted to $\boldsymbol{\Omega}_{TTI}$. TTI_{min} is set to 2, 3, 4, and 5. The results presented in Figure 3.7 demonstrate that increasing the TTI_{min} decreases the accuracy and increases the EST of the speller using the greedy EDG adaptive paradigm. With significant TTI_{min} restrictions such as $TTI_{min} = 4$ or 5, the greedy EDG adaptive paradigm still yields a higher accuracy in the range of $0.5 < d' < 1.5$ and lower EST across $0.5 < d' < 2.5$.

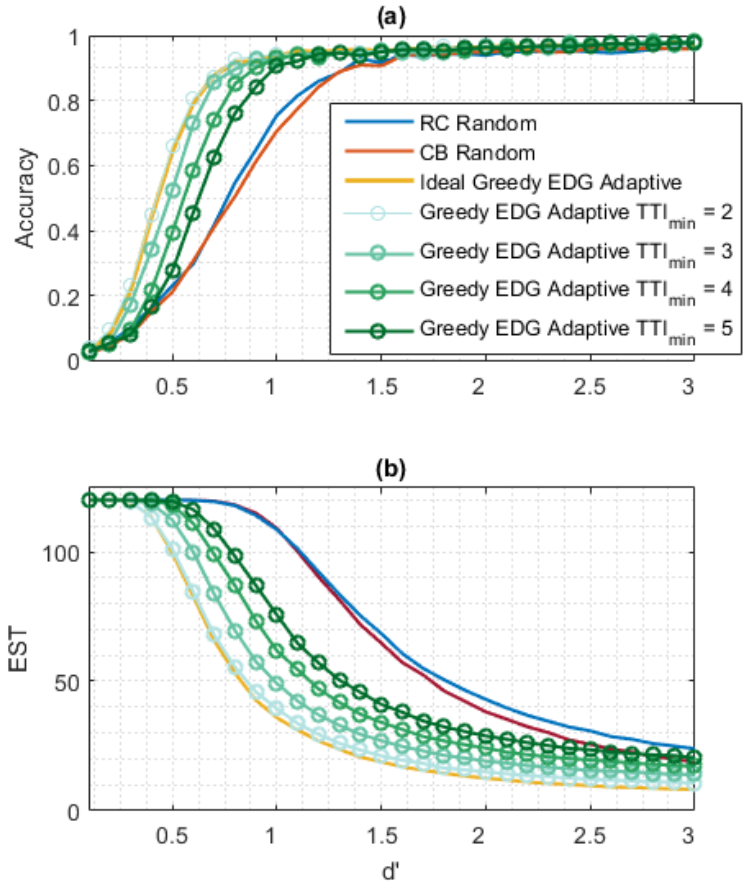

Figure 3.7 - Simulation results of comparison between RC, CB, and various greedy EDG adaptive paradigms with various TTI_{min} restrictions. (a) shows the accuracy and (b) shows the EST of a stimulus presentation paradigm as a function of d'. As TTI_{min} increases, the accuracy decreases and the EST increases across all d'.

3.3.2.4 Combined Constraints

In the prior sections, various individual restrictions such as TTI, OD, and character presentation limits are discussed, and simulations for those individual restrictions are performed. In this section, a simulation is performed for the greedy EDG adaptive

paradigm that jointly enforces TTI, OD, and character presentation limit restrictions and compares its performance to the CB and RC paradigms. The search space for the combined adaptive paradigm is the intersection of the search space with TTI and character limit constraints: $\boldsymbol{\Omega}_C = \boldsymbol{\Omega}_\alpha \cap \boldsymbol{\Omega}_{TTI}$, where $\boldsymbol{\Omega}_C$ is the new combined constraint search space. The EDG is estimated as described in section 3.3.2.1 and flash group $\mathcal{F}^*_{t+1+\delta}$ after each flash index t to account for OD. The notation to describe the restrictions in the greedy EDG adaptive paradigm will be $COMB(\alpha_{max}, OD, TTI_{min})$.

The simulation described in section 3.3 is performed, however, combined restrictions are used. The TTI_{min} is set to 3 as a TTI of 1 or 2 yield low detectibly as shown in Figure 2.9. The OD was set to 6, as that is the approximate OD expected in online simulations. Finally, the character limit, α_{max} is set to 9 as that is the maximum number of characters presented in the RC paradigm. The results of the simulation comparing the performance of the adaptive paradigm with the combined restrictions and independent restrictions are shown in Figure 3.8.

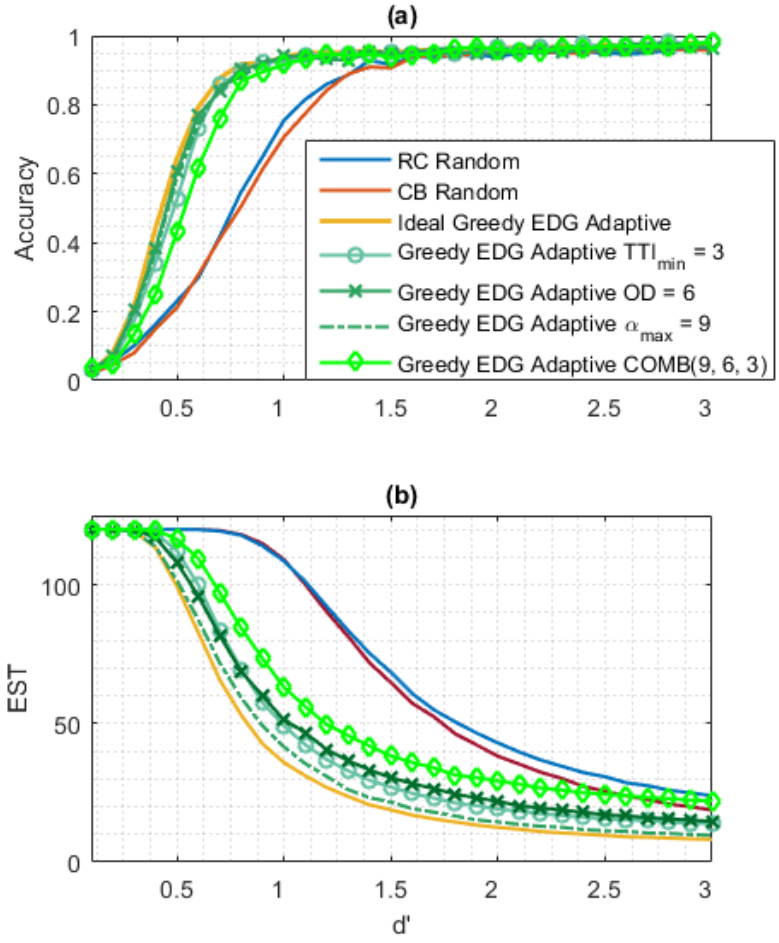

Figure 3.8 - Simulation results of comparison between RC, CB, and greedy EDG adaptive paradigms with independent restrictions and combined restrictions. (a) shows the accuracy and (b) shows the EST of a stimulus presentation paradigm as a function of d'. As TTI_{min} increases, the accuracy decreases and the EST increases across all d'.

Though the greedy EDG adaptive COMB(9, 6, 3) paradigm yields slightly worse performance than can be achieved with the independent constraints, the COMB(9, 6, 3) paradigm has significantly higher accuracy and lower EST than the RC and CB random

paradigms across most d'; specifically, the accuracy d'<1.5 is higher and the EST is lower at $0.5 < d' < 2.5$.

The results of the greedy EDG adaptive paradigm with the combined constraints are promising, however, this approach leads to a speller that presents the target character at a fixed TTI_{min} repetitively. Figure 3.9 shows the distribution of the TTI of a target presentation for $d' = 0.5$, 1.0, and 1.5 for $TTI_{MIN} = 1$ and 3 with a COMB(9,6,TTI_{min}) constraint. Regardless of d', most of the target presentations (≈80%) occur with a TTI of TTI_{min}. As shown in Figure 3.11, this leads to highly correlated flash groups, which will decrease the amount of information gained after each presentation. Furthermore, most P300 speller paradigms are randomized or pseudo-randomized [23], [25], [37], [38], [85], [86], [102] as there is a concern that the user will recognize a pattern in a non-randomized paradigm, and thus P300 will no longer be elicited when a target is presented.

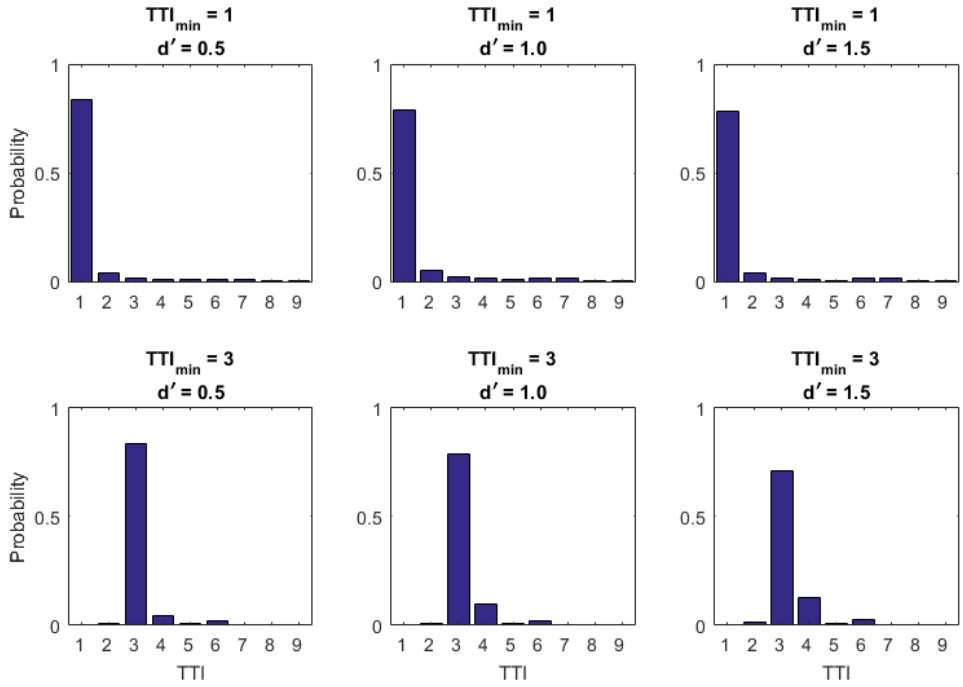

Figure 3.9 – Distribution of TTI for greedy EDG adaptive paradigm with COMB(9,6,TTI_{min}) constraints. In the top row TTI_{min} = 1, and in the bottom row TTI_{min} = 3. The columns correspond to a d' of 0.5, 1.0, and 1.5 from left to right. Regardless of d', most of the target presentations occur at TTI_{min}.

3.3.3 Stochastic TTI Constraints

A stochastic TTI_{min} is proposed to decrease the correlation of the flash groups as well as to decrease the amount of repetitive TTI presentations. Instead of setting a fixed TTI_{min}, the TTI_{min} is ddrawn from a discrete distribution. At each flash index, t, a TTI_{min} is independently drawn, from a discrete distribution, for all presented characters. This occurs at all flash indices, and for all characters. Due to the independent draw, characters

presented at the same t can have different TTI_{min}; furthermore, the same characters at different t can have a different TTI_{min}.

The distribution used in this work for the stochastic TTI constraint is the TTI distribution from the performance based paradigm (PBP) [37]. The PBP paradigm is used as it has shown an improvement in accuracy and EST compared to the random RC and CB paradigms. The PBP TTI distribution is shown in Figure 3.10 below.

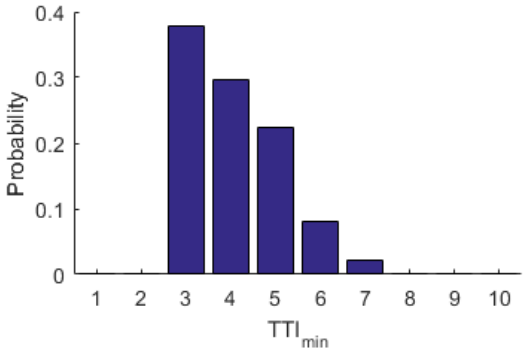

Figure 3.10 – Distribution used for stochastic TTI constraints for TTI_{min} in this work, the PBP TTI distribution.

A numerical simulation is performed as a function of d' to compare RC and CB random, ideal greedy EDG adaptive, greedy EDG adaptive COMB(9, 6, 3), and greedy EDG adaptive COMB(9, 6, PBP). Figure 3.13 shows the TTI distributions obtained at d' = 0.5, 1.0, and 1.5 for $TTI_{min} \sim PBP$. As with a fixed TTI_{min}, the stochastic TTI_{min} yields a TTI distribution similar to the respective TTI_{min} distribution. This result shows that TTI distribution of the greedy EDG adaptive paradigm can be approximately set with the

TTI_{min} distribution. This may be helpful if the user has a preference to how often they want the target character to be presented as well as what kind of uncertainty they prefer in the TTI.

Examples of stimulus presentations for the ideal greedy EDG adaptive paradigm, greedy EDG adaptive paradigm with COMB(9,6,3) restrictions and greedy EDG adaptive paradigm with COMB(9,6,PBP) restrictions are shown in Figure 3.11. The ideal adaptive paradigm presents some characters repetitively with a TTI of 1. The greedy EDG adaptive paradigm with COMB(9,6,3) restrictions does not present characters at a TTI of 1 or 2, however, there is a high correlation between flash groups. Furthermore, many characters are repetitively presented at a TTI of 3, which may eliminate the ERP as the target presentation will no longer be a surprise to the user. The greedy adaptive paradigm with COMB(9,6,PBP) restrictions does not present characters at a TTI of 1 or 2, and thus decreases the correlation between the flash groups that are presented.

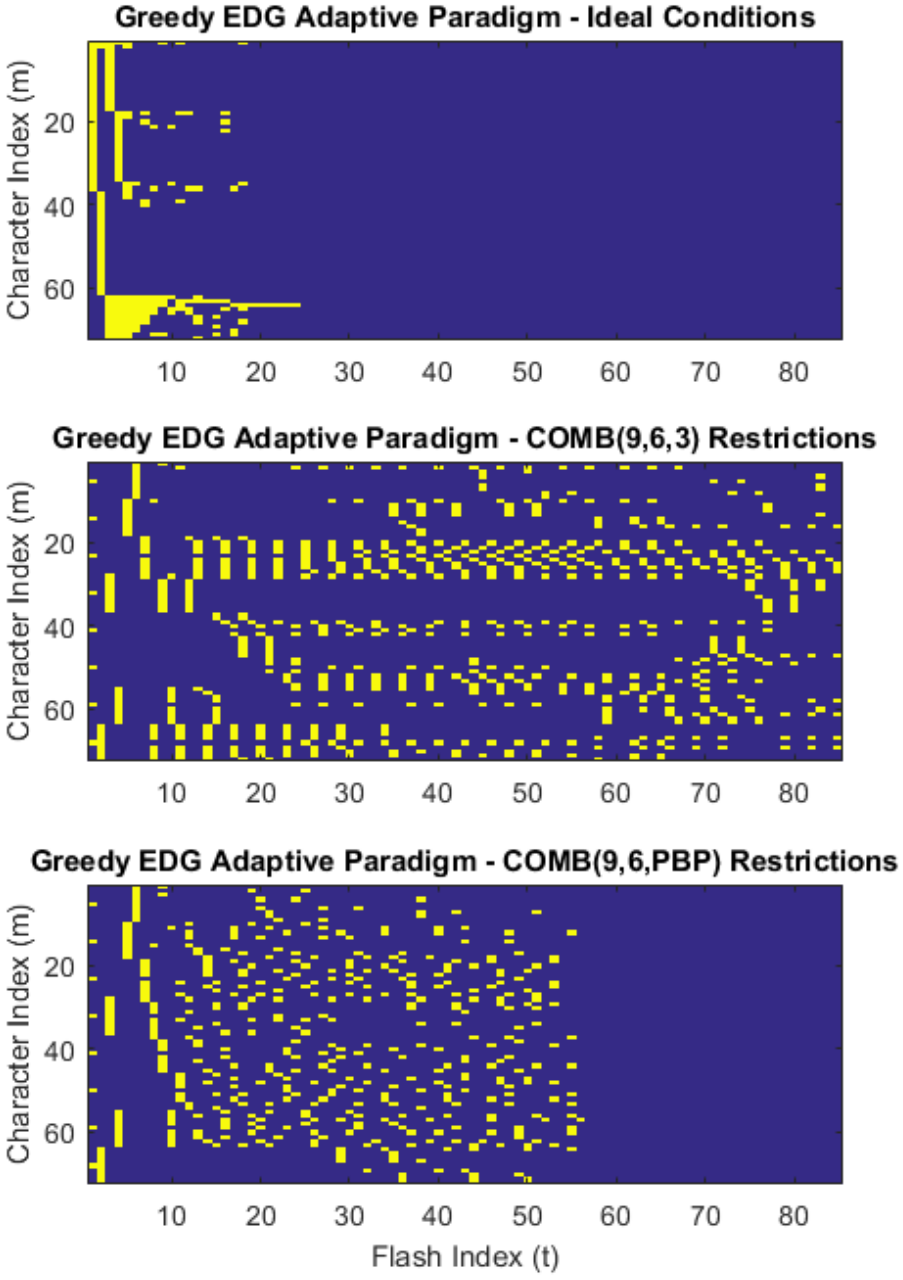

Figure 3.11 - Examples of stimulus presentations for the ideal greedy EDG adaptive paradigm (top), greedy EDG adaptive paradigm with COMB(9,6,3) constraints (middle) and greedy EDG adaptive paradigm with COMB(9,6,PBP) constraints (top). Yellow indicates that a character *m* is presented at flash index *t*.

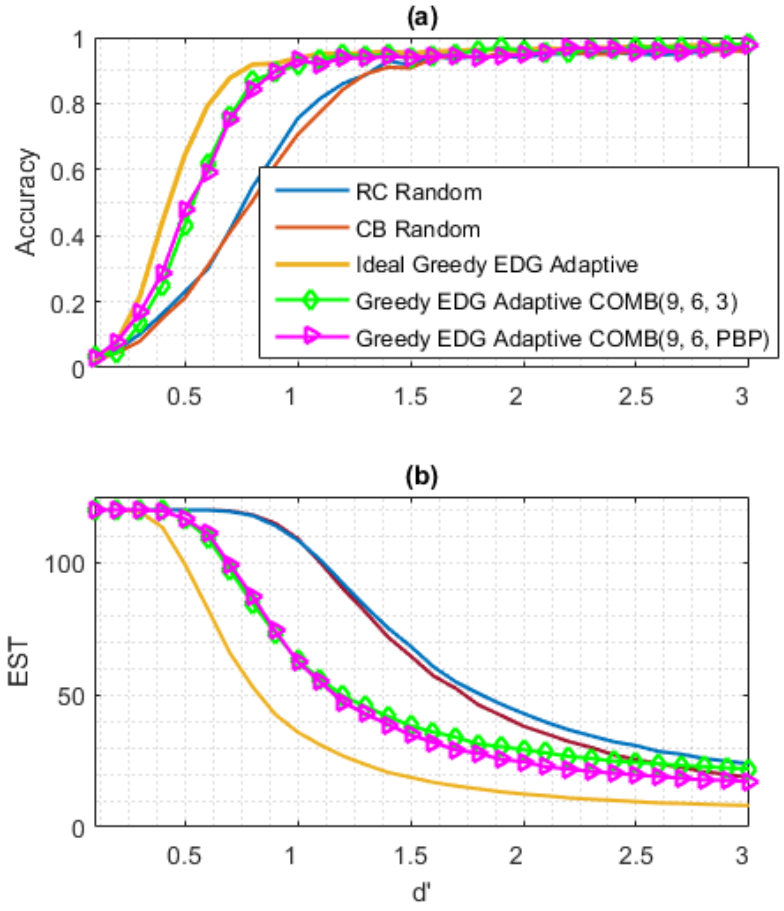

Figure 3.12 - Simulation results of comparison between RC, CB, and greedy EDG combined restrictions with fixed and stochastic $TTI_{min} \sim PBP$. (a) shows the accuracy and (b) shows the EST of a stimulus presentation paradigm as a function of d'. At a d' of 0.1 to 0.3, there is a slight increase in accuracy when using a stochastic TTI_{min}. At a $d' > 1$, there is a decrease in EST when using a stochastic $TTI_{min} \sim PBP$.

Figure 3.12 shows the accuracy and EST of the simulations as a function of d'. The adaptive paradigm that used the stochastic TTI_{min} had a lower EST than the fixed TTI_{min} at a $d' > 1$. Otherwise, the adaptive paradigm with the fixed TTI_{min} = 3 constraint yields similar accuracy and EST to the adaptive paradigms with the stochastic TTI_{min}.

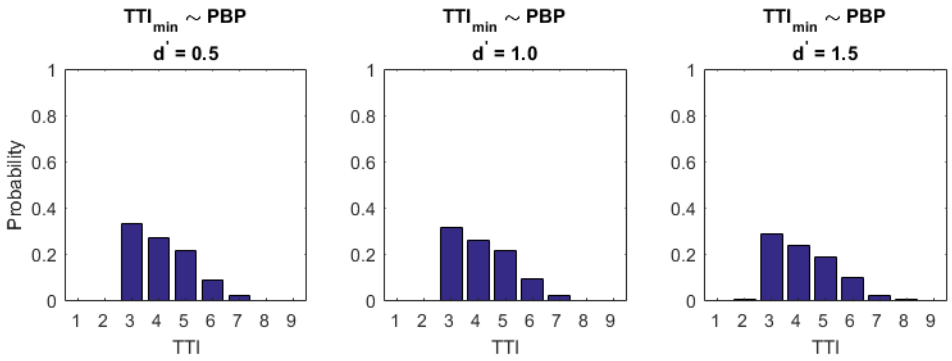

Figure 3.13 - Distribution of TTI for greedy EDG adaptive paradigm with a COMB(9,6,TTI_{min}) restriction. The columns correspond to a d' of 0.5, 1.0, and 1.5 from left to right. Regardless of d', the distribution of TTI is similar to the PBP TTI_{min} distribution.

The PBP TTI distribution was used for TTI_{min} in this work to show that a stochastic TTI constraint can be beneficial in simulations, as well as likely to be beneficial in practice. It is important to note, however, that the work in this section does not suggest that the PBP TTI distribution is the optimal distributions for TTI_{min} for the greedy EDG adaptive paradigm.

3.4 Online Data Collection

The simulations presented above demonstrated an improvement in accuracy and EST when using the greedy EDG adaptive paradigm, with realistic restrictions, compared the RC and CB paradigms. Though the adaptive paradigm shows improvement results in performance improvements in simulation, it remains unclear if it will continue to outperform a random paradigm in an online setting. In this section, an online data

collection is performed to compare both the greedy EDG adaptive paradigm and the CB paradigm.

3.4.1 Methodology

For the online experiment, eight able-bodied participants were recruited from Duke University. All data was collected in a sound-proof booth at Duke University. The participants gave informed consent, and the study was approved by the Duke University Institutional Review Board (IRB). Two paradigms were compared: the CB random and greedy EDG adaptive paradigm with $\alpha_{max} = 9$, and $TTI_{min} \sim PBP$ restrictions. The maximum number of presentations in a single trial was set to 144, and the stopping threshold, $P_{thresh} = 0.9$.

One session of data collection was completed for each participant, and each session consisted of a training task (see section 3.4.1.1) followed by a testing task for one paradigm, and then a training task followed by a testing task for the second paradigm. The order of the paradigms used were randomized to avoid biases that may arise because of the temporal non-stationarity typically found in EEG data [103]. Three, six-letter, words were spelled during a training task and five, six-letter words were spelled during a testing task. The words being spelled by the participants were randomly chosen from a subset of English words from the English Lexicon Project [104].

3.4.1.1 Training task for Greedy EDG Adaptive Paradigm

As with any supervised learning algorithm, it is crucial that the training data is drawn from a similar underlying distribution as the testing data. To ensure that the EEG data between training and testing tasks are similar, the stimulus presentations between the training and testing tasks must also be similar.

During the training task for a random paradigm, flash groups are randomly selected and presented to the user as they would be during the testing task. The training task for an adaptive paradigm, however, is less straightforward. The greedy EDG adaptive paradigm chooses stimuli to present based on the target character probabilities obtained by using a trained classifier and conditional classifier score PDFs. During the training task, however, the classifier and conditional classifier score PDFs are not yet learned.

Therefore, in this work, synthetic classifier scores were drawn using predefined target and non-target conditional classifier score PDFs. The greedy EDG adaptive paradigm was then used to select stimuli to present based on the drawn classifier scores and estimated likelihood. The conditional classifier score PDFS used during the training task were: $\ell 0 \sim \mathcal{N}(0,1)$ and $\ell 1 \sim \mathcal{N}(1,1)$, yielding a $d' = 1$.

3.4.2 Online Results

The results from the online data collection are shown in Figure 3.14. Figure 3.14 (a) shows the accuracy, (b) shows the average number of presentations, and (c) shows the

bit-rate for eight subjects sorted by accuracy of the CB paradigm. The CB paradigm yields a higher accuracy for six of the eights subjects and a higher average accuracy across all subjects. The average number of presentations is lower using the greedy EDG adaptive paradigms versus the CB paradigm for seven of eight subjects. The bit-rate using the greedy EDG adaptive paradigm is higher for six of the eight subjects than the CB paradigm. Using the 2-sided Wilcoxon signed-rank test [105], only the decrease in the average number of presentations is statistically significant.

With the greedy EDG adaptive paradigm, many of the subjects did not reach 90% accuracy and had an EST far fewer than the maximum number of presentations (144). This indicates that the target character update model does not fit the data collected during the testing phase of the greedy EDG adaptive paradigm experiment. Anecdotally, many subjects complained about a high number of character presentations that were adjacent to the target character; this likely led to adjacency distractions that were not explicitly modeled in the numerical or MC simulations.

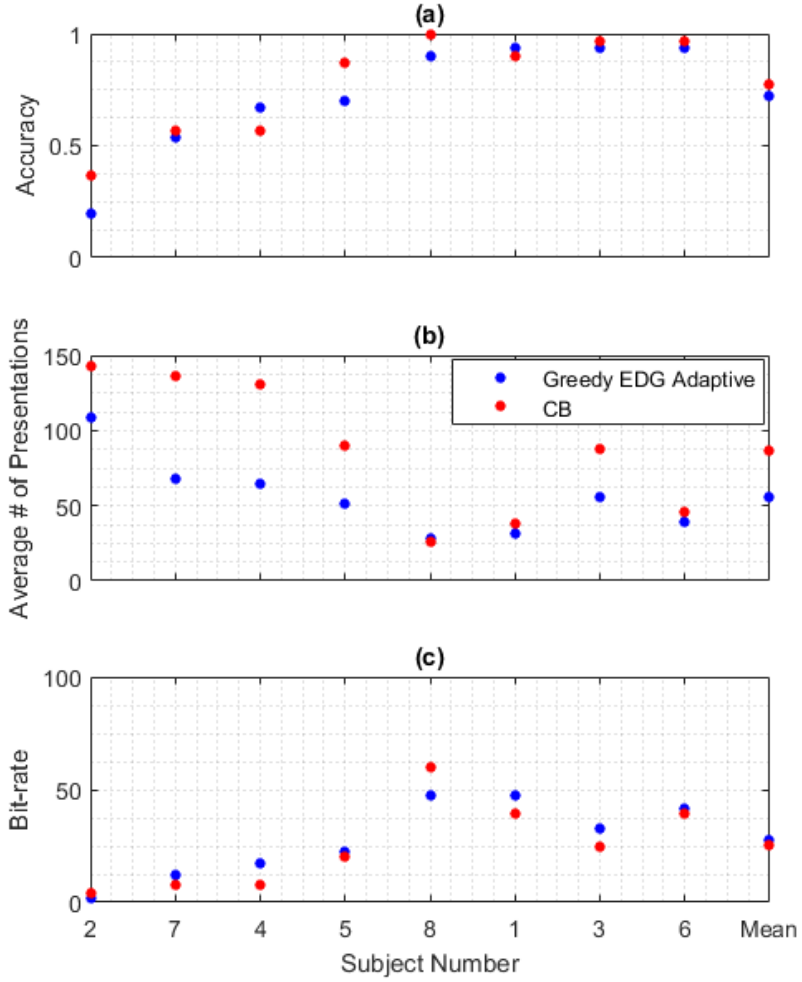

Figure 3.14 – Results from the online data collection to compare CB and greedy EDG adaptive paradigms. (a) shows the accuracy, (b) shows the average number of presentations, and (c) shows the bit-rate for eight subjects sorted by accuracy of the CB paradigm. The mean of the respective performance metrics are shown in the last column.

3.5 Discussion

In this chapter, a framework for a greedy adaptive stimulus presentation paradigm using EDG was introduced. In designing the greedy EDG adaptive paradigm, physiological phenomena were considered by enforcing restrictions on TTI and character presentations. System constraints were also considered by modeling OD in the numerical simulations. The numerical simulations showed that even with restrictions set on the greedy EDG adaptive paradigm, it has the potential to yield higher accuracy, and fewer presentations than the random RC and CB paradigm. It was shown further that a stochastic TTI_{min} constraint can minimize the correlation between flash groups, reducing the number of presentations required for users with a $d' > 1$. In this work only a single distribution for stochastic TTI_{min} was considered; in future work, other distributions should be considered to select an optimal distribution for a stochastic TTI_{min}. The results for the numerical simulations from this chapter were published in [106].

The online results showed the potential improvement in the EDG adaptive paradigm as there were far fewer presentations required to spell a character than when using the CB random paradigm. Furthermore, the average bit-rate for the EDG adaptive paradigm was higher than that of the CB random paradigm. The online results, however, also showed that the accuracy tends to decrease for the greedy EDG adaptive paradigm compared to the CB random paradigm.

One potential reason for the low accuracy when using the EDG adaptive paradigm is that the training data was collected under a slightly different paradigm than the testing data. The training data used synthetic data to choose the stimulus presentation, ignoring potential physiological phenomena. This could cause the training and testing data to be drawn from different underlying distributions. This difference can cause the trained classifier score and conditional classifier score PDFs to be sub-optimal for determining the true target character probabilities in the testing phase. It is possible that this issue can be resolved by modifying the training procedure when using an adaptive paradigm to include an intermediate training phase. For this proposed intermediate phase, the trained classifier, conditional classifier score PDFs, and EDG curves learned from the initial training phase can be used for stimulus selection using EEG data collected during the intermediate training phase. The data collected during the intermediate training phase can then be used to re-train a classifier, learn new conditional classifier score PDFs, and learn a new EDG curves to be used during the testing phase.

Another potential reason for the low accuracy in the online greedy EDG adaptive paradigm is the high number of character presentations adjacent to the target character, as noted by several study participants. These adjacency errors can erroneously elicit a P300, causing an incorrect character selection. In future work, adjacency distractions models should be developed and incorporated in the target character probability update

equation. If adjacency distractions can be modelled accurately in the Bayesian update, the accuracy is more likely to reach the threshold if the presentation limit is not reached.

4 Exploiting Target-to-target Interval Information to Mitigate Refractory Effects

The negative impact of refractory effects on P300 speller performance has been demonstrated in previous BCI literature [42], [43], [82], [88], [89], [107]. In the literature, refractory effects are typically quantified by the time interval between target character presentations or the target-to-target interval (TTI). Classification performance has been shown to increase with increasing TTI [42], [43], [82], [88], [89], [107], as illustrated in Figure 2.9. However, most P300 speller implementations do not incorporate the temporal relationship between TTI and classifier scores as the classifier is typically trained by aggregating the EEG responses of all target stimulus presentations into a single class label.

Work in [84], [107] showed that refractory effects can be mitigated by specifying multiple target classes based on the state of the previous stimulus event presentation. However, the approaches in [84], [107] only consider short-term refractory effects as the TTI information provided is limited to the previous stimulus event, i.e. a TTI of 1.

To model longer-term refractory effects, it is proposed that a longer history of previous stimulus events be considered during classification: this can be achieved by incorporating information on TTI values greater than 1. It also hypothesized that TTI information can also be exploited to enhance the data-driven stimulus selection, presented in Chapter 3, by biasing the selection process towards flash groups that will minimize target character presentations with TTIs.

This chapter presents approaches to incorporate TTI information in the P300 speller to improve BCI performance with the Bayesian algorithm. Results from P300 speller simulations are presented using stimulus presentation paradigms that are susceptible to refractory effects to demonstrate the utility of incorporating TTI information to mitigate these effects.

4.1 Incorporating TTI-information in the Bayesian Algorithm

In the standard Bayesian algorithm proposed in Throckmorton et al. [72], no TTI information is used when updating the target character probabilities (i.e., posterior probability). In section 4.1.1, the Bayesian update equation from Throckmorton et al. is modified such that instead of using a single target classifier score PDF, multiple target classifier score PDFs, conditioned on the TTI of the character presentation, are used.

In the previous implementation of the adaptive paradigm proposed in Chapter 3, a minimum TTI was imposed to limit refractory effects. However, imposing a TTI_{min} is an ad-hoc restriction that does not explicitly model the refractory effects present in the EEG data. The greedy EDG adaptive stimulus paradigm will be modified in section 4.1.2 to use the Bayesian update equation that incorporates TTI information.

4.1.1 Character Probability Updates

Assuming TTI-specific classes for target classifier scores, the Bayesian update equation used to update the target character probability (3.5) is modified to the following:

$$\mathcal{p}_{t,m} = \frac{\ell_{t,m}(z_t, i_{t,m})\mathcal{p}_{t-1,m}}{\sum_{c=1}^{M} \ell_{t,m}(z_t, i_{t,m})\mathcal{p}_{t-1,c}} \tag{4.1}$$

$$\ell_{t,m}(z_t, i_{t,m}) = \begin{cases} \ell 1(z_t, i_{t,m}), & \mathcal{f}_{t,m} = 1 \\ \ell 0(z_t), & \mathcal{f}_{t,m} = 0 \end{cases} \tag{4.2}$$

where $\mathcal{p}_{t-1,m}$ and $\mathcal{p}_{t,m}$ are the prior and posterior probabilities, respectively, for character m at time index t; $\ell_{t,m}(z_t, i_{t,m})$ is the data likelihood of character m at t; $i_{t,m}$ is the TTI of character m assuming that m is the target character at t; $\ell 1(z_t, i_{t,m})$ is the likelihood of a classifier score, conditioned on $i_{t,m}$, if m is presented at t; and $\ell 0(z_t)$ is the likelihood of a classifier score if m is not presented at t.

Modeling TTI-specific target classes minimizes the penalty during the probability update process for target character presentations with low TTI values that are more likely to generate low classifier scores. Figure 4.1 shows an example of target and non-target conditional classifier score PDFs without TTI information, and also target classifier score PDFs conditioned on TTI 1 and 2. In this example, the TTI = 1 target classifier score PDF is more similar to the non-target classifier score PDF than the target classifier PDF. If, for example, a target is presented with a TTI of 1, it would likely lower the target character's probability when using the naïve (i.e., no TTI information) Bayesian update. Modeling the

refractory effects, however, allows the Bayesian algorithm to weight stimulus presentations differently based on every presented character's TTI.

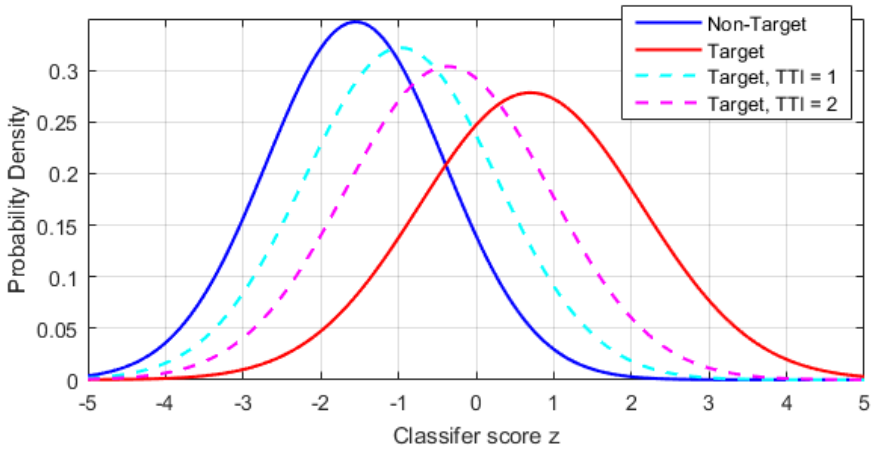

Figure 4.1 – Example of target and non-target classifier score PDFs without TTI information in red and blue curves, respectively. The target classifier score PDFs conditioned on TTI 1 and 2 are shown with dotted cyan and magenta curves. As TTI decreases, the target conditional classifier score PDFs with TTI information shifts closer to the non-target classifier score PDF.

4.1.2 Expected Discrimination Gain

Assuming character probability updates with TTI information according to (4.1), the EDG can be rewritten to incorporate TTI information as shown below:

$$\Delta TD_{KL}(\mathcal{F}_{t+1}^{*}) = \int_{-\infty}^{\infty} \sum_{m=1}^{M} p_{t,m}\ell_{t+1,m}(z_{t+1}^{*}, i_{t+1,m}) \log\left(\frac{\ell_{t+1,m}(z_{t+1}^{*}, i_{t+1,m})}{p(z_{t+1}^{*}|Z_t)}\right) dz_{t+1}^{*} \tag{4.3}$$

where $\Delta TD_{KL}(\mathcal{F}^*_{t+1})$ is the *TTI-EDG* (i.e., EDG with TTI information), given the flash group $\mathcal{F}^*_{t+1}$. Grouping characters in the flash group $\mathcal{F}^*_{t+1}$ with the same TTI values, i.e. $i_{t+1,m}$ according to (4.2), the EDG can be simplified accordingly:

$$\Delta TD_{KL}(\mathcal{F}^*_{t+1}) \tag{4.4}$$

$$= \int_{-\infty}^{\infty} \sum_{\forall m: \ell^*_{t+1,m}=1} p_{t,m} \ell_{t+1,m}\left(z^*_{t+1}, i_{t+1,m}\right) \log\left(\frac{\ell_{t+1,m}\left(z^*_{t+1}, i_{t+1,m}\right)}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1}$$

$$+ \int_{-\infty}^{\infty} \sum_{\forall m: \ell^*_{t+1,m}=0} p_{t,m} \ell_0(z^*_{t+1}) \log\left(\frac{\ell_0(z^*_{t+1})}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1}$$

$$= \int_{-\infty}^{\infty} \sum_{\forall TTI} P1_{t,TTI} \ell 1(z^*_{t+1}, TTI) \log\left(\frac{\ell 1(z^*_{t+1}, TTI)}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1} \tag{4.5}$$

$$+ \int_{-\infty}^{\infty} P0_t \ell 0(z^*_{t+1}) \log\left(\frac{\ell 0(z^*_{t+1})}{p(z^*_{t+1}|Z_t)}\right) dz^*_{t+1}$$

where $P1_{t,TTI}$ and $P0_t$ are defined accordingly:

$$P1_{t,TTI} = \sum_{\forall c:\, \ell^*_{t+1,c}=1 \,\cap\, i_{t+1,c}=TTI} p_{t,c} \tag{4.6}$$

Using a similar process as in (3.18) to group characters, the conditional data probability, $p(z^*_{t+1}|Z_t)$, can be expressed as:

$$p(z_{t+1}^*|Z_t) = \sum_{\forall TTI} \sum_{\forall c:\, \ell_{t+1,c}^*=1 \cap i_{t+1,c}=TTI} \mathcal{P}_{t,c}\ell 1(z_{t+1}^*, TTI) + \sum_{\forall c:\, \ell_{t+1,c}^*=0} \mathcal{P}_{t,c}\ell 0(z_{t+1}^*) \tag{4.7}$$

$$= \sum_{\forall TTI} P1_{t,TTI}\ell 1(z_{t+1}^*, TTI) + P0_t\ell 0(z_{t+1}^*) \tag{4.8}$$

Using (4.5), (4.9), and exploiting the sum of probabilities (i.e. $\sum_{\forall TTI} P1_{t,TTI} + P0_t = 1$), the $\Delta TD_{KL}(\mathcal{F}_{t+1}^*)$ is:

$$\begin{aligned}
&\Delta TD_{KL}(\mathcal{F}_{t+1}^*) \\
&= \int_{-\infty}^{\infty} \sum_{\forall TTI} P1_{t,TTI}\ell 1(z_{t+1}^*, TTI) \log\left(\frac{\ell 1(z_{t+1}^*, TTI)}{\sum_{\forall TTI} P1_{t,TTI}\ell 1(z_{t+1}^*, TTI) + \left(1 - \sum_{\forall TTI} P1_{t,TTI}\right)\ell 0(z_{t+1}^*)}\right) dz_{t+1}^* \\
&+ \int_{-\infty}^{\infty} \left(1 - \sum_{\forall TTI} P1_{t,TTI}\right) \\
&\qquad * \ell 0(z_{t+1}^*) \log\left(\frac{\ell 0(z_{t+1}^*)}{\sum_{\forall TTI} P1_{t,TTI}\ell 1(z_{t+1}^*, TTI) + (1 - \sum_{\forall TTI} P1_{t,TTI})\ell 0(z_{t+1}^*)}\right) dz_{t+1}^*
\end{aligned} \tag{4.9}$$

Assuming fixed conditional classifier score PDFs shown in (4.2), the classifier score z_{t+1}^* can be marginalized out in (4.9) to yield a TTI-EDG function that depends only on $P1_{t,TTI}$ for all unique TTIs.

Assuming TTI values exist from 1 to 3 for simplicity, three cases of varying discriminability between the TTI-specific and the non-target classifier score PDFs are presented in Figure 4.2 to illustrate how stimulus selection is affected by incorporating TTI information in the EDG. In Figure 4.2(a), the TTI-specific conditional classifier score

PDFs are identical. From the EDG function shown in Figure 4.2(a), the TTI-EDG is maximized when $\sum_{TTI} P1_{t,TTI} = 0.5$. Note that in this scenario, the stimulus selection process is similar to the case where there is only one target classifier score PDF (see Figure 2.6). In Figure 4.2(b), the TTI-specific PDFs for the TTI 2 and 3 have higher discriminability from the non-target classifier score PDF compared to the PDF for TTI 1. From the EDG function shown in Figure 4.2(b), the TTI-EDG is maximized when $\sum_{TTI=2}^{3} P1_{t,TTI} = 0.5$ and $P1_{t,1} = 0$. In this scenario, stimulus selection will favor characters with TTIs of 2 or 3. In Figure 4.2(c), the TTI-specific PDFs are distinct from each other, with increasing discriminability from the non-target pdf as TTI increases. From the EDG function in Figure 4.2(c), the TTI-EDG is maximized when $P1_{t,3} = 0.5$ and $P1_{t,2} = P1_{t,1} = 0$. In this scenario, stimulus selection will favor characters with a TTI of 3. The example shows that the effect of explicitly modeling TTI-related effects in the EDG function inherently biases the selection process towards characters with a higher discriminability.

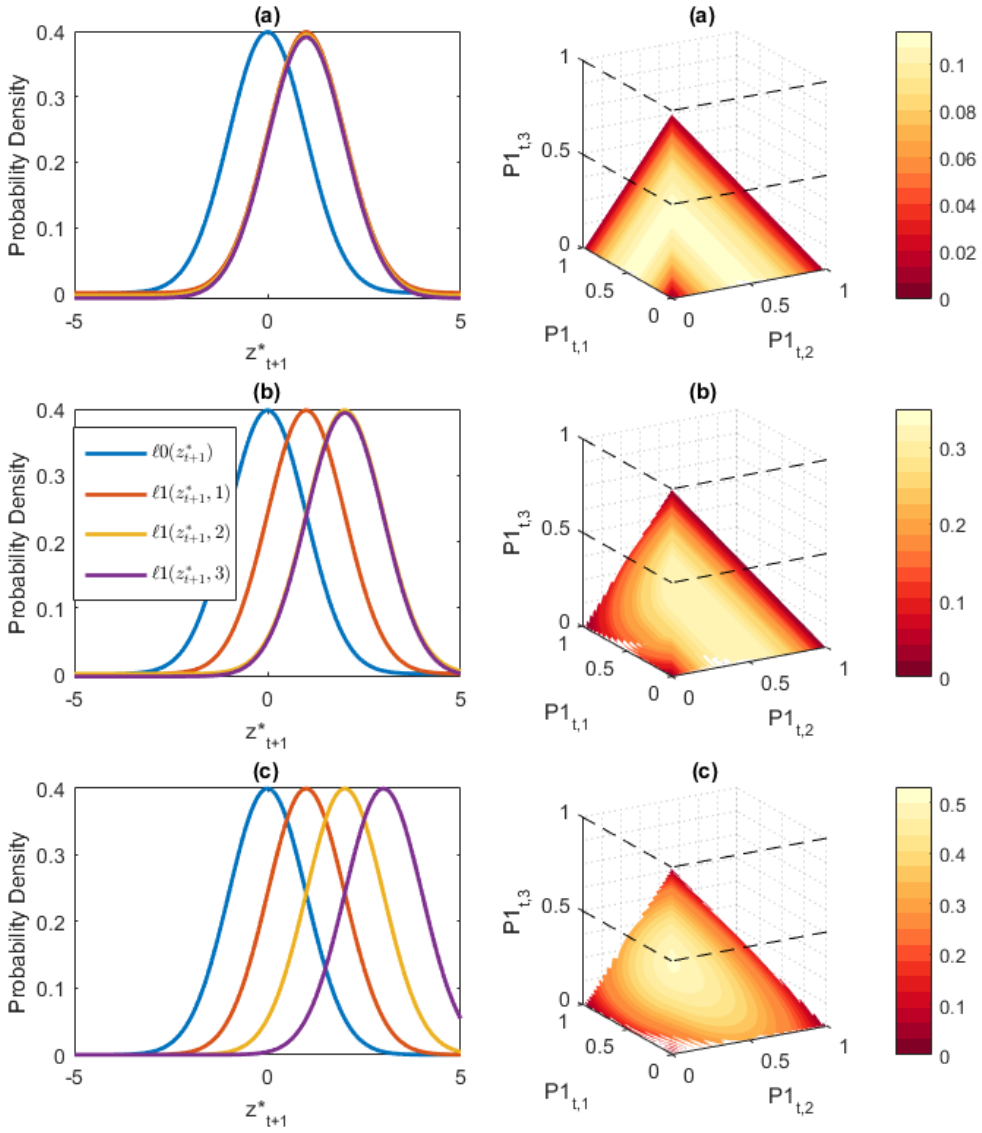

Figure 4.2 – Example of conditional classifier score PDFs for non-target and TTI 1, 2, and 3 presentations and their corresponding TTI-EDG volumes. Each row shows the conditional classifier score PDFs on the left, and TTI-EDG volumes for the corresponding conditional classifier score PDFs on the right. In example (a), the target conditional classifier score PDFs for all TTI-bins are equivalent. In example (b), the TTI-specific pdfs for TTI 2 and 3 are the same. In example (c), all three TTI bins have different conditional classifier score PDFs.

4.2 Performance Evaluation with Simulations

Two types of simulations will be performed in this section. In section 4.2.1, MC simulations will be used to evaluate the potential benefit of using TTI information to update the target character probabilities (see 4.1.1) for the RC paradigm. In 4.2.2 MC simulations will be used to evaluate the potential benefit of using the TTI-EDG metric (see 4.1.2) for a greedy, adaptive, stimulus presentation paradigm.

4.2.1 Row-column Paradigm

Refractory effects can occur quite often in the RC paradigm due to the greater proportion of short TTIs, as shown in Figure 2.9. As a result, the RC paradigm is likely to benefit from incorporating TTI-information during target character estimation. Marten et al., [84] developed a graphical model for character estimation that incorporated refractory effects. In simulation, Marten et al. showed no improvement between the graphical model that incorporated refractory effects versus the model that did not incorporate the refractory effects when using the RC paradigm. However, the graphical model framework in Martens et al. only modeled short term memory refractory effects (i.e., $TTI = 1$ vs. $TTI \neq 1$), and did not consider long term memory refractory effects (i.e., $TTI = 1$ vs. $TTI = 2$ vs $TTI = 3$ etc.). In this work, long term refractory effects are modeled in the Bayesian algorithm.

Estimates of the TTI-specific conditional classifier score PDFs may be poor if a low number of samples are used for estimation. The example TTI distribution in Figure 2.10

shows that on average, most TTI values in the RCP have fewer than 20 samples. Furthermore, pre-computing the TTI-EDG is intractable if there are many TTI groupings because the number of precomputed TTI-EDG values grows exponentially with the number of TTI groupings.

To improve the TTI-specific conditional classifier score PDF estimation and to increase tractability of the TTI-EDG pre-computation, the data from TTI values with similar conditional classifier score PDFs are binned together to generate a new set of "virtual" TTI conditional classifier score PDFs. These virtual TTIs obtained by binning a range of TTI values will be referred to as *binned-TTI*. Based on analyzing the similarity of TTI-specific pdfs of participants in [72], the TTI groupings shown in Table 4.1 were determined to provide the most suitable binned-TTIs. Figure 4.3 shows example conditional classifier score PDFs using (a) non-binned and (b) binned TTI groupings for a single participant in [72].

4.2.1.1 Methods

Using data the data from [72], collected using the RC paradigm, MC simulations (see 2.4.2) are performed to compare the naïve Bayesian algorithm (no TTI information), the TTI Bayesian algorithm (no binning), and the binned Bayesian algorithm. The MC simulations are slightly modified to account for TTI effects; the target classifier scores are drawn such that the TTI of the presentation in the simulation matches the TTI of the

presentation in the real dataset. The KDE is used to estimate the conditional classifier score PDFs in the simulations.

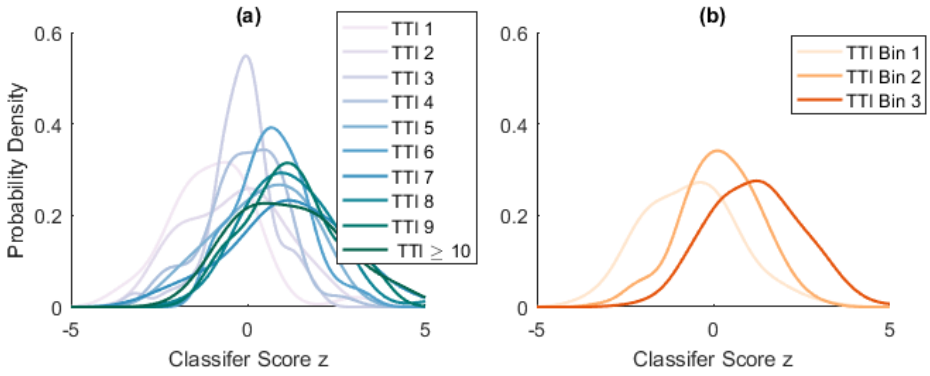

Figure 4.3 – conditional classifier score PDFs using (a) non-binned and (b) binned TTI groupings for a single participant in [72].

Table 4.1 – TTI to TTI-bin assignments

Binned-TTI	Range of TTI values in bin
1	1, 2
2	3, 4, 5, 6
3	≥7

4.2.1.2 Results

Figure 4.3 below shows the average accuracy, EST, and bit-rate obtained for each participant in the dataset by running MC simulations with the RC paradigm. 'Naïve

Bayesian Update' corresponds to the Bayesian algorithm whose update equation does not use TTI information (see section 2.2.4.2); 'TTI Bayesian update' corresponds to the Bayesian algorithm whose update equation groups TTIs without binning (see section 4.1.1); and 'Binned TTI Bayesian update' corresponds to the Bayesian algorithm whose update equation groups TTIs by binning using the bins shown in Table 4.1. Table 4.2 shows the average accuracy, EST, and bit-rate of the simulations with astrisks denoting statistical significance using the two-sided Wilcoxin signed-rank test.

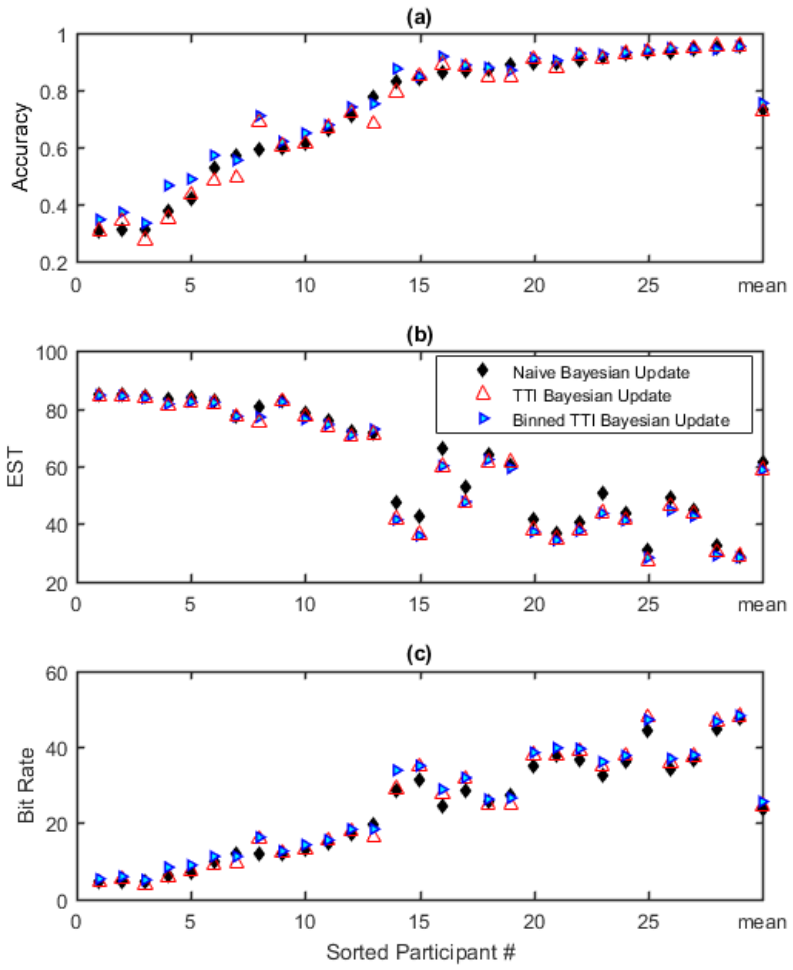

Figure 4.4 - The average accuracy (a), EST (b), and bit-rate (c) obtained for each participant by running the MC simulation. The results are sorted by the naïve Bayesian Update accuracy.

Table 4.2 – Average accuracy, EST, and bit-rate for MC simulations comparing the Naïve Bayesian update, TTI Bayesian update, and binned TTI Bayesian update. Bold font represents the highest accuracy, lowest number of stimulus presentations, or highest bit-rate of the methods compared. Conversely, the italics-underlined font represents the lowest accuracy, highest number of stimulus presentations, or lowest bit-rate of the methods compared. The asterisks indicate pairwise statistical significance with the Naïve Bayesian Update using the two-sided Wilcoxin signed-rank test.

Bayesian Update Type	Accuracy	EST	Bit-rate
Naïve Bayesian Update	0.731	*61.131*	*25.090*
TTI Bayesian Update	*0.726*	**58.675***	25.990*
Binned TTI Bayesian Update	**0.755***	58.743*	**27.177***

The simulations show that using the TTI Bayesian update or the Binned TTI Bayesian update significantly decreases the EST and significantly increases the bit-rate compared to the Naïve Bayesian update; this is especially evident in subjects with high accuracy (>80%). The binned TTI Bayesian update has a significantly higher accuracy than both the TTI Bayesian update and the Naïve Bayesian update. The low accuracy subjects (sorted subjects 1-10) experience the highest accuracy increase from the Binned TTI Bayesian update compared to the TTI or Naïve Bayesian updates.

These simulations show that using TTI information in the Bayesian algorithm can improve the accuracy, decrease the EST, and increase the bit-rate of a P300 speller using the RC paradigm. To increase the accuracy of the Naïve Bayesian update, TTI binning is necessary. For the remainder of the work, TTI binning will be used when using TTI information in the Bayesian update.

4.2.2 Greedy Adaptive Paradigm

In this section, MC simulations will be used to evaluate the potential benefit of using the TTI-EDG metric within the context of the greedy adaptive stimulus presentation paradigm. It is important to note that the data used in the MC simulation in this section was collected using the RC paradigm, however, the paradigm used in the simulation is an adaptive paradigm. The physiological effects such as refractory effects or adjacency distractions can result in different target and non-target data distributions between the data used in the simulation and data collected using the proposed greedy EDG-TTI adaptive paradigm. Therefore, the results obtained in this section are for illustrative purposes and only aim to show the potential benefit of using the greedy TTI-EDG adaptive paradigm.

4.2.2.1 Methods

MC simulations using data from [72] are used to compare the greedy naïve EDG adaptive (no TTI information) and greedy TTI-EDG adaptive paradigms. As with the MC simulation discussed in 4.2.1.1, the target classifier scores are drawn such that the TTI of the stimulus presentation in the simulation matches the TTI of the presentation in the real dataset; this allows the simulation to account for refractory effects. Three different algorithms are considered: the greedy naïve EDG adaptive paradigm with naïve Bayesian updates; the greedy naïve EDG adaptive paradigm with TTI Bayesian updates; and the greedy TTI-EDG adaptive paradigm with TTI Bayesian updates. For simplicity, the

algorithms will be referred to as naïve EDG – naïve update, naïve EDG – TTI update, and TTI EDG – TTI update.

The purpose of the naïve EDG – TTI update algorithm is to provide a baseline/control case; otherwise, if the TTI EDG – TTI update outperforms the naïve EDG - naïve update, it becomes unclear if the stimulus selection, target character probability update, or both stimulus selection and target character probability update results in the performance improvement.

The character presentation limit is set to 9, and the OD is set to 6 for all simulations in this section (identical to restrictions used in section 3.3.2.4). For each algorithm, three different TTI_{min} restrictions are considered for stimulus selection: no TTI_{min} restrictions, $TTI_{min} = 3$, and $TTI_{min} \sim$ PBP. The TTI bins shown in Table 4.1 are used for the simulations that incorporate the TTI information. The non-target and target conditional classifier score PDFs are fit to a Gaussian PDF in order to increase the tractability of precomputing the TTI-EDG. Stimuli are selected using a greedy search.

4.2.2.2 Results

Figure 4.5 shows the average accuracy, EST, and bit-rate for naïve EDG – naïve update, naïve EDG – TTI update, and TTI EDG – TTI update algorithms with no TTI_{min} constraints, $TTI_{min} = 3$, and $TTI_{min} \sim PBP$ restrictions. Table 4.3 shows the bit-rate results from Figure 4.7 in the form of numerical values.

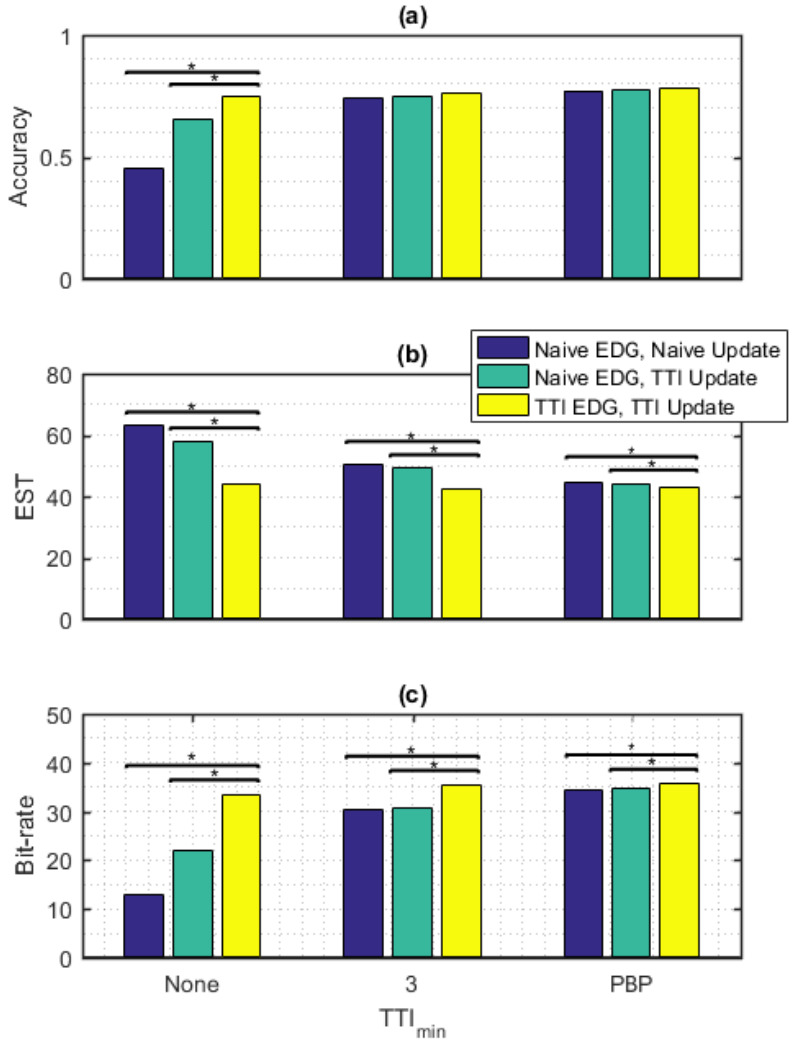

Figure 4.5 - Average (a) accuracy, (b) EST, and (c) bit-rate for naïve EDG – naïve update, naïve EDG – TTI update, and TTI EDG – TTI update algorithms with no TTI_{min} constraints, $TTI_{min} = 3$, and $TTI_{min} \sim PBP$ constraints. Asterisks represent pair-wise statistical significance. The results show statistically significant improvement in EST and bit-rate, for all three TTI restrictions, between TTI EDG - TTI Update and Naïve EDG - TTI Update algorithms.

Table 4.3 – Average bit-rates from the MC simulations comparing adaptive stimulus presentations. The first column shows the type of algorithm and the first row shows the TTI_{min} used in the simulation. Underlined and italicized values indicate the lowest bit-rate and bold values indicate the highest bit-rate for each TTI_{min} type.

TTI_{min} / Algorithm	None	3	PBP
Naïve EDG - Naïve Update	*12.78*	*30.27*	*34.37*
Naïve EDG - TTI Update	21.85	30.82	34.56
TTI EDG - TTI Update	**33.53**	**35.43**	**35.71**

The results show statistically significant improvement in EST and bit-rate, for all three TTI restrictions, between Naïve EDG - TTI Update and Naïve EDG - Naïve Update algorithms. The results also show statistically significant improvement in EST and bit-rate, for all three TTI restrictions, between TTI EDG - TTI Update and Naïve EDG - TTI Update. There is a statistically significant increase in accuracy between the Naïve EDG - TTI Update and Naïve EDG - Naïve Update algorithms as well as between TTI EDG - TTI Update and Naïve EDG - TTI Update algorithms when there are no TTI_{min} restrictions, however, there is no significant improvement in accuracy for $TTI_{min} = 3$ or $TTI_{min} \sim PBP$ when for any of the algorithms with TTI information.

Table 4.3 indicates that incoprorating TTI information in the Bayesian update and the EDG metric provides the highest accuracy, EST, and bit-rate improvement to the

algorithm with no TTI_{min} restriction. Consider an example participant from the dataset used in this section whose conditional classifier score PDFs and corresponding EDG-TTI volumes are shown in Figure 4.6. There is a significant difference between discriminability at high TTIs (TTI bin 2 and 3) than at low TTIs (TTI bin 1) as shown in the conditional classifier score PDFs. Therefore, it is expected that the greedy TTI-EDG adaptive paradigm will select fewer low TTI presentations than the Naïve-EDG adaptive paradigms. Figure 4.7 shows the TTI distributions for the participant comparing TTI EDG - TTI Update and Naïve EDG - TTI Update algorithms. These TTI distributions indicate that the Naïve EDG - TTI Update almost exclusively presents TTI = 1 presentations, however, the TTI EDG - TTI Update algorithm can account for the low discriminability at low TTI presentation and instead chooses most presentations at TTI = 3. This reduction in low-TTI presentations when using the greedy TTI-EDG versus the greedy naïve EDG adaptive stimulus selection paradigm likely accounts for the signficant increase in accuracy and bit-rate and significant decrease in EST when no TTI_{min} restrictions are set for the adaptive stimulus selection paradigm.

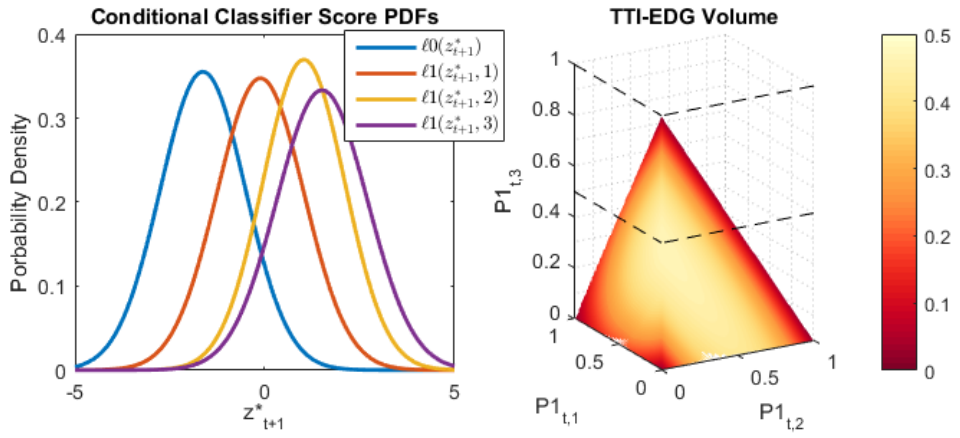

Figure 4.6 – (Left) presents the trained conditional classifier score PDFs for non-target and target binned TTI 1, 2, and 3 presentations for a single participant. (Right) is the corresponding TTI-EDG volume for the conditional classifier score PDFs.

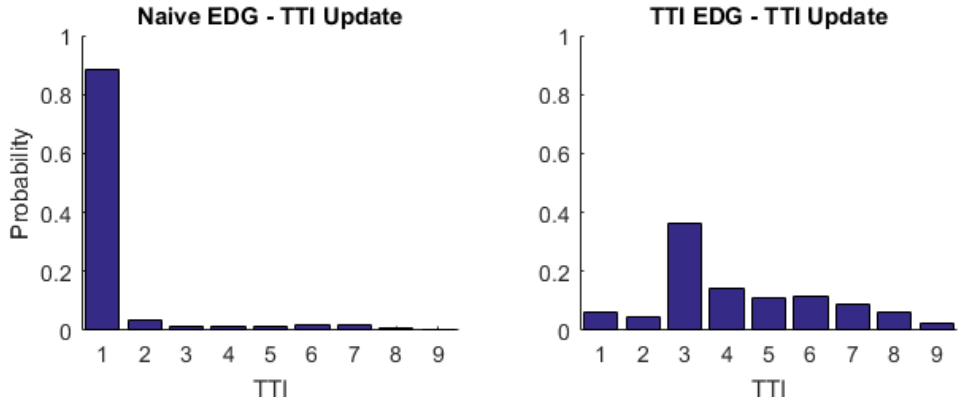

Figure 4.7 –TTI distributions for the naïve EDG – TTI update (left) and TTI EDG - TTI update (right) algorithms with no TTI_{min} restrictions for the participant whose trained conditional classifier score PDFs are shown in Figure 4.6.

4.3 Discussion

In this chapter, TTI information was used to model refractory effects in the EEG data with the goal of increasing the accuracy, decreasing the EST, and increasing the bit-

rate of the P300 speller. First, the target character probability update was modified such that the conditional classifier score PDFs were conditioned on the TTI of the characters being presented. Next, the greedy EDG adaptive paradigm was modified to incorporate the modified target character proabability update to yield the greedy TTI-EDG adaptive paradigm.

It was shown that using TTI information in the Bayesian update equation when using the RC paradigm can improve accuracy, EST, and bit-rate, for users of the P300 speller, as was shown in simulation (section 4.2.1.2). The MC simulations indicated that in order to increase the accuracy of the speller when using TTI information, TTI binning should be used to group TTIs when estimating the conditional classifier score PDF. TTI binning is likely beneficial because the number of training examples for all individual TTIs is limited; therefore, it may be possible to remove the need for TTI binning if more training data is collected.

In section 4.2.2.2, it was shown that using the greedy TTI-EDG adaptive paradigm can further improve the performance of the P300 speller compared to the naïve greedy EDG adaptive paradigm that was proposed in Chapter 3. For all TTI_{min} restrictions used in the simulation, the greedy TTI-EDG adaptive paradigm yields a higher average bit-rate than the naïve greedy EDG adaptive paradigm. The simulations with no TTI_{min} restrictions obtains the highest accuracy, EST, and bit-rate improvement when incorporating TTI information in the form of the Bayesian update equation and the

adaptive paradigm. This shows that modelling refractory effects in the Bayesian update and the adaptive stimulus selection can reduce the need for TTI restrictions.

The $TTI_{min} \sim PBP$ and $TTI_{min} = 3$ restrictions, however, yield a higher overall bit-rate when the TTI EDG - TTI Update was used compared to using no TTI_{min} restrictions. This likely occurs because the greedy EDG paradigms are non-optimal; they are myopic and do not look ahead to future stimulus presentations. For example, there may be a case where presenting a character at $t+1$ will barely maximize $\Delta TD_{KL}(\mathcal{F}_{t+1}^{*})$, however, presenting that character at $t+2$ instead will yield a significantly higher $\Delta TD_{KL}(\mathcal{F}_{t+2}^{*})$ than $\Delta TD_{KL}(\mathcal{F}_{t+1}^{*})$. In practice, it would be preferable to wait until $t+2$ to present the character, however, the TTI-EDG paradigm will instead present that character at $\Delta TD_{KL}(\mathcal{F}_{t+1}^{*})$.

The results in this chapter show the potential for improvement of the speller when TTI information is used in the Bayesian update equation as well as in the adaptive stimulus presentation paradigm. However, it will be important to perform online experiments prior to making concrete conclusions because of physiological phenomena that are not explicitly accounted for in the simulation, such as adjacency distraction errors or fatigue.

5 Probabilistic P300 and Eye-tracker Hybrid System

In this chapter, a novel probabilistic hybrid P300 and eye-gaze system is proposed, and the results of the hybrid system are presented. The goal of the hybrid speller is to alleviate the individual disadvantages of both the P300 and eye-tracker spellers as well as improve upon the performance of the individual spellers.

This chapter is outlined as follows: In section 5.1, a novel, standalone probabilistic eye-gaze typing system for spelling using a P300 display will be introduced. In section 5.2, a probabilistic hybrid speller is introduced and both online and offline experiments are described that evaluate the performance and robustness of the system. In section 5.3 possible avenues for future work are discussed.

5.1 Probabilistic Eye Gaze Spelling

The eye-gaze spelling algorithm that was utilized in this work is a probabilistic, dynamic stopping algorithm based on the Bayesian dynamic stopping algorithm proposed for processing P300 data presented in Throckmorton et al. [72] (see section 2.2.4). Developing an eye-typing framework similar to that of the P300 speller allows for a natural fusion of eye-gaze and P300 spellers. The flow-chart presented in Figure 5.1 below describes the probabilistic eye-gaze spelling framework.

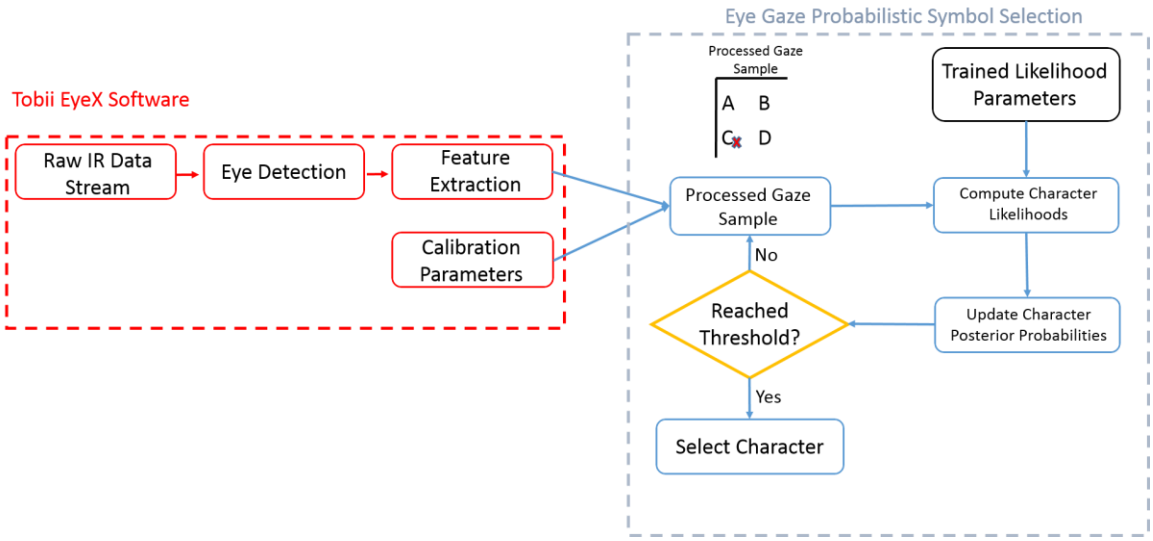

Figure 5.1 – Flow chart for choosing the target character using the probabilistic speller. The portion of the flow chart in the dashed-red rectangle denotes the processing performed by the Tobii EyeX software. This software is propriety; thus, the exact processing methods are unknown. As the Tobii EyeX software outputs data at a sampling rate of 30Hz, the data is used to estimate the character likelihood given trained character parameters. The likelihoods are then used to estimate the character posteriors based on the eye-gaze data.

First, eye-tracker calibration is performed by the user (see Section 2.1). Next, the training parameters are learned using training data, as will be discussed in section 5.1.1. After the training parameters are learned, an estimate of the target character is made during a testing phase trial. After each observation of gaze data is collected during the testing phase, the character likelihood and posterior probabilities are computed as discussed in section 5.1.2. Finally, when any character's posterior probability exceeds a predefined threshold, the character with the highest posterior probability is selected, and the trial is concluded.

5.1.1 Eye Gaze Training Character Parameters

After the eye-tracker has been calibrated for the user, the data likelihood parameters must be estimated in order to compute each character's posterior probability during the testing phase. An assumption is made that the eye-gaze data is drawn from an identically independently distributed (iid) multivariate Gaussian distribution centered at the position of the target character $\boldsymbol{\mu}_m \in \Re^2$and there exists some covariance $\boldsymbol{\Sigma}_m \in \Re^{2x2}$. Therefore, the eye gaze data likelihood can be defined by:

$$p(Y_t|c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) = \mathcal{N}(\mathrm{Y_t}|\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{5.1}$$

where $\boldsymbol{\mu}_m \in \mathcal{R}^2$ is the estimated position of the target character on the screen, $\boldsymbol{\Sigma}_m \in \mathcal{R}^{2x2}$ is the estimated eye-gaze covariance matrix for the target character. Given this distribution of the data, $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ can be estimated using training data collected in parallel to the EEG data during the training phase of the P300 speller.

During the training task, the subject is instructed to gaze at a specific character, c_*, and the likelihood parameters $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ are estimated by finding the mean and covariance of the eye-gaze data. In an optimal training scenario, both $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ can be estimated for each character on the grid. However, this amount of training would be prohibitively time consuming. Thus, rather than measure $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ as a function of each individual character, $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ are instead estimated from the eye gaze data that was collected during the EEG classifier training process. For the variance, a simplifying

assumption is made that the covariance of all of the characters in the grid is identical and equal to $\mathbf{\Sigma}$. $\mathbf{\Sigma}$ is assumed to be a diagonal matrix parameterized by σ_v^2, the variance in the vertical axis, and σ_h^2, the variance in the horizontal axis:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_h^2 \end{bmatrix} \tag{5.2}$$

To estimate σ_v^2 and σ_h^2, the variance in both the horizontal and vertical axis is estimated for all target characters that have training data, and the maximum variance in the horizontal and vertical directions are set to σ_v^2 and σ_h^2 respectively. The max operator was selected to avoid underestimating the variance in the signal model because underestimating the variances can dramatically degrade the performance of the speller, as will be demonstrated in the offline simulations presented in section 5.2.4.

Since the characters are evenly spaced on the grid for the horizontal and vertical direction of the speller matrix, a least squares linear regression (LSLR) approach was used to learn $\boldsymbol{\mu}_m$ for all characters, c_m. Let the training labels $\mathbf{\Psi}_T \in R^{Tx3}$ where T is the total number of training examples (i.e., observations) collected in the training task. Each row of $\mathbf{\Psi}_T$ is termed Ψ_t and is set to $[1, row - 1, col - 1]$, where the first term is the bias term, row is the row, and col is the column of the target character at presentation t. For example, in the 9x8 matrix speller grid, the upper-left-most character takes on the index [1, 0, 0] and the bottom right-most character takes on the label [1, 8, 7].

The first step of the LSLR algorithm is to learn the weights $\boldsymbol{W}_{eye} \in R^{3x2}$ using the training gaze $\boldsymbol{Y}_T \in R^{Tx2}$ and its corresponding training labels $\boldsymbol{\Psi}_T$. The weights are learned using the normal equation below [108]:

$$\boldsymbol{W}_{eye} = (\boldsymbol{\Psi}_T{}'\boldsymbol{\Psi}_T)^{-1}\boldsymbol{\Psi}_T{}'\boldsymbol{Y}_T \tag{5.3}$$

Once the weights are learned, $\boldsymbol{\mu}_m$ is learned for all characters by using the label associated with character c_m using the following:

$$\boldsymbol{\mu}_m = \Psi_m \boldsymbol{W}_{eye} \tag{5.4}$$

where Ψ_m is the label associated with character c_m.

Examples of the training gazes $\boldsymbol{Y}_T$ from one training session for a single participant are plotted overlaying the speller matrix shown in Figure 5.2. The colors/character combination corresponds to the character that the participant was instructed to focus on. Most of the gazes overlap the target character, but there do exist gazes that are far away from the target character. For example, there are gazes at 'O' and 'X' when attention should have been on 'H', and gazes at '2' and 'S' when attention should have been on 'U'.

After collecting training data, $\boldsymbol{W}_{eye}$ and $\boldsymbol{\mu}_m$ are estimated for all of the characters, even those that had no corresponding observations in training; an example of an estimate of $\boldsymbol{\mu}_m$ for all characters for one participant are shown in Figure 5.3. Some $\boldsymbol{\mu}_m$, especially

those far away from characters that were used as target characters in training, are far away from the characters themselves. Since only words are spelled during testing for this experiment, the bias in the characters at the bottom of the screen (e.g. 'Ctrl','Caps', etc.) will not affect the accuracy of the system, since those characters are never spelled. If it is expected that all characters on the grid will be used during testing, it is important to collect training data for characters throughout the grid (e.g, 'Caps', 'Esc', 'Sleep', 'Bs', etc.).

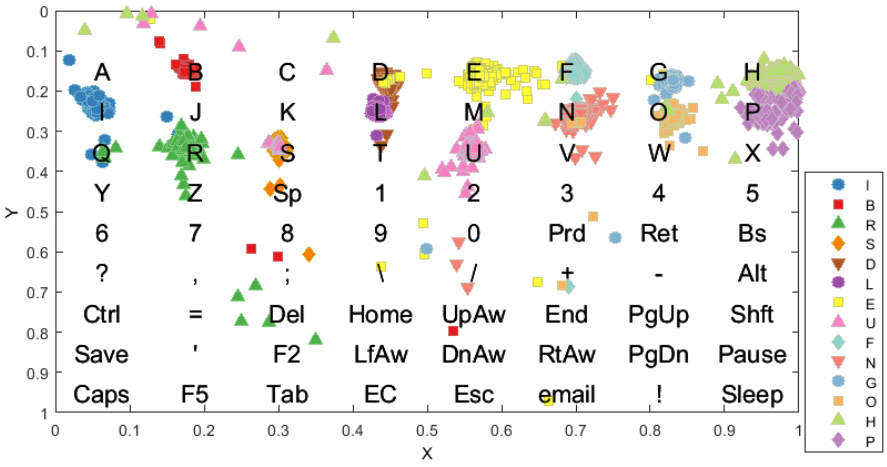

Figure 5.2 – Illustration of the grid used for the spelling experiments, along with an example of the training gazes, Y_T measured for a single participant plotted on a speller matrix. The color and character combination corresponds to the character that the participant was instructed to focus on. There are 168 eye-gaze data points for each character being trained on, and there are 18 characters being trained on. In all, there are 3024 data points that are gathered during the training phase.

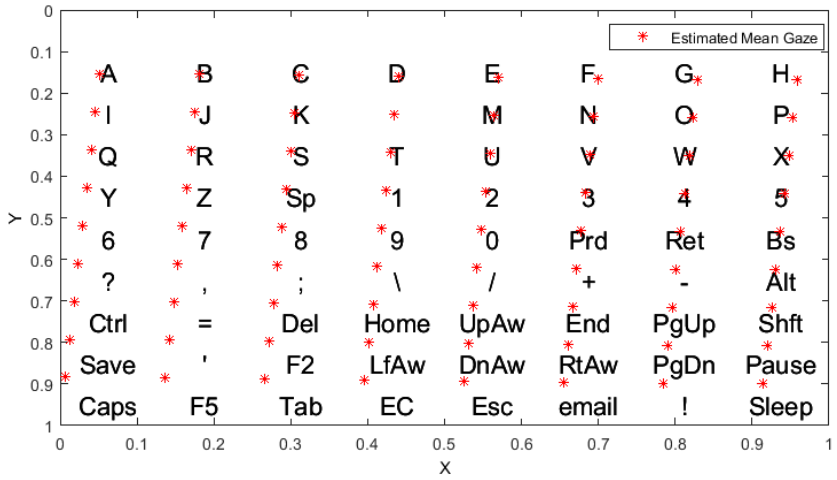

Figure 5.3 - Trained character position means, μ_m, for all characters for a single participant, are plotted in this figure as red asterisks. Some μ_m, especially characters' means are far away from the characters. This bias occurs because training data is limited to characters.

5.1.2 Eye Gaze Posterior Update

After the parameters of the likelihood are estimated with the training data, the posterior probability for every character can be learned using Bayes rule:

$$p(c_m = c_* | \boldsymbol{Y}_t, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) = \frac{p(\boldsymbol{Y}_t | c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) p(c_m = c_*)}{\sum_{i=1}^{M} p(\boldsymbol{Y}_t | c_i = c_*, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) p(c_i = c_*)} \tag{5.5}$$

The likelihood in the numerator in equation (5.5) can then be decomposed further towards developing an update equation such that the posterior probability can be updated after every presentation, t. Given that each observation is i.i.d.:

$$p(\boldsymbol{Y}_t|c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) = p(Y_t|c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)\, p(\boldsymbol{Y}_{t-1}|c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{5.6}$$
$$\propto p(Y_t|c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)\, p(c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m|\boldsymbol{Y}_{t-1})$$

The expression in equation (5.6) that is proportional to the likelihood of equation (5.5) can then be substituted for the likelihood and re-normalized resulting in the following update equation:

$$p(c_m = c_*|\boldsymbol{Y}_t, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{5.7}$$
$$= \frac{p(Y_t|c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)\, p(c_m = c_*, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m|\boldsymbol{Y}_{t-1}) p(c_m = c_*)}{\sum_{i=1}^{M} p(Y_t|c_i = c_*, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)\, p(c_i = c_*, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i|\boldsymbol{Y}_{t-1}) p(c_i = c_*)}$$

Using this posterior probability update equation enables the system to learn the probability that character c_m is the target character after presentation t in the testing phase.

5.2 Eye Gaze and EEG Fusion

In sections 2.2 and 5.1, probabilistic P300 and eye-gaze spellers were discussed individually, respectively. Eye-gaze spellers are used because of their high accuracy and speed. P300 spellers are used in cases where eye-gaze is unreliable as is the case for some people with ALS ([9], [14]–[17]). To alleviate the disadvantages of the individual spellers, a probabilistic framework that uses data from the EEG-based P300 speller as well as the eye-tracker is developed; the hybrid speller will be introduced and described in this section.

In section 5.2.1, the algorithm used to fuse the probabilistic eye-gaze typing system and the P300 speller will be discussed. Sections 5.2.2, 5.2.3, and 5.2.4 will discuss the details of the experiment performed to evaluate the performance of the hybrid speller. Sections 5.2.3.1 and 5.2.4.1 will show the results of the experiments, and section 5.2.5 will discuss the significance of the results obtained in the experiments.

5.2.1 Eye Gaze and EEG Fusion Algorithm

The posterior probability (i.e., target probability) for each character, c_m can be learned using Bayes rule where:

$$p(c_m = c_* | \boldsymbol{z}_t, \boldsymbol{Y}_t, \boldsymbol{\Phi}_{m,t}) = \frac{p(\boldsymbol{z}_t, \boldsymbol{Y}_t | c_m = c_*, \boldsymbol{\Phi}_{m,t}) p(c_m = c_*)}{\sum_{i=1}^{M} p(\boldsymbol{z}_t, \boldsymbol{Y}_t | c_i = c_*, \boldsymbol{\Phi}_{i,t}) p(c_i = c_*)} \tag{5.8}$$

where $\Phi_{m,t} = \{\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m, \mathcal{F}_t\}$ is the set of eye-gaze parameters for character m and flash group at t and $\boldsymbol{\Phi}_{m,t} = \{\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m, \boldsymbol{\mathcal{F}}_t\}$ is the vector of the set of eye-gaze parameters for character m and all flash groups up to t. Similar to the derivation of the posterior probability update equation for the probabilistic eye-gaze speller in section 5.1, the joint likelihood can be rewritten as:

$$\begin{aligned} p(\boldsymbol{z}_t, \boldsymbol{Y}_t | c_m = c_*, \boldsymbol{\Phi}_{m,t}) &= p(z_t, Y_t | c_m = c_*, \Phi_{m,t}) p(\boldsymbol{z}_{t-1}, \boldsymbol{Y}_{t-1} | c_m = c_*, \boldsymbol{\Phi}_{m,t-1}) \\ &\propto p(z_t, Y_t | c_m = c_*, \Phi_{m,t}) p(c_m = c_* | \boldsymbol{z}_{t-1}, \boldsymbol{Y}_{t-1}, \boldsymbol{\Phi}_{m,t-1}) \end{aligned} \tag{5.9}$$

replacing the likelihood in (5.8) with (5.9) and normalizing yields:

$$p(c_m = c_* | \boldsymbol{z}_t, \boldsymbol{Y}_t, \boldsymbol{\Phi}_{m,t}) = \frac{p(z_t, Y_t | c_m = c_*, \Phi_{m,t}) p(c_m = c_* | \boldsymbol{z}_{t-1}, \boldsymbol{Y}_{t-1}, \boldsymbol{\Phi}_{m,t-1})}{\sum_{i=1}^{M} p(z_t, Y_t | c_i = c_*, \Phi_{i,t}) p(c_i = c_* | \boldsymbol{z}_{t-1}, \boldsymbol{Y}_{t-1}, \boldsymbol{\Phi}_{i,t-1})} \tag{5.10}$$

As can be seen in the update equation (5.10), the a-priori probability is the posterior probability learned from the previous presentation. A uniform distribution was assumed for the a-priori probability for the first presentation in a trial, but a language model can be used [e.g. 69] to initialize the a-priori probability.

An assumption is made that there is complete independence between both EEG and eye-gaze likelihood. This assumption is not completely accurate given that the P300 speller performs better when the participant is looking directly at the character being spelled [109]. However, this assumption greatly simplifies the problem formulation and will therefore be assumed to be true for this work. Given this assumption, the posterior probability after every presentation can be learned using equation (5.11) below:

$$p(c_m = c_* | \boldsymbol{z}_t, \boldsymbol{Y}_t, \boldsymbol{\Phi}_{m,t}) \tag{5.11}$$
$$= \frac{p(z_t | c_m = c_*, \Phi_{m,t}) p(Y_t | c_m = c_*, \Phi_{m,t}) p(c_m = c_* | \boldsymbol{z}_{t-1}, \boldsymbol{Y}_{t-1}, \boldsymbol{\Phi}_{m,t-1})}{\sum_{i=1}^{M} p(z_t | c_i = c_*, \Phi_{i,t}) p(Y_t | c_m = c_*, \Phi_{i,t}) p(c_i = c_* | \boldsymbol{z}_{t-1}, \boldsymbol{Y}_{t-1}, \boldsymbol{\Phi}_{i,t-1})}$$

5.2.2 Experimental Methods

For the online experiment, sixteen able-bodied participants were recruited from Duke University. One participant was unable to complete the experiment. Incorrect training data was used to train the online classifier for another participant; thus, only offline simulation results for which training could be corrected are presented for this participant. For the remaining fourteen participants, both online and offline results are presented. All data was collected in a sound-proof booth at Duke University. The participants gave informed consent, and the study was approved by the Duke IRB.

Both EEG and eye-gaze data were collected while the participant was using the P300 speller, including experiments that only used the EEG data for decision-making; this eye-gaze data was used in the offline experiments discussed in section 5.2.4. Characters were presented using the checkerboard (CB) paradigm [25]. Adjacent characters were presented using different colors as color has been shown to improve speller performance [110].

5.2.2.1 Data Collection

EEG and eye-gaze data were collected in parallel during both training and testing phases of the experiment. EEG data were collected in 800ms windows after each presentation, and the eye gaze position at the end of each window was recorded. Therefore, the number of eye-gaze data points is equivalent to the number of

presentations, and each 800ms EEG sample has a single corresponding eye-gaze data point.

One session of data collection was completed for each participant, and each session included one training and four online testing tasks. Three words were spelled during the training task, and this data was used to train the P300 classifier, and five words were spelled during each of the four online testing tasks. The words being spelled by the participants were randomly chosen from a subset of English words from the English Lexicon Project [104]. The subset consists of 400 six-character words with the highest frequency of occurred in written communication as measured by the HAL corpus frequency [111]. The algorithms used for processing the data from the four testing tasks are described in detail in section 5.2.3. The order of the online testing tasks were randomized to avoid biases that may arise because of the temporal non-stationarity typically found in EEG data [103]. During training, 144 presentations of data collected were collected for each character. For testing, the dynamic stopping data collection algorithm was used with the maximum number of presentations set to 144.

5.2.3 Experiment 1: Online Test of the System

This experiment was performed to assess the performance of the hybrid speller. After training data was collected and the EEG classifier score and eye-gaze data likelihood parameters were learned, all participants completed four online testing spelling tasks. An EEG-alone (P300 speller) task was run as a baseline, and the three additional tasks used

the same classifier with different **Σ** estimates. In the first task, **Σ** was estimated from training data as described previously. However, the estimated covariance, **Σ,** was expected to be very low (<0.01) according to the technical information provided with the Tobii Inc. eye-tracker [54] due to the subject pool consisting of participants without disabilities. This has the potential to decrease the robustness of the hybrid system since the low variance of the eye tracker data would result in the system placing extremely high confidence in the eye gaze data relative to the EEG data. A foreseeable issue that can arise with using a small estimated variance is a bias between the estimated and actual character means (i.e. the location of the character on the screen). One way to potentially alleviate this problem is by artificially increasing the variance being used by the algorithm, giving more weight to the EEG data. Robustness as a function of the variance was investigated by setting **Σ** to $\begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$ and $\begin{bmatrix} 1.0 & 0 \\ 0 & 1.0 \end{bmatrix}$ in the final two tasks; these two conditions were termed medium and high variance.

5.2.3.1 Experiment 1 Results

In obtaining the online results, it was hypothesized that the hybrid system (EEG and eye-gaze) would outperform the EEG-alone (P300 speller) system. Figure 5.4 shows the accuracy, Figure 5.4(a), average number of stimulus presentations, Figure 5.4(b), and bit-rate, Figure 5.4(c), for the four different online testing tasks performed by participants who completed the experiment. Participant numbers are listed on the x-axis. An incorrect set of classifier scores (the BCI2000 default classifier scores) were inadvertently used

during online spelling for participant 1, therefore those results are not presented. Participants were sorted in ascending order based on their EEG accuracy.

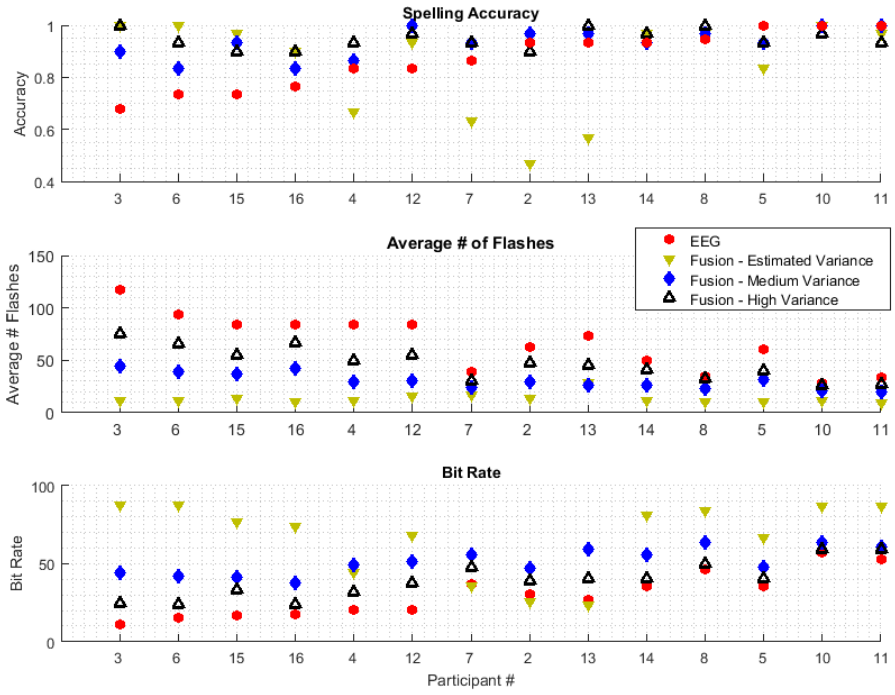

Figure 5.4 – A plot of the spelling accuracy, average number of stimulus presentations, and bit-rate for the 4 different online testing tasks performed by the participants. The participants are sorted by EEG-alone accuracy in ascending order.

The results in Figure 5.4 indicate that the bit rate of the fusion algorithm under all of the variance conditions tends to be higher than that for the EEG-alone algorithm. In all cases except for participant 5, there is a variance condition that achieves a higher accuracy

than the EEG-alone algorithm. In addition, the average number of stimulus presentations is highest for the EEG-alone algorithm across all participants. The fusion algorithm, particularly with the estimated variance for the eye gaze, performs poorly in terms of accuracy for participants 2, 4,6, 7, and 13. The estimated variance condition has the worst bit-rate compared to all other conditions in participants 2, 7, and 13. These results show that the fusion algorithm with estimated variance typically outperforms any other condition, however, in some cases the estimate variance condition is not as robust as the other conditions, leading to a lower bit rate.

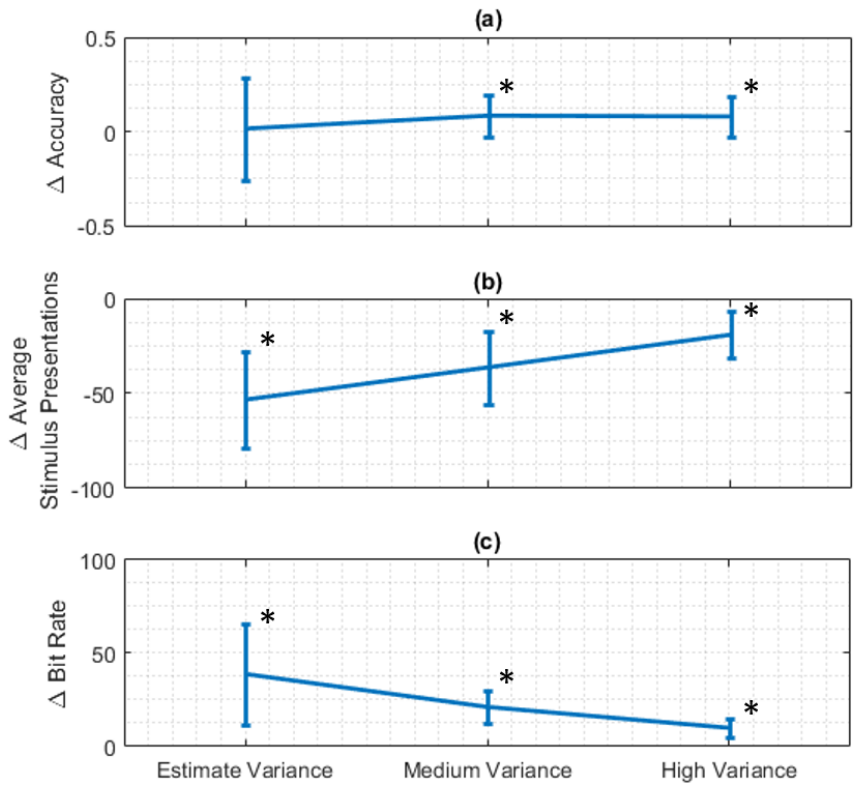

Figure 5.5 - This figure plots the difference between the EEG-alone results and the three fusion algorithm results that are shown in Figure 5.4. Subplots (a), (b), and (c) plot the average and standard deviation of the difference between the accuracy, average number of stimulus presentations, and bit rate of the EEG-alone algorithm and the fusion algorithm for all three variance conditions. The x-axis corresponds to the fusion algorithm that is being compared to the EEG-alone algorithm, and the y-axis is the difference between the EEG-alone algorithm and the corresponding fusion algorithm. In addition to the average, this figure shows ± 1 standard deviation across participants in the form of error bars. The asterisks denote statistical significance (i.e., $p<0.05$ using the two-sided Wilcoxon signed rank test).

Figure 5.5 aggregates the online results shown in Figure 5.4 and examines the difference in the performance metrics between the baseline algorithm and the fusion

algorithm. For each participant, the accuracy, average number of stimulus presentations, and bit rate results for the EEG-alone algorithm were subtracted from the results for the fusion algorithm under the three variance conditions. The differences were then averaged across participants to investigate the relative performance of the fusion algorithm compared to using EEG-alone. The results indicate that on average, the fusion algorithm has a higher accuracy, lower average number of stimulus presentations, and higher bit rate than the baseline algorithm across variance conditions. These results show that the fusion algorithm does improve the performance of the speller, however, the estimate variance condition isn't as robust as the medium and high variance conditions as shown by the larger error bars in the accuracy, average number of stimulus presentations, and bit-rate differences. Using the two-sided Wilcoxon signed rank test [105], the performance improvement for the hybrid speller was statistically significant for accuracy, average number of stimulus presentations, and bit-rate for all variance conditions with the exception of accuracy at the estimate variance condition.

5.2.4 Experiment 2: Offline Test of the System

In addition to the online experiments, several offline simulations were performed to further investigate the fusion algorithm's robustness to potential mismatches between the training and testing eye-gaze parameters. Two types of errors were considered: a horizontal bias and less accurate eye gaze. Horizontal bias might be considered when there is a mismatch between the trained $\boldsymbol{\mu}_m$, and $\boldsymbol{\mu}_m$ observed during testing. This can

occur if the user of the system has oculomotor abnormalities that doesn't allow them to gaze directly at one or more of the target character (i.e., $\boldsymbol{\mu}_m$ cannot be represented $\boldsymbol{\mu}_m = \Psi_m \boldsymbol{W}_{eye}$). Less accurate eye gaze, or higher variance eye gaze, would be considered if the user has difficulty using an eye tracker (e.g., eyelid drooping, nystagmus, etc.), and $\boldsymbol{\Sigma}_m$ is incorrectly estimated during training. These error types were tested under the same variance conditions as the online testing (EEG-alone, fusion with estimated variance, medium variance, and high variance) as well as under eye-gaze alone (see 5.1.2).

The goal of the offline experiments was to observe the robustness of the EEG-alone, eye-gaze alone, and fused systems for these two error types (bias and variance). All of the offline simulations were performed using a MC simulation in order to decrease the high variance that exists in the online accuracy results as described in section 2.4.2.

The data used for the offline simulation were the EEG and eye-gaze data collected during the EEG-alone online experiment; as mentioned previously, eye-gaze was collected in the EEG-alone online experiment even though it was unused in the online setting. The data from the EEG-alone algorithm were used because there were more stimulus presentations on average for the EEG-alone algorithm than any of the other fusion algorithms. First, noise was added to the eye-gaze data depending on the parameters of the experiment being performed (i.e., white noise or bias). For each of the 30 characters spelled during the experiment, the EEG and noisy eye-gaze data used to spell the character was first extracted. Next, the data was sampled with replacement until

the posterior probability of any character reached a threshold of 0.9. An underlying, simulated, checkerboard paradigm was used to determine whether to draw a target or non-target EEG data sample. A simulation of spelling 30 characters was performed 50 times for each simulation condition (1,500 total MC iterations).

For the bias, a horizontal shift was added to each eye gaze observation, represented as a proportion of the total screen. For example, a bias of 0.5 indicates that the gaze is shifted to the right, halfway across the screen. In the second simulation, zero-mean Gaussian noise with an isotropic covariance matrix (i.e., $\mathbf{\Sigma}_{noise} = \begin{bmatrix} \sigma_{noise} & 0 \\ 0 & \sigma_{noise} \end{bmatrix}$ was added to the data. The gaze bias and variance was incrementally increased until the bit-rate (a performance metric incorporating accuracy and spelling speed) of all of the fusion algorithms were below the EEG-alone algorithm, indicating that eye-gaze data was no longer improving the system. As the results presented Figure 5.6 and Figure 5.7 indicate, this happened at a gaze bias just prior to 0.5, or at a noise variance of approximately $\sigma_{noise} = 2$.

5.2.4.1 Experiment 2 Results

Figure 5.6 provides a plot of the bit-rate of the EEG alone algorithm and the fusion algorithm under three eye-gaze variance conditions (estimated, medium, and high variance) as a function of the horizontal gaze bias and additional gaze noise, σ_{noise}^2, averaged over all fifteen participants; the participant whose data was excluded from the online study due to inadvertently using the incorrect classifier was included in the offline

study with the use of that participant's correct classifier. Figure 5.6 (a) shows the bit-rate as a function of gaze bias, and Figure 5.6 (b) shows the bit-rate as a function of σ^2_{noise}. As expected, the EEG alone algorithm results are relatively flat across bias conditions since the algorithm is not dependent on eye gaze. Theoretically, the EEG alone algorithm results should be completely flat, however, the simulation is not deterministic since a MC simulation was used, resulting in some jitter in the bit-rate results. The left-most parts of Figure 5.6 (a) and (b) correspond to bias-free simulations. Figure 5.6 (a) and (b) suggest that the lower the assumed variance in the fusion algorithm, the quicker and steeper the drop off in bit-rate performance as a function of both bias and σ^2_{noise}. This is consistent with the expectation that low variance in the algorithm would be highly susceptible to errors since the fusion algorithm would place a high confidence on eye gaze.

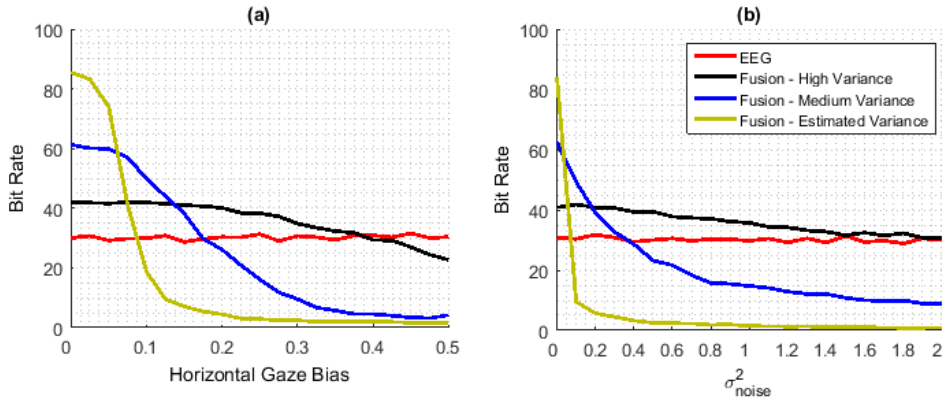

Figure 5.6 - A plot of the bit-rate of the EEG alone algorithm and the fusion algorithm under three eye-gaze variance conditions (estimated, medium, and high variance) as a function of the horizontal gaze bias and additional gaze noise, σ^2_{noise}, averaged over all fifteen participants. (a) shows the bit-rate as a function of gaze bias, and (b) shows the bit-rate as a function of σ^2_{noise}.

Figure 5.7 provides a plot of the bit-rate of the fusion algorithm and the eye-gaze only algorithm with all variance conditions as a function of the horizontal gaze bias and additional gaze noise, σ^2_{noise}, averaged over all fifteen participants Figure 5.7 (a) shows the bit-rate as a function of gaze bias, and Figure 5.7 (b) shows the bit-rate as a function of σ^2_{noise}. For all horizontal gaze biases and σ^2_{noise} the eye-gaze alone algorithms have a lower bitrate than their fusion counterpart. These results show that adding EEG-data improves the bit-rate of the speller regardless of the mismatch between the $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ in training and testing. Furthermore, using eye-gaze data without EEG data is not robust to the training and testing mismatch.

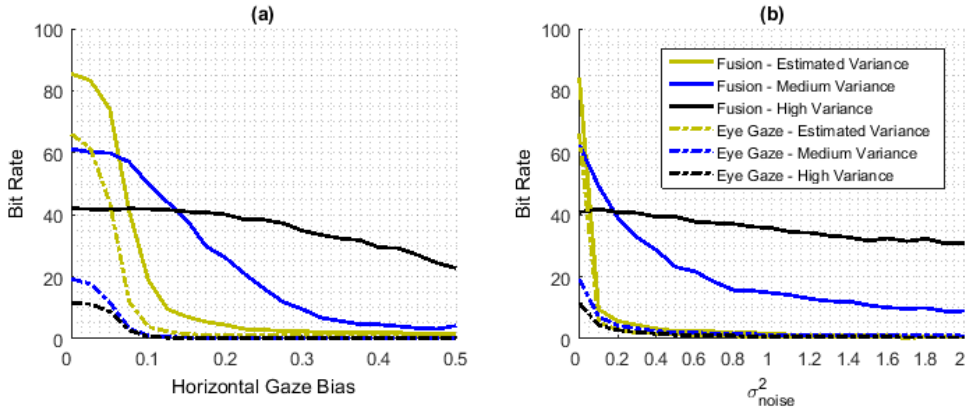

Figure 5.7 - A plot of the bit-rate of the fusion algorithm and the eye-gaze only algorithm with all variance conditions as a function of the horizontal gaze bias and additional gaze noise, σ^2_{noise}, averaged over all fifteen participants. (a) shows the bit-rate as a function of gaze bias, and (b) shows the bit-rate as a function of σ^2_{noise}.

5.2.5 Experiment 1 & 2 Discussion

The work in this chapter explored the possible effectiveness of developing a hybrid P300 speller using both EEG and eye-gaze data. While the results will not transfer directly to the target population since these subjects had near perfect eye gaze, they do provide a mechanism by which to assess efficacy and investigate robustness issues. The fusion algorithm was tested under three gaze-based variance conditions (estimated variance, medium variance, and high variance) and compared to the original EEG-alone, or P300 speller, algorithm.

Online and offline simulation results showed that for non-disabled participants, using eye-gaze in a Bayesian framework with EEG data improved accuracy, reduced the average number of stimulus presentations required to spell a character, and improved bit-rate over the EEG-based Bayesian framework proposed in [72]. As shown in Figure 5.5, the improvements in the bit-rate of the fusion algorithm with all variance conditions were statistically significant compared to the EEG-alone algorithm. In an offline study, noise was added to the eye-gaze data to assess the robustness of the hybrid speller; the fusion algorithm with all variance conditions except for the estimated (i.e. low) eye-gaze variance condition remained robust and outperformed the standard EEG only and eye-gaze only systems in terms of bit-rate.

Examining the offline results shown Figure 5.6 and Figure 5.7 further, it can be seen that when considering the fusion algorithm, the variance conditions have a trade-off

between performance and robustness. The conditions that perform better prior to adding noise tend to perform worse after noise is added. For example, in Figure 5.6(c), without any bias, the estimated-variance condition outperforms (i.e. has a higher bit-rate) all other conditions for both the fusion and eye-gaze only algorithm. However, the estimated-variance condition performs worse than all other conditions after some bias is added. The high-variance condition for the fusion algorithm only performs better than the EEG-alone algorithm prior to the addition of any bias, but performs better than the estimated- and medium-variance fusion algorithm after some bias is added. The high-variance condition for the fusion algorithm performs better than the EEG alone algorithm prior to the addition of any bias, and also performs better than the estimated- and medium-variance fusion algorithm after some bias or σ_{noise}^2 is added. The offline experiments show that the variance conditions used for the speller algorithm can have a significant effect on the bit-rate of the speller; however, the optimal method to determine the variance condition remains unclear and is therefore something worth investigating in the future. All of the results in this chapter were published in [112].

5.3 Future Work

While the work in this chapter proposes estimating the mean locations of the characters on screen using a least-squares method based on the P300 speller training data, other methods may be worth investigating. Rather than estimating the $\boldsymbol{\mu}_m$ parameters, a more obvious approach would be to simply use the position of the characters on the

screen, after the initial calibration using the Tobii software, as the estimate. For example, if the position of character 'A' is known to be [0.1, 0.1], $\boldsymbol{\mu}_{(`A')}$ can be set to [0.1, 0.1]. The advantage of utilizing the method proposed in this work is that eye gaze data can be collected during the training phase of the basic P300 speller, so using that data to estimate character positions ($\boldsymbol{\mu}_m$) would likely improve the accuracy of the estimated character position, particularly if the baseline Tobii calibration has errors.

The disadvantage in using the method proposed in this work is that characters that are far away from characters that were spelled during training can have an inaccurate $\boldsymbol{\mu}_m$. The issue can be observed in Figure 5.3, where the characters at the bottom of the screen appear to have a character mean far from what was expected after the eye-tracker calibration. This can be an issue if the characters spelled during testing are far away from those during training; in the experiments performed in this work, the bottom half of the grid was unused, so it is unlikely that the results were affected. In the future, it is recommended that characters that encompass all sections of the grid are spelled during training.

The work in this chapter assumed that $\boldsymbol{\Sigma}_m$ is identical for all characters. It is, however, known that the precision of the eye-tracker decreases as the participant looks away from the center of the screen, a phenomenon that can be seen in Figure 5.2. Characters 'H', 'I', and 'P' appear to have higher variance in the horizontal direction than characters 'L', 'U', and 'S'. Ideally, $\boldsymbol{\Sigma}_m$ can be learned independently for all characters and

determined based on the characters' position on the screen. One such method that can be used to learn $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ jointly is the Gaussian process regression (GPR) [113]. GPR is a probabilistic, non-linear regression model, that jointly learns regression weights, as well as the variance in those weights. In the case of eye-gaze data, the GPR model can estimate $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ directly when training data is available and estimate $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ based on nearby character data for characters without eye-gaze training data. This enables estimating a $\boldsymbol{\Sigma}_m$ that varies by character. Furthermore, it may be beneficial to examine the use of distributions other than the Gaussian distribution (e.g., student's t or Laplacian) for the eye-gaze likelihood for a character.

6 Conclusion

There were two major contributions made in this thesis towards improving the P300 speller communication device. First, a novel data-driven adaptive stimulus selection paradigm, based on maximizing the EDG metric, was developed. The novel adaptive stimulus selection paradigm was then further improved upon by incorporating a model of refractory effects that exist in the EEG data (see section 2.3.1). Second, an algorithm for a hybrid P300 and eye-gaze speller was developed to potentially improve the communication rate of an individual EEG- or eye-gaze-based speller. The potential benefits of the adaptive stimulus selection paradigm and hybrid speller were explored using both online and offline analyses.

The first contribution of this thesis was to improve the accuracy and decrease the spelling time required for a P300 speller BCI by using an adaptive stimulus presentation paradigm that aims to maximize the amount of information gained after every stimulus presentation. A novel adaptive stimulus presentation framework was designed that greedily selected characters that maximized the EDG metric. In designing the greedy EDG adaptive paradigm, physiological phenomena and system constraints were considered by enforcing various restrictions on the flash groups that can be selected for a stimulus presentation. Offline simulations showed that even with the restrictions set on the greedy EDG adaptive paradigm, it has the potential to yield higher accuracy, and fewer presentations than a CB or RB random stimulus presentation paradigm. Online results

demonstrated a decrease in spelling time and increase in bit-rate when using the greedy EDG adaptive paradigm versus the CB random paradigm. However, the accuracy of the speller suffered when using the greedy EDG adaptive paradigm compared to the CB paradigm.

In addition to exploring the adaptive paradigm, improvements were made to the P300 speller BCI by modeling refractory effects in the target character estimation algorithm. Simulations showed that using TTI information in the target character probability update for the RC paradigm can improve accuracy, EST, and bit-rate significantly. The MC simulations showed that in order to increase the accuracy of the speller when using TTI information, TTI binning should be used to group TTIs when estimating the conditional classifier score PDFs.

An improvement to the greedy EDG adaptive paradigm was also investigated, such that the EDG metric explicitely modelled refractory effects; this paradigm was termed greedy TTI-EDG adaptive paradigm. In offline MC simulations, there were significant improvements in accuracy and EST that were observed in the greedy TTI-EDG adpative paradigm in comparison to the naïve EDG adaptive paradigm when no TTI_{min} restrictions were set. This result indicates that it may not be entirely necessary to set ad-hoc TTI_{min} restrictions if refractory effects are modelled in the Bayesian update equation and EDG.

The second contribution of this thesis was the development of an algorithm for a hybrid speller that uses eye-gaze and EEG data for character selection. Online simulation results showed that for non-disabled participants, using eye-gaze in a Bayesian framework with EEG data improved accuracy, reduced the average number of flashes required to spell a character, and improved bit-rate over the EEG alone Bayesian framework proposed in Throckmorton et al. [72]. In offline simulations, noise was added to the eye-gaze data to assess the robustness of the hybrid speller; and the fusion algorithm with all variance conditions except for the estimated, i.e. low, eye-gaze variance condition remained robust and outperformed the standard EEG alone and eye-gaze alone systems in terms of bit-rate. It was shown that when considering the fusion algorithm, the variance conditions exhibit a trade-off between performance and robustness.

There are many promising avenues for continuing the work presented in this thesis. First, the training methods used for the greedy EDG adaptive paradigm can be improved in an online setting to ensure that the training and testing data distributions are similar when using an adaptive paradigm. A large scale online data collection can then be performed to compare the greedy EDG adaptive paradigm with the improved training methods and a random paradigm. An online data collection can also be performed for the TTI-EDG adaptive paradigm to analyze the potential performance improvement gained by incorporating refractory effects in the stimulus selection paradigm.

Future work can also be pursued in order to investigate the greedy EDG adaptive paradigm that incorporates the eye-gaze data in a hybrid eye-gaze and P300 speller. The adaptive stimulus selection can then be used to select the flash group that maximizes the hybrid EEG and eye-gaze EDG. It will be important to consider the computational complexity of the hybrid EDG calculation as a triple integral will be required for each proposed flash group as the EEG classifier scores and both dimensions of the eye-gaze data need to be marginalized.

Appendix A – Expected Discrimination Gain and Mutual Information

In section 3.2 it was noted that the expression for expected discrimination gain given conditionally independent observations as defined in Kolba [100], is equivalent to the conditional mutual information [99]. The proof for this equivalency will be shown in this appendix. For simplicity, this proof will be shown using the P300 speller framework and notation.

As described in section 3.2, the expected discrimination gain for presentation $t+1$ can be written as follows:

$$\Delta D_{KL}(t+1) \tag{A.1}$$
$$= \int \sum_{\forall m} p(c_m = c_*|z_{t+1}, Z_t) \log\left(\frac{p(c_m = c_*|z_{t+1}, Z_t)}{\mathrm{p}(c_m = c_*|\mathrm{Z}_t)}\right) p(z_{t+1}|Z_t) dz_{t+1}$$

where the notation is identical to the notation used in the document (see sections 2.2.4 and 3.2). Bayes rule states that the expression inside the logarithm can be rewritten as follows:

$$\frac{p(c_m = c_*|z_{t+1}, Z_t)}{p(c_m = c_*|\mathrm{Z}_t)} = \left(\frac{p(z_{t+1}|c_m = c_*, Z_t)p(c_m = c_*|Z_t)}{p(c_m = c_*|Z_t)p(z_{t+1}|Z_t)}\right) = \left(\frac{p(z_{t+1}, c_m = c_*|Z_t)}{p(c_m = c_*|Z_t)p(z_{t+1}|Z_t)}\right) \tag{A.2}$$

Replacing the logarithm in (A.1) with (A.2) yields the following:

$$\Delta D_{KL}(t+1) = \int_{z_{t+1}} \sum_{m} p(c_m = c_*|z_{t+1}, Z_t)\, p(z_{t+1}|Z_t) \log\left(\frac{p(z_{t+1}, c_m = c_*|Z_t)}{p(c_m = c_*|Z_t) p(z_{t+1}|Z_t)}\right) dz_{t+1} \quad \text{(A.3)}$$

Which can be further simplified using the chain rule:

$$\Delta D_{KL}(t+1) = \int_{z_{t+1}} \sum_{m} p(z_{t+1}, c_m = c_*|Z_t) \log\left(\frac{p(z_{t+1}, c_m = c_*|Z_t)}{p(c_m = c_*|Z_t) p(z_{t+1}|Z_t)}\right) dz_{t+1} \quad \text{(A.4)}$$

It can be seen that (A.4) is equivalent to the mutual information [99] between random variables z_{t+1} and c_* given Z_t as shown in (A.5) below:

$$I(z_{t+1}; c_*|Z_t) = \Delta D_{KL}(t+1) = \int_{z_{t+1}} \sum_{m} p(z_{t+1}, c_m = c_*|Z_t) \log\left(\frac{p(z_{t+1}, c_m = c_*|Z_t)}{p(c_m = c_*|Z_t) p(z_{t+1}|Z_t)}\right) dz_{t+1} \quad \text{(A.5)}$$

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, Jun. 2002.

[2] L. Ball, A. Nordness, S. Fager, K. Kersch, G. Pattee, and D. Beukelman, "Eye-gaze access of AAC technology for persons with amyotrophic lateral sclerosis.," *J. Med. Speech Lang. Pathol.*, vol. 18, pp. 11–23, 2010.

[3] CDC, "May is ALS Awareness Month," *Centers for Disease Control and Prevention*, 02-May-2016. [Online]. Available: http://www.cdc.gov/features/alsregistry/index.html. [Accessed: 21-Mar-2017].

[4] "Prevalence of Amyotrophic Lateral Sclerosis — United States, 2010–2011." [Online]. Available: https://www.cdc.gov/mmwr/preview/mmwrhtml/ss6307a1.htm. [Accessed: 18-Jan-2017].

[5] C. Gibbons and E. Beneteau, "Functional Performance Using Eye Control and Single Switch Scanning by People With ALS," *Perspect. Augment. Altern. Commun.*, vol. 19, no. 3, p. 64, Sep. 2010.

[6] C.-S. Hwang, H.-H. Weng, L.-F. Wang, C.-H. Tsai, and H.-T. Chang, "An Eye-Tracking Assistive Device Improves the Quality of Life for ALS Patients and Reduces the Caregivers' Burden," *J. Mot. Behav.*, vol. 46, no. 4, pp. 233–238, Jul. 2014.

[7] E. Pasqualotto *et al.*, "Usability and Workload of Access Technology for People With Severe Motor Impairment A Comparison of Brain-Computer Interfacing and Eye Tracking," *Neurorehabil. Neural Repair*, p. 1545968315575611, Mar. 2015.

[8] R. Spataro, M. Ciriacono, C. Manno, and V. La Bella, "The eye-tracking computer device for communication in amyotrophic lateral sclerosis," *Acta Neurol. Scand.*, vol. 130, no. 1, pp. 40–45, Jul. 2014.

[9] A. Calvo *et al.*, "Eye Tracking Impact on Quality-of-Life of ALS Patients," in *Computers Helping People with Special Needs*, K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, Eds. Springer Berlin Heidelberg, 2008, pp. 70–77.

[10] A. Duchowski, *Eye Tracking Methodology: Theory and Practice*. Springer Science & Business Media, 2007.

[11] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, *Eye Tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[12] P. Majaranta, U.-K. Ahola, and O. Špakov, "Fast Gaze Typing with an Adjustable Dwell Time," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2009, pp. 357–360.

[13] D. M. Stampe and E. M. Reingold, "Selection By Looking: A Novel Computer Interface And Its Application To Psychological Research," in *Studies in Visual Information Processing*, vol. 6, R. W. and R. W. K. John M. Findlay, Ed. North-Holland, 1995, pp. 467–478.

[14] P. A *et al.*, "Eye movement in amyotrophic lateral sclerosis: a longitudinal study.," *Ger. J. Ophthalmol.*, vol. 4, no. 6, pp. 355–362, Nov. 1995.

[15] C. Donaghy, M. J. Thurtell, E. P. Pioro, J. M. Gibson, and R. J. Leigh, "Eye movements in amyotrophic lateral sclerosis and its mimics: a review with illustrative cases," *J. Neurol. Neurosurg. Psychiatry*, vol. 82, no. 1, pp. 110–116, Jan. 2011.

[16] J. Marti-Fàbregas and C. Roig, "Oculomotor abnormalities in motor neuron disease," *J. Neurol.*, vol. 240, no. 8, pp. 475–478, Aug. 1993.

[17] S. Shaunak *et al.*, "Oculomotor function in amyotrophic lateral sclerosis: Evidence for frontal impairment," *Ann. Neurol.*, vol. 38, no. 1, pp. 38–44, Jul. 1995.

[18] A. Palmowski *et al.*, "Eye movement in amyotrophic lateral sclerosis: a longitudinal study," *Ger. J. Ophthalmol.*, vol. 4, no. 6, pp. 355–362, Nov. 1995.

[19] H. E. Moss *et al.*, "Cross-sectional evaluation of clinical neuro-ophthalmic abnormalities in an amyotrophic lateral sclerosis population," *J. Neurol. Sci.*, vol. 314, no. 1–2, pp. 97–101, Mar. 2012.

[20] C. Donaghy, "PATH47 Slow saccades in bulbar-onset motor neuron disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 81, no. 11, pp. e20–e20, Nov. 2010.

[21] E. W. Sellers, T. M. Vaughan, and J. R. Wolpaw, "A brain-computer interface for long-term independent home use," *Amyotroph. Lateral Scler.*, vol. 11, no. 5, pp. 449–455, Oct. 2010.

[22] E. W. Sellers and E. Donchin, "A P300-based brain–computer interface: Initial tests by ALS patients," *Clin. Neurophysiol.*, vol. 117, no. 3, pp. 538–548, Mar. 2006.

[23] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.

[24] E. W. Sellers, A. Kubler, and E. Donchin, "Brain-computer interface research at the university of south Florida cognitive psychophysiology laboratory: the P300 speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 221–224, Jun. 2006.

[25] G. Townsend *et al.*, "A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 121, no. 7, pp. 1109–1120, Jul. 2010.

[26] A. Kübler, A. Furdea, S. Halder, E. M. Hammer, F. Nijboer, and B. Kotchoubey, "A Brain–Computer Interface Controlled Auditory Event-Related Potential (P300) Spelling System for Locked-In Patients," *Ann. N. Y. Acad. Sci.*, vol. 1157, no. 1, pp. 90–100, Mar. 2009.

[27] S. Sutton, M. Braren, J. Zubin, and E. R. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, no. 3700, pp. 1187–1188, Nov. 1965.

[28] B. Z. Allison and J. A. Pineda, "ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 110–113, Jun. 2003.

[29] L. Bianchi, S. Sami, A. Hillebrand, I. P. Fawcett, L. R. Quitadamo, and S. Seri, "Which Physiological Components are More Suitable for Visual ERP Based Brain–Computer Interface? A Preliminary MEG/EEG Study," *Brain Topogr.*, vol. 23, no. 2, pp. 180–185, Apr. 2010.

[30] P. Cipresso *et al.*, "The use of P300-based BCIs in amyotrophic lateral sclerosis: from augmentative and alternative communication to cognitive assessment," *Brain Behav.*, vol. 2, no. 4, pp. 479–498, Jul. 2012.

[31] F. Nijboer *et al.*, "A P300-based brain-computer interface for people with amyotrophic lateral sclerosis," *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 119, no. 8, pp. 1909–1916, Aug. 2008.

[32] S. Silvoni *et al.*, "P300-based brain-computer interface communication: evaluation and follow-up in amyotrophic lateral sclerosis," *Front. Neurosci.*, vol. 3, 2009.

[33] D. B. Ryan, K. A. Colwell, C. S. Throckmorton, L. M. Collins, K. Caves, and E. W. Sellers, "Evaluating Brain-Computer Interface Performance in an ALS Population: Checkerboard and Color Paradigms," *Clin. EEG Neurosci.*, p. 155005941773744, Oct. 2017.

[34] D. B. Ryan, G. Townsend, N. A. Gates, K. Colwell, and E. W. Sellers, "Evaluating brain-computer interface performance using color in the P300 checkerboard speller," *Clin. Neurophysiol.*, vol. 128, no. 10, pp. 2050–2057, Oct. 2017.

[35] Q. Li, S. Liu, J. Li, and O. Bai, "Use of a Green Familiar Faces Paradigm Improves P300-Speller Brain-Computer Interface Performance," *PLOS ONE*, vol. 10, no. 6, p. e0130325, Jun. 2015.

[36] T. Kaufmann, S. M. Schulz, C. Grünzinger, and A. Kübler, "Flashing characters with famous faces improves ERP-based brain–computer interface performance," *J. Neural Eng.*, vol. 8, no. 5, p. 056016, 2011.

[37] B. O. Mainsah, G. Reeves, L. M. Collins, and C. S. Throckmorton, "Optimizing the stimulus presentation paradigm design for the P300-based brain-computer interface using performance prediction," *J. Neural Eng.*, vol. 14, no. 4, p. 046025, 2017.

[38] J. Hill, J. Farquhar, S. Martens, F. Biessmann, and P. B. Schölkopf, "Effects of Stimulus Type and of Error-Correcting Code Design on BCI Speller Performance," in *Advances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds. Curran Associates, Inc., 2009, pp. 665–672.

[39] J. Park and K. E. Kim, "A POMDP Approach to Optimizing P300 Speller BCI Paradigm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 4, pp. 584–594, Jul. 2012.

[40] R. Ma, N. Aghasadeghi, J. Jarzebowski, T. Bretl, and T. P. Coleman, "A Stochastic Control Approach to Optimally Designing Hierarchical Flash Sets in P300 Communication Prostheses," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 1, pp. 102–112, Jan. 2012.

[41] K. Kastella, "Discrimination gain to optimize detection and classification," *IEEE Trans. Syst. Man Cybern. - Part Syst. Hum.*, vol. 27, no. 1, pp. 112–116, Jan. 1997.

[42] M. G. Woldorff, "Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction," *Psychophysiology*, vol. 30, no. 1, pp. 98–119, Jan. 1993.

[43] S. M. M. Martens, N. J. Hill, J. Farquhar, and B. Schölkopf, "Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential," *J. Neural Eng.*, vol. 6, no. 2, p. 026003, 2009.

[44] R. Fazel-Rezai, "Human Error in P300 Speller Paradigm for Brain-Computer Interface," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 2516–2519.

[45] B. Z. Allison *et al.*, "Toward smarter BCIs: extending BCIs through hybridization and intelligent control," *J. Neural Eng.*, vol. 9, no. 1, p. 013001, 2012.

[46] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general," *J. Neural Eng.*, vol. 8, no. 2, p. 025005, 2011.

[47] X. Yong, M. Fatourechi, R. K. Ward, and G. E. Birch, "The Design of a Point-and-Click System by Integrating a Self-Paced Brain #x2013;Computer Interface With an Eye-Tracker," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 1, no. 4, pp. 590–602, Dec. 2011.

[48] T. O. Zander, M. Gaertner, C. Kothe, and R. Vilimek, "Combining eye gaze input with a brain–computer interface for touchless human–computer interaction," *Intl J. Human–Computer Interact.*, vol. 27, no. 1, pp. 38–51, 2010.

[49] H. J. Hwang, J. H. Lim, J. H. Lee, and C. H. Im, "Implementation of a mental spelling system based on steady-state visual evoked potential (SSVEP)," in *2013 International Winter Workshop on Brain-Computer Interface (BCI)*, 2013, pp. 81–83.

[50] J.-H. Lim, J.-H. Lee, H.-J. Hwang, D. H. Kim, and C.-H. Im, "Development of a hybrid mental spelling system combining SSVEP-based brain–computer interface and webcam-based eye tracking," *Biomed. Signal Process. Control*, vol. 21, pp. 99–104, Aug. 2015.

[51] P. McCullagh, L. Galway, and G. Lightbody, "Investigation into a Mixed Hybrid Using SSVEP and Eye Gaze for Optimising User Interaction within a Virtual Environment," in *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion*, C. Stephanidis and M. Antona, Eds. Springer Berlin Heidelberg, 2013, pp. 530–539.

[52] L. Galway, C. Brennan, P. McCullagh, and G. Lightbody, "BCI and Eye Gaze: Collaboration at the Interface," in *Foundations of Augmented Cognition*, D. D. Schmorrow and C. M. Fidopiastis, Eds. Springer International Publishing, 2015, pp. 199–210.

[53] "OpenBCI - Open Source Biosensing Tools (EEG, EMG, EKG, and more)." [Online]. Available: http://openbci.com/. [Accessed: 19-Jul-2016].

[54] X. Dong, H. Wang, Z. Chen, and B. E. Shi, "Hybrid Brain Computer Interface via Bayesian integration of EEG and eye gaze," in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2015, pp. 150–153.

[55] J.-S. Choi, J. W. Bang, K. R. Park, and M. Whang, "Enhanced Perception of User Intention by Combining EEG and Gaze-Tracking for Brain-Computer Interfaces (BCIs)," *Sensors*, vol. 13, no. 3, pp. 3454–3472, Mar. 2013.

[56] P. Majaranta and K.-J. Räihä, "9. Text Entry by Gaze: Utilizing Eye-Tracking," *ResearchGate*, Jan. 2007.

[57] L. A. Frey, K. P. White, and T. E. Hutchison, "Eye-gaze word processing," *IEEE Trans. Syst. Man Cybern.*, vol. 20, no. 4, pp. 944–950, Jul. 1990.

[58] D. W. Hansen, J. P. Hansen, M. Nielsen, A. S. Johansen, and M. B. Stegmann, "Eye typing using Markov and active appearance models," in *Sixth IEEE Workshop on Applications of Computer Vision, 2002. (WACV 2002). Proceedings*, 2002, pp. 132–136.

[59] J. P. Hansen, D. W. Hansen, and A. S. Johansen, "Bringing gaze-based interaction back to basics," in *ResearchGate*, 2001, pp. 325–329.

[60] T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Trans. Syst. Man Cybern.*, vol. 19, no. 6, pp. 1527–1534, Nov. 1989.

[61] P. Majaranta, I. S. MacKenzie, A. Aula, and K.-J. Räihä, "Effects of feedback and dwell time on eye typing speed and accuracy," *Univers. Access Inf. Soc.*, vol. 5, no. 2, pp. 199–208, Jul. 2006.

[62] P. Majaranta and K.-J. Räihä, "Twenty Years of Eye Typing: Systems and Design Issues," in *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, New York, NY, USA, 2002, pp. 15–22.

[63] "How do Tobii Eye Trackers work?," 10-Aug-2015. [Online]. Available: http://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/. [Accessed: 02-Nov-2016].

[64] "American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. [Review]," *J. Clin. Neurophysiol.*, vol. 8, no. 2, pp. 200–202, Apr. 1991.

[65] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward Enhanced P300 Speller Performance," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 15–21, Jan. 2008.

[66] D. J. Krusienski *et al.*, "A comparison of classification techniques for the P300 Speller," *J. Neural Eng.*, vol. 3, no. 4, pp. 299–305, Dec. 2006.

[67] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components — A tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, May 2011.

[68] D. G. Childers, N. W. Perry, I. A. Fischler, T. Boaz, and A. A. Arroyo, "Event-related potentials: a critical review of methods for single-trial detection," *Crit. Rev. Biomed. Eng.*, vol. 14, no. 3, pp. 185–200, 1987.

[69] T. C. Handy, *Event-related Potentials: A Methods Handbook*. MIT Press, 2005.

[70] M. Schreuder, J. Höhne, B. Blankertz, S. Haufe, T. Dickhaus, and Michael Tangermann, "Optimizing event-related potential based brain–computer interfaces: a systematic evaluation of dynamic stopping methods," *J. Neural Eng.*, vol. 10, no. 3, p. 036025, 2013.

[71] P.-J. Kindermans, M. Tangermann, K.-R. Müller, and B. Schrauwen, "Integrating dynamic stopping, transfer learning and language models in an adaptive zero-training ERP speller," *J. Neural Eng.*, vol. 11, no. 3, p. 035005, 2014.

[72] C. S. Throckmorton, K. A. Colwell, D. B. Ryan, E. W. Sellers, and L. M. Collins, "Bayesian approach to dynamically controlling data collection in P300 spellers," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 21, no. 3, pp. 508–517, May 2013.

[73] B. O. Mainsah, K. A. Colwell, L. M. Collins, and C. S. Throckmorton, "Utilizing a Language Model to Improve Online Dynamic Data Collection in P300 Spellers," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 837–846, Jul. 2014.

[74] J. M. Clements, E. W. Sellers, D. B. Ryan, K. Caves, L. M. Collins, and C. S. Throckmorton, "Applying dynamic data collection to improve dry electrode system performance for a P300-based brain–computer interface," *J. Neural Eng.*, vol. 13, no. 6, p. 066018, 2016.

[75] W. Speier, N. Chandravadia, D. Roberts, S. Pendekanti, and N. Pouratian, "Online BCI typing using language model classifiers by ALS patients in their homes," *Brain-Comput. Interfaces*, vol. 4, no. 1–2, pp. 114–121, Apr. 2017.

[76] W. Speier, A. Deshpande, L. Cui, N. Chandravadia, D. Roberts, and N. Pouratian, "A comparison of stimulus types in online classification of the P300 speller using language models," *PLOS ONE*, vol. 12, no. 4, p. e0175382, Apr. 2017.

[77] V. Bostanov, "BCI competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1057–1061, Jun. 2004.

[78] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1073–1076, Jun. 2004.

[79] A. Rakotomamonjy and V. Guigue, "BCI Competition III: Dataset II- Ensemble of SVMs for BCI P300 Speller," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1147–1154, Mar. 2008.

[80] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang, "BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1067–1072, Jun. 2004.

[81] U. Orhan, K. E. Hild, D. Erdogmus, B. Roark, B. Oken, and M. Fried-Oken, "RSVP keyboard: An EEG based typing interface," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 645–648.

[82] B. O. Mainsah, "Adaptive Brain-Computer Interface Systems For Communication in People with Severe Neuromuscular Disabilities," 2016.

[83] V. Mayya, B. Mainsah, and G. Reeves, "Information-Theoretic Analysis of Refractory Effects in the P300 Speller," *ArXiv170103313 Cs Math*, Jan. 2017.

[84] S. M. M. Martens, J. M. Mooij, N. J. Hill, J. Farquhar, and B. Schölkopf, "A Graphical Model Framework for Decoding in the Visual ERP-Based BCI Speller," *Neural Comput.*, vol. 23, no. 1, pp. 160–182, Oct. 2010.

[85] J. Geuze, J. D. R. Farquhar, and P. Desain, "Dense codes at high speeds: varying stimulus properties to improve visual speller performance," *J. Neural Eng.*, vol. 9, no. 1, p. 016009, 2012.

[86] T. Verhoeven, P. Buteneers, J. R. Wiersema, J. Dambre, and P. J. Kindermans, "Towards a symbiotic brain–computer interface: exploring the application–decoder interaction," *J. Neural Eng.*, vol. 12, no. 6, p. 066027, 2015.

[87] L. Citi, R. Poli, and C. Cinel, "Exploiting P300 amplitude variations can improve classification accuracy in Donchin's BCI speller," in *2009 4th International IEEE/EMBS Conference on Neural Engineering*, 2009, pp. 478–481.

[88] J. Jin, E. W. Sellers, and X. Wang, "Targeting an efficient target-to-target interval for P300 speller brain–computer interfaces," *Med. Biol. Eng. Comput.*, vol. 50, no. 3, pp. 289–296, Feb. 2012.

[89] C. Polprasert, P. Kukieattikool, T. Demeechai, J. A. Ritcey, and S. Siwamogsatham, "New stimulation pattern design to improve P300-based matrix speller performance at high flash rate," *J. Neural Eng.*, vol. 10, no. 3, p. 036012, 2013.

[90] N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide*. Taylor & Francis, 2004.

[91] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface," *Biol. Psychol.*, vol. 102, pp. 118–129, Oct. 2014.

[92] M. A. S. Boksem, T. F. Meijman, and M. M. Lorist, "Effects of mental fatigue on attention: An ERP study," *Cogn. Brain Res.*, vol. 25, no. 1, pp. 107–116, Sep. 2005.

[93] S. K. L. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biol. Psychol.*, vol. 55, no. 3, pp. 173–194, Feb. 2001.

[94] S. Makeig and T.-P. Jung, "Tonic, phasic, and transient EEG correlates of auditory awareness in drowsiness," *Cogn. Brain Res.*, vol. 4, no. 1, pp. 15–25, Jul. 1996.

[95] A. Aref and J. Huggins, "Combining Methods To Predict Accuracy of Individual Brain-Computer Interface Selections," in *Proceedings of the 6th International Brain-Computer Interface Meeting*, Asilomar, CA.

[96] K. Colwell, C. Throckmorton, L. Collins, and K. Morton, "Projected Accuracy Metric for the P300 Speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 5, pp. 921–925, Sep. 2014.

[97] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, "Brain–computer interface (BCI) operation: optimizing information transfer rates," *Biol. Psychol.*, vol. 63, no. 3, pp. 237–251, Jul. 2003.

[98] C. Z. Mooney, *Monte Carlo Simulation*. SAGE Publications, 1997.

[99] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, 2 edition. Hoboken, N.J: Wiley-Interscience, 2006.

[100] M. P. Kolba, "Information-Based Sensor Management for Static Target Detection Using Real and Simulated Data," 2009.

[101] C. Cai, "Information-driven sensor path planning and the treasure hunt problem," Duke University, 2008.

[102] C. Guan, M. Thulasidas, and J. Wu, "High performance P300 speller for brain-computer interface," in *IEEE International Workshop on Biomedical Circuits and Systems, 2004.*, 2004, p. S3/5/INV-S3/13-16.

[103] W. Klonowski, "Everything you wanted to ask about EEG but were afraid to get the right answer," *Nonlinear Biomed. Phys.*, vol. 3, p. 2, May 2009.

[104] D. A. Balota *et al.*, "The English Lexicon Project," *Behav. Res. Methods*, vol. 39, no. 3, pp. 445–459, Aug. 2007.

[105] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biom. Bull.*, vol. 1, no. 6, pp. 80–83, 1945.

[106] D. Kalika, L. M. Collins, C. Throckmorton, and B. O. Mainsah, "Adaptive Stimulus Selection in ERP-Based Brain-Computer Interfaces by Maximizing Expected Discrimination Gain," presented at the IEEE Systems, Man, and Cybernetics (SMC, Banff, Alberta, Ca, 2017.

[107] S. M. M. Martens and J. M. Leiva, "A generative model approach for decoding in the visual event-related potential-based brain–computer interface speller," *J. Neural Eng.*, vol. 7, no. 2, p. 026003, 2010.

[108] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. 2006. Corr. 2nd printing. Springer, 2007.

[109] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, "Does the 'P300' speller depend on eye gaze?," *J. Neural Eng.*, vol. 7, no. 5, p. 056013, 2010.

[110] D. Ryan, "Improving Brain-Computer Interface Performance: Giving the P300 Speller Some Color.," *Electron. Theses Diss.*, Aug. 2011.

[111] K. Lund and C. Burgess, "Producing high-dimensional semantic spaces from lexical co-occurrence," *Behav. Res. Methods Instrum. Comput.*, vol. 28, no. 2, pp. 203–208, Jun. 1996.

[112] D. Kalika, L. M. Collins, K. Caves, and C. S. Throckmorton, "Fusion of P300 and Eye-tracker Data for Spelling Using BCI2000 (Submitted)," *J. Neural Eng.*

[113] C. E. Rasmussen, "Gaussian processes for machine learning," 2006.

Biography

Dmitry Kalika was born in Moscow, Russia. He received a Bachelor of Science degree in Electrical and Computer Engineering at the University of Pittsburgh, Pittsburgh, PA, USA in 2012. He then went on to receive a Master of Science Degree in Electrical and Computer Engineering at Duke University, Durham, NC, USA, under the supervision of Dr. Leslie Collins, in 2014. He was the recipient of the Wireless Intelligent Sensor Networks (WISeNet) NSF IGERT fellowship in 2014-2016. He received a Doctorate of Philosophy degree in Electrical and Computer Engineering from Duke University in January of 2018.

He has authored the following journal and conference publications:

[1] D. Kalika, L. M. Collins, C. Throckmorton, and B. O. Mainsah, "Adaptive Stimulus Selection in ERP-Based Brain-Computer Interfaces by Maximizing Expected Discrimination Gain," presented at the IEEE Systems, Man, and Cybernetics (SMC), Banff, Alberta, Ca, 2017.

[2] Kalika, D., Collins, L. M, Caves, K., & Throckmorton, C. S. (2017). "Fusion of P300 and eye-tracker data for spelling using BCI2000". Journal of Neural Engineering.

[3] Kalika, D., Knox, M. T., Collins, L. M., Torrione, P. A., & Morton, K. D. (2015, May). "Leveraging robust principal component analysis to detect buried explosive threats

in handheld ground-penetrating radar data". In SPIE Defense+ Security (pp. 94541D-94541D). International Society for Optics and Photonics.

[4] Kalika, D., Morton, K. D., Collins, L. M., & Torrione, P. A. (2014, May). "Hyperbolic and PLSDA In SPIE Defense+ Security. International Society for Optics and Photonics.

[5] Sejdic, E., Kalika, D., & Czarnek, N. (2013). "An analysis of resting-state functional transcranial doppler recordings from middle cerebral arteries". PloS one.

He has presented the following posteriors at a conference:

[1] Kalika, D., Collins, L. M., Caves, K, & Throckmorton, C. S., "Fusion of P300 and eye tracker data for spelling using BCI2000", in Proceedings of the 6th International Brain-Computer Interface Meeting, organized by the BCI Society, Asilomar, CA, 2016.

[2] Kalika, D., Collins, L. M., & Throckmorton, C. S., "Fusion of EEG and eye tracking for spelling", in the International Conference on Basic and Clinical Multimodal Imaging, Utrecht, Netherlands, 2015.