UNIVERSITY OF CALIFORNIA SAN DIEGO

Analysis of Behavioral and Autonomic States in Unstructured Multiday Human Intracranial Electrophysiology

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Electrical Engineering
(Signal and Image Processing)

by

Abdulwahab A. Alasfour

Committee in charge:

Professor Vikash Gilja, Chair
Professor Eric Halgren
Professor Kenneth Kreutz-Delgado
Professor Piya Pal
Professor Bhaskar Rao

2020

The dissertation of Abdulwahab A. Alasfour is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

2020

EPIGRAPH

One has to strive with all his diligence and he does not have to be successful, for success is in Allah's hand.

—My Father

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It has been a long and arduous journey that has culminated in this thesis. I came to San Diego 10 years ago as a freshman in the ECE department, not knowing where the road will take me and not quite sure what I am going to end up doing with my life. Without the guidance of God, the people around me, and some luck along the way, this thesis would not have come to fruition. I want to take this opportunity to thank everybody who has made it possible for me to complete my Ph.D.

Firstly, I would like to thank my mentor and Ph.D. advisor Professor Vikash Gilja. Vikash's tutelage and mentorship throughout my graduate journey have been invaluable. He always made me feel like a colleague and pushed me to explore ideas that pushed the boundaries in the field of neural engineering. Vikash gave me space to grow and never forced me to focus on a particular subject. I heard many stories about students being disenfranchised with academia and the doctorate degree as a whole. With Vikash, he made me realize how much I loved exploring, learning, and figuring things out. I hope this thesis is just a start of many collaborations for us in the future.

Secondly, I would like to sincerely thank Professor Eric Halgren for welcoming me into his lab and being my mentor for two years in my doctorate study. I was an engineer without any neurophysiology background and felt utterly out of my depth when I started working in his lab. His patience and advice slowly gave me confidence and the belief that I can tackle any problem, no matter the subject, given time and effort. Working with Eric has broadened my view on how to approach problems from a neuroscientific point of view. As a researcher, I grew tremendously under his mentorship, and I am eternally grateful for being part of his lab.

Thank you to all the Translational Neuroengineering Lab members, past and present, for making my journey both mentally stimulating and enjoyable. The camaraderie between us was tangible, and there was always a sense of family and community. Thank you to Abdullah Alothman, Pablo Tostado, Daril Brown, Tejaswi Pailla, Aashish Patel, Kenny Chen, Paolo Gabriel, and to all TNEL members for being the most amazing labmates a graduate student can ever hope for, and to Nikhil Das for always providing me with a coffee run so we can vent about our Ph.D. woes. Also, thank you for all the members of the Radiology Imaging Laboratory for expanding my horizons.

I want to thank my family, for, without them, this would have been impossible. Thank

you, Father, for ingraining in me the idea that the pursuit of knowledge is the ultimate honor, and for pushing me to pursue a doctorate degree. Thank you to my brother, Mahmoud, for constantly checking up on me and always making sure I took care of myself. Thank you, Mother, for your patience and sacrifice. Without your undying love and support, I am not quite sure that I could have made this far. Finally, I would like to thank my beautiful wife, Dalal, for your relentless love, support, compassion, and empathy. You left home to spend three years abroad with me, and they were the best three years of my life. I am eternally indebted to you. Thank you, my love.

The chapters of this dissertation consist of a published journal article and manuscripts to be submitted to journals. The dissertation author was the primary investigator for all of the works listed:

- Chapter 2 is a reprint of the material as it appears in Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan, Daniel Friedman, Werner Doyle, Orin Devinsky, David Gonda, Shifteh Sattar, Sonya Wang, Eric Halgren, and Vikash Gilja. "Coarse behavioral context decoding." Journal of Neural Engineering 16, no. 1 (2019): 016021.

- Chapter 3 is a reprint of the material as it appears in Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan, Daniel Friedman, Werner Doyle, Orin Devinsky, David Gonda, Shifteh Sattar, Sonya Wang, Eric Halgren, and Vikash Gilja. "Identifying neural signal characteristics that discriminate between naturalistic behavioral states." Manuscript in preparation, to be submitted to PLOS Computational Biology.

- Chapter 4 is a reprint of the material as it appears in Alasfour, Abdulwahab, Xi Jiang, Jorge Gonzalez-Martinez, Vikash Gilja, Eric Halgren "High gamma activity in cortex and hippocampus is correlated with autonomic tone during sleep." Manuscript in preparation, to be submitted to the Journal of Neuroscience.

VITA

2014 - Bachelor of Science in Electrical and Computer Engineering, University of California San Diego

2016 - Master of Science in Electrical Engineering (Signal and Image Processing), University of California San Diego

2020 - Doctor of Philosophy in Electrical Engineering (Signal and Image Processing), University of California San Diego

2020 - Assistant Professor, Kuwait University

PUBLICATIONS

Chen, Kenny, Paolo Gabriel, **Abdulwahab Alasfour**, Chenghao Gong, Werner K. Doyle, Orrin Devinsky, Daniel Friedman et al. "Patient-specific pose estimation in clinical environments." *IEEE Journal of Translational Engineering in Health and Medicine* 6 (2018): 1-11.

Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan et al. "Coarse behavioral context decoding." *Journal of Neural Engineering* 16, no. 1 (2019): 016021.

Gabriel, Paolo Gutierrez, Kenny Chen, **Abdulwahab Alasfour**, Tejaswy Pailla, Werner Doyle, Orrin Devinsky, Daniel Friedman et al. "Neural correlates of unstructured motor behaviors." *Journal of Neural Engineering* (2019).

Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan et al. "Identifying neural signal characteristics that discriminate between naturalistic behavioral states." (in preparation to be submitted to PLOS Computational Biology, 2020).

Alasfour, Abdulwahab, Xi Jiang, Jorge Gonzalez-Martinez, Vikash Gilja, Eric Halgren "High gamma activity in cortex and hippocampus is correlated with autonomic tone during sleep." (in preparation to be submitted to the Journal of Neuroscience, 2020).

ABSTRACT OF THE DISSERTATION

Analysis of Behavioral and Autonomic States in Unstructured Multiday Human Intracranial Electrophysiology

by

Abdulwahab A. Alasfour

Doctor of Philosophy in Electrical Engineering
(Signal and Image Processing)

University of California San Diego, 2020

Professor Vikash Gilja, Chair

Systems neuroscience and neural engineering have relied mainly on the task-based experimental paradigm to understand brain activity. While this method has proved fruitful, it fails to capture the full variability of unstructured and naturalistic neural activity. In this thesis, we explore the value of investigating multiday unstructured intracranial electrophysiology recordings in humans. Using stereotactic-electroencephalography (sEEG) and electrocorticography (ECoG) electrodes, we analyze days of neural recordings to investigate how internal and external states, such as autonomic tone and behavior, correlate to neural activity. Firstly, we determine whether coarsely labeled unstructured behavioral contexts or states are discriminable in the neural activity space. Subjects were not instructed to perform any task; therefore, only spontaneous behaviors were analyzed. Controls to determine whether temporal correlations and time of day effects impact the separability of behavioral states were investigated, concluding that both the time of day and behavior have a combined

effect on neural activity. Secondly, once we determined that these behavioral states are separable, the cause of this separability was further investigated. In other words, what neural signal characteristics are responsible for our ability to decode abstract behavioral states? Both long term signal characteristics and spatiotemporal dynamics contribute information regarding naturalistic behavior, showing that outside the lab, neural activity has multiple axes of variability that could be used to discriminate between different states. In the final section of this work, we investigate the neural correlates to autonomic tone during sleep, leveraging multiple days of unstructured neural activity to make physiological conclusions regarding the connection between the central and autonomic nervous systems.

Chapter 1

Introduction

1.1 Background

The tried and tested method for studying the brain relies on the traditional experimental paradigm. A patient or subject is instructed to engage in a specific set of tasks. These tasks could be designed to involve the subject passively, such as showing the subject a series of images to study the visual perception [1]. On the other hand, an experimental design could involve active tasks such as engaging in a motor movement to understand the sensorimotor cortex [2, 3, 4, 5, 6]. An experimental study designed the neural correlates to language, and speech generation could involve uttering a particular phoneme, phrase, or sentence [7, 8, 9, 10, 11]. Neural correlates to visual stimuli have also been studied using the experimental trial structure [12]. The trial-based experimental paradigm has been the gold standard in advancing our understanding of the brain, and rightfully so. This understanding has also pushed the field towards developing neuro-technologies such as deep brain stimulators (DBS) that alleviate motor tremors plaguing Parkinson's Disease patients and brain-computer interfaces (BCIs) to help patients with Amyotrophic Lateral Sclerosis (ALS, or Locked-in Syndrome) to communicate with the outside world [13, 14].

Studying the brain in a structured trial-based setting simplifies our understanding of how neural activity is represented outside the lab. The assumptions made in an experimental paradigm ignores the complexity of natural internal and external states, their interactions, and their representation in the neural activity space. Additionally, brain activity is known

to follow time scales longer than a single stimulus or task [15], and that the brain state, whether it is due to externally visible factors (such as walking) or internally driven (such as hunger or thirst), can have profound effects on brain activity [16]. Spontaneous activity is also shown to have brain-wide representation, where multiple latent dimensions spanning different areas of mouse cortex are correlated to behavior [17]. In the experimental setting, neural activity is associated with externally measured variables, and this reliance causes experimenters to attribute some of the neural variability as noise [18]. Studies such as [19, 17, 20] show that this noise could be attributed to internal states and do correlate with behavior. There has been extensive research on brain activity during rest or non-engagement in tasks or stimuli. These studies have shown that the brain is highly active during resting state and have investigated the roles of different functional networks such as the Default Mode Network [21]. These studies' temporal scales still only span minutes to hours and are constrained to an experimental setting, where the task is to be at rest. To understand the brain in its naturalistic state and explore every axis of variability, which is the broader goal of neuroscientists and neuro engineers, we need to move outside the lab and into the wild.

1.2 Objective

With the dawn of the Information Age, there have been massive leaps in recording hardware quality and data storage capabilities, giving us the capability to collect a colossal amount of data. The time is ripe to move away from studying small sets of data spanning minutes to hours of brain activity and towards understanding the statistical and dynamical structure of brain activity across multiple days. In this thesis, we analyze days of brain activity collected in the epilepsy monitoring unit (EMU) from electrodes planted in and on the brain. Alongside brain activity, we have data modalities collected synchronously to investigate internal states, such as the autonomic tone using electrocardiograms (ECG), or external states such as behavioral states using video. We show that by leveraging days worth of brain activity, it is possible to gain novel insights about how the brain behaves in its natural environment without an experimenter applying additional constraints.

1.3 Challenges

The notion of an experiment to study the brain has been created to simplify the statistical analysis of the data and remove confounding noise that could jeopardize the investigation's success. Since only a handful of researchers have investigated naturalistic states in humans, the area of studying naturalistic states has not been investigated thoroughly. The tried and true methods that we wouuld apply on data collected in a structured trial based setting will fail miserably in this scenario. Therefore, we need to show that investigating naturalistic states is a feasible approach to investigating neural correlates to external and internal states. This must be done by applying controls and analyses that take into account the timescale of the data. Additionally, to the best of our abilities, we must separate the signal from the noise, and this challenge is even more profound when the data is not collected in a structured methodology.

1.4 Thesis Overview

We hope that this thesis serves as a starting point for many researchers to start investigating the statistical structure of brain activity outside the lab. We separate the body of this thesis into three main chapters that investigate the following:

In Chapter 2, we show that abstract behavioral states, or contexts, collected across days can be separated from the neural data alone. These behavioral contexts are coarsely labeled 5-minute segments with labels such as 'engaging in dialogue' and 'using electronics.' Even though these labels are coarse and vague compared to a traditional experimental paradigm, we show that we are able to have a statistically significant decoding performance using the high gamma band activity as a feature in a classifier. Additionally, we investigate a novel confound, such as the issue of long-range temporal correlations. We show that the time of day does have a significant effect on brain activity and behavioral state decodability. Chapter 2 is a reprint of the material as it appears in Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan et al. "Coarse behavioral context decoding." Journal of Neural Engineering 16, no. 1 (2019): 016021.

In Chapter 3, we leverage the previous chapter's data to ask an obvious next question, why are these abstract behavioral states separable? Why is it that labels generated in such an unstructured fashion can be used in a classifier and generate classification performance far exceeding chance? There could be many reasons for this phenomenon to occur, and in this chapter, we investigate why. We show that these abstract behavioral states can be separated using a myriad of different signal features such as long-term mean and variance shifts, differences in class covariances, and fast and slow changes in the high gamma band amplitude. Additionally, we use gaussian process factor analysis (GPFA) to show that there exist behavioral state-specific spatiotemporal patterns that separate these abstract behavioral states. Chapter 3 is a reprint of the material being prepared for submission as it appears in Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan et al. "Identifying neural signal characteristics that discriminate between naturalistic behavioral states."

In Chapter 4, we move beyond external states, such as behavior, to internal states, such as autonomic tone. We investigate whether neural recording spanning days can be used to gain novel insights about the neural correlates to autonomic tone during sleep. The autonomic tone is estimated from ECG recordings. In this work, we used sleep data to ensure that our analysis is not impacted by other states that are correlated with autonomic activity, such as movement. This chapter shows that high gamma band activity in different cortical and hippocampal regions is highly correlated with autonomic tone during sleep and that these correlations could be either positive or negative in the same region. Chapter 4 is a reprint of the material being prepared for submission as it appears in Alasfour, Abdulwahab, Xi Jiang, Jorge Gonzalez-Martinez, Eric Halgren, "High gamma activity in cortex and hippocampus is correlated with autonomic tone during sleep."

References

[1] Kai J. Miller, Gerwin Schalk, Dora Hermes, Jeffrey G. Ojemann, and Rajesh P. N. Rao. Spontaneous Decoding of the Timing and Content of Human Object Perception from Cortical Surface Recordings Reveals Complementary Information in the Event-Related Potential and Broadband Spectral Change. *PLOS Computational Biology*, 12(1):e1004660, January 2016. Publisher: Public Library of Science.

[2] Vikash Gilja, Paul Nuyujukian, Cindy A Chestek, John P Cunningham, Byron M Yu, Joline M Fan, Mark M Churchland, Matthew T Kaufman, Jonathan C Kao, Stephen I Ryu, and Krishna V Shenoy. A high-performance neural prosthesis enabled by control algorithm design. *Nature Neuroscience*, 15(12):1752–1757, December 2012.

[3] Mark M. Churchland, John P. Cunningham, Matthew T. Kaufman, Justin D. Foster, Paul Nuyujukian, Stephen I. Ryu, Krishna V. Shenoy, and Krishna V. Shenoy. Neural population dynamics during reaching. *Nature*, 487(7405):51–56, 2012.

[4] Kai J. Miller, Eric C. Leuthardt, Gerwin Schalk, Rajesh P.N. Rao, Nicholas R. Anderson, Daniel W. Moran, John W. Miller, and Jeffrey G. Ojemann. Spectral changes in cortical surface potentials during motor movement. *Journal of Neuroscience*, 27(9):2424–2432, 2007.

[5] Andrew J. Peters, Simon X. Chen, and Takaki Komiyama. Emergence of reproducible spatiotemporal activity during motor learning. *Nature*, 510(7504):263–267, 2014. Publisher: Nature Publishing Group.

[6] Kai J Miller, Gerwin Schalk, Eberhard E Fetz, Marcel den Nijs, Jeffrey G Ojemann, and Rajesh P N Rao. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proceedings of the National Academy of Sciences*, 107(9):4430 LP – 4435, March 2010.

[7] Gopala K. Anumanchipalli, Josh Chartier, and Edward F. Chang. Speech synthesis from neural decoding of spoken sentences. *Nature*, 568(7753):493–498, 2019.

[8] Miguel Angrick, Christian Herff, Emily Mugler, Matthew C. Tate, Marc W. Slutzky, Dean J. Krusienski, and Tanja Schultz. Speech synthesis from ECoG using densely connected 3D convolutional neural networks. *Journal of Neural Engineering*, 16(3), 2019. Publisher: IOP Publishing.

[9] Kristofer E. Bouchard, Nima Mesgarani, Keith Johnson, and Edward F. Chang. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441):327–332, 2013. Publisher: Nature Publishing Group.

[10] Nima Mesgarani and Edward F. Chang. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485(7397):233–236, May 2012.

[11] Brian N. Pasley, Stephen V. David, Nima Mesgarani, Adeen Flinker, Shihab A. Shamma, Nathan E. Crone, Robert T. Knight, and Edward F. Chang. Reconstructing speech from human auditory cortex. *PLoS Biology*, 10(1), 2012.

[12] Kai J Miller, Christopher J Honey, Dora Hermes, Rajesh PN Rao, Marcel denNijs, and

Jeffrey G Ojemann. Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *New Horizons for Neural Oscillations*, 85:711–720, January 2014.

[13] Leigh R. Hochberg, Daniel Bacher, Beata Jarosiewicz, Nicolas Y. Masse, John D. Simeral, Joern Vogel, Sami Haddadin, Jie Liu, Sydney S. Cash, Patrick van der Smagt, and John P. Donoghue. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375, May 2012.

[14] Chethan Pandarinath, Paul Nuyujukian, Christine H Blabe, Brittany L Sorice, Jad Saab, Francis R Willett, Leigh R Hochberg, Krishna V Shenoy, and Jaimie M Henderson. High performance communication by people with paralysis using an intracortical brain-computer interface. *eLife*, 6:e18554, February 2017. Publisher: eLife Sciences Publications, Ltd.

[15] Alexander Huk, Kathryn Bonnen, and Biyu J. He. Beyond trial-based paradigms: Continuous behavior, ongoing neural activity, and natural stimuli. *Journal of Neuroscience*, 38(35):7551–7558, August 2018. Publisher: Society for Neuroscience.

[16] David A. McCormick, Dennis B. Nestvogel, and Biyu J. He. Neuromodulation of Brain State and Behavior. *Annual Review of Neuroscience*, 43(1):391–415, 2020. _eprint: https://doi.org/10.1146/annurev-neuro-100219-105424.

[17] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D. Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893, April 2019.

[18] Alexander C. Huk and Eric Hart. Parsing signal and noise in the brain. *Science*, 364(6437):236, April 2019.

[19] Jan Gründemann, Yael Bitterman, Tingjia Lu, Sabine Krabbe, Benjamin F. Grewe, Mark J. Schnitzer, and Andreas Lüthi. Amygdala ensembles encode behavioral states. *Science*, 364(6437):eaav8736, April 2019.

[20] William E. Allen, Michael Z. Chen, Nandini Pichamoorthy, Rebecca H. Tien, Marius Pachitariu, Liqun Luo, and Karl Deisseroth. Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science*, 364(6437), April 2019. Publisher: American Association for the Advancement of Science Section: Research Article.

[21] William W. Seeley. The Salience Network: A Neural System for Perceiving and Responding to Homeostatic Demands. *The Journal of Neuroscience*, 39(50):9878–9882, December 2019. Publisher: NLM (Medline).

Chapter 2

Coarse Behavioral Context Decoding

Abstract

Objective. Current brain-computer interface (BCI) studies demonstrate the potential to decode neural signals obtained from structured and trial-based tasks to drive actuators with high performance within the context of these tasks. Ideally, to maximize utility, such systems will be applied to a wide range of behavioral settings or contexts. Thus, we explore the potential to augment such systems with the ability to decode abstract behavioral contextual states from neural activity. *Approach.* To demonstrate the feasibility of such context decoding, we used electrocorticography (ECoG) and stereo-electroencephalography (sEEG) data recorded from the cortical surface and deeper brain structures, respectively, continuously across multiple days from three subjects. During this time, the subjects were engaged in a range of naturalistic behaviors in a hospital environment. Behavioral contexts were labeled manually from video and audio recordings; four states were considered: *engaging in dialogue*, *rest*, *using electronics*, and *watching television*. We decode these behaviors using a factor analysis and support vector machine (SVM) approach. *Main Results.* We demonstrate that these general behaviors can be decoded with high accuracies of 73% for a 4-class classifier for one subject and 71% and 62% for a 3-class classifier for two subjects. *Significance.* To our knowledge, this is the first demonstration of the potential to disambiguate abstract naturalistic behavioral contexts from neural activity recorded throughout the day from implanted electrodes. This work motivates further study of context decoding for BCI

applications using continuously recorded naturalistic activity in the clinical setting.

2.1 Introduction

Brain-computer interface (BCI) research has experienced an impressive growth since the first demonstration of direct control of a robotic manipulator by employing electrical activity generated by ensembles of cortical neurons [1]. Now, BCIs designed for experimental and clinical studies can translate raw neuronal signals into control signals for various applications including, but not limited to, motor prostheses [2, 3, 4, 5, 6, 7], speech decoding [8, 9], and functional clinical mapping [10]. For many of these studies, performance is optimized by constraining the particular application to be highly structured, so as to reduce behavioral complexity and variability due to exogenous factors. These structured task-based studies, particularly those with human subjects, use limited datasets, typically on the order of minutes to tens of minutes. Limited dataset size often constrains the complexity and generalizability of BCI decoding strategies explored. In an effort to vastly augment data collection, preliminary work has been conducted to study neuronal signals in a more naturalistic, unconstrained manner by leveraging hours of recordings from monkey [11] and patients taken from the epilepsy monitoring unit (EMU) [12, 13].

In that regard, we labeled unstructured behavior from patients in the EMU. As a complement to trial-based experiments, these data are comprised of hand labeled recordings from multiple hours of video recorded in the EMU during which the subjects are engaged in naturalistic activities and are not instructed to engage in specific tasks. These labels are paired with simultaneously recorded electrocorticography (ECoG) or stereo-electroencephalography (sEEG) signals. ECoG captures electrical activity directly across a significant portion of the cortical surface, while sEEG utilizes depth electrodes that record from deeper brain structures as well as the cortical surface. We utilized these datasets to examine the potential to decode abstract behavioral contextual states from recorded neural activity. We classified different broad behavioral descriptions, which are *engaging in dialogue*, *watching TV*, *using electronics*, or *being at rest*, using a dataset of at least 9 hours per subject of neural data with at least 40

minutes of data for each behavioral label. For future BCI applications, it is critical to have reliable prosthetic performance across different behavioral contexts [14, 11]. We show in this paper that these complex and abstract behavioral contexts are separable. In practical terms, these separable contexts can be used to drive a context switch (Figure 2.1) in future, more generalizable and scalable BCIs. These future BCIs must work well across varied conditions automatically [14, 11]. A switch in context may change the dynamics of the observed neural signals and the intended range of control signals; ideally, a BCI will recognize and respond to those changes [11]. For example, a BCI that can detect behavioral context changes could engage context-specific and task-specific decoding algorithms, as shown in Figure 2.1. Such a BCI would, for example, be able to interpret current behavioral contexts to efficiently use resources, going into a power saving mode when the user is at rest, or to prevent generating unwanted actions, such as moving a limb when the subject intends to speak [14, 11].

Furthermore, a context decoder could be used in conjunction with a deep brain stimulator (DBS) in a closed-loop system for patients with movement disorders. Rather than continuously supplying the brain with current, which could induce unwanted side effects, such as speech and balance impairment [15, 16], and impact battery life [17], a closed loop DBS with a context switch implemented could be an adaptable system that would choose the most optimal stimulation parameters according to context. For example, after patient programming in the clinic, two sets of parameters could be determined, one that would best suppress hand and arm tremors, while another would suppress less of the tremor, but would cause less speech impairment. In a fully integrated DBS and context decoder system, the system would choose the latter set of stimulation parameters when the patient is determined to be engaging in dialogue. Such future potential BCI and DBS designs motivate investigating the separability of behavioral contexts within the recorded neural activity space. The purpose of this study is to provide a proof of concept that these naturalistic behavioral contexts can be decoded from recorded neural activity.

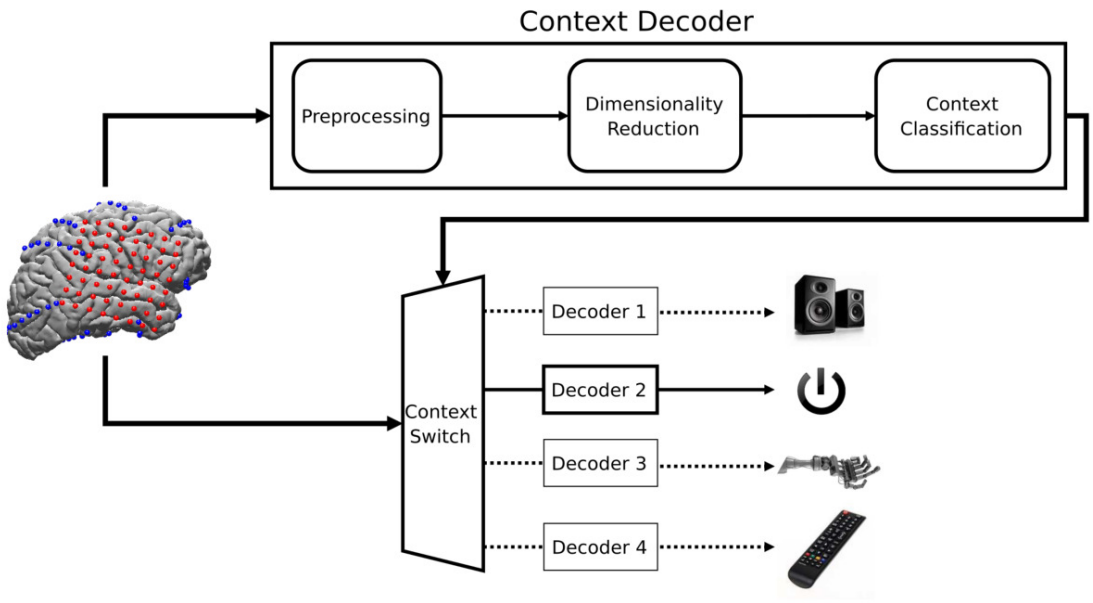


Figure 2.1: **Brain-computer interface (BCI) application for a context decoder.** Neural signals recorded from ECoG and sEEG are processed to drive both a context decoder and a context-dependent BCI application. Current BCIs extract specific features and apply unique decoding strategies on recorded neural activity to achieve high performance in specific tasks. To generalize the capabilities of dedicated BCI applications, we evaluate the possibility of a context decoding layer whose output would be used as a switch between different decoders. A BCI that is able to operate on natural behavior should be able to distinguish between different contexts and select the appropriate task-specific decoder.

2.2 Methods

2.2.1 Description of Data

Subjects

Three subjects with intractable epilepsy with hospital ID codes corresponding to NY394, RCH1, and RCH3 participated in this study. For simplicity, they will be respectively referred as Subjects 1, 2 and 3 for the remainder of this work. All subjects underwent invasive monitoring to localize epileptogenic zones before surgical resection. During their stays at New York University (NYU) Langone Comprehensive Epilepsy (Subject 1) and Rady Children's Hospital, San Diego (RCH) Pediatric Epilepsy centers (Subjects 2 and 3), the subjects gave their informed consent and were enrolled in the study. The studies were approved by the Internal Review Board of both institutions.

A Microsoft Kinect v2 was used to record both audio and video for each subject's

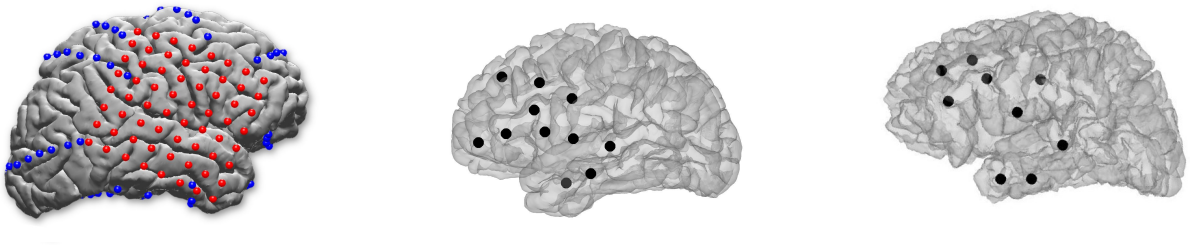
stay using multiple modalities such as RGB, IR, and depth [12]. However, only audio and RGB video was used for this study. The audio channel was used to synchronize the video and neural streams .

Recording

Subject 1 was implanted in the subdural space with platinum electrode arrays (Ad-Tech Medical Instrument Corporation, Oak Creek, WI) using a combination of linear arrays consisting of 4 to 10 electrode contracts, and a single 8x8 contact grid. Over 100 clinical subdural electrode contacts were embedded in SILASTIC sheets (2.3mm exposed diameter, 10 mm center-to-center spacing) [18] were implanted directly on the right cortical surface across multiple brain regions as shown in Figure 2.2a. The placement of electrodes was based on clinical considerations for the identification of seizure foci. EEG activity was recorded in the frequency range from 0.1 to 230 Hz using Nicolet clinical amplifiers, digitized and sampled at 512 Hz, and referenced to a 2 contact electrode array with electrodes screwed into the skull at a cm distance from the craniotomy edge under the scalp. For this study, we focused on recordings from the 8x8 grid given its coverage of frontal and temporal lobes (Figure 2.2a). These cortical regions of interest have been shown to be engaged in motor [19] and language [20] related behaviors, respectively. Subjects 2 and 3 were implanted with stereo-electroencephalography (sEEG) electrodes via the ROSA robotic surgical implantation (Medtech Surgical Inc., USA) in various orientations to allow for intracranial recording from lateral, intermediate, and/or deep cortical and subcortical structures in a 3-D arrangement [21]. sEEG electrode contacts had a 0.8 mm diameter, were 2mm long, and spaced 1-1.5 mm apart. Electrodes were sampled at 2000 Hz using Xltek 128Fs clinical amplifiers. For Subject 2, the sEEG implants were placed on the left frontal and temporal lobes (Figure 2.2b) while for Subject 3 the implants were placed on both the left and right frontal and temporal lobes. Only electrodes implanted in the left hemisphere for Subject 3 were used (Figure 2.2c).

Table 2.1: Description of data

Subject ID	Hospital ID	Age	Sex	Duration (hrs)	Implant	Location	Electrodes used
1	NYU394	30	Male	36	ECoG	Right Hemisphere	51
2	RCH1	18	Female	9	sEEG	Left Frontal and Temporal	71
3	RCH3	19	Male	10	sEEG	Bilateral Frontal and Temporal	68



(a) *Subject 1:* ECoG grid with broad right hemisphere coverage

(b) *Subject 2:* 11 sEEG shanks placed on the left frontal and temporal lobes

(c) *Subject 3:* 9 sEEG shanks placed on the left frontal and temporal lobes

Figure 2.2: **ECoG grid and sEEG shanks locations.**

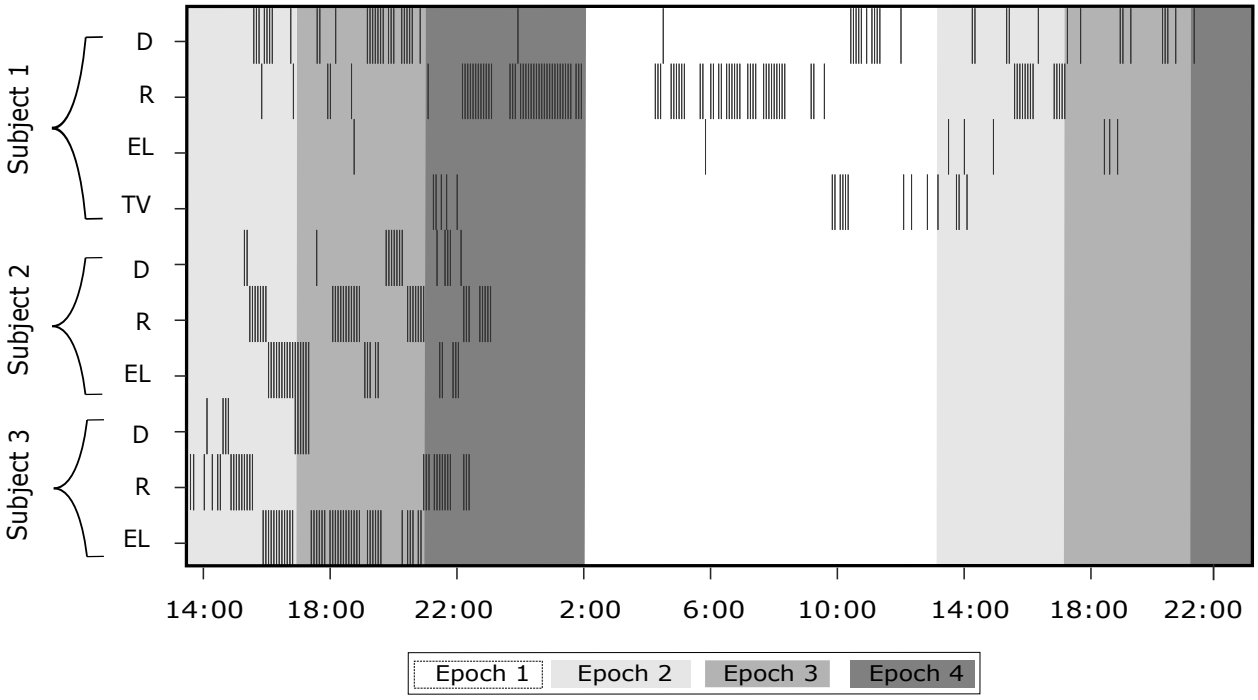


Figure 2.3: **Raster plot showing the locations of each context in time.** Each behavior label represents 5-minutes of manually annotated video. Epochs were segmented around periods where behavior labels were more prevalent throughout the day.

Table 2.2: Labeling criteria for behavior

Label	**Description** (subject is...)	**Mins S1**	**Mins S2**	**Mins S3**
(D) Dialogue	engaging in dialogue (including when the subject is talking on the phone)	275	75	50
(EL) Electronics	using an electronics device (phone, remote, tablet, video games, etc).	40	130	210
(TV) Television	watching television (does not include using remote or watching video on a tablet or phone)	90	N/A	N/A
(R) Rest	awake but resting, either with eyes open or closed, and either with or without visitors in the room	470	165	140

Table 2.3: Labeling criteria for time of day

Label	**Time**	**Mins S1**	**Mins S2**	**Mins S3**
(E1) Epoch 1	2:00-12:59	340	N/A	N/A
(E2) Epoch 2	13:00-16:59	150	125	185
(E3) Epoch 3	17:00-20:59	160	155	165
(E4) Epoch 4	21:00-1:59	225	90	50

Labeling

For each subject, we analyzed simultaneously recorded neural signals and clinical video/audio recording began 1-3 days (post-implant) of the subjects' stay in the hospital. Hours of continuous recordings were divided into 5-minute blocks and manually labeled according to the predominant behavior observed during each video segment. A behavioral rubric was designed for this study, which categorized waking periods according to whether or not the subject was active and, if so, who or what they were interacting with. For this analysis, a subset of behavior labels were selected to represent 1 inactive ('Rest') and 3 active ('Dialogue', 'Electronics', 'Television') behaviors. Electronics use included using the phone, remote, tablet or playing video games. Additionally, when a subject is talking to someone on the phone, that behavior is labeled as 'Dialogue', not 'Electronics'. Also, the 'Television' label does not include the times when the subject is using the remote. We used 3 active states and 1 'Rest' state in our context decoding scheme for Subject 1, but only 2 active states ('Dialogue', 'Electronics') and 1 rest state for Subject 2 and 3. This is due to the lack of the 'Television' context when labeling Subject 2 and 3's behaviors throughout the day. We also observed that different contexts tend to happen at specific times. For example, when considering time after 22:00, there is a very high chance that the subject is at 'Rest'. Therefore, we applied a second set of labels to relate behaviors to time of day. Note that the non-uniform segmentation of time is due to the nature of the dataset in which interesting contexts are more prevalent at different times throughout the day. Tables 2.2 and 2.3 show a summary of behavior and time labels. Subject 1's behavioral labels were verified by a second author and Subject 2 and 3's behavioral labels were not verified by a second author.

2.2.2 Neural Signal Conditioning

Prior to analysis, ECoG channels that were over seizure foci were down-selected manually by a clinician based upon visual inspection for Subject 1, while the sEEG channels for Subjects 2 and 3 did not undergo manual down-selection. Furthermore, we have rejected sEEG channels that were outside the cortical surface. Additional ECoG and sEEG channels were rejected if their mean of the squared signal value exceeded a threshold of 3 or 1.5 standard

deviations (for ECoG and sEEG respectively) above the mean squared signal value across all channels, indicative of potential noise corruption. sEEG channels were then down-sampled from 2000 Hz to 500 Hz for faster processing. The remaining channels were high-pass filtered with a cutoff frequency of 1 Hz to remove any underlying DC signal and then notch filtered with center frequencies at 60, 120, 180, and 240 Hz to remove line noise and its harmonics. Finally, electrode channels were re-referenced, either to a common average of all channels to remove shared noise (ECoG) or in a pair-wise manner with respect to neighboring electrodes (sEEG).

Spectral features from the 8–25 Hz and 70–110 Hz band of the ECoG signal were extracted by band-passing each channel. We have only utilized the high frequency band for the sEEG signal due the bipolar referencing scheme that was implemented which is likely to suppress spatially broad low frequency activity. Previous BCI studies motivate the selection of these frequency bands, as they demonstrate power modulation with respect to sensory-motor behavior [22, 23, 24]. The signal envelope was then calculated by taking the amplitude of the analytic signal of the Hilbert Transform. The resulting envelopes were then binned every 3 seconds and the average signal amplitude was calculated as an estimate of the square root of the power in each bin. Each channel was then z-scored across time.

2.2.3 Dimensionality Reduction: Factor Analysis

In the factor analysis model, we assume that there exists a multivariate latent random variable that has a Gaussian distribution with zero mean and identity covariance. The model assumes that each high-dimensional neural sample x is generated by sampling a lower dimensional multivariate Gaussian z. Then z is mapped to a higher k-dimensional space by computing $\mu + \Lambda z$, where μ is the observed mean. Lastly, x is generated by adding a strictly diagonal covariance noise Ψ to $\mu + \Lambda z$. We therefore obtain the following probability

distributions:

$$z \sim \mathcal{N}(0, I) \tag{2.1}$$

$$x|z \sim \mathcal{N}(\mu + \Lambda z, \Psi) \tag{2.2}$$

Factor analysis is a dimensionality reduction tool similar to principle component analysis (PCA). However, unlike PCA, factor analysis takes into account sensor noise (in this case ECoG or sEEG electrode noise) that is modeled by the diagonal matrix Ψ and allows the scale of sensor noise to vary across sensors. If sensor noise is non-uniform, the principle components found by PCA are biased towards the axes of the noisiest sensors. Thus, when sensor noise can vary, like in ECoG or sEEG, factor analysis can model this non-uniformity and avoid this biasing issue to more effectively find common structure across sensors. We used the Laurens van der Maaten dimensionality MATLAB toolbox for factor analysis [25] which implements the expectation-maximization based algorithm described in [26] to find Λ and Ψ. The joint distribution of x and z is as follows:

$$\begin{bmatrix} z \\ x \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ \mu \end{bmatrix}, \begin{bmatrix} I & \Lambda^T \\ \Lambda & \Lambda\Lambda^T + \Psi \end{bmatrix} \right) \tag{2.3}$$

We projected the electrode space into a factor space of 30 dimensions by calculating the conditional expectation $E(Z|X)$. Since both x and z are random variables obtained from a Gaussian distribution, the conditional distribution is also Gaussian and can be calculated analytically based upon the joint distribution above, thus $E(Z|X)$ evaluates to $\Lambda^T(\Lambda\Lambda^T + \Psi)^{-1}x$.

2.2.4 Classification

These projected factors form the input features for both a 4-class support vector machine (SVM) classifier with labels 'Dialogue', 'Electronics', 'TV', and 'Rest' for Subject 1. A 3-class support vector machine with labels 'Dialogue', 'Rest', and 'Electronics,' was used for Subjects 2 and 3. We also use a 4-class SVM in classifying time of day labels, 'Epoch

1,' 'Epoch 2,' 'Epoch 3,' and 'Epoch 4' for Subject 1 and a 3-class SVM using time of day labels 'Epoch 2,' 'Epoch 3,' and 'Epoch 4' for Subject 2 and 3. The multi-class SVM uses a one-vs-one coding design and a linear kernel. SVM classification was implemented using the MATLAB machine learning toolbox. We generated confusion matrices for the 4-class and 3-class classifiers depending on the number of contexts available for each subject. We applied a 7-fold cross validation on the data, where the test and training data were each taken from a continuous time segment rather than randomized points across the entire dataset. This was done to minimize the chance that classification was achieved due to the temporal locality of the testing and training data points. The number of samples for each class in the training set is randomly downsampled to the number of samples in the class with the least number of samples to ensure that the SVM does not skew in favor of predicting the majority class. The test set class samples were also balanced to simplify comparison to chance performance. For example, looking at Subject 1's data in Table 2.3, the label with the least amount of data is 'EL' with 40 minutes. Therefore, each fold for electronics would have around 5.7 minutes worth of neural data, which means that we have around 34.2 minutes per class for training and 5.7 minutes per class for testing in each fold.

2.2.5 Electrode Analysis

To motivate future BCI design in terms of implant location on the brain, we ranked the electrodes for each subject by sorting the single channel context decoding performance for each electrode using the classification scheme described in Section 2.2.4. The ranked channels for each subject were then used to investigate the effect of an incremental addition of the best performing electrodes. The accuracy vs. the number of top channels used were plotted for each subject. Finally, we investigated if neighboring electrodes carry similar information that discriminate between different contexts. This was done by segmenting the electrodes into groups belonging to either 2x2 clusters (for Subject 1 who is implanted with an 8x8 ECoG array, giving us 16 different clusters), or shanks (for Subjects 2 and 3 who were implanted with sEEG shanks). We then found the decoding performance for each cluster/shank and compared them with each single electrode that belongs to that cluster or shank. We plotted

the single channel performance versus the cluster/shank performance for each subject. If neighboring electrodes carry the same information regarding context separability, then we should see single channels perform as well as the clusters or shanks that they belong in.

2.2.6 Modeling Neural Activity Dependence on Time

Having observed that different behavioral contexts tend to happen at specific times of day, a possible confound to the context decoder described is that it is not truly decoding context from neural activity, but instead the decoder could rely upon neural activity variation that is primarily coupled by time of day. Thus, we investigated the influence of *temporal distance* and *time of day*. Figure 2.4 is a simple illustration of three possible graphical models that could describe the relationship between time of day, behavioral context, and neural activity generation and their affects on context classification performance. Model 1 suggests that the ability to decode natural behaviors is completely dependent upon the time of day confound rather than behavioral contexts, as it suggests that neural activity is generated based upon time of day and not context. Model 2 and 3 both suggest that context can be disambiguated from neural activity, without a reliance on time of day based modulation. In these two models neural activity is directly context dependent, with time of day having a modulatory effect on neural activity in Model 3. These models do not make any claims on how the brain actually works, but rather highlight the mechanisms that are driving our context classifier's performance. Two control analyses were conducted to investigate the potential time of day confound with respect to these three models.

Temporal Distance vs. Performance

Temporal distance is defined by the time separation between the training and testing data. A main concern when analyzing this type of data is the effect of temporal distance on the context decoder performance. We investigated the possibility that decoder performance would decrease as temporal distance increases. If this phenomenon proves to be true, it would highly suggest that Model 1 in Figure 2.4 is the underlying model explaining the separability of neural data in the context decoder. It would also suggest that time dependent neural

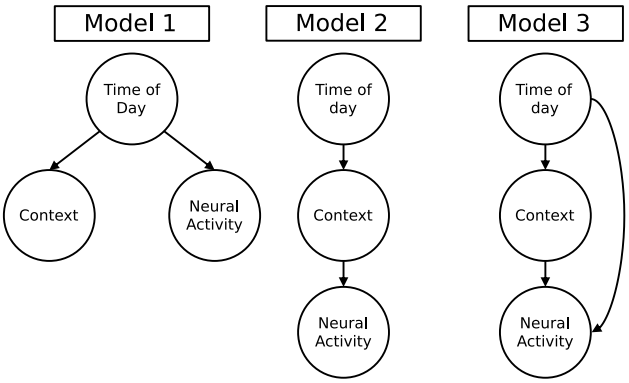


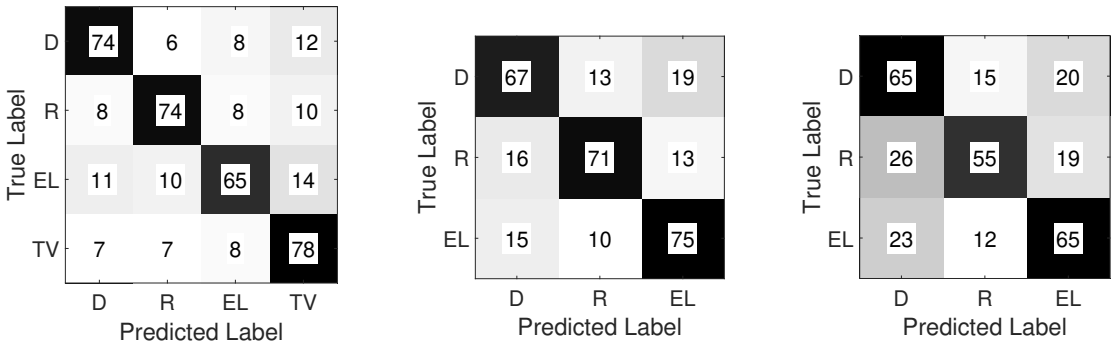
Figure 2.4: **Possible models of time dependencies for behavior context and neural activity that explain context decoding performance.** Model 1 assumes that neural activity and behavioral contexts are independent of each other, but both are dependent on time. Model 2 assumes that the neural activity is dependent only on context, while the context is dependent on the time of day in which it occur. Model 3 assumes that neural activity is dependent on both context and time of day. We did not illustrate these models to provide a generalized explanation of the mechanisms underlying neural activity, but rather to elaborate possible reasons for classification performances when applying a context decoder on a set of neural data.

activity is being decoded, rather than context dependent neural activity. We ran two analyses to explore this idea. Firstly, we organized our dataset so that each context was sorted by time stamp. We then trained a model using a temporally adjacent 20 minute segment per class and tested on 5 minute segments per class that varied in temporal distance from the training set. For example, the classifier could be trained from a time segment from 15:00 to 15:20 and tested using a time segment from 16:00, 18:00 and 20:00. We tested the hypothesis that temporal locality is being leveraged by evaluating whether a statistically significant downward trend was found between the temporal distance between the training and test sets and performance. This process was repeated 30 times for different temporal distances. Additionally, to prevent over-fitting due to the small amount of training data compared to context decoding using the entire dataset, we determined the top performing electrodes by sorting the single channel performance using the method described in Section 2.2.4. We ran the temporal analysis using the top 5 electrodes. Linear regression was used to fit a line on the data for each iteration of the temporal analysis to quantify the existence of any statistically significant trend between accuracy and average temporal distance. For this specific analysis, we z-score the test data using the test data's mean and standard deviation

to prevent trending effects influencing our results. If the p-value is greater than 0.05, we can accept the null hypothesis that the slope of the line being fitted is zero and reject the hypothesis that there is a linear trend relating temporal distance and classifier performance. Finally, we investigated the effects of dropping the 'Rest' state from this analysis to see if it alleviated any negative trending effects correlating performance and temporal distance. This is due to the fact that the 'Rest' state is not as well defined as other contexts. For example, a subject being inactive alone or with the presence of people would both be labeled as 'Rest'.

Time of Day and Context Maximum Likelihood Estimation

We further explored the dependence of neural activity on time of day by exploring three separate classification schemes. We implemented a similar classification approach as mentioned in Section 2.2.4: 3 second time windows and 30 latent factors were used with a linear, multi-class SVM classifier, where behavior labels were substituted by the time of day (Table 2.3). This was compared to a time of day maximum likelihood estimator from the context labels. For this estimator, only the context and time of day labels for the 5-minute blocks are used as the features. This was done by calculating the probability of a specific time of day given a context label and choosing the time of day with the highest probability given a specific context. The performance of the maximum likelihood estimator serves as a conservative baseline to determine the amount of information shared between the time of day and context labels. Additionally, the maximum likelihood estimator serves as a biased estimator (favoring the maximum likelihood estimator) since the same dataset is used for training and testing. We compared the performances of the SVM time of day classifier to the time of day maximum likelihood estimator described above to investigate to what extent the time of day affects the context decoder performance. Finally, we constructed another maximum likelihood estimator that predicts context given the time of day using the context and time of day labels for each 5-minute block as features. This estimator is compared to the context decoder described in the Section 2.2.4. For both of the maximum likelihood estimators, we randomly sub-sampled each class label to have the same number of labels for each class for an easier comparison to chance level. The maximum likelihood time of day



(a) *Subject 1* (b) *Subject 2* (c) *Subject 3*

Figure 2.5: **Confusion matrices for multi-class context decoding using 7-fold cross validation**. Each number in the confusion matrices corresponds to percentages. Rows of the confusion matrices are normalized. Rounding errors will cause some of the rows' sum to deviate slightly from 100%. Both training and test sets are balanced by randomly subsampling the same number of data points in each test fold for each class. For Subject 1, Figure 2.5a, 'Dialogue', 'Electronics', 'TV', and 'Rest' behaviors were classified by a linear, multi-class SVM classifier, using spectral features from the 8-25 Hz and 70-110 Hz bands of ECoG signals collected simultaneously across the brain in 3-second time bins. For Subjects 2 and 3 in Figure 2.5b and Figure 2.5c, only the 70-110 Hz band of the sEEG signal was used. For all three Subjects, behavioral labeling was done in 5 minute blocks. For Subject 3, 'Dialogue', 'Electronics', and 'TV' show a high rate of confusion with 'Rest.' The mean and standard deviation of context decoding accuracy across folds for all subjects can be seen in Table 2.4.

and context estimators are governed by the following equations (D, C, and T denote decoder, context and time respectively) :

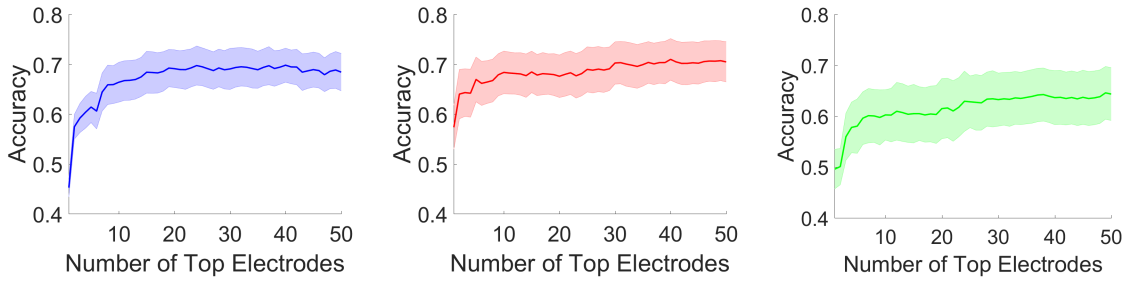
$$D_{context}(j) = \underset{i}{\text{argmax}} P(C_i|T_j) \tag{2.4}$$

$$D_{time}(i) = \underset{j}{\text{argmax}} P(T_j|C_i) \tag{2.5}$$

2.3 Results

2.3.1 Performance of Classifier

The mean of the confusion matrices for all subjects across folds are shown in Figure 2.5. The linear SVM-based classifier was able to decode different neural states with accuracies significantly above chance performance for all subjects. The classification accuracies when all contexts were used (4-class for Subject 1 and 3-class for Subjects 2 and 3) were 73%, 71%



(a) *Subject 1* (b) *Subject 2* (c) *Subject 3*

Figure 2.6: **Accuracy vs. number of top electrodes.** Each bold line is the mean across 7 folds, and the shaded region is one standard error from the mean across folds. For all 3 subjects, only a small number of top electrodes are needed to achieve adequate context classification performance. The chance level is 25% for Subject 1 and 33% for Subjects 2 and 3.

and 62% for Subjects 1, 2 and 3 respectively. The standard deviation across folds are 9.5%, 10.9%, and 13.5% for Subjects 1, 2 and 3 respectively. We also observe that for Subject 3 in Figure 2.5c, the active states are highly confused with the 'Rest' state.

2.3.2 Electrode Analysis

Looking at Figure 2.6, we observe that the performance of the classifier saturates after a small number of top electrodes are used. Figures 2.7, 2.8 and 2.9 indicate that the top 10 electrodes that contribute most in context decoding occupy multiple regions of the cortical surface (for all subjects) as well as deeper brain structures (for Subjects 2 and 3). Additionally, Figure 2.10 compares the performance when a single electrode is used vs when the entire cluster or shank that electrode is in is used. We observed that for most cases, the 2x2 clusters in the ECoG grid and the sEEG shanks generally outperformed single channels.

2.3.3 Temporal Distance vs. Performance

In Figure 2.11, we observe that there is a negative trend relating performance and temporal distance for Subjects 1 and 2, with only Subject 1 having a statistically significant negative trend ($p<0.05$). For Subject 3, however, we observe that the fitted slope is almost zero, and becomes slightly positive when we drop the 'Rest' state. Temporal distance has a lesser affect when only the active states were decoded ('Rest' is dropped) for Subject 1

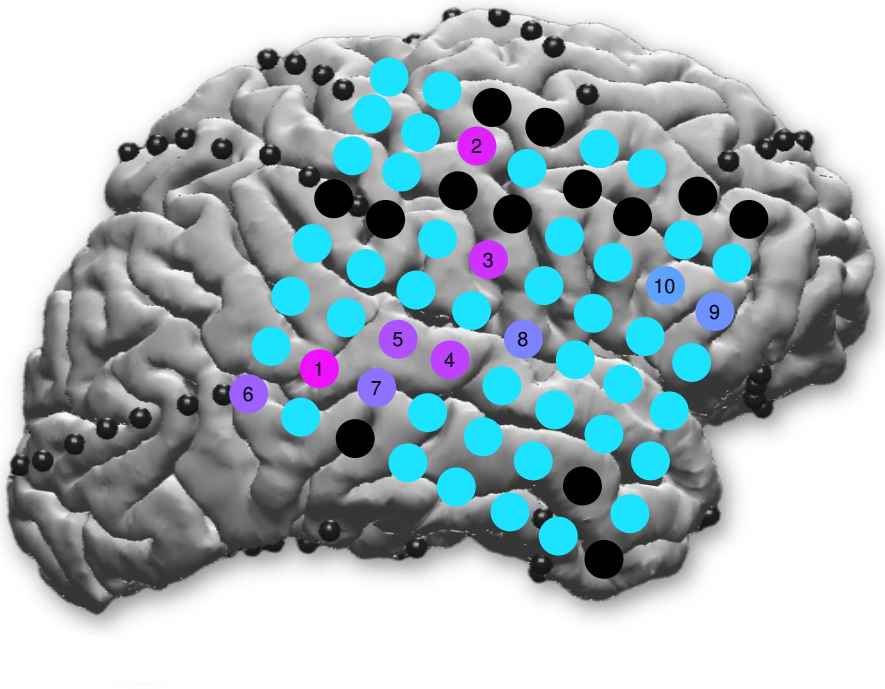
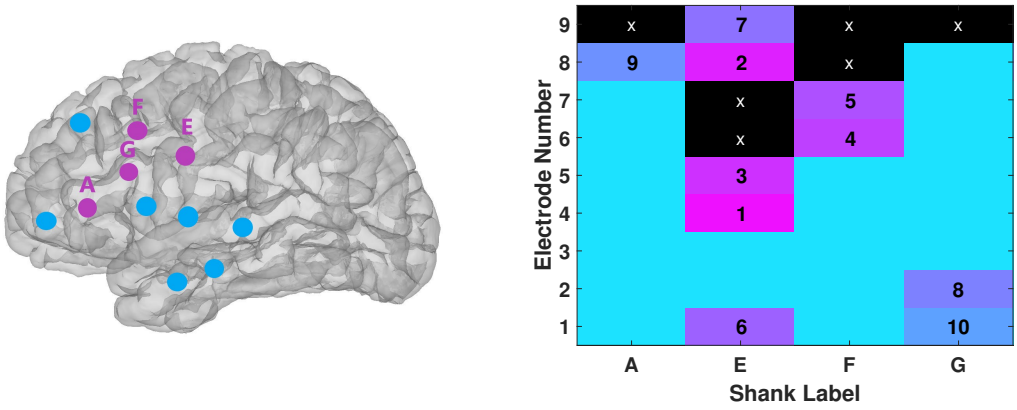


Figure 2.7: **Subject 1: Locations of the top 10 performing electrodes are numbered.** Colors closer to bright pink symbolizes the highest performing electrodes while colors closer to light blue symbolizes the lowest performing electrodes. The electrodes labeled as 'rejected' are colored black.

and 2. The fitted slope for Subject 1 when 'Rest' is dropped has a high enough p-value, 0.093, that indicates that the fitted slope is not statistically significant. For all Subjects 1 and 2, the fitted slope is closer to zero and the p-value is greater for when the active contexts are classified compared to when all contexts are classified. We also notice that for certain pairings of training and test sets, the classifier performance is lower than chance.

2.3.4 Time of Day and Context Maximum Likelihood Estimation

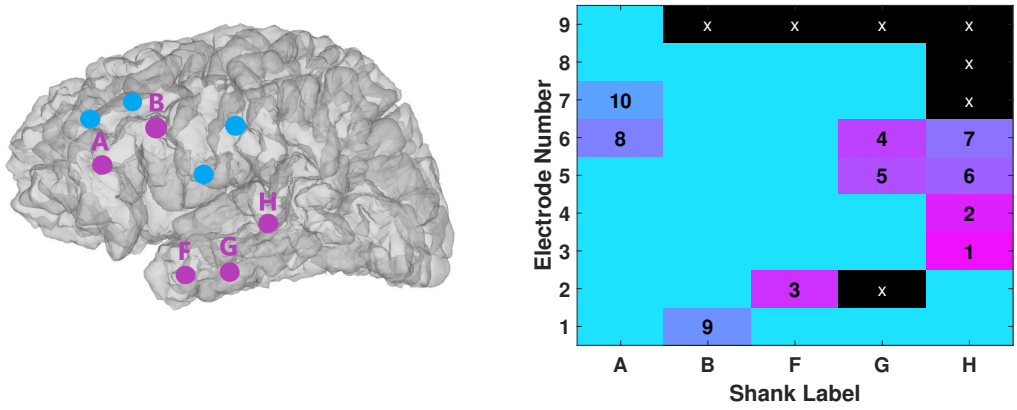
Figure 2.12 shows the confusion matrices of each of the classifiers described in Section 2.2.6 for all subjects. Table 2.4 summarizes the classification accuracies of each type of decoder. The estimator accuracies for the maximum likelihood context estimator given time of day are 38%, 49%, and 51% for Subjects 1, 2 and 3 respectively. In comparison to the



(a) (b)

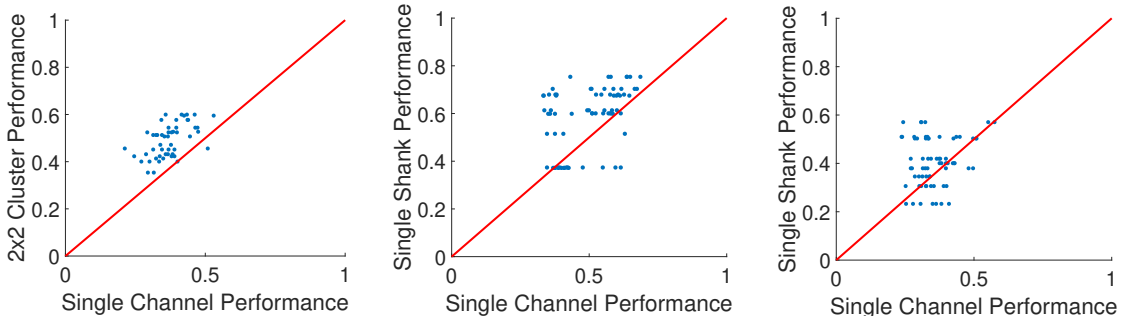
Figure 2.8: **Subject 2: Figure 2.8a shows the locations of the sEEG electrode shanks that include the top 10 performing electrodes.** These shanks are labeled and colored in pink. Figure 2.8b displays the locations of the top 10 performing electrodes across all shanks. Colors closer to bright pink symbolizes the highest performing electrodes while colors closer to light blue symbolizes the lowest performing electrodes. The electrodes labeled as 'rejected' are colored black. Electrode number 9 corresponds to the electrode that is closest to the cortical surface. The electrode contacts are spaced 1-1.5mm apart, are 2 mm long, and have a 0.8 mm diameter.

context decoder from neural activity accuracies of 73%, 71%, 62% for Subjects 1, 2 and 3 respectively; the context decoder from neural activity outperforms the maximum likelihood context estimator, this suggests that context is separable from neural activity. For the time of day decoders, we observe that the classification accuracies of the maximum likelihood time of day estimator given context are 47%, 49%, 53% and the mean classification accuracies of the time of day given neural activity decoder are 57%, 68%, and 74% with standard deviations across folds of 10.5%, 14.5%, and 7.4% for Subjects 1, 2 and 3 respectively. Subject 1 exhibits the best performance that supports the claim that context-relevant information is being decoded, as the context decoder from neural data outperforms the maximum likelihood context estimator and that the time of day given neural activity decoder does not out perform the maximum likelihood time of day estimator. For Subject 1, the time of day from neural activity decoder performs worse in comparison with other subjects. We theorize that because Subject 1 has the most data, where the distribution of contexts across time is more uniform, as seen in Figure 2.3, the performance of the time of day classifier given neural activity is significantly less than the other two subjects. In Figure 2.3, for Subjects 2 and 3, the



(a) (b)

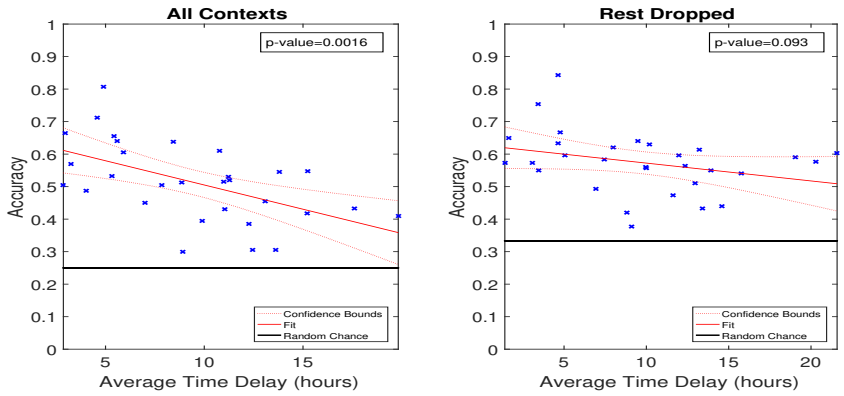
Figure 2.9: **Subject 3: Figure 2.9a shows the locations of the sEEG electrode shanks that include the top 10 performing electrodes across all shanks.** These shanks are labeled and colored in pink. Figure 2.9b displays the locations of the top 10 performing electrodes. Colors closer to bright pink symbolizes the highest performing electrodes while colors closer to light blue symbolizes the lowest performing electrodes. The electrodes labeled as 'rejected' are colored black. Electrode number 9 corresponds to the electrode that is closest to the cortical surface. The electrode contacts are spaced 1-1.5mm apart, are 2 mm long, and have a 0.8 mm diameter.



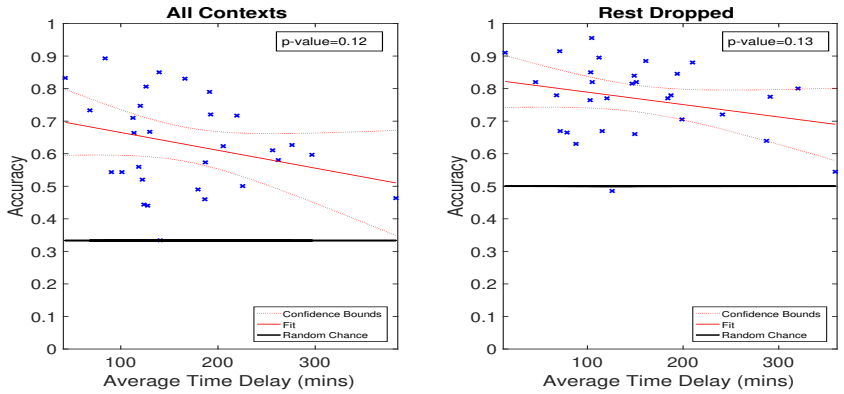
(a) *Subject 1* (b) *Subject 2* (c) *Subject 3*

Figure 2.10: **Shank/cluster vs. single channel performance.** For all subjects, the 2x2 ECoG clusters and sEEG shanks out perform single channel performance. This indicates the presence of additive information in neighboring channels, which motivates the use of focal electrode coverage in future BCI design.

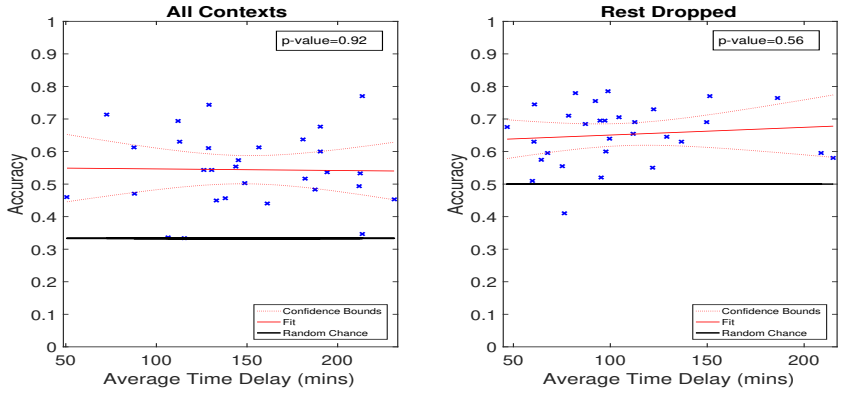
distribution of contexts in time is far from uniform, as we notice a significant clustering of the contexts in specific times of day. For example, we observe that for Subject 3, the 'Epoch 2' time label is also exclusively the 'Electronics' context label.



(a) *Subject 1*



(b) *Subject 2*



(c) *Subject 3*

Figure 2.11: **Classifier accuracy vs. average temporal distance between training and test sets.** For each subject, we show the possibility of context decoding even when the temporal distance between the training and test set is large. We notice mainly that for Subject 1 in Figure 2.11a, the fitted negative slope statistical significance is greatly alleviated when the 'Rest' state is dropped ($p>0.05$).

Table 2.4: **Performance of different decoders**: Mean accuracies and standard deviations are shown for each decoder for all subjects. Note that for the maximum likelihood estimators, the features were a context or time of day label of a 5-minute segment. Therefore, due to a much reduced sample size as compared to using 3 second bins, k-fold cross validation is not used, and no standard deviations are shown.

Subject	Chance	Context from Neural	MLE-Context	Time from Neural	MLE-Time
1	25%	73% ± 9.5%	38%	57% ± 10.5%	47%
2	33%	71% ± 10.9%	49%	68% ± 14.5%	49%
3	33%	62% ± 13.5%	51%	74% ± 7.4%	53%

2.4 Discussion

2.4.1 Context Decoding

In this study we demonstrate the potential to decode abstract behavioral contexts from neural activity by achieving classification accuracies significantly above chance for all three subjects as shown in Table 2.4. To our knowledge, this work is the first step in human behavior context decoding from neural signals collected from naturalistic behavior. We discuss in Section 2.1 the practical applications of behavioral context decoding for augmenting BCI and DBS designs. However, the behavioral contexts that we have analyzed in this work is by no means the optimal contexts to implement in a context decoder for a BCI or DBS design scheme. We would like this work to serve as a proof of concept that neural data collected from a naturally behaving subject could facilitate the augmentation of current BCI and DBS designs.

2.4.2 Electrode Analysis

The results shown in Section 2.3.2 supports our intuition that to decode complex activities such as *'engaging in dialogue'*, *'using electronics'*, or *'watching TV'*, coverage from multiple regions of the brain is ideal since across three subjects, the top performing electrodes were present in all major regions of the brain in which electrodes were included in the analyses. However, in order to form an adequate statistical analysis of the relationship between different brain regions and contexts, more subjects are needed with broad electrode coverage. We also found that with only a small set of electrodes, we are able to decode context from both the cortical surface and deeper brain regions. In addition to that, we have also found that

clusters or shanks generally outperform single electrodes. This indicates that in the space of context decoding, neighboring electrodes in a confined space generate distinctive information that aid in classifying different contexts. Both of these findings suggest that in a practical BCI design, it is beneficial to have broad coverage of different regions of the brain. However, if one has only access to a small portion of the cortex, then a dense electrode coverage would enhance performance.

2.4.3 Temporal Analysis

In Figure 2.11, we show that when 'Rest' is removed, there is less of a negative correlation between temporal distance and performance. This suggests that the 'Rest' state is highly variable and more sensitive to the time of day than other states. Therefore, it is paramount to investigate if there exists sub-states within the 'Rest' state to explore why 'Rest' is highly variable. For all three subjects, we observe many instances where the performance of the classifier is very close to chance or even below chance. This could be due to the fact that we are using a very small dataset for training (20 mins) that is localized in time, which could cause the classifier to overfit for that specific time. However, for many instances, even with limited training data size per iteration, context decoding is possible even when the temporal distance is large. Also, the precise temporal location of the training and test set might impact this measure.

Our findings in Section 2.2.6 indicate that there is discriminative context-dependent information in the neural signal, which supports Models 2 and 3 rather than Model 1, since the context from neural activity decoder outperforms the maximum likelihood context from time of day estimator for all subjects. However, we also observe that the time of day from neural activity decoder outperforms the maximum likelihood time of day from context estimator for all three subjects as well. Along with the fact that there is a slight negative trend relating decoder accuracy and temporal distance, we conjecture that discriminative information in the neural data is dependent on both time of day and context, which supports Model 3 as the true underlying model that is driving the context decoder's performance. Therefore, additional days of recording are needed to more fully examine the impact of time of day

on context decoding. If the performance of the time of day classifier using neural activity was considerably higher than the maximum likelihood time of day from context estimator, then it suggests that the neural activity provides information about time of day beyond the information due to the dependence of context on time of day.

In aggregate, the results of these control analyses support Model 3 over Model 1 and Model 2, as described in Figure 2.4. Due to the fact that, first of all, the context decoder from neural activity outperformed the maximum likelihood context estimator from time, and the time of day from neural activity decoder also outperformed the maximum likelihood time of day estimator from context labels. These two observations support the fact that there is both context and time of day information in the neural activity. To fully explore candidate dependencies between neural activity, time of day, and behavioral context, a richer dataset would need to be constructed. Ideally, such a dataset could be sub-sampled to yield datasets in which there is a uniform distribution of different behavioral contexts throughout all time of day labels.

2.5 Conclusion

We have demonstrated the ability to decode coarse and broad behavioral contexts from ECoG and sEEG based neural signals. With a relatively simple machine learning pipeline, we were able to decode behavioral contexts for three subjects with an accuracy significantly above chance. More in-depth analysis and modeling of neural activity is essential to fully understand why these contexts are separable, and to advance our understanding of the underlying physiology. Such analyses could motivate specific electrode coverage paradigms for future BCIs. Furthermore, analysis of context decoders could be used by researchers and clinicians to augment functional mapping of the cortical surface for both clinical and neural engineering applications with minimal interference to the patient's daily routine. The results obtained from our simple behavioral labeling and decoding scheme motivate further exploration into the methods employed for context decoding. The application of more advanced machine learning methods, such as those that can better leverage temporal and

spatial relationships in recorded neural activity, could improve classification performance. This work also motivates the development of richer and more comprehensive datasets of neural activity and naturalistic behavior. Behavioral labeling is a key challenge towards developing such a dataset, as manual annotation is time consuming and subjective. Therefore, there is a need to create an objective and quantitative way to label patient behavior, possibly through video and audio analysis of recorded behavior. Advanced computer vision and speech detection algorithms could be used to generate more descriptive and precise labels, on a timescale of seconds rather than minutes. The development of such datasets would allow for the exploration of richer definitions of contextual state collected across many subjects. In summary, as a proof of concept this work motivates further exploration of behavioral context decoding methods and their potential to advance BCI systems development and clinical neuroscience.

Author Contributions

BCI research is highly collaborative by design, requiring expertise from the fields of clinical neurology, neuroscience, and engineering. To enable this work, the authors contributed in the following ways: Implantation of electrodes was performed via surgery by Drs. Devinsky (NYU) and Gonda (RCH). Clinical monitoring was provided by Drs. Dugan (NYU), Friedman (NYU), Doyle (NYU), Sattar (RCH) and Wang (RCH). Data acquisition was coordinated by Melloni (NYU), Thesen (NYU), Gabriel (NYU,RCH), and mentored by Gilja (NYU, RCH). Behavior labels were annotated by Jiang, Shamie, Alasfour, and mentored by Gilja and Halgren. Analysis and writing were led by Alasfour, assisted by Gabriel, and mentored by Gilja and Halgren.

2.6 Acknowledgements

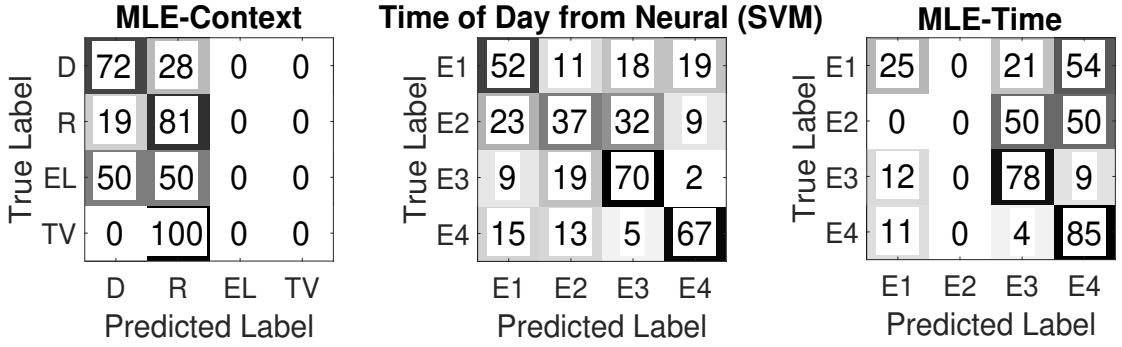
This work has been supported, in part, by the UCSD ECE Department Medical Devices & Systems Initiative, the UCSD Centers for Human Brain Activity Mapping (CHBAM) and Brain Activity Mapping (CBAM), the UCSD Frontiers of Innovation Scholars Program, the

Qualcomm Institute Calit2 Strategic Research Opportunities (CSRO) program, the Hellman Fellowship, the Institute of Engineering in Medicine Graduate Student Fellowship, and the Clinical and Translational Research Institute at UC San Diego. We would like to thank the patients and clinicians who contributed to this study at UC San Diego, Rady Children's Hospital of San Diego, and the Comprehensive Epilepsy Center at NYU Langone Medical Center. We specifically thank Preet Minas, and Hugh Wang for their contributions to data collection and annotation. This chapter is a reprint of the material as it appears in Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan, Daniel Friedman, Werner Doyle, Orrin Devinsky, David Gonda, Shifteh Sattar, Sonya Wang, Eric Halgren, and Vikash Gilja. "Coarse behavioral context decoding." Journal of neural engineering 16, no. 1 (2019): 016021.

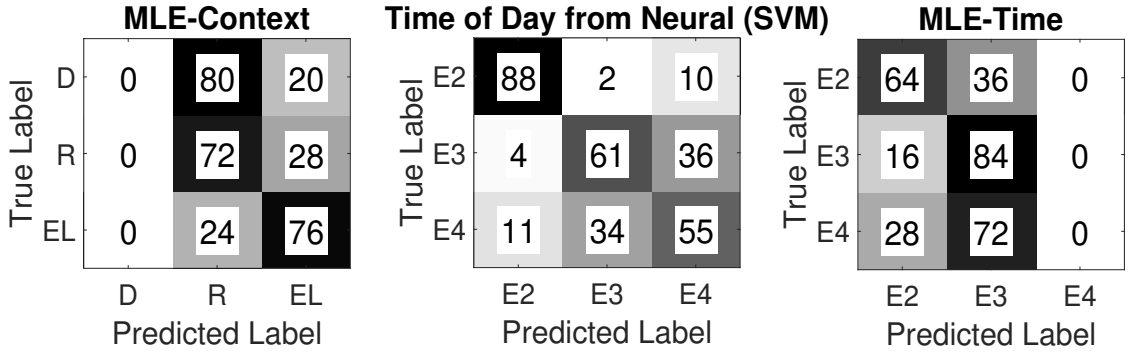
2.7 Supplementary Data

Table 2.5: This table shows the amount of data in the training and testing folds for the SVM decoders when applying 7-fold cross-validation (duration in minutes) after randomly down-sampling to have an equal representation of each class in training and testing. See Table 1 for the total amount of data for each class.

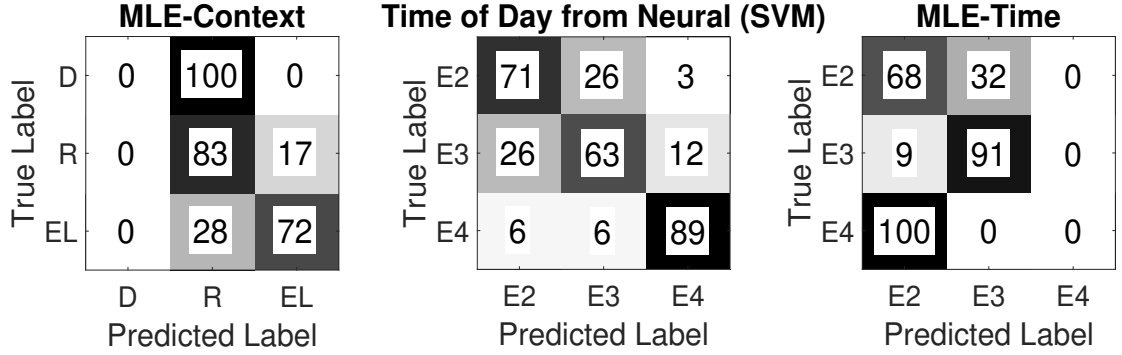
Subject ID	**Train/Context**	**Test/Context**	**Train/Epoch**	**Test/Epoch**
1	34.3	5.7	128.6	21.4
2	64.3	10.7	77.1	12.9
3	42.9	7.1	42.9	7.1



(a) *Subject 1*



(b) *Subject 2*



(c) *Subject 3*

Figure 2.12: **Confusion matrices for decoding context from time of day, time of day from neural data, and time of day and from context for all three subjects.** Rows of the confusion matrices are normalized. Rounding errors will cause some of the rows' sum to deviate slightly from %100. For all three subjects, (performance summarized in Table 2.4), the context from neural activity out performs the maximum likelihood context from time of day estimator, and the time of day from neural activity decoder also out performs the maximum likelihood time of day from context estimator. However we note that for Subject 1, where context labels are most uniformly spread through time, the time of day from neural activity decoder out performs the maximum likelihood time of day from context estimators the least out of all three subjects. We also note that for maximum likelihood estimators, some labels are completely predicted incorrectly.

References

[1] Fetz E. Operant conditioning of cortical unit activity. *Science*, **163**(3870):955–958, 1969.

[2] Wang W, Collinger J, Degenhart A, Tyler-Kabara E, Schwartz A, Moran D, Weber D, Wodlinger B, Vinjamuri R, Ashmore R, Kelly J, and Boninger M. An Electrocorticographic Brain Interface in an Individual with Tetraplegia. *PLoS ONE*, **8**, 2013.

[3] Carmena J, Lebedev M, Crist R, O'Doherty J, Santucci D, Dimitrov D, Patil P, Henriquez C, and Nicolelis M. Learning to control a brain–machine interface for reaching and grasping by primates. *PLoS Biology*, **1**(2), 10 2003.

[4] Chaplin J, Moxon K, Markowitz R, and Nicolelis M. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature*, **2**:664–670, 1999.

[5] Gilja V, Nuyujukian P, Chestek C, Cunningham J, Krishna S, et al. A high-performance neural prosthesis enabled by control algorithm design. *Nature Neuroscience*, **15**:1752–1758, 2012.

[6] Hochberg L, Serruya M, Friehs G, Mukand J, Saleh M, Caplan A, Branner A, Chen D, Penn R, and Donoghue J. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, **442**:164–171, 2006.

[7] Velliste M, Perel S, Spalding C, Whitford A, and Schwartz A. Cortical control of a prosthetic arm for self-feeding. *Nature*, **458**:1098–1101, 2008.

[8] Bouchard K, Mesgarani N, Johnson K, and Chang E. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, **495**(7441):327–332, March 2013.

[9] Mesgarani N and Chang E. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, **485**(7397):233–6, May 2012.

[10] Korostenskaja M, Wilson A, Rose D, Brunner P, Schalk G, Leach J, Mangano F, Fujiwara H, Rozhkov L, Harris E, et al. Real-time functional mapping with electrocorticography in pediatric epilepsy: comparison with fmri and esm findings. *Clinical EEG and Neuroscience*, **45**(3):205–211, 2014.

[11] Santatham G, Linderman M, Gilja V, Afshar A, Ryu S, Meng T, and Shenoy K. Hermesb: A continuous neural recording system for freely behaving primates. *IEEE Transactions on Biomedical Engineering*, **54**(11):2037–2050, 2007.

[12] Gabriel P, Doyle W, Devinsky O, Friedman D, Thesen T, and Gilja V. Neural correlates to automatic behavior estimations from rgb-d video in epilepsy unit. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 3402–3405. IEEE, 2016.

[13] Wang N, Olson J, Ojemann J, Rao R, and Brunton B. Unsupervised decoding of long-term, naturalistic human neural recordings with automated video and audio annotations. *Frontiers in human neuroscience*, **10**, 2016.

[14] Linderman M, Santhanam G, Kemere C, Gilja V, Shenoy K O'Driscoll S, Meng T, et al.

Signal processing challenges for neural prostheses: a review of state-of-the-art systems. *IEEE signal processing magazine*, pages 18–28, 2008.

[15] Baizabal-Carvallo J, Kagnoff M, Jimenez-Shahed J, Fekete R, and Jankovic J. The safety and efficacy of thalamic deep brain stimulation in essential tremor: 10 years and beyond. *Journal of Neurology, Neurosurgery & Psychiatry*, 2013.

[16] Guehl D, Edwards R, Cuny E, Burbaud P, Rougier A, Modolo J, and Beuter A. Statistical determination of the optimal subthalamic nucleus stimulation site in patients with parkinson disease. *Journal of Neurosurgery JNS*, **106**(1):101 – 110, 2007.

[17] Houston B, Thompson M, Ojemann J, Ko A, and Chizeck H. Classifier-based closed-loop deep brain stimulation for essential tremor. In *8th International IEEE EMBS Conference on Neural Engineering*, pages 316–221. IEEE, 2017.

[18] Quinn B, Carlson C, Doyle W, Cash S, Devinsky O, Spence C, Halgren E, and Thesen T. Intracranial cortical responses during visual-tactile integration in humans. *The Journal of Neuroscience : the official journal of the Society for Neuroscience*, **34**(1):171–181, January 2014.

[19] Miller K, Schalk G, Fetz E, Den Nijs M, Ojemann J, and Rao R. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proceedings of the National Academy of Sciences*, **107**(9):4430–4435, 2010.

[20] Pasley B, David S, Mesgarani N, Flinker A, Shamma S, Crone N, Knight R, and Chang E. Reconstructing speech from human auditory cortex. *PLoS Biology*, **10**(1):e1001251, 2012.

[21] Gonzalez-Martinez J, Bulacio J, Thompson S, et al. Robotic epilepsy surgery: technique, results and complications related to robotic assisted seeg. *Neurosurgery*, **78**(2):169–179, 2016.

[22] Arroyo S, Lesser R, Gordon B, Uematsu S, Jackson D, and Webber R. Functional significance of the mu rhythm of human cortex: an electrophysiologic study with subdural electrodes. *Electroencephalography and clinical neurophysiology*, **87**(3):76–87, 1993.

[23] Crone N, Miglioretti D, Gordon B, and Lesser R. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. ii. event-related synchronization in the gamma band. *Brain*, **121**(12):2301–2315, 1998.

[24] Aoki F, Fetz E, Shupe L, Lettich E, and Ojemann G. Increased gamma-range activity in human sensorimotor cortex during performance of visuomotor tasks. *Clinical Neurophysiology*, **110**(3):524–537, 1999.

[25] Van Der Maaten L, Postma E, and Van den Herik J. Dimensionality reduction: a comparative review. *J Mach Learn Res*, **10**:66–71, 2009.

[26] Ghahramani Z, Hinton G, et al. The em algorithm for mixtures of factor analyzers. Technical report, Technical Report CRG-TR-96-1, University of Toronto, 1996.

Chapter 3

Identifying Neural Signal Characteristics that Discriminate Between Naturalistic Behavioral States

Abstract

The experimental paradigm in neuroscience relies on trials within a limited time window. While this approach has been the gold standard, it fails to investigate the complex and rich neural activity that occurs in a natural, unconstrained environment. Improved recording, computing, and data storage capabilities in human neuroscience allow us to amass large-scale datasets spanning days rather than minutes. These developments also grant us the ability to investigate the neural correlates of unstructured and naturalistic behavioral states and deduce the statistical structure and sources of discriminability between different states. In this work, we analyze neural activity from two human subjects recorded from electrocorticography (ECoG) and stereo-electroencephalography (sEEG) and show that multiple neural signal characteristics exist that discriminate between unstructured and naturalistic behavioral states such as *'engaging in dialogue'* and *'using electronics'*. Using the high gamma amplitude estimate, we demonstrate that behavioral states in a naturalistic setting are discriminable based on long-term mean shifts, variance shifts, and differences in the covariance structure of the neural activity. Both the rapid and slow changes in high

gamma band activity separate unstructured behavioral states. Additionally, we use gaussian process factor analysis (GPFA) to show the existence of salient spatiotemporal structures with variable smoothness in time. We demonstrate that both the temporally smooth and stochastic spatiotemporal activity can be used to separate unstructured behavioral states. To our knowledge, this is the first attempt to elucidate how different neural signal features contain information about behavioral states collected outside the conventional experimental paradigm.

Author summary

Neuroscience research generally relies on the experimental trial-based paradigm to uncover how the brain works. While this methodology has been both tractable and fruitful for decades, there is a need to move towards studies that leverage unstructured and naturalistic behaviors to uncover the statistical and dynamic structure of neural activity when not constrained to a controlled experimental setting. In this work, we show how different signal features of neural activity recorded from intracranial electrodes can differentiate naturalistic behavioral states. These signal features include both static and dynamic aspects of the spatiotemporal neural activity.

3.1 Introduction

The overwhelming majority of experiments in the neurosciences involve a trial-based experimental protocol. A patient or subject is instructed to complete a specific set of actions, engage in cognitive tasks, or passively react to external stimuli. This approach has yielded an appreciable understanding on how the brain controls different external behaviors such as motor activity [1, 2] and language [3, 4, 5, 6], or internal states such as thirst, learning and sensorimotor strategies [7, 8, 9, 10]. The trialized experimental approach to neuroscience serves as a proxy for understanding the brain in its natural state due to its tractable nature. However, these neural correlates may not be directly equivalent to spontaneously occurring behaviors or states that are not manifested from a structured, trial-based experimental

paradigm. Continuous and naturalistic sensory-cognitive-motor loops are not only tractable but should be thought of as the main framework in which behavior, perception, and cognition should be studied [11]. With the advent of increased computational power, storage, and the ability to collect large scale multimodal data that span days rather than minutes, the time is ripe to move towards studying the brain without external constraints.

The brain has varying temporal dynamics that could span tens of seconds to hours [12], far exceeding the timespans investigated in a typical experimental setting. It has been shown that band-limited power of the local field potential display fluctuations at many timescales, with the highest power present in the very low frequencies (<0.1 Hz) [13]. Neural activity across hemispheres [14], and is directly related to information accumulation [15]. Additionally, differences in temporal timescales of neural activity are directly related to cortical hierarchy, with lower and higher-order areas having faster and slower timescales, respectively [16]. These slower timescales could also be related to brain state, which is controlled by slow-activating neuromodulators and strongly influences behavior [17]. In mice, spontaneous behaviors exhibit multidimensional, brainwide activity [18]. Therefore, collecting long-term unstructured neural activity could help us understand both the spatial and temporal time scales that are relevant to describe the variability present in naturalistic behavior.

Previous works have demonstrated that the neural correlates to naturalistic behaviors such as numerical processing [19] and arm movements [20, 21, 22] display behaviorally correlated neural activity. The ability to decode unstructured and naturalistic internal states such as mood has also been previously established [23]. Additionally, in a previous study, we decoded abstract unstructured behavioral states from neural activity [24]. In this study, we leverage hand-labeled data collected from human subjects in the epilepsy monitoring unit (EMU) spanning multiple days to fully understand the spatial distribution and temporal timescales relevant to describing abstract behavioral states. Behavioral state labels such as "engaging in dialogue", and "using electronics," are paired with simultaneously recorded electrocorticography (ECoG) or stereo-electroencephalography (sEEG) signals. ECoG captures electrical activity directly across a significant portion of the cortical surface, while sEEG utilizes depth electrodes that record from deeper brain structures as well as

the cortical surface. While our previous work [24] demonstrated the potential to identify behavioral states using conventional neural signal features used in trial-based studies, here we show that multiple neural signal features describe the neural variability present in unstructured behavior. Therefore, customary signal processing pipelines driven from trial based studies might not be ideal for capturing the full scope of neural variability outside the experimental protocol.

To illustrate how different sources of information could influence the separability of neural activity features, Fig 3.1A illustrates a toy example of two epochs spanning 60 seconds of two behavioral states. In this illustration, we display only two simulated electrodes for simplicity. Both states are separable by multiple signal components such as channel mean, variance, covariance, and temporal dynamics. In this example, State 1 has a lower channel mean but higher variance than State 2. The covariance of both electrodes is positive in State 1 but negative in State 2. Finally, both electrodes covary slowly together in State 1, but much faster in State 2. Fig 3.1B displays the electrode values as a scatter plot to illustrate the differences in mean, variance, and covariance between the two states. It is imperative to investigate all of these components in the neural signal to fully understand the sources of information that discriminate between behavioral or internal states. We demonstrated that all of these components differentiate abstract and unstructured behavioral context states.

3.2 Materials and methods

3.2.1 Description of Data

Subjects

Three subjects with intractable epilepsy with hospital ID codes corresponding to NY394, RCH1, and RCH3 participated in this study. For simplicity, they will be respectively referred as Subjects 1–3 for the remainder of this work. All subjects underwent invasive monitoring to localize epileptogenic zones before surgical resection. During their stays at New York University (NYU) Langone Comprehensive Epilepsy (Subject 1) and Rady Children's Hospital, San Diego (RCH) Pediatric Epilepsy centers (Subjects 2 and 3), the subjects gave

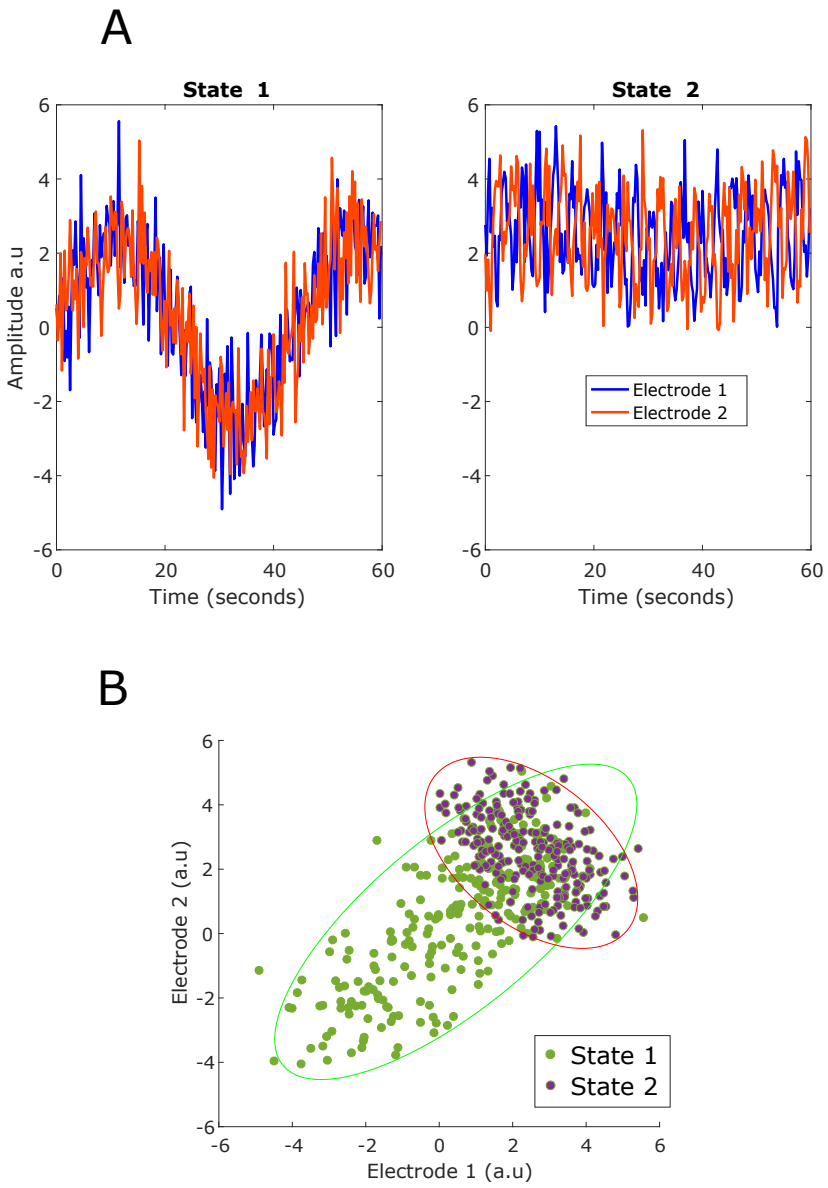


Figure 3.1: **Simulated electrodes visualized for two different behavioral states.** (A) Two simulated electrodes are shown to visualize different multi-variate signal components that could differentiate between two behavioral states. There is a difference between the electrodes' mean and variance when comparing states 1 and 2. Two spatial-temporal features are present at different rates in both states. There is a slow component where the electrodes are positively correlated and a fast component where the electrodes are negatively correlated. In state 1, the slow component is more prominent, while in state 2 the fast component is. (B) The values of the electrodes across both states are shown in a scatter plot. This method of visualizing electrodes could tell us a lot about the differences between the two states spatially but lacks any information regarding the temporal dynamics.

their informed consent and were enrolled in the study. The studies were approved by the Internal Review Board of both institutions. A Microsoft Kinect v2 was used to record both

audio and video for each subject's stay using multiple modalities such as RGB, IR, and depth [25]. However, only audio and RGB video was used for this study. The audio channel was used to synchronize the video and neural streams.

Recording

Subject 1 was implanted in the subdural space with platinum electrode arrays (Ad-Tech Medical Instrument Corporation, Oak Creek, WI) using a combination of linear arrays consisting of 4–10 electrode contracts and a single 8 x 8 contact grid. Over 100 clinical subdural electrode contacts were embedded in SILASTIC sheets (2.3 mm exposed diameter, 10 mm center-to-center spacing) [26] were implanted directly on the right cortical surface across multiple brain regions as shown in [24]. The placement of electrodes was based on clinical considerations for the identification of seizure foci. EEG activity was recorded in the frequency range from 0.1 to 230 Hz using Nicolet clinical amplifiers, digitized and sampled at 512 Hz, and referenced to a two contact electrode array with electrodes screwed into the skull at a cm distance from the craniotomy edge under the scalp. For this study, we focused on recordings from the 8x8 grid, given its frontal and temporal lobes coverage. These cortical regions of interest have been shown to be engaged in motor [27] and language [28] related behaviors, respectively. Subjects 2 and 3 were implanted with stereo-electroencephalography (sEEG) electrodes via the ROSA robotic surgical implantation (Medtech Surgical Inc., USA) in various orientations to allow for intracranial recording from lateral, intermediate, and/or deep cortical and subcortical structures in a 3D arrangement [29]. sEEG electrode contacts had a 0.8 mm diameter, were 2 mm long, and spaced 1–1.5 mm apart. Electrodes were sampled at 2000 Hz using Xltek 128Fs clinical amplifiers. For Subject 2, the sEEG implants were placed on the left frontal and temporal lobes, while for Subject 3, the implants were placed on both the left and right frontal and temporal lobes.

3.2.2 Neural Signal Conditioning

Prior to analysis, ECoG channels that were over seizure foci were down-selected manually by a clinician based upon visual inspection for Subject 1, while the sEEG channels

for Subjects 2 and 3 did not undergo manual down-selection. sEEG channels were then referenced in a pair-wise manner with respect to neighboring electrodes. Furthermore, we have rejected sEEG channels that were outside the cortical surface. Additional ECoG and sEEG channels were rejected if their mean of the squared signal value exceeded a threshold of 3 or 1.5 standard deviations (for ECoG and sEEG respectively) above the mean squared signal value across all channels, indicative of potential noise corruption. Channel values that were equal to zero were also rejected. sEEG channels were then down-sampled from 2000 Hz to 500 Hz for faster processing (after anti-aliasing filtering). The remaining channels were then notch filtered with center frequencies at 60, 120, 180, and 240 Hz to remove line noise and its harmonics. Finally, ECoG channels were re-referenced to a common average of all channels to remove shared noise. Spectral features from the 70–110 Hz band of the ECoG signal were extracted by band-passing each channel. Previous studies motivate the selection of these frequency bands, as the high frequency activity in local field potential is heavily correlated to local neuronal spiking activity [30, 2, 31, 32]. Additionally, the HFB band-limited power demonstrate power modulation with respect to sensory-motor behavior [33, 34].

The signal envelope was then calculated by taking the amplitude of the analytic signal of the Hilbert Transform. The resulting envelopes were then binned every 250 ms, and the average signal amplitude was calculated as an estimate of the square root of the power in each bin. For each continuous segment of neural activity and each channel, we remove timepoints where the value is seven times above the channel median and replace that value using linear interpolation. We visually inspect all channels after the above preprocessing steps and reject channels that have excess noise or variance. Finally, we investigated whether the power spectral densities of the band-limited power has a 1/f fall off, as has been shown in [13, 12, 14]. The PSD's of the band-limited amplitude estimates for Subject 3's recorded exhibited heightened activity in the high frequency component of the binned band-limited power. As this is atypical relative to the previous literature, this heightened activity may represent either noise artifacts or pathological activity. Thus, we have removed Subject 3 from subsequent analyses.

3.2.3 Labeling of Behavioral States

For each subject, we analyzed simultaneously recorded neural signals and clinical video/audio recording began 1-3 days (post-implant) of the subjects' stay in the hospital. Hours of continuous recordings were divided into 5-minute blocks and manually labeled according to the predominant behavior observed during each video segment. A behavioral rubric was designed for this study, which categorized waking periods according to whether or not the subject was active and, if so, who or what they were interacting with. For this analysis, a subset of behavior labels were selected to represent one inactive ('Rest') and three active ('Dialogue', 'Electronics', 'Television') behaviors. Electronics use included using the phone, remote, tablet or playing video games. Additionally, when a subject is talking to someone on the phone, that behavior is labeled as 'Dialogue', not 'Electronics'. Also, the 'Television' label does not include the times when the subject is using the remote. We used three active states and one 'Rest' state in our context decoding scheme for Subject 1, but only two active states ('Dialogue', 'Electronics') and one rest state for Subject 2. This is due to the lack of the 'Television' context when labeling Subject 2's behaviors throughout the day. For more details on labeling see our previous work [24].

3.2.4 Classification of Slowly Varying Mean and Variance

Firstly, we examined if the abstract behavioral contexts could be classified based upon long term changes in either mean or variance. The mean and variance of the high gamma band amplitude estimate across 30-second bins for each channel were calculated to form the features in the decoder. A four-class SVM with labels 'Dialogue', 'Rest,' 'Electronics' and 'TV', and a three-class SVM with labels 'Dialogue', 'Rest' and 'Electronics' were used for Subjects 1 and 2, respectively. The multi-class SVM uses a one-versus-one coding design and a linear kernel. SVM classification was implemented using the MATLAB machine learning toolbox. We applied a 7-fold cross-validation on the data, where the test and training folds were each taken from continuous time segments rather than randomized points across the entire dataset. The training and testing sets are class-balanced to ensure not learning a model biased to the most representative class and easier comparison to chance. We additionally

applied a 1-fold buffer between the training and test set. This was done to reduce the potential impact of temporally local correlations between the testing and training data. Testing and training data are at least 5-minutes apart.

3.2.5 Classification Due to Changes in Power at Different Time Scales

As an extension to the analysis described in the previous section, we investigated features at two time scales, a slow timescale in which the temporal dynamics are slower than 3 seconds and a fast timescale in which the temporal dynamics are faster than 3 seconds. Mainly, we filter the 250 ms binned high gamma band amplitude estimate using both a low pass and a high pass filter with a cut off frequency of 0.333 Hz (3 second period). We then generate the signal envelope of the filtered signal and take the average across 30-second bins. Prior to filtering, we z-score each 5-minute segment to ensure that discriminability between behavioral states is due to varying temporal dynamics that are not caused by ultraslow mean and variance shifts. Cross-validation and an SVM classification paradigm were then applied to the low pass and high pass filtered features using the same scheme mentioned in the previous section. Additionally, we calculated the power spectral density of each class for each channel using Welch's method. A Hanning window was used along with a 256 point FFT on each 30-second segment with 75% overlap. The power spectral densities of exemplar channels for both subjects are shown to highlight the differences in the temporal dynamics for different behavioral contexts.

3.2.6 Classification Due to Changes in Covariance Structure

Other than shifts in signal amplitude or power in different frequency bands, the correlation or connectivity between different channels could be class-specific. To test this, we first z-score the data across 5-minute segments to remove any class-specific differences in long-term channel variances. Then, for each 30-second segment, we calculate the covariance matrix. We then apply 7-fold cross-validation with a 1-fold buffer and use a minimum distance to mean classification algorithm to classify each class's covariance matrices, as shown in [35]. To simplify, we chose the Euclidean mean and distance as our metrics in the decoder. Other

mean and distance metrics did not show a significant improvement in classification accuracy. Training and test sets are class balanced, similarly to the previous sections.

3.2.7 Gaussian Process Factor Analysis

Motivation and Description

In the previous analyses, we investigate differences in the spatial and temporal dynamics for each behavioral context. However, these methods rely on investigating the differences in the spatial and temporal domains separately. Therefore, it is imperative to apply a data-driven method to extract spatiotemporal patterns that are relevant for each behavioral state. In the spatial domain, frequently used dimensionality reduction techniques such as Principle Component Analysis (PCA) or Factor Analysis (FA) look for subspaces where different channels coactivate together. PCA looks for the subspaces that exhibit the highest variance, while FA considers private channel noise and more effectively finds common structure across sensors. However, each point in time is assumed to be independent for both methods, and no temporal structure is determined from the data. In this section, we motivate the use of Gaussian Process Factor Analysis (GPFA) as described in [36], to find spatial factors with varying autocorrelation functions. Using GPFA, one can determine each spatial component's smoothness in a data-driven way without making many assumptions about the temporal dynamics.

In GPFA, the neural activity is assumed to be generated from set of latent factors described in the following equation,

$$y_{:,t}|x_{:,t} \quad \sim \mathcal{N}(Cx_{:,t} + d, R) \tag{3.1}$$

Where $y_{:,t} \in \mathbb{R}^{q\times 1}$ is the neural activity of q channels at time point t, and $x_{:,t} \in \mathbb{R}^{p\times 1}$ is the latent neural state with dimensionality p at time point t (p<q). Additionally, $C \in \mathbb{R}^{q\times p}$ is the factor loading matrix that maps the latent neural states into the observed neural activity. $d \in \mathbb{R}^{q\times 1}$ is the channel mean and $R \in \mathbb{R}^{q\times q}$ is the channel private variance.

In a standard factor analysis model, each latent dimension $x_{i,:}$ $(i = 1, \ldots, p)$ is assumed

be generated from an iid Gaussian distribution. Therefore, there is no correlation between each consecutive time point. The GPFA augments the standard FA model by adding the assumption that the underlying latent neural activity can be correlated across different time points. GPFA defines each latent dimension across time to be generated from a gaussian process (GP):

$$x_{i,:} \sim \mathcal{N}(0, K_i) \tag{3.2}$$

Where $x_{i,:} \in \mathbb{R}^{1\times T}$ is the activity of latent state i across time. In this GP, the covariance matrix across time has the following structure

$$K_i(t_1, t_2) \approx e^{\frac{-(t_1-t_2)^2}{2\tau_i^2}} \tag{3.3}$$

The full equation approximated above is shown in [36]. The correlation between time points t_1 and t_2 are related by a decaying exponential function. The parameter tau controls how fast the exponential decays. As tau increases, adjacent time points further away are correlated together, and the latent state is smoother in time. In other words, the decaying exponential controls the autocorrelation of each latent neural state.

GPFA is fit on the data using the EM algorithm [36]. Once all the parameters, $\theta = \{C, d, R, \tau_1, \ldots, \tau_p\}$, are fit to the data, the p latent neural states are evaluated by determining the posterior expectation $E[x|y]$ [36].

Reduced GPFA

The GPFA EM algorithm looks for latent neural states with varying temporal autocorrelation functions that best fit the data. However, multiple factors could be similar in the spatial domain but vary in their temporal dynamics. Additionally, the percentage of the variance explained by each latent neural state of the data isn't explicitly determined by the algorithm. To generate latent factors with distinct spatial structures and to understand their importance describing the data, we apply an orthonormalization procedure on the latent neural states described in [36]. The factor loading matrix C is decomposed using the singular

value decomposition such that

$$C = UDV' \tag{3.4}$$

where $U \in \mathbb{R}^{q \times p}$, $V \in \mathbb{R}^{p \times p}$, and $R \in \mathbb{R}^{p \times p}$. U and V have orthonormal columns, and D is a diagonal matrix. Since the columns of C describe the vector space in which the latent neural states move in, we can rewrite the projection of the latent neural states to the channel space using an orthonormal vector space by the following operation:

$$Cx_{:,t} = U(DV^{'}x_{:,t}) = U(\tilde{x_{:,t}}) \tag{3.5}$$

A new set of orthonormal neural trajectories $\tilde{x_{:,t}}$ is generated since the space that $\tilde{x_{:,t}}$ spans is determined by the columns of U. Additionally, the columns of U are ordered by the amount of covariance explained in the data. Therefore, one can rank the latent neural states by their importance in explaining the actual neural activity. Another advantage of orthonormalization is that the orthonormal factor loading matrix U in equation 3.5 is insensitive to rotations applied to the latent dimensions $\tilde{x_{:,t}}$, in contrast to factor matrix C shown in equation 3.1, which is unique up to an arbitrary rotation matrix.

Prediction Error

To determine whether GPFA provides insightful information regarding the temporal dynamics of different latent neural states, we compare the goodness-of-fit of GPFA in comparison with FA. If GPFA provides a better fit than FA, then the addition of the GP temporal smoothness constraint results in factors that better explain the recorded neural activity. To test this, we calculate the leave-one-out root mean squared error, similar to the method described in [36], using a different number of latent states. Additionally, we implement 7-fold cross-validation with a 1-fold buffer between the training and test sets. In this analysis, the training set is balanced to ensure that the model is not biased to the class

with the most samples. However, we use all of the test set in each fold to reduce the variability due to the random sampling of the test set. We calculated RMSE of each behavioral context separately and summed them together to generate a class-balanced RMSE. We also calculate the leave-one-out class-balanced RMSE of reduced GPFA to determine whether a smaller set of orthonormal latent factors is a better fit for the data.

Classification of Latent Neural States Determined from GPFA

The GPFA EM algorithm described in the previous section is an unsupervised method to extract spatiotemporal patterns in the data. In this work, we aim to find spatiotemporal patterns that are salient to specific unstructured behavioral states. Thus we developed a GPFA based classification method. To achieve this, we first segment the neural activity into 30-second windows of 250 ms binned high gamma amplitude estimates that are z-scored for each 5-minute segment. Z-scoring is applied to remove class-specific long term mean and variance changes that would bias the EM algorithm. The dimensionality of each epoch is q x T, where q is the number of channels and T is the number of timesteps. The epochs are ordered chronologically in time for each behavioral state and segmented into K-folds. In the training phase, we extract a balanced number of samples for each behavioral state in the training folds to not bias the EM algorithm to look for spatiotemporal patterns specific to a certain class. We add a 1-fold buffer between the training and testing folds to minimize the effects of long term correlations. The EM algorithm is fitted on the data to extract the GPFA parameters $\theta = \{C, d, R, \tau_1, \ldots, \tau_p\}$. Once the parameters are determined, the latent neural states in both the training and test sets are determined using the posterior expectation. Each epoch of neural state activity has dimensionality p x T, where p is the number of latent states or factors.

The epochs are then vectorized to a dimensionality of (p x T) x 1. A quadratic discriminant analysis (QDA) classifier is trained on the data and tested on a balanced testing set to facilitate a straightforward comparison to chance prediction. Since the dimensionality of vectorized data (p x T) is much higher than the number of samples, a parametrization of the fitted class mean and covariance is necessary. This is achieved by calculating the following

class mean and covariance.

$$\mu_C = \begin{bmatrix} \mu_{1c} \\ \vdots \\ \mu_{pc} \\ \vdots \\ \mu_{1c} \\ \vdots \\ \mu_{pc} \end{bmatrix} \tag{3.6}$$

$$\Sigma_C = \begin{bmatrix} \Sigma_c & \dots & \dots & 0 \\ \vdots & \Sigma_c & 0 & \vdots \\ \vdots & 0 & \Sigma_c & \vdots \\ 0 & \dots & \dots & \Sigma_c \end{bmatrix} \tag{3.7}$$

Here, μ_{ic} is the mean of the latent neural state i for class c. μ_C is the class mean of the vectorized epochs. The mean is assumed to be static across time. Σ_c is the covariance of the latent neural states for class c determined from concatenating all the epochs together, forming a matrix of dimensionality p x (T*E) where E is the number of epochs for class C. The covariance Σ_C of the vectorized epochs for class C is constructed such that the covariance of the latent neural states at each time point is assumed to be equivalent and that time points are assumed to be independent. Even though GPFA explicitly finds temporal correlations for each neural state, this independence assumption achieves a tractable covariance matrix estimate. In each fold, we train and test using the slowest latent neural state and incrementally add faster states until we include all latent states and vice versa. This is done to determine whether separability is due to slow or fast states and whether there is additive information when including both. Additionally, we train and test using a linear discriminant analysis where the class covariances are assumed to be equal as a control. Finally, we train and

test using one latent neural state at a time to determine how much each latent neural state contributes to behavioral state separability.

3.3 Results

3.3.1 Decoding Performance Using First and Second-Order Statistics

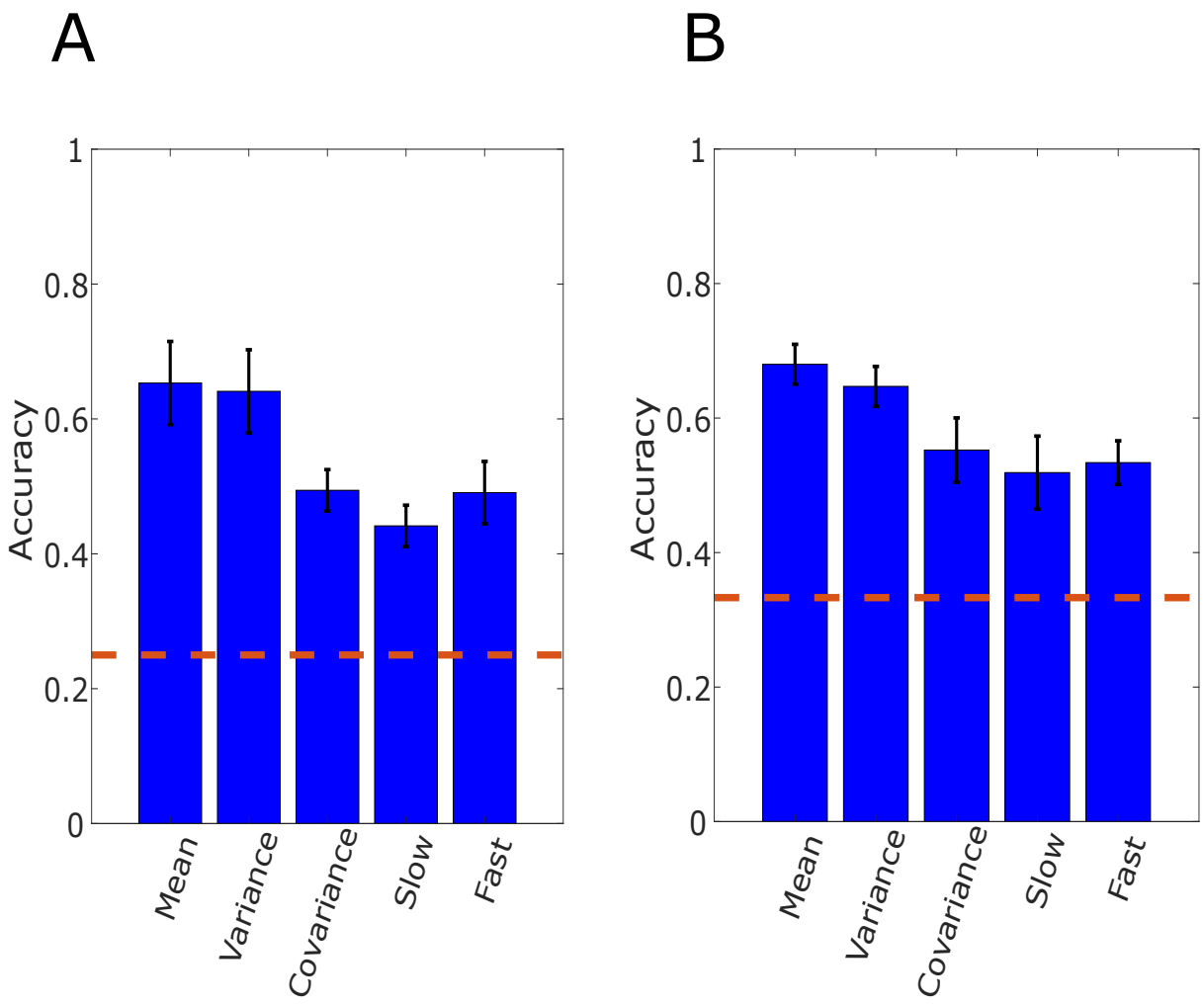


Figure 3.2: **Decoding performance using first and second-order statistics.** The decoding performance for Subjects 1 and 2 are shown in (A) and (B), respectively. All five different spatial and temporal components in the neural activity are able to discriminate between abstract behavioral states. The slow varying mean and variance achieves the highest decoding performance. The covariance matrix is also able to separate behavioral states without leveraging single channel variances. Therefore, inter-electrode correlations contribute to behavioral state separability. Finally, both the high gamma band amplitude estimate's slow and fast components contain information regarding the behavioral state.

Fig 3.2 shows the classification performance using the mean, variance, covariance of

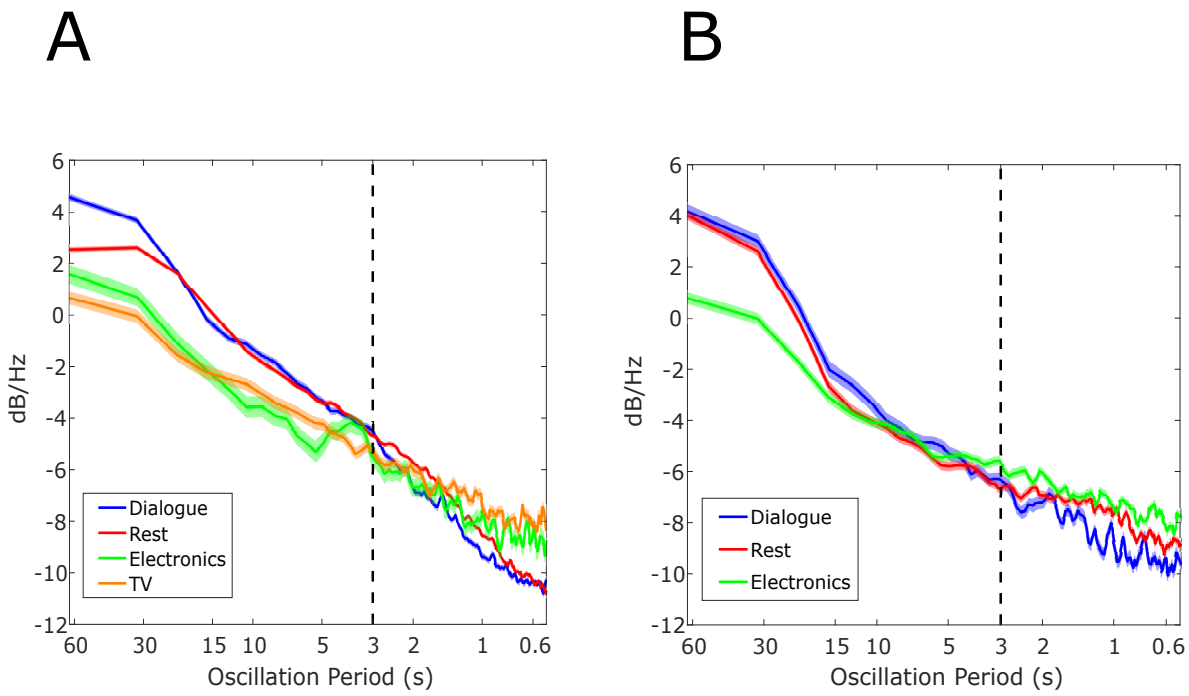


Figure 3.3: **Power spectral density estimates of binned high gamma amplitude estimate.** Sample electrodes are shown for Subjects 1 and 2 in (A) and (B), respectively. The power spectral densities in both sample electrodes show that different behavioral contexts exhibit varying temporal dynamics at different timescales. For example, in (A), 'Dialogue' and 'Rest' have more power in comparison to 'Electronics' and 'TV' for temporal dynamics slower than 3 seconds but vice versa for dynamics faster than 3 seconds. Most of the signal variance is concentrated in the slower components for both subjects.

the high gamma amplitude estimate, and the mean of the slow and fast components. The error bars of the 7-fold cross-validation are shown. Slow and fast components are oscillations in the high gamma amplitude estimate that is slower or faster than 3 seconds. For both subjects, using any of the five signal characteristics, we are able to achieve a classification performance well above chance level (25% and 33% for Subjects 1 and 2, respectively). This is evidence that in these naturalistic behavioral contexts, the differences in the spatial and temporal dynamics of neural activity most likely resemble the example displayed in Fig 3.1. To investigate the differences in temporal dynamics of the high gamma amplitude estimate, we show the power spectral densities of exemplar channels for each behavioral state for both subjects in Fig 3.3. Firstly, the power spectral density decreases monotonically as a function of frequency, corroborating previous studies [13, 12]. Most of the signal variance is clustered in the slower frequencies. Therefore, applying a decoding algorithm on the unfiltered high gamma amplitude would focus on the separability shown in the high gamma envelope's slower

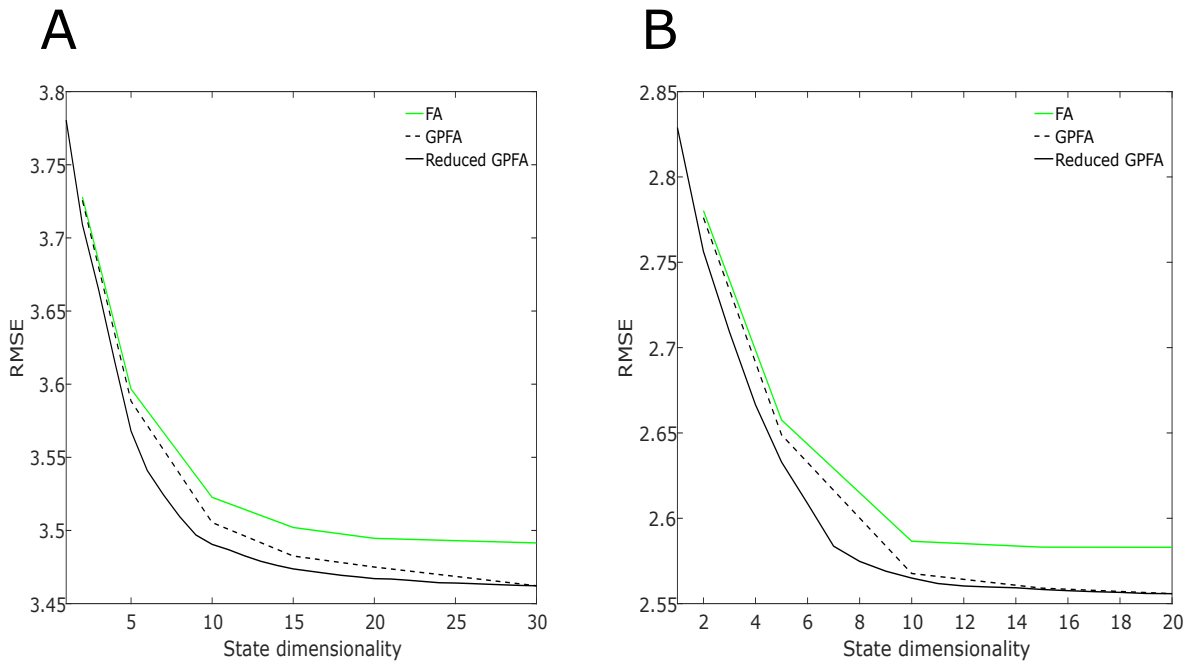


Figure 3.4: **Root mean square error vs. latent state dimensionality determined from FA, GPFA, and reduced GPFA.** (A) and (B) show the RMSE using the leave-one-electrode-out method for Subjects 1 and 2, respectively. For both subjects, the error decays as a function of state dimensionality, indicating the presence of complex spatial components in the neural activity. Additionally, for both subjects, the RMSE in FA is higher than GPFA for all dimensions, and the difference in RMSE increases as more latent states are learned. The orthogonal factors determined using reduced GPFA further reduces the error and shows that there exists a smaller set of spatial components that could reconstruct the neural activity as well as its higher dimension non-orthogonal counterparts. For Subject 1 in (A), reduced GPFA reconstructs the neural activity using 20 components instead of 30 via GPFA. The same phenomenon is shown in (B) for Subject 2 (around 12 factors in reduced GPFA has around the same RMSE value as 20 factors using GPFA.)

frequency components. This is illustrated in Fig 3.3, where, for example, 'Dialogue' has higher power in the slower frequencies but lower power in the higher frequencies in comparison to the other behavioral states. However, multiple channels do not exhibit a monotonically decreasing power as a function in frequency, as shown in Supplementary Fig 3.9 and 3.10.

3.3.2 Reconstruction Error vs. Dimensionality

To assess whether the GPFA model displayed in the Methods section is an accurate representation of the data, and therefore, the neural activity is generated from a latent subspace with variable temporal smoothness, we compared the reconstruction error vs. number of latent dimensions of FA, GPFA, and reduced GPFA. In Fig 3.4, we show that

for both subjects and across a 7-fold cross-validation paradigm, the root mean square error (RMSE) monotonically decreases as a function of the latent space dimensionality. This shows that there exists a high dimensional subspace that could reliably describe neural activity occurring during unstructured activity. There is a clear drop in RMSE when comparing FA performance versus that of GPFA when the number of latent dimensions increases. This provides us with evidence that several spatial components in the neural activity with variable smoothness in time exist in the data, and that these spatial components, or latent factors, are repeatable and are not limited to a local temporal window. Reduced GPFA shows that many of the spatial components found in the data are non-orthogonal, and after fitting 30 or 20 factors on the data for Subject 1 and 2, respectively. We show that there is a smaller orthonormal latent space that could reconstruct the data as well as a larger non-orthogonal latent space. We also compare GPFA and Factor Analysis's performances with varying degrees of smoothness applied to the data before fitting. This is done as a control to test whether there is a single autocorrelation function common across all latent neural states and that there are latent neural states with varying temporal dynamics. We found that there does not exist any universal autocorrelation function that led to Factor Analysis having the same or better reconstruction error as GPFA. This control provides further evidence that both slow and fast varying latent factors are repeatable in the data.

3.3.3 Decoding Performance using Gaussian Process Factors

Once we have established that latent dimensions with variable temporal smoothness exist in the neural activity, a natural extension determines whether these latent dimensions contain information regarding unstructured behavioral states. Fig 3.5 displays the decoding performance using a QDA classifier and 7-fold cross-validation with a 1-fold buffer. We sorted the factors from slowest to fastest and vice versa to test whether behavioral state-specific information is present at either extreme. The decoder's performance increases monotonically for both subjects, whether starting from the slowest factor or the fastest. This shows that there is relevant information at both scales regarding behavioral state and confirms the results we described earlier. A key difference from our earlier analyses of the first and second-order

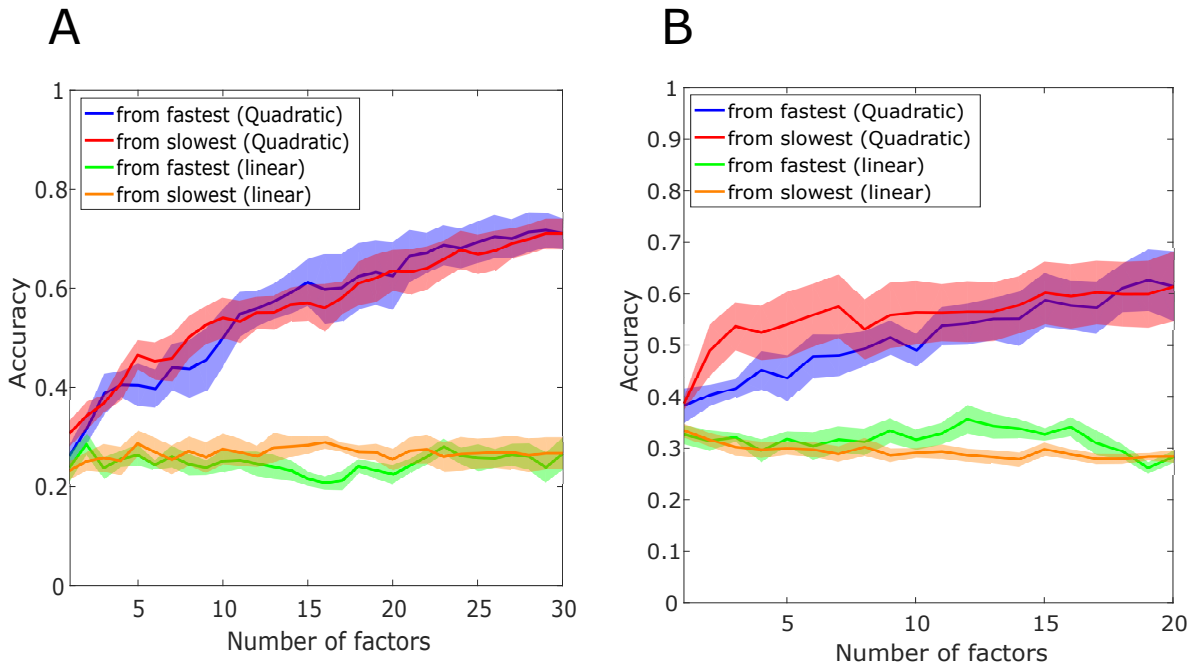


Figure 3.5: **Number of latent factors vs. decoding accuracy.** The decoding accuracy as a function of the number of factors generated by GPFA used (by incremental addition from slowest to fastest factors and vice versa) is shown in (A) and (B) for Subjects 1 and 2, respectively. Using a quadratic discriminant analysis as the classification algorithm, the separability of 30-second epochs increases as a function of the number of factors used. Additionally, for both subjects, there is additive information in both fast and slow factors. As a control, we used an LDA to show that separability is due to differences in the latent neural trajectories' power rather than their mean (which would indicate a behavioral state-specific shift in variance or mean in the data).

statistics is that these fast and slow neural activity aspects are synchronized activity between multiple channels. Additionally, as a control, we applied a Linear Discriminant Analysis (LDA) classifier to test whether the separability is due to a behavioral state-specific shift in the latent factors' mean. The LDA classifiers' performance is close to chance, suggesting that the behavioral states' separability is due to differences in the variance and covariance of the latent Gaussian process factors. In other words, the power of each factor is what is separating behavioral states. The importance of each latent factor is evident in Fig 3.6. We use each latent factor as a single feature in the QDA. The slow and fast factors contain information regarding behavioral states for both subjects, confirming our results earlier. Additionally, there is not a single factor contributing to most of the decoding performance, showing that there are multiple dimensions containing relevant information. Not surprisingly, multiple factors (three for Subject 1, and one for Subject 2) contain almost no behavioral

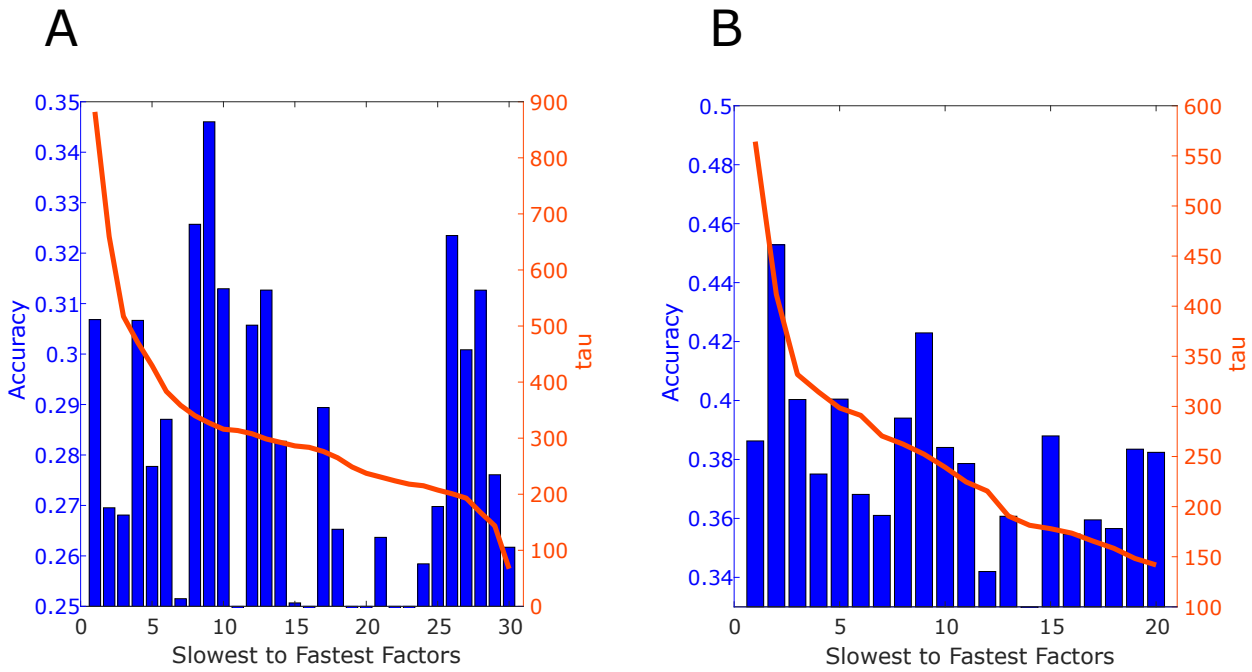


Figure 3.6: **Decoding accuracy of single factors.** The decoding accuracy of single factors generated by the GPFA algorithm is shown for Subjects 1 and 2 in (A) and (B), respectively. The factors are sorted from slowest to fastest. We show that there is relevant information in latent factors spanning different timescales. Additionally, several factors are extracted from the data that does not contain any information regarding behavioral states. The unit of tau is milliseconds.

state information. This is not surprising as there must be brain states independent of the behavioral states we have labeled in this dataset.

3.3.4 Visualization of Latent Factors

Fig 3.7 shows the latent factor weights for Subjects 1 and 2 in subsections B and D, respectively. The latent factors determined from GPFA are sorted from slowest to fastest. Firstly, since we know that many factors contribute to behavioral state separability, as shown in Fig 3.6 and Fig 3.7, it is clear from Fig 3.7B and D that these factors span multiple regions in the brain, indicative that multidimensional brain-wide activity is activated during unstructured behavioral states. Furthermore, many factors show both positive and negative values for different electrodes, indicating that synchronized activation and deactivation of different regions of the brain are descriptive of naturalistic and unstructured behavioral states. Fig 3.8 displays a sample epoch where the fastest and slowest factor fitted from GPFA is

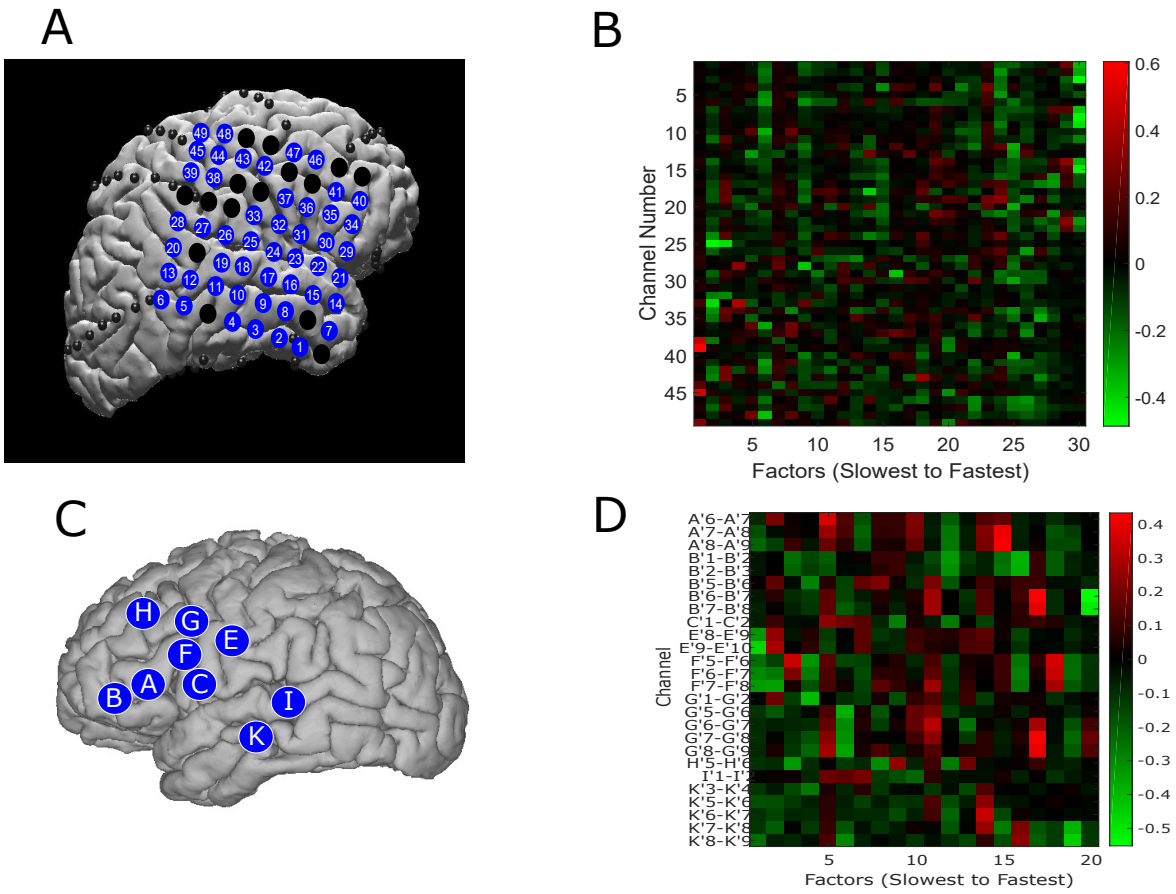


Figure 3.7: **Visualizations of latent factors extracted from GPFA.** We show the values of the factor weights extracted from GPFA for Subject 1 and 2 in (B) and (D), respectively. (A) shows the coverage of the ECoG grid in the cortex for Subject 1 as well as the corresponding channel numbers. (C) shows the location of where the sEEG shanks penetrated the cortex for Subject 2. The latent factors' values show that there is synchronized activity corresponding to different areas of the cortex. Additionally, some factors have both positive and negative values assigned to different channels, indicating that some factors exhibit synchronized activation and deactivation of certain brain regions.

shown. We only display the projections for Subject 1 since the ECoG cortical coverage is easier to visualize. For these two sample latent factors, most of the electrode projection values are all positive or all negative, therefore, the sign does not provide any additional information about the spatial structure of these factors. We multiply the fastest factor values by -1 for visualization purposes (in order to observe the synchronization between the single channel activity and the latent factor). Fig 3.8B and D show that the fastest latent factors track multiple high frequency spikes in neural activity, while the slowest latent factor tracks the slow varying component. This visualization illustrates the existence of coherent and informative spatiotemporal patterns of unstructured behavioral states. It is important to

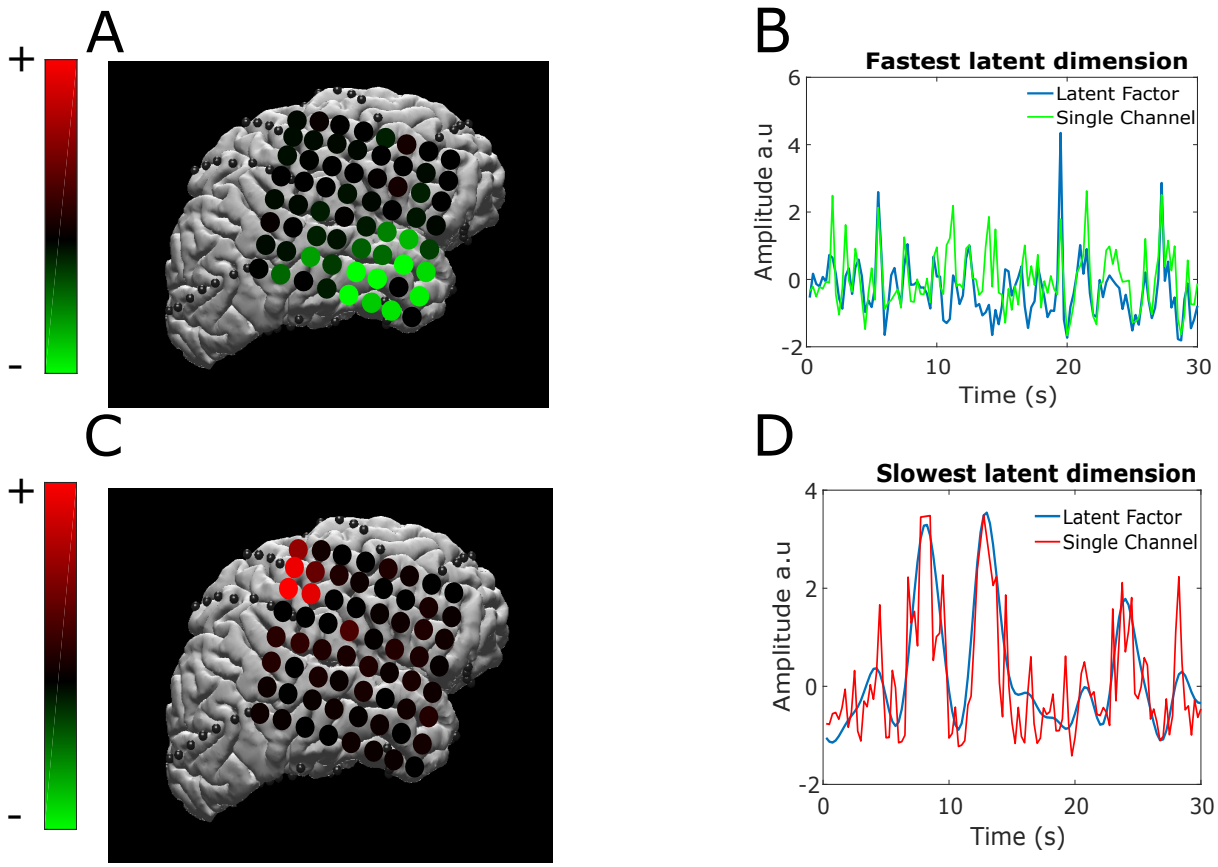


Figure 3.8: **Sample latent neural trajectories with variable temporal dynamics.** For Subject 1, we show the fastest and slowest latent factors extracted from the data. The factors' spatial weights are shown in the cortex in (A) and (C). Sample trajectories of the fastest and slowest factor of a single epoch are shown in (B) and (D) overlaid with the channel's value that contributes most to the corresponding factor. The factor in (B) is inverted in order to match the sign of the channel for visualization purposes. In (B), the latent factors track many high frequency peaks in the single channel amplitude. Additionally, (D) shows that the slow latent factor tracks the slow moving component of the channel amplitude.

note that these two latent factors are displayed here to show the value in using GPFA to visualize factors with different timescales. These two latent factors need not be the most important features in separating behavioral states. Additionally, these latent states are learned from a downsampled training set to ensure an even distribution of behavioral context, and therefore, do not exploit the full power of the data. Finally, we show in Supplemental Fig 3.13 an alternative approach in visualizing the latent factors, where we applied SVD to orthonormalize the latent factors and rank them according to the amount of variance

explained (as shown in [36]). We can also visualize the PSDs of these orthonormal factors to understand how each orthonormal factor's temporal dynamics differ according to behavioral state, as seen in Fig 3.12 and 3.13.

3.4 Discussion

This paper shows that multiple sources of information in the neural activity could be leveraged to discriminate between behavioral states labeled in an unstructured setting. The mean and variance of 30 seconds of the high gamma band amplitude could be used to achieve a decoding performance significantly higher than the chance level. This separability could be due to a myriad of factors such as slow shifts in brain states or slight changes in recording quality along with behavioral state-specific neural activity. The ability to discriminate states using the covariance matrix of the high gamma band amplitude after normalization means that decoding is still possible without any knowledge about the power in the neural activity. This confirms that the connectivity between different brain regions could be used to understand the differences between behavioral states. The filtered and normalized high gamma band amplitude produces both slow and fast features and provides information regarding the behavioral state. Even though the decoding performance is less than using the unfiltered and unnormalized mean and variance, this result shows that both the long term fluctuations of neural activity and faster responses separate behavioral states. The slower responses could be attributed to information accumulation of slow fluctuations in brain state, such as changes in hunger or attentiveness, driving behavior. Faster responses could be due to the brain's reaction to external stimuli.

Our results from the GPFA analyses confirm that naturalistic and unstructured behavior could be used to extract relevant spatiotemporal patterns in the neural activity. Regions of the brain moving in synchrony with a specific temporal structure provide meaningful information regarding the behavioral state. These spatiotemporal patterns are extracted in the absence of long term mean and variance shifts, therefore, we are confident that these patterns are not due to some noise related brain-wide change in activity. If we applied

GPFA on unnormalized neural data, an alternative scenario could be that one context has a predominantly high activity for most spatiotemporal patterns, therefore, understanding how these patterns relate to behavior would not be straightforward.

Previous studies have developed novel algorithms to extract spatiotemporal patterns in neural activity [9, 37, 38]. However, these algorithms do not consider the many sources of variability present in long-term neural activity collected in an unstructured setting. Additionally, the presence of non-behavior related neural activity and the lack of trial structure make this avenue of research challenging. Not only that, but relevant neural activity need not be related to externally observed changes in behavior but could be due to unobserved internal activity such as changes in mood, hunger, thirst, or mental engagement. In this work, we leverage GPFA due to the lax enforcement of numerous assumptions, such as sinusoidality, trial locking, or presence of sequences. However, GPFA and other algorithms that were used in similar contexts are unsupervised methods to fit a model to describe the neural activity. While applying supervised deep learning methods to analyze spatiotemporal patterns in neural activity is undoubtedly a thriving avenue of research, these methods suffer from being a black box making it difficult to understand how the algorithm models the neural activity. Therefore, it is imperative to develop supervised, neuroscience-based models, accounting for multiple variability sources, to truly understand how the brain behaves when unhindered by experimental constraints.

Trial-based experiments are designed in order to find the neural correlates to specific behaviors, in other words, to separate the signal from the noise. However, in this work, we have shown that multiple features of neural activity, spanning different time scales and spatial distributions, correlate to spontaneous behaviors. Huk and Hart [39] make the case that researchers need to move away from linking neural activity to externally measured variables and that neural variability that could be attributed as noise is actually linked to internal states that might influence behavior. In our work, we provide evidence to support this claim. Slow spatiotemporal dynamics clearly contribute to the separability of observable behavioral states in two subjects. In a controlled trial-based experimental paradigm, these slower dynamics could be ignored as noise. Additionally, Fig 3.7 show that factors learned from GPFA have

brain-wide representation, confirming a previous study that showed that brain-wide latent states are correlated to spontaneous observable behaviors in mice [18]. We show that in studying brain activity generated from an unconstrained, spontaneous setting across hours and days, multiple spatiotemporal features at different time scales correlate to behavior in humans. We aim to motivate researchers to understand the brain representation of behavior and internal states and supplement the traditional trial-based experimental paradigm with unconstrained longitudinal analysis of brain activity.

3.5 Acknowledgments

This work has been supported, in part, by the UCSD ECE Department Medical Devices & Systems Initiative, the UCSD Centers for Human Brain Activity Mapping (CHBAM) and Brain Activity Mapping (CBAM), the UCSD Frontiers of Innovation Scholars Program, the Qualcomm Institute Calit2 Strategic Research Opportunities (CSRO) program, the Hellman Fellowship, the Institute of Engineering in Medicine Graduate Student Fellowship, and the Clinical and Translational Research Institute at UC San Diego. We would like to thank the patients and clinicians who contributed to this study at UC San Diego, Rady Children's Hospital of San Diego, and the Comprehensive Epilepsy Center at NYU Langone Medical Center. We specifically thank Preet Minas and Hugh Wang for their contributions to data acquisition. This chapter is a reprint of material currently being prepared for submission in Alasfour, Abdulwahab, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan, Daniel Friedman, Werner Doyle, Orin Devinsky, David Gonda, Shifteh Sattar, Sonya Wang, Eric Halgren, and Vikash Gilja. "Identifying neural signal characteristics that discriminate between naturalistic behavioral states."

3.6 Supplementary Data

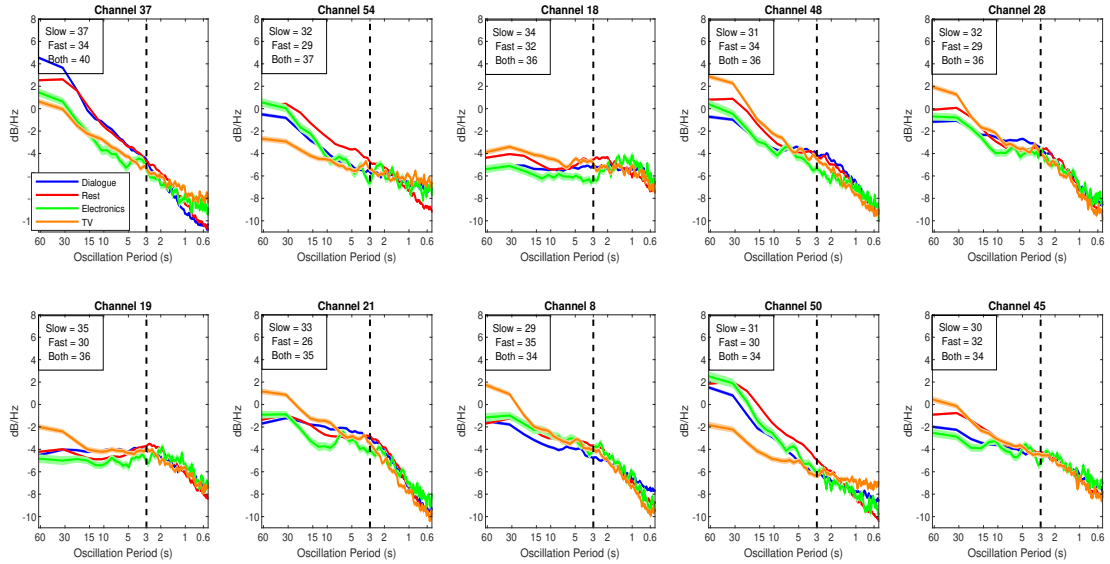


Figure 3.9: **Power spectral density of top 10 performing channels using slow and fast temporal dynamics as features for Subject 1.**

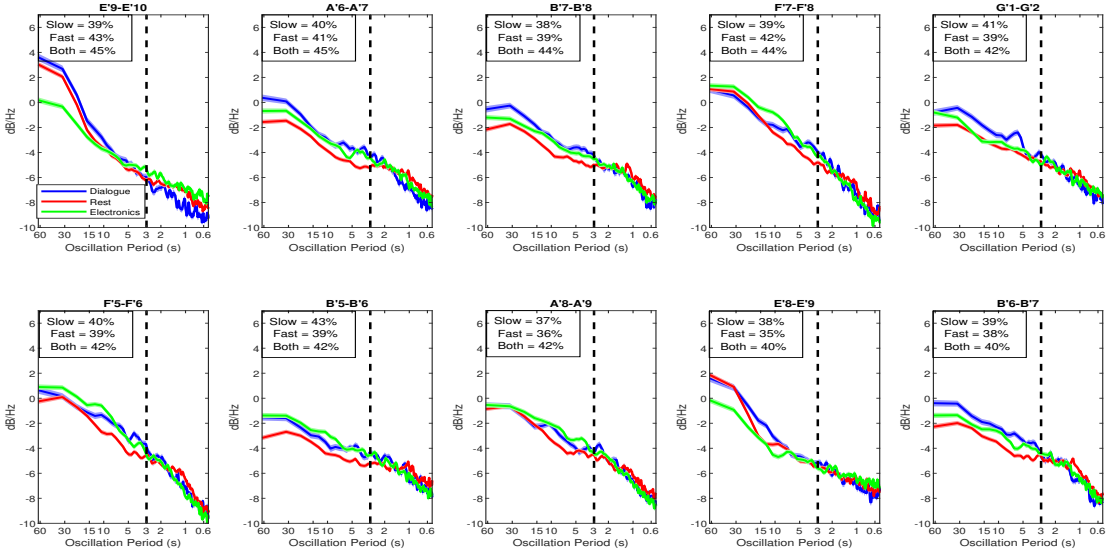


Figure 3.10: **Power spectral density of top 10 performing channels using slow and fast temporal dynamics as features for Subject 2.**

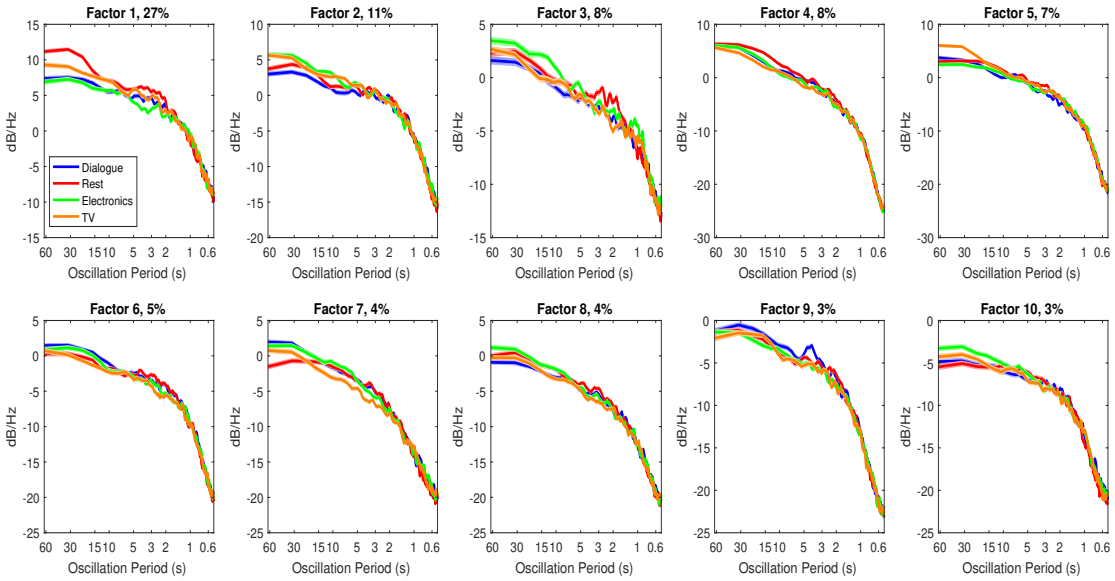


Figure 3.11: **Power spectral density of top 10 orthonormal factors for Subject 1.** The percentage of variance explained for each factor is displayed in the title of each factor

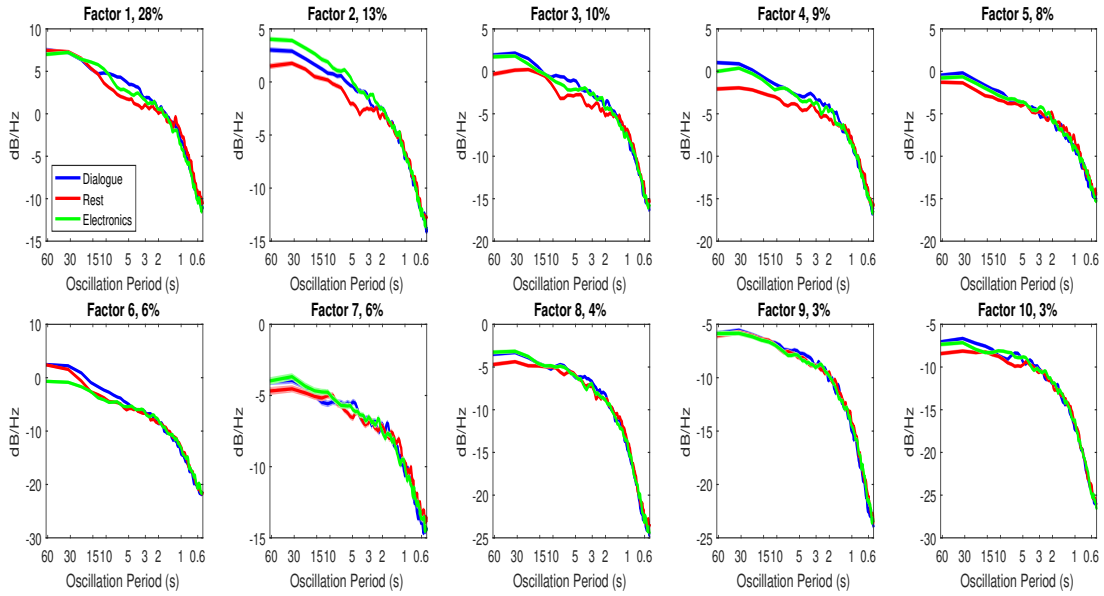


Figure 3.12: **Power spectral density of top 10 orthonormal factors for Subject 2.** The percentage of variance explained for each factor is displayed in the title of each factor

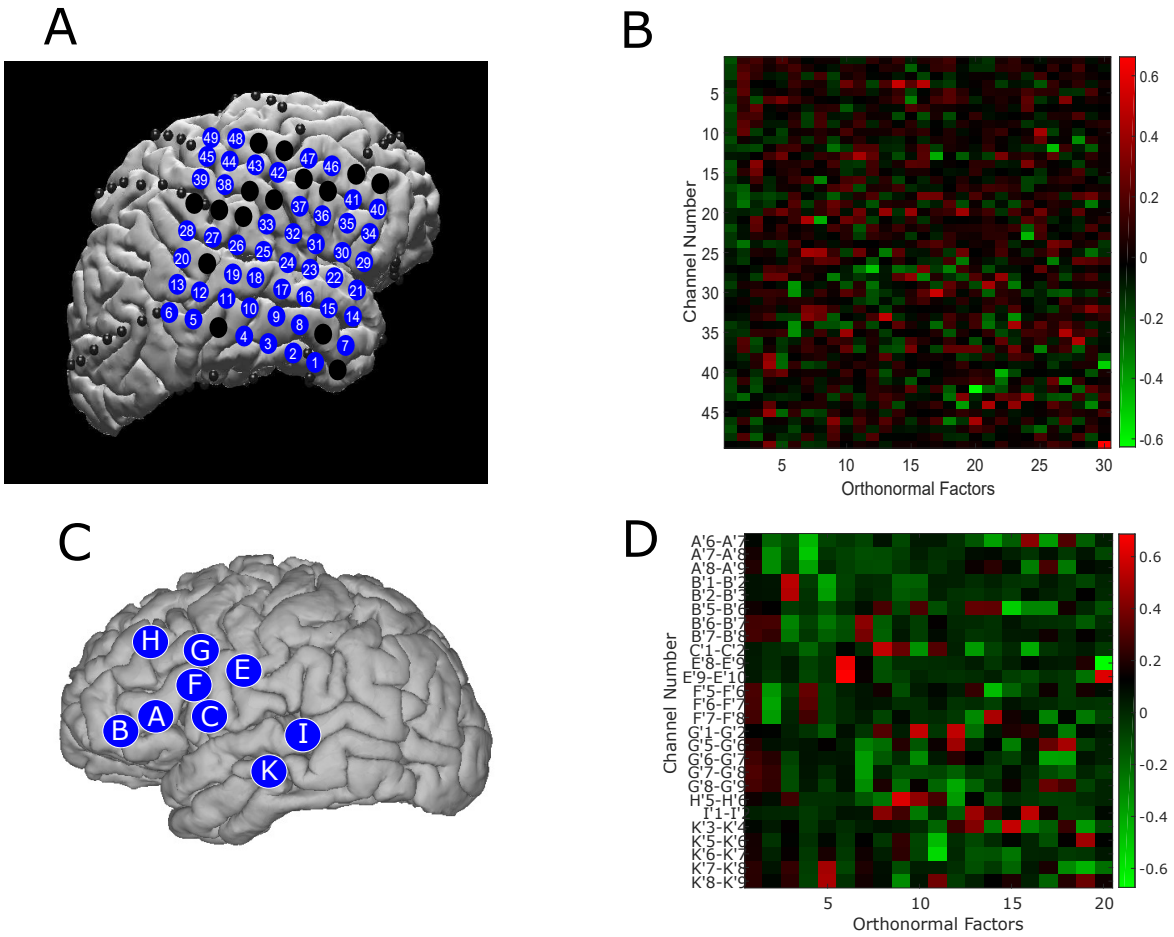


Figure 3.13: **Orthonormal factor weights for both subjects.** Locations of ECoG channels and sEEG shank entry points are shown for Subject 1 and 2 in subpanels A and C respectively.

References

[1] Kai J. Miller, Eric C. Leuthardt, Gerwin Schalk, Rajesh P.N. Rao, Nicholas R. Anderson, Daniel W. Moran, John W. Miller, and Jeffrey G. Ojemann. Spectral changes in cortical surface potentials during motor movement. *Journal of Neuroscience*, 27(9):2424–2432, 2007.

[2] Kai J Miller, Christopher J Honey, Dora Hermes, Rajesh PN Rao, Marcel denNijs, and Jeffrey G Ojemann. Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *New Horizons for Neural Oscillations*, 85:711–720, January 2014.

[3] Kristofer E. Bouchard, Nima Mesgarani, Keith Johnson, and Edward F. Chang. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441):327–332, 2013. Publisher: Nature Publishing Group.

[4] Miguel Angrick, Christian Herff, Emily Mugler, Matthew C. Tate, Marc W. Slutzky, Dean J. Krusienski, and Tanja Schultz. Speech synthesis from ECoG using densely connected 3D convolutional neural networks. *Journal of Neural Engineering*, 16(3), 2019. Publisher: IOP Publishing.

[5] Spencer Kellis, Kai Miller, Kyle Thomson, Richard Brown, Paul House, and Bradley Greger. Decoding spoken words using local field potentials recorded from the cortical surface. *Journal of Neural Engineering*, 7(5), 2010.

[6] Gopala K. Anumanchipalli, Josh Chartier, and Edward F. Chang. Speech synthesis from neural decoding of spoken sentences. *Nature*, 568(7753):493–498, 2019.

[7] William E. Allen, Michael Z. Chen, Nandini Pichamoorthy, Rebecca H. Tien, Marius Pachitariu, Liqun Luo, and Karl Deisseroth. Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science*, 364(6437), April 2019. Publisher: American Association for the Advancement of Science Section: Research Article.

[8] Andrew J. Peters, Simon X. Chen, and Takaki Komiyama. Emergence of reproducible spatiotemporal activity during motor learning. *Nature*, 510(7504):263–267, June 2014. Number: 7504 Publisher: Nature Publishing Group.

[9] Alex H. Williams, Tony Hyun Kim, Forea Wang, Saurabh Vyas, Stephen I. Ryu, Krishna V. Shenoy, Mark Schnitzer, Tamara G. Kolda, and Surya Ganguli. Unsupervised Discovery of Demixed, Low-Dimensional Neural Dynamics across Multiple Timescales through Tensor Component Analysis. *Neuron*, 98(6):1099–1115.e8, June 2018. Publisher: Cell Press.

[10] Adam J. Calhoun, Jonathan W. Pillow, and Mala Murthy. Unsupervised identification of the internal states that shape natural behavior. *Nature Neuroscience*, 22(12):2040–2049, December 2019. Publisher: Nature Research.

[11] Alexander Huk, Kathryn Bonnen, and Biyu J. He. Beyond trial-based paradigms: Continuous behavior, ongoing neural activity, and natural stimuli. *Journal of Neuroscience*,

38(35):7551–7558, August 2018. Publisher: Society for Neuroscience.

[12] Yuxiao Yang, Omid G. Sani, Edward F. Chang, and Maryam M. Shanechi. Dynamic network modeling and dimensionality reduction for human ECoG activity. *Journal of Neural Engineering*, 16(5), August 2019. Publisher: Institute of Physics Publishing.

[13] David A Leopold, Yusuke Murayama, and Nikos K Logothetis. Very Slow Activity Fluctuations in Monkey Visual Cortex: Implications for Functional Brain Imaging. *Cerebral Cortex*, 13(4):422–433, April 2003.

[14] Yuval Nir, Roy Mukamel, Ilan Dinstein, Eran Privman, Michal Harel, Lior Fisch, Hagar Gelbard-Sagiv, Svetlana Kipervasser, Fani Andelman, Miri Y. Neufeld, Uri Kramer, Amos Arieli, Itzhak Fried, and Rafael Malach. Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. *Nature Neuroscience*, 11(9):1100–1108, September 2008.

[15] Christopher J. Honey, Thomas Thesen, Tobias H. Donner, Lauren J. Silbert, Chad E. Carlson, Orrin Devinsky, Werner K. Doyle, Nava Rubin, David J. Heeger, and Uri Hasson. Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron*, 76(2):423–434, 2012. Publisher: Elsevier Inc.

[16] Greg J. Stephens, Christopher J. Honey, and Uri Hasson. A place for time: the spatiotemporal structure of neural dynamics during natural audition. *Journal of Neurophysiology*, 110(9):2019–2026, November 2013.

[17] David A. McCormick, Dennis B. Nestvogel, and Biyu J. He. Neuromodulation of Brain State and Behavior. *Annual Review of Neuroscience*, 43(1):391–415, 2020. _eprint: https://doi.org/10.1146/annurev-neuro-100219-105424.

[18] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D. Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893, April 2019.

[19] Mohammad Dastjerdi, Muge Ozker, Brett L. Foster, Vinitha Rangarajan, and Josef Parvizi. Numerical processing in the human parietal cortex during experimental and natural conditions. *Nature Communications*, 4, 2013. Publisher: Nature Publishing Group.

[20] Nancy X.R. Wang, Jared D. Olson, Jeffrey G. Ojemann, Rajesh P.N. Rao, and Bingni W. Brunton. Unsupervised decoding of long-term, naturalistic human neural recordings with automated video and audio annotations. *Frontiers in Human Neuroscience*, 10(APR2016):1–13, 2016.

[21] Paolo G. Gabriel, K. J. Chen, A. Alasfour, T. Pailla, W. K. Doyle, O. Devinsky, D. Friedman, P. Dugan, L. Melloni, T. Thesen, D. Gonda, S. Sattar, S. G. Wang, and V. Gilja. Neural correlates of unstructured motor behaviors. *Journal of Neural Engineering*, 16(6), October 2019. Publisher: Institute of Physics Publishing.

[22] Steven M. Peterson, Satpreet H. Singh, Nancy X. R. Wang, Rajesh P. N. Rao, and Bingni W. Brunton. Behavioral and neural variability of naturalistic arm movements. *bioRxiv*, page 2020.04.17.047357, 2020.

[23] Omid G Sani, Yuxiao Yang, Morgan B Lee, Heather E Dawes, Edward F Chang, and Maryam M Shanechi. Mood variations decoded from multi-site intracranial human brain activity. *Nature Biotechnology*, 36(10):954–961, November 2018.

[24] Abdulwahab Alasfour, Paolo Gabriel, Xi Jiang, Isaac Shamie, Lucia Melloni, Thomas Thesen, Patricia Dugan, Daniel Friedman, Werner Doyle, Orin Devinsky, David Gonda, Shifteh Sattar, Sonya Wang, Eric Halgren, and Vikash Gilja. Coarse behavioral context decoding. *Journal of Neural Engineering*, 16(1), February 2019. Publisher: Institute of Physics Publishing.

[25] Paolo Gabriel, Werner K. Doyle, Orrin Devinsky, Daniel Friedman, Thomas Thesen, and Vikash Gilja. Neural correlates to automatic behavior estimations from RGB-D video in epilepsy unit. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016-Octob:3402–3405, 2016. ISBN: 9781457702204 Publisher: IEEE.

[26] Brian T. Quinn, Chad Carlson, Werner Doyle, Sydney S. Cash, Orrin Devinsky, Charles Spence, Eric Halgren, and Thomas Thesen. Intracranial cortical responses during visual-tactile integration in humans. *Journal of Neuroscience*, 34(1):171–181, 2014.

[27] Kai J Miller, Gerwin Schalk, Eberhard E Fetz, Marcel den Nijs, Jeffrey G Ojemann, and Rajesh P N Rao. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proceedings of the National Academy of Sciences*, 107(9):4430 LP – 4435, March 2010.

[28] Brian N. Pasley, Stephen V. David, Nima Mesgarani, Adeen Flinker, Shihab A. Shamma, Nathan E. Crone, Robert T. Knight, and Edward F. Chang. Reconstructing speech from human auditory cortex. *PLoS Biology*, 10(1), 2012.

[29] Jorge González-Martínez, Juan Bulacio, Susan Thompson, John Gale, Saksith Smithason, Imad Najm, and William Bingaman. Technique, Results, and Complications Related to Robot-Assisted Stereoelectroencephalography. *Neurosurgery*, 78(2):169–180, February 2016.

[30] Roy Mukamel, Hagar Gelbard, Amos Arieli, Uri Hasson, Itzhak Fried, and Rafael Malach. Neuroscience: Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science*, 309(5736):951–954, August 2005.

[31] Jeremy R. Manning, Joshua Jacobs, Itzhak Fried, and Michael J. Kahana. Broadband Shifts in Local Field Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans. *Journal of Neuroscience*, 29(43):13613–13620, October 2009. Publisher: Society for Neuroscience Section: Articles.

[32] Supratim Ray and John H. R. Maunsell. Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLOS Biology*, 9(4):e1000610, April 2011. Publisher: Public Library of Science.

[33] F. Aoki, E. E. Fetz, L. Shupe, E. Lettich, and G. A. Ojemann. Increased gamma-range activity in human sensorimotor cortex during performance of visuomotor tasks. *Clinical Neurophysiology*, 110(3):524–537, 1999.

[34] N. Crone, Diana L. Miglioretti, Barry Gordon, and Ronald P. Lesser. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. *Brain*, 121:2301–2315, 1998. ISBN: 0006-8950 (Print).

[35] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass Brain–Computer Interface Classification by Riemannian Geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928, April 2012. Conference Name: IEEE Transactions on Biomedical Engineering.

[36] Byron M. Yu, John P. Cunningham, Gopal Santhanam, Stephen I. Ryu, Krishna V. Shenoy, and Maneesh Sahani. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology*, 102(1):614–635, July 2009.

[37] Bingni W. Brunton, Lise A. Johnson, Jeffrey G. Ojemann, and J. Nathan Kutz. Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, 258:1–15, January 2016. Publisher: Elsevier B.V. _eprint: 1409.5496.

[38] Emily L Mackevicius, Andrew H Bahle, Alex H Williams, Shijie Gu, Natalia I Denisenko, Mark S Goldman, and Michale S Fee. Unsupervised discovery of temporal sequences in high-dimensional datasets, with applications to neuroscience. *eLife*, 8:e38471, February 2019. Publisher: eLife Sciences Publications, Ltd.

[39] Alexander C. Huk and Eric Hart. Parsing signal and noise in the brain. *Science*, 364(6437):236, April 2019.

Chapter 4

High Gamma Activity in Cortex and Hippocampus is Correlated with Autonomic Tone During Sleep

Abstract

Studies in animals have demonstrated a strong coupling between cortical and hippocampal activity, and autonomic tone. However, the extent, distribution, and nature of this coupling have not been investigated with intracranial recordings in humans during sleep. Cortical and hippocampal population neuronal firing was estimated from high gamma band activity (HG) in local field potentials recorded from 15 subjects (9 females) during non-rapid eye movement (NREM) sleep. Autonomic tone was estimated from heart rate variability (HRV). HG and HRV were significantly correlated in both anterior and posterior hippocampus and multiple cortical sites in NREM stages N1-3. Coupling was related to delta power, was weaker in N1 and N2 as compared to N3, and was strongest in regions that have previously been associated with autonomic processes (such anterior hippocampus and insular cortex). The correlation between HG and HRV could be positive or negative across patients for a given anatomical location and sleep stage and was overall negative in the insula during N2/N3, suggestive of greater cortical activity associated with sympathetic tone. Additional studies are needed to define whether these patterns also occur during waking, and to disentangle causal

relationships between autonomic state and cortical and hippocampal activation. However, this study demonstrates that fluctuations in autonomic state are associated with significantly different patterns of cortical and hippocampal activation.

Significance Statement

Studies in animals have shown that the autonomic nervous system sets the operating mode of all the organ systems in the body, including the central nervous system. We show here that high gamma activity in widespread cortical and hippocampal regions is correlated with heart rate variability in humans during sleep. The correlation was especially profound in sites which have previously been associated with autonomic and emotional regulation. The direction of change varied between forebrain locations, indicating the existence of sympathetic and parasympathetic modulating structures. We found that the coupling between autonomic tone and cortical activity was greatest in the deepest stages of slow-wave sleep. Overall, this study characterizes in humans a foundational link in the unity of mind and body.

4.1 Introduction

The general state of the human organism is regulated by the autonomic nervous system (ANS) through afferent and efferent pathways. These pathways modulate various functions in the human body, such as digestion, blood pressure, heart rate, urination, sexual arousal, and others. A healthy autonomic system is key to maintain the balance of all these functions, adapt to different environmental stimuli, and keep the body in homeostasis. Two main divisions ensure this, the sympathetic and parasympathetic divisions. These divisions work together to keep our physiological processes at an appropriate level, with the sympathetic system responsible for priming our bodies for a fight or flight response. In contrast, the parasympathetic system promotes energy conservation and digestion. These systems are not only influencing the waking brain but are responsible for a healthy and regenerative sleep cycle. Within the brain, ascending pathways modulate the level of cortical activity, notably via noradrenergic fibers from the locus coeruleus and cholinergic fibers from the nucleus

basalis. These nuclei are functionally connected with the autonomic nervous system, which thus may treat the cerebral cortex as another internal organ to be maintained at an optimal level. In mice and monkey, pupillary dilation, which could be used as an index of autonomic modulation, was shown to be correlated with the firing of noradrenergic and cholinergic neurons; phasic pupillary dilations track the firing of noradrenergic axons, whereas sustained dilations track the firing of cholinergic axons [1, 2].

In humans, multiple studies have established a link between the central and autonomic nervous systems, mainly by examining correlations in task-induced fMRI and autonomic measurements. Meta-analyses [3, 4, 5, 6, 7] showed that the BOLD response in multiple cortical, limbic and hippocampal structures as well as the default mode network (DMN) are correlated with heart rate variability when observed through fMRI task-based and task-free experimental paradigms. However, the origin of this correlation is unclear. It could be a modulation of cortical firing by ascending noradrenergic and/or cholinergic pathways associated with the autonomic system, modulation of brainstem autonomic structures by corticofugal influences, or a viscerosensory response by cortical neurons, especially in the insula. Alternatively, the cortical BOLD modulation may not be due to neural activity, but rather to direct effects of the autonomic system on blood flow mediated by its well-known effects on blood pressure, heart rate, and vasodilation [8]. This potential confound does not arise in electrophysiological studies. A recent study found that the firing rate of human cingulate and parahippocampal gyrus neurons may show a negative or positive correlation with the heartrate [9]. There have also been studies that investigate the CNS-ANS coupling using EEG recordings throughout sleep [10, 11]. However, due to the limited temporal resolution of fMRI studies, the possibility that the BOLD signal can be influenced by local neurovascular modulation, and the lack of spatial resolution with EEG-sleep studies, little is known about the ANS-CNS coupling and the interplay between cortical/hippocampal structures with the ANS during sleep.

In this study, we aimed to investigate the correlation of heart rate variability, which is used as a metric for the autonomic tone, with intracranial recordings of cortical and hippocampal activity as indexed by high gamma band activity, which is known to be

positively correlated with neuronal firing rate [12, 13, 14, 15], during sleep. We focused on sleep because it is relatively free of other activities (such as eating, talking, or moving) which could influence either forebrain activity, autonomic state, or both. Additionally, we investigated the effect of the sleep stage on the correlation between autonomic tone and neural activity in different anatomical regions, since sleep stage and autonomic control are known to be correlated [16, 17]. To our knowledge, this is the first attempt to investigate the connection between the ANS and CNS during sleep using intracranial EEG recordings.

4.2 Methods

4.2.1 Patient Selection

Sixteen patients with long-standing drug-resistant partial seizures underwent stereo-electroencephalography (sEEG) depth electrode implantation to localize seizure onset and thus direct surgical treatment. We selected the 16 patients (10 female) from a group of 54 for this study that displayed minimal hippocampal pathology and had electrocardiogram (ECG) recordings. The average age was 30 and ranged from 16 to 58 years old. sEEG implantation was based completely on clinical needs [18]. All patients gave fully informed consent for data usage as monitored by the local Institutional Review Board, following clinical guidelines and regulations at Cleveland Clinic.

4.2.2 Electrode Localization

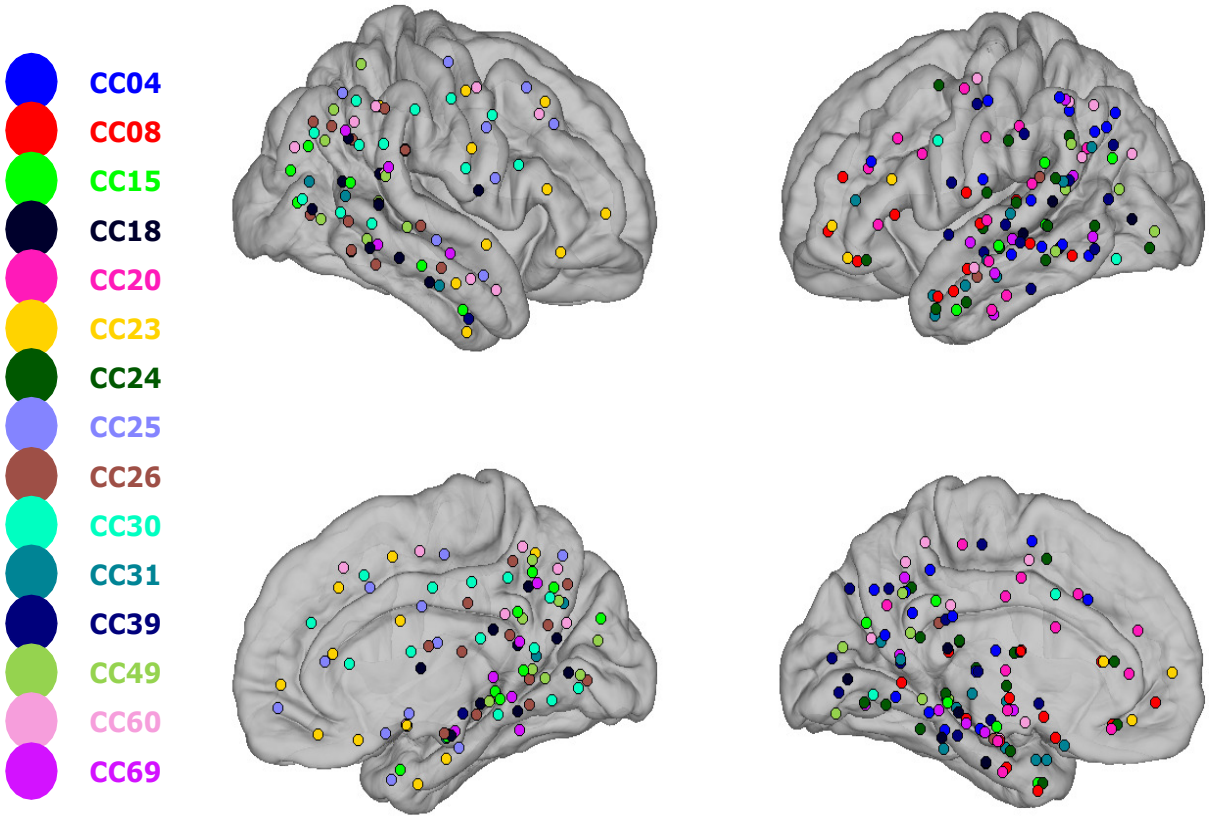
After implantation, electrodes were located by aligning postimplant CT to preoperative 3D T1-weighted structural MRI with 1 mm3 voxel size [4] using 3D Slicer (RRID: SCR-005619). The assignment of depth contacts to the anterior or posterior hippocampus was made with the posterior limit of the uncal head as boundary [19, 20]. The distance of each hippocampal contact from the anterior limit of the hippocampal head was obtained in Freesurfer (RRID:SCR-001847). The CT-visible cortical contacts were then identified as previously described [21], to ensure that activity recorded by bipolar transcortical pairs is locally generated [22]. Electrode contacts were rejected from analysis if they were involved

in the early stages of the seizure discharge or had frequent interictal activity or abnormal spontaneous local field potentials (LFPs.) A total of 368 bipolar pairschannels across 16 patients were accepted for further analysis (Figure 4.1a). Freesurfer [23, 24] was used to reconstruct from individual MRI scans the cortical pial and inflated surfaces, as well as automatic parcellation of the cortical surface into anatomical areas [25], after a sulcal-gyral alignment process. Additionally, standard FreeSurfer regions of interest (ROIs) were combined into 12 composite ROIs (Figure 4.1b), as well as different functional networks defined in [26] and adapted from network cllustering on resting-state fMRI data [27]. An average surface generated from previous work [28] of 20 patients, that included the 16 patients used in this study, served as the basis of all 3D maps. While each cortical sEEG electrode contact's location was obtained through direct correlation of CT and MRI as described earlier in this section, we obtained the cortical parcellation labels corresponding to each contact by morphing the right-anterior superior-oriented anatomical coordinates from individual surfaces to the average surface space [29]. All visualizations were created with custom scripts in MATLAB 2016b (The MathWorks). For the majority of this study, we focus our analyses on the 12 ROIs in Figure 4.1b.

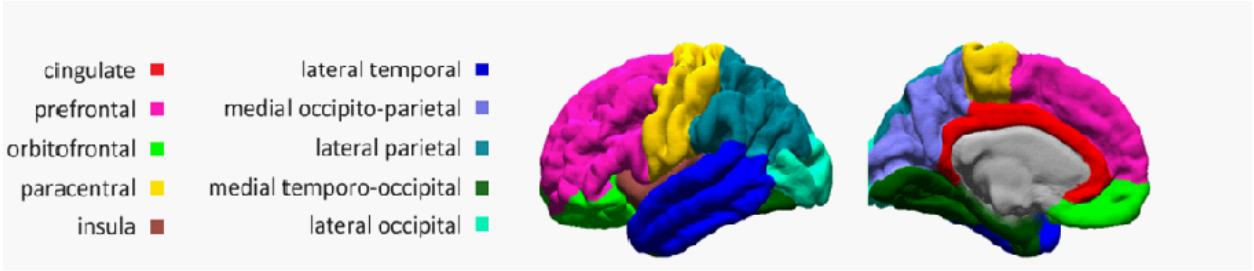
4.2.3 Data Collection and Preprocessing

sEEG preprocessing and sleep stage detection

Continuous recordings from SEEG depth electrodes were made with a cable telemetry system (JE-120 amplifier with 128 or 256 channels, 0.016 –3000 Hz bandpass, Neurofax EEG1200, Nihon Kohden) across multiple nights over the course of clinical monitoring for spontaneous seizures, with 1000 Hz sampling rate. The total NREM sleep durations vary across patients; while some difference is expected given intrinsic variability of normal human sleep duration (Carskadon and Dement, 2010) and sleep deprivation in a clinical environment. Recordings were anonymized and converted into the European Data Format (EDF). Separation of patient NREM sleep/wake states from intracranial LFP alone was achieved by previously described methods using clustering of first principal components of delta-to-spindle and delta-to-gamma power ratios across multiple LFP derived signal vectors



(a) **Electrode locations.** A: Anatomical locations of recording sites across 15 patients that are closest to the cortical and medial surfaces. Each dot location indicates the sEEG depth electrode entry point through the cortex or exit point though the medial view, therefore the locations are not exact as sEEG electrodes aren't necessarily directly on the surface. Dot color indicates patient ID.



(b) **Regions of interest (ROI)** Anatomical map that was used to group channel pairs for analysis

[30, 31], with the addition that separation of N2 and N3 was empirically determined by the proportion of down states (DSs) that are also part of slow oscillations (at least 50% for N3) [32] since isolated DSs in the form of K-complexes are predominantly found in Stage 2 sleep [33]. Due to difficulty in distinguishing between REM sleep and waking in this dataset, only NREM stages were used in this study.

The neural signals were then notched filtered at 60 Hz and its harmonics to remove line

noise. Channels were rejected if they hit white matter only, display epileptic activity, or do not lie in the cortex or hippocampus. The accepted bipolar channels were then filtered from 70-110 Hz using a zero-phase Chebychev type II IIR bandpass filter to extract the high gamma activity, which is shown to be highly correlated to neural spiking activity[12, 13, 14, 15]. The Hilbert envelope was obtained and then averaged across 1 minute non-overlapping windows to give an estimate of high gamma activity. Resulting band-passed and averaged bipolar channels were visually inspected to reject channels with noisy high gamma band activity. The same procedures were followed to investigate the effects of delta band mediation (see the 'Correlation Analysis' section below), with the exception of applying a 1-4 Hz zero-phase Chebychev type I IIR bandpass filter. The total NREM sleep durations vary across patients; beyond the normal variability (Carskadon and Dement, 2010), sleep may be disrupted in a clinical environment. Recordings were anonymized and converted into the European Data Format (EDF). Segmentation of patient NREM sleep/wake states from intracranial LFP alone was achieved by previously described methods using clustering of first principal components of delta-to-spindle and delta-to-gamma power ratios across multiple LFP derived signal vectors [30, 31], with the addition that separation of N2 and N3 was empirically determined by the proportion of down states (DSs) that are also part of slow oscillations (at least 50% for N3) [32] since isolated DSs in the form of K-complexes are predominantly found in Stage 2 sleep [33]. Due to difficulty in distinguishing between REM sleep and waking in this dataset, only NREM stages were used in this study.

Artifact Rejection

A dual artifact rejection criterion was used on the neural signals to identify and remove outlier 1-minute epochs. Firstly, any 1-minute epoch that exceeded 3 standard deviations above the channel high gamma band mean for more than 20 percent of the channels was considered as artifactual and rejected (comprising $1.5 \pm 1.5\%$ of 1-minute epochs across all patients and days). Additionally, for each minute the power spectral density was obtained to identify the presence of any peaks in the high gamma band. We only accepted epochs for which the modulation in the high gamma band presented itself as a broadband shift

that is added to the 1/f noise characteristic of neural signals, rather than an oscillatory-like bump. Broadband shifts in the high gamma band are most likely due to an overall increase in spiking activity, and it has been shown that asynchronous broadband shifts are linked to the fMRI BOLD response [34]. Narrowband peaks in the high gamma band would be evidence of synchronized oscillatory behavior [35] or noise. For this study, we disregarded any epochs that displayed evidence of high gamma oscillations or noise and only accepted epochs that show evidence of a broadband shift in power. To do this we fitted a line across the power spectral density (PSD) calculated for each 1-minute epoch across the 70-110 Hz frequency band. We then subtracted the fitted line from the PSD and look for the maximum value. If the max exceeded 3 dB across 20% of the channels, then that epoch is considered artifactual (10.8±12% of 1-minute epochs across all patients and days). Finally, any day's worth of recording where the number of total artifacts exceeded 50% was removed from subsequent analyses (this resulted in a single subject being rejected from further analyses and no days rejected for the other subjects).

Z-score Normalization

For each day, the 1-minute bipolar pairs were separated according to the sleep stage label assigned to that 1 minute time bin. Then, the 1-minute high gamma amplitude estimate recorded by each bipolar pair in each sleep stage for each sleep period was z-scored with its mean and standard deviation to arrive at the HGnorm. The same process was applied to the HRV vector to arrive at the HFnorm (see below). The HGnorm and HFnorm for each sleep stage were concatenated from each sleep period. This ensured that any linear correlations derived in subsequent analyses were due to within-sleep period within-sleep stage variations and not to a change in state that could happen across sleep periods or sleep stages within a specific sleep period.

ECG preprocessing and HRV frequency-domain analysis

Electrocardiogram recordings were also acquired throughout the days in which the patient was in the hospital. The ECG recordings were visually inspected and artifacts were

rejected if the raw value of the recording exceeded 3 standard deviations above the median, which indicates a movement artifact (0.34±0.67% across all patients and days). Furthermore, the ECG recordings were analyzed initially using the Kubios Premium HRV software [36]to detect the QRS wave and pinpoint the location of the R peaks (Figure 4.2 A). Then, the RR interval was generated by finding the time between adjacent R peaks, and artifacts were corrected for missed and ectopic beats using an automatic artifact rejection algorithm [37]. The RR-interval was then interpolated using a cubic spline interpolation and then sampled at 4 Hz to ensure even sampling of the signal, as this enabled us to extract accurate frequency-domain metrics from the data.

The resulting interpolated RR interval was then segmented into 1-minute bins and detrended using the smoothness priors algorithm that removed very low frequency (<0.035 Hz) and non-stationary trends in the RR interval, which could distort frequency domain analyses that require the signal to be stationary [38](Figure 4.2B). The mechanisms that drive the very low-frequency component of the RR-signal could be due to thermoregulatory cycles or changes in plasma renin activity, however, the mechanisms driving these slow oscillations are not well understood and therefore removed from any subsequent frequency domain analyses [39].

We then calculated the power spectral density of each detrended 1-minute segments of RR intervals using Welch's method with 30-second windows and 75% overlap. We used a hamming window to prevent spectral leakage. The power in the low-frequency (LF) band, from 0.04 Hz to 0.15 Hz, and the high-frequency (HF) band, from 0.15 Hz to 0.4 Hz of the RR interval are calculated by estimating the area under the power spectral density in those bands using the trapezoidal method. The LF band is primarily modulated by both the sympathetic and parasympathetic nervous systems. On the other hand, the HF band is modulated by the parasympathetic nervous system and mainly through respiratory sinus arrhythmia [40]. Once the power of each of the two bands was calculated (Figure 4.2B), we estimated the normalized high-frequency component (HFnorm) by dividing the high-frequency power by the total power in the two bands HF/(LF+HF) (Figure 4.2C). The HFnorm reflects the proportion of parasympathetic to sympathetic activity and is used as a marker of sympathovagal balance

[41]. A higher HFnorm value indicates a shift of the ANS towards a parasympathetic state.

4.2.4 Experimental Design and Statistical Analysis

Correlation Analyses

For each patient, we calculated the Spearman correlation of the high gamma band activity estimate for each bipolar pair in each NREM sleep stage with the normalized high-frequency power of the RR-interval. We then pooled together all of the R^2 values for all channels in its defined region of interest (ROI). To determine if there is an overall patient-wide correlation between the high gamma band and HFnorm in each ROI/sleep stage, we applied a one-sample Wilcoxon signed-rank test to evaluate the null hypotheses that the median of the pooled R^2 values in each ROI and sleep stage does not deviate from zero. Additionally, for each bipolar pair, we calculated the partial Spearman correlation between the HGnorm and the HFnorm conditioned on the delta band activity (partialcorr function in MATLAB). We reapplied the one-sample Wilcoxon signed-rank test to determine if the pooled partial correlation R^2 values deviate from zero for each ROI/sleep stage. This experimental design investigates if the overall correlation between HG and HRV is mediated by the delta band since studies have shown that delta band upstates and downstates modulate high gamma band activity in humans during sleep [42, 43].

Bootstrapped Coupling Analysis

Additionally, for each sleep stage, we calculated the proportion of statistically significant channels in each ROI to determine the overall coupling between the ANS-CNS. This is done with a bootstrapped analysis by randomly subsampling the HGnorm from each bipolar channel-pair and the HFnorm with replacement and then repeating the Spearman correlation analyses described in the previous section. We set the number of samples in each correlation analysis to be equal (n=50) to ensure that an uneven number of samples did not influence the correlation p-values. We repeated the correlation analyses for a total of N=1000 iterations, and in each iteration, we calculated the percentage of statistically correlated bipolar

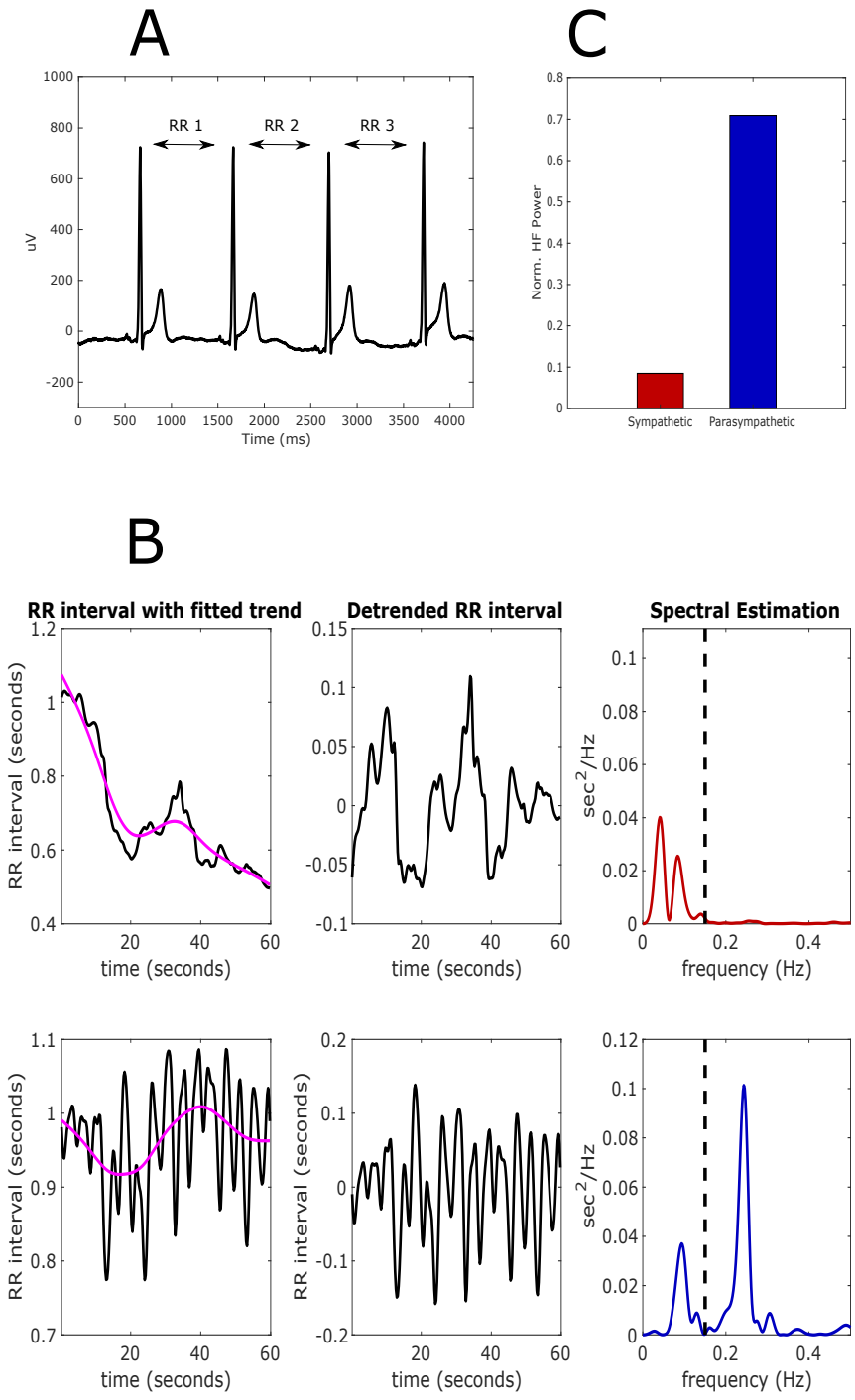


Figure 4.2: **Heart rate variability calculations.** A) Sample 4 second ECG signal; the RR interval is determined from the time difference between subsequent R-peaks. B) Example calculations of HFnorm for two one-minute interval, one with overall sympathetic tone (top row) and one with overall parasympathetic tone (bottom). The very low frequency trend (magenta line) is removed from the raw RR intervals (left-most plots), resulting in the detrended time-series (middle plots). Power spectral density shows most power is 'high-frequency' (0.14-0.4Hz) in the parasympathetic interval (right-most plots) C) Calculated high frequency normalized values for the two intervals displayed in figure 4.2B

channel-pairs in each ROI and sleep stage separately (FDR corrected with α=0.05). This generated a distribution of ANS-CNS coupling as a function of the percentage of channels that have a statistically significant correlation in each ROI and sleep stage combination (for both the high gamma band correlation and the partial correlation mediated by delta band).

Statistical Analysis

For each of the statistical tests, we used an FDR correction with α =0.05 [44]. The FDR correction procedure was implemented as follows: for a total of N hypotheses, each with corresponding p values $P_i(i \subseteq N)$, which are sorted in ascending order to identify P_k (k being the largest i for which $P_i \leq (i/N) * \alpha$), all hypotheses with p values less than or equal to Pk would be rejected. We used the Wilcoxon signed-rank test to determine if the overall correlation in each ROI/Sleep stage deviated from zero.In the bootstrapped coupling analysis, the FDR correction to determine the percentage of statistically significant channel-pairs in each iteration was applied on each ROI separately, therefore, the number of comparisons is the number of channel-pairs in each ROI. The FDR correction for the Wilcoxon sign-rank tests was applied for each correlation measure separately for a total of 36 comparisons per analysis (12 regions, 3 sleep stages).

To investigate whether the overall patient-wise correlation in each ROI is affected by the sleep stage, we applied an N-way ANOVA using the ROI and sleep stage as the grouping variable and the Spearman correlation coefficient as the response variable. Finally, to determine whether the overall patient wise correlation is due to patient-specific variation, we fit and compared linear mixed-effects models. Specifically, we compared an LME model that incorporated the ROI and Sleep stage as predictor variables with a simplified model that only used the patient ID as a predictor variable. In the bootstrapped coupling analysis, we applied a one-way ANOVA using the mean of the percentage of correlated channels as the response variable and sleep stage as the grouping variable to determine if sleep stage has a significant effect on the ANS-CNS coupling percentage.

Finally, we tested whether the direction of the correlation between HGnorm across hippocampus and cortex and HFnorm is consistent in sleep stage transitions (agnostic of

location). We first selected the channels where HGnorm and HFnorm were determined to be statistically correlated, as separately tested in each of the sleep stages. Then, we determined the number of channels that have the same direction of correlation across all sleep stages, i.e., positive correlations in N1, N2, and N3, or negative correlations in N1, N2, and N3. In this analysis, if we assume that each channel's HGnorm has an equal probability of being either positive or negatively correlated with HFnorm, then the probability that that channel has the same correlation direction for all sleep stages is 0.25. We then applied a binomial test with the probability of success being 0.25 to determine whether a statistically correlated channel is more likely to have the same direction of correlation across all sleep stages or not.

4.3 Results

4.3.1 Summary of Data

A total of 368 stereo-electroencephalography (sEEG) sites across the 15 patients were included for further analysis (for demographic and clinical information, see Table 4.1). Figure 4.1a shows the distribution of cortical penetration sites of the sEEG shanks. There is broad coverage across multiple cortical sites in both left and right hemispheres. Figure 4.1b shows the mapping of neocortical regions of interest that we apply to cluster the sEEG site locations for further analyses. Anterior and posterior hippocampus, as well and lateral temporal and parietal lobe coverages are present in the most number of subjects (<11). Also, there is coverage in cingulate and insular cortices, as well as lateral-occipital, paracentral, prefrontal, medial occipito-parietal, and medial temporo-occipital lobes. For each region of interest, we have at least 5 subjects with coverage in that region of interest. Across multiple days we collected on average (mean $\pm$ std) 170 ± 98 minutes for N1, 526 ± 295 for N2, and 328 ± 215 minutes for N3, for a total of an average of 1024 minutes ($\sim$17 hours) of sleep data per patient.

Table 4.1: **Patient and recording information.**

						Number of 1-min Samples		
Patient	Age (mean±std)	Sex	Handedness	Number of Channels	Number of Sleep periods	N1	N2	N3
CC04	20	M	Right	22	3	141	178	175
CC08	58	F	Right	30	3	164	998	88
CC15	42	M	Left	29	4	159	148	497
CC18	18	F	Left	30	3	142	197	246
CC20	22	M	Left/Right	22	3	150	297	336
CC23	40	F	Right	16	5	154	588	879
CC24	43	F	Right	18	3	202	219	223
CC25	16	M	Right	20	5	144	504	536
CC26	32	F	Right	39	3	91	507	197
CC30	36	M	Left	28	4	162	820	420
CC31	21	F	Left	23	3	105	922	121
CC39	21	F	Right	23	8	472	805	577
CC49	29	F	Right	23	4	279	351	226
CC60	24	M	Right	24	3	56	492	170
CC69	31	F	Right	21	4	120	886	234
Mean ± Std	30.2 ± 11.8			22.3 ± 5.2	3.6 ± 1.5	170 ± 98	526 ± 295	328 ± 215

4.3.2 Overall correlation effect present in multiple sites across different NREM sleep stages

Figures 4.3A-C show the distribution of R^2 values derived from the Spearman correlation between the 1-minute averaged high gamma band activity estimate (HGnorm) with the normalized high-frequency power of the RR intervals (HFnorm). It is evident that there is a wide range of effects for each region of interest across all NREM sleep stages as the R^2 values range from -0.55 to 0.54. However, it is also evident that in some ROIs, the distribution of R^2 values is consistently either greater than or less than zero. Positive correlations suggest that increased neural activity is associated in that contact with an overall shift of the ANS towards a parasympathetic state, and conversely, negative correlations suggest that increased neural activity is associated in that contact with an overall shift of the ANS towards sympathetic state.

The overall inter-patient trends are summarized in Figures 4.3D-F. Regions of interests for each sleep stage that have an overall median R^2 that is statistically different from zero are indicated with an asterisk. The means and p-values generated from the Wilcoxon signed-rank test (to test whether the mean is statistically different from zero) for each ROI/NREM combination are listed in Table 4.2. Both the anterior and posterior hippocampus show an overall positive trend in all NREM stages..Additionally, there is an overall statistically significant positive trend in lateral temporal, lateral parietal, orbitofrontal and prefrontal

lobes in N1, orbitofrontal lobe in N2, and prefrontal lobe in N3. Higher neural activity in these regions is directly correlated with a shift to a parasympathetic autonomic state, with each region showing a profound effect according to the sleep stage. Finally, the lateral temporal lobe shows a negative trend in N3 that is statistically significant ($R^2 = 0.14$). Overall, of the 12 regions and 3 states, 14 means were significantly different from zero, 12 in the positive direction, and 2 in the negative, indicating that greater neural activity is more often associated with parasympathetic tone (binomial test, 2-tailed, p=0.0018).

Furthermore, we applied a 2-way ANOVA to determine if the correlation between neural activation and autonomic balance shows a main effect of anatomical location, sleep stage, or their interaction. The ANOVA analysis shows that there is a main effect of both anatomical location (F=9.22, p=0) and sleep stage (F=4.01 p=0.018) on the correlations, as well as the interaction between location and sleep stage (F=2.01 p=0.003). Finally, to take into account patient wise variability, we fitted two linear mixed-effects models on the data. Firstly, we fitted a simplified model that only takes into account the patient's ID as a random effect, Corr $\sim$ 1 +(1|Patient-ID). This model would be appropriate if the patient-wise variability is the main factor that contributes towards the correlation values. The model we compare it to is a more complex model that takes into account anatomical location and sleep stage and their interaction as fixed effects and the patient ID as the random effect, Corr $\sim$ 1+ AL*SS+(1|Patient-ID). Using a theoretical likelihood ratio test, we found that indeed the more complex model is a better fit to the data that the simplified model (p-value=0). Along with the 2-way ANOVA analysis, this is strong evidence that there is a clear effect of the anatomical location, sleep stage, and their interaction on the overall correlation between the high gamma band activity and the HFnorm.

Next, we investigated whether the direction of correlation of correlated channels is consistent across sleep stages. 30 channels were found to be statistically correlated across all sleep stages, and 28 of those 30 channels have the same direction of correlation across all sleep stages. We then applied a binomial test to determine whether a statistically correlated channel is more likely to have the same direction of correlation across all sleep stages. In the binomial test, the probability of success (success being the direction of correlation across all

Table 4.2: **Average correlation of high gamma and heart rate variability within region.** The correlation (R^2) was calculated between the normalized high gamma and heart rate variability measures (HGnorm and HFnorm) within each sleep stage, and averaged across all channels within each ROI. The Wilcoxon sign rank test was used to determine if the mean of the distribution of the R^2 values differs from zero. Bold and underligned numbers are statistically significant (FDR corrected p<0.05)

Location	Number of Bipolar Channels	Number of Patients	Sleep Stage 1		Sleep Stage 2		Sleep Stage 3	
			Mean R^2	p-value	Mean R^2	p-value	Mean R^2	p-value
Anterior Hippocampus	48	12	0.14	0	0.1	0	0.15	0
Posterior Hippocampus	39	11	0.058	0.01	0.06	0.0072	0.1	0.0016
Cingulate	8	6	0.0031	0.92	-0.0026	0.96	0.08	0.21
Insula	5	4	0.042	0.56	-0.048	0.62	-0.17	0.33
Lateral Temporal	46	13	0.078	0.0067	-0.058	0.0083	-0.13	0
Lateral Occipital	10	6	0.029	0.38	-0.031	0.52	0.04	0.63
Lateral Parietal	52	14	0.082	0.0001	-0.019	0.41	0.032	0.22
Medial Occipito-Parietal	28	11	-0.0136	0.62	0.011	0.97	0.065	0.032
Medial Temporo-occipital	24	10	0.018	0.59	-0.031	0.23	-0.052	0.022
Orbitofrontal	13	5	0.089	0.011	0.1	0.0003	0.04	0.54
Paracentral	19	9	-0.014	0.73	-0.039	0.32	0.0011	0.98
Prefrontal	38	8	0.06	0.036	0.0009	0.15	0.066	0.005

sleep stages is the same) is 0.25, the number of trials is 30, and the number of successful trials is 28. We reject the null hypothesis that the probability of success is 0.25 (p=0), meaning that a correlated channel is more likely to have the same direction of correlation across all sleep stages. Thus, although the correlation between HGnorm and HGnorm could be either positive or negative, in 93% of the channels the direction was consistent across stages.

4.3.3 Correlation of high gamma band with the autonomic response is related to delta band activity.

NREM stages N2 and N3 are characterized by high levels of delta band activity, which powerfully modulates high gamma activity in humans [42, 43], raising the possibility that the correlations found between high gamma band and HFnorm are mediated by variations in delta activity within each sleep stage. Figures 4.4 shows the partial correlation of the high gamma band activity estimate with HFnorm conditioned on the delta band activity estimate for each anatomical location and sleep stage combination. Overall, the number of statistically significant locations was lower (from n=14 to n=10, total across N1, N2 and N3) when removing the effect of delta band activity variation. A key difference is a decrease in anterior hippocampal correlations when the effects of delta band activity were removed in N1 and N3. The other sites becoming statistically significant or insignificant had only very small increases or decreases in the absolute value of the average R^2. We conclude that the overall

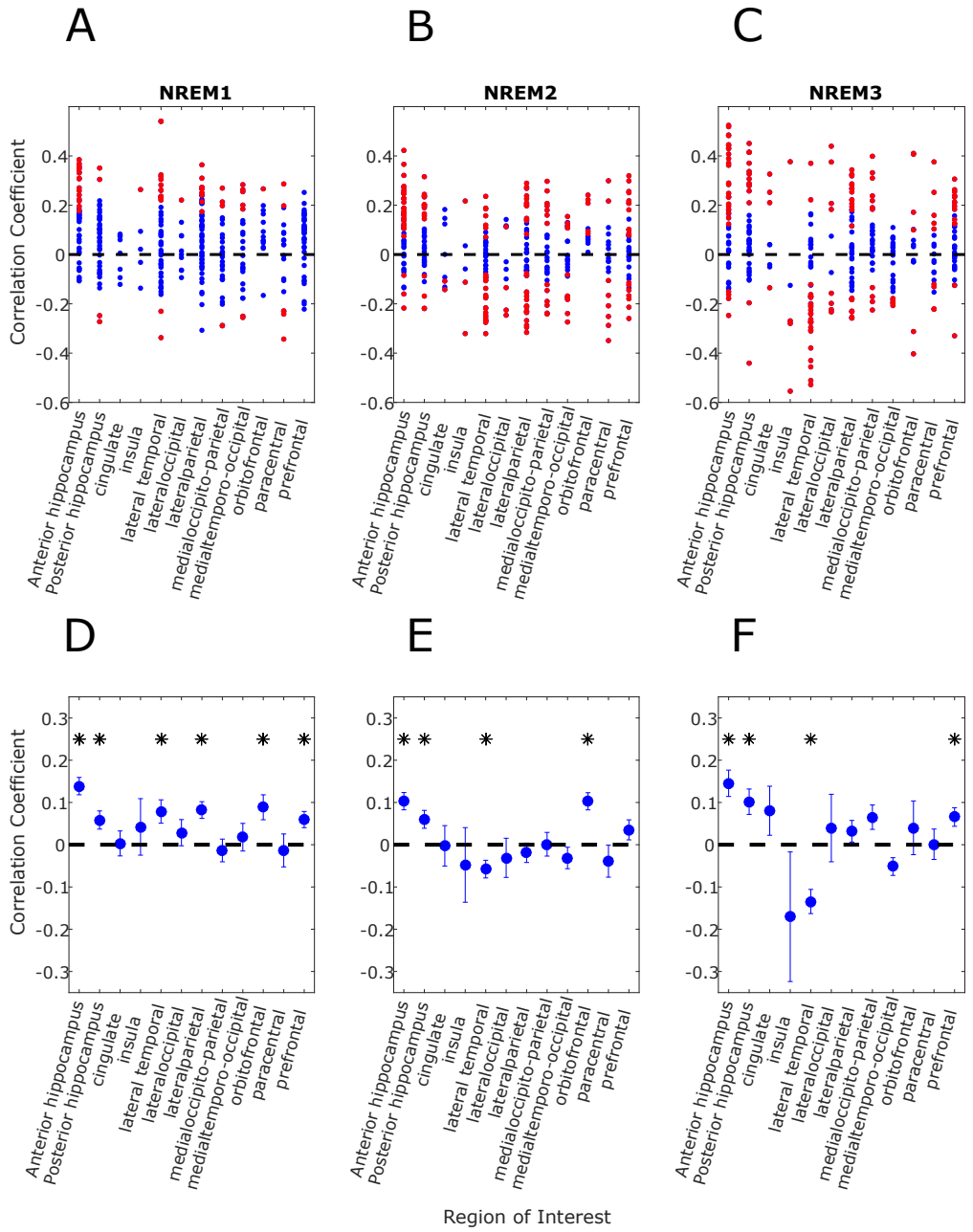


Figure 4.3: **Correlation of population neural activity (HGnorm) and estimated parasympathetic balance (HFnorm).** The distribution of R^2 values for each ROI in each sleep stage is displayed in the topmost plots (A-C), The red dots are statistically significant channel pairs whereas the blue dots are not (FDR corrected $p<0.05$). The mean and standard error of the R^2 values are displayed in panels D-F. Regions where the mean R^2 is significantly differ from zero are labeled with an asterisk.

pattern of correlations in the cortex is little affected by variations in delta band activity.

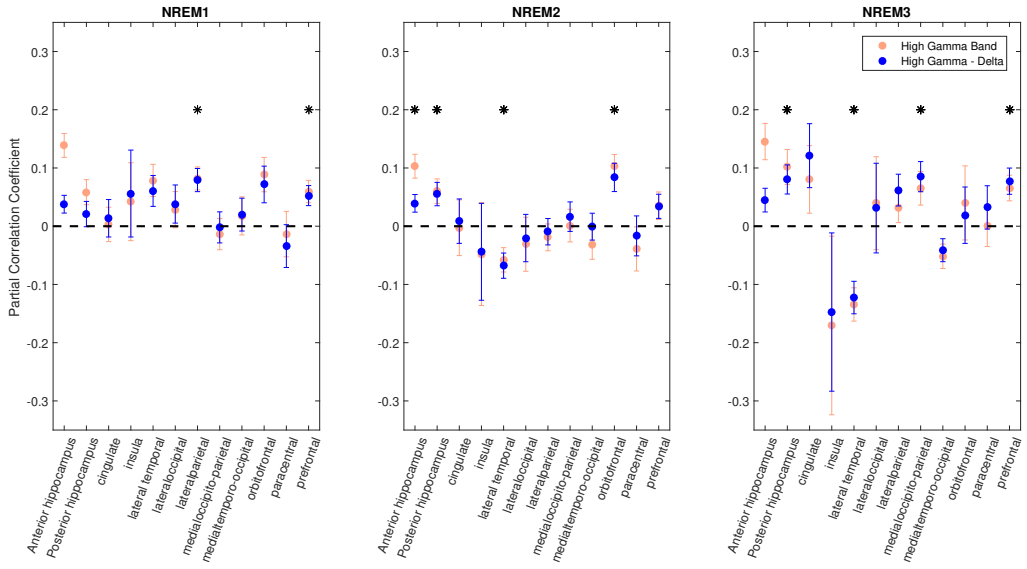


Figure 4.4: **Correlation of population neural activity and autonomic state after conditioning on delta.** Similarly to Figures 4.3D-F, the mean and standard error of the R^2 values calculated from the correlation between HGnorm and HFnorm, but after conditioning on delta band activity (blue). The unconditioned mean and standard errors shown in Figures 4.3D-F are superimposed in orange to facilitate their comparison.

4.3.4 Cortico-hippocampal-autonomic coupling generally increases during deep sleep.

In Figures 3A-C, as for almost all regions and all sleep stages, it is evident that within each anatomical location, some channel pairs across patients show a negative correlation while others show a positive correlation. Therefore, it is important to not only investigate whether the overall patient wide correlation was statistically different from zero, but also the strength of the CNS-ANS coupling measured by the percentage of statistically correlated channel pairs in each anatomical location and sleep stage.

Firstly, a key finding in Figure 4.5 is that for many anatomical locations, the overall percentage of statistically correlated channel pairs increases in N3 as compared to N1 and N2, indicating that there is a stronger CNS-ANS coupling in deeper sleep. This was verified by applying a one-way ANOVA using the mean of the percentage of correlated channels as the response variable and sleep stage as the grouping variable. Sleep stage has a significant

effect on ANS-CNS coupling (F=6.24, p=0.005). This phenomenon is mainly observed in both anterior and posterior hippocampus, insula, lateral temporal, lateral occipital and orbitofrontal lobes, with the insula having the strongest coupling in N3 at 53.1% of channels (bootstrapped) having a statistically significant correlation, even though the insula did not show a statistically significant positive or negative trend in the previous analyses. This indicates that within the same region of the brain, the coupling between the autonomic system and neural activity could be in opposite directions.

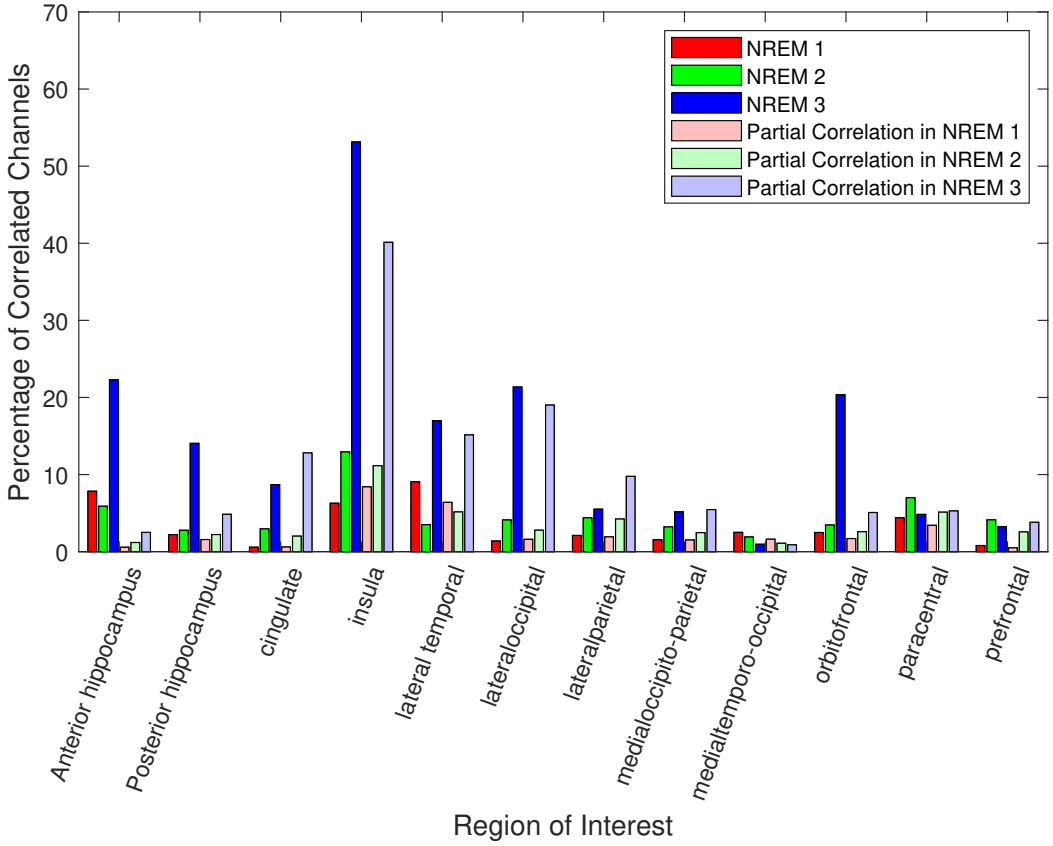


Figure 4.5: **Distribution of responsive channels across locations and sleep stages.** The percentage of channels where the correlation of HGnorm and HFnorm is statistically significant in each ROI are plotted for each sleep stage, with conditioning on delta activity (light colors) and without (dark colors). Percentage values for each ROI/sleep stage/correlation measure are shown in Supplementary Table 4.3

Figure 4.5 also displays the percentage of partially correlated (delta band conditioned) channel pairs in the lighter colors. Both the anterior and posterior hippocampal-ANS coupling after removing the effect of delta band dropped significantly, most notably in N3, from 22.3%

to 2.5% and 14.0% to 4.8% respectively. The percentage of correlated channels in insula and orbitofrontal dropped as well in N3, from 53.1% to 40.1% and 20.4% to 5.1% respectively. Lateral temporal and lateral occipital lobes show little to no change in ANS coupling when mediating for the delta band effect. Additionally, the cingulate, as well as the lateral parietal lobe, show an increase in ANS coupling in N3 when applying the partial correlation. We reapplied the one-way ANOVA analysis on the percentage of partially correlated channels and found that sleep stage still has a significant effect on the ANS-CNS coupling even after delta band mediation ($F=5.06$, $p=0.0121$). The values of the mean correlation and partial correlation coupling percentages for all ROI's, sleep stages, and correlation measures are shown in Supplementary Table 4.3.

4.4 Discussion

In this study, we have collected ECG and neural sEEG recordings from 15 subjects during sleep and quantified the relationship in terms of linear correlations and coupling strengths between high gamma band activity in 12 distinct cortical and hippocampal locations and autonomic tone as indexed by the normalized high-frequency component of the heart rate. To our knowledge, this is the first attempt to understand the brain-heart interaction during sleep using intracranial recordings. We have found a significant correlation between the high gamma and autonomic tone in distinct anatomical locations across multiple patients that differ according to the sleep stage. In some locations (e.g., the anterior hippocampus in all sleep stages as well as the orbitofrontal cortex in N2) these correlations were positive, indicating increased activity during a higher parasympathetic tone. In others (e.g., the lateral temporal lobe in N3), the correlation was negative, suggesting higher cortical activity during a greater sympathetic tone. However, the direction of the correlation varied across the electrodes in each structure, indicating a significant but variegated response. Overall, predominantly positive correlations were more common, and a global response was evident as a function of sleep stage. The coupling indicated by these correlations could serve many roles (Figure 4.6): (A) modulation by brainstem autonomic efferents of both cortical/hippocampal tone and HRV, as

part of an overall modulation of all internal organs, including the brain and heart ; (B) control by cortical/hippocampal regions of brainstem autonomic centers which project to the heart ; and/or (C) viscerosensory responses by cortical/hippocampal regions to changes in the internal milieu . Indeed, there is ample anatomical and physiological evidence for all three functional relationships, and they may all be reflected in the current results. Overall, our results demonstrate and characterize a tight coupling between visceral and cortical/hippocampal state. We did not find a significant effect when investigating the correlation between HGnorm and HFnorm using functional networks as labels. Previous studies have theorized that the default mode network (DMN) is a major regulator of parasympathic tone [3]. However, we found that neural activity in the DMN is only correlated with parasympathetic modulation in N1, and not in the deeper sleep stages (Supplementary Figure 4.7). This raises the question whether the regulation of the DMN of parasympathetic tone is mainly a property of waking and light sleep. Furthermore, similar to a previous finding [9], where it is shown that the correlation between neural firing rate and the cardiac cycle length could be positive or negative for different units in medial structures of humans during waking state, we found both positive and negative correlation coefficients for different channel pairs in these same structures during sleep (Figure 4.3 A-C).

4.4.1 Role of Hippocampus

The hippocampus has been shown to modulate sympathetic contributions in the central autonomic network during waking fMRI studies [3]. However, little is known about the role of the hippocampus in autonomic modulation in sleep. We have shown that, during sleep, the high gamma band activity in the anterior hippocampus has a stronger correlation with parasympathetic modulation as compared to the posterior hippocampus. This could be because the anterior hippocampus has stronger connections with areas of the brain that is known to modulate the autonomic system as compared to the posterior hippocampus, with direct connections found in anterior hippocampus with amygdalar nuclei [45, 20, 46]. Most interestingly, this overall effect in the anterior hippocampus is significantly reduced when conditioning the correlation of high gamma and autonomic tone on delta band activity

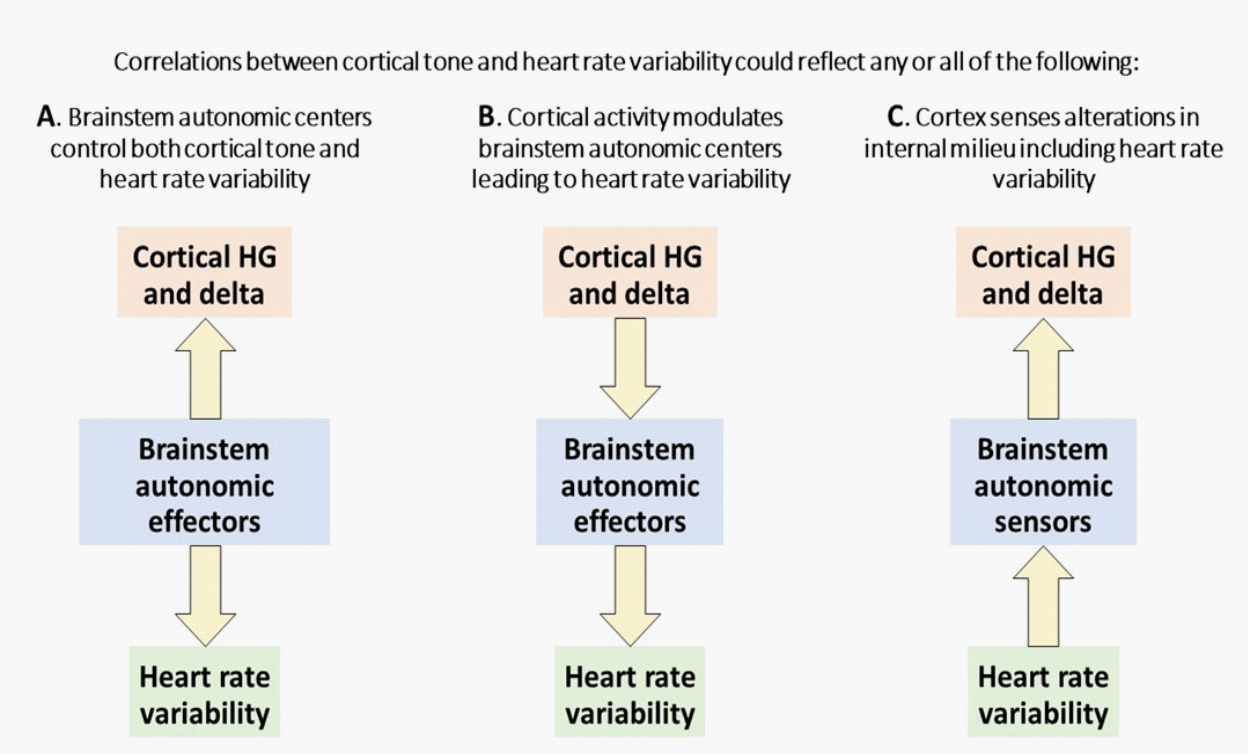


Figure 4.6: **Possible functional relationships underlying the correlation between cortical and hippocampal activity and heart rate variability.**

in all sleep stages (Figure 4.4). The posterior hippocampus's overall effect is less affected when conditioning on delta band activity. The overall coupling strength, as a percentage of statistically correlated channels, is significantly reduced in both the anterior and posterior hippocampus, as shown in Figure 4.5. Previous studies have shown that delta band modulates activity in high gamma band during sleep [42, 43], and we have shown in this work that this modulation affects both hippocampal-ANS correlation and coupling strength. Further analysis is needed to understand the directionality and time scale of the hippocampal-ANS connection during sleep.

4.4.2 Cortical Interactions

The central autonomic network in humans has been shown to consist of several cortical regions such as the anterior cingulate cortex (ACC), ventromedial prefrontal cortex, amygdala, and insula during waking studies [3]. The insula, in particular, has been studied extensively in its involvement in autonomic arousal [47, 48, 49, 50, 51], and is also correlated with interocep-

tive awareness [52]. Even though Figure 4.3 D-F did not show a statistically significant overall trend in the correlation between insula high gamma activation and autonomic modulation, we did observe that the overall coupling in the insula, most notably in N3, is substantially higher than all other regions (with 5 channel pairs collected across 4 patients), and is not heavily suppressed by removing the effect of delta band activity. This supports the claim that the insula is a major hub in the central autonomic network (CAN), especially during deep sleep. Figure 4.3 A-C shows that within insula (like other areas) bipolar channel-pairs could have either sympathetic and parasympathetic coupling with high gamma activity. Clearly, further work is required to understand how different areas within each subregion of the cortex are related to autonomic tone.

Many other cortical sites have been established to modulate autonomic tone, such as the orbitofrontal cortex [53], lateral temporal lobe [3], and prefrontal lobe [5]. We have shown that these areas are heavily involved in autonomic modulation during sleep and that their effect is affected by the sleep stage. We have also shown that both the lateral parietal and medial occipito-parietal lobes could be considered part of the CAN during sleep, as they have an overall effect observed in N1 and N3, respectively. The CAN has also been established to include the cingulate cortex [54, 55], yet we did not observe an overall statistically significant correlation in any sleep stage, whether mediating for delta band activity or not (Figures 4.3 and 4.4). Figure 4.5 shows that the coupling between cingulate cortex high gamma activity and autonomic tone is more pronounced after removing the effect of the delta band, yet overall, the percentage of correlated channels is low in comparison to other anatomical regions (12% when removing delta band effect).

4.4.3 Future Work

We hope that this paper serves as a starting point for future researchers to understand CNS-ANS interaction during sleep. Since our method has low temporal resolution (in the order of minutes), as well as limited cortical coverage, it is difficult to infer directionality. Therefore, further analysis is needed to understand the dynamics of the interaction between cortical or hippocampal structures and the ANS by implementing measures of autonomic modulations

that operate in the order of seconds or hundreds of milliseconds, such as a transient increase in heart rate, as well as utilizating other measures such as galvanic skin response. Having a higher temporal resolution will enhance our understanding of what processes are generating the correlation and coupling between neural activity and heart rate variability, as shown in Figure 4.6. In this study, we focused our analyses mainly on sleep periods, due to the lack of any potentially confounding behavioral or cognitive activity that could influence autonomic tone, such as moving or being in an anxious state. It would be important to extend these findings to the waking state using long-term intracranial recordings. This would lead to a direct comparison to fMRI studies that leveraged either short-term resting state analyses or a trial-based experimental approach. Through understanding the interaction between the brain and the autonomic state, therapeutic modalities such as deep brain stimulation could be enhanced. Autonomic conditions that could be potentially managed through DBS include hypertension, asthma, and obstructive sleep apnea [56]. Better understanding of CAN would facilitate the modulation of blood pressure, respiration, and heart rate by stimulating areas in the brain that are part of the CAN. To our knowledge, this is the first work to investigate the autonomic correlates to high gamma activity using sEEG electrodes during sleep. Further research is required to understand the directionality of the CNS-ANS coupling as well as a more nuanced analysis of multiple structures, such as the insula.

4.5 Acknowledgements

This chapter is a reprint of material currently being prepared for submission in Alasfour, Abdulwahab, Xi Jiang, Jorge Gonzalez-Martinez, Vikash Gilja, and Eric Halgren. “High gamma activity in cortex and hippocampus is correlated with autonomic tone during sleep." We would like to thank Burke Rosen for his insight and providing us with functional network labels for the sEEG channels used in this work.

4.6 Supplementary Data

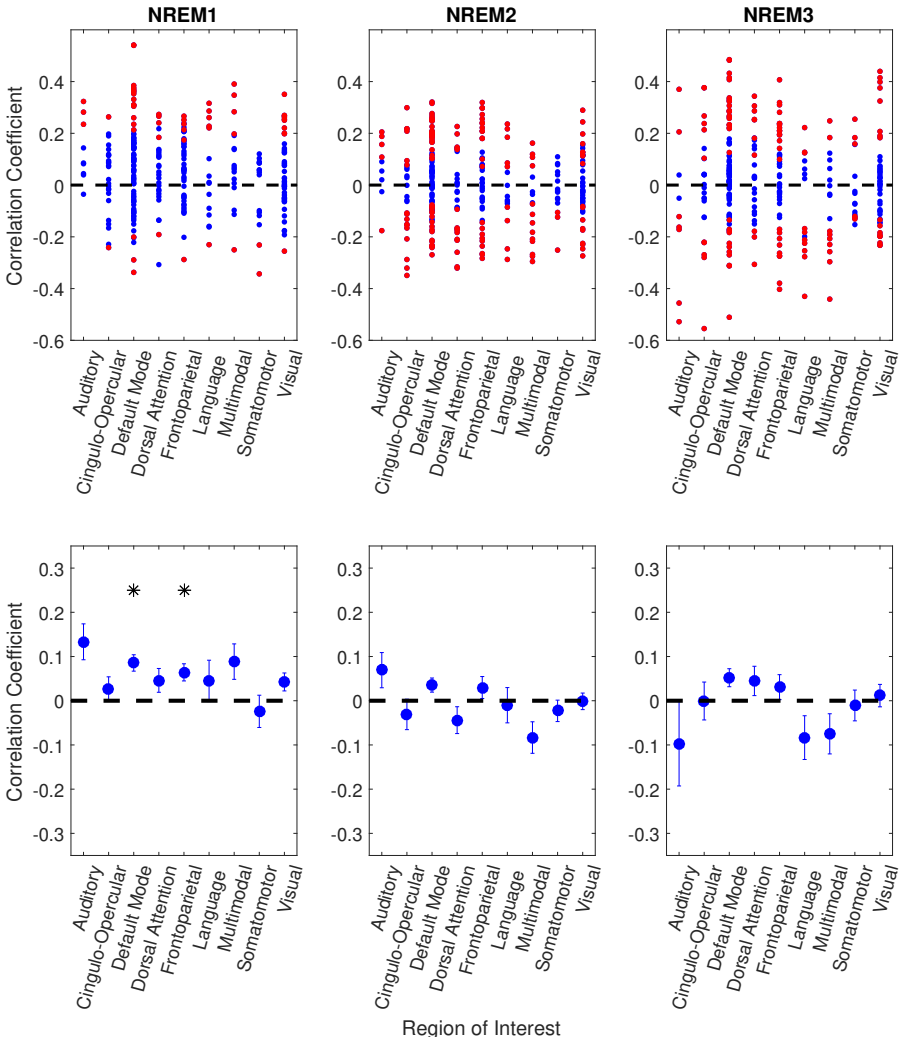


Figure 4.7: **Spearman correlation analysis of HGnorm and HFnorm in different functional networks.** The distribution of R^2 values for each functional network in each sleep stage is displayed in the topmost plots; the red dots are statistically significant channel pairs, whereas the blue dots are not (FDR corrected $p<0.05$). The mean and standard error of the R^2 values are displayed in the bottom three panels. Regions where the mean R^2 significantly differ from zero are labeled with an asterisk.

Table 4.3: **Mean correlation and partial correlation coupling percentages.**

	Mean % of statistically correlated channel pairs.			Mean % of statistically partially correlated channel pairs.		
Location	N1	N2	N3	N1	N2	N3
Anterior Hippocampus	7.8	5.9	22.3	0.6	1.2	2.5
Posterior Hippocampus	2.2	2.8	14	1.6	2.2	4.8
Cingulate	0.6	3	8.7	0.6	2	12.8
Insula	6.3	12.9	53.1	8.4	11.1	40.1
Lateral Temporal	9.1	3.5	17	6.4	5.5	15.1
Lateral Occipital	1.4	4.1	21.4	1.6	2.8	19
Lateral Parietal	2.1	4.4	5.5	1.9	4.2	9.8
Medial Occipito-Parietal	1.5	3.2	5.2	1.5	2.4	5.5
Medial Temporo-occipital	2.5	1.9	1	1.6	1.1	0.9
Orbitofrontal	2.5	3.5	20.4	1.7	2.5	5.1
Paracentral	4.4	7	4.8	3.4	5.1	5.2
Prefrontal	0.8	4.1	3.3	0.5	2.6	3.8

References

[1] Jacob Reimer, Matthew J. McGinley, Yang Liu, Charles Rodenkirch, Qi Wang, David A. McCormick, and Andreas S. Tolias. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, 7, November 2016. Publisher: Nature Publishing Group.

[2] Siddhartha Joshi, Yin Li, Rishi M. Kalwani, and Joshua I. Gold. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, 89(1):221–234, January 2016.

[3] F. Beissner, K. Meissner, K.-J. Bar, and V. Napadow. The Autonomic Brain: An Activation Likelihood Estimation Meta-Analysis for Central Processing of Autonomic Function. *Journal of Neuroscience*, 33(25):10503–10511, June 2013. Publisher: Society for Neuroscience.

[4] Andrew R. Dykstra, Alexander M. Chan, Brian T. Quinn, Rodrigo Zepeda, Corey J. Keller, Justine Cormier, Joseph R. Madsen, Emad N. Eskandar, and Sydney S. Cash. Individualized localization and cortical surface-based registration of intracranial electrodes. *NeuroImage*, 59(4):3563–3570, February 2012. Publisher: Academic Press.

[5] Julian F. Thayer, Fredrik Ahs, Mats Fredrikson, John J. Sollers, and Tor D. Wager. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience and Biobehavioral Reviews*, 36(2):747–756, February 2012.

[6] Paul M. Macey, Jennifer A. Ogren, Rajesh Kumar, and Ronald M. Harper. Functional Imaging of Autonomic Regulation: Methods and Key Findings. *Frontiers in Neuroscience*, 9, 2016. Publisher: Frontiers.

[7] Miriam Sklerov, Eran Dayan, and Nina Browner. Functional neuroimaging of the central autonomic network: recent developments and clinical implications. *Clinical Autonomic Research*, 29(6):555–566, December 2019.

[8] Pinar Senay Özbay, Catie Chang, Dante Picchioni, Hendrik Mandelkow, Miranda Grace Chappel-Farley, Peter van Gelderen, Jacco Adrianus de Zwart, and Jeff Duyn. Sympathetic activity contributes to the fMRI signal. *Communications Biology*, 2(1):421, November 2019.

[9] Kayeon Kim, Josef Ladenbauer, Mariana Babo-Rebelo, Anne Buot, Katia Lehongre, Claude Adam, Dominique Hasboun, Virginie Lambrecq, Vincent Navarro, Srdjan Ostojic, and Catherine Tallon-Baudry. Resting-state neural firing rate is linked to cardiac-cycle duration in the human cingulate and parahippocampal cortices. *Journal of Neuroscience*, 39(19):3676–3686, May 2019. Publisher: Society for Neuroscience.

[10] Massimiliano de Zambotti, Adrian R. Willoughby, Peter L. Franzen, Duncan B. Clark, Fiona C. Baker, and Ian M. Colrain. K-Complexes: Interaction between the Central and Autonomic Nervous Systems during Sleep. *Sleep*, 39(5):1129–1137, April 2016. Publisher: Oxford University Press (OUP).

[11] Massimiliano de Zambotti, John Trinder, Alessandro Silvani, Ian Colrain, and Fiona C. Baker. Dynamic coupling between the central and autonomic nervous systems during sleep: a review. *Neuroscience and biobehavioral reviews*, 90:84–103, July 2018.

[12] Roy Mukamel, Hagar Gelbard, Amos Arieli, Uri Hasson, Itzhak Fried, and Rafael Malach. Neuroscience: Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science*, 309(5736):951–954, August 2005.

[13] Kai J Miller, Christopher J Honey, Dora Hermes, Rajesh PN Rao, Marcel denNijs, and Jeffrey G Ojemann. Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *New Horizons for Neural Oscillations*, 85:711–720, January 2014.

[14] Jeremy R. Manning, Joshua Jacobs, Itzhak Fried, and Michael J. Kahana. Broadband Shifts in Local Field Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans. *Journal of Neuroscience*, 29(43):13613–13620, October 2009. Publisher: Society for Neuroscience Section: Articles.

[15] Supratim Ray and John H. R. Maunsell. Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLOS Biology*, 9(4):e1000610, April 2011. Publisher: Public Library of Science.

[16] J. Trinder, J. Kleiman, M. Carrington, S. Smith, S. Breen, N. Tan, and Y. Kim. Autonomic activity during human sleep as a function of time and sleep stage. *Journal of Sleep Research*, 10(4):253–264, 2001.

[17] A. Monti, C. Medigue, H. Nedelcoux, and P. Escourrou. Autonomic control of the cardiovascular system during sleep in normal subjects. *European Journal of Applied Physiology*, 87(2):174–181, June 2002.

[18] Jorge Gonzalez-Martinez, Juan Bulacio, Andreas Alexopoulos, Lara Jehi, William Bingaman, and Imad Najm. Stereoelectroencephalography in the "difficult to localize" refractory focal epilepsy: Early experience from a North American epilepsy center. *Epilepsia*, 54(2):323–330, February 2013.

[19] Song Lin Ding and Gary W. Van Hoesen. Organization and detailed parcellation of human hippocampal head and body regions based on a combined analysis of Cyto- and chemoarchitecture. *Journal of Comparative Neurology*, 523(15):2233–2253, October 2015. Publisher: Wiley-Liss Inc.

[20] Jordan Poppenk, Hallvard R. Evensmoen, Morris Moscovitch, and Lynn Nadel. Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, 17(5):230–240, May 2013. Publisher: Elsevier Current Trends.

[21] Xi Jiang, Jorge Gonzalez-Martinez, Eric Halgren, John Gale, Qianqian Deng, Charles Dickey, Darlene Evardone, Zach Fitzgerald, Chris Gonzalez, Don Hagler, Milan Halgren, Erik Kaestner, Rachel Mak-Mccully, Adam Niese, Burke Rosen, and T G Venti. Posterior hippocampal spindle-ripples phase-locked with parietal spindles during NREM sleep in humans. *Journal of Neuroscience*, 2019.

[22] Rachel A. Mak-Mccully, Matthieu Rolland, Anna Sargsyan, Chris Gonzalez, Michel

Magnin, Patrick Chauvel, Marc Rey, Hélène Bastuji, and Eric Halgren. Coordination of cortical and thalamic activity during non-REM sleep in humans. *Nature Communications*, 8, May 2017. Publisher: Nature Publishing Group.

[23] Anders M. Dale, Bruce Fischl, and Martin I. Sereno. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, 9(2):179–194, February 1999.

[24] B. Fischl, M. I. Sereno, and A. M. Dale. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, February 1999.

[25] Rahul S. Desikan, Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, Anders M. Dale, R. Paul Maguire, Bradley T. Hyman, Marilyn S. Albert, and Ronald J. Killiany. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980, July 2006.

[26] Burke Q. Rosen and Eric Halgren. A whole-cortex probabilistic diffusion tractography connectome. preprint, Neuroscience, June 2020.

[27] Jie Lisa Ji, Marjolein Spronk, Kaustubh Kulkarni, Grega Repovš, Alan Anticevic, and Michael W. Cole. Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*, 185:35–57, January 2019.

[28] Xi Jiang, Jorge Gonzalez-Martinez, Eric Halgren, John Gale, Qianqian Deng, Charles Dickey, Darlene Evardone, Zach Fitzgerald, Chris Gonzalez, Don Hagler, Milan Halgren, Erik Kaestner, Rachel Mak-Mccully, Adam Niese, Burke Rosen, and T G Venti. Coordination of human hippocampal sharpwave-ripples during NREM sleep with cortical theta bursts, spindles, downstates and upstates. *Journal of Neuroscience*, 2019.

[29] B. Fischl, M. I. Sereno, R. B. Tootell, and A. M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284, 1999.

[30] Damien Gervasoni, Shih Chieh Lin, Sidarta Ribeiro, Ernesto S. Soares, Janaina Pantoja, and Miguel A.L. Nicolelis. Global forebrain dynamics predict rat behavioral states and their transitions. *Journal of Neuroscience*, 24(49):11137–11147, December 2004.

[31] Xi Jiang, Isaac Shamie, Werner K. Doyle, Daniel Friedman, Patricia Dugan, Orrin Devinsky, Emad Eskandar, Sydney S. Cash, Thomas Thesen, and Eric Halgren. Replay of large-scale spatio-temporal patterns from waking during subsequent NREM sleep in human cortex. *Scientific Reports*, 7(1), December 2017. Publisher: Nature Publishing Group.

[32] Michael H. Silber, Sonia Ancoli-Israel, Michael H. Bonnet, Sudhansu Chokroverty, Madeleine M. Grigg-Damberger, Max Hirshkowitz, Sheldon Kapen, Sharon A. Keenan, Meir H. Kryger, Thomas Penzel, Mark R. Pressman, and Conrad Iber. The visual scoring of sleep in adults. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine*, 3(2):121–131, March 2007. Place: United States.

[33] Sydney S Cash, Eric Halgren, Nima Dehghani, Andrea O Rossetti, Thomas Thesen, ChunMao Wang, Orrin Devinsky, Ruben Kuzniecky, Werner Doyle, Joseph R Madsen, Edward Bromfield, Loránd Erőss, Péter Halász, George Karmos, Richárd Csercsa, Lucia Wittner, and István Ulbert. The Human K-Complex Represents an Isolated Cortical Down-State. *Science*, 324(5930):1084, May 2009.

[34] Jonathan Winawer, Kendrick N. Kay, Brett L. Foster, Andreas M. Rauschecker, Josef Parvizi, and Brian A. Wandell. Asynchronous broadband signals are the principal source of the bold response in human visual cortex. *Current Biology*, 23(13):1145–1153, July 2013.

[35] Matar Haller, Thomas Donoghue, Erik Peterson, Paroma Varma, Priyadarshini Sebastian, Richard Gao, Torben Noto, Robert T. Knight, Avgusta Shestyuk, and Bradley Voytek. Parameterizing neural power spectra. *bioRxiv*, page 299859, January 2018.

[36] Mika P Tarvainen, Juha-pekka Niskanen, Jukka A Lipponen, Perttu O Ranta-aho, and Pasi A Karjalainen. Kubios HRV – Heart rate variability. *Computer Methods and Programs in Biomedicine*, 113(1):210–220, 2013. Publisher: Elsevier Ireland Ltd.

[37] Jukka A. Lipponen and Mika P. Tarvainen. A robust algorithm for heart rate variability time series artefact correction using novel beat classification. *Journal of Medical Engineering and Technology*, 43(3):173–181, April 2019. Publisher: Taylor and Francis Ltd.

[38] Mika P. Tarvainen, Perttu O. Ranta-aho, and Pasi A. Karjalainen. An advanced detrending method with application to HRV analysis. *IEEE Transactions on Biomedical Engineering*, 49(2):172–175, 2002.

[39] Gary G. Bernston, Bigger Jr. Thomas, Dwain L. Eckberg, Paul Grossman, Peter G. Kaufman, Marek Malik, Haikadi N Magaraja, Stephen W. Porges, J. Philip Saul, Peter H. Stone, and Maurots W. Van Der Molen. Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, 34(6):623–648, November 1997. Publisher: John Wiley & Sons, Ltd.

[40] Fred Shaffer and J. P. Ginsberg. An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, 5, September 2017. Publisher: Frontiers Media SA.

[41] James A. J. Heathers. Everything Hertz: methodological issues in short-term frequency-domain HRV. *Frontiers in Physiology*, 5:177, 2014.

[42] Richárd Csercsa, Balázs Dombovári, Dániel Fabó, Lucia Wittner, Loránd Erss, László Entz, András Sólyom, György Rásonyi, Anna Szcs, Anna Kelemen, Rita Jakus, Vera Juhos, László Grand, Andor Magony, Péter Halász, Tamás F. Freund, Zsófia Maglóczky, Sydney S. Cash, László Papp, György Karmos, Eric Halgren, and István Ulbert. Laminar analysis of slow wave activity in humans. *Brain*, 133(9):2814–2829, 2010. Publisher: Oxford University Press.

[43] Milan Halgren, Daniel Fabó, István Ulbert, Joseph R. Madsen, Lorand Eross, Werner K. Doyle, Orrin Devinsky, Donald Schomer, Sydney S. Cash, and Eric Halgren. Superficial Slow Rhythms Integrate Cortical Processing in Humans. *Scientific Reports*, 8(1),

December 2018. Publisher: Nature Publishing Group.

[44] Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995. _eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1995.tb02031.x.

[45] Charan Ranganath and Maureen Ritchey. Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13(10):713–726, October 2012.

[46] Bryan A. Strange, Menno P. Witter, Ed S. Lein, and Edvard I. Moser. Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, 15(10):655–669, October 2014.

[47] H. D. Critchley, D. R. Corfield, M. P. Chandler, C. J. Mathias, and R. J. Dolan. Cerebral correlates of autonomic cardiovascular arousal: A functional neuroimaging investigation in humans. *Journal of Physiology*, 523(1):259–270, 2000.

[48] H. D. Critchley, R. Elliott, C. J. Mathias, and R. J. Dolan. Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(8):3033–3040, April 2000.

[49] Ronald M. Harper, Richard Bandler, David Spriggs, and Jeffry R. Alger. Lateralized and widespread brain activation during transient blood pressure elevation revealed by magnetic resonance imaging. *Journal of Comparative Neurology*, 417(2):195–204, February 2000.

[50] Hugo D. Critchley. Neural mechanisms of autonomic, affective, and cognitive integration. In *Journal of Comparative Neurology*, volume 493, pages 154–166, December 2005. ISSN: 00219967 Issue: 1.

[51] Edmund T. Rolls. Functions of the anterior insula in taste, autonomic, and related functions. *Food for thought: The functional and neural mechanisms of food perception and choice*, 110:4–19, December 2016.

[52] Hugo D. Critchley, Stefan Wiens, Pia Rotshtein, Arne Öhman, and Raymond J. Dolan. Neural systems supporting interoceptive awareness. *Nature Neuroscience*, 7(2):189–195, February 2004.

[53] Morten L. Kringelbach. The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, 6(9):691–702, September 2005.

[54] Hugo D. Critchley, Christopher J. Mathias, Oliver Josephs, John O'Doherty, Sergio Zanini, Bonnie Kate Dewar, Lisa Cipolotti, Tim Shallice, and Raymond J. Dolan. Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. *Brain*, 126(10):2139–2152, 2003.

[55] Hugo D. Critchley, Joey Tang, Daniel Glaser, Brian Butterworth, and Raymond J. Dolan. Anterior cingulate activity during error and autonomic response. *NeuroImage*, 27(4):885–895, October 2005.

[56] Jonathan A. Hyam, Morten L. Kringelbach, Peter A. Silburn, Tipu Z. Aziz, and Alexander L. Green. The autonomic effects of deep brain stimulation—a therapeutic opportunity. *Nature Reviews Neurology*, 8(7):391–400, July 2012. Number: 7 Publisher: Nature Publishing Group.

Chapter 5

Conclusion

The work presented in Chapters 2-4 serves multiple purposes for future researchers. We hope that in the culmination of this thesis, future researchers are encouraged to investigate naturalistic and unstructured neural activity, gaining insight that could serve both in neural engineering applications and furthering our understanding of brain activity outside the lab. To create neural technologies that can operate in a myriad of behavioral states, we must consider internal and external brain states that could influence behavior. Changes in brain state, both internally and externally driven, affect information processing, decision making, and action [1]. We have shown that neural activity in an unstructured setting has multiple axes of variability that span both the spatial and temporal domains. By the end of Chapter 3, we have shown that abstract behavioral states can be identified by spatiotemporal patterns independent of long term signal statistics, providing evidence that these features are salient to naturalistic behavioral states. Chapter 4 shows that autonomic tone significantly correlates to neural activity during sleep using intracranial electrophysiology. By investigating days worth of recordings from multiple patients, we can reach conclusions regarding the physiology of the connection between the autonomic and central nervous systems. While some of the results match those of previous studies, others do not. Therefore, we implore future researchers to challenge and reproduce our work using electrophysiological recordings.

5.1 Challenges

Tackling neural data collected in a long term setting comes with its own set of challenges. Mainly, the standard neural signal processing pipeline applied to data collected in discrete trials cannot be used in this setting. Therefore, using a single trial or epoch based analyses on unstructured data such as GPFA or LFADS [2, 3] is an avenue to circumvent the lack of trial structure in analyzing the data. Additionally, an extensive amount of neural signal conditioning is necessary before making any conclusions in the data. For example, in Chapter 3, we needed to z-score each 5-minute segment to remove long-term mean or variance shifts that would influence the data's GPFA factorization. If these shifts are not removed, then some spatiotemporal factors will arise due to these slow shifts in neural activity rather than to spatiotemporal activity independent of these shifts.

Additionally, in Chapter 4, an extensive amount of z-scoring was necessary to remove long-term data shifts. We hypothesis that long term shifts in high gamma band power could be due to changes in brain state and/or noise, while shifts in heart rate variability must be due to actual changes in autonomic state since the metric is derived from visually verifiable RR-intervals. Without correcting for these shifts, we would generate correlations that are due to changes in state from one sleep period to another, which would lead us to make dubious conclusions regarding the relationship between the central and autonomic nervous systems. Finally, understanding the behavior of the data in both the temporal and frequency domain is essential. In Chapter 4, we realized that in many of our epochs, high gamma band activity modulation is shown as a narrow peak in the frequency domain rather than a broadband shift. While synchronized high gamma oscillations are a thriving research area, it is a separate mechanism to a broadband shift in high gamma band. Oscillations indicate synchronized activity [4], while a broadband shift in high gamma band correlates to an asynchronous increase in neuronal firing rate [5, 6, 7, 8]. Without checking both sides of the coin, the temporal and frequency sides, we might have missed this confound in the data.

It is imperative to extensively visualize neural activity to understand where the correlations or separability is coming from before drawing conclusions. As engineers, we are

generally attracted to applying novel and interesting algorithms on our datasets and then try to make deductions depending on the output of the used algorithms. Without visualizing the data, it is much more challenging to truly understand the mechanisms and origins of the variability in the neural data. Throughout multiple stages in this dissertation, we have tried to implement different types of algorithms with varying degrees of complexities, from generative models to deep learning frameworks, to extract spatiotemporal representations in neural activity. With explicitly defined models such as tensor decomposition methods or hidden Markov models, we have often reached a roadblock, where the complexity of the algorithm clouds our interpretation of the data. We have also tried to apply deep learning frameworks to determine whether salient behavioral state discriminating spatiotemporal patterns exist in unstructured neural data. The answer to that question at the time was yes, but unfortunately, we were never able to "open the hood" of the neural network black box to understand what these spatiotemporal patterns were. Supervised deep learning methods are useful in finding non-linear aspects in the signal to achieve the desired output, such as classification or regression. However, it is much more challenging to understand the process between input and output. Therefore, in approaching highly variable neural activity, with absolutely no notion of ground truth, we encourage anyone willing to tackle unstructured neural data by applying the most straightforward methods first, such as analyzing basic first and second-order statistics, and move on from there. As you go through the process, slowly but carefully add complexity to the algorithm design. It is vital to have a solid foundation for understanding why each part of the algorithm is essential and what neuroscience theory supports it.

5.2 Future Work

In the intersection of data science, engineering, and neuroscience, many algorithms have been developed to extract meaningful information from neural data, whether generated from individual neuron spikes or aggregate local field potentials. Unsupervised algorithms such as sequential non-negative matrix factorization [9] and tensor decomposition [10] look for

spatiotemporal patterns in neural activity. Dynamic mode activity also looks for spatiotemporal patterns but assumes sinusoidality [11]. Deep learning methods such as LFADS [3] investigate spiking activity's neural population dynamics without trial averaging. Although each of these algorithms has been quite useful in investigating trialized datasets, there is a lack of specialized algorithms to investigate highly variable unstructured neural activity.

Additionally, there is also a void in supervised algorithms that identify spatiotemporal patterns that best discriminate between different classes. Since spatiotemporal patterns are a hallmark of brain activity, it would be incredibly beneficial to have a mathematical algorithm to extract specific patterns that are truly salient to one particular class. In Chapter 2, we used GPFA due to the simplicity of its assumptions. However, a supervised GPFA algorithm that looks for spatiotemporal patterns that best separate between different classes is a powerful tool that could enhance our understanding of neural activity driving naturalistic behavior.

When investigating the neural correlates to autonomic tone, we have certainly only scratched the surface in Chapter 4. We investigated sleep due to the lack of activity and cognitive states that could impact autonomic and neural activity. A natural progression is to examine the neural correlates to autonomic tone during waking rest state. One approach is to assume that the majority of waking activity could be categorized as 'rest' state. Therefore, by the law of averages, we will eventually make reliable conclusions about the correlation of rest state and autonomic tone given enough data. One can use a camera and microphone in a more structured approach, similar to our studies in Chapters 2-3 to identify 'rest' periods. In analyzing the correlation between rest and autonomic tone using intracranial electrophysiology, one can compare whether the results from fMRI based studies are reproducible. Additionally, future work could involve investigating the ANS-CNS connection using a finer temporal resolution to infer directionality and causality.

References

[1] David A. McCormick, Dennis B. Nestvogel, and Biyu J. He. Neuromodulation of Brain State and Behavior. *Annual Review of Neuroscience*, 43(1):391–415, 2020. _eprint: https://doi.org/10.1146/annurev-neuro-100219-105424.

[2] Byron M. Yu, John P. Cunningham, Gopal Santhanam, Stephen I. Ryu, Krishna V. Shenoy, and Maneesh Sahani. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology*, 102(1):614–635, July 2009.

[3] Chethan Pandarinath, Daniel J. O'Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D. Stavisky, Jonathan C. Kao, Eric M. Trautmann, Matthew T. Kaufman, Stephen I. Ryu, Leigh R. Hochberg, Jaimie M. Henderson, Krishna V. Shenoy, L. F. Abbott, and David Sussillo. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*, 15(10):805–815, October 2018.

[4] Matar Haller, Thomas Donoghue, Erik Peterson, Paroma Varma, Priyadarshini Sebastian, Richard Gao, Torben Noto, Robert T. Knight, Avgusta Shestyuk, and Bradley Voytek. Parameterizing neural power spectra. *bioRxiv*, page 299859, January 2018.

[5] Kai J Miller, Christopher J Honey, Dora Hermes, Rajesh PN Rao, Marcel denNijs, and Jeffrey G Ojemann. Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *New Horizons for Neural Oscillations*, 85:711–720, January 2014.

[6] Jeremy R. Manning, Joshua Jacobs, Itzhak Fried, and Michael J. Kahana. Broadband Shifts in Local Field Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans. *Journal of Neuroscience*, 29(43):13613–13620, October 2009. Publisher: Society for Neuroscience Section: Articles.

[7] Roy Mukamel, Hagar Gelbard, Amos Arieli, Uri Hasson, Itzhak Fried, and Rafael Malach. Neuroscience: Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science*, 309(5736):951–954, August 2005.

[8] Supratim Ray and John H. R. Maunsell. Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLOS Biology*, 9(4):e1000610, April 2011. Publisher: Public Library of Science.

[9] Emily L Mackevicius, Andrew H Bahle, Alex H Williams, Shijie Gu, Natalia I Denisenko, Mark S Goldman, and Michale S Fee. Unsupervised discovery of temporal sequences in high-dimensional datasets, with applications to neuroscience. *eLife*, 8:e38471, February 2019. Publisher: eLife Sciences Publications, Ltd.

[10] Alex H. Williams, Tony Hyun Kim, Forea Wang, Saurabh Vyas, Stephen I. Ryu, Krishna V. Shenoy, Mark Schnitzer, Tamara G. Kolda, and Surya Ganguli. Unsupervised Discovery of Demixed, Low-Dimensional Neural Dynamics across Multiple Timescales through Tensor Component Analysis. *Neuron*, 98(6):1099–1115.e8, June 2018. Publisher: Cell Press.

[11] Bingni W. Brunton, Lise A. Johnson, Jeffrey G. Ojemann, and J. Nathan Kutz. Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, 258:1–15, January 2016. Publisher: Elsevier B.V. _eprint: 1409.5496.