A Functional-Near Infrared Spectroscopy Investigation of Mental Workload

by
Srinidhi Parshi

A Thesis submitted to the Department of Electrical and Computer Engineering,

Cullen College of Engineering

in partial fulfillment of the requirements for the degree of

Master of Science
in Computer and Systems Engineering

Chair of Committee: Rose T Faghih

Committee Member: David Mayerich

Committee Member: Yuhua Chen

University of Houston
May 2020

DEDICATION/EPIGRAPH

This thesis is dedicated to,

my father; for the unwavering faith and support he has shown me,

my loving mother; to whom I owe my intellectual acumen,

and my little brother who has been the light and joy of my life.

Srinidhi Parshi

May 2020

ACKNOWLEDGMENTS

I would like to express my deepest respect and gratitude to my advisor, Dr.Rose Faghih, whose knowledge and guidance have helped me successfully complete my journey to a Master's Thesis degree. The Computational Medicine Laboratory has been a source of strength and support for me during the last two years. As such, I would like to thank her for providing this platform to me even though I had no prior research experience. In spite of my constant misinterpretations and several errors on my part, she has been patient and understanding, for which I am eternally grateful. During the period when I experienced an intense longing for my home back in India, she has shown me acceptance and gentle guidance, which have inspired me to pursue my degree more enthusiastically. Her expansive knowledge, coupled with her professional guidance has helped me realize my potential, and it is my dearest hope that I can live up to the standards she has taught me.

I also would like to extend my sincere gratitude towards my mentors Md. Rafiul Amin, and Hamid Fekhri Azgomi. They have helped me shape my research, and have always provided expert and invaluable advice on how to write professional research papers. It turns out, that writing research-grade papers is my kryptonite. So I'm glad to have mentors with such fortitude. They have always provided me with timely and professional feedback, and I'm indebted to them. I consider myself very fortunate to have known them and worked with them. I would also like to thank Dilranjan Wickramasuriya for helping me polish the experiment, and providing crucial feedback for my papers and presentations. I would also like to acknowledge the support given by my labmates Divesh Deepak Pednekar, Jon Genty, Luciano Branco, and Sankalp Parekh.

My family has provided me with constant support and their faith in me has inspired me to achieve new heights. I cannot even begin to express my thanks to my father, Padmanabha Rao, and my mother Revathi. My father's belief in me has helped me achieve my dreams, and I can only hope that I have made him proud. My mother's work ethic has always pushed me to test my limits. Her love and encouragement play a great role in my achievements.

I'm fortunate to have an affectionate little brother, Pradyumna, who thinks the world of me, and whose love and support have been the light and joy in my life. I would also like to thank my grandmother, aunts, uncles, and cousin brothers for all of their warmhearted efforts of encouragement.

I would like to acknowledge the guidance of my mentor Professor K.R.Sarma, whose lectures on electronics have inspired me to pursue higher education and a research-oriented degree. I would like to thank my dearest friends Ashwini Chaudhari, Manojna Sistla, and Nikitha Kakani for their incredible support. I have relied on them in my lowest points, and I'm grateful to have friends who can pull me out of endless spirals of anxiety. I would also like to thank all my friends in both India and Houston, for their support and encouragement. Last, but not least I would like to express my sincere thanks to my colleagues and manager at University IT. They have actively motivated me to keep pursuing my dreams and I'm truly blessed to have such positive mentors in them.

ABSTRACT

Mental workload assessment is a critical aspect of human-computer interfaces. Mental workload which is either underwhelming or overwhelming will result in a negative influence on one's performance. Enhancing existing classification techniques of mental workload holds great potential in improving our understanding of cognitive engagement. Accurate detection of mental workload can have wide-ranging applications such as improving cognitive engagement or cognitive workload during an online class, which in turn affects the student's performance and learning. In this research, we utilize functional near-infrared spectroscopy (fNIRS) to obtain insights into the dynamic functional connectivity of the brain as a function of the mental workload. Variations in the brain's blood oxygenation and deoxygenation reflect neuronal activation patterns and can be measured using fNIRS. Interpreting connectivity in the brain using noisy fNIRS data with low signal to noise ratio is challenging. To overcome the challenges with fNIRS data, we use a hierarchical latent dictionary learning approach. This approach provides covariance matrices to obtain the dynamic functional connectivity and neuronal activation patterns that change over time. We use features from the dynamic functional connectivity of the brain reflected in fNIRS data collected from the prefrontal cortex to investigate mental workload. An analysis of two different datasets, which use memory task consisting of different levels of difficulty, is presented. Covariance matrices are obtained for each difficulty level of the memory-based task by modeling the underlying neural signals as Gaussian processes and applying a latent factor model to the observed hemodynamic data, which in turn reflects the dynamic functional connectivity of the brain. The obtained covariance matrices are inputted to three machine learning algorithms, namely, Support Vector Machines, K-Nearest Neighbors, and Linear Discriminant Analysis, in order to evaluate classification accuracies in discerning the levels of difficulty. This hierarchical latent dictionary learning approach is implemented on an open-access dataset as well as a novel memory-related mental workload experiment conducted as part of this research. To elucidate the effects of music on mental workload, the novel experiment included calming and vexing music sessions. Our classification accuracies

verify the viability of hierarchical latent dictionary learning approach to obtain functional connectivity and also expound the effects of music on mental workload.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

1.1 Quantifying Mental Workload: Impact and General Assessment

1.1.1 Significance of Mental Workload and its Assessment

Mental workload, or cognitive workload is an important aspect of ergonomics and comprehending human factors that influence performance [2]. Intuitively, it can be understood as the amount of mental work required for a person to complete a task in a specified interval of time. However, it is not an isolated property, rather it is a result of interaction required while performing the task, skills, environment, behaviors and other influences [3]. Due to the dependence on several factors, different people exhibit different ability to handle mental workload while performing tasks. For instance, consider the interaction between a human and a computer also known as Human Computer Interactions (HCI). This interaction can be considered as a task, or a mental workload. The technological advancement of recent years has led to the ubiquitous nature of HCI. Therefore, an important aspect of designing a successful HCI that can execute its task efficiently, is to improve the performance of the system and its human operator [4]. In recent years, the prevalence of HCI and the advancements in neuroscience have led to an emerging new field of Brain Computer Interfaces (BCI) where computers are given input commands via neural signals [5]. Seamless integration and interaction between the BCI system and operator leads to improvised performance and efficiency, and one of the factors that heavily influences this performance is the mental workload or the mental cost of operating the system [6]. Due to several elements such as prior experience, aptitude, stress, and many external factors, an individual's quality of interaction with the computer may decrease, leading to negative performance metrics. This is true for both underload and an overload of mental workload. An underload leads to feelings of boredom, whereas an overload leads to increased frustration, thereby increasing chances of committing errors [3]. Thus, obtaining accurate insights into the mental workload provides essential understanding of whether changes in workload lead to a reduced

performance [7]. Such an understanding of the dynamics of mental workload and its influence on performance lead to improved working conditions, aimed at optimal performance output.

There are several ways to assess the effects of mental workload on a task. Some of the most well-studied effects are workload experienced by an individual while driving. In the study conducted in [8], the authors have assessed mental workload while driving using a driving simulator. Research in [9], is another assessment of mental workload on visual search and decision making while driving, and it provides essential knowledge in the deteriorating effects of performing tasks while driving. Other studies such as [10] evaluate the cognitive workload experienced by air-traffic controllers. Most recently, there have been several studies conducted to understand the affect of mental workload on BCIs [11]. Roy *et al.* [12] have explored the workload induced by a working memory experiment, called n-back task, to be utilized in BCI. Due to its widespread use in BCIs, we analyze the n-back task and the mental workload that it induces.

Apart from detecting and understanding workload, several studies have explored the idea of introducing music to reduce the mental workload, thereby increasing performance. Research conducted in [13] and [14] utilizes music tracks to influence driving performance, and [15], [16] shows the positive influence of music on concentration and the viability of using music to reduce stress. In this thesis, we are interested in mental workload and its impact on HCIs, and more specifically, BCIs. There exists a gap in how music influences working memory, therefore we detail an experiment conducted to analyze the effects of different types of music on mental workload, and performance of users in n-back tasks.

1.1.2 A Brief Summary of Existing Mental Workload Detection and Analysis Frameworks

In addition to the various experiments developed to assess mental workload, recent studies have analyzed and developed methods to measure brain activity using a variety of

neuroimaging technologies such as Magnetoencephalography (MEG), Electroencephalography (EEG), functional Magnetic Resonance Imaging (fMRI) and functional Near Infrared Spectroscopy (fNIRS) [17], [18], [19], [20]. Though [21] and [22] show that a variety of physiological signals that can be used to determine mental workload, neuroimaging techniques such as fNIRS, and the recent fNIRS-EEG hybrid systems remain popular tools to quantify it [1], [23]. To accurately analyze how different regions in the brain process working memory task having different mental workload, we need to utilize a neuroimaging technique that has relatively higher spatial resolution. When compared with EEG, fNIRS has more spatial resolution, and fNIRS-EEG hybrid systems are not as portable and easy to implement as simple fNIRS headcaps. Furthermore, previous functional imaging studies of n-back task using fMRI [24], show significant spatio-temporal changes for working memory tasks in the prefrontal cortex using fNIRS. Later analysis such as work in [1] confirm that fNIRS is uniquely suited to assess mental workload, especially in the prefrontal cortex region of the brain. An fNIRS headcap is easy to wear and it measures the hemodynamic changes within minutes, while guaranteeing high data quality. Thus, fNIRS-based measurements of mental workload in memory tasks are ideal for a passive BCI used to monitor workload levels.

Mental workload classification is improved as a result of better algorithms and toolboxes developed to analyze neuroimaging data. One such toolbox which uses data from fNIRS to construct a functional connectivity map of the brain is the Brain AnalyzIR [25]. Functional connectivity pertains to the covariances between different brain regions during a mental workload task. The Brain AnalyzIR utilizes the generalized linear models (GLM) technique to model the fNIRS data as a linear regression [25]. This regression model in time can be used to extract the functional connectivity maps of the brain during the task. However, the Brain AnalyzIR toolbox does not model the time evolving or *dynamic* functional connectivity of the brain. Thus, there exists a gap in our knowledge of understanding how the mental workload affects different regions in the brain as it processes information. A model that can describe this dynamic functional connectivity, while incorporating the

underlying neural signals can vastly improve our understanding of how the brain processes cognitive stress, and it can lead to better classification level of workloads on the brain. Furthermore, a model that builds these connectivity maps from an underlying neural signal can inherently incorporate the similarities between two trials of the same task, and the differences between two trials of two separate tasks. To capture time evolving functional connectivity and to reduce the dimensionality, one could use dynamic latent factor Model. However, such a model assumes Markov dynamics which lacks long range dependencies. In Markov dynamics based models, the time-invariance nature of the state transition matrix yields homoscedasticity (i.e. the covariance does not evolve in time). To consider the aforementioned aspects of the brain recordings, Fysche *et al.* [17] proposed to use Hierarchical Latent Dictionary Model (HLDM) in MEG data. Inspired by their approach, we propose to model the fNIRS recordings with HLDM approach. The main features are extracting the time evolving dynamic functional connectivity of the brain and including the influence of the underlying neural signal specific to a task irrespective of the trials. Thus, using HLDM we extract the time evolving dynamic functional connectivity maps of the brain which can be used to better understand how the brain processes specific tasks.

1.2 Thesis Outline

This thesis presents an approach to obtain functional connectivity maps from fNIRS data to improve our understanding of mental workload. The first part focusses on our novel experiment that is conducted for recording neurophysiological data during memory tasks in presence of music. The goal is to investigate the neurophysiological response corresponding to different mental workload in presence of different types of music. The second part details obtaining dynamic functional connectivity on both an open access dataset and those obtained from our experiment. A comparison of the classification accuracies thus obtained is also presented.

1.2.1 A Working Memory Neurophysiological Dataset Incorporating the Influence of Music

An underload or an overload of mental workload can directly result in reduced performance in everyday life. Several studies have explored the idea of introducing music to improve performance of various tasks, driving performance, and concentration during various cognitive engagement tasks [13], [14], [15], [16]. Music especially has been shown to influence workload in a variety of HCIs. Studies such as [15], [16] use specific classical tracks to discern mental stress in cognitive engagement tasks. However, similar music tracks might not effect different individuals equally, since taste in music can be varied from person to person. Experiments conducted in [13], [14] use pre-selected participant specific music tracks to be played during a driving task. They show the positive influence of music on concentration and the viability of using music to reduce stress. Influence of music has also been explored in working memory tasks [26] by listening to specific classical tracks before performing the task. Several people utilize music background as an effective means while performing cognitive tasks (e.g. learning, driving). To understand how music influences cognitive performance, we designed a novel experiment with working memory tasks similar to [1] under two different background music influences, i.e. calming and vexing specific to each participant. Though physiological data such as electrodermal activity and heart rate have been used before in conjunction with neuroimaging data to operate BCI [22], [21], a collection of all the physiological signals to form a more comprehensive understanding of stress during a mental workload task has not been conducted. It has been shown that various biological signals such as skin conductance, and neural spiking activity can be used to assess the emotional and cognitive state of the brain [27], [28], [29], [30], [31], [32]. Therefore, to gain a more comprehensive understanding of mental workload we collect data from a broad variety of physiological signals in addition to fNIRS. This chapter details the methods of collecting the experiment, along with the protocols and setup followed.

1.2.2 Experimental Validation of Mental Workload Assessment Using Hierarchical Latent Dictionary Method

Enhancing existing models and classification techniques of mental workload holds great potential in improving our understanding of cognitive engagement and user performance. In this chapter, we propose to model the fNIRS recordings with a latent factor model in order to understand and model the neural signal aspects, where factors are modeled as Gaussian processes, during a mental workload experiment. fNIRS is a method that determines neuronal activation patterns by detecting the amount of oxygenated and deoxygenated hemoglobin levels in the brain. In this chapter, we propose to model fNIRS data with HLDM for analysis so that we can capture the dynamic functional connectivity during different mental workload tasks. We first obtain the functional connectivity maps on a publicly available fNIRS dataset. Secondly, we analyze our experimental dataset that uses fNIRS on n-back memory task to explore the influence of two different types of personalized music, i.e. calming and vexing, on cognitive stress.

By modeling the underlying neural signal factors as Gaussian processes and applying the latent factor model to the observed hemodynamic data, we obtain a covariance matrix representing functional connectivity of a difficulty level of the n-back task, which in turn reflects the dynamic functional connectivity of the brain. We then utilize the extracted covariance matrices as inputs to machine learning algorithms namely, Support Vector Machines (SVM), K-Nearest Neighbors (KNN) and Linear Discriminant Analysis (LDA) to evaluate classification accuracies in distinguishing different levels of difficulty and observe the affects of music on the performance of these tasks. Our classification accuracies verify the viability of the model. The functional connectivity in mental workloads can be applied to better understand cognitive engagement and stress, and could result in improved HCI design.

2 A Working Memory Neurophysiological Dataset Incorporating the Influence of Music

2.1 A Synopsis of Mental Workload Experiments

Mental workload is the measure of cognitive stress during varied working environments. Performing different mental workload tasks requires different levels of cognitive engagement and memory resources in the brain [24]. Recently, several works have explored this idea of quantifying mental workload by implementing different neuroimaging techniques on participants performing cognitive engagement tasks [1], [23]. Popular neuroimaging techniques to detect cognitive stress are functional Near Infrared Spectroscopy (fNIRS), which detects blood oxygenation levels by distinguishing absorption spectrum of oxygenated and deoxygenated blood, and Electroencephalography (EEG), which detects electrical signals of neuronal activation in the brain. Herff *et al.* [1] performed a mental workload experiment where they collected fNIRS dataset. On the other hand, Shin *et al.* [23] incorporated EEG along with fNIRS in their experiment to retain both temporal and spatial resolution. However, hybrid fNIRS-EEG systems have the added effect of inducing participant fatigue due to the prolonged experimental duration. For these reasons, experiments conducted with simple fNIRS headcaps are preferred. Research conducted in [22] and [21] use physiological signal such as heart rate and skin conductance to assess mental workload. But, a complete dataset which includes multiple physiological signals along with neuroimaging methods that can be used to gain a more thorough understanding of the body's response to cognitive stress, is absent. In this research, we present a dataset obtained by conducting an experiment that collects data from multiple physiological signals along with fNIRS to further our understanding of the affects of mental workload.

Apart from assessing mental workload using different neuroimaging techniques, several studies have explored different means of mitigating the effects of cognitive stress and improve performance. One factor which has been researched extensively is the influence of

music in working environments, as it has been shown to have quantifiable effect in stress reduction [33]. Several studies have explored the idea of introducing music to improve performance while doing various tasks such as driving, and concentration during various cognitive engagement tasks [13], [14], [15], [16] [34]. The results in these studies show the positive influence of music on concentration and its viability to reduce stress. While [15], [16] used specific classical tracks to discern mental stress for all participants, studies in [13], [14] have personalized tracks chosen from a playlist for each participant. Personalizing music tracks for each participant ensures that the music does have the desired calming or vexing effect on them. For this reason, we have included participant-specific tracks as the background music. Though several studies have been conducted on the influence of mental workload while performing tasks such as driving and arithmetic tasks, its influence on working memory has not been explored. Therefore, in the experiment performed, we include the effects of two different sessions of music namely; calming and vexing music, and discern their influence on cognitive stress. Similar to [13], [14], participants in this experiment chose personal instrumental music tracks for both their calming and vexing music sessions.

The following sections detail the experimental protocol, and setup along with preliminary visualization and a discussion of data obtained.

2.2 Methods

2.2.1 Experimental Setup

The experimental setup used to record physiological data is described in this section.

Functional Near Infrared Spectroscopy: The NIRSport 2 noninvasive sensors are placed according to the positions depicted in Figure 1 on a head cap worn by the participant. Near infrared region of the electromagnetic spectrum (620-1000nm) is scattered by biological tissue, but it is absorbed by hemoglobin [35]. Using the amount of absorbed near infrared light and modified Beer Lambert Law [36], fNIRS estimates oxygenated and deoxygenated hemoglobin (HbO and HbR, respectively). This hemodynamic response reflects neuronal

activation patterns of the brain and fNIRS uses this phenomenon to estimate the levels of HbO and HbR using modified Beer Lambert Law [36]. fNIRS demonstrates excellent spatial resolution with the downside of having poor temporal resolution [37]. We use the spatial resolution of fNIRS to obtain the functional connectivity of the different regions of brain. Optodes of fNIRS can be placed according to the 10-5 international system [38] such that measurements can be taken from the entire scalp. Spanning the pre-frontal cortex and the occipital areas of the brain, the fNIRS channels placed during this experiment collected hemodynamic data pertaining to working memory capacity. The sampling frequency is 7.81 Hz.

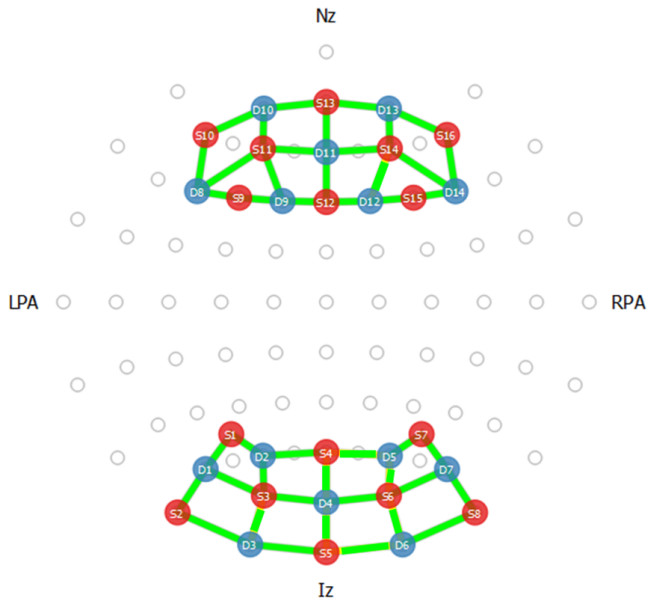

Figure 1: Optode layout of the fNIRS sources (Red), detectors (Blue) and channels (Green) used during the experiment.

Electrocardiography (ECG): Sensors placed on the torso as shown in Figure 2 collect the ECG data using the MP160 BioPac system with the BioNomadix wireless devices. The EL 503 BioPac general purpose disposable electrodes are used on the torso region. ECG is the process of obtaining an electrocardiogram, which is the electrical signal from the heart muscle, used to detect heart problems. The sampling frequency is 2000Hz.

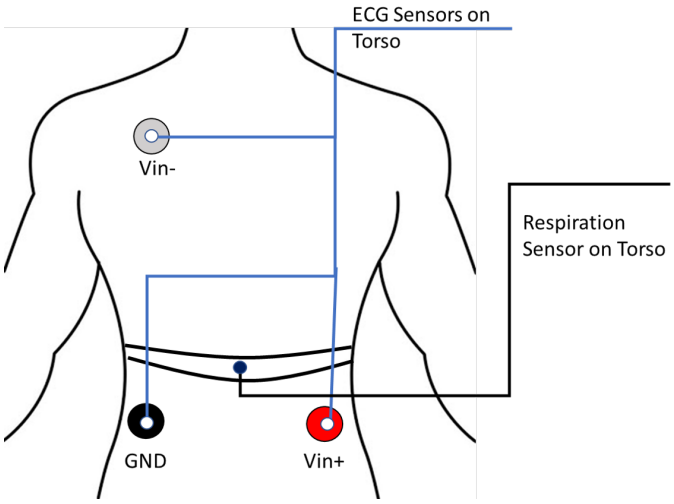

Figure 2: Electrode placements of Electrocardiogram sensors and respiration belt on the front torso.

Respiration: Respiration belt sensor of the MP160 BioPac system is placed on the abdomen of the participant in contact with the torso, as depicted in Figure 2. The extent of inhaling and exhaling that the participant experiences during this experiment is captured by the belt sensor. The sampling frequency is 2000Hz.

Skin Surface Temperature: As portrayed in Figure 2, the skin surface temperature data is collected from the little finger of the non-dominant hand using the MP160 BioPac system with the BioNomadix wearable device coupled with BN-TEMP-A-XDCR BioPac sensor. The Empatica E4 wearable wristband worn by the participant also collected skin temperature data. The sampling frequency for BioPac is 2000Hz, and for Empatica E4, is 4 Hz.

Electrodermal Activity (EDA): Sensors from both MP160 BioPac system and Empatica E4 wearable wristband are used to collect skin conductance recordings. The Empatica E4 wearable wristband is worn on the wrist of the participant. The M160 BioPac system sensors are placed over the ring and middle fingers of the participant as shown in Figure 2. The BioPac EL507 disposable electrodes are used as the leads for EDA. EDA signals

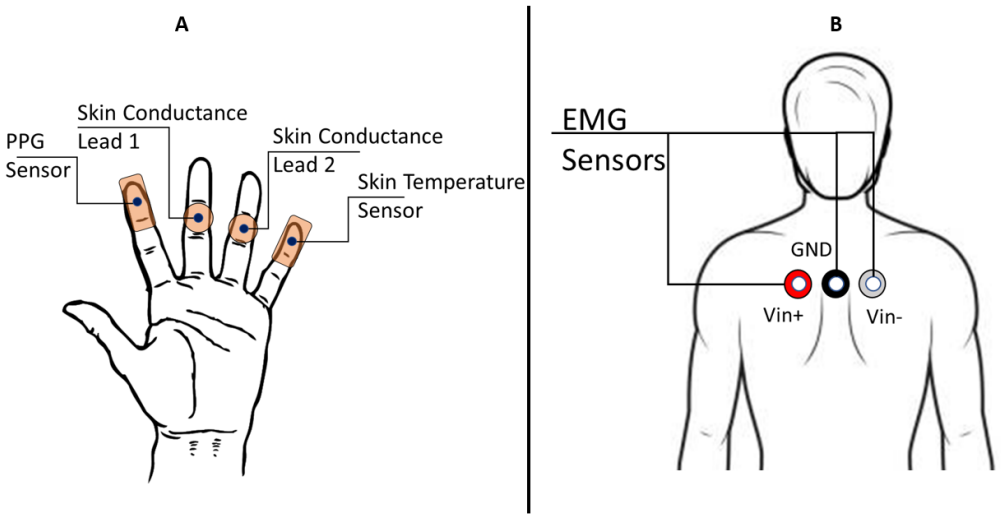

Figure 3: **A)** Sensor placements on the hand for PPG, EDA, and Skin Temperature. **B)** EMG sensors placed on the participant's trapezius muscles to detect stress states from muscle movements.

are the changes in skin's conductance level in response to emotional arousal. The sampling frequency for BioPac is 2000Hz, and for Empatica E4 is 4 Hz.

Photoplethysmography (PPG): Wearable physiological sensor, BN-PULSE-XDCR coupled with BioNomadix unit, is placed on the index finger of the non dominant hand, as described in Figure 3 to obtain photoplethysmography (PPG) data with the M160 BioPac system. The Empatica E4 wearable wristband, worn on the wrist of the participant also collects PPG data. PPG is the optical means to detect changes in blood volume in a tissue. PPG is generally used to monitor cardiac health and heart rate. The sampling frequency for the BioPac system is 2000Hz. Sampling frequency for the Empatica E4 wristband is 4 Hz.

Electromyogram (EMG): Sensors from the MP160 BioPac system for Electromyogram (EMG) recordings are placed on the participant's trapezius muscle, as depicted in Figure 2. The EL503 general purpose electrodes are used on the trapezius muscle, which

has been shown to detect mental stress [39]. EMG is used to detect the health of muscles and the nerves that control them. In this experiment, placement of EMG electrodes provides data about the tensing of a participant's shoulders and back while performing a cognitive stress task. The sampling frequency is 2000Hz.

Eye Gaze Tracking: Tobii Pro glasses 2 are used for eye gaze tracking. participant's wore the glasses similar to regular eye glasses. These glasses track the gaze of the participant, along with changes to pupil diameter.

Facereader: Facereader software, which detects emotional arousal and valence along with emotional state of the participant based on facial expressions was used to track emotional changes.

Synchronization of Multi-modal Data Recording: The experiment's design, timing and triggers for different equipment is executed using the Chronos input device and E-Prime software. Thus, a recording of each participant's performance, i.e. number of correct and incorrect responses to the stimuli presented, and response time data is collected. This data can prove valuable to measure performance with different kinds of music playing in the background.

Along with the aforementioned signals, the Empatica E4 also collects 3-axis accelerometer data which is the indication of the movement in the right hand. These signals can be used to negate motion artifacts from the data collected by Empatica E4. The sampling frequency for this data is 32 Hz.

2.2.2 Experimental Protocol

The experiment conducted in this research is approved by the Institutional Review Board at the University of Houston, Houston, Texas, USA. There are 11 participants (five male and six female) in the experiment. All participants are between the ages of 22-25. With the aim of observing the influence of music on cognitive engagement, two sessions of similar tasks are conducted for each participant. Session 1 is conducted with the participants choice of

calming music playing in the background and session 2 is with participant's choice of vexing music playing in the background. After each session, 2 minutes resting period is provided.

***n*-back Task:** The n-back task performed by the participants in this experiment is a working memory experiment which measures the mental workload that the participant experiences during a simple memory recalling task. In an n-back task, the participant is presented with a series of stimuli displayed one at a time, and the participant has to identify if the currently displayed stimuli is the same as the n^{th} previous one. As is evident, the difficulty level increases with n, since the participant has to recall more of the stimuli with larger values of n.

Each session included 8 trials each of the two n-back tasks. In this experiment, we use letters as stimulus, similar to [1], [23]. While studies [1], [23] perform 10 trials for each n-back task for working memory experiments, we have the participants performing 8 trials of each n-back task. During the initial trials of our experiment, the participants displayed signs of fatigue after 30 minutes of performing the experiments, and having 10 trials in our experiment considerably exceeded the total time. Hence, to avoid participant fatigue, we designed the experiment with 8 trials of each task. The two tasks in this experiment are the *1*-back, and the *3*-back tasks. Since in *1*-back task participant has to recall only the previous stimuli, the expected cognitive stress of the participant is low for this task. Whereas the stress while performing the *3*-back task is high as the participant has to recall 3^{rd} previous stimuli. We chose to perform multiple trials of these two n-back tasks in this experiment to make sure that the participant does not experience fatigue over the course of the experiment. Choosing these two tasks also has the added advantage of showing different physiological reactions, as their corresponding stress levels are varied. At the beginning of each trial, we show a 5 seconds instruction describing the task. This display described the task (e.g. *1*-back or *3*-back task) that the participant would be performing. Every trial consisted of 22 stimuli. For each stimuli, a letter is displayed for 0.5 seconds followed by a resting cross for 1.5 seconds. Hence, each trial had a total duration of 49 seconds

[instruction time 5 seconds + (stimulus display time 0.5 seconds + resting cross display 1.5 seconds)×22 trials = 49 seconds]. In each trial, 30% of the stimuli are a target. The trials of each n-back task are randomized. At the end of each trial, a 10 second RELAX segment is presented where a resting cross is displayed on the screen. After 8 trials (halfway mark for each session), a 20 second RELAX section is presented where a resting cross is displayed on a smart 65 inch TV screen connected via HDMI to a laptop. The data is collected on a separate laptop connected to all the devices. The entire duration of each session is 964 seconds [each trial duration 49× number of trials 16+ inter trial rest period 10× number of trials16+ RELAX period 20 = 964 seconds] which is approximately 16 minutes. After each session, there is a 2 minute relaxation break where the participant is allowed to repose. A relaxation cross is displayed on the screen during this time. Figure 4 describes the timing of one session with 16 randomized trials.

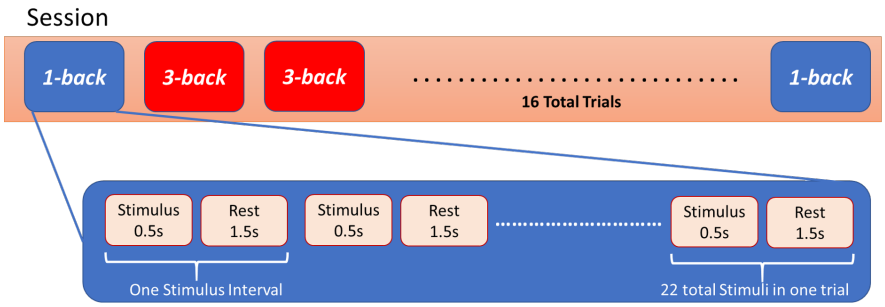

Figure 4: Session representation of n-back tasks with multiple trials. Each trial will consist of 22 stimuli presented to the participant.

The participants are seated comfortably wearing the required noninvasive sensors, and a display screen placed approximately 1-2 meters in front of them. The only movement required of the participant is to indicate whether the stimulus is a target or not, by pressing two buttons on a Chronos Keypad, the target button and the non-target button. The participants are required to press either one of the buttons for each displayed stimulus. The correct and incorrect responses are also recorded.

In total the experiment took, a total of 2168 seconds [calming session duration 964+ inter

session break 120+ vexing session duration 964+ after session resting period 120 = 2168 seconds], i.e. approximately half an hour.

2.3 Results

During the experiment, we successfully obtained all the physiological recordings with fNIRS analysis. Figure 5 is a representation of the deoxygenated and oxygenated hemoglobin concentration for participant 6, in the calming music session for channel 2. The red line indicates in concentration of HbO and the blue line indicates concentration of HbR. Figure 6

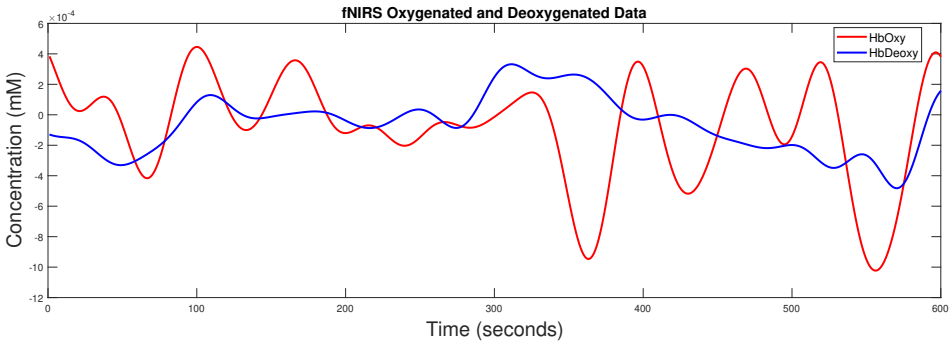

Figure 5: Portion of fNIRS oxygenated and deoxygenated hemoglobin concentration data from channel 2 during the calming music session.

shows three physiological signals, namely respiration, skin temperature, and EDA collected using Biopac MP160 from participant 6 as an example. The blue background indicates the initial baseline period, The green background indicates the calming music session, and the red background represents the vexing music session. The figure shows variations of the signals during different periods of the experiments. Figure 7 shows 20 seconds of ECG and EMG signal collected for participant 6 during the experiment. The signals displayed are extracted from the initial baseline period before the starting of the experiment. We can observe the ECG signals corrupting the collected EMG signals.

We consider and analyze physiological data from nine of the eleven participants, since two participants did not complete the entire experiment. Figure 8 shows all the EDA

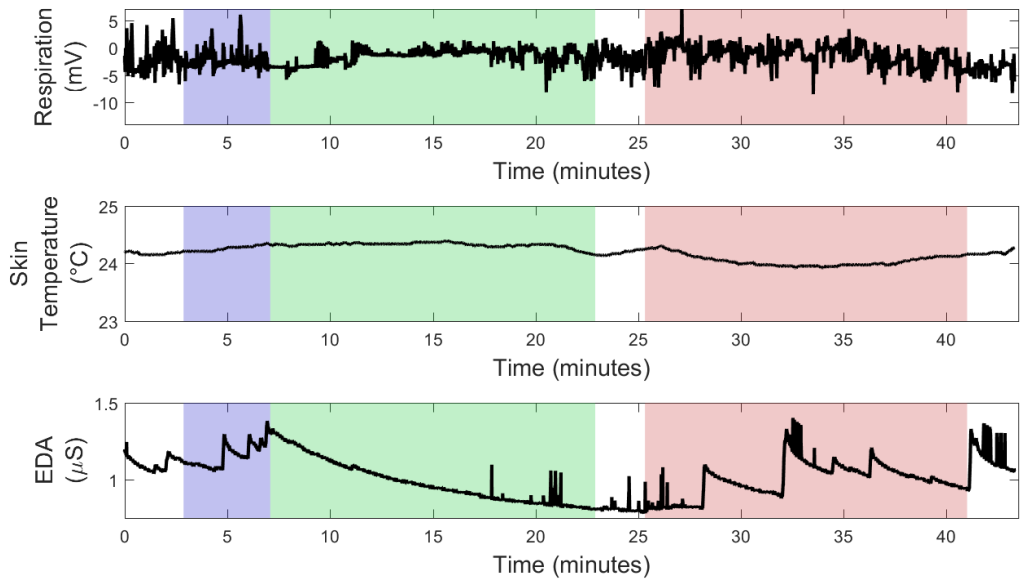

Figure 6: The sub-panels from top to bottom shows respiration signal, skin temperature, and EDA signal for participant 6, respectively. Color coded backgrounds represent the initial baseline period (blue), calming (green), and vexing (red) music sessions.

signals from nine participants out of eleven during calming and vexing music sessions. We obtain the mean EDA signal by taking the mean of these nine EDA signals. We discarded EDA signals from participants 1 and participant 4 in this figure because of incomplete data collection. Figure 8 also represents the mean EDA signal. Similarly, Figure 9 shows all the skin temperature signals from nine participants out of eleven participants. We discard the same two participants for similar reason. Figure 9 also shows the mean skin temperature signal obtained from these nine participants.

Figure 10 represents box-plots for means and variances obtained from different physiological signals from the nine participants in three different sessions, i.e. calming music session, vexing music session, and two minute relaxation period in between two music sessions. In Figure 10, the first subplot represents the box-plot denoting the distribution of the means of the EDA signals obtained from each of the nine participants. Similarly, the

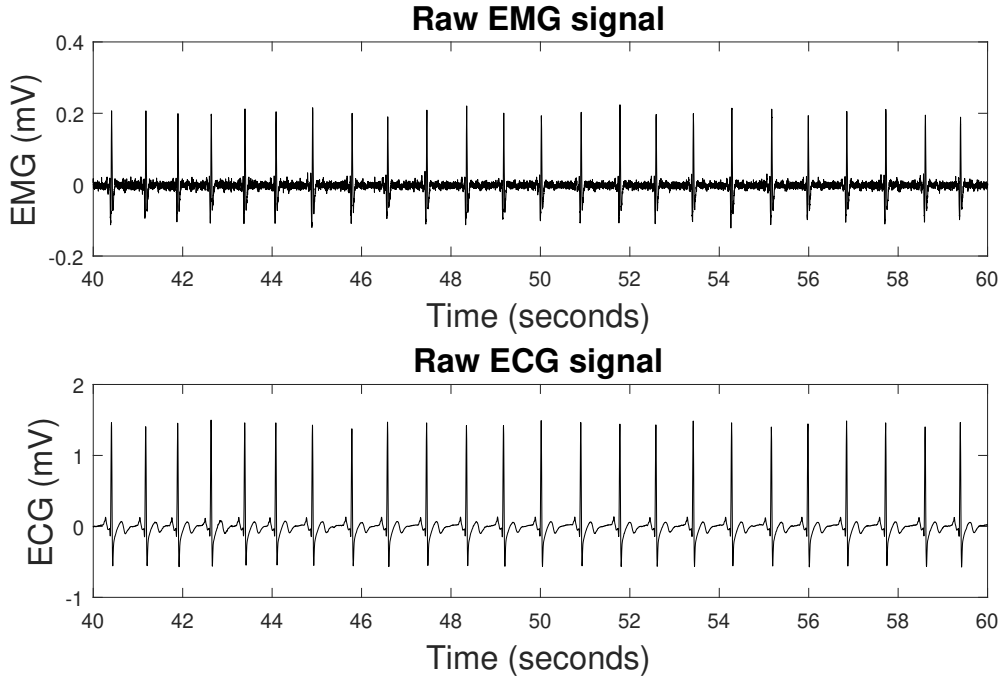

Figure 7: Example of collected raw EMG and ECG signals for participant 6, during the initial baseline period displayed for a duration of 20 seconds for participant 6.

second subplot in Figure 10 represents the box-plot denoting the distribution of the means of the skin temperature signals obtained from each of the nine participants during the three sessions of interest. The third and fourth subplots in Figure 10 represent the distributions of the variances of EDA and skin temperatures in different sessions. The first and third box-plot in Figure 10 visualizes how EDA, i.e. sweat secretion activity is changing in order to cool the body in response to varying metabolism during the three sessions, i.e. calming music session, vexing music session and two minute relaxation period in between two music sessions. On the other hand, the second and third box-plot in Figure 10 represent the body temperature variations as a combined result of variations in the body's metabolism as well as the body's cooling effect. The fifth subplot in Figure 10 corresponds to the distribution of the respiration variances in all nine participants for all three sessions. This is a representation of how the variations in the amplitude of the respiration are distributed in different

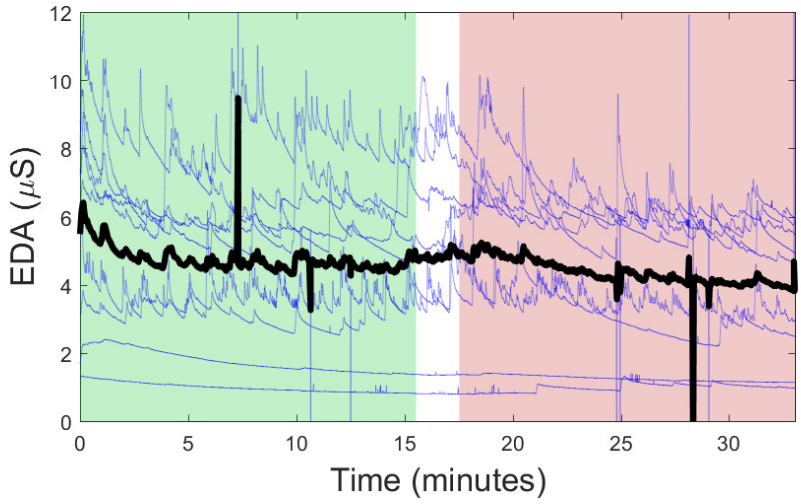

Figure 8: EDA data for the calming (green background), RELAX (white background), and vexing (red background) music sessions. Blue lines are data from all the participants, the black signal represents the mean EDA from all the participants.

sessions. The last subplot in Figure 10 represents the variance in the EMG signal collected from the back of the participants during the three different sessions.

Usually, there are some differences in the distribution plots in Figure 10. However, exclusive visual interpretation from the box-plot might be misleading. Therefore, we performed statistical analysis among all the distributions. We performed two sample t-tests between calming music vs relaxation, calming music vs vexing music, and between vexing music vs relaxation for each of the subplots. The t-test is performed considering the null hypothesis that the two distributions come from independent random samples from normal distributions with equal means and equal but unknown variances. None of the statistical analyses could reject the null hypothesis except for the two distribution of the variances in the skin temperature in calming music and relaxation. Our statistical analysis shows that the two distribution of the variances in the skin temperature in calming music and that of relaxation come from different distributions with a p-value of 0.0029. Figure 11 also shows the paired differences of means of EDA, variances of EDA, means of skin temperature, and

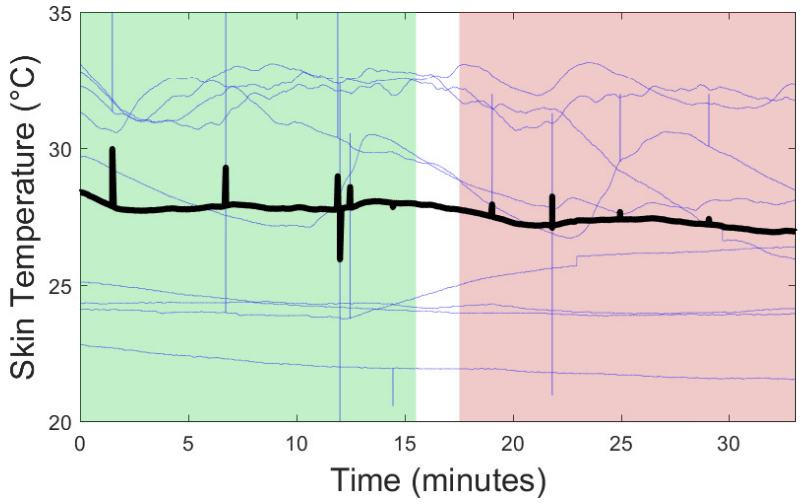

Figure 9: Skin temperature data for the calming (green background), RELAX (white background), and vexing (red background) music sessions. Blue lines are data from all the participants, the black signal represents the mean skin temperature from all the participants.

variances of the skin temperature between the calming and the vexing music sessions for each participant. We performed the Wilcoxon signed-rank t-test, and found a significant difference between the means of EDA during calming music and the means of EDA during vexing music with p-value of 0.0156 in the paired test.

2.4 Discussion

In this section, we discuss the preliminary analysis results on the neurophysiological signals obtained from our experiment.

We note that Figure 5 shows the changes in concentration in the HbO and HbR during a portion of the calming music session. There is also an oscillating change between the HbO and HbR concentrations, indicating periodic hemodynamic changes. Higher HbO level at the beginning might indicate more oxygenated blood than deoxygenated blood at that time. While the lower concentration of HbO towards the end indicates lower oxygenated level. In Figure 6, we observe the EDA signal being more tapered during the calming music session,

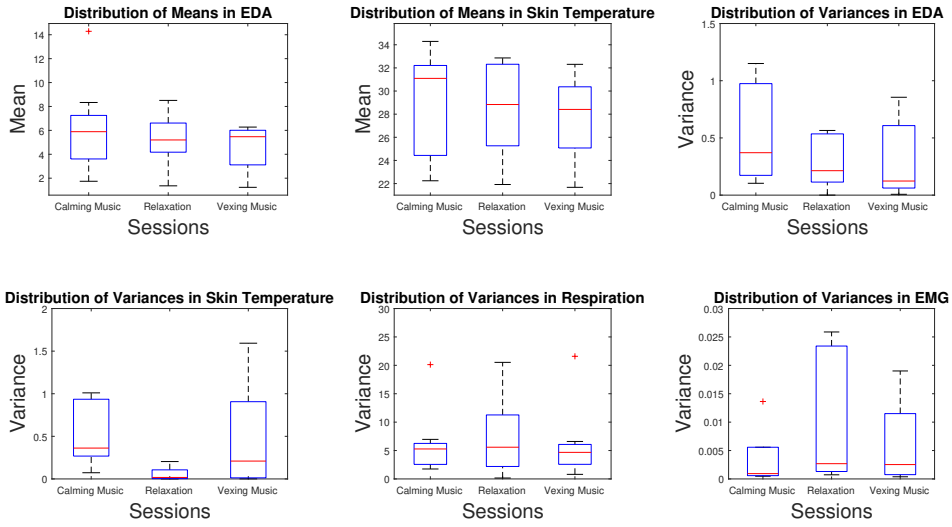

Figure 10: Clockwise from top; distributions of means of EDA, means of skin temperature, variances in EDA, variances in skin temperature, variances in respiration, and variances in EMG for the calming music, relaxation, and vexing music sessions.

whereas it presents several spikes during the vexing music session for the specific participant 6. This is consistent with the expected results, as performing the n-back tasks with vexing music playing in the background can increase the mental workload. However, this might not be a general case for all the participants. In order to check the consistency of this pattern, we plotted all the EDA signals from nine participants in Figure 8. While there is a difference in the patterns of EDA in a different session, it is not consistent among all the participants. For instance, for some participants, the EDA activity is higher in the vexing music session, when compared to the calming music session. Though this consistency is not immediately visible, this might be an indication of interpersonal variability in the EDA response. We also derived the mean of the EDA signals. We can see that the mean signal is smaller in the vexing music session than the calming music session. This might indicate that the participants are less likely to concentrate on the task resulting in reduced cognitive engagement and stress. Figure 10 shows the distribution of the mean and variances of EDA

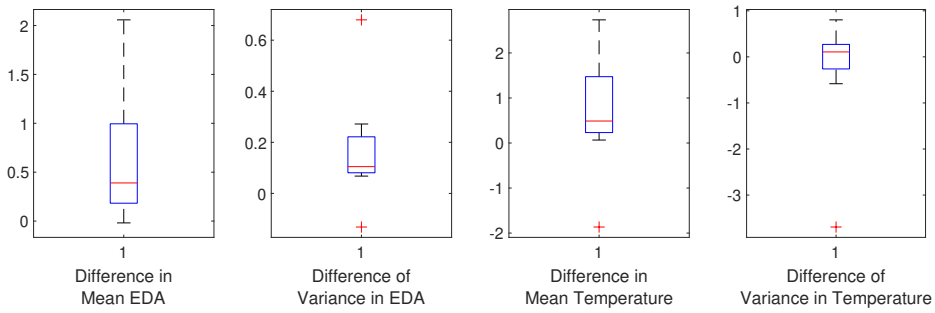

Figure 11: Each subplot from left to right shows the paired differences of means of EDA, variances of EDA, means of skin temperature, and variances of the skin temperature between the calming and the vexing music sessions within each participant.

in three different sessions. We performed pooled t-test between the different distributions. However, no significant difference is found. As pooled t-test might not reflect the changes due to the interpersonal variability, we performed the paired t-test on the means of EDA during calming and vexing music. We observe a significant difference in the means of EDA during calming and vexing music. From the box-plot in Figure 11, we note that the mean EDA in calming session is different than the mean in vexing music.

In Figure 6, we observe that the skin temperature signal is slightly higher during the calming music session than that of the vexing music session. One might think this is counter-intuitive as more temperature could be an indication of increased heat generation due to metabolism corresponding to the stress that has been induced in the session. However, the temperature is the combined result of increased metabolism and the cooling effect. As there is less EDA activity during the calming session, the less cooling gave rise to higher body temperature. Moreover, this pattern might not be consistent with different participants. Therefore, we plot all the temperature signals in Figure 9 and derive the mean signal. It shows the mean temperature during vexing music is slightly lower. We also observe the distribution of the skin temperature during three sessions. In Figure 10, the second and the fourth subplots correspond to the distributions of the temperature. From the pooled

statistical analysis, we observe the two distributions of the variances in the skin temperature during calming music and that of relaxation come from two different distributions. Any statistical significance has not been observed in the paired analysis. This is an indication of maintaining the body's homeostasis during different stress conditions. Although the difference has not been observed between the calming and vexing music sessions, which are more longer in duration, the difference has been observed with statistical significance between calming music sessions and relaxation period. This indicates a sudden decrease in temperature. One possible explanation is that the sudden decrease in the mental workload task leads to decreased metabolism. However, higher cooling rates caused by higher skin moisture resulted during the calming music session lead to a sudden drop in the skin temperature.

The respiration signal also shows a cursory difference between the two music sessions for participant 6 in Figure 6. During a majority of the calming music session, the respiration signal stays nearly constant, whereas in the vexing music session, we observe a marked increase in the fluctuations. This indicates that the participant had even breaths during the calming music session and more erratic breathing during the vexing music session. We calculated the variances during calming music session, vexing music session, and the relaxation to calculate this fluctuation. However, a significant difference has not been observed between calming and vexing music sessions in both the pooled t-test and the Wilcoxon signed-rank test. We also did not find any statistical difference in the EMG signal in our preliminary statistical analysis. The future directions include a more systematic analysis for respiration, EMG, and ECG data to obtain meaningful information from them.

3 Experimental Validation of Mental Workload Assessment Using Hierarchical Latent Dictionary Method[1]

3.1 Background of Mental Workload Analysis

Human-computer interfaces (HCI) have been extensively researched due to their myriad applications [41, 42]. The idea of interacting with systems directly via neural signal factors opens avenues of applications such as improving communication with the differently-abled, detecting cognitive engagement in online learning [43]. An important aspect of any HCI is to assess user's mental workload [7]. Mental workload can have direct negative consequences in reducing the performance of the HCI [44]. It is therefore important to understand and classify the amount of mental workload during a cognitive task [7].

One of the most widely used methods to quantify mental workload is Functional Near Infrared Spectroscopy (fNIRS), which obtains amounts of deoxygenated and oxygenated blood data in brain regions in a non-invasive way. Biological tissues scatter light from the near infrared region of the electromagnetic spectrum (620-1000nm), but it is absorbed by hemoglobin [35]. Using the amount of absorbed near infrared light and modified Beer Lambert Law [36], fNIRS estimates oxygenated and deoxygenated hemoglobin (HbO and HbR, respectively). This measurement of hemodynamic responses indicates neurons in regions of the brain that are consuming more oxygen because of neuronal activation. Therefore fNIRS measurement can be considered as an indirect way of measuring neuronal activations in different regions of the brain. fNIRS uses this phenomenon to estimate the levels of HbO and HbR using modified Beer Lambert Law [36]. The non-invasive nature of fNIRS coupled with the its portability make it one of the few easy and functional neuroimaging technique to obtain mental workload datasets.

Established fNIRS analysis techniques such as generalized linear models do not model

[1] **This chapter was presented in part at the proceedings of IEEE Biology and Health Informatics Conference [40].**

the dynamic functional connectivity and implement other properties such as heteroscedasticity and obtaining a functional connectivity map for one specific n-back task. In this study, we apply the hierarchical latent dictionary method (HLDM) [17] to model the time varying functional connectivity. Additionally, we incorporate several characterstics with HLDM, such as reducing data from spatially localized redundant sensors, incorporate trial to trial variability while also modeling the underlying neural signal factors for one specific n-back task. We model fNIRS with HLDM to improve classifications of between different mental workload tasks. We obtain hidden latent factors from the hemodynamic response of fNIRS on two datasets. Firstly, it is applied on the open-access dataset from [1]. The second dataset on which we applied HLDM, is data collected from a similar experiment that we conducted, which includes two different types of music, namely relaxing and disturbing music and to perceive their impact on the mental workload (see Chapter 2 for details). The shared covariance matrix for each of the n-back tasks, for each participant is extracted to obtain definite results of differences in the neuronal pathways when various n-back are performed. Multiple machine learning algorithms namely; support vector machines, k-nearest neighbours method and linear discriminant analysis are used to classify these tasks using the features extracted from the functional connectivity maps derived by the application of HLDM.

3.2 Methods

In this section, we first briefly discuss the datasets and our novel experiments. We outline the mathematical formulation and introduce the algorithm to analyze the hemoglobin data from two datasets.

3.2.1 Experiments

Experiment 1 (*Dataset 1*): We use the publicly available fNIRS data collected from human participants while performing 3 n-back tasks [1]. There are 10 trials for each task with

the main focus on the cognitive task. Each trial consisted of a 5-second instruction, followed by 22 letters with 0.5 seconds for displaying the letter and 1.5 seconds for displaying the resting cross and for the participant to respond to the given stimulus [1]. This resulted in a total of 44 seconds of stimulus time and a 15 second relaxation break. Each trial had 3 ± 1 stimulus and after every 15 trials a break period of 150 seconds is incorporated. Hemodynamic responses are measured using an Oxymon Mark III by Artinis Medical Systems and two wavelengths of 765 nm and 856 nm are used to measure concentration changes in HbO and HbR, respectively [1]. The experiment made use of four transmitter and four receiver optodes placed on the prefrontal cortex region of the brain such that each detector measures time multiplexed data from two sources thus resulting in a total of 8 channels of HbO and HbR. The sampling frequency is 25 Hz. Here we refer to the data obtained in [1] as *Dataset 1*. The analysis conducted on this dataset is an extension of the study in [40].

Experiment 2 (*Dataset 2*): This dataset is from the experiment we conducted to collect fNIRS and physiological signals while performing n-back tasks (see Chapter 2 for complete details). To explore the effects of different kinds of music on the mental workload, background music of the calming and vexing type are played while the participants perform the n-back tasks. Only *1*-back and *3*-back tasks are performed in this experiment in order to further comprehend the difference between the lowest and highest difficulty levels. Two sessions of n-back experiment are conducted on 11 participants with five male and six female participants. Though fNIRS data has been collected from 11 participants, only 8 are available since two participants did not complete the experiment and the fNIRS data was corrupted for the third participant. There are two sessions in our experiment, session 1 and session 2, which are conducted similarly except for the different music playing in the background. Session 1 had calming music, and Session 2 had vexing music in the background. Each session had a total of 16 trials; 8 of *1*-back and 8 of *3*-back. Further details of the protocol, and setup of this experiment are described in Chapter 2. The dataset thus generated from this experiment is referred to as Dataset 2. We asked each participant to select

calming music tracks and vexing music tracks for their experiment. Thus we are making sure that the music tracks have the desired effect on the participants. We also collect other physiological signal data apart from fNIRS, such as electrodermal activity, respiration, heart rate, skin temperature, etc (see Chapter 2 for further details).

Table 1 presents the information of the two datasets in a compact manner. Each n-back task has tr number of trials and in both datasets, for p number of channels, the observation matrix $y^{tr,n}(t)$ is $p \times 1$ and the task-specific covariance matrix has a dimension of $p \times p$. The mean dimensions are trial-specific and hence are of dimension $tr \times p$, i.e. tr trials per n-back task for p channels.

Table 1: Dataset Information [1].

Dataset	participants	Channels	$y^{tr,n}(t)$ Dimension	Functional Connectivity Dimension
Dataset 1	10	8	8×1	8×8
Dataset 2	8	22	22×1	22×22

3.2.2 Modeling

Processing data from a large number of channels can be challenging. The computational load and time increase as more channels are considered. Additionally, there are redundancies in the information in large number of channels that can be reduced. HLDM decreases the computational load by hypothesizing that the observed signals originate from hidden latent variables that are of lower dimension as compared to the observation matrix. We also obtain covariance matrices that evolve in time. This describes the dynamic functional connectivity and ameliorates the time-varying connections between different regions in the brain for different tasks. HLDM incorporates this time evolving dynamic nature of the covariance matrices by considering that the latent factors also evolve in time. Another important consideration is obtaining the underlying neural signal factors that do not vary from trial to trial and specific to a particular mental workload task. HLDM obtains this by considering them as Gaussian processes and linking them to the hidden latent factors hierarchically. The

following subsections describe in detail how we model the aforementioned properties.

Gaussian Processes: Gaussian process is a probability distribution over a function $f : \mathbb{R} \rightarrow \mathbb{R}$, belongs to a Gaussian process $G(m, q)$ if

$$p(f(t_1), f(t_2), f(t_3), \cdots, f(t_u)) \sim \mathcal{N}(\mu, Q), \tag{1}$$

where $\mathbb{R}$ represents real numbers and $\mathcal{N}$ is the normal distribution [45],
$\mu = [m(t_1), m(t_2), \cdots, m(t_u)]$ and Q are the mean and kernel functions of the Gaussian process. In HLDM, underlying neural signal factors are modelled as Gaussian processes. For the Gaussian process f, $B = [t_1, t_2, \cdots, t_u]$ are values where $t_1, t_2, \cdots, t_u$ are arbitrary points in $\mathbb{R}$, and the vector $[f(t_1), f(t_2), \cdots, f(t_u)]$ forms a multivariate normal distribution of dimension u. If we consider the mean of the process to be 0, then

$$f(t_w) \sim \mathcal{N}(0, Q_{w,e}), \tag{2}$$

$Q_{w,e}$ is the $u \times u$ *Gram Matrix* or *covariance matrix* between all possible pairs of (t_w, t_e) [45]. These covariance matrices reflect the extent of similarities between different points in a signal.

Mean and Kernel Functions: Based on equation (2), by assuming zero-mean Gaussian process, the Gaussian process will be entirely defined by its variance, which can be represented as a kernel function Q. The kernel function essentially gives us the covariance between values at (t_w, t_e). A Gaussian process can have a kernel function like squared exponential kernel that results in a smooth, infinitely differentiable functions, or have a non-differentiable function from a Brownian kernel [46]. The kernel function describes how similar one data point is with another data point. Exponential kernel function is defined as follows:

$$q_h(t_w, t_e) = d_h \exp(-\kappa \parallel (t_w - t_e) \parallel_2^2), \ \forall \ h \in \{0, 1\}, \tag{3}$$

where q_0, q_1 are kernel functions, d_0, d_1 are scale hyperparameter, and κ is bandwidth parameter.

Latent Factor Analysis: The latent factor model relates a set of observable variables to a set of latent factors with a lower dimension [47]. Modelling the observed data to be expressed as the functions of a number of latent or hidden factors helps in reducing the high dimensionality of the fNIRS data. It also isolates the unobserved latent variables which define as brain's neural signal factors. The observed multi-channel fNIRS signals are modeled as linear combinations of latent factors in addition to an error term. This will result in finding the covariance matrices that reflect the relationships among many variables in terms of a few underlying hidden random quantities. Designing the observation matrix $y^{tr,n}(t_d)$ as a $p \times 1$ matrix wherein each element represents data from one sensor at a particular discrete point in time $t_d, \forall d \in \{1, ..., \tau\}$ where τ is the final time for one trial of a n-back experiment and the time points being represented by t_d. Assuming $y^{tr,n}(t_d) \sim \mathcal{N}_p(0, \Sigma)$, the latent factor model can be written as

$$y^{tr,n}(t_d) = C_n(t_d)x^{tr,n}(t_d) + \epsilon(t_d), \tag{4}$$

where $x^{tr,n}$ is the k dimensional latent factor with $k << p$, C_n is the factor loading matrix, and ϵ is the Gaussian noise factor, $\epsilon \in \mathcal{N}_p(0, \Sigma_0)$, Σ_0 is a diagonal matrix with variance as the diagonal elements. $\Sigma(t_d) = C_n(t_d) \times C_n^T(t_d) + \Sigma_0(t_d)$ is the $p \times p$ covariance matrix of $y^{tr,n}(t)$ that denotes the time evolving functional connectivity [48], [49], [50].

Derivation of Σ- From equation (4), covariance of $y^{tr,n}(t_d)$; Σ, will be

$$\Sigma = E[(C_n(t_d)x^{tr,n}(t_d) + \epsilon(t_d)) \times (C_n(t_d)x^{tr,n}(t_d) + \epsilon(t_d))^T], \tag{5}$$

where E stands for the expectation operation. Further solving the equation, we get,

$$\begin{aligned}\Sigma = E[C_n(t_d)x^{tr,n}(t_d)x^{tr,n}(t_d)^T C_n(t_d)^T + C_n(t_d)x^{tr,n}(t_d)\epsilon(t_d) + \epsilon(t_d)C_n(t_d)^T x^{tr,n}(t_d)^T \\ + \epsilon(t_d)^T \epsilon(t_d)].\end{aligned} \tag{6}$$

Since $x^{tr,n}(t_d)$ is independent and identically distributed, $E[x^{tr,n}(t_d)x^{tr,n}(t_d)^T]$ is a $k \times k$ identity matrix and $E[x^{tr,n}(t_d)\epsilon(t_d)]$ is reduced to zero. Therefore, equation (6) reduces to,

$$\Sigma = C_n \times C_n^T + \Sigma_0, \tag{7}$$

where $\Sigma_0 = E[\epsilon(t_d)^T \epsilon(t_d)]$.

Dynamic Latent Factor Model: A dynamic latent factor model assumes Markov evolution for the latent factors $x^{tr,n}$, with time-invariant parametrization [51], [52]. Hence, the stationary process $x^{tr,n}(t_d)$ follows the equation:

$$x^{tr,n}(t_d) = A^{tr,n}(t_d)x^{tr,n}(t_{d-1}) + v^{tr,n}(t_d), \tag{8}$$

where $A^{tr,n}$ is the transition matrix, $v^{tr,n}(t) \in \mathcal{N}_k(0, I_k)$ is the Gaussian noise factor, I_k is a k dimensional identity matrix. Characterizing the observed signal with a dynamic latent factor model provides us with the mathematical framework to derive the time-evolving covariance matrices.

Hierarchical Latent Factor Model: Building upon the latent factor model, the hierarchical latent factor model represents the factors so that they evolve non-parametrically, thus capturing the long range dependencies between data points [17]. This will also result in a time-evolving covariance matrix which represents the dynamic functional connectivity of different brain regions when participants are performing memory-related tasks. The shared covariance obtained for a single n-back experiment, irrespective of the number of trials, reflects the functional connectivity map of the brain for that specific experiment. The p dimensional observation matrix $y^{tr,n}(t_d)$ for trial number tr, task n, at discrete point time t_d with trial specific mean $\mu^{tr,n}(t_d)$ and task specific covariance $\Sigma^n(t_d)$ can be written as,

$$y^{tr,n}(t_d) \sim \mathcal{N}_p(\mu^{tr,n}(t_d), \Sigma^n(t_d)). \tag{9}$$

Instead of equation (8) the latent factors $x^{tr,n}(t_d)$ are modelled as k latent dictionary functions $\phi^{tr,n}(t_d)$ each of which is the summation of a trial specific random function [53], [48], and a k dimensional noise component $v^{tr,n}(t_d)$:

$$x^{tr,n}(t_d) = \phi^{tr,n}(t_d) + v^{tr,n}(t_d). \tag{10}$$

Knowledge sharing within this latent space is done by hierarchically coupling the trial specific dictionary functions around the mean $\phi_0^{tr,n}(t_d)$:

$$\phi^{tr,n}(t) = [\phi_1^{tr,n}(t), \cdots, \phi_k^{tr,n}(t)]. \tag{11}$$

Here, we consider the squared exponential kernel function defined in (3). The following equations describe the hierarchical coupling by equating the mean of the child processes to the parent process $\phi_0^{tr,n}(t)$ [17]:

$$\phi_j^{tr,n}(t) \sim GP(\phi_0^{tr,n}(t), q_1), \tag{12}$$

$$\phi_0^{tr,n}(t) \sim GP(0, q_0). \tag{13}$$

where GP refers to a Gaussian process, q_0, q_1 are the corresponding kernel functions. This hierarchical linking is used to obtain the underlying neural signal that is assumed to be similar across trials. Thus, the obtained covariance matrices will reflect the underlying neural signal that is hierarchically coupled with the latent dictionary functions. Latent factors do not evolve at a constant rate throughout the duration of the task. This evolving variability across time, also known as heteroscedasticity, is captured by making the factor loading matrix C_n to be the time-evolving and weighted combination of smaller dimensional latent dictionary functions:

$$C_n(t) = \Theta^n(t)\zeta^n(t), \tag{14}$$

where $\Theta^n \in \mathbb{R}^{p \times L}$ is distributed according to a conditionally Gaussian shrinkage prior and $\zeta^n(t)$ is a Gaussian process with zero mean and kernel function q_0.

The resulting trial specific mean $\mu^{tr,n}(t_d)$ and task specific covariance $\Sigma^n(t_d)$ will be as,

$$\mu^{tr,n}(t_d) = \Theta^n \zeta^n(t_d) \phi^{tr,n}(t_d), \tag{15}$$

$$\Sigma^n(t_d) = \Theta^n \zeta^n(t_d) \zeta^{n^T}(t_d) \Theta^{n^T} + \Sigma_0^n, \tag{16}$$

where Σ_0^n is a diagonal matrix with diagonal elements being the covariances. As described in [54] and [55], the estimation algorithm using Bayesian non-parametric modeling is further explored for the fNIRS application.

3.2.3 Estimation

The following steps detail the formulation of the algorithm. We first assume a prior distribution over Θ^n from equation (14). To achieve reduced computational complexity, we impose a prior that can reduce a majority of the elements to zero. We then model and sample the dictionary function and the noise parameter $\phi^{tr,n}$ and $v^{tr,n}(t_d)$ from equation (10). Then, using these dictionary function we model and sample the underlying neural signal as the child process $\phi^{0,n}(t_{1:\tau})$ from equation (13). Then, we model the latent covariance dictionary function ζ^n from equation (14). Now that we have all the parameters in equation (13), we can find Σ_0^n. The last step is to update the Θ^n values.

Step 1: Selecting Priors- Our first step is to model Θ^n, from equation (14) by imposing a Gamma distribution prior [17] on Θ^n such that a larger number of elements of Θ^n will be close to zero. This helps in reducing the computational complexity of the model. We then chose L, i.e. the number of columns in Θ^n, to be greater than the number of latent factors k. As suggested in [48], we use the shrinkage prior

$$\Theta_{i,l}^n | \, \chi_{i,l}, \theta_l \sim \mathcal{N}(0, \chi_{i,l}^{-1}, \theta_l^{-1}); \chi_{i,l} \sim \text{Ga}(3/2,3/2), \tag{17}$$

$$\delta_1 \sim \text{Ga}(a_1,1), \delta_h \sim \text{Ga}(a_2,1), h \geqslant 2, \theta_l = \prod_{h=1}^{l} \delta_h, \tag{18}$$

where Ga refers to the Gamma distribution. Thus, taking $a_2 > 1$ leads to δ_h being greater than 1 in general, which leads to θ_l tending to infinity with greater l[17], [48]. This process shrinks the elements of $\Theta_{i,l}$ to zero for larger l.

We also specify an independent inverse Gamma prior on the diagonal elements of Σ_0^n

$$\sigma_{i,n}^{-2} \sim \ \text{Ga}(a_\sigma, b_\sigma); i \in \{1,2,3,...,p\}. \tag{19}$$

Step 2: Block Sampling $\phi^{tr,n}$ and $v^{tr,n}(t_d)$**-** After modeling Θ^n we model and update each dictionary function $\phi^{tr,n}$ from its posterior by marginalizing out $v^{tr,n}(t_d)$ and cycling through each $\phi_m^{tr,n}, m \in \{1,2,3,...,k\}$. This step calculates the posterior computations to obtain the dictionary functions. For a single trial tr of total J_n trials of one n-back task, we calculate each dictionary function $\phi_m^{tr,n}$ from the conditional posterior given $y^{tr,n}(t), \phi_{\neq m}^{tr,n}(t_{1:\tau}), \phi_m^{0,n}(t_{1:\tau}), \Theta^n, \zeta^n, \Sigma_0^n$ [where $\phi_{\neq m}^{tr,n}(t_{1:\tau})$ is the set of latent dictionary functions $\phi_e^{tr,n}$ such that $e \neq m$] as,

$$\phi_m^{tr,n}(t_{1:\tau}) | y^{tr,n}(t), \phi_{\neq m}^{tr,n}(t_{1:\tau}), \phi_m^{0,n}(t_{1:\tau}), \Theta^n, \zeta^n, \Sigma_0^n \sim \mathcal{N}_n\left(\begin{bmatrix} B_1 \\ \vdots \\ B_\tau \end{bmatrix}, \hat{\Sigma}_\phi^n\right), \tag{20}$$

$$\hat{\Sigma}_\phi^{-n} = Q_1^{-1} + \text{diag}(D_1, \cdots, D_\tau), \tag{21}$$

where

$$B_d = [C(t_d)]'_{.m} \Sigma^{-n}(t_d) \hat{y}^{tr,n}(t_d) + \phi^{0,n}(t_d); \ \forall d \in \{1, \cdots, \tau\}, \tag{22}$$

$$\hat{y}^{tr,n}(t) = y^{tr,n}(t) - \sum_{r \neq l} [C_n(t)]'_{.r} \phi_r^{tr,n}, \tag{23}$$

and

$$D_d = [C_n(t_d)]'_{.m}\Sigma^{-n}(t)[C_n(t_d)]_{.m};\ \forall d \in \{1, \cdots, \tau\}. \tag{24}$$

Based on $y^{tr,n}(t), \phi^{tr,n}(t), \Theta^n, \zeta^n, \text{and } \Sigma_0^n$ we independently compute $v^{tr,n}(t)$ as,

$$v^{tr,n}(t)|y^{tr,n}(t), \phi^{tr,n}(t), \Theta^n, \zeta^n, \text{and } \Sigma_0^n \sim \mathcal{N}_k(C_n(t)\Sigma_0^{-n}y^{tr,n}(t), I + C_n^{'}(t)\Sigma_0^{-n}C_n(t)), \tag{25}$$

where

$$\hat{y}^{tr,n}(t) = y^{tr,n}(t) - C_n(t)\phi^{tr,n}(t). \tag{26}$$

Step 3: Sampling $\phi^{0,n}(t_{1:\tau})$**-** Now, the underlying neural signal factors that is common for all trials (i.e. the mean of the child processes, $\phi_m^{0,n}(t_{1:\tau})$) is calculated as,

$$\phi_m^{0,n}(t_{1:\tau}) \sim \mathcal{N}_n(\Omega^{-n}\psi_m^n, \Omega^n), \tag{27}$$

$$\Omega^n = Q_0^{-1} + J_n \times Q_1^{-1}, \tag{28}$$

$$\psi_m^n = Q_1^{-1}\sum_{tr=1}^{J_n}\phi_m^{tr,n}(t_{1:\tau}). \tag{29}$$

Step 4: Sampling ζ^n**-** Model the $L \times k$ latent covariance dictionary functions ζ^n from equation (14) on the basis of $y^{tr,n}(t), \phi^{tr,n}(t), v^{tr,n}(t), \Theta^n, \zeta^n_{\neq l,m}, \Sigma_0^n$ ($\zeta^n_{\neq l,m}$ being the set of latent covariance dictionary elements $\zeta^n_{e,g}$ such that $e, g \neq l, m$),

$$\zeta^n_{l,m}(t_{1:\tau})|y^{tr,n}(t), \phi^{tr,n}(t), v^{tr,n}(t), \Theta^n, \zeta^n_{\neq l,m}, \Sigma_0^n \sim \mathcal{N}_n\left(\hat{\Sigma}_\zeta \times \sum_{tr=1}^{J_n}\begin{bmatrix} H_1 \\ . \\ . \\ H_d \\ . \\ . \\ H_\tau \end{bmatrix}, \hat{\Sigma}_\zeta\right), \tag{30}$$

where

$$H_d = x_{d,m}^{tr,n} \sum_{i=1}^{p} (\Theta_{i,l}^n) \sigma_{i,n}^{-2} \hat{y}_i^{tr,n}(t_d);\ \forall d \in \{1, \cdots, \tau\}, \tag{31}$$

$$\hat{\Sigma}_\zeta^{-1} = Q_1^{-1} + \sum_{tr=1}^{J_n} \text{diag}\bigg((x_m^{tr,n}(t_1))^2 \times O, \cdots (x_m^{j,n}(t_\tau))^2 \times O \bigg), \tag{32}$$

$$O = \sum_{i=1}^{p} (\Theta_{i,l}^n)^2 \sigma_{i,n}^{-2}, \tag{33}$$

and

$$\hat{y}_i^{tr,n}(t_d) = y_i^{tr,n}(t_d) - \sum_{(r,s) \neq (l,m)} \Theta_{i,r}^n \zeta_{r,s}^n(t_d); i = 1 : p. \tag{34}$$

Step 5: Sampling Σ_0^n- Now, to find Σ_0^n, consider

$$\Theta_{i.}^n = [\Theta_{i,1}^n, \cdots, \Theta_{i,L}^n], \tag{35}$$

and

$$x^{tr,n}(t) = \phi^{tr,n}(t) + v(t). \tag{36}$$

Then, as the diagonal elements of Σ_0^n have an inverse Gamma prior according to equation (16), the standard conjugate posterior can be derived as,

$$\sigma_{i,n}^{-2} \sim Ga(a_\sigma \frac{n \times tr}{2}, F), \tag{37}$$

where

$$F = b_\sigma + \frac{1}{2} \sum_{tr=1}^{J_n} \sum_{t_d=t_1}^{t_\tau} (y_i^{tr,n}(t) - \Theta_{i.}^n \zeta^n(t) x^{tr,n}(t))^2. \tag{38}$$

Derivation of Inverse Gamma Posteriors- We derive the conditional posteriors of the diagonal elements of Σ_0^n having an inverse Gamma prior [56]. For an independent and identically distributed x_i the conditional posterior can be derived as,

$$x_i | \mu, \sigma^2 \sim \mathcal{N}(\mu, \sigma^2), \tag{39}$$

where

$$\mu|\sigma^2 \sim \mathcal{N}(\mu_0, n_0\sigma^2), \tag{40}$$

and

$$\sigma^2 \sim \text{Ga}(\alpha, \beta), \tag{41}$$

where μ_0 is the initial mean, and $n_0\sigma^2$ is the initial variance.

As derived in [56], expressing the joint density $P(\sigma^2, \mu|x)$ as,

$$P(\sigma^2, \mu|x) \propto P(\sigma^2) \times P(\mu|\sigma^2) \times P(x|\sigma^2, \mu). \tag{42}$$

Equation (41) leads to a Gamma posterior for σ^2 as,

$$\sigma^2|x \sim \text{Ga}\Big(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum(x_i - \bar{x})^2 + \frac{nn_0}{2(n+n_0)}(\bar{x} - \mu_0)^2\Big). \tag{43}$$

For a complete derivation refer to Chapter 9 in [56]. Equation (37) and (38) can be derived in a similar manner.

Derivation of Gibbs Sampler- We derive the posterior distributions for sampling Gaussian Process dictionary elements as detailed in [55]. Combining equations (4) and (14) we have,

$$y_{tr} = \Theta \begin{bmatrix} \zeta_{11}(x_i)\zeta_{12}(x_i)\cdots\zeta_{1k}(x_i) \\ \zeta_{21}(x_i)\zeta_{22}(x_i)\cdots\zeta_{2k}(x_i) \\ \vdots \\ \zeta_{L1}(x_i)\zeta_{L2}(x_i)\cdots\zeta_{Lk}(x_i) \end{bmatrix} x_{tr} + \epsilon_{tr}. \tag{44}$$

Further expanding, we get the double summation

$$y_{tr} = \sum_{l=1}^{L}\sum_{m=1}^{k} \theta_{i,l} x_{tr,m}\zeta_{lm} + \epsilon_{tr,i}. \tag{45}$$

Conditioning on one single value ζ_{lm} equation (44) transforms to,

$$y_{tr} = x_{tr,m} \begin{bmatrix} \Theta_{1,l} \\ \vdots \\ \Theta_{p,L} \end{bmatrix} \zeta_{lm} + \tilde{\epsilon}_{tr}, \tag{46}$$

where

$$\tilde{\epsilon}_{tr} \sim \mathcal{N}\Bigg(\mu_{lm} \triangleq \begin{bmatrix} \sum_{(r,s)\neq(l,m)} \Theta_{1r} x_{tr,s} \zeta_{rs} \\ \vdots \\ \sum_{(r,s)\neq(l,m)} \Theta_{pr} x_{tr,s} \zeta_{rs} \end{bmatrix}, \Sigma_0 \Bigg). \tag{47}$$

Let $\Theta_{.l} = [\Theta_{1l} \cdots \Theta_{pl}]'$. Then, substituting in equation (46) we get

$$\begin{bmatrix} y(t_1) \\ \vdots \\ y(t_\tau) \end{bmatrix} = \begin{bmatrix} x_m(t_1)\Theta_{.l} 0 \cdots 0 \\ 0 x_m(t_2)\Theta_{.l} \cdots 0 \\ \vdots \\ 00 \cdots x_m(t_\tau)\Theta_{.l} \end{bmatrix} \begin{bmatrix} \zeta_{lm}(t_1) \\ \vdots \\ \zeta_{lm}(t_\tau) \end{bmatrix} + \begin{bmatrix} \tilde{\epsilon(t_1)} \\ \vdots \\ \tilde{\epsilon(t_\tau)} \end{bmatrix}. \tag{48}$$

We define $A_{lm} = \text{diag}(x_m(t_1)\Theta_{.l}, \cdots, x_m(t_\tau)\Theta_{.l})$. Thus, the conditional posterior distribution for ζ_{lm} based on $y^{tr,n}(t), x, \Theta^n, \zeta^n, \Sigma_0^n$ will be,

$$\begin{bmatrix} \zeta_{lm}(t_1) \\ \vdots \\ \zeta_{lm}(t_\tau) \end{bmatrix} | y^{tr,n}(t), x, \Theta^n, \zeta^n, \Sigma_0^n \sim \mathcal{N}\Big(\tilde{\Sigma} A'_{lm} diag(\Sigma_0^{-1} \cdots \Sigma_0^{-1}) \begin{bmatrix} \tilde{y}_{lm}(t_1) \\ \vdots \\ \tilde{y}_{lm}(t_\tau) \end{bmatrix}, \tilde{\Sigma} \Big), \tag{49}$$

where $\tilde{y} = y - \mu_{lm}(t_d)$ and if Q is the *Gram matrix*, then

$$\tilde{\Sigma}^{-1} = Q^{-1} + A'_{lm} \text{diag}(\Sigma_0^{-1} \cdots \Sigma_0^{-1}) A_{lm}. \tag{50}$$

Equation (48) can be derived by the standard derivations of posterior mean and variance in [56]. As derived in [56], conditional posterior for mean of a Gaussian distribution $x \sim$

$\mathcal{N}(\mu, \sigma^2)$, assuming a gauusian prior for the mean as $\mu \sim \mathcal{N}(\mu_0, \sigma_0^2)$ will be,

$$\mu_{post} = \mu_0 + \frac{\sigma^2}{\sigma^2 + \sigma_0^2}(x - \mu_0). \tag{51}$$

Similarly, the posterior for variance (σ^2_{post}) will be,

$$\frac{1}{\sigma^2_{post}} = \frac{1}{\sigma^2} + \frac{1}{\sigma_0^2}. \tag{52}$$

The equations (48) and (49) can be derived in a similar manner to that of equations (50) and (51). The derivations for the Latent dictionary function $\phi_m^{tr,n}(t_{1:\tau})$ can also be obtained in a similar manner to the above stated derivations of the latent covariance dictionary elements $\zeta_{lm}(t_1)$; equations(44) to (51).

Step 6: Sampling Θ^n- Based on latent precision matrices $\chi^n, y_i^{tr,n}(t), \phi^{tr,n}, v^{tr,n}(t_{1:\tau})$, and ζ^n according to the shrinkage prior stated in [48], $\Theta_{i.}^n$ is updated as,

$$\Theta_{i.}^n | \chi^n, y_i^{tr,n}(t), \phi^{tr,n}, v^{tr,n}(t_{1:\tau}), \text{ and } \zeta^n \sim \mathcal{N}_L(\sigma_{i,n}^{-2}\hat{\Sigma}_\Theta^{-n}\hat{x}^{n'}y_{.,i}, \hat{\Sigma}_\Theta^{-n}), \tag{53}$$

where

$$(\hat{\Sigma}_\Theta^{-1})^n = \sigma_{i,n}^{-2}\hat{x}^{n'}\hat{x}^n + \text{diag}(\chi_{i,1}^n\theta_1^n, \cdots \chi_{i,L}^n\theta_L^n), \tag{54}$$

$$\hat{x}^{tr'} = [G_{1,1}, \cdots, G_{n,1}, \cdots, G_{1,tr}, \cdots, G_{n,tr}], \tag{55}$$

$$G_{d,s} = \zeta^n(t_d)x_d^{s,n}; \forall d \in \{1, \cdots, \tau\}. \tag{56}$$

and

$$s \in \{1, \cdots, tr\}, \tag{57}$$

$$y_i^{tr,n} = [y_i^{1,n}(t_{1:\tau}), \cdots y_i^{tr,n}(t_{1:\tau})]'. \tag{58}$$

The above steps are represented in the algorithm as follows,

Algorithm 1 Sampling algorithm run for J_n trials, of each n-back experiment with p channels, $t_1, t_2, \cdots, t_\tau$ discrete time points in each trial tr, and L and k as the dictionary function dimensions. Let Q_i denote the Gram matrix where each element represents $q_i(t, t')$ for $i \in \{0, 1\}$.

for $j = 1 : J_n$ **do**
  for $l = 1 : L$ **do**
    Calculate $\phi^{tr,n}$ based on the equation (20) to equation (24)
  end for
  for $t = t_1 : t_\tau$ **do**
    Calculating $v^{tr,n}(t)$ using equation (25)
  end for
end for
for $l = 1 : L$ **do**
  Calculate $\phi^{0,n}$ using equation (27)
end for
for $l = 1 : L$ **do**
  for $m = 1 : k$ **do**
    Calculate the $L \times k$ Latent covariance dictionary
    function ζ^n using equation (30)
  end for
end for
for $i = 1 : p$ **do**
  Using equation (35) to equation (38) find the diagonal
  elements of Σ_0^n
end for
for $i = 1 : p$ **do**
  Sample Θ^n as in equation (53)
end for

3.2.4 Classification

In this section, we outline how binary classification is performed using the covariance matrices. We obtain and analyze the covariance matrix for one participant for one task. Recall that the covariance $\Sigma^n(t)$ is task-specific whereas the mean $\mu^{tr,n}(t)$ is trial-specific. The final result of a *1*-back task with 10 trials for one participant will be a $p \times p \times N$ covariance matrix. This matrix represents changes in the dynamic functional connectivity of different brain regions with respect to a given task. The obtained covariance matrix for one task at one time point k is a $p \times p$ matrix with diagonal values as 1. For N time points, we build a 3-dimensional covariance matrix. Since the covariance between channels 1 and 2

is the same as the covariance between channels 2 and 1, the matrix is symmetric in nature, i.e out of the $p \times p$ values obtained from the datasets [1] and [23], $\frac{1}{2} \times (p^2 - p)$ are unique. We take the mean, variance and third order central moment of these $\frac{1}{2} \times (p^2 - p)$ values across the N time points. Thus, we have a $3 \times \frac{1}{2} \times (p^2 - p)$ features that are used for classification. Each n-back task results in one covariance matrix.

To classify the obtained covariance matrices, we use three different supervised machine learning algorithms namely; support vector machine (SVM), k-nearest neighbors (KNN) and linear discriminant analysis (LDA) [45].

SVM: The SVM classifier represents the data as points in space such that separate categories are divided as clusters with wide gaps between them. For classification, we use the quadratic SVM classifier with 5 fold cross validation.

KNN: The KNN classifier measures similarity of the data point with its "k" nearest neighbors. For classification of the covariance matrices, we used a weighted KNN classifier with the k value set to 10. For all the classifications, a 5 fold cross validation is applied.

LDA: LDA, which uses a linear combination of features to separate two labels. LDA assumes both classes are distributed normally and that the covariances are identical for both classes. This leads to a linear combination of observations for classification. A 5 fold cross validation is used for LDA.

3.3 Results

We highlight the dimensions of covariance matrix elements denoting covariance between channel pairs for memory related tasks like the n-back. Figures 12, 13, and 14 show a slice of the 8×8 plot of covariance matrices for all participants in *Dataset 1*, 10 seconds after the start of the trial for *1*-back, *2*-back, and *3*-back, respectively. We see a marked difference in the *1*-back, *2*-back, and *3*-back, tasks. The covariances for the *1*-back task indicate similar functional connectivity across the majority of the channels. Comparing the *1*-back task and

2-back task, the changes in covariances between sensors indicate lesser covariance in the *1*-back task to the *2*-back task indicating that the neuronal connectivity is no longer similar and that different regions of the prefrontal cortex are functionally connected in processing the *2*-back task. The matrix for *3*-back task shows a digression from the other two tasks, indicating complex pathways that are dissimilar to each other.

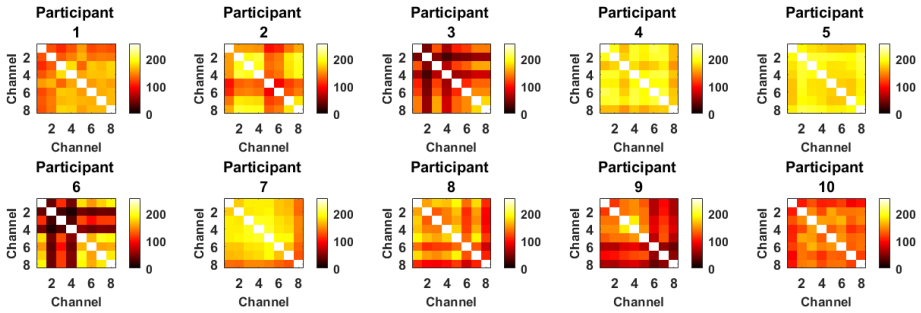

Figure 12: The 8×8 plot of covariance matrix for *Dataset 1* [1] representing the *1*-back task for all 10 participants, 10 seconds after the start of the task.

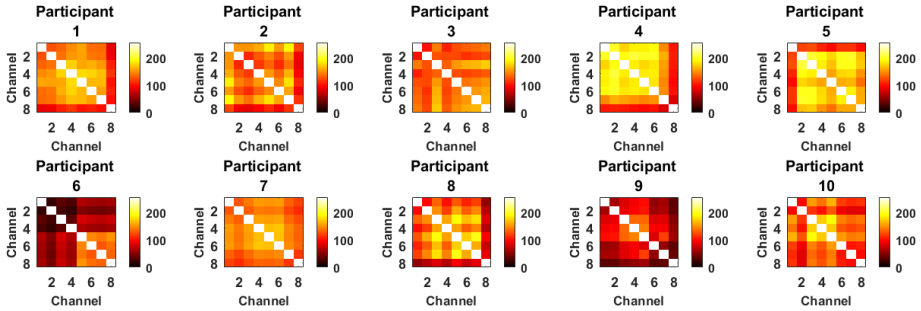

Figure 13: The 8×8 plot of covariance matrix for *Dataset 1* [1] representing the *2*-back task for all 10 participants, 10 seconds after the start of the task.

Figures 15, and 16 represent the matrices from the calming music session of *Dataset 2*, 10 seconds after the start of the trial, for *1*-back and *3*-back, respectively. Moreover Figures 17, and 18 represent the matrices from the vexing music session of *Dataset 2*, 10 seconds after the start of the trial, for *1*-back and *3*-back, respectively. We see a marked difference between the *1*-back and *3*-back tasks for both music sessions, but especially for the calming

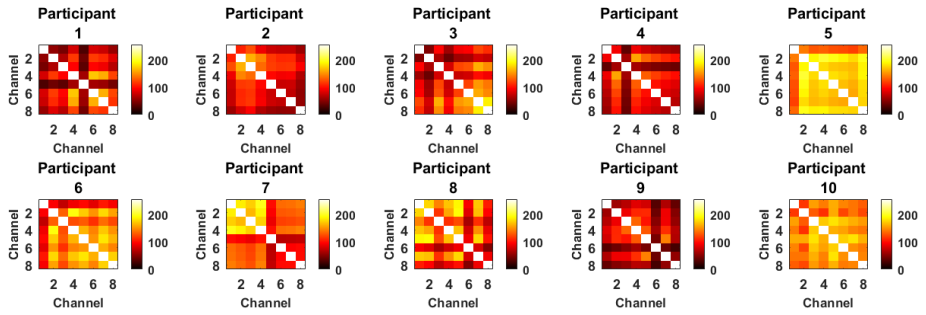

Figure 14: The 8×8 plot of covariance matrix for *Dataset 1* [1] representing the *3*-back task for all 10 participants, 10 seconds after the start of the task.

music session. The *3*-back task map exhibits more similarities than the *1*-back task. When compared to calming music maps, vexing music has less channels with similar covariance. Classification accuracies of the different n-back tasks in *Dataset 1* between the tasks and between each task and RELAX are summarized in Table 2 and Table 3, respectively.

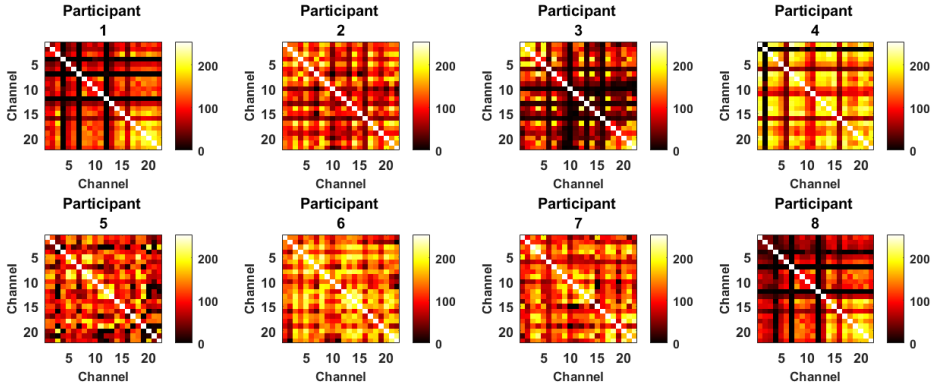

Figure 15: The 22×22 covariance matrix for the calming music session in *Dataset 2* representing the *1*-back task for all 8 participants, at 10 seconds after the start of the task.

Figure 19 offers a clear comparison between the three n-back tasks for participant 4, 10 seconds after the beginning of the trial. Similarly, Figure 20 is a comparison between the covariance maps obtained for the two n-back tasks for the calming and vexing music sessions.

Table 2 refers to the classification accuracies obtained for all the 3 tasks in the dataset

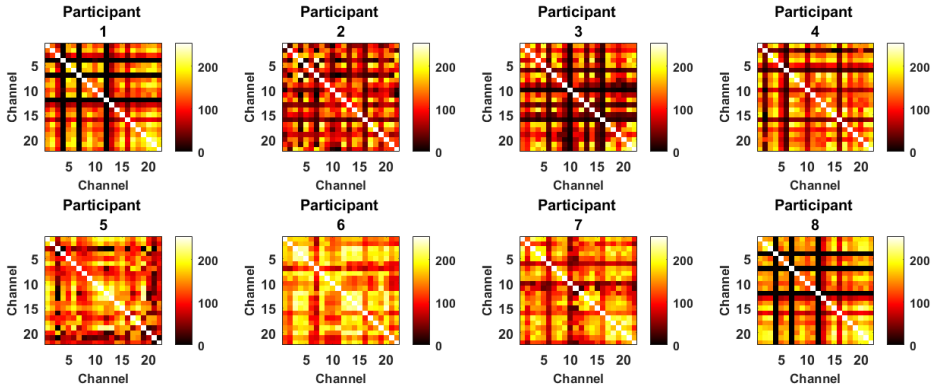

Figure 16: The 22×22 covariance matrix for the calming music session in *Dataset 2* representing the *3*-back task for all 8 participants, at 10 seconds after the start of the task.

Table 2: Classification accuracies for *1-back, 2-back, 3-back* tasks using mean, variance and third order moment of Σ_t^n.

Task	Accuracy in [1]	HLDM-SVM
1-back, 2-back	58.5%	65.6%
2-back, 3-back	61%	67.5%
1-back, 3-back	78.0%	75.0%

[1] and the accuracies obtained with our HLDL-SVM approach. Table 3 reports accuracies for the three n-back tasks against the RELAX signal where the participant's readings are taken as baseline.

Table 4 describes the classification accuracies between the two n-back tasks in *Dataset 2* with the backdrop of different music sessions. The KNN results outperformed other machine learning algorithms' results, hence we selected KNN for this classification. Table 5 depicts accuracies for the two n-back tasks in the calming music session of *Dataset 2* compared with Relax trials. In this case, all three of the classification accuracies, i.e. SVM, KNN and LDA have produced comparable results.

3.4 Discussion

Figure 12 depicts the covariance matrices of all 10 participants for the *1*-back task, 10 seconds after the start of the trial. We note that though there are a few uneven arbitrary

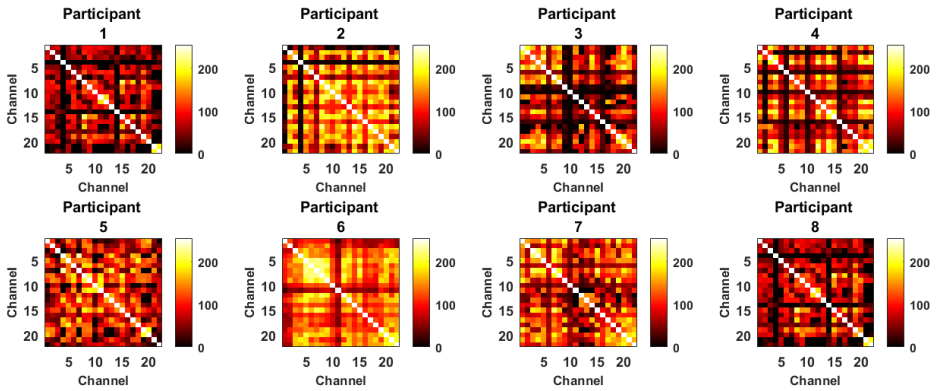

Figure 17: The 22×22 covariance matrix for the vexing music session in *Dataset 2* representing the *1*-back task for all 8 participants, at 10 seconds after the start of the task.

Table 3: Classification accuracies for *Dataset 1*, between *1*-back, *2*-back, *3*-back, tasks against RELAX task using mean, variance and third order moment of $\Sigma^n(t)$

Task	Accuracy in [1]	HLDM-SVM	HLDM-LDA	HLDM-KNN
1-back-RELAX	71.5%	75%	75%	65%
2-back-RELAX	80.3%	87.5%	95%	80%
3-back-RELAX	80.5%	87.5%	70%	75%

patterns of lower and higher covariances seen between the channels in participant 6, participant 3, and participant 9, the overall covariance matrices are even with highly similar values. In Figure 13, the covariance matrices of all 10 participants for the *2*-back task, taken 10 seconds after the beginning of the trial, are represented. Though the matrices for participant 1, and participant 4 are even, similar to the *1*-back task, we also see more patterns of lower covariance and higher covariance regions. This difference in covariance is starkly visible in the patterns formed for the covariance matrix of participant 6, participant 8, and participant 9. Figure 14 represents the covariance matrices for the *3*-back task for all 10 participants, 10 seconds after the beginning of the trial. We note that, when compared with Figure 12 and Figure 13, Figure 14 has more dissimilar values of covariance. We also see the emergence of more uneven arbitrary patterns in the covariance maps, such as seen

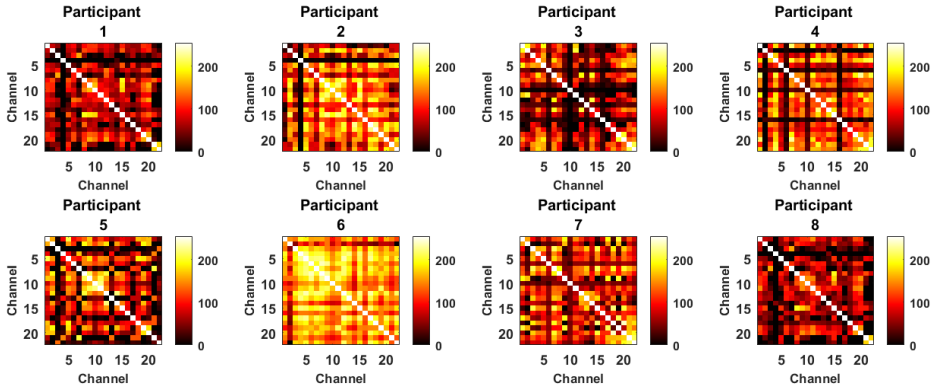

Figure 18: The 22×22 covariance matrix for the vexing music session in *Dataset 2* representing the *3*-back task for all 8 participants, at 10 seconds after the start of the task.

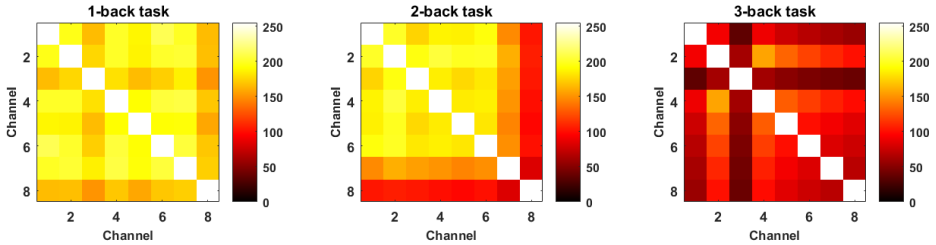

Figure 19: The 8×8 plot of covariance matrix for *Dataset 1* [1] comparing between the *1*-back, *2*-back, and *3*-back tasks for participant 4, 10 seconds after the start of the task.

for participant 1, participant3, participant 4, participant 8, and participant 9. Thus, from the aforementioned figures from *Dataset 1*, we see a steady increase in uneven arbitrary pattern formation and differences in the covariance values between channels as the n-back task's difficulty level increases. The *3*-back task has more patterns exhibited in the covariance maps compared with the evenness of the *1*-back task.

This increase in different covariance values between channels is also seen in the results obtained from *Dataset 2*. Figure 15 represents the covariance maps obtained for the *1*-back task for all 8 participants, during the calming music session, 10 seconds after the start of the trial. We note that the covariance values in *3*-back task for calming music session

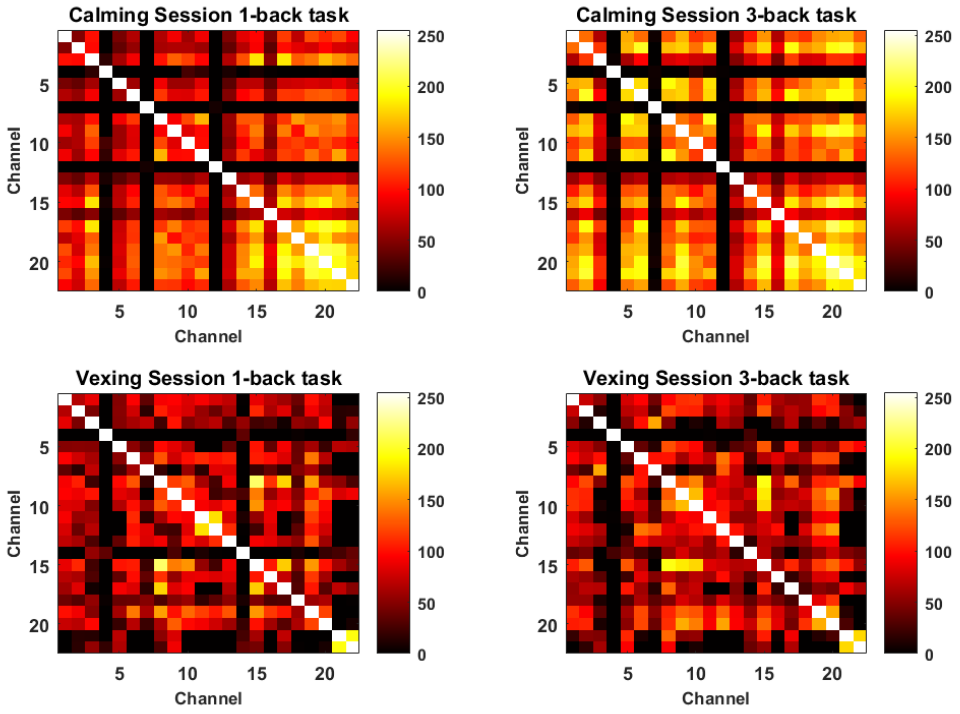

Figure 20: The 22 × 22 plot of covariance matrix for *Dataset 2* comparing between the calming music session *1*-back, and *3*-back tasks (top row) and the vexing music session *1*-back, and *3*-back tasks (bottom row) for participant 1, 10 seconds after the start of the task.

represented in Figure 16, are uneven when compared with the *1*-back task in calming music session's covariance maps represented in Figure 15. This can be seen in the covariance maps of participant 3, participant 4, participant 5, participant 6, and participant 7. Figure 16, which represents the covariance maps obtained for the *3*-back task, during the calming music session, 10 seconds after the beginning of the trial, exhibits more varied covariance maps than in Figure 15. This could be due to the difference in the difficulty level of performing *1*-back and the *3*-back tasks.

Figure 17 represents the covariance maps of *1*-back task for all 8 participants, 10 seconds after the start of the trial. We note that, when compared with the calming music *1*-back covariances represented in Figure 15, the vexing music *1*-back covariances represented in Figure 17 have many more patterns in them. Figure 18 is the depiction of covariance

Table 4: Classification accuracies for *Dataset 2* between different n-back tasks using mean and variance and different machine learning algorithms.

Task	HLDM-KNN
Calm *1*-back-Calm *3*-back	81.3%
Vex*1*-back-Vex *3*-back	62.5%

Table 5: Classification accuracies for *Dataset 2* between different n-back tasks and RELAX trials using mean and variance and different machine learning algorithms taking.

Task	HLDM-SVM	HLDM-LDA	HLDM-KNN
Calm *1*-back-RELAX	87.5%	87.5%	68.8%
Calm *3*-back-RELAX	75%	65%	75%

matrices for the vexing music session for all 8 participants, 10 seconds after the beginning of the trial. Similar to the differences observed between the calming music *1*-back from Figure 15 and the vexing music *1*-back from Figure 17, the vexing music *3*-back covariances from Figure 18 shows more varied covariance maps than the calming music *3*-back seen in Figure 16. This increase in difference and patterns could be due to the influences of different music backgrounds, i.e. the calming session *1*-back and *3*-back tasks shown in Figure 15 and Figure 16, respectively are more even due to the calming music influence whereas Figure 17, and Figure 18 have more variations because of the influence of vexing music.

Similarly, we note the increased differences between the two different tasks' covariance maps in the vexing music session. Though it is harder to discern than in the calming music session, we see the increased variations in participant 5, and participant 7 covariance maps between Figure 17 and Figure 18. This increased variations could be due to the difference in performing the n-back tasks. We see the comparisons more accurately in Figure 19, and Figure 20. Figure 19 depicts the three n-back tasks for participant 4 in *Dataset 1*. We see that the *1*-back task is relatively even, whereas the *2*-back task has some uneven patterns and lower covariance values. The *3*-back task has more arbitrary patterns. Likewise, from Figure 20 we note that the calming music session's n-back tasks are more uniform, whereas

the vexing music session's n-back tasks have more uneven patterns among the channel covariances. This could indicate the disrupting effect of vexing music on the participant's brain functional connectivity. We observe the increase in unevenness between the *1*-back and *3*-back task during the calming music session. However, the same cannot be observed for *1*-back and *3*-back tasks in vexing music session, as the unevenness overwhelms the difference. Uneven covariances could be the indications of the increase in difficulty level between the tasks.

Table 2 depicts the classification accuracies between the three n-back tasks. The *1*-back task and the *3*-back task have the highest accuracy, as is expected due to the differences in their difficulty levels. The *2*-back and *3*-back tasks are more closer in difficulty levels, hence the accuracy of 67.5% between them. The observed increase in unevenness of the covariance matrices with increasing difficulty level of the n-back task, could result in these classification accuracies. In Table 3, we classify the 3 n-back tasks against the RELAX signal where the rest period is taken as baseline. In this case, we are able to achieve better accuracy while distinguishing *1*-back, *2*-back, and *3*-back, tasks against RELAX compared with results reported in [1]. An explanation for this could be that the unevenness of the covariance maps during the n-back tasks is more than that of the RELAX covariance maps.

Table 4 describes the classification accuracies obtained for *Dataset 2* between the *1*-back and *3*-back tasks for each session. As is evident from the difference in classification between tasks during calming music session and vexing music session, there is a higher ability to differentiate between tasks during the calming music and the vexing music session. This could also be due to the increased unevenness in the vexing session's covariance maps disrupting the classification. The accuracies obtained by using KNN for the calming music session are similar to the ones obtained from *Dataset 1*. In presence of vexing music, the differences in the hemodynamic covariances between *1*-back and *3*-back tasks, as seen in Figure 17 and Figure 18, are not as evident as they are during calming music session's *1*-back and *3*-back task, as seen in Figure 15 and Figure 16. Therefore, the classification

accuracies for vexing music session are lower than that of calming music session. Influence of calming music could be a possible reason for the increased accuracy between *1*-back and *3*-back tasks, comparing Table 4 with Table 2.

In Table 5, the results depict the difference between the n-back tasks and the RELAX trials by using 3 different machine learning algorithms; namely SVM, KNN and LDA. While the accuracies are higher between *1*-back and RELAX than *3*-back and RELAX during the calming music, they are comparable to the task and RELAX classification values from *Dataset 1* in Table 3. This could also be the result of increasing unevenness in the covariance matrices as the difficulty level n, increases. The classification accuracies obtained verify the viability of our method for classifying mental workload. The HLDM method can be applied to other mental workload and cognitive engagement tasks to distinguish cognitive engagement in BCI. Though the results might be of varying strength depending on the participant's preference of music, an overall improvement of classification accuracy between *1*-back and *3*-back tasks while using calming music is evident. Although we discuss the increase in unevenness with the increase in the difficulty level of the n-back task, and the difference in evenness between the calming music session covariances and the vexing music session covariances, this is an observation for limited participants, and might not represent a general scenario. As is the case for participant 3, participant 6, and participant 9 in the *1*-back task for *Dataset 1* in Figure 12, which have uneven covariance values between channels. The same interpretation can be made for the even covariance maps of participant 1, participant 4 and participant 7 for the *2*-back task in Figure 13. Figure 14 also has even covariance maps seen in participant 3, and participant 7. These inconsistencies might be because different people experience mental workload differently. Due to several factors such as interpersonal variability, sensor placement variability, quality of signal acquisition, and day to day variation of a person's brain activity, the functional connectivity corresponding to specific mental workload condition may vary. To overcome this, future studies should include more accurate ways of quantifying mental workload on a case by case basis. A

larger dataset with several participants will also lead to a better understanding of the changes in covariance maps. This will provide a stronger understanding of the person to person variations in mental workload experiences. Other drawbacks of this study are that it is an offline approach. Real time fNIRS analysis will require modifications to the HLDM method. In future this approach can be extended to a real time version such that it can be applied to detect levels of cognitive engagement while learning in online classrooms, or to quantify workload during a memory task being performed.

4 Conclusion and Future Work

4.1 Conclusion

Variations in the brain's blood oxygenation and deoxygenation reflect neuronal activation patterns and can be measured using fNIRS. In this thesis, we utilized fNIRS to obtain insights into the dynamic functional connectivity of the brain as a function of the mental workload. We performed a memory task-based experiment, which explores the influence of different types of music such as calming and vexing, on different mental workloads defined by the difficulty of the memory task. We also collect multiple physiological signals along with fNIRS recordings in our experiment to gain a more comprehensive understanding of the effects of mental workload. In order to obtain the evolving functional connectivity from the experimental recordings, we model the fNIRS recordings with the hierarchical latent dictionary model approach to extract the hemodynamic covariances in different regions of the brain for different time instances throughout the experiment. We derived features from the dynamic functional connectivity of the brain reflected in fNIRS data collected. Along with our experimental data, we analyzed a publicly available dataset with fNIRS recordings from the prefrontal cortex to investigate mental workload as. We also compared our study with previous results.

4.1.1 A Working Memory Neurophysiological Dataset Incorporating the Influence of Music

Quantifying mental workload can lead to a better understanding of the effects that underwhelming or overwhelming stress has on cognitive engagement. It is therefore important to detect, classify, and improve the mental workload of an individual. There exists the need to develop a complete understanding of how human physiology responds while experiencing cognitive stress. Therefore, in this study, we performed a mental workload experiment to collect multi-modal physiological signals along with brain hemodynamic response signals

from participants. Participants are asked to perform a working memory task with different difficulty levels. The difficulty levels were included to ensure different levels of cognitive engagement during the experiment. We used fNIRS to record the brain hemodynamic response to retain a high spatial resolution. By incorporating multiple physiological signals in the experimental data collection, we derive a rich neurophysiological dataset that can offer a more complete comprehension of the body's reaction to cognitive stress. Additionally, there is a need to quantify the influence of music on cognitive tasks. Several studies theorize the use of music as a positive influence that can reduce cognitive stress. In this dataset, we incorporate the influence of music by conducting the cognitive engagement task in two sessions-one with calming music, and the other with vexing music playing in the background. We decided to use personalized music selected by the participants for calming and vexing music. We performed preliminary statistical analysis on the physiological signals as well in order to obtain an understanding of the physiological response. Our preliminary analysis of the on the physiological signals shows a significant difference between different conditions. We also conclude that the systematic analysis of different physiological signals can lead to further insights into the physiological response.

4.1.2 Experimental Validation of Mental Workload Assessment Using Hierarchical Latent Dictionary Method

Analysis of the fNIRS signal in order to obtain insight into the brain is a challenging problem. One popular approach to obtain insight is to determine how the functional connectivity of different regions of the brain change depending on the mental workload of interest. In this research, we apply a novel method to model fNIRS signal with hidden low dimensional neural factors of the brain, to investigate how the dynamic functional connectivity evolves in the brain during a mental workload task. We utilize hidden latent factor models with modeling underlying neural signal factors as Gaussian processes to obtain time-varying covariance matrices of the memory-related tasks known as the n-back tasks. The

covariance matrices obtained using hierarchical latent dictionary model on data collected have demonstrated the effect of calming and vexing music on mental workload. We also verify the variations in the different difficulty levels of the tasks reflected in the obtained dynamic functional connectivity represented with covariance matrices, both in the publicly available dataset and in our experimental dataset. We extracted features from the obtained dynamic functional connectivity, i.e. time-evolving covariance matrix, to train the machine learning model. We then use the features obtained from the covariance matrices in order to classify different mental workloads with several machine learning algorithms, namely SVM, KNN, and decision trees. We show our approach can successfully classify different difficulty levels of the n-back memory tasks and discern the influence of music on these memory tasks. Our results verify the applicability of the method and its usefulness in classifying mental workload tasks.

4.2 Future Work

EDA, skin temperature, ECG, EMG, and respiration are the most widely used physiological markers for estimating the emotional and cognitive stress of an individual. Future avenues of work derived from this study include expanding the classification of mental workload by incorporating several physiological signals in addition to fNIRS. The experimental dataset obtained and expounded in this thesis also includes mental workload tasks conducted under the influence of two styles of music. Since the participants chose the tracks that they find calming and vexing themselves, this dataset can be reliably used as a comprehensive physiological investigation of mental workload and its effects. Additionally, other neuroimaging techniques such as EEG-fNIRS hybrid systems can also provide a rich dataset to further extend our understanding of cognitive stress and its effect on performance. A future potential direction is extending the current work by joint modeling of both EEG and fNIRS recordings with hierarchical latent dictionary model so that both higher temporal and spatial resolution can be obtained.

Bibliography

[1] Christian Herff, Dominic Heger, Ole Fortmann, Johannes Hennrich, Felix Putze, and Tanja Schultz. Mental workload during n-back task quantified in the prefrontal cortex using fnirs. *Frontiers in Human Neuroscience*, 7:935, 2014.

[2] Mark S Young, Karel A Brookhuis, Christopher D Wickens, and Peter A Hancock. State of science: mental workload in ergonomics. *Ergonomics*, 58(1):1–17, 2015.

[3] Luca Longo. A defeasible reasoning framework for human mental workload representation and assessment. *Behaviour & Information Technology*, 34(8):758–786, 2015.

[4] Brad Cain. A review of the mental workload literature. Technical report, Defence Research And Development Toronto (Canada), 2007.

[5] Sarah N Abdulkader, Ayman Atia, and Mostafa-Sami M Mostafa. Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, 16(2):213–230, 2015.

[6] A Marinescu, S Sharples, AC Ritchie, T Sánchez López, M McDowell, and H Morvan. Exploring the relationship between mental workload, variation in performance and physiological parameters. *IFAC-PapersOnLine*, 49(19):591–596, 2016.

[7] Haibin Zhu and Ming Hou. Restrain mental workload with roles in hci. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, pages 387–392. IEEE, 2009.

[8] Jan Jarvis, Felix Putze, Dominic Heger, and Tanja Schultz. Multimodal person independent recognition of workload related biosignal patterns. In *Proceedings of the 13th international conference on multimodal interfaces*, pages 205–208, 2011.

[9] Miguel A Recarte and Luis M Nunes. Mental workload while driving: effects on visual search, discrimination, and decision making. *Journal of experimental psychology: Applied*, 9(2):119, 2003.

[10] Hasan Ayaz, Ben Willems, B Bunce, Patricia A Shewokis, Kurtulus Izzetoglu, Sehchang Hah, A Deshmukh, and Banu Onaral. Cognitive workload assessment of air traffic controllers using optical brain imaging sensors. *Advances in understanding human performance: Neuroergonomics, human factors design, and special populations*, pages 21–31, 2010.

[11] Elizabeth A Felton, Justin C Williams, Gregg C Vanderheiden, and Robert G Radwin. Mental workload during brain–computer interface training. *Ergonomics*, 55(5):526–537, 2012.

[12] Raphaëlle N Roy, Stephane Bonnet, Sylvie Charbonnier, and Aurélie Campagne. Mental fatigue and working memory load estimation: interaction and implications for eeg-based passive bci. In *2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6607–6610. IEEE, 2013.

[13] Ayca Berfu Ünal, Linda Steg, and Kai Epstude. The influence of music on mental effort and driving performance. *Accident Analysis & Prevention*, 48:271–278, 2012.

[14] Ayça Berfu Ünal, Dick de Waard, Kai Epstude, and Linda Steg. Driving with music: Effects on arousal and performance. *Transportation research part F: traffic psychology and behaviour*, 21:52–65, 2013.

[15] Maxim S Kuschpel, Shuyan Liu, Daniel J Schad, Stephan Heinzel, Andreas Heinz, and Michael A Rapp. Differential effects of wakeful rest, music and video game playing on working memory performance in the n-back task. *Frontiers in psychology*, 6:1683, 2015.

[16] Rong-Hwa Huang and Yi-Nuo Shih. Effects of background music on concentration of workers. *Work*, 38(4):383–387, 2011.

[17] Alona Fyshe, Emily Fox, David Dunson, and Tom Mitchell. Hierarchical latent dictionaries for models of brain activation. In Neil D. Lawrence and Mark Girolami, editors, *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, volume 22 of *Proceedings of Machine Learning Research*, pages 409–421, La Palma, Canary Islands, 21–23 Apr 2012. PMLR.

[18] Katrina Wendel, Outi Väisänen, Jaakko Malmivuo, Nevzat G Gencer, Bart Vanrumste, Piotr Durka, Ratko Magjarevic, Selma Supek, Mihail Lucian Pascu, Hugues Fontenelle, et al. Eeg/meg source imaging: methods, challenges, and open issues. *Computational intelligence and neuroscience*, 2009:13, 2009.

[19] Chris Berka, Daniel J. Levendowski, Michelle N. Lumicao, Alan Yau, Gene Davis, Vladimir T. Zivkovic, Richard E. Olmstead, Patrice D. Tremoulet, and Patrick L. Craven. Research article eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks.

[20] Carryl L. Baldwin and B.N. Penaranda. Adaptive training using an artificial neural network and eeg metrics for within- and cross-task workload classification. *NeuroImage*, 59(1):48 – 56, 2012. Neuroergonomics: The human brain in action and at work.

[21] Angel Jimenez-Molina, Cristian Retamal, and Hernan Lira. Using psychophysiological sensors to assess mental workload during web browsing. *Sensors*, 18(2):458, 2018.

[22] Andreas Henelius, Kati Hirvonen, Anu Holm, Jussi Korpela, and Kiti Muller. Mental workload classification using heart rate metrics. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1836–1839. IEEE, 2009.

[23] Jaeyoung Shin, Alexander von Lühmann, Do-Won Kim, Jan Mehnert, Han-Jeong Hwang, and Klaus-Robert Müller. Simultaneous acquisition of eeg and nirs during cognitive tasks for an open access dataset. In *Scientific data*, 2018.

[24] Yoko Hoshi, Brian H Tsou, Vincent A Billock, Masato Tanosaki, Yoshinobu Iguchi, Miho Shimada, Toshikazu Shinba, Yoshifumi Yamada, and Ichiro Oda. Spatiotemporal characteristics of hemodynamic changes in the human lateral prefrontal cortex during working memory tasks. *Neuroimage*, 20(3):1493–1504, 2003.

[25] Hendrik Santosa, Xuetong Zhai, Frank Fishburn, and Theodore Huppert. The nirs brain analyzir toolbox. *Algorithms*, 11(5):73, 2018.

[26] Janina AM Lehmann and Tina Seufert. The influence of background music on learning in the light of different theoretical perspectives and the role of working memory capacity. *Frontiers in psychology*, 8:1902, 2017.

[27] Rose T Faghih, Patrick A Stokes, Marie-France Marin, Rachel G Zsido, Sam Zorowitz, Blake L Rosenbaum, Huijin Song, Mohammed R Milad, Darin D Dougherty, Emad N Eskandar, et al. Characterization of fear conditioning and fear extinction by analysis of electrodermal activity. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 7814–7818. IEEE, 2015.

[28] D S Wickramasuriya, C Qi, and R. T. Faghih. A state-space approach for detecting stress from electrodermal activity. In *Conference proceedings:... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, volume 2018, pages 3562–3567, 2018.

[29] Dilranjan S Wickramasuriya and Rose T Faghih. Online and offline anger detection via electromyography analysis. In *Healthcare Innovations and Point of Care Technologies (HI-POCT), 2017 IEEE*, pages 52–55. IEEE, 2017.

[30] Xinyi Deng, Rose T Faghih, Riccardo Barbieri, Angelique C Paulk, Wael F Asaad, Emery N Brown, Darin D Dougherty, Alik S Widge, Emad N Eskandar, and Uri T Eden. Estimating a dynamic state to relate neural spiking activity to behavioral signals during cognitive tasks. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 7808–7813. IEEE, 2015.

[31] Rose T Faghih. From physiological signals to pulsatile dynamics: a sparse system identification approach. In *Dynamic Neuroscience*, pages 239–265. Springer, 2018.

[32] Md. Rafiul Amin and Rose T Faghih. Inferring autonomic nervous system stimulation from hand and foot skin conductance measurements. In *Asilomar Conference on*

Signals, Systems, and Computers, 2018.

[33] Myriam V Thoma, Roberto La Marca, Rebecca Brönnimann, Linda Finkel, Ulrike Ehlert, and Urs M Nater. The effect of music on the human stress response. *PloS one*, 8(8), 2013.

[34] ARVIND Gupta. The interesting connection between math and music. *The Vancouver Sun*, 2009.

[35] Arno Villringer, J Planck, C Hock, L Schleinkofer, and U Dirnagl. Near infrared spectroscopy (nirs): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neuroscience letters*, 154(1-2):101–104, 1993.

[36] Angelo Sassaroli and Sergio Fantini. Comment on the modified beer–lambert law for scattering media. *Physics in Medicine Biology*, 49(14):N255, 2004.

[37] Siamac Fazli, Jan Mehnert, Jens Steinbrink, Gabriel Curio, Arno Villringer, Klaus-Robert Müller, and Benjamin Blankertz. Enhanced performance by a hybrid nirs–eeg brain computer interface. *NeuroImage*, 59(1):519 – 529, 2012. Neuroergonomics: The human brain in action and at work.

[38] Robert Oostenveld and Peter Praamstra. The five percent electrode system for high-resolution eeg and erp measurements. *Clinical Neurophysiology*, 112(4):713 – 719, 2001.

[39] Jacqueline Wijsman, Bernard Grundlehner, Julien Penders, and Hermie Hermens. Trapezius muscle emg as predictor of mental stress. In *Wireless Health 2010*, pages 155–163. 2010.

[40] S. Parshi, R. Amin, H. F. Azgomi, and R. T. Faghih. Mental workload classification via hierarchical latent dictionary learning: A functional near infrared spectroscopy study. In *2019 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 1–4, May 2019.

[41] Xiaozheng Zhang, Charles C Broun, Russell M Mersereau, and Mark A Clements. Automatic speechreading with applications to human-computer interfaces. *EURASIP Journal on Advances in Signal Processing*, 2002(11):240192, 2002.

[42] Jan Van Erp, Fabien Lotte, and Michael Tangermann. Brain-computer interfaces: beyond medical applications. *Computer*, 45(4):26–34, 2012.

[43] Noman Naseer and Keum-Shik Hong. fnirs-based brain-computer interfaces: a review. *Frontiers in human neuroscience*, 9:3, 2015.

[44] Jialin Fan and Andrew P Smith. The impact of workload and fatigue on performance. In *International Symposium on Human Mental Workload: Models and Applications*, pages 90–105. Springer, 2017.

[45] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer, 2003.

[46] Stefan Bergman. *The kernel function and conformal mapping*, volume 5. American Mathematical Soc., 1970.

[47] JM Bernardo, MJ Bayarri, JO Berger, AP Dawid, D Heckerman, AFM Smith, and M West. Bayesian factor regression models in the "large p, small n" paradigm. *Bayesian statistics*, 7:733–742, 2003.

[48] A Bhattacharya and D B Dunson. Sparse bayesian infinite factor models. *Biometrika vol. 98,2 (2011): 291-306.*

[49] Carlos M. Carvalho, Jeffrey Chang, Joseph E. Lucas, Joseph R. Nevins, Quanli Wang, and Mike West. High-dimensional sparse factor modeling: Applications in gene expression genomics. *Journal of the American Statistical Association*, 103(484):1438–1456, 2008. PMID: 21218139.

[50] David Knowles and Zoubin Ghahramani. Infinite sparse factor analysis and infinite independent components analysis. In *Proceedings of the 7th International Conference on Independent Component Analysis and Signal Separation*, ICA'07, pages 381–388, Berlin, Heidelberg, 2007. Springer-Verlag.

[51] Hedibert Freitas Lopes, Esther Salazar, and Dani Gamerman. Spatial dynamic factor analysis. *Bayesian Anal.*, 3(4):759–792, 12 2008.

[52] Mike West. Bayesian factor regression models in the "large p, small n" paradigm. In *Bayesian Statistics*, pages 723–732. Oxford University Press, 2003.

[53] Emily B. Fox and David B. Dunson. Bayesian nonparametric covariance regression. *Journal of Machine Learning Research*, 16:2501–2542, 2015.

[54] Anirban Bhattacharya and David B Dunson. Sparse bayesian infinite factor models. *Biometrika*, pages 291–306, 2011.

[55] Emily Fox and David Dunson. Bayesian nonparametric covariance regression. *arXiv preprint arXiv:1101.2017*, 2011.

[56] Christian Robert. *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, 2007.