Probabilistic approaches for tracking physiological states in the cortex through sleep and seizures

by

Vera Marie Dadok

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

in

Engineering – Mechanical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Andrew J. Szeri, Chair
Professor J. Karl Hedrick
Professor Laurent El Ghaoui
Professor Heidi E. Kirsch

Fall 2013

Probabilistic approaches for tracking physiological states in the cortex through sleep and seizures

Abstract

Probabilistic approaches for tracking physiological states in the cortex through sleep and seizures

by

Vera Marie Dadok

Doctor of Philosophy in Engineering – Mechanical Engineering

University of California, Berkeley

Professor Andrew J. Szeri, Chair

We present work in this dissertation on methods to map measured electrode signals from human subjects into the physiologically relevant parameter space of a mathematical model of the cortex, approaching two specific dynamical brain phenomena: sleep and seizures.

In the context of sleep, we develop a probabilistic method for mapping human sleep electroencephalogram (EEG) signals onto a state space based on a biologically plausible mathematical model of the cortex. From a noninvasive EEG signal, this method produces physiologically meaningful pathways of the cortical state over a night of sleep. We propose ways in which these pathways offer insights into sleep-related conditions, functions, and complex pathologies. To address explicitly the noisiness of the EEG signal and the stochastic nature of the mathematical model, we use a probabilistic Bayesian framework to map each EEG epoch to a distribution of likelihoods over all model sleep states. A Hidden Markov Model (HMM) is incorporated to improve the path results using the prior knowledge that cortical physiology has temporal continuity.

Next we adapt our probabilistic methodology to infer the parameter region in the mathematical model of the cortex most likely to be producing seizures observed in an electrocorticogram (ECoG) signal. This method produces a probabilistic pathway of the temporal evolution of physiological state in the cortex over the course of individual seizures. We describe ways in which these methods and results offer insights into seizure etiology and suggest potential new treatment options.

Once again, a probabilistic Bayesian framework is used to map features of EEG or ECoG segments onto a distribution of likelihoods over physiological parameter states. And again, a Hidden Markov Model (HMM) is introduced to incorporate the belief that cortical physiology has both temporal continuity and also a degree of reproducibility between individual seizures. However, for seizures, we additionally inspect the ratio of likelihoods between HMMs run under two possible parameter regions, both of which produce seizures in the model, to determine which physiological parameter regions are more likely to be causing the observed seizures. We show that between individual seizures, there is consistency in these

likelihood ratios between hypothesized regions, in the temporal pathways calculated, and in the separation of seizure from non-seizure time segment likelihood maps. We also improve upon several of the underlying techniques for sampling the parameter state space, feature selection, and probability density estimation.

I dedicate this dissertation to my family and friends

Contents

List of Figures

List of Tables

Acknowledgments

I could not have done this research without the help of my advisor Andrew Szeri and his excellent guidance. I also want to thank my other dissertation committee members for their time, thoughts, and suggestions: Karl Hedrick, Laurent El Ghaoui, and Heidi Kirsch. The support of my collaborators as I conducted my PhD research was invaluable and I would like to extend my thanks to Jamie Sleigh and Beth Lopour. I would also like to thank Kevin Haas and Prashanth Selvaraj for their suggestions and insights.

This work was partially supported by an NSF Graduate Research Fellowship to Vera Dadok and in part by the National Science Foundation through the research grant CMMI 1031811.

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

Chapter 1

Introduction

Approximately 0.4 to 1% of the world's population has epilepsy [65], a clinical condition characterized by recurring seizures. Although epilepsy has been diagnosed and studied for hundreds of years, many of the causes of both epilepsy and the onset of individual seizures are not well understood [12; 41; 73].

1.1 Epilepsy Background

Epilepsy is a pathological condition in which chronic, recurrent seizures occur, and epileptogenesis is the way in which a normal brain develops epilepsy [72]. Seizures are each a single event in which a neuronal population in the brain "exhibits abnormal synchronous excitation" [72].

There are many different classifications of epilepsies based on clinical symptoms, etiology, and mechanisms of action. Often a clinical classification is based on the development of the abnormal brain patterns. Partial or focal epilepsy is a seizure that originates at a small region of the brain, the 'focus' of the seizure. The focal seizure may have a secondary generalization, in which the abnormal synchronous behavior spreads throughout the brain. This differs from generalized seizures in which the initial onset is generalized throughout both hemispheres of the brain [37].

There are many pathological conditions that are known or hypothesized to be caused by or correlated with seizures in different types of epilepsy. Typically focal seizures are thought to be brought on by increased excitability of neurons which may be caused by changes in cellular properties, alterations in synaptic connections (possibly due to scarring, tumors, blood clots, abnormal cell populations, abnormal glial cell functionality, etc) [37]. Comparatively, in the case of generalized seizures, which have a more heterogeneous presentation of brain activity, hyperexcitability of the cortex is still expected, but the thalamocortical pathways are thought to play a large role in the rapid generalization to the entire cortex [37].

Although some general characteristics are expected or hypothesized to be present in the development of seizure onset (ictogenesis), the true mechanisms underlying the transition

from normal brain behavior into seizure are not well understood in many cases [19; 93]. Examples include contention over whether focal seizures develop due to focuses or networks, whether hyposynchrony may be essential to ictal onset, and if excessive inhibition can trigger seizures [25].

1.1.1 Treatments

Treatment options for epilepsy patients are limited. Generally, anti-seizure medications are tried first, and prove effective in approximately two thirds of cases [63]. These anti-epileptic drugs (AEDs) have numerous risks and side effects associated with their use and the reasons for their efficacy or failure are often unclear [63]. Epilepsies that do not respond to AEDs are referred to as refractory epilepsy.

After the failure of medication to treat seizure symptoms, refractory epilepsy patients typically pursue resection surgery if seizures are focal. If resection surgery is not viable or seizures are not focal, patients may choose to have an implanted Vagus Nerve Stimulator (VNS) device which is less than 3% effective in entirely preventing seizures but reduces seizure frequency by about 50% on average in patients [9; 21; 20].

Two new deep brain stimulation (DBS) devices, developed by NeuroPace and Medtronic, have recently been moving through the FDA approval process and show promise [82; 29; 35]. The Neuropace device is just now receiving pre-market FDA approval. Overall, treatment options are limited with few viable selections for patients, each with significant risks and side effects.

1.1.2 Sleep and seizures

For certain types of seizure, sleep has a significant impact, and literature shows that sleep stage may be playing an important role in determining the onset of seizures [41; 23] and plays a significant role in the failure of seizure prediction algorithms [73].

Thus a better understanding of both seizure and sleep physiology dynamics is crucial for a better understanding and treatment of pathologies such as epilepsy that interact with sleep. Full measurement of the physiological state of the brain during sleep would require invasive measurements and is poorly suited for typical human subjects research. Thus researchers have long used noninvasive measurements of electrical brain signals for traditional sleep staging. Traditional sleep stages were not developed to capture and thus do not capture all of the physiological information and microstructures available in EEG and other electrical signals. However sleep staging does provide a rich standard for classifying sleep and serves as a foundation to inform additional more detailed representations of sleep and its microstructure such as those described in [8; 83; 11; 58; 51; 31] and this work. Recently Lopour et. al [51] worked quantitatively to relate experimentally measured human sleep EEG data to a continuous set of sleep states based on a dynamical mathematical model of the brain.

1.1.3 Tools for study of epilepsy and treatment development

Investigations of epilepsy and seizures typically fall into the following categories: experimental based animal studies, clinical data collection of human subjects and data mining of clinical data, mathematical model-based investigations of modeled seizures, and investigations of experimental data leveraging mathematical models.

Experimental animal studies, including studies of cultured tissue, brain slices, and in-vivo animal studies, are often referred to as 'animal models' [88]. In each of these cases, the applications of the study results are limited due to the constraints of the experiment. Clearly, the behavior of cultured tissue and brain slices in isolation may differ greatly from that of whole-brain in-vivo. Similarly, animal models and in-vivo experiments are applicable to the types of seizures generated or found in those animals. This encompasses only a limited set of epilepsy types that often only loosely relate to a few epilepsy types found in humans. For instance, seizures are often artificially introduced into the animals via electroshock, chemical compounds, genetic modifications, and other methods [33; 12; 4]. Though the results of these often share certain characteristics with clinically observed seizures in humans, and offer insights into certain types of epilepsies, the mechanisms of epileptogenesis and resulting mechanisms of seizure evolution and propagation do not survey the full spectrum of human epilepsies. Animal models are used extensively for studying the efficacy and safety of new antiepileptic drugs (AEDs) [33; 12], in fact [5] shows which AEDs are effective in several different animal model classes.

There are also a number of clinical and data-mining studies of epilepsy and seizures using non-invasively measured human subject data and data gathered in the course of medically necessary treatments. These studies may investigate in-vivo properties such as electrode measurements (electroencephalogram (EEG), electrocorticogram (ECoG), or depth electrodes) or other time series data collected (e.g. heart rate, blood oxygen content, video). Many seizure detection and prediction papers (e.g. [27; 1; 16; 57]) use such data-mining methods on EEG to build empirically verified detection or prediction algorithms that could be applied for medical or clinical applications in the future. In recent years, more databases of such clinically-relevant collected data sets have been established or are being established (e.g. [34; 28]) and will enable cheaper and more reliable access to data.

Additionally, studies are conducted on other human subjects data such as the properties of resected tissue samples after medically necessary surgery, genetic samples, and MRI data. Studies using surgically resected tissues from focal seizures often investigate the histopathology, or microscopic and cellular properties, of the tissues (e.g. in cortical dysplasia [14]).

We also call attention to a body of literature on models of epilepsy and primarily model-based analysis of seizures. There are numerous computational models at various scales from single-neuron, to neural network, to meso-scale mean-field models of the brain [88]. Many of these model-centric studies focus on modeled seizures with only a brief or qualitative discussion of experimentally measured data. These studies often offer important insights into what *types* of mechanisms and models may be associated with seizures (e.g. [59]), develop new mathematical models that show certain important properties, analyze model

seizures with the hope of understanding mechanisms or testing new treatments (e.g. [50]), or conduct other exploratory investigations that are not practical for experimental studies by using a model as a testbed.

The final type of analysis are studies that attempt to bridge the gap between mathematical modeling and experimental data in a quantitative way. The goal of these studies is to leverage the insights gained from mathematical models to explain or clarify the hidden processes in experimental data. Verified mathematical models also hold promise to extend experimental knowledge beyond understanding of past experiments and into prediction of new phenomena. Many papers have identified the benefits of such analyses and their potential applications range from aid in developing new seizure treatments to new understanding of physiological mechanisms in seizures [88; 3; 42; 43]. Such studies can avoid the prohibitive costs both financially and ethically that are involved in experimental studies and are more flexible for conducting more exploratory analyses.

These are all promising areas, however, all have their drawbacks. Animal-based studies are costly, and only investigate certain types of seizures that are possible to trigger in animal models. Clinical data collection is essential, but studies based only on clinical data do not always fully interpret the physiological meaning underlying the results and do not often offer the potential to predict new results. Pure model-based studies are not always grounded in reality, and are challenging to defend. However it is quite difficult to relate models to reality, as no model is a full description of the human brain. However, we believe that this leveraging of models to understand better the physiology of seizures is essential to the development of improved treatment strategies for epilepsy.

1.2 Contributions of this dissertation

In this work we explore the intersection of experimental cortical measurements and a mathematical model of cortical dynamics. Experimental human measurements are vital for a patient-specific, individual-based understanding of pathologies and functionality of brain states. Mathematical representations are an excellent non-invasive platform we leverage to understand better true brain dynamics. By using models to understand seizures and sleep, we can expect to improve prediction, treatment, and investigations into these brain states above what we could accomplish with empirical models and experimental data alone. Also, as research progresses in finding stronger connections between such models and brain measurements, we may be able to work with less invasive measurements and run preliminary experiments via computation, using promising model-based results to choose better the direction of costly experiments.

With this broad aim, we present work in this thesis on promising ways to map measured electrode signals from human subjects into the physiologically relevant parameter space of a mathematical model. With these methods we approach two specific dynamical brain phenomena: sleep and seizure.

1.2.1 Probabilistic sleep evolution

The first major contribution to the field is the development of a probabilistic method for mapping human sleep electroencephalogram (EEG) signals onto a state space based on a biologically plausible mathematical model of the cortex. From a noninvasive EEG signal, this method produces physiologically meaningful pathways of the cortical state over a night of sleep. We propose ways in which these pathways offer insights into sleep-related conditions, functions, and pathologies. To address explicitly the noisiness of the EEG signal and the stochastic nature of the mathematical model, we use a probabilistic Bayesian framework to map each EEG epoch to a distribution of likelihoods over all model sleep states. We show that the mapping produced from human data robustly separates rapid eye movement sleep (REM) from slow wave sleep (SWS). A Hidden Markov Model (HMM) is incorporated to improve the path results using the prior knowledge that cortical physiology has temporal continuity.

This sleep mapping is proposed as a more detailed and physiologically relevant way to understand sleep via EEG recordings. It is an improvement over classic R&K sleep staging often used in clinical practice due to its physiological insights. This work also extends [51], which produced a mapping into cortical states via a dimensionality reduction method. The method presented in this dissertation is an improvement as it handles the stochastic nature of the system explicitly with probabilistic methods and also incorporates a method to handle temporal continuity of the physiological states.

1.2.2 Probabilistic seizure evolution

The second major contribution is the application of this probabilistic method to inferring the parameter region in a biologically plausible mathematical model of the cortex most likely to be producing seizures observed in an electrocorticogram (ECoG) signal. Additionally, this method produces a probabilistic pathway of the temporal evolution of physiological state in the cortex over the course of individual seizures, leveraging a model of the cortex to describe cortical physiology. We describe ways in which these methods and results offer insights into seizure etiology and suggest potential new treatment options.

Leveraging the work in mapping sleep to a mathematical model, we again account for the stochastic and noisy nature of the mathematical model and the ECoG signal, and use a probabilistic Bayesian framework to map features of ECoG segments onto a distribution of likelihoods over physiological parameter states. A Hidden Markov Model (HMM) is then introduced to incorporate the belief that cortical physiology has both temporal continuity and also a degree of reproducibility between individual seizures. By inspecting the ratio of likelihoods between HMMs run under two possible parameter regions, both of which produce seizures in the model, we determine which physiological parameter regions are more likely to be causing the observed seizures. We show that between individual seizures, there is consistency in these likelihood ratios between hypothesized regions, in the temporal pathways calculated, and in the separation of seizure from non-seizure time segment likelihood maps.

One of the major improvements of this approach over current literature in this area is the usage of this physiologically based model to understand electrical brain signals, rather than using only signal processing techniques with no eye to the biological context. Much literature analyzing EEG and ECoG signals in seizures limits itself to pure signal processing approaches without full consideration of the biological basis of these signals. A more full description of the surrounding literature and the differences in this approach are described in Chapter 5.

1.3 Map of this dissertation

In Chapter 2 we describe the mathematical model for sleep and seizure and motivate its usage in the following studies. Next, in Chapter 3 we introduce some of the probabilistic methods used in our analyses. After these introductory chapters, Chapter 4 contains the methods for and experimental results of mapping sleep EEG data into a parameter space of the model. We then describe in Chapter 5 the methods for choosing more likely seizure parameter regions and demonstrate the results tracking parameter changes over the course of seizures. We then conclude and offer insights into future extensions in Chapter 6.

Chapter 2

Cortical model

We use a mean field model of the cortical dynamics developed in [47; 77; 48] for the waking cortex and in [76] for the sleeping cortex as a platform to explore the physiological evolution of seizures and of sleep. Such spatially-continuous models have been proposed in literature for some time to model mesoscale and macroscale phenomena (e.g. EEG, ECoG) with less computational cost than neural networks. This mesoscale model of the cortex treats the brain as an active medium in which average neuron populations of inhibitory and excitatory cells interact with subcortical inputs and short and long range connections between populations. The model is derived from a physiological understanding of cortical connectivity and its parameters are based on experimentally measured physiological quantities. A sketch of the principles operating in this dynamical model is shown in Figure 2.1.

This type of mesoscale model is well-suited for comparison to EEG or electrocorticogram (ECoG) data, due to the fact that EEG electrodes measure the aggregate behavior of neuron populations as a result of the large electrode spatial scale. In fact, simulations of this model produce ECoG-like data, approximated by either dimensionless variable $\tilde{h}_e$ (for sleep analysis) or $\tilde{h}_m$[1] (for seizure analysis), which we will leverage to understand how physiological states in the model produce brain waves with different properties.

This model also has the advantage of significant literature displaying its applicability to modeling both seizures and sleep as well as other phenomena such as anesthesia and comas [42; 79; 90; 46; 7].

In mathematical terms, this model is a set of stochastic nonlinear partial differential equations displayed in full in Section 2.3 along with brief descriptions of the parameters and variables.

2.1 Seizures in the cortical model

In a variety of parameter settings, this model exhibits seizure-like behavior [42; 43; 74]. These seizures exhibit sharp rhythmic spiking waves which propagate through the simulated

[1] This metric is derived and described in [50]

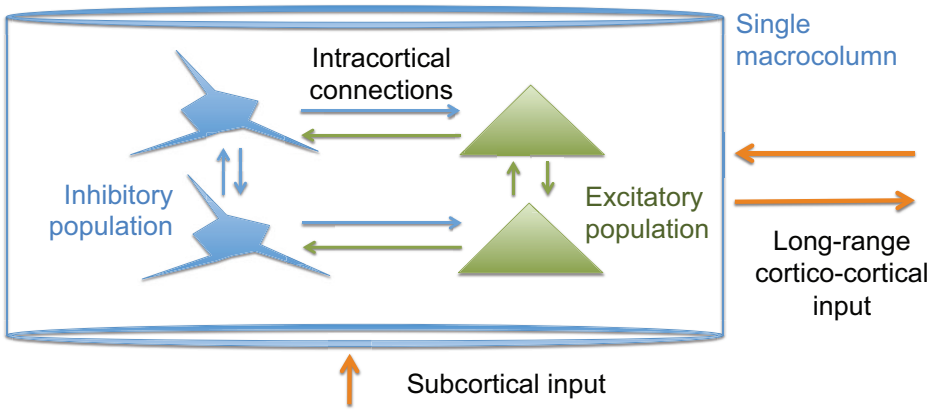

Figure 2.1: A conceptual sketch of the biological principles underlying system of partial differential equations governing the cortical model. Neuronal populations of excitatory and inhibitory cells are arranged in macrocolumns that interact with populations in neighboring columns and within each column. Additionally there is a stochastic subcortical input to the cortex populations, modeling the inputs from other brain structures. A single macrocolumn is depicted here with implied neighboring columns.

cortex. We delve further into particular seizure parameter regimes and examples in Chapter 5.

2.2 Cortical model of sleep

This particular cortical model is adapted to the sleeping brain by the presence of two parameters L and Δh_e^{rest}, presenting a natural continuous 2-dimensional state space for sleep [76; 51]. The parameter L describes the connection strength of excitatory neuron populations as a gain on the excitatory postsynaptic potentials, and reflects the biology of synaptic downscaling that is believed to be an important function of sleep [13; 49; 18]. The other parameter Δh_e^{rest} represents a change in the resting potential of the excitatory population of neurons in the cortex. This is the primary target for the various brain-stem neuromodulators that control arousal (acetylcholine and amines), and also the various somnogens such as adenosine that increase sleepiness. Thus the model is able to link processes of interest at the cellular level with changes in neuronal excitability and connectivity and hence with changes in the EEG. More information on this model and its derivation can be found in [76; 91].

Table 2.1: Model variables

Variable	Description
$\tilde{h}_{e,i}$	Dimensionless mean soma potential
$\tilde{I}_{ee,ie,ei,ii}$	Dimensionless postsynaptic activation
$\tilde{\phi}_{e,i}$	Dimensionless distant input to neuron populations
$\tilde{t}$	Dimensionless time
$\tilde{x}$	Dimensionless space

2.3 Model equations and default parameters

The dimensionless version of the model developed in [76] detailed in [42] for seizures and in [51] for sleep is presented in Equations 2.1 to 2.8. In Table 2.1 the definitions of the variables are described while the parameter definitions and nominal values for sleep and seizure cases are contained in Tables 2.3 and 2.2 respectively. For further information on these parameters see [48; 78] for general parameters and [76] for sleep parameters.

$$\frac{\partial \tilde{h}_e}{\partial \tilde{t}} = 1 - \tilde{h}_e + \frac{\Delta h_e^{rest}}{h^{rest}} + L\Gamma_e(h_e^0 - \tilde{h_e})\tilde{I}_{ee} + \Gamma_i(h_i^0 - \tilde{h_e})\tilde{I}_{ie} \tag{2.1}$$

$$\frac{\partial \tilde{h}_i}{\partial \tilde{t}} = 1 - \tilde{h}_i + L\Gamma_e(h_e^0 - \tilde{h_i})\tilde{I}_{ei} + \Gamma_i(h_i^0 - \tilde{h_i})\tilde{I}_{ii} \tag{2.2}$$

$$\left(\frac{1}{T_e}\frac{\partial}{\partial \tilde{t}} + 1\right)^2 \tilde{I}_{ee} = N_e^\beta \tilde{S}_e\left[\tilde{h}_e\right] + \tilde{\phi}_e + P_{ee} + \tilde{\Gamma}_1 \tag{2.3}$$

$$\left(\frac{1}{T_e}\frac{\partial}{\partial \tilde{t}} + 1\right)^2 \tilde{I}_{ei} = N_e^\beta \tilde{S}_e\left[\tilde{h}_e\right] + \tilde{\phi}_i + P_{ei} + \tilde{\Gamma}_2 \tag{2.4}$$

$$\left(\frac{1}{T_i}\frac{\partial}{\partial \tilde{t}} + 1\right)^2 \tilde{I}_{ie} = N_i^\beta \tilde{S}_i\left[\tilde{h}_i\right] + P_{ie} + \tilde{\Gamma}_3 \tag{2.5}$$

$$\left(\frac{1}{T_i}\frac{\partial}{\partial \tilde{t}} + 1\right)^2 \tilde{I}_{ii} = N_i^\beta \tilde{S}_i\left[\tilde{h}_i\right] + P_{ii} + \tilde{\Gamma}_4 \tag{2.6}$$

$$\left(\frac{1}{\lambda_e}\frac{\partial}{\partial \tilde{t}} + 1\right)^2 \tilde{\phi}_e = \frac{1}{\lambda_e^2}\frac{\partial^2 \tilde{\phi}_e}{\partial \tilde{x}^2} + \left(\frac{1}{\lambda_e}\frac{\partial}{\partial \tilde{t}} + 1\right) N_e^\alpha \tilde{S}_e\left[\tilde{h}_e\right] \tag{2.7}$$

$$\left(\frac{1}{\lambda_i}\frac{\partial}{\partial \tilde{t}} + 1\right)^2 \tilde{\phi}_i = \frac{1}{\lambda_i^2}\frac{\partial^2 \tilde{\phi}_i}{\partial \tilde{x}^2} + \left(\frac{1}{\lambda_i}\frac{\partial}{\partial \tilde{t}} + 1\right) N_i^\alpha \tilde{S}_e\left[\tilde{h}_e\right] \tag{2.8}$$

The sigmoid transfer functions that transform the mean soma voltage to the dimensionless firing rates ($\tilde{S}_e,\ \tilde{S}_i$) for the population are shown in Equations 2.9 and 2.10.

$$\tilde{S}_e\left[\tilde{h}_e\right] = \frac{1}{1+\exp\left[-\tilde{g}_e(\tilde{h}_e-\tilde{\theta}_e)\right]} \tag{2.9}$$

$$\tilde{S}_i\left[\tilde{h}_i\right] = \frac{1}{1+\exp\left[-\tilde{g}_i(\tilde{h}_i-\tilde{\theta}_i)\right]} \tag{2.10}$$

The stochastic input from subcortical sources are shown in Equations 2.11 to 2.14.

$$\tilde{\Gamma}_1 = \alpha_{ee}\sqrt{P_{ee}}\xi_1[\tilde{x},\tilde{t}] \tag{2.11}$$

$$\tilde{\Gamma}_2 = \alpha_{ei}\sqrt{P_{ei}}\xi_2[\tilde{x},\tilde{t}] \tag{2.12}$$

$$\tilde{\Gamma}_3 = \alpha_{ie}\sqrt{P_{ie}}\xi_3[\tilde{x},\tilde{t}] \tag{2.13}$$

$$\tilde{\Gamma}_4 = \alpha_{ii}\sqrt{P_{ii}}\xi_4[\tilde{x},\tilde{t}] \tag{2.14}$$

Variables ξ_k represent white Gaussian noise with zero mean, δ-function correlations and approximated numerically by:

$$\xi_k[\tilde{x},\tilde{t}] = \frac{R(m,n)}{\sqrt{\Delta\tilde{x}\Delta\tilde{t}}} \tag{2.15}$$

and $[\tilde{x},\tilde{t}] = [m\Delta\tilde{x},\ \ n\Delta\tilde{t}]$–integer coordinates on a grid spaced by $[\Delta\tilde{x},\ \ \Delta\tilde{t}]$. This PDE was simulated in MATLAB with an Euler predictor-corrector method designed in [51].

For only the sleep model, Equation 2.16 is used for α_{xx}, all subcortical inputs, and the form of this equation is from [51]. In this equation μ_e is based on the values of the equilibrium points h_e of the PDE, scaled so that μ_e varies between 0 and 1, where 0 is the minimum over this region of the manifold, and 1 is the maximum over the manifold. For further information, please see [51].

$$\alpha_{ee,ei,ie,ii} = \hat{\alpha}(-4\mu_e+5) \tag{2.16}$$

For the seizure model, a measurement variable $\tilde{h}_m$, developed in [50], derived from other state variables, and displayed in Equation 2.17 is used as an ECoG-like measurement.

$$h_m = (h_o^e) - \tilde{h}_e\tilde{I}_m = \frac{(45-h_e)}{-70}\tilde{I}_m \tag{2.17}$$

Parameter		Definition	Value
$\tilde{\Gamma}_{e,i}$	Influence of input on mean soma potential	$\frac{\Gamma_{e,i}\exp(1)S^{max}}{\gamma_{e,i}\left\|h^{rev}_{e,i}-h^{rest}\right\|}$	1.42×10^{-3}, 0.0774
$h^0_{e,i}$	Dimensionless reversal potential	$h^{rev}_{e,i}/h^{rest}_{e,i}$	-0.643, 1.29
$T_{e,i}$	Dimensionless neurotransmitter rate constant	$\tau\gamma_{e,i}$	12.0, 2.60
$\lambda_{e,i}$	Dimensionless corticocortical inverse length scale	$\tau v\Lambda_{ee,ei}$	11.2, 18.2
$P_{ee,ie}$	Subcortical input to e populations	$p_{ee,ie}/S^{max}$	11.0, 16.0
$P_{ei,ii}$	Subcortical input to i populations	$p_{ei,ii}/S^{max}$	16.0, 11.0
$N^{\alpha}_{e,i}$	Number of distant connections	$N^{\alpha}_{e,i}$	4000, 2000
$N^{\beta}_{e,i}$	Number of local connections	$N^{\beta}_{e,i}$	3034, 536
$\tilde{g}_{e,i}$	Dimensionless sigmoid slope at inflection point	$g_{e,i}h^{rest}$	-19.6, -9.80
$\tilde{\theta}_{e,i}$	Dimensionless inflection point for sigmoid	$\theta_{e,i}/h^{rest}$	0.857, 0.857
h^{rest}_e	Resting potential		-70 mV
$\alpha_{ee,ei,ie,ii}$	Stochastic input gain		1.6
L	Strength of excitatory connections		1
Δh^{rest}_e	Change in resting potential		0

Table 2.2: Nominal model parameters for seizures. Note that sleep variables L and Δh^{rest}_e are chosen to have no effect in this parameter set for seizures. These default parameter values are taken from [42].

Parameter		Definition	Value
$\tilde{\Gamma}_{e,i}$	Influence of input on mean soma potential	$\frac{\Gamma_{e,i}\exp(1)S^{max}}{\gamma_{e,i}\left\|h^{rev}_{e,i}-h^{rest}\right\|}$	4.6875×10^{-4}, .0105
$h^0_{e,i}$	Dimensionless reversal potential	$h^{rev}_{e,i}/h^{rest}_{e,i}$	0, 1.0938
$T_{e,i}$	Dimensionless neurotransmitter rate constant	$\tau\gamma_{e,i}$	12, 3.6
$\lambda_{e,i}$	Dimensionless corticocortical inverse length scale	$\tau v\Lambda_{ee,ei}$	11.2, 11.2
$P_{ee,ie}$	Subcortical input to e populations	$p_{ee,ie}/S^{max}$	25, 25
$P_{ei,ii}$	Subcortical input to i populations	$p_{ei,ii}/S^{max}$	25, 25
$N^{\alpha}_{e,i}$	Number of distant connections	$N^{\alpha}_{e,i}$	3710, 3710
$N^{\beta}_{e,i}$	Number of local connections	$N^{\beta}_{e,i}$	410, 800
$\tilde{g}_{e,i}$	Dimensionless sigmoid slope at inflection pt	$g_{e,i}h^{rest}$	-29.021, -19.347
$\tilde{\theta}_{e,i}$	Dimensionless inflection point for sigmoid	$\theta_{e,i}/h^{rest}$	0.91406, 0.91406
h^{rest}_e	Resting potential		-64 mV
L	Strength of excitatory connections		
Δh^{rest}_e	Change in resting potential		
$\hat{\alpha}$	Minimum gain on subcortical input		

Table 2.3: Nominal model parameters for sleep simulations. Values for all parameters in this default set were taken from [51].

Chapter 3

Probabilistic Methods

In this section, we present two of the probabilistic models that will enable our analysis: a Naive Bayes model for inference of state given stochastic observations and a Hidden Markov model for incorporating dynamical state evolution through time.

We describe them in depth here and use them later in Chapters 4 and 5 to analyze data from the stochastic dynamical system of the cortex.

3.1 Naive Bayes for likelihood estimation

3.1.1 Inference problem

Suppose we are given a system with a set $\mathbf{s}$ of K discrete states s_k, $k \in [1, \ldots, K]$, each of which produce M stochastic observations $\mathbf{e} = [e_1, \ldots, e_M]$ via some unknown probabilistic relationship, defined by a probability distribution conditioned on each state $p(e_1, \ldots, e_M|\mathbf{s} = s_k)$. Additionally, we have been given a set of training data (labeled samples): known states paired with their observations $(s_k, \mathbf{e})_i$. The problem we seek to address is to infer what the underlying state is for a new set of observations; equivalently to estimate the posterior probability of each state given the observations $p(s_k|\mathbf{e})$. This is a classic inference problem. As an example, the states could be a binary option on the statement 'this email is spam' and the observations could be measurements such as words present, email address, and time of day. We additionally have a historical data set of samples from past emails. Then, for each email, the problem becomes to infer the state (whether it is spam or not) given our observations of the current email combined with knowledge of the system.

One way to address this inference problem is by using Bayes' rule, shown in Equation 3.1. The training data is used to estimate a likelihood function $p(e_1, \ldots, e_M|s_k)$. The prior probability $p(s_k)$ is estimated from the data or assigned values based on assumptions. The marginal likelihood $p(\mathbf{e})$ in the denominator is a summation of the numerator over all states.

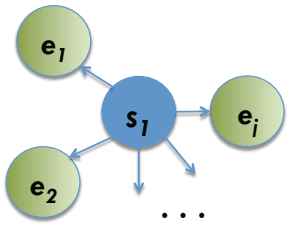

Figure 3.1: Graphical model of a Naive Bayes structure. The parameter state $\mathbf{s}$ of the model directly influences the probable outcome of the observations, and this state is the only parent to each feature. This structure implies that the observations are independent of one another given the state. More information on these models can be found in [10]

$$p(\mathbf{s} = s_k|\mathbf{e}) = \frac{p(\mathbf{e}|\mathbf{s} = s_k)p(s_k)}{p(\mathbf{e})} \tag{3.1}$$

Returning to the earlier example, upon encountering a new email with observed features $\mathbf{e}$, if we estimate from historical, labeled data the distribution of what has been observed in the past from the states $p(\mathbf{e}|\mathbf{s} = s_k)$, assume a prior on states $p(s_k)$, and normalize with $p(\mathbf{e})$, we can take *past, known* observations and build a model to infer the probability over all states s_k of the *unknown* state underlying new observations. This is a very powerful tool and Bayes' rule is used across all disciplines. It describes how to leverage observed information to better estimate for an unknown state that is probabilistically linked to those observations the probability across possible states .

3.1.2 Likelihood function estimation

For a very simple problem with a large number of samples and only a few binary observations and a binary state, estimating $p(\mathbf{e}|\mathbf{s} = s_k)$ is a straightforward process. However, as the number of states K and observations M increase, the problem of estimating the likelihood function with no assumptions about its structure becomes intractable as both the necessary amount of training data and computational time increases exponentially.

To reduce the amount of time and number of data samples necessary to estimate $p(\mathbf{e}|\mathbf{s} = s_k)$, one can make an assumption that the data has an underlying 'Naive Bayes' structure, shown as a graphical model in Figure 3.1. Note that this structure implies that the features are independent of one another conditioned the state $p(e_i, e_j|\mathbf{s} = s_k) = p(e_i|\mathbf{s} = s_k)p(e_j|\mathbf{s} = s_k)$, which is assumed to be a 'naive' assumption, hence the name of the structure. Under this structure, the joint conditional probability distribution $p(\mathbf{e}|\mathbf{s} = s_k)$ can be simplified to a product of one-variable conditional distributions as shown in Equation 3.2.

$$p(\mathbf{e}|\mathbf{s} = s_k) = \prod_i^M p(e_i|\mathbf{s} = s_k) \tag{3.2}$$

Naive Bayes structures are often applied to classification problems and can also be used for probability density function estimation. They are often excellent classifiers [69] in spite of their simplicity even if the assumption of independence is incorrect. Some empirical evidence that Naive Bayes classifiers can perform well for probability density estimation is shown in [52].

3.2 Hidden Markov Models

Given a sequence of stochastic observations in a time series $[\mathbf{e}^1, \ldots, \mathbf{e}^T]$, derived from an underlying, but unobserved state trajectory $[\mathbf{s}^1, \ldots, \mathbf{s}^T]$, one could treat the observations as independent events when trying to estimate the underlying state from the observations. However, if one assumes that there is an underlying structure to the evolution of state in time, ignoring the temporal aspects in this way can lead to incorrect results and ignore important correlations due to the dynamics.

One way to incorporate temporal dynamics evolution is to use a Hidden Markov Model (HMM) which is used to model stochastic processes with Markovian dynamics in which a state at time t depends on the state at time $t-1$ and have the property that $p(\mathbf{s}^t|\mathbf{s}^{[1,\ldots,t-1]}) = p(\mathbf{s}^t|\mathbf{s}^{t-1})$. This property can be seen via the graphical representation of an HMM shown in Figure 3.2. Thus it incorporates a notion of temporal dynamics, and is a relatively simple model with well-known statistical algorithms. HMMs are very versatile and used to model such things as human speech, DNA, and other non-stationary processes [86].

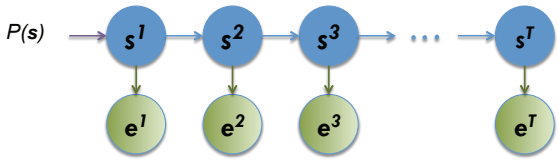

Figure 3.2: A graphical representation of the structure of a Hidden Markov Model. The arrows denote conditional dependencies.

HMMs are known as 'Hidden' Markov Models because they additionally assume that the states are unobserved, and only the noisy measurements of the states can be observed. Hidden Markov models consist of two levels, as seen in Figure 3.2 a 'hidden' level (blue in the figure) in which the state $\mathbf{s}^t$ of the system of interest evolves in a Markovian manner, and an observed level (green in the figure), containing some directly measurable observations or evidence, $\mathbf{e}^t$ that is dependent on the hidden state at the same time. An overview of Hidden Markov Models can be found in [86].

Each arrow in Figure 3.2 represents a conditional probability relationship. The arrow from each state to its observation represents the probability distribution of observing $\mathbf{e}^t$ given state $\mathbf{s}^t$, and is defined by the function $p(e_i|\mathbf{s} = s_k)$. The arrows between hidden states represent a transition matrix or function $A_{mn} = p(\mathbf{s}^{t+1} = s_n|\mathbf{s}^t = s_m)$ which assigns the probabilities of switching from state s_m to every other state s_n at the consecutive time instant.[1] Note that $\sum_n p(\mathbf{s}^{t+1} = s_n|\mathbf{s}^t = s_m)$ =1, as probability must be conserved between time instants and thus we assume that the state always transitions to some state in the model and does not exit the set of states. Finally, the arrow leading into the first hidden state s^1 represents a prior distribution for the first hidden state $p(\mathbf{s}^1)$.

Implementation of the HMM requires that these three components A_{mn}, $p(\mathbf{s}^1)$, and $p(e_i|\mathbf{s} = s_k)$, be chosen or estimated. The set of these HMM assumptions are indicated in the conditions of the following probability distributions as $\mathcal{M} = \{A_{mn},\ p(\mathbf{s}^1),\ p(e_i|\mathbf{s} = s_k)\}$.

3.2.1 HMM algorithm

Under these assumptions, we are prepared to examine the probability of being at any state $\mathbf{s}^{t'} = s_k$ at time t' given an entire sequence of observations $[\mathbf{e}^1,\ \ldots,\ \mathbf{e}^T] = \mathbf{e}^{\mathcal{T}}$. The joint probability distribution of the full HMM $p(\mathbf{s}^{\mathcal{T}}, \mathbf{e}^{\mathcal{T}}|\mathcal{M})$ and is shown in Equation 3.3. The posterior probability $p(\mathbf{s}^{\mathcal{T}}|\mathbf{e}^{\mathcal{T}})$ is shown in Equation 3.4. Finally, the posterior marginal probability at time instant t', $p(\mathbf{s}^{t'}|\mathbf{e}^{\mathcal{T}}, \mathcal{M})$, is shown in Equations 3.5 and 3.6.

$$p(\mathbf{s}^{\mathcal{T}}, \mathbf{e}^{\mathcal{T}}|\mathcal{M}) = P_1 \prod_{t=1}^{T-1} A_{t,t+1} \prod_{t=1}^{T} p\left(\mathbf{e}^t|\mathbf{s}^t\right) \tag{3.3}$$
$$P_1 = p(\mathbf{s}_1|\mathcal{M})$$
$$A_{t,t+1} = p\left(\mathbf{s}^{t+1}|\mathbf{s}^t, \mathcal{M}\right)$$

$$p(\mathbf{s}^{\mathcal{T}}|\mathbf{e}^{\mathcal{T}}, \mathcal{M}) = \frac{p(\mathbf{s}^{\mathcal{T}}, \mathbf{e}^{\mathcal{T}}|\mathcal{M})}{p(\mathbf{e}^{\mathcal{T}}|\mathcal{M})} \tag{3.4}$$

$$p(\mathbf{s}^{t'}|\mathbf{e}^{\mathcal{T}}, \mathcal{M}) = \sum_{\mathbf{s}_{\{t \neq t'\}}} p(\mathbf{s}^{\mathcal{T}}|\mathbf{e}^{\mathcal{T}}, \mathcal{M}) \tag{3.5}$$

Due to the Markovian properties and structure of the HMM, Equation 3.5 can be equivalently written as Equation 3.6, written in terms of functions α and β.

[1] One can consider this transition function essentially a lookup table of transition probabilities between every two states, thus this is a lookup table of size N^2, with each entry a probability (a number between 0 and 1).

$$p(\mathbf{s}^{t'}|\mathbf{e}^{\mathcal{T}}, \mathcal{M}) = \frac{\alpha(\mathbf{s}^{t'}|\mathcal{M})\beta(\mathbf{s}^{t'}|\mathcal{M})}{p(\mathbf{e}^{\mathcal{T}}|\mathcal{M})} \tag{3.6}$$

$$\alpha(\mathbf{s}^{t'}|\mathcal{M}) = p(\mathbf{e}^{1}, \ldots, \mathbf{e}^{t'}, \mathbf{s}^{t'}|\mathcal{M})$$

$$\beta(\mathbf{s}^{t'}|\mathcal{M}) = p(\mathbf{e}^{t'+1}, \ldots, \mathbf{e}^{T}|\mathbf{s}^{t'}, \mathcal{M})$$

$$p(\mathbf{e}^{\mathcal{T}}|\mathcal{M}) = \sum_{s^{t'}} \alpha(\mathbf{s}^{t'})\beta(\mathbf{s}^{t'}),\ \forall \mathrm{t}' \tag{3.7}$$

The alpha-beta algorithm (also known as the forward-backward algorithm), efficiently calculates the α and β terms found in Equation 3.6, and a simple summation over all states at any time t' yields the normalization factor in the denominator (calculation shown in Equation 3.7). For each time t', this algorithm calculates the influence of all past to current observations $[1,\ \ldots,\ t'-1, t']$ in α and all future observations $[t'+1,\ \ldots,\ T]$ in β and incorporates these into the posterior marginal probabilities of each of the states at time t'. Because both current and future time is used in each calculation, this is known as a smoothing operation.

The normalization factor on the posterior marginal probabilities $p(\mathbf{e}^{\mathcal{T}}|\mathcal{M})$ is the same at every instant t' in time and is given in Equation 3.7, and is simply the summation over all states of the alpha beta product.

Chapter 4

A probabilistic framework for a physiological representation of dynamically evolving sleep state

4.1 Introduction

All humans sleep and approximately 10% of the population [66] has sleep disorders which include conditions as varied as insomnia, restless leg syndrome, and sleep apnea. Such pathologies, and other temporary conditions that disrupt sleep cause significant problems in individual emotional, psychological, and physical well-being. In turn, such widespread sleep disruptions cause broader economic impacts and demonstrate the significant research that must be conducted to investigate the properties of sleep. Thus a better understanding of sleep dynamics is crucial for developing more knowledge about the fundamental dynamics and treatment of pathologies that interact with sleep. Full measurement of the neurophysiological state of the brain during sleep would require invasive measurements and is poorly suited for typical human subjects research. Thus researchers have long turned to noninvasive measurements of electrical brain signals for traditional sleep staging. Traditional sleep stages were not developed to capture and thus do not capture all of the physiological information and microstructures available in EEG and other electrical signals. However sleep staging does provide a good standard for classifying sleep and serves as a foundation to inform additional more detailed representations of sleep and its microstructure such as those described in [8; 83; 11; 58; 51; 31; 64] and this work.

Recently Lopour et al. [51] worked quantitatively to relate experimentally measured human sleep EEG data to a continuous set of sleep states based on a dynamical mathematical model of the brain which has been extensively studied and developed for sleep, seizures, anesthesia and other applications [76; 77; 79; 43; 79; 80; 75; 81; 91; 89]. This and other dynamical models with phase transitions have been applied to understand better mammalian sleep and anesthesia qualitatively and experimentally; the related literature includes work

exploring the foundations of the K-complex [91], understanding up and down state transitions of cortical oscillations [92], and other applications. For a more complete introduction to these models we refer the reader to [75]. Previous work in directly mapping EEG data to a dynamical model of sleep [51] presented a novel and practical application of a mathematical model; however [51] used a deterministic and temporally static method that did not fully account for noise in the data, nor use temporal information directly.

In this chapter we address these issues and demonstrate a method of mapping EEG sleep data that elegantly handles the stochastic and temporal nature of the problem. We use a Bayesian framework to map sleep states onto a continuous two-dimensional representation of sleep derived from a physiologically plausible dynamical model of the cortex. This continuous representation of sleep allows for inference of physiological information. Bayes' rule allows us to take probability of A given B and use that to calculate the likelihood of B having observed A. In this case, we sample a stochastic dynamical model of the cortex multiple times to estimate the probability distribution of EEG features at each possible brain state in the model. Then upon observing a new EEG signal from an unknown cortical state, we leverage the state-to-EEG distributions within a Bayesian framework to estimate the likelihood that this signal was generated from each brain state given that observation. Thus, with a model that generates EEG-like data, we create a way to map probabilistically a night of sleep epochs onto the continuous space of brain states in the model.

Using a probabilistic framework addresses the uncertainty which appears frequently when classifying noisy measurements into disjoint classes. For example, with a noisy EEG measurement that appears to be 51% likely to be in state A, but 49% likely to be in state B, where A and B could be stages or points in a continuous representation of sleep, a probabilistic framework allows that information to be preserved, while a non-probabilistic framework might force this measurement to be considered state A. Also, after using a probabilistic framework on data, a threshold (or maximum function) can be applied to force the classified data into disjoint sets. Note that this probabilistic approach allows for the use of confidence intervals on classifications.

We next incorporate temporal information into this continuous representation using a Hidden Markov Model (HMM) inference algorithm, which reduces noise when tracking the sleep state through time. We do this based upon the expectation that sleep state at time t is correlated to sleep state at time $t+1$, and assuming that the brain is a dynamical system with a state that evolves through time. This HMM is a more complex statistical model and an extension of the previously mentioned static likelihood calculations, in essence linking the static likelihood calculations together in a temporal string. When evaluating the probability of being at a brain state at time t, the HMM includes knowledge of observations from time t, but also prior and later observations from times before and after t. This is similar to what the human mind does when faced with a bad telephone connection–any particular word may be difficult to understand if heard in isolation, but after a full sentence is spoken, the temporal information combined with grammar allows one to decipher the full sentence.

The purpose of this work is to offer a method for mapping noninvasive electrical measurements of sleep to a representation that offers additional physiological insights. The results

presented here also demonstrate a rich representation of EEG sleep data that offers insight into sleep dynamics, and a visualization of sleep EEG data that can be created in real-time for each epoch or to track the physiological trajectory. Our work both explicitly handles the noise inherent in EEG and the stochasticity of the model, and also incorporates the idea of temporal continuity and uses the time series sequence order as information into the algorithm. These extensions are significant improvements over previous algorithms and offer a more robust framework for this mapping process and future work. This paper demonstrates results that serve to validate the underlying mathematical model by quantitatively linking sleep EEG to states in the model. Due to the general nature of the mapping procedure, this method can be adapted to other sleep models for comparisons between representations. Such a tool will be useful for both theoretical and experimental research.

In this chapter we are solving the inverse problem: given measured signals from the brain, can we identify in a continuous model the sleep state that produced the observation. This work begins with a description of the continuous representation of sleep and the dynamical model underpinning that representation in Section 4.2. Next, in Section 4.3 we explain the foundations of the probabilistic framework, which is followed by a discussion of feature selection in Section 4.4. Extending that framework to incorporate sequential information, in Section 4.5 we introduce the temporal HMM model and methods for tracking sleep states using knowledge of the entire night of sleep, both future and past epochs, at every epoch calculation. Results from the mapping from EEG to the continuous state space with static likelihoods are presented in Section 4.6 most significantly including a demonstration of the separation of slow wave sleep and REM sleep using this representation. Results using the temporal information and HMM information are presented in Section 4.7, showing examples of model parameter trajectories through time for human subject data.

4.2 Sleep Model

4.2.1 Sleep Stages

Sleep is typically measured by recording electroencephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG) time series signals sampled from a subject over the course of a night. Often additional measurements such as eye movement or breathing patterns are monitored, especially for clinical assessment of sleep disorders. Using these electrical signals measured from the brain activity, eye movements, and muscle movements the night of sleep is typically parsed into 20 or 30 second segments which are classified by an expert into the different sleep stages: wake, stage 1 (S1), stage 2 (S2), stage 3 (S3), stage 4 (S4), and rapid eye movement (REM) sleep. For the purposes of this work, slow wave sleep (SWS), will be defined as sleep with slow waves including spindles and thus be comprised of S2, S3, and S4 sleep stages.

These sleep stages are characterized by certain patterns described by classification rules that are apparent in the EEG, EOG, and EMG signals with EEG being of primary impor-

tance in differentiating non-REM stages. The Rechtschaffen and Kales (R&K) sleep stages are based on visual patterns in the time domain and are thus not a complete description of the underlying biological processes during sleep; rather these are useful tools that provide international standards for use in sleep research but leave room for augmentation via new representations. The original work developing the R&K criteria notes in the conclusion that the staging criteria 'should be viewed as a working instrument rather than a statute' [70; 67].

Typical EEG segments from the various stages are depicted in Figure 4.1. Note that there is a significant amount of microstructure to the typical EEG segments, an indication that sleep stages are functional labels only, but do not serve as descriptors of the physiology underlying the sleeping brain.

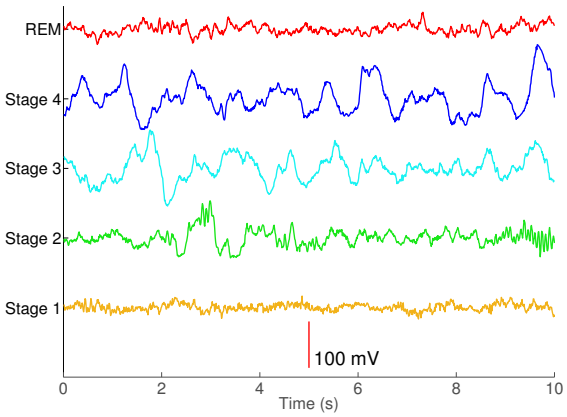

Figure 4.1: A snapshot of typical EEG patterns found in sleep stages. Note the changes in the microstructure as the sleep stages move from light (S1, REM) to slow wave sleep (S3, S4).

4.2.2 Dynamical Cortical Model

The cortical model described in Chapter 2 is used as a platform to create a continuous state space representation of sleep as in [51].

These two sleep parameters L and Δh_e^{rest} define the sleep state in this model.

4.2.3 Sleep Manifold

Given a pair of sleep parameters $(L, \Delta h_e^{rest})$, we can solve for the equilibrium point(s) of the set of partial differential equations given in Section 2.3. These equilibrium points are fixed points in the PDE model in the absence of stochastic forcing. That is, these points are sets of variables $[\tilde{h}_{e,i},\ \tilde{I}_{ee,ei,ie,ii},\ \tilde{\phi}_{i,e}]^{\mathrm{T}} \in \mathcal{R}^{[8\times 1]}$, which cause all time and space derivatives of the PDE equations to be zero given no subcortical input ($\xi_k = 0$). They are called equilibrium

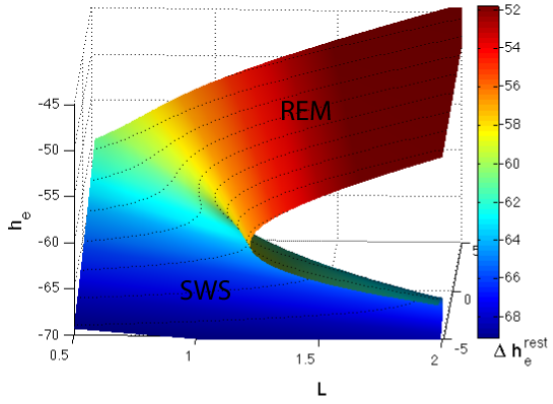

Figure 4.2: Sleep manifold of equilibrium states of h_e over a two-dimensional parameter space. A view of this manifold from above is shown in Figure 4.3.

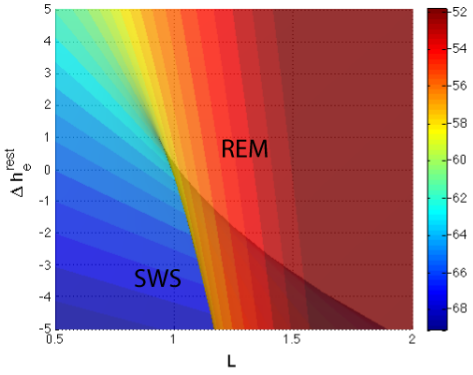

Figure 4.3: Sleep manifold projected into two-dimensional parameter space. The transparency of this manifold highlights the parameter area with multiple equilibrium states. The scale shows the values of the mean soma potential, h_e in mV.

points or fixed points because with no perturbations to the system, the variables do not change if the system of equations is integrated forward in time at such a point.

The exact values of these points depends on the set of sleep parameters and the other parameters defined in Section 2.3. As we vary the sleep parameters (L, Δh_e^{rest}), both the exact values of these equilibrium points and the *number* of equilibrium points change. Thus equilibrium points are not unique at all sleep parameter values. To depict this change, we show one variable, the mean soma potential $\tilde{h}_e$, from these 8-vector equilibrium points plotted over a range of sleep parameters in Figures 4.2 and 4.3. We refer to this surface, shown from two viewpoints in the two figures, as the sleep manifold because it is a two-dimensional surface embedded in three-dimensional space.

From Figure 4.3 it is particularly clear that there is a wedge-shaped region of sleep parameters in which there are three equilibrium points. The theory of this model suggests that the area of multiple equilibria is a location at which a rapid transition occurs from SWS stages (S4, S3) to REM sleep stages [76]. The lines defining this multiple equilibria region will be referred to as the fold lines, and the region as the cusp region due to the cusp point at one end. On several other plots of parameter space, these curves will be drawn to remind the reader of this region.

Over the course of a typical sleep cycle, a qualitative trajectory through this sleep parameter space has been explored conceptually by [76]. Initially, the brain state moves to deeper slow wave sleep gradually, with a rapid jump to REM sleep, and then gradual descent back into SWS. The fast transition from SWS to REM is thought to be due to a fast jump (first order transition) from the lower branch of the sleep manifold to the upper branch of the sleep manifold when the sleep parameters are such that there are multiple stable states in the sleep

manifold (see Figure 4.2) [76]. After this rapid transition, the sleeping cortex evolves slowly through higher Δh_e^{rest} and then back to slow wave sleep via a trajectory down the continuous slope from REM to SWS. Then for the next sleep cycle, this counter-clockwise cyclic behavior of the manifold and in the cortex is repeated.

4.2.4 Sleep hormones and parameters over a typical night

Sleep is a complex physiological phenomenon involving many brain structures. However, EEG measures the summation of voltage fluctuations due to cortical neuronal activity. The mathematical model chosen for this work is parsimoniously based on a section of undifferentiated cortical tissue and is representative of general qualitative aspects of sleep and the effects of the most influential neurotransmitters [76; 51; 91]. Other descriptions of sleep and models that incorporate additional brain structures such as the thalamus can be found in [37; 71].

The parameter Δh_e^{rest} (and to a lesser extent the parameter L) reflects the relative strengths of somnogen and arousal neuromodulators. These are numerous compounds, but for simplicity may be represented by actions of adenosine and acetylcholine. During REM sleep and wakefulness acetylcholine levels are high, which increases Δh_e^{rest}, depolarizing the membrane potential due to shutting of the potassium leak currents. This depolarization causes neurons to be more likely to fire by decreasing the amount of excitatory neuronal input needed to fire. The opposite effects occur when acetylcholine decreases during SWS. This fluctuation in acetylcholine input to the cortex from the brain-stem occurs over approximately 90 minute cycles through the night in humans. Early in the night the adenosine levels are high causing a tendency to neuronal hyperpolarization by opening potassium leak channels [76]; through the course of the night the adenosine progressively decreases, resulting in a progressive depolarizing drift. Thus changes in Δh_e^{rest} primarily show 90 minute cycles overlaid on a slow (about 6hr) depolarization. The sleep parameter L is modestly influenced by the aforementioned neuromodulator milieu, but is more strongly a direct indicator of the processes of sleep-induced synaptic homeostasis. At the start of the night the synaptic strength, and glutamate levels are high and both progressively decrease through the night, if the nature of the sleep pattern is effective.

4.2.5 Model EEG data

The mathematical model of the cortex can generate EEG-like model data via electrode-like measurements of the cortical state gathered during simulations of the full stochastic PDE model [51]. This model EEG-like data is a measure of the dimensionless variable $\tilde{h}_e$ which corresponds to a dimensionless mean soma potential, and is theoretically proportional to measurements that would be made by an electrode on the simulated cortex[1]. For the

[1]In previous work a measure other than $\tilde{h}_e$ was used for feedback control to reflect more accurately an ECoG measurement [50], however for the purposes of this analysis, which does not require electrical input based on precise measurements, $\tilde{h}_e$ will serve well as it does in [51]

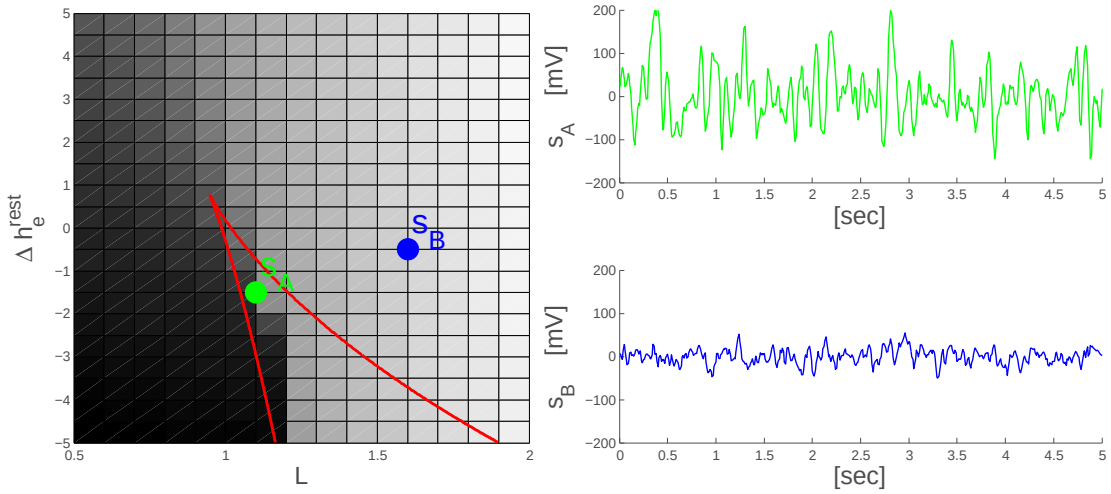

Figure 4.4: **Left:** the state space of the model (equivalently the parameter space or parameter grid) with two example states, s_A and s_B, highlighted. **Right:** examples of EEG-like data from states s_A and s_B in the parameter space. State s_A is a parameter state from the SWS portion of the manifold while state s_B is at a parameter set that contains both REM and SWS equilibrium points in the manifold. The manifold area with 3 stable states is shown in red.

purposes of this work, we will be converting the mean soma potential $\tilde{h}_e$ into millivolts, which will be considered model EEG-like data and comparing features of this model data to features of patient EEG-data measured from human subjects. Figure 4.4 shows two EEG-like signals generated from the model at two different states.

4.3 Probabilistic method for mapping EEG epochs to a sleep state

Problem statement: Given a patient EEG time series over sleeping epochs, find the dynamical evolution of sleep parameters in the cortical model that is most likely to underlie this observed signal. Equivalently, at each time t, find the most likely pair of sleep parameters causing the current EEG output given the entire night of sleep EEG.

These pairs of parameters are referred to as the states of the system, and each state $s_k = [L_k,\ \Delta h_{e,k}^{rest}]$ is a point in the two dimensional sleep parameter space. Figure 4.4 shows the gridded parameter space and sample EEG-like data from two states.

This is an inverse problem–in this algorithm input is an EEG time series and output is the underlying parameters that generated the EEG signal.

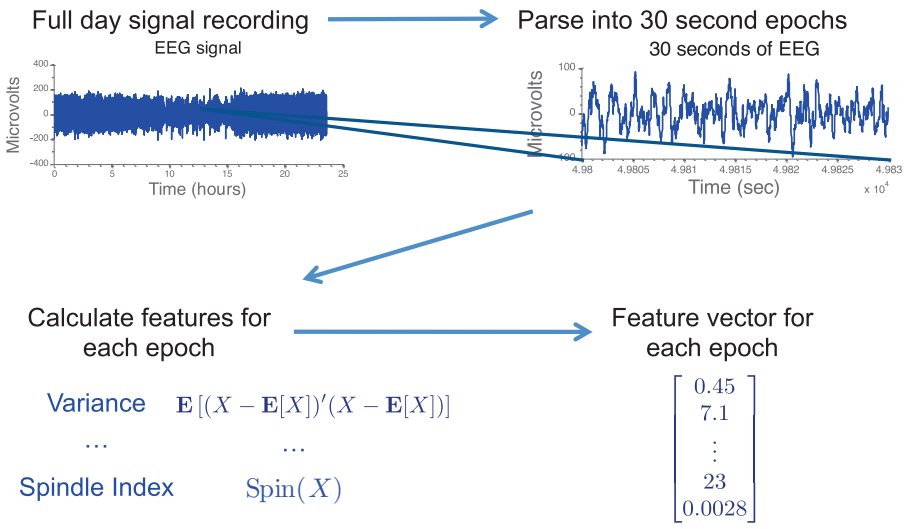

Figure 4.5: A sketch of EEG processing: epoch feature extraction

To address this problem, we use a three-step approach. First, using the mathematical model of the cortex we run multiple simulations at each state and use these to create a probabilistic model of the feature value distributions over the state space. This first step is completed only once to set up the probabilistic framework to map patient EEG data onto the model state space. After this probabilistic model has been created, patient EEG data is parsed into epochs (Figure 4.5) and the next two steps calculate the likelihood of being at each state s_k given the observed data.

Overview of the algorithm for a new patient data set

1. First the EEG signal is segmented into 30 second epochs, and feature vectors $\mathbf{e}$ are calculated at each epoch as shown in Figure 4.5 and described in Section 4.3.2.

2. Next, for each isolated time epoch, the likelihood of being at each model state given the patient feature vector is calculated with the probabilistic framework as depicted in Figure 3.1 and described in detail in Section 4.3.3.

3. Finally temporal evolution of the cortex is incorporated using a Hidden Markov Model (HMM) framework as described in Section 4.5.

4.3.1 Notation

A state in the model s_k within the state space $\mathbf{s} = [s_1, \ldots, s_K]$ is a particular pair of parameters $s_k = [L, \Delta h_e^{rest}]_k$, treating $k \in [1, K]$ as an index to a particular state in the state space of size K, where each state is a 2-tuple. The continuous state space has been

gridded into a finite number of points covering the parameter space of interest (see Figure 4.4). For this application, in which general evolution of parameters is concerned, the grid size is sufficient, however note that the coarseness of the grid can be increased or decreased depending on the application.

Features that are extracted from a particular segment of EEG data are $\mathbf{e} = [e_1,\ e_2,\ \ldots,\ e_M]$, where the letter e represents evidence and can be thought of as an observation of the underlying brain state.

When including the temporal evolution of sleep, a state at time t is $\mathbf{s}^t = s_k$, and the entire state space at time t would be $\mathbf{s}^t$. Similarly, a set of observed features from time t is $\mathbf{e}^t = [e_1^t,\ e_2^t,\ \ldots,\ e_M^t]$. Time over an entire night is also discretized so that there are T time instances $\mathcal{T} = [1,\ \ldots,\ T]$, each separated by time step Δt. The entire set of observations is $\mathbf{e}^{\mathcal{T}} = \mathbf{e}^{[1,\ \ldots,\ T]}$, and the states over the entire set of time is $\mathbf{s}^{\mathcal{T}} = \mathbf{s}^{[1,\ \ldots,\ T]}$.

4.3.2 EEG time series preprocessing

Feature extraction from a time series To compare two stochastic non-stationary time series, it is often not sufficient to compare directly the values measured through time–consider one time series that is then offset by a second compared to itself. Although we might consider these time series almost the same except for an offset of a second, direct comparison of the values measured over each sampling time could easily show no relationship.

Instead of direct comparison, scalar features may be calculated from the time series and then these features may be compared in lieu of comparing time series points themselves. This reduces complexity and parses the data in a straightforward manner. These features can include any number of mathematical operations on the time series that operate on a time series vector and produce a scalar value. In this particular work we consider features that are meaningful to sleep EEG signals such as the power in different frequency bands or a measure of the number of sleep spindles in a segment. A full list of features and descriptions is in Appendix A.1.

The full set of features to extract was initially developed heuristically from patient data and domain knowledge, then a subset of features was refined by measurements of success of mappings to the parameter space as discussed in sections 4.3.3, 4.4, and 4.6.

Preprocessing algorithm The input to this algorithm is an EEG time series measured from a sleeping patient over the course of a night. The first few steps of processing this input into a more useful form are shown in Figure 4.5. As shown in this figure, first the algorithm takes the input of an EEG signal and parses it into T segments of time (epochs), which discretizes the path of physiological parameters. Next, features e_i^t are calculated for each segment time t, and assembled into a feature vector for that epoch $\mathbf{e}^t = [e_1^t,\ e_2^t,\ \ldots,\ e_M^t]$. After feature calculation, each feature is normalized by the root mean square (RMS) over all sleep epochs.

Human subject data source The data used in this analysis is from the Sleep-EDF database, experimentally measured EEG sleep data from four normal human subjects obtained from from a publicly available de-identified data source PhysioBank [28] in the Sleep-EDF Database [40; 39] and are labeled sc4002, sc4012, sc4112, and sc4102 which we will refer to in that order as Subjects 1 through 4.

As an example of the features and results of this time series processing, Figure 4.6 shows two of the calculated features varying over all epochs in the course of the night for Subject 1.

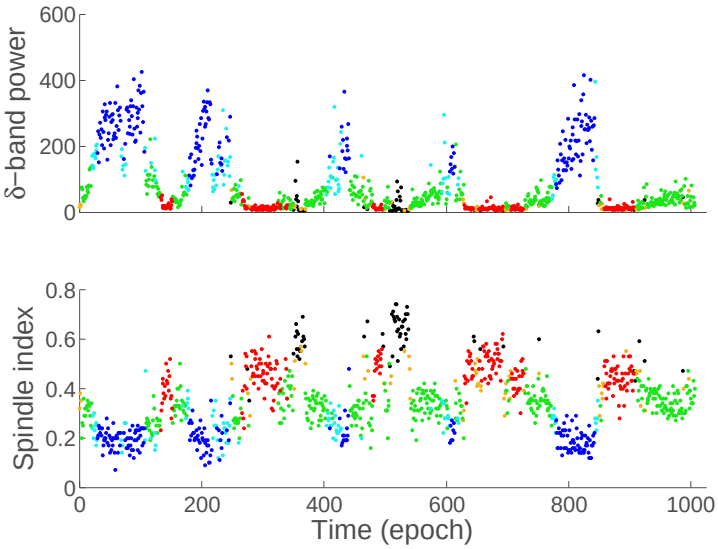

Figure 4.6: Human subject data: subplots show the evolution of two example features over the course of one night of sleep. The colors correspond to different sleep stages to illustrate how features vary between stages. Black, red, yellow, green, light-blue, dark blue respectively indicate wake, REM, S1, S2, S3, and S4. Appendix A.1 contains more information about these features.

4.3.3 Building a static likelihood map from features to parameter state

The cortical model is a stochastic model, meaning that each time it is run at a particular set of parameters, it generates a different EEG-like signal due to the Gaussian noise subcortical input. For a more complete picture of the typical output from a fixed state, the model is simulated multiple times at each parameter pair in the gridded sleep manifold. Features calculated for each of the EEG-like outputs at a single state can be viewed as a histogram, or a sample of the distribution of all features that might be calculated at that state. The

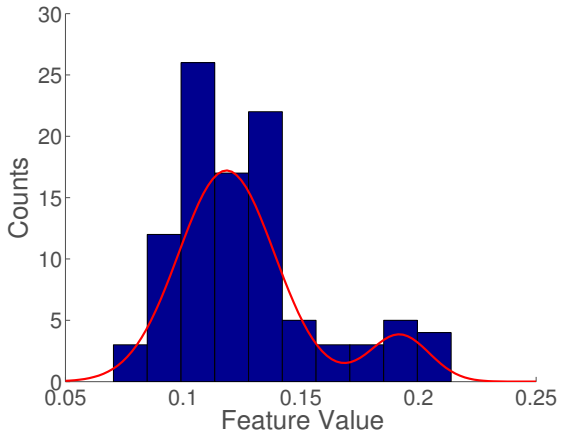

Figure 4.7: Example histogram of one feature at one parameter state generated from many runs of the model with different stochastic input. The red line is the probability distribution function of a Gaussian mixture model curve fitted to this data.

cortical model was run 100 times at each parameter state s_k in the grid. An example of a histogram for a single feature at a single parameter state is shown in Figure 4.7.

Each feature is then approximated with a one dimensional probability distribution function which is fitted to its histogram. A Gaussian mixture model distribution is fitted to each feature histogram individually, such that the probability of observing a feature e_i at state s_k is $p(e_i|\mathbf{s} = s_k)$ shown in Equation 4.1. Variables $\sigma_{h,i,k}$ and $\mu_{h,i,k}$ are the sample variance and mean of feature i at state s_k on the h-index component of the Gaussian mixture model. M indicates the number of components and w_h the weight of each component.

$$p(e_i|\mathbf{s} = s_k) = \sum_{h=1}^{M} \frac{w_h}{\sigma_{h,i,k}\sqrt{2\pi}} \exp\left(\frac{(e_i - \mu_{h,i,k})^2}{-2\sigma_{h,i,k}^2}\right) \tag{4.1}$$

There are many such probability density function estimation methods and for this work a Gaussian mixture model (GMM) was flexible enough to cover the sometimes bimodal feature distributions. The GMM was also used to prevent over-fitting due to the relatively small number of samples–a histogram or kernel method could work as well if the number of samples was increased significantly. The mixture model was limited to contain between one and three components, and the number of components was chosen to minimize the Akaike's information criterion [2]. Choosing up to three components is expected to capture the model behavior as we know that the theoretical sleep manifold has between one and three equilibrium points at any state.

At this point, the system has been simulated multiple times at each state, but to use these feature distributions to estimate probabilities and likelihoods, assumptions must be made about the structure of the probabilistic system. Here we assume that the the graphical model

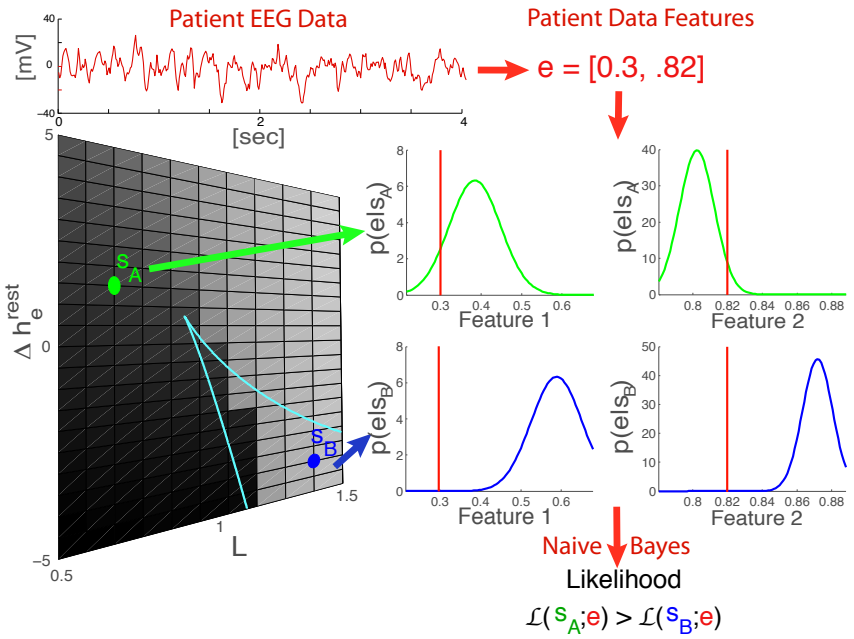

Figure 4.8: At each state s_k in a 2-dimensional state space representation of sleep, sampling multiple stochastic simulations produced probability distributions $p(\mathbf{e}|\mathbf{s} = s_k)$ of observed features $\mathbf{e}$ at s_k. The left grid shows the full state space and two example states that point to example distributions for two features. The top red signal represents a new patient EEG signal from which two features are calculated and red lines show these feature values in the $p(\mathbf{e}|s_k)$ distributions. For each state, the likelihood $\mathcal{L}(\mathbf{s} = s_k; \mathbf{e})$ of being at state k given the features is calculated using a Naive Bayes graphical model (see Figure 3.1) In this particular example, the likelihood of the unknown signal is greater at s_A than s_B.

shown in Figure 3.1 and described in Section 3.1 represents the underlying structure of the probabilistic relationship between the features and the state of the sleep parameters. Note that this structure implies that the features are independent of one another given the state.

We now have a set of probability distributions across each state for each feature. This can directly be used to estimate the probabilities $p(\mathbf{e}|\mathbf{s} = s_k)$ of observing a feature given a parameter state. More importantly, using Bayes' rule, we can leverage this probabilistic structure to calculate the likelihood $\mathcal{L}(\mathbf{s} = s_k; \mathbf{e})$ of being at a state s_k given a feature vector $\mathbf{e} = [e_1,\ e_2,\ \ldots,\ e_M]$ of an EEG epoch generated from an unknown state.

For this simple graphical model the likelihood $\mathcal{L}(\mathbf{s} = s_k; \mathbf{e})$ is calculated using Bayes rule and is shown in Equation 4.2. This corresponds to a Naive Bayes structure, in which each feature is independent given the state, and the result is that the overall likelihood for the

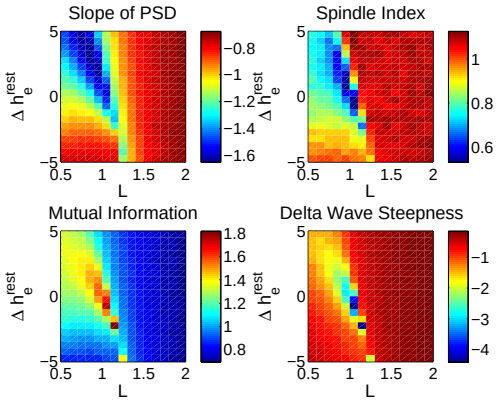

Figure 4.9: Model data: Mean features calculated over the samples at each state in the sleep parameter space.

feature vector is simply the multiplication of the individual likelihoods of each feature.

$$\mathcal{L}(\mathbf{s} = s_k; \mathbf{e}) = p(\mathbf{e}|\mathbf{s} = s_k) = \prod_i p(e_i|\mathbf{s} = s_k) \tag{4.2}$$

$$\mathcal{L}(\mathbf{s} = s_k; \mathbf{e}) \propto \prod_i \frac{1}{\sigma_{i,k}\sqrt{2\pi}} \exp\left(\frac{-(e_i - \mu_{i,k})^2}{2\sigma_{i,k}^2}\right) \tag{4.3}$$

With this probabilistic structure built from samples of model EEG data, we have created the tool necessary to calculate the likelihood at each state of a new EEG epoch. These tools require time-consuming simulations, but must only be created once and then may be quickly and efficiently applied to generate likelihood maps from patient EEG data.

4.3.4 Calculating static likelihoods from an EEG epoch

The question this work seeks to answer is what is the likelihood of being at state s_k given an EEG epoch measured from a human patient and thus generated from an unknown sleep state. The procedure for this calculation is shown in Figure 4.8.

First, features $\mathbf{e}$ are calculated from the EEG epoch. Then Equation 4.2 is used to calculate the likelihoods $\mathcal{L}(\mathbf{s} = s_k; \mathbf{e})$ at every state s_k in the state space. This simply consists of evaluating the GMMs at each feature and taking the product.

The result of this calculation on a single epoch is a likelihood at each grid point that can be viewed via a color plot or surface over the two-dimensional state space.

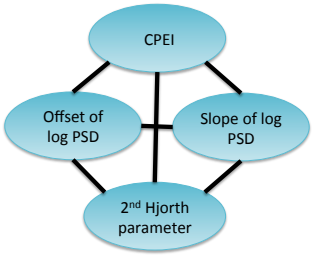

Figure 4.10: An example complete subgraph of redundant features in model space

4.4 Feature selection

Using the entire set of features will not necessarily give a robust or better mapping than using a subset of features. For instance, some features may be noisy, inconsistent between patients, or redundant, and using all features can lead to overfitting, excess complexity, noisy results, and biases. In fact, with this particular algorithm, there are several clear goals that a final feature set should satisfy. At the very least, the features used in this algorithm should meet three requirements. (1) Features should contain meaningful variations between SWS and REM in both human patient data and the model. (2) Features should vary in the same qualitative direction between SWS and REM, which is a first indication as to whether the feature is captured by the model or relies on unmodeled cortical dynamics. (3) The feature set should not contain redundant features due to the Naive Bayes independence conditional on state assumption.

To achieve the optimal set of features under a certain algorithm and cost function criterion, one typically must iterate through every possible subset of features, each time running the full algorithm and evaluating the cost function to find the minimizer. However, the number of possible subsets grows exponentially and this approach becomes intractable for large numbers of features. Good results can be achieved by using heuristics or scoring functions to prune features from the initial set. The full procedure for feature selection is described in Appendix A.2.

The final feature set chosen is the slope of the power spectral density, spindle index, weighted delta wave steepness, and the equiprobable mutual information. Figure 4.9 shows these chosen features varying over the feature set and Appendix A.1 includes descriptions of these features.

Pruning procedure summary

1. Remove features without meaningful variations between SWS and REM in the human subject data using the Spearman correlation coefficient, a nonlinear measure of correlation.

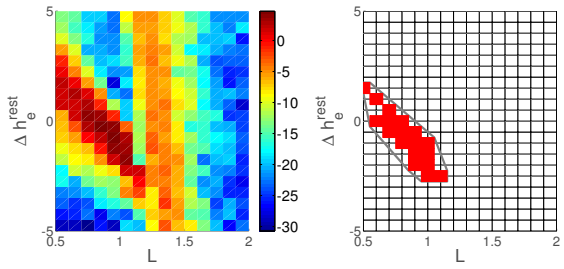

Figure 4.11: The left plot shows an example of a single epoch log likelihood log $(\mathcal{L}(\mathbf{s} = s_k; e_k))$ with the strength of the log likelihood indicated by the colors. The right image then shows shows the 95% likelihood region of the right likelihood map. The states marked by red squares on the right are the states that contain 95% of the likelihood in the left log-likelihood map. All of the unshaded states together contain the remaining 5% of likelihood in the right map. The outline around the red squares is the convex hull of this region, which we use for visualization.

2. Remove features in which the model has different trends than human subject data. This was accomplished by comparing the direction of trends (increase/decrease) of each feature in both the subject data and the model data from SWS to REM.

3. Test for linear redundancy in the features using the Pearson correlation coefficient, and only include up to one feature from any full connected set of correlated features. See Figure 4.10 for an example fully connected set of features found in this algorithm.

4. Do not include feature sets in which any two features are highly correlated given the state.

Evaluating pruned feature subsets The sets generated after the pruning procedure were evaluated with a cost function that penalized the overlap of SWS with REM. This overlap was defined using the overlap of regions in parameter space that contained 95% of the likelihood for average likelihood for each stage. Figure 4.11 shows this procedure.

4.5 Method for tracking the temporal evolution of sleep

The static likelihood mapping developed in Section 4.3.3 and with results in Section 4.6 is applicable to a single segment of EEG, or a single epoch in time, but does not incorporate the expected continuous time evolution of the system or the information from the entire night of sleep signals. That is to say, it does not use the information that might be contained in

earlier and later observations to inform the likelihood at a particular point in time. Up to this point, the likelihood has implicitly been $\mathcal{L}(\mathbf{s}^t = s_k; \mathbf{e}^t)$, the likelihood of being at state s_k at time t given the observation of feature vector $\mathbf{e}^t$ at time t. To incorporate information from all segments of time, the posterior marginal probability of being at state s_k at time t given all observed features throughout the night, $p(\mathbf{s}^t = s_k|\mathbf{e}^{[1,\ \ldots,T]})$, must be calculated.

One way to calculate this type of likelihood is to use a Hidden Markov Model (HMM) as described in Chapter 3.2. Recall that Figure 3.2 shows a graphical representation of a Hidden Markov Model. In this model, there is a hidden state $\mathbf{s}^t$ which cannot be directly observed–this is the parameter state. The observations $\mathbf{e}^t$ are the feature vectors calculated at each time segment, one feature vector per time step up to time T. The arrow from each state to its observation is the probability distribution of observing $\mathbf{e}^t$ given state $\mathbf{s}^t$, and is defined by the time-independent measures we calculated for each epoch: $p(e_i|\mathbf{s} = s_k)$ in Equation 4.1.

We choose the transition function $p(\mathbf{s}^{t+1} = s_n|\mathbf{s}^t = s_m)$ to be a two-dimensional Gaussian distribution in parameter space with a variance of three grid points and mean centered on each originating point, Equation 4.4, effectively favoring nearby transitions[2]. The initial prior distribution $p(\mathbf{s})$ is assumed to be uniform over the entire parameter space. Altering this prior can be straightforward and would not change the methods presented here.

$$
\begin{aligned}
p(\mathbf{s}^{t+1} = s_n|\mathbf{s}^t = s_m) &= p\left([L,\ \Delta h_e^{rest}]_n^{t+1}\ \middle|\ [L,\ \Delta h_e^{rest}]_m^{t}\right) \\
&= \mathcal{N}([L,\ \Delta h_e^{rest}]_n^t,\ \Sigma) && (4.4) \\
&= \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(s_n - s_n)^{\mathrm{T}}\Sigma^{-1}(s_n - s_n)\right) && (4.5) \\
\Sigma &= \begin{bmatrix} .3 & 0 \\ 0 & 1.5 \end{bmatrix}
\end{aligned}
$$

Using this framework and assumptions, the alpha-beta algorithm (also known as forward-backward algorithm), calculates the posterior marginals $p(\mathbf{s}^t = s_k|\mathbf{e}^{\mathcal{T}})$, Equation 3.5. For each time t, this algorithm calculates the influence of all past observations $[1,\ \ldots,\ t-1]$ and future observations $[t+1,\ \ldots,\ T]$ and incorporates that into the likelihoods of each of the states at time t.

4.5.1 Verification of tracking method on temporal evolution of model data

Before investigating the results of the epoch and HMM posterior marginals on patient EEG data, these methods were tested on model EEG-like data, so that comparisons could be

[2]This transition matrix and prior assumption can be relaxed or altered without a significant change to the algorithm. For instance, a transition matrix may be learned using the Baum-Welch algorithm given sufficient data and under certain assumptions [86].

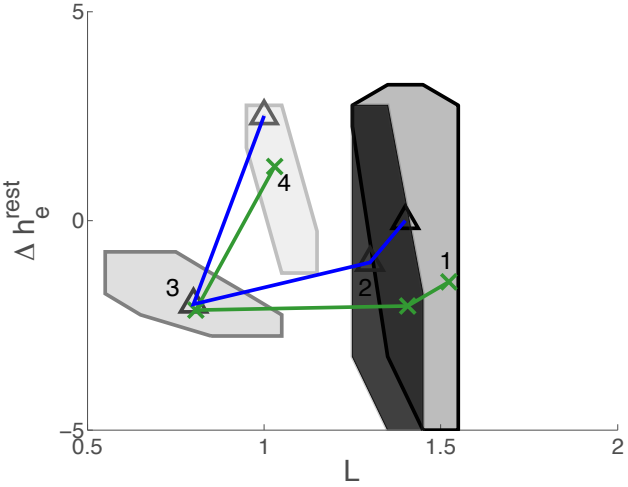

Figure 4.12: Model pseudo-EEG data: using model data, a known path through the state space generated EEG-like data. This true path is shown via the green 'x' marks in the plot. Next, the EEG-like data was put through the algorithm and the HMM posterior marginal at each point was calculated. The black triangles show the path of maximum HMM posterior marginals $\max_{s_k} p\left(\mathbf{s}^t = s_k | \mathbf{e}^{[1,\ 2,\ \ldots,T]}\right)$, and the polygons show the convex hulls containing 80% of the posterior marginal at each time epoch. This a segment of a longer overall path illustrating typical results. The numbers labeling each convex hull indicates the epoch number to clarify the ordering of the convex hulls which progress in time from darkest to lightest gray.

made between the known parameter state of the model data and the results of the posterior maps. Figure 4.12 shows the calculated maximum state points of the posterior marginals for a known path. Note that the maxima depicted are noisy measurements of the full posterior marginal probability maps which have some value across every state, thus convex hulls of the regions containing 80% of the posterior marginal probability are included.

The depicted path is created from a segment of a pseudo-EEG path generated with 800 epochs, each at different parameter states in the model. The posterior marginal probability at each epoch was calculated after applying the method to the entire path of 800 epochs, however only a few epochs are shown for clarity. The probability regions and maximum probability locations show results that are typical throughout the entire path. Several other paths were also generated with similar results. The pseudo-EEG signals underlying these paths were generated using the assumptions made in the HMM, with the only difference being that actual states were not required to lie on grid points, but were allowed to be generated from any location within a continuous space with the same bounds. First an initial state was chosen from a uniform distribution across a continuous state space. Then, for the rest of the epochs, the next state was chosen from a Gaussian distribution of its neighbors as defined

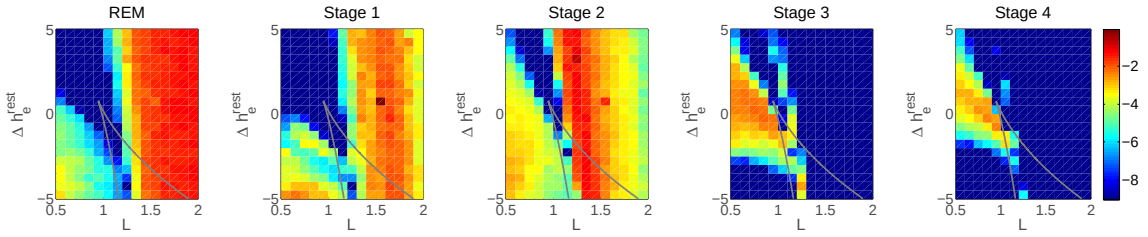

Figure 4.13: Human subject data: these plots depict the logarithm of the average static likelihoods of each sleep stage, using the feature set containing the slope of the power spectral density, spindle index, delta wave steepness, and the mutual information. These plots are an average over all subjects.

by the transition matrix in Equation 4.4. In this way, the purpose was not to investigate a path similar to a typical night's sleep cycle but rather test the efficacy of this algorithm on model-generated data using the assumptions. Generating data from known states is the only way to create a signal in which there is a known 'ground-truth' and testing this assures us that this inverse problem is solvable. By showing that this method works for model data, it implies that it is more feasible that the method may work on patient data and serves as a low-level proof-of-concept.

4.6 Results of probabilistic likelihood mapping of patient data

With the chosen set of features, we now probabilistically map the patient EEG data epochs to the state space using the static likelihood mapping method. After this procedure, for each patient we have a set of likelihood maps, with one map for each time epoch from the original data set.

To summarize the resulting mapping in a way that allows us to leverage accepted sleep knowledge, we examine this mapping in the context of classical sleep stages. We collect the static likelihood maps $\mathcal{L}(\mathbf{s}^t; \mathbf{e}^t)$ of all epochs of all subjects that correspond to each stage and average the likelihoods within each stage to produce average log likelihood maps shown in Figure 4.13 and calculated using equation 4.6. Relative differences in likelihood reflect the relative differences in $p(e|s)$, for instance SWS stages appear to have overall slightly lower likelihoods, perhaps indicating that REM features are relatively better captured by our mathematical model.

$$\log \frac{1}{N_{StageA}} \left(\sum_{t \in StageA} \mathcal{L}(\mathbf{s}^t | \mathbf{e}^t) \right) \tag{4.6}$$

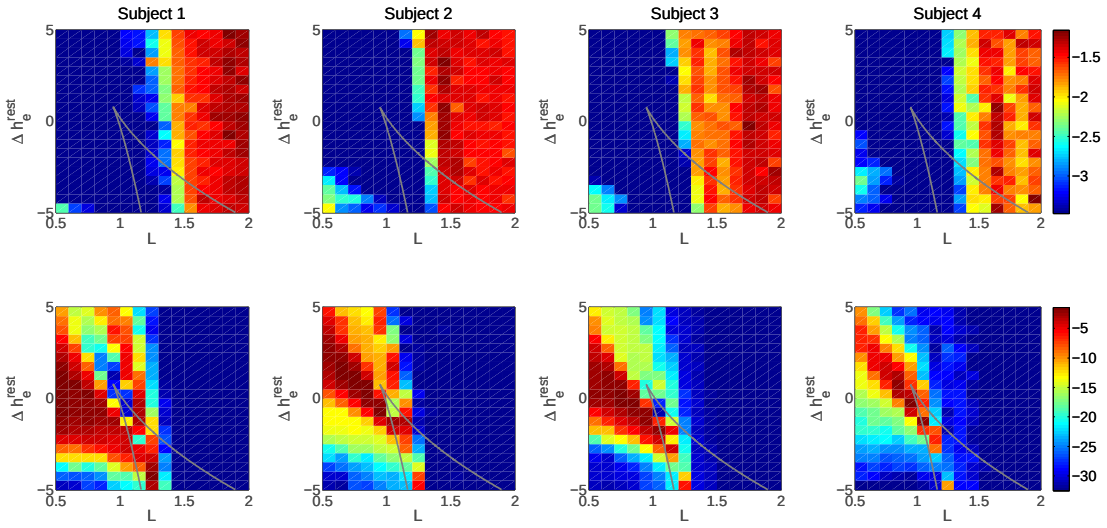

Figure 4.14: Human subject data: these plots depict the logarithm of the average static likelihoods of REM (top row) and Stage 3 (bottom row), using the feature set containing the slope of the power spectral density, spindle index, delta wave steepness, and the mutual information. Each column shows data from a different human subject. These plots demonstrate the consistency in location of stages between different subjects.

Additionally, Figure 4.14 shows the log of the average likelihood under this feature set for each of four human subjects for two representative stages: REM and Stage 3. This demonstrates the consistency of this mapping across different subjects for the same stages, with some minor variations in location that may indicate differences in individual sleep patterns.

As a result of the feature selection process, the average locations of 95% likelihood regions of Stage 4 sleep and REM sleep are fully separated in the manifold. To show that this is robust for individual epochs, which are noisier than averages, Figure 4.15 shows results from all individual epoch mappings. This figure shows the ratio of the numbers of SWS to REM epochs whose maximum likelihoods are mapped to each state. There is very good separation of SWS (which is both Stage 4 and Stage 3 in this figure) and REM maximum likelihood locations, and only a few states in the manifold in which there is overlap. With noise in the EEG and an overlap of SWS and REM expected in certain areas of the manifold, this is quite good separation. This is one of the principal findings of this work. In fact, the fraction of REM overlapping SWS-majority locations is 2.2% (17 out of a total 741 REM epochs) and the fraction of SWS overlapping REM-majority is 4.1% (22 out of a total 529 SWS epochs).

This result for average sleep likelihoods is a successful mapping of measured sleep onto this two-dimensional physiologically meaningful parameterization of sleep. Using just this

probabilistic method on each epoch, EEG sleep can be mapped in real time[3] to this state space and provides a set of information unavailable in typical sleep stages. This could be performed automatically and the results used for sleep studies and investigations into the hormones of sleeping patients. Seizures and sleep disorders could be investigated using this additional sleep information based on a dynamical model for further insights into causes and potential cures.

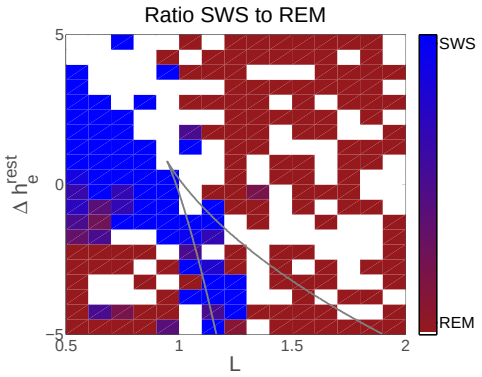

Figure 4.15: Human subject data: this plot demonstrates the robustness of the separation of REM and SWS on the manifold. At each SWS and REM epoch in all subjects, the maximum likelihood state was identified and plotted on this plot. States that were never a maximum likelihood area for SWS nor REM are white. States that were only maximum likelihood locations for REM and SWS are pure red and pure blue respectively. The shade of purple indicates the ratio of the count of SWS epochs that had maximum likelihood at that state to the count of REM epoch that had maximum likelihood at that state. Note that very few states have any overlap.

4.7 Results of HMM temporal evolution of patient data

We show typical results of applying the HMM methods to the data from one human subject in Figures 4.16 and 4.18, demonstrating the temporal evolution of sleep parameters.

The state corresponding to the maximum posterior marginal at a particular time are a coarse and noisy measure of these probabilistic results; thus Figure 4.16 additionally shows

[3]This method is a real-time method in the context of each 30-second epoch constituting an instant of time. However note that minor changes would need to be incorporated into the implementation of this method to achieve effective real time evaluation–primarily, the RMS normalization of the human subject data would need to be calculated using a real-time approximation of the normalization parameters (an approximation of the RMS). We would not expect this approximation to change significantly the results.

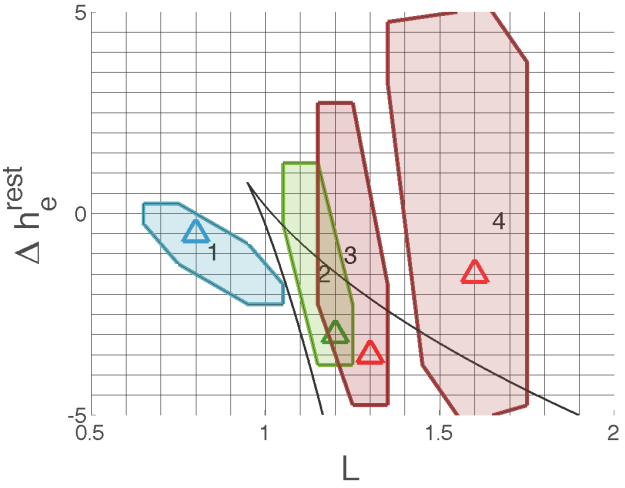

Figure 4.16: Human subject data: HMM posterior marginal convex hulls containing 80% of posterior marginal evolving over time epochs A to D (each Δt between steps is 60 seconds, showing every other epoch) for a single subject. The maximum likelihood locations at each epoch is shown as a triangle of the same color. The colors represent the sleep stage at each time. REM, Stage 2, and Stage 3, are red, green, and blue respectively. This sample set of epochs was chosen to demonstrate an example of a transition from SWS to REM.

posterior marginal outlines of 80% of non-contiguous probability evolving in time for HMM posterior marginals. These outlines encapsulate the maximum posterior marginal and all the highest probability points until 80% of the posterior marginal probability is accounted for at a particular time step. This set of epochs consist of Subject 3 epochs 432, 434, 436, and 438 after sleep onset, which occurred at epoch 1196, these epochs are identified as A, B, C and D respectively. This transition is typical of the movement into REM from SWS and these epochs chosen to capture such a transition.

For comparison to the static likelihood method, Figure 4.17 shows the same epochs using simple static likelihood rather than the full HMM implementation.

For a more comprehensive view of a typical night of sleep, all epochs in Subject 3 data are displayed in Figure 4.18. This figure shows the evolution of the locations (average weighted by posterior marginal probability) in the two coordinates L and Δh_e^{rest} of the HMM posterior marginals colored by sleep stage. It does appear from these plots that major oscillations in L on the order of 100 minutes are present in this patient data as is expected. In particular note that in the top plot containing L through time, troughs occur around epoch 50, then around epoch 150, and finally around epoch 290, similarly peaks occur at around 0, 100, and 230 epochs. With this current data set, early to late night changes are not immediately apparent.

We also show in Figure 4.19, the results without applying the HMM. Note that these

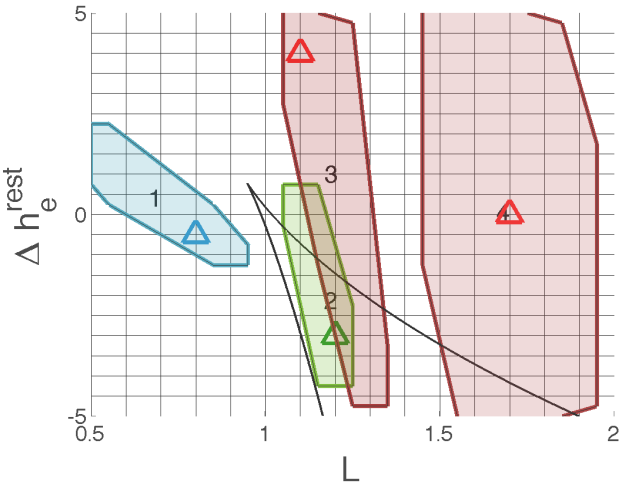

Figure 4.17: Human subject data: for comparison with Figure 4.16 this plot shows the same epochs static likelihood (no temporal continuity assumed via HMM) convex hulls containing 80% of likelihood evolving over time steps. Note that these epoch regions cover larger areas than the HMM likelihood plot. This is an indication of the utility of including the HMM algorithm.

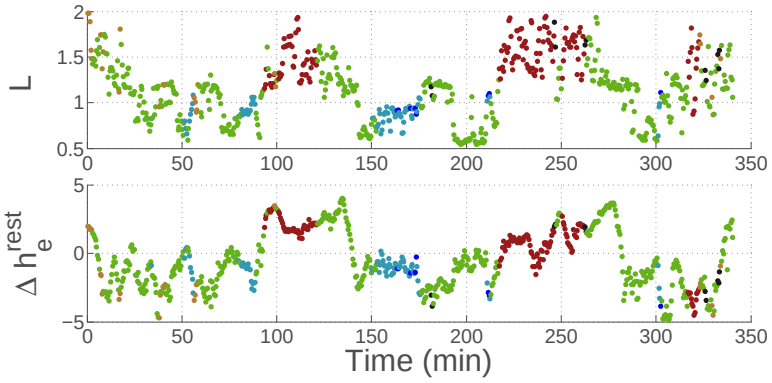

Figure 4.18: Human subject data: HMM posterior marginal centroid locations over entire night of sleep for a single subject. This figure presents the sleep of subject 3, each point indicating an epoch. The colors represent the sleep stage at each time. REM, Stage 1, Stage 2, Stage 3, and Stage 4 are red, yellow, green, light blue, and dark blue respectively.

are simply centroid locations, and as these are weighted averages they are not as drastically impacted by using the HMM as the 2-dimensional plots over each epoch would be. However, one can see that the Δh_e^{rest} is very significantly impacted by the HMM application and that L is somewhat impacted. Recall though that these are portions of a 2-dimensional representation so the overall trajectory is a composition of both.

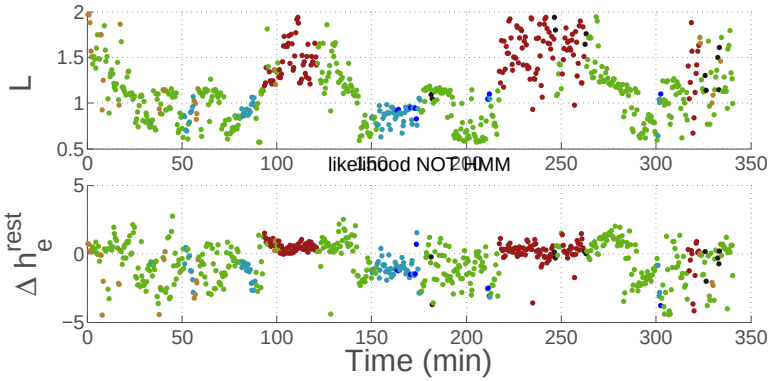

Figure 4.19: Human subject data: Likelihood centroid locations over entire night of sleep for a single subject. This figure presents the sleep of subject 3, each point indicating an epoch. The colors represent the sleep stage at each time. REM, Stage 1, Stage 2, Stage 3, and Stage 4 are red, yellow, green, light blue, and dark blue respectively.

Finally, we turn to a consideration of the placement on the sleep manifold of the sleep trajectory over the course of a night. Although one might expect lower L values over the course of REM stage sleep to emerge, that is not a strong tendency in the present results; in that respect our results agree with [51]. However, we do see indications of decreases in Δh_e^{rest} over SWS when compared with REM. The choice of the regions (and the NREM-REM cycles) in [76; 91] were based on information from other studies that suggested the change in acetylcholine levels would result in predominantly antagonism of the hyperpolarizing (potassium channel opening) effects of adenosine. The effects of acetylcholine to decrease the magnitude of the EPSP are relatively minor (approximately 10%). However these assumptions were partially based on intuition as to what is reasonable. The purpose of the method presented in this paper is to let the EEG data speak for itself and define what regions of the manifold are occupied in a less biased manner. From Figure 4.16 we see that the NREM-REM transition seems to have more to do with the changes in synaptic (EPSP) gain than with the resting membrane potential (which covers a wide range in both states). This is a novel conclusion from this study and would agree with suggestions [22; 18] that the NREM state is associated primarily with a reduction of synaptic EPSP impact (the idea of "cortical block") as compared to the REM and awake states – rather than primarily a change in resting membrane potential mediated by intrinsic currents.

Another note to make here as we examine these results is that it can be seen in Equations 2.1 and 2.2 that L is multiplied by a number of terms (e.g. Γ_e) to control/describe the excitatory activity of the neuronal population. The parameter L is a dimensionless gain term, and there is a possibility that some of the changes in L seen in these results could potentially be influenced by changes in the other multiplicative terms (e.g. Γ_e) that may occur in the complex dynamics of human sleep.

Also, although in the model equations the Δh_e^{rest} and L are independent parameters, it is possible that the values of the two parameters are conflated to some degree. It must also be acknowledged that factors influencing the membrane potential and the synaptic gain are not necessarily physiologically completely independent [32]. However our model does emphasize the importance of changes in synaptic gain in the transition between REM to NREM sleep over the changes in Δh_e^{rest}. Thus the model supports many of the assertions made by Tononi's group about importance of sleep stages in modifying synaptic gain and activity; they have measured both fluctuations in extracellular glutamate and direct measurements of EPSP [49].

4.8 Potential Extensions

While this method was applied to a particular mean-field model, the algorithm described could be adapted for other models of the brain as well, and potentially used to compare models of sleep. In this Bayesian framework, comparing the overall likelihood of a Hidden Markov Model can inform which model is more likely to have generated the data.

In addition, although the HMM posterior marginal calculation uses both future and past data points in its calculation via the alpha-beta algorithm, thus smoothing and filtering the observed data, one could use only past and current data and calculate a filtered likelihood in an online algorithm in real-time using all previous epochs (but not future epochs). This might be useful in applications that are required to be executed in real-time and is essentially an intermediate step in the full HMM calculation (the alpha portion of the alpha-beta algorithm).

Chapter 5

A probabilistic method for determining cortical dynamics during seizures

The goal of this work is to develop a better understanding of the physiology underlying or occurring in conjunction with seizure onset and evolution by leveraging a physiologically meaningful mathematical model of mesoscale cortical dynamics. If we engineer a way to take EEG or ECoG measurements of seizures and learn how the corresponding physiology is changing in a model of the cortex, there is great potential to gain understanding of seizures that among other implications may lead to new treatment strategies for individual seizures or epilepsy as a condition, including such options as new drug development to target certain physiological changes, new implantable therapies, or new algorithms for existing devices [42; 43; 88].

We note that in many patients, individual seizures produce patterns with repeated motifs, indicating a potentially repetitive pathological brain state evolution. If this is the case, then perhaps causes or correlates to that pathological pathway can be discovered via a model and new treatment options developed to render that pathway to seizure impassable, or perhaps speed up such a trajectory so that seizures last only a fraction of a second rather than tens of seconds or minutes. To investigate the physiological state at the root of seizure evolution, we first assume seizures from the neocortex can be modeled with a stochastic nonlinear meso-scale description of undifferentiated cortex, in which the state of the model parameters (e.g. strength of subcortical input or measures of synaptic connectivity strength) can lead to non-seizure or seizure-like brain activity like that which is observed with ECoG measurements. This is effectively an inference problem, and we can use Bayes theorem to decode what is happening in the underlying system (the physiology) after observing what is communicated from this underlying physiology over the noisy channel of the cortical dynamics [53].

There have been several studies in associating cortical states (including seizures) measured in patients with these types of meso-scale cortical models and neural mass models for both seizures [42; 43; 60; 87] and sleep [51; 17]. In [42; 43], pathways to seizure regions

identified with bifurcation analysis through the particular mesoscale model used in this work were explored qualitatively to determine potential directions through seizure and enumerate seizure regions. Using a different but related mathematical model, [60] explores a more quantitative approach to determining parameter pathways, concentrating on time domain waveforms and attempting to match specific waveforms in measured EEG to waveforms generated by a neural mass model. Similarly [87] uses a minimalist model to analyze seizure waves.

With an application to sleep rather than seizures, in [51; 17], sleep states were associated with sleep physiological parameters in such a meso-scale cortical model. The methods presented in this work most closely resemble the methods used in [17] and in Chapter 4 with a new application to seizures, an extension to comparing model regions rather than analysis within a single region, and technical improvements including improved sampling and a simplified feature selection algorithm.

However the eventual goal of the work presented here is aligned closely to one of the goals of [60]: to determine quantitatively underlying parameters. In comparison to [60], we present a probabilistic approach to handle a stochastic model and use a wider variety of features in the feature set encompassing not only time domain features but also frequency domain features, wavelet features, and features of neighboring ECoG channels.

It is also important to note that there are several other areas of active research in understanding the physiology of the brain by leveraging models. Related research and methods include explorations of seizures in mesoscale model under the theoretical impact of certain anesthetic drugs [46; 91; 81; 24] and Dynamic Causal Modeling [26; 56] which looks at coupling of brain areas from neuroimaging time series.

In this paper we present a probabilistic method to map ECoG data onto the physiological parameter states of a cortical model over the course of pre-seizure, seizure, and post-seizure activity on a subject-by-subject basis. In addition, we show that given two possible parameter state regions (among many) in which pathological trajectories could be lying, we can correctly choose the region that is more likely given the observed evidence. First, the state trajectory tracking and parameter region identification is shown to work robustly with model-generated data. Next, a proof-of-concept example is shown with human subject ECoG data containing multiple seizures. The reliability of the region identification and trajectories produced by this method is demonstrated across multiple seizures. Although this work does not test all possible paths to seizure, it demonstrates how this method could easily be extended to multiple seizure models

The principal idea is that if this model of seizures is in some way faithful, and if true seizures come from trajectory paths within the modeled parameter spaces, we can determine along which path seizures evolve and the quantitative physiological trajectory. And a key ingredient is that the parameters are connected to meaningful physiological measures, like the strength of the post-synaptic response from excitatory neurons or the strength of subcortical inputs. That means if we find a parameter region or a trajectory through parameter space, and if the model is faithful, then we can infer something about the underlying physiological evolution.

This work begins with an overview of the cortical model and probabilistic methods which are presented in Section 5.1. Results are then presented in Section 5.2. A discussion of the implications of the results in the context of recent literature is discussed in Section 5.3.

5.1 Methods

5.1.1 Overview of method

In this paper, we first assume that the human cortex and its measured ECoG signals are well-modeled by the chosen cortical model[1] and its ECoG-like output, described in detail in Section 5.1.2. Given this assumption, we can leverage this model to understand how the physiological parameter state of the cortex translates into ECoG observations. Then we can aim to take human subject ECoG observations and solve the inverse problem to infer what underlying parameter states are most likely changing to cause the observed ECoG signal and therefore also infer the underlying physiological states.

We approach this inference problem in several steps. As a preliminary step, we choose two possible cortical parameter regions in which the true seizures may lie and hypothesize that seizures may lie on trajectories generated by one or the other. These parameter regions are sets of physiological values bounded within certain ranges. The purpose of this is two-fold. First, the aim of this paper is to demonstrate that this method works robustly under the assumptions made and also has promise for human subjects research, in which case some of these assumptions do not hold. Thus we do not survey all possible parameter sets in the full cortical model, but rather choose subsets of typical parameters with seizure regions. Secondly, as human subject seizures are investigated, it may be that in different subjects, different parameters (or different parameter ranges) may be of primary importance, and we may be able to classify such seizures into different categories based on which parameter regions are most likely to encompass the trajectory. Keeping multiple parameter regions has the potential to reduce the computational complexity while allowing multiple hypotheses to be compared against one another. In this work we use two planar regions, but one could choose many multi-dimensional regions.

After choosing two parameter regions to investigate, we present a process for mapping ECoG onto either of them via a general procedure similar to that in [17]. To account for the innate stochasticity and noisiness of this system, we simulate the stochastic cortical model multiple times at each model parameter state to estimate the probability distribution functions of ECoG features at each state. Then, when observing features from a new ECoG signal generated at an unknown cortical state, we estimate the likelihood that this signal was generated from each possible cortical state given the observation by using the state-to-ECoG probability distributions in a Bayesian likelihood function. This likelihood mapping is fully described in Section 5.1.5.

[1] Note that other cortical models could be substituted into this methodology with few changes

This likelihood mapping from observed ECoG features to likelihoods over each cortical state treats each ECoG segment as probabilistically independent. However, the ECoG epochs of a trajectory are observed in sequence through time and contain temporal information, which we incorporate by modeling the sequence of underlying cortical states and observed features as a Hidden Markov Model (HMM). Calculating the HMM posterior marginal probabilities of each cortical state at each epoch in time given a sequence of features evolving over time provides a probabilistic trajectory over cortical states. Given two potential seizure parameter regions, we can compare the posterior probability of the hypothesized parameter regions given the data and choose the more probable region. The HMM model and methods are described in Section 5.1.6.

5.1.2 Cortical Model

We use a mean field model of cortical dynamics described in Chapter 2 and developed in [47; 77; 48] as a platform to explore the physiological evolution of seizures. The parameters and variables are also described in Chapter 2 and several references [48; 78]. Simulations of this model produce ECoG-like data, dimensionless variable $\tilde{h}_m$, which we will leverage to understand how physiological states in the model produce brain waves with different properties. A few examples of simulated ECoG-like data (model-generated data) are shown in the right portion of Figure 5.1. In addition, this model has been covered extensively in literature and has proven its ability to model seizures and other brain states such as sleep [42; 79; 90; 46; 7].

5.1.2.1 Submodel parameter planes

We have made a preliminary assumption on the structure of this problem: that the brain is well modeled by this cortical PDE model, and additionally as a proof-of-concept demonstration in this paper, we offer two sub-hypotheses which we will use Bayes theorem to test against one another. For convenience in the writing, we will conflate the hypotheses with the planes in parameter space. $\mathcal{H}_1$ is that all parameters remain at nominal values with the exception of P_{ee} and $\tilde{\Gamma}_e$ which vary within the bounds shown in Figure 5.1. The second hypothesis is $\mathcal{H}_2$, that all parameters remain at nominal values with the exception of $\tilde{g}_e$ and $\tilde{\Gamma}_i$ which vary within the bounds shown in Figure 5.1. In both parameter planes, there are pathways to seizures, however, these pathways have different physiological meanings due to the locations in different parameter planes.

Both parameter planes include parameters that impact excitatory and inhibitory characteristics of the neuronal populations. It is often hypothesized that increases in excitation or decreases in inhibition lead to seizures, and typically anti-convulsant drugs block excitation or enhance inhibition [37].

In $\mathcal{H}_1$, P_{ee}, ranging between 0 and 531.25, is the dimensionless subcortical input to excitatory neurons and higher P_{ee} implies increased excitation in excitatory populations, often thought to be a typical factor in inciting seizures. Increases in P_{ee} may be due to

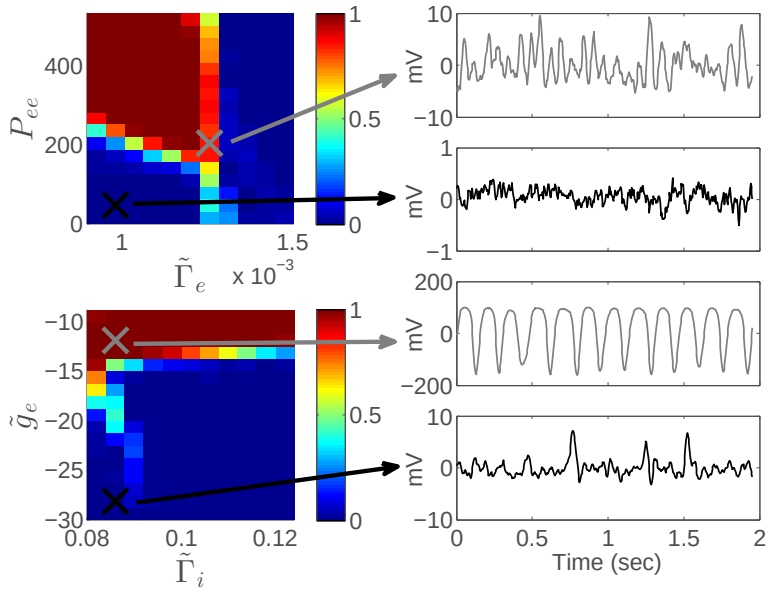

Figure 5.1: Parameter planes $\mathcal{H}_1$ and $\mathcal{H}_2$ top left and bottom left respectively, the color shows the fraction of simulation epochs that contained seizures, red indicates 100% of simulation epochs contained seizures, blue indicates 0% of simulations epochs contained seizures. To the right are two example trajectories h_m from each parameter plane in mV, the parameter state at which they were generated is shown on the left with an 'X' marker.

increases in p_{ee} or decreases in S_{max} in dimensional parameters, described in Table 5.1. $\tilde{\Gamma}_e$, ranging from 9.0e-4 to 1.5e-3, is a dimensionless measure of how input influences the mean soma potential of excitatory neurons and also impacts the excitation. It has a more complex relation to several dimensional parameters shown in Appendix 2.3. Further discussion of these two parameters can be found in [42; 43].

In $\mathcal{H}_2$, $\tilde{g}_e$ ranges between -30 and -8.75 and relates to the sigmoid slope of the transfer function between the average membrane voltage and the firing rate. This change in the firing rate of excitatory neurons is another mechanism to increasing or decreasing the excitability of the cortex[2]. This dimensionless parameter depends on the dimensional parameters g_e and h_e^{rest}, described in Table 5.1. $\tilde{\Gamma}_i$, ranging from 0.08 to 0.124, is a dimensionless measure of how input influences the mean soma potential of inhibitory neurons. By including this parameter, we explore whether changes in inhibition can exacerbate or ameliorate seizure formation.

In each of these parameter planes, we grid the space into discrete states shown as indi-

[2]This particular parameter plane uses the sigmoid curve properties to change the excitatory behavior rather than P_{ee}, a departure from previous literature [42; 43] that yields an alternative set of pathways for exploration

Table 5.1: Selected dimensional parameter definitions. Dimensional parameters are fully described in [78]

Parameter	Description
h_e^{rest}	resting potential
g_e	sigmoid slope at inflection point of function from mean soma potential to average firing rate
p_{ee}	subcortical spike input to neuronal populations
S^{max}	sigmoid maximum firing rate
$\Gamma_{e,i}$	amplitude of excitatory/inhibitory post-synaptic potential
$h_{e,i}^{rev}$	reversal or Nernst potential
$\gamma_{e,i}$	rate constant for excitatory/inhibitory post-synaptic potential neurotransmitters

vidual blocks in Figure 5.1. These discrete parameter states, often referred to as states, are identified as s_k, where k is simply an index identifying the specific state out of all possible states. The goal of this work will be to take a sequence of features calculated from a sequence of epochs and estimate (a) which parameter plane is most likely to have generated that sequence of features and (b) what states are probable to have generated each epoch's data on a particular parameter plane, in other words determine the underlying trajectory through state space that generated the observed ECoG signal.

With the data given and these assumptions, we can compare which hypothesis the data (measured ECoG signal) supports more strongly and thus which parameter plane is more likely to have created the observed data. In the future we can extend these hypotheses to be larger sets of parameters of interest. For now, we are interested in whether one of these parameter planes offers a better explanation for the development of the human subject seizures we analyze from a particular subject and what that may tell us of the pathology and the system.

We first must make sure this is a feasible comparison to make by comparing known trajectories through these parameter planes and making sure they are distinguishable; i.e. that this problem is invertible in the way we pose it and that given correct assumptions this problem produces appropriate answers. Secondly, we test this problem on real data points and determine the more likely set of parameters for an example dataset.

5.1.3 Human subjects data and protocol

To augment a model-based discussion, we consider human subject ECoG data collected from a human patient undergoing routine medically-necessary monitoring at UCSF prior to surgical resection. The data was collected in accordance with UCSF and UC Berkeley human subjects guidelines and proper IRB protocols. A total of 16 seizures were collected over the course of 1 day of monitoring under a 64 electrode ECoG grid over the left frontoparietal region. We used two of the electrode measurements that were identified as pathological

by clinicians. The final clinical diagnosis was ILAE type IIA focal cortical dysplasia using classification system of [6].

5.1.4 ECoG time series processing

The procedure for analyzing the data involves a preprocessing step, then feature extraction, and finally an estimation of probabilistic physiological state.

Preprocessing First we preprocess the ECoG data by removing the 60 Hz line noise and resampling the signal to 256 Hz. We then subtract the mean voltage over the entire time series. We then parse the full ECoG time series into epochs of 3 seconds, overlapping by 40%. This is now a sequence of epochs from $t = 1$ to T.

Features Each epoch is a shorter time series, however, instead of directly comparing the values composing these time series, we calculate a variety of features–scalar measurements of properties of the signal such as variance, powers in different frequency bands, and coherence metrics between two neighboring electrodes. We refer to individual features as e_i, a vector of features as $\mathbf{e} = [e_1, \ \ldots, \ e_N]$, a set of features at time t as $\mathbf{e}^t$ and the sequence of the set of features over all time as $\mathbf{e}^{\,[1, \ \ldots, \ T]} = \mathbf{e}^{\mathcal{T}}$. See Figure 5.2 for a display of several individual features over the course of a seizure.

Probabilistic state Finally, we apply the probabilistic methods described in Sections 5.1.5 and 5.1.6 to estimate the likelihoods and probabilities of this epoch being generated from each parameter state s_k considered.

5.1.5 Likelihood function of state given observations

We assume that for each epoch of observed ECoG data, there is an underlying cortical state that generated that data. The problem is thus how to infer which state in our cortical model is most closely associated with the observed data, knowing that the features observed are an indirect and noisy measurement of the underlying physiology.

To address this problem, we first need to learn how the underlying state translates to the output ECoG signal analyzed by examining features. In probabilistic terms, we need to know the probability of observing different values of feature i given each underlying state s_k, $p(e_i|\mathbf{s} = s_k)$. We generate samples of $p(e_i|\mathbf{s} = s_k)$ by running the cortical model 75 times[3] within each state s_k of interest and calculating features over each epoch.

We then estimate the one-dimensional probability distribution functions $p(e_i|\mathbf{s} = s_k)$ at each state using kernel density estimates[4].

[3]These 75 runs were uniformly sampled within the state region and thus not from a single point in parameter space. Also, 5 overlapping epochs per 10 second run were extracted from each sample.

[4]In particular, the `ksdensity` function from `MATLAB` was used for pdf estimation

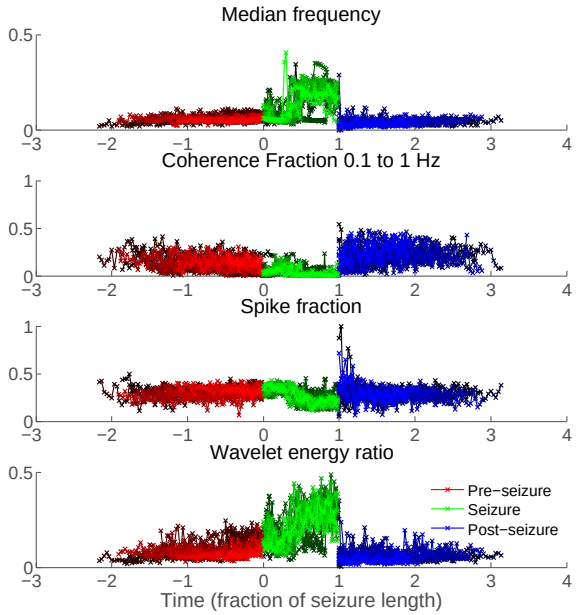

Figure 5.2: This figure shows four features calculated from human subject data over 16 seizures plotted with time scaled by the length of the seizure. Note that the seizure points in green lie between 0 and 1 due to this scaling. These features demonstrate the significant changes in features from pre-seizure to seizure to post-seizure epochs. This is an indication of the usefulness of a feature-based approach to understanding seizures. The four features shown are from top to bottom, the median frequency over the epoch, the ratio of the coherence in the band 0.1 to 1 Hz to the total coherence, the spike fraction (ratio of points outside of 1 standard deviation), and the wavelet energy ratio. Note that the bottom 3 features are part of the final feature set.

Equipped with better knowledge of how the state of the cortical model is communicated to the individual observed features, we now need a likelihood function $\mathcal{L}(\mathbf{s} = s_k; \mathbf{e}, \mathcal{H}_j)$ that takes as an argument a set of newly observed features $\mathbf{e}$ and estimates the likelihood at every state s_k. The likelihood function is defined by the first equality in Equation 5.1, has its foundation in Bayes rule. To find the likelihood of a set of features $\mathbf{e}$, we use a Naive Bayes structure which assumes that the individual features' probability distribution functions are independent. The second equality in Equation 5.1 shows that this assumption leads to a simple product for calculating the likelihood function at each state from a set of features. A summary of our methods to calculate likelihoods is shown in Figure 5.3.

$$\mathcal{L}(\mathbf{s} = s_k; \mathbf{e}, \mathcal{H}_j) = p(\mathbf{e}|\mathbf{s} = s_k, \mathcal{H}_j) = \prod_i p(e_i|\mathbf{s} = s_k, \mathcal{H}_j) \tag{5.1}$$

Evaluating this likelihood function at each state at a time epoch yields a likelihood map

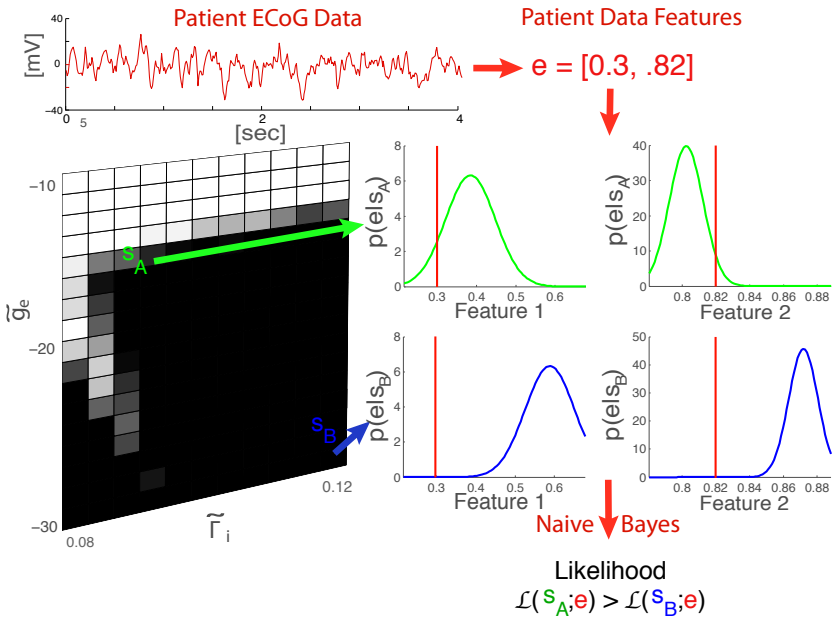

Figure 5.3: This figure shows a summary of the methods of calculating the likelihood of an epoch over a particular parameter plane. Here we show $\mathcal{H}_2$ on the left bottom as the example parameter plane containing states s_k onto which we map epochs. The process is as follows: given a new epoch of ECoG (top left, red), features are calculated (top right, red), then at each state (here two states are shown s_A in green and s_B in blue), the likelihood is calculated using the estimated distributions of features at that state combined with a Naive Bayes structure. In this case, considering only two states and two features, it is clear that the likelihood of s_A would be larger than the likelihood of s_B given the feature set observed.

over the state space that can be used to highlight maximum likelihood regions, to note unlikely regions, or begin comparisons of the relative likelihood of cortical states.

Note that the likelihood function is proportional to the posterior marginal probability $p(\mathbf{s}|\mathbf{e}) = \frac{p(\mathbf{e}|\mathbf{s})p(\mathbf{s})}{p(\mathbf{e})}$ if we assume a uniform prior distribution over all states $p(\mathbf{s} = s_k)$ is constant.

5.1.6 HMM-augmented trajectories

In the calculation of likelihood, each epoch is considered independently, and ignores the fact that the data epochs come from a time sequence. To incorporate the assumption that there is temporal continuity to the underlying physiology, we use a Hidden Markov Model framework as described in Section 3.2.

HMMs require the definition of a transition matrix $A_{mn} = p(\mathbf{s}^{t+1} = s_n|\mathbf{s}^t = s_m)$ and a prior distribution for the first hidden state $p(\mathbf{s}^1)$. For this work, the prior distribution

is assumed to be uniform across the state space. The transition matrix is assumed to be a two-dimensional Gaussian, which is generated by the rules shown in Equations 5.2 and 5.3 to model the idea that closer state transitions are favored over large jumps in cortical state[5]. These HMM assumptions are indicated in the conditions of the following probability distributions as $\mathcal{M}$.

$$p(\mathbf{s}^{t+1} = s_n | \mathbf{s}^t = s_m) = \mathcal{N}([P_{ee},\ \tilde{\Gamma}_e]^t_m,\ \Sigma_1), \qquad \Sigma_1 = \begin{bmatrix} 125, & 0 \\ 0, & 2.2e{-}4 \end{bmatrix} \tag{5.2}$$

$$p(\mathbf{s}^{t+1} = s_n | \mathbf{s}^t = s_m) = \mathcal{N}([\tilde{g}_e,\ \tilde{\Gamma}_i]^t_m,\ \Sigma_2), \qquad \Sigma_2 = \begin{bmatrix} 5.0, & 0 \\ 0, & 1.6e{-}2 \end{bmatrix} \tag{5.3}$$

Under these assumptions, we are prepared to examine the probability of being at any state $\mathbf{s}^{t'} = s_k$ at time t' given the entire sequence of features over all epoch times observed $[\mathbf{e}^1,\ \ldots,\ \mathbf{e}^T] = \mathbf{e}^{\mathcal{T}}$. The joint probability distribution of the full HMM $p(\mathbf{s}^{\mathcal{T}}, \mathbf{e}^{\mathcal{T}} | \mathcal{H}_j, \mathcal{M})$ and is shown in Equation 5.4. The posterior probability $p(\mathbf{s}^{\mathcal{T}} | \mathbf{e}^{\mathcal{T}}, \mathcal{H}_j)$ is shown in Equation 5.5. However, the quantity of interest as we track seizures through time is the posterior marginal probability at time instant t', $p(\mathbf{s}^{t'} | \mathbf{e}^{\mathcal{T}}, \mathcal{H}_j, \mathcal{M})$, which is shown in Equations 5.6 and 3.6.

$$\begin{aligned} p(\mathbf{s}^{\mathcal{T}}, \mathbf{e}^{\mathcal{T}} | \mathcal{H}_j, \mathcal{M}) &= P_1 \prod_{t=1}^{T-1} A_{t,t+1} \prod_{t=1}^{T} p\left(\mathbf{e}^t | \mathbf{s}^t, \mathcal{H}_j\right) \\ P_1 &= p(\mathbf{s}_1 | \mathcal{H}_j, \mathcal{M}) \\ A_{t,t+1} &= p\left(\mathbf{s}^{t+1} | \mathbf{s}^t, \mathcal{H}_j, \mathcal{M}\right) \end{aligned} \tag{5.4}$$

$$p(\mathbf{s}^{\mathcal{T}} | \mathbf{e}^{\mathcal{T}}, \mathcal{H}_j, \mathcal{M}) = \frac{p(\mathbf{s}^{\mathcal{T}}, \mathbf{e}^{\mathcal{T}} | \mathcal{H}_j, \mathcal{M})}{p(\mathbf{e}^{\mathcal{T}} | \mathcal{H}_j, \mathcal{M})} \tag{5.5}$$

$$p(\mathbf{s}^{t'} | \mathbf{e}^{\mathcal{T}}, \mathcal{H}_j, \mathcal{M}) = \sum_{\mathbf{s}_{\{t \neq t'\}}} p(\mathbf{s}^{\mathcal{T}} | \mathbf{e}^{\mathcal{T}}, \mathcal{H}_j, \mathcal{M}) \tag{5.6}$$

Due to the Markovian properties and structure of the HMM, Equation 5.6 can efficiently calculated using the alpha-beta algorithm described in Section 3.2.

5.1.7 Hypothesis comparison

One can use the results of the HMM alpha-beta algorithm to calculate the ratio of the posterior probabilities of the two hypotheses given the data (Equation 5.7).

[5]Although there is transition matrix for each parameter space, these are chosen to have the same effect in coordinates normalized over the range of each parameter plane, in essence the standard deviation covers the equivalent of 4 neighboring states

$$\frac{p(\mathcal{H}_1|\mathbf{e}^{\mathcal{T}},\mathcal{M})}{p(\mathcal{H}_2|\mathbf{e}^{\mathcal{T}},\mathcal{M})} = \frac{p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_1,\mathcal{M})p(\mathcal{H}_1)}{p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_2,\mathcal{M})p(\mathcal{H}_2)} \tag{5.7}$$

Recall these hypotheses are that the seizures lie in one of two parameter planes. That is, we can calculate the posterior probability that the observed features sequence over a seizure was generated by the HMM $\mathcal{M}$ on each hypothesized parameter plane (Equation 5.8) and then compare the two probabilities to determine which hypothesis is more likely to be correct.

$$p(\mathcal{H}_j|\mathbf{e}^{\mathcal{T}},\mathcal{M}) = \frac{p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_j,\mathcal{M})p(\mathcal{H}_j)}{p(\mathbf{e}^{\mathcal{T}}|\mathcal{M})} \tag{5.8}$$

$$p(\mathcal{H}_j|\mathbf{e}^{\mathcal{T}},\mathcal{M}) = \frac{p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_j,\mathcal{M})p(\mathcal{H}_j)}{\sum_{\mathcal{H}_j} p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_j,\mathcal{M})p(\mathcal{H}_j)} \tag{5.9}$$

$$p(\mathcal{H}_j|\mathbf{e}^{\mathcal{T}},\mathcal{M}) \propto p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_j,\mathcal{M})p(\mathcal{H}_j) \tag{5.10}$$

To compare these two hypothesized parameter planes, we must make an assumption about the prior on the two hypotheses. As there is no reason to favor one over the other, we assume that they are equiprobable $p(\mathcal{H}_1) = p(\mathcal{H}_2)$. With this assumption, the ratio simplifies to Equation 5.11 which is the ratio of the two normalization constants under each hypothesis, shown before in Equation 3.7.

$$\frac{p(\mathcal{H}_1|\mathbf{e}^{\mathcal{T}},\mathcal{M})}{p(\mathcal{H}_2|\mathbf{e}^{\mathcal{T}},\mathcal{M})} = \frac{p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_1,\mathcal{M})}{p(\mathbf{e}^{\mathcal{T}}|\mathcal{H}_2,\mathcal{M})} \tag{5.11}$$

With this ratio, we can easily compare the hypotheses and suggest which parameter plane is more likely to have produced the data[6]. We will refer to this as the hypothesis ratio in the following results and discussion.

5.1.8 Selected feature set

The final feature set is a set of seven features, listed and defined in Table 5.2.

Three of these features are shown over the course of human subject seizures in Figure 5.2. The method for determining the feature set is fully described in Appendix B.

[6]Note that this hypothesis ratio is only one way to evaluate and compare models. Often a more involved method called Bayesian model comparison is used to sweep across model parameters and assess model complexity in addition to comparing models [53]. Here both hypothesized planes are of the same complexity so our comparison is valid. For certain types of future analysis one may wish to investigate models with differing numbers of parameters or complexity, in which Bayes factors may be appropriate

Table 5.2: List of features used in the results of this work.

Feature
Log of the ratio of coherence in the band containing frequencies 15 to 26 Hz and 100 Hz and above
The fraction of the coherence in band from 0.1 to 1 Hz compared to the total coherence across all frequencies
The correlation coefficient of the Hilbert amplitude between channels at 0 time delay.
The auto-bicorrelation at a delay of 4 time steps (1/64 sec). Implemented with the MATS toolbox described in [84].
The wavelet energy ratio of the second leaf in a wavelet packet decomposition using a Daubechies wavelet, with 4 levels[7] Including this feature was inspired by features in [44]
The spike fraction, measured as the number of points over 1 standard deviation in amplitude within the epoch

5.2 Results

5.2.1 Demonstrating method effectiveness with model data

Before applying these techniques to human subject data, we must demonstrate that this method yields good results when run purely on model data. With simulated model data we know the true underlying cortical states and can demonstrate typical results and quality metrics under the conditions regulated by the assumptions. The goal is to assess whether it is possible to ascertain the correct parameter plane and path from observed feature sets.

Two trajectories of 1000 epochs were generated using the transition matrices between states shown in Equations 5.2 and 5.3. The first trajectory Tr_1 was generated from states within parameter plane 1 $\mathcal{H}_1$, and the second Tr_2 from within $\mathcal{H}_2$.

With these trajectories, we show metrics of the success of this algorithm. First, we show that the ratio of the posterior probability of the two hypotheses using this data favors, as expected, the true parameter plane underlying the trajectory.

For the full 1000 epochs sequences, Tr_1 is predicted as exp(3270) times more likely to have been generated from $\mathcal{H}_1$ than $\mathcal{H}_2$, and Tr_2 is predicted as exp(2127) times more likely to have been generated from $\mathcal{H}_2$ than $\mathcal{H}_1$. In addition, we test shorter and more diverse sub-trajectories, and show the results as fractions of runs in which the hypothesis ratio correctly identifies the generating parameter plane in Table 5.2.1. The hypothesis ratio is tested for the following cases (i) 10 sequences of 100 epochs, (ii) 100 sequences of 10 epochs, (iii) 20 shuffled sequences of 50 epochs, (iv) 100 shuffled sequences of 10 epochs.

These results show robust reliability in identifying the correct parameter plane underlying a simulated trajectory. Even in the case of shuffled data, in which we expect that our assumptions about the HMM transitions to be incorrect, we can reliably identify the correct parameter plane that generated the simulated data.

Having shown the ability to identify correctly the generating parameter plane, we next

Plane	(i)	(ii)	(iii)	(iv)
$\mathcal{H}_1$	1.0	1.0	1.0	0.95
$\mathcal{H}_2$	1.0	1.0	1.0	0.85

Table 5.3: Fraction of correct identification of underlying parameter plane that generated the simulated data under different cases.

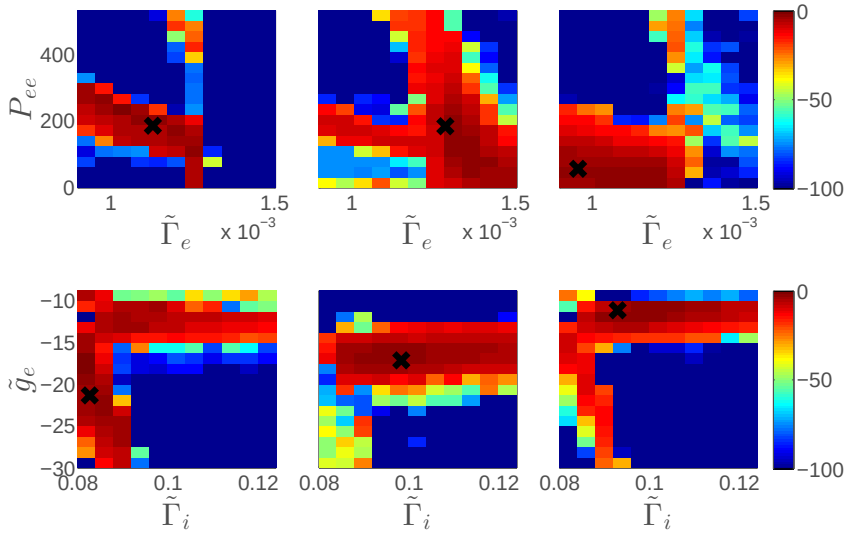

Figure 5.4: Simulated model data from known parameter states are processed using the algorithms described in this work to demonstrate typical results under known conditions. Each plot shows results from a single epoch: the true parameter location that was used to generate the model data in a simulation is shown as a black 'X', and the underlying colormap is the log posterior marginal probability created using the methods described above. The top row shows typical results from simulations within $\mathcal{H}_1$, and the bottom row shows typical results from simulations within $\mathcal{H}_2$. Note that the most probable states are localized around the true state that generated the epoch data.

test the quality of the posterior marginal probabilistic trajectories by comparing these to the true state trajectories. Table 5.2.1, shows for each parameter plane the following metrics: (a) the average normalized distance from the centroid of the HMM posterior marginal probabilities at an epoch to the true state of the generated data and (b) the mean fraction of states that had lower probability than the true state. Normalized distances are calculated by dividing the absolute parameter differences by the ranges of the parameters before calculating the 2-norm.

Typical trajectories and log posterior marginal results from a few epochs are shown in Figure 5.4 for two different trajectories, one in each parameter plane. Note that although

Plane	(a) Distance	(b) Fraction of States
$\mathcal{H}_1$	0.15	0.95
$\mathcal{H}_2$	0.20	0.92

Table 5.4: This table shows two metrics of the quality of the probabilistic trajectory produced using the methods described above.

only one parameter plane is shown for each known trajectory in these figures, log-probability maps are generated for both parameter planes.

5.2.2 Results on human subject data

Having shown that with model-generated data from known state trajectories, the algorithm correctly choses the proper plane and produces reliable trajectory tracking, we test the algorithm on human subject data.

The data for subject 1 contained 16 seizures. For the purposes of this work, pre-seizure and post-seizure epochs of up to 2 minutes were included. For each individual seizure, the HMM algorithm was run to generate the posterior marginal probabilities. Seizure boundaries were assessed by a board-certified clinical neurophysiologist[8].

5.2.2.1 Hypothesis Ratios

For this feature set, we find that for all individual seizures, $\mathcal{H}_1$ was the more likely parameter plane. The hypothesis ratio (Equation 5.11) ranged from exp(469) to exp(107) with a mean log hypothesis ratio of 303.

5.2.2.2 Probabilistic Trajectories

The average locations of pre-seizure, seizure, and post seizure epochs for this subject is shown in Figure 5.5 for both parameter planes. Although we have shown that under our assumptions the evidence supports that these seizures are more likely to have been generated in the parameter plane of $\mathcal{H}_1$, we show results from both hypothesized planes.

To augment the average locations of seizure epochs and delve into the dynamic evolution throughout seizures, Figure 5.6 shows one metric of the estimated evolution of seizures. In this figure, the centroid position of the posterior marginal probability in parameter space is used as a representation of where in parameter space the epoch probability lies. Of course, these centroids are only averages of the posterior marginal probabilities and do not have the detail of full mappings across all states, but they provide a useful visualization to detect large trends over all the seizures.

[8]Dr. Heidi E. Kirsch at the University of California, San Francisco assessed these seizure boundaries.

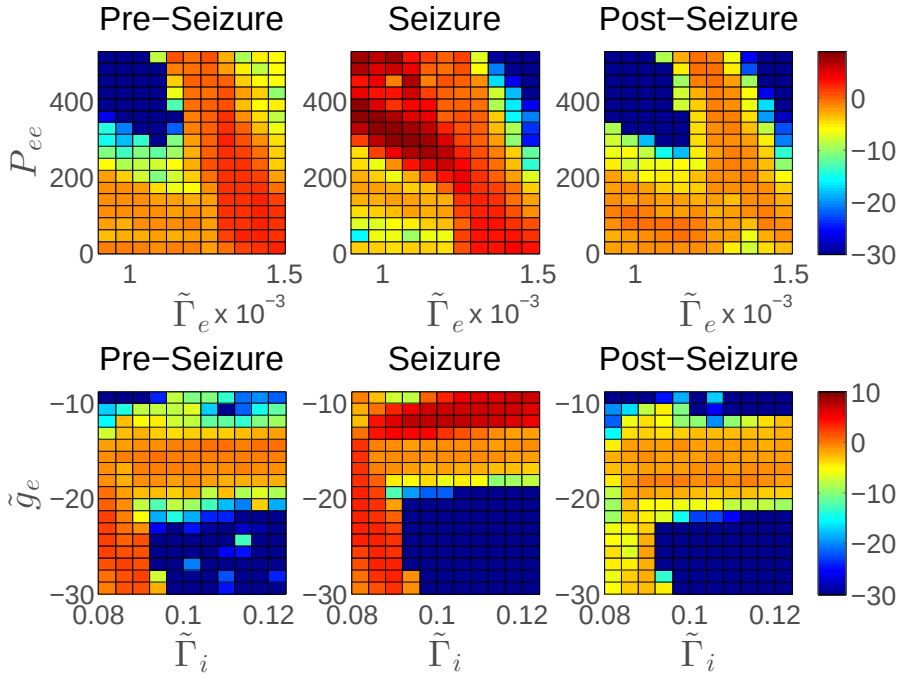

Figure 5.5: Summary of likelihood locations for parameter plane $\mathcal{H}_1$ (top row) and $\mathcal{H}_2$ (bottom row). The information displayed is the logarithm of the average likelihoods of all epochs falling into three different categories: pre-seizure, seizure, and post-seizure.

Additionally, Figure 5.7 shows a brief summary of how pre- and post- seizure locations change over the course of the epochs.

In summary, we do not know that this parameter plane is where the physiological seizure trajectory lies for this human subject data, but of these two parameter planes, $\mathcal{H}_1$ is the more likely.

5.3 Discussion

5.3.1 Implications of clinical pathology

Focal cortical dysplasia is one of the most common types of refractory epilepsy pathologies [14; 36] and is characterized by abnormal cortical cells and malformations of the cortical tissue [62]. Type IIA cortical dysplasia is classified according to [6], and presents as abnormal dysmorphic neurons, lack of cortical layer differentiation, and blurring between the gray and white matter transitions. In particular, the neuronal changes often include increased diameter of both cell nucleuses and pyramidal or interneuron cells, changes in Nissl substance, and changes in neurofilaments [6]. Although the histopathological features of cortical dysplasia

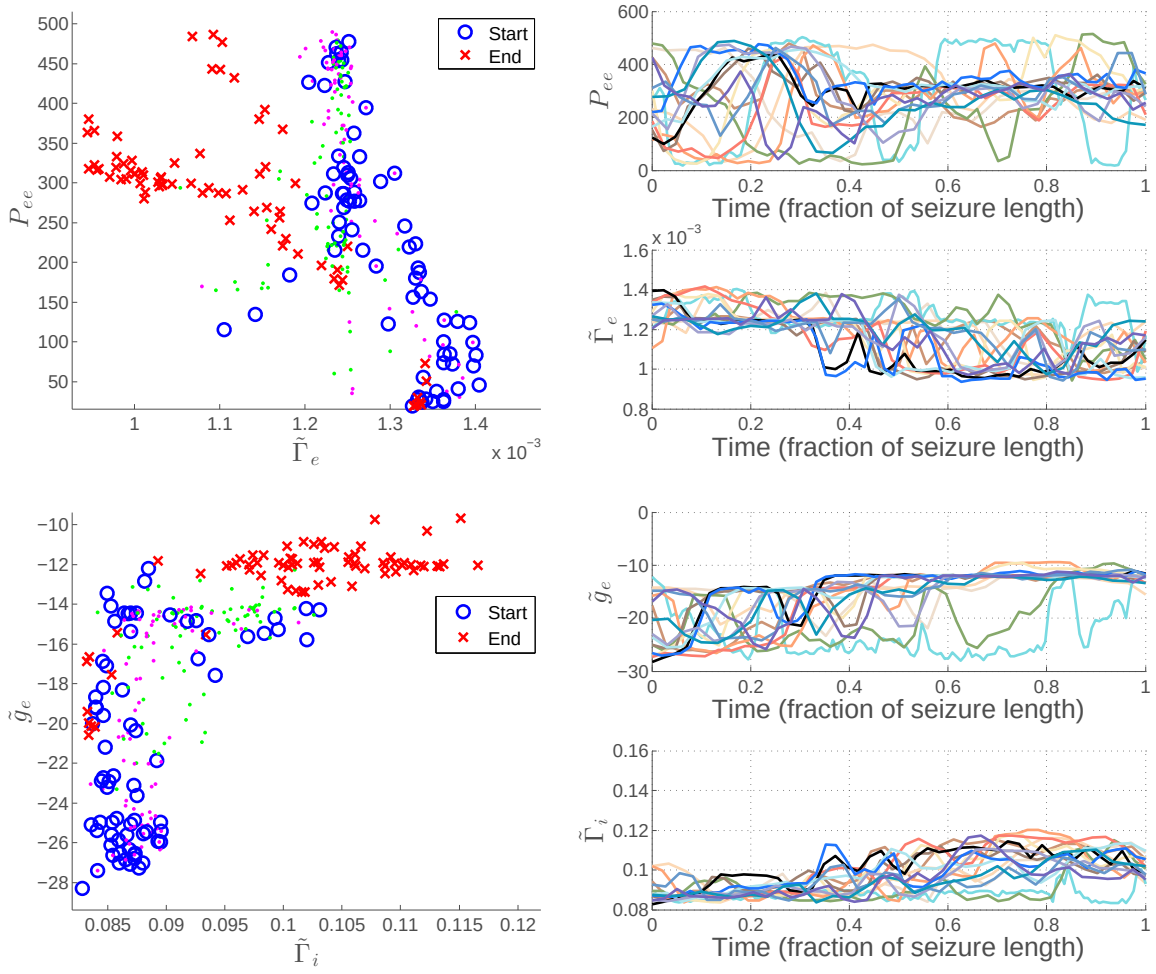

Figure 5.6: Left plots show the first and last 10% of epochs plotted as posterior marginal probability centroids in $\mathcal{H}_1$ (top) and $\mathcal{H}_2$ (bottom). Note that start points (blue) and end points (red) appear to be concentrated in different regions of the parameter planes. The magenta and green points show the centroid of the posterior marginal probability just before and just after the secondary generalization of the seizure respectively. These points were included to investigate whether significant changes occurred in parameter space at the onset of the secondary spread of the seizure to regions beyond the focus. On the right, two plots per parameter plane show the temporal evolution of seizure epoch posterior marginal probability centroid positions of each parameter, where time has been scaled according to the length of the seizure. With these two plots, one can see large differences between the beginnings and ends of seizures and the typical evolution in parameter space over the course of seizures.

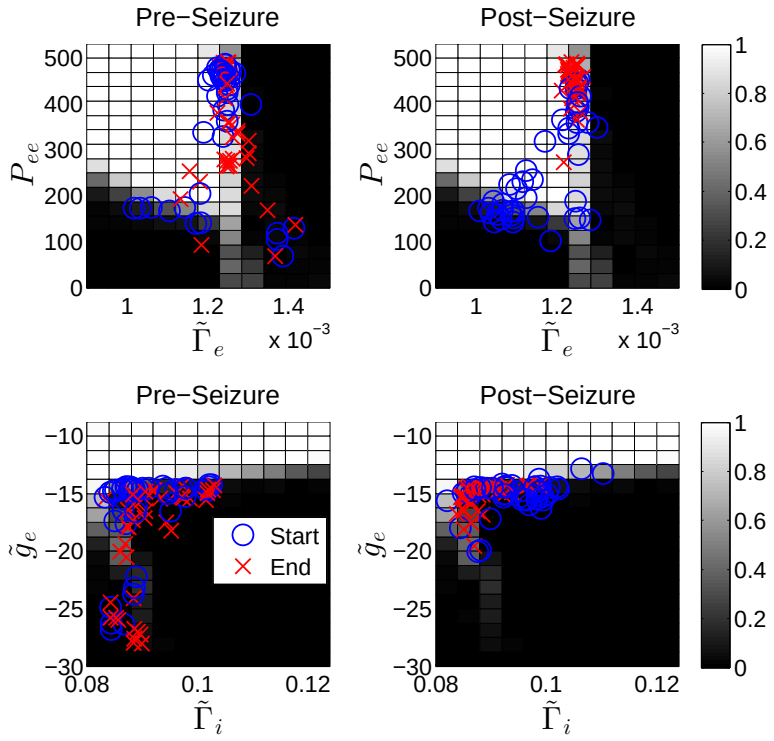

Figure 5.7: Left plots show the first and last 5% of pre-seizure epochs plotted as posterior marginal probability centroids in $\mathcal{H}_1$ (top) and $\mathcal{H}_2$ (bottom). Right plots show the first and last 5% of post-seizure epochs. The background colormap shows the fractions of seizure epochs located in each state region, where 1 is all seizures and 0 is no seizures. Note that start points (blue) and end points (red) appear to be concentrated in different regions of the parameter planes for $\mathcal{H}_2$ (bottom) pre-seizure epochs, but seem to remain localized in one region for post-seizure epochs.

are well known, it is not generally well understood what causes the abnormal tissue to lead to seizures [14; 38] although there are many diverse theorized mechanisms described in literature including hyperexcitability of certain abnormal cells, reduction in inhibitory neurons, and abnormal and aberrant synaptic connectivity [14; 30; 15; 85].

We ask how these changes in histopathological features may translate into the mesoscale model and trajectories found. There may be many ways, and we discuss potential implications of pathways through both parameter spaces.

5.3.1.1 Parameter Plane 1

In $\mathcal{H}_1$ results shown in the top row of Figure 5.6, seizures evolve from higher $\tilde{\Gamma}_e$ counter clockwise to lower $\tilde{\Gamma}_e$, along the bottom edge of seizure regions. At times just preceding

the secondary spread of the seizure (magenta points), parameters appear to be in the same locations as when the seizure began, while in times immediately post-generalization (green points), there is a slight decrease in $\tilde{\Gamma}_e$, perhaps coupled with a slight increase in P_{ee}. There is some noisy evidence that trajectories exit seizure regions at lower $\tilde{\Gamma}_e$ and P_{ee} than at entry, as shown in top row of Figure 5.7 (comparing early post-seizure and late pre-seizure points), although it appears that in the post-seizure time frame, there may be a migration to higher P_{ee} and $\tilde{\Gamma}_e$ over time (compare locations of blue (starting 5%) and red (ending 5%).

The importance of P_{ee} in these seizures may demonstrate the role of subcortical brain structures in initiating or correlating with initiation of seizures via changes in dimensional variable p_{ee}. Perhaps an increase in excitatory input from other brain substructures is correlated with one of the mechanisms that cause the abnormal region of cortical dysplasia to enter a seizure state. An alternative interpretation may be that the connections from the cortical dysplasia to the subcortical structures are abnormal and there may be a constant state of higher subcortical excitation to these regions (note that the last 5% of post-seizure states seem have higher P_{ee} than the initial 5%). Note that another explanation is via the dependence of P_{ee} on S^{max}, and thus these changes seen in P_{ee} could be due to abnormal changes in the maximum firing rate, perhaps related to cortical dysplasia.

Even more noticeable are large trends in $\tilde{\Gamma}_e$, which could arise from a number of dimensional physiological parameters, as shown in its definition in Table 2.2. A decrease in $\tilde{\Gamma}_e$ could be caused by a decrease in Γ_e or S^{max}, or by an increase in γ_e or the quantity $|h_e^{rev} - h^{rest}|$, all of which are defined in Table 5.1[9]. As this is a measure of influence of the excitatory synaptic inputs into mean soma potentials, changes in the cortex due to cortical dysplasia synaptic connectivity may be associated with these large swings in excitatory synaptic strength parameters.

Trajectories across seizures show marked trends to reductions in $\tilde{\Gamma}_e$ over the course of the seizure, agreeing with the qualitative findings of [42] for the subject data analyzed in that report. This agreement using different features and a different method is interesting and may imply that there may be a consistency to mechanisms or correlates of mechanisms of seizure evolution.

5.3.1.2 Parameter Plane 2

In $\mathcal{H}_2$, we see in the bottom row of Figure 5.6 that seizure evolution correlates with a trajectory that begins at a lower $\tilde{\Gamma}_i$ and $\tilde{g}_e$, and then over the course of seizure moves to higher values of $\tilde{\Gamma}_i$ and $\tilde{g}_e$.

The increase in $\tilde{g}_e$ over the course of seizures may relate to the abnormality in the cells in cortical dysplasia regions. The parameter $\tilde{g}_e$ is related to the slope of the sigmoid function that determines the average firing rate given the mean soma voltage. Higher $\tilde{g}_e$ (less negative) implies lower g_e (less positive) in the dimensional model or lower h_e^{rest}. Here we focus on the potential implications of lower values of g_e which indicates a less steep slope in the

[9] In some sense this shows the power of the dimensionless approach, as any of these changes could lead to the change in the dimensionless parameter associated with the observed dynamics.

sigmoid function. When related back to the basic properties of the neurons composing the neuron populations, this sigmoid slope may be thought of as a product of the variance in the firing rate thresholds in individual neurons–lower g_e implying higher variance. In fact [55] suggests that g_e may be roughly inversely proportional to the variance in the firing rate thresholds of the neuron, but may also be due to the diversity of neuronal states in underlying neuron populations. Perhaps seizure onset is correlated with a more synchronous state (lower variance, higher g_e, lower $\tilde{g}_e$), which then over the course of seizure evolution gives way to more variance in the neuronal states or firing rate thresholds by the end of seizures (lower g_e, higher $\tilde{g}_e$). It is possible that this occurs just after the onset of secondary spread (green points), perhaps with neuronal variance occurring due to the interaction of pathological tissue and normal tissue.

In addition, we see increases in $\tilde{\Gamma}_i$ over the course of seizure epochs, which could be caused by an increase in Γ_i or S^{max}, or by a decrease in γ_e or the quantity $|h_e^{rev} - h^{rest}|$, defined in Table 5.1. Many of the histopathological characteristics of cortical dysplasia could influence these different parameters contributing to the dimensionless influence of input on the mean inhibitory soma potential. It is also a possibility that this increase in inhibitory synaptic activity is an indication that the cessation of seizure is correlated to or caused by increases in variability or inhibitory activity.

These discussions of cortical dysplasia with the results shown in the mapping from human subject data to these parameter planes represent just a few preliminary thoughts on links between this meso-scale model and seizure pathologies that may be further investigated with additional analysis.

5.3.2 Potential applications of methodology

We propose this type of analysis as a way to leverage a mathematical model of the cortex to understand better the physiology of seizure evolution. For instance, in our example human subject data, we see certain consistent markers over seizure evolution in the parameter pathways which suggest possible interpretations of the underlying physiology leading to the changes seen. Perhaps further study may show that the general pathway in parameter space may be prevented from occurring under the proper application of novel drug or other treatments, or perhaps the cycle could be accelerated so that the duration of seizures is short enough to avoid impacting quality of life. The idea that such tools could provide the direction and incentive to develop new treatment options is a powerful motivator for this research.

Additionally, as we find parameter pathways such as this one in the model parameter space, we may be able to incorporate theorized or empirical models of drug action or other stimuli (electrical, optogenetic) that would either mitigate the brain wave behavior (for instance electrical inputs as in [50]), or alter the cortical parameter state itself (pharmaceutical, optogenetic inputs [74], etc. may change neuron population behaviors) and test the efficacy in model space.

Finally, although we have discussed this as a testbed for pathway analysis and understanding specific classes of seizure etiologies, in the future it might be useful for diagnosis and classification of etiology. In particular, if this method is adapted for EEG signals, it may augment non-invasive diagnosis and help with determining the optimal choice from treatment options.

5.3.3 Practical implications

Note that the HMM algorithm is a smoothing operation and thus cannot be performed in real time. However the likelihood calculations, a precursor to the HMM runs, can easily be implemented in real time to track roughly trajectories of seizures as they evolve[10].

5.3.4 Limitations of analysis

For the work with human subject data, our two parameter planes may be too limited to capture the true physiological evolution for this subject's seizure; however it provides a necessary proof-of-concept to show the potential for the application of this method. With success and consistency at this level, we propose that in the future pathological parameter regions of higher dimensionality or specific sets of parameters corresponding to a particular hypothesized pathology may be tested by leveraging this mesoscale model and examining what parameter regions and pathways are most likely. In fact, one need not compare multiple regions at all–the pathway tracking results could be valuable in studying seizure evolution through a particular region of interest in parameter space.

[10]Note that in addition, the intermediate component α of the alpha–beta algorithm for calculating the full HMM posterior marginal probability, can be used as a filtered version of the likelihood and run in real time.

Chapter 6

Conclusion

6.1 Summary of results

In this work we have explored the intersection of experimental cortical measurements and a physiologically meaningful mathematical model of cortical dynamics. In two different applications areas, sleep and seizure, we have explored how signals measured from human subjects (EEG and ECoG) can be related to a physiologically based model of cortical dynamics.

First, given a signal measured from the brain (EEG), we have successfully demonstrated a method that infers the cortical evolution in terms of physiological parameters in a continuous two-dimensional representation of sleep. This procedure can be performed in real time using a static likelihood method, or post-data-collection using a HMM posterior marginal that accounts for temporal information. The mappings were shown to separate robustly SWS and REM epochs and produce results that augmented classical sleep stage knowledge. We also presented results of tracking human subject data through an entire night while incorporating the entire night of sleep information into each individual epoch calculation.

One of the key advantages of representing sleep with the two-dimensional sleep parameter space in this model is that these parameters tie directly to some of the most important biological processes governing sleep. In fact, although this model lacks a thalamic component to drive changes in neuromodulators, successfully mapping sleep into these two parameters may allow noninvasive techniques that could lend insight into the evolution of sleep neuromodulators.

Next, we have shown a quantitative way to map ECoG data from epilepsy patients onto physiologically meaningful parameter states in a stochastic mesoscale model of the cortex to understand better the evolution of individual seizures. We have shown that this quantitative link is possible to make and successful at tracking seizure parameter pathways under certain assumptions. The method is successful at both distinguishing the correct one of two parameter planes and probabilistically tracking a known parameter pathway when ECoG-like data is generated using the mathematical model.

We have also tested our analysis on human subject data, and find consistent trends in

parameter plane likelihood ratios and pathways over each seizure.

Overall we have presented a method to link quantitatively a mesoscale model with experimental measurements using methods that explicitly handle stochastic systems.

6.2 Future directions

The work presented here has contributed important components to linking experimentally measured cortical signals to mathematical models, however there is a wealth of future studies needed to extend and further validate this approach for future practical applications.

Future work will include leveraging the results of this work in a similar probabilistic framework to investigate cortical seizure evolution during sleep as well as investigating the similarities and differences in sleep, anesthesia, and comas. The ability of this meso-scale PDE framework to model key aspects of these different conditions is an important feature in its applicability [42; 77; 79; 90]. Bringing the two-dimensional representation of sleep and the tracking of seizure states together into a single technique would allow these methods to be applied broadly in practice. Additionally other pathologies or sleep research could be investigated using this methodology to gain insight into different dynamics of sleep that may appear under different conditions. This method offers new ways to use noninvasive measurements of the cortex.

In the future, these methods should be thoroughly tested and extended. The ways in which we would suggest testing the results of this method depend on the available experimental techniques. A first pass would be to compare the results of this method across many subjects who have been classified into different seizure categories by their drug efficacy history, known etiology, or other classifications that provide insight into the underlying pathological physiology of seizure evolution. With such a set of classified subjects, one would expect similar underlying behavior from subjects within the same class, and such an experiment could help refine the methodology. The problem with this approach is that such classifications may cover broad categories that encompass multiple types of pathways or influence them in different ways. For instance, one drug may have multiple actions that propagate into different aspects of a meso-scale model.

This should be paired with large studies of the efficacy of certain treatment types or etiology to validate the results with a ground truth. If a certain class of seizures are treated effectively by drug A, we might hope to see similarities in the physiological pathways when compared with seizures that do not respond to that drug. This type of analysis has begun to be explored in [46; 91] for anesthesia drugs.

Ideally, one would be able to measure the physiological changes and create a ground truth of the physiological parameters while simultaneously recording via EEG or ECoG–this has little likelihood of being possible or feasible as an experiment to run, however perhaps partial measurements of the physiology may become possible in animal studies or as scientific equipment advances.

In addition to testing this method and its results in future experiments, the underlying parameter space should be extended to parameter regions of interest to specific problems or hypotheses. One could imagine that a set of perhaps five parameters over certain ranges may be of particular interest in a particular etiology or in a particular animal model. In such cases, this particular region could be investigated in isolation, or compared to multiple other regions. Also, extensions to incorporate sleep parameter regimes and investigate the interactions of sleep and seizures using this model would be an excellent opportunity to leverage this multi-purpose platform. For certain types of seizure, sleep has a significant impact, and literature shows that sleep stage may be playing an important role in determining the onset of seizures [41; 23] and plays a significant role in the failure of seizure prediction algorithms [73]. We hope that the methodology proposed in this work gives insight to researchers tackling these challenging extensions.

.

Bibliography

[1] Acharya UR, Molinari F, Sree SV, Chattopadhyay S, Ng KH, Suri JS (2012) Automated diagnosis of epileptic EEG using entropies. Biomedical Signal Processing and Control 7(4):401–408

[2] Akaike H (1974) A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19(6):716–723

[3] Aradi I, Érdi P (2006) Computational neuropharmacology: Dynamical approaches in drug discovery. Trends in pharmacological sciences 27(5):240–243

[4] Battaglia G, Becker AJ, LoTurco J, Represa A, Baraban SC, Roper SN, Vezzani A, et al (2009) Basic mechanisms of MCD in animal models. Epileptic Disorders 11(3):206–14

[5] Bialer M, White H (2010) Key factors in the discovery and development of new antiepileptic drugs. Nature Reviews Drug Discovery 9(1):68–82

[6] Blümcke I, Thom M, Aronica E, Armstrong DD, Vinters HV, Palmini A, Jacques TS, Avanzini G, Barkovich AJ, Battaglia G, et al (2011) The clinicopathologic spectrum of Focal Cortical Dysplasias: A consensus classification proposed by an ad hoc task force of the ILAE Diagnostic Methods Commission. Epilepsia 52(1):158–174

[7] Bojak I, Liley DTJ (2007) Self-organized 40 Hz synchronization in a physiological theory of EEG. Neurocomputing 70(10):2085–2090

[8] Bonnet M, Carley D, Carskadon M, Easton P, Guilleminault C, Harper R, Hayes B, Hirshkowitz M, Ktonas P, Keenan S, Pressman M, Roehrs T, Smith J, Walsh J, Weber S, Westbrook P (1992) EEG arousals: Scoring rules and examples. Sleep 15(2):173–184

[9] Boon P, Vonck K, Vandekerckhove T, D'have M, Nieuwenhuis L, Michielsen G, Vanbelleghem H, Goethals I, Caemaert J, Calliauw L, Reuck JD (1999) Vagus nerve stimulation for medically refractory epilepsy; efficacy and cost-benefit analysis. Acta Neurochirurgica 141(5):447–453

[10] Borgelt C, Steinbrecher M, Kruse RR (2009) Graphical models: Representations for learning, reasoning and data mining, vol 704. Wiley, Chichester, UK

[11] Brandenberger G, Ehrhart J, Buchheit M (2005) Sleep stage 2: An electroencephalographic, autonomic, and hormonal duality. Sleep 28(12):1535

[12] Brodie MJ, Covanis A, Gil-Nagel A, Lerche H, Perucca E, Sills GJ, White HS (2011) Antiepileptic drug therapy: Does mechanism of action matter? Epilepsy & Behavior 21(4):331–341

[13] Bushey D, Tononi G, Cirelli C (2011) Sleep and synaptic homeostasis: Structural evidence in drosophila. Science 332(6037):1576–1581

[14] Cepeda C, André VM, Levine MS, Salamon N, Miyata H, Vinters HV, Mathern GW (2006) Epileptogenesis in pediatric cortical dysplasia: The dysmature cerebral developmental hypothesis. Epilepsy & Behavior 9(2):219–235

[15] Cepeda C, André VM, Vinters HV, Fisher RS, Levine MS, Mathern GW (2010) Epileptogenesis and cortical dysplasias. In: Atlas of Epilepsies, Springer, pp 353–357

[16] Chisci L, Mavino A, Perferi G, Sciandrone M, Anile C, Colicchio G, Fuggetta F (2010) Real-Time Epileptic Seizure Prediction Using AR Models and Support Vector Machines. Biomedical Engineering, IEEE Transactions on 57(5):1124–1132

[17] Dadok VM, Kirsch HE, Sleigh JW, Lopour BA, Szeri AJ (2013) A probabilistic framework for a physiological representation of dynamically evolving sleep state. Journal of Computational Neuroscience DOI 10.1007/s10827-013-0489-x

[18] Dash MB, Douglas CL, Vyazovskiy VV, Cirelli C, Tononi G (2009) Long-term homeostasis of extracellular glutamate in the rat cerebral cortex across sleep and waking states. The Journal of Neuroscience 29(3):620–629

[19] De Curtis M, Gnatkovsky V (2009) Reevaluating the mechanisms of focal ictogenesis: The role of low-voltage fast activity. Epilepsia 50(12):2514–2525

[20] Elliott RE, Morsi A, Kalhorn SP, Marcus J, Sellin J, Kang M, Silverberg A, Rivera E, Geller E, Carlson C, Devinsky O, Doyle WK (2011) Vagus nerve stimulation in 436 consecutive patients with treatment-resistant epilepsy: Long-term outcomes and predictors of response. Epilepsy & Behavior 20(1):57–63

[21] Englot DJ, Chang EF, Auguste KI (2011) Vagus nerve stimulation for epilepsy: A meta-analysis of efficacy and predictors of response. Journal of Neurosurgery 115(6):1248–1255

[22] Esser SK, Hill S, Tononi G (2009) Breakdown of effective connectivity during slow wave sleep: Investigating the mechanism underlying a cortical gate using large-scale modeling. Journal of neurophysiology 102(4):2096–2111

[23] Foldvary-Schaefer N, Grigg-Damberger M (2006) Sleep and epilepsy: What we know, don't know, and need to know. Journal of Clinical Neurophysiology 23(1):4–20

[24] Foster BL, Bojak I, Liley DTJ (2008) Population based models of cortical drug response: Insights from anaesthesia. Cognitive Neurodynamics 2(4):283–296

[25] Frei MG, Zaveri HP, Arthurs S, Bergey GK, Jouny CC, Lehnertz K, Gotman J, Osorio I, Netoff TI, Freeman WJ, et al (2010) Controversies in epilepsy: debates held during the fourth international workshop on seizure prediction. Epilepsy & Behavior 19(1):4–16

[26] Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19(4):1273–1302

[27] Ghosh-Dastidar S, Adeli H, Dadmehr N (2007) Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. Biomedical Engineering, IEEE Transactions on 54(9):1545–1551

[28] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation 101(23):e215–e220

[29] Good LB, Sabesan S, Marsh ST, Tsakalis K, Treiman D, Iasemidis L (2009) Control of synchronization of brain dynamics leads to control of epileptic seizures in rodents. International Journal of Neural Systems 19(03):173–196

[30] Hamiwka LD, Wirrell EC (2007) Epilepsy in patients with cerebral malformations. Handbook of Clinical Neurology 87:387–407

[31] Hangya B, Tihanyi BT, Entz L, Fabó D, Eróss L, Wittner L, Jakus R, Varga V, Freund TF, Ulbert I (2011) Complex propagation patterns characterize human cortical activity during slow-wave sleep. The Journal of Neuroscience 31(24):8770–8779

[32] Hardie JB, Pearce RA (2006) Active and passive membrane properties and intrinsic kinetics shape synaptic inhibition in hippocampal CA1 pyramidal neurons. The Journal of neuroscience 26(33):8559–8569

[33] Holmes G, Zhao Q (2008) Choosing the correct antiepileptic drugs: From animal studies to the clinic. Pediatric neurology 38(3):151–162

[34] Ihle M, Feldwisch-Drentrup H, Teixeira CA, Witon A, Schelter B, Timmer J, Schulze-Bonhage A (2012) EPILEPSIAE–a european epilepsy database. Computer methods and programs in biomedicine 106(3):127–138

[35] Jobst BC (2010) Electrical stimulation in epilepsy: Vagus nerve and brain stimulation. Current Treatment Options in Neurology 12(5):443–453

[36] Kabat J, Król P (2012) Focal cortical dysplasia–review. Polish Journal of Radiology 77(2):35

[37] Kandel E, Schwartz J, Jessell T (2000) Principles of neural science, vol 4. McGraw-Hill New York

[38] Kaya M, Becker AJ, Gürses C (2012) Blood–brain barrier, epileptogenesis, and treatment strategies in cortical dysplasia. Epilepsia 53(s6):31–36

[39] Kemp B (2009) The Sleep-EDF Database. URL `www.physionet.org/physiobank/database/sleep-edf`, accessed September, 2012

[40] Kemp B, Zwinderman AH, Tuk B, Kamphuisen HA, Oberye JJ (2000) Analysis of a sleep-dependent neuronal feedback loop: The slow-wave microcontinuity of the EEG. Biomedical Engineering, IEEE Transactions on 47(9):1185–1194

[41] Kostopoulos GK (2009) Brain mechanisms linking epilepsy to sleep. In: Schwartzkroin PA (ed) Encyclopedia of basic epilepsy research, Academic Press, pp 1327–1336

[42] Kramer MA, Kirsch HE, Szeri AJ (2005) Pathological pattern formation and cortical propagation of epileptic seizures. Journal of the Royal Society Interface 2(2):113–127

[43] Kramer MA, Szeri AJ, Sleigh JW, Kirsch HE (2007) Mechanisms of seizure propagation in a cortical model. Journal of Computational Neuroscience 22(1):63–80

[44] Kuhlmann L, Burkitt AN, Cook MJ, Fuller K, Grayden DB, Seiderer L, Mareels IMY (2009) Seizure detection using seizure probability estimation: Comparison of features used to detect seizures. Annals of Biomedical Engineering 37(10):2129–2145

[45] Leslie K, Sleigh JW, Paech MJ, Voss L, Lim CW, Sleigh C (2009) Dreaming and electroencephalographic changes during anesthesia maintained with propofol or desflurane. Anesthesiology 111(3):547

[46] Liley DTJ, Bojak I (2005) Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. Journal of Clinical Neurophysiology 22(5):300–313

[47] Liley DTJ, Cadusch PJ, Wright JJ (1999) A continuum theory of electro-cortical activity. Neurocomputing 26:795–800

[48] Liley DTJ, Cadusch PJ, Dafilis MP (2002) A spatially continuous mean field theory of electrocortical activity. Network: Computation in Neural Systems 13(1):67–113

[49] Liu ZW, Faraguna U, Cirelli C, Tononi G, Gao XB (2010) Direct evidence for wake-related increases and sleep-related decreases in synaptic strength in rodent cortex. The Journal of Neuroscience 30(25):8671–8675

[50] Lopour BA, Szeri AJ (2010) A model of feedback control for the charge-balanced suppression of epileptic seizures. Journal of Computational Neuroscience 28(3):375–387

[51] Lopour BA, Tasoglu S, Kirsch HE, Sleigh JW, Szeri AJ (2011) A continuous mapping of sleep states through association of EEG with a mesoscale cortical model. Journal of Computational Neuroscience 30(2):471–487

[52] Lowd D, Domingos P (2005) Naive Bayes models for probability estimation. In: Proceedings of the 22nd international conference on Machine learning, ACM, pp 529–536

[53] MacKay DJC (2010) Information Theory, Inference, and Learning Algorithms. Cambridge University Press

[54] MacKay EC, Sleigh JW, Voss LJ, Barnard JP (2010) Episodic waveforms in the electroencephalogram during general anaesthesia: A study of patterns of response to noxious stimuli. Anaesthesia and Intensive Care 38(1):102–112

[55] Marreiros AC, Daunizeau J, Kiebel SJ, Friston KJ (2008) Population dynamics: Variance and the sigmoid activation function. Neuroimage 42(1):147–157

[56] Moran R, Pinotsis DA, Friston K (2013) Neural masses and fields in dynamic causal modeling. Frontiers in Computational Neuroscience 7(57):1–12

[57] Mormann F, Andrzejak RG, Elger CE, Lehnertz K (2007) Seizure prediction: The long and winding road. Brain 130(2):314

[58] Müller B, Gäbelein WD, Schulz H (2006) A taxonomic analysis of sleep stages. Sleep 29(7):967

[59] Netoff TI, Clewley R, Arno S, Keck T, White JA (2004) Epilepsy in small-world networks. The Journal of Neuroscience 24(37):8075–8083

[60] Nevado-Holgado AJ, Marten F, Richardson MP, Terry JR (2012) Characterising the dynamics of EEG waveforms as the path through parameter space of a neural mass model: Application to epilepsy seizure evolution. Neuroimage 59(3):2374–2392

[61] Olofsen E, Sleigh JW, Dahan A (2008) Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. British Journal of Anaesthesia 101(6):810–821

[62] Palmini A, Najm I, Avanzini G, Babb T, Guerrini R, Foldvary-Schaefer N, Jackson G, Lüders HO, Prayson R, Spreafico R, Vinters HV (2004) Terminology and classification of the cortical dysplasias. Neurology 62(6 suppl 3):S2–S8

[63] Perucca E (2002) Patient-tailored antiepileptic drug therapy: Predicting response to antiepileptic drugs. In: International Congress Series, Elsevier, vol 1244, pp 93–103

[64] Phillips AJ, Robinson PA, Kedziora DJ, Abeysuriya RG (2010) Mammalian sleep dynamics: How diverse features arise from a common physiological framework. PLoS computational biology 6(6):e1000,826

[65] Picot MC, Baldy-Moulinier M, Daurès JP, Dujols P, Crespel A (2008) The prevalence of epilepsy and pharmacoresistant epilepsy in adults: A population-based study in a western European country. Epilepsia 49(7):1230–1238

[66] Ram S, Seirawan H, Kumar SK, Clark GT (2010) Prevalence and impact of sleep disorders and sleep habits in the United States. Sleep and Breathing 14(1):63–70

[67] Rechtschaffen A, Kales A (1968) A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. US Government Printing Office, US Public Health Service, Washington, DC, 204th edn

[68] Riedner BA, Vyazovskiy VV, Huber R, Massimini M, Esser S, Murphy M, Tononi G (2007) Sleep homeostasis and cortical synchronization: III. A high-density EEG study of sleep slow waves in humans. Sleep 30(12):1643

[69] Rish I (2001) An empirical study of the naive Bayes classifier. In: IJCAI 2001 workshop on empirical methods in artificial intelligence, vol 3, pp 41–46

[70] Rodenbeck A, Binder R, Geisler P, Danker-Hopfe H, Lund R, Raschke F, Weeß HG, Schulz H (2006) A review of sleep EEG patterns. Part I: A compilation of amended rules for their visual recognition according to Rechtschaffen and Kales. Somnologie 10(4):159–175

[71] Saper CB, Scammell TE, Lu J (2005) Hypothalamic regulation of sleep and circadian rhythms. Nature 437(7063):1257–1263

[72] Scharfman H (2007) The neurobiology of epilepsy. Current neurology and neuroscience reports 7(4):348–354

[73] Schelter B, Winterhalder M, Maiwald T, Brandt A, Schad A, Timmer J, Schulze-Bonhage A (2006) Do false predictions of seizures depend on the state of vigilance? A report from two seizure-prediction methods and proposed remedies. Epilepsia 47(12):2058–2070

[74] Selvaraj P, Sleigh JW, Freeman WJ, Kirsch HE, Szeri AJ (2013) Open loop optogenetic control of cortical epileptiform activity. Journal of Computational Neuroscience DOI 10.1007/s10827-013-0484-2

[75] Steyn-Ross DA, Steyn-Ross M (eds) (2010) Modeling phase transitions in the brain. Springer, New York

[76] Steyn-Ross DA, Steyn-Ross ML, Sleigh JW, Wilson MT, Gillies IP, Wright JJ (2005) The sleep cycle modelled as a cortical phase transition. Journal of Biological Physics 31(3):547–569

[77] Steyn-Ross ML, Steyn-Ross DA, Sleigh JW, Liley DTJ (1999) Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. Physical Review E 60(6):7299

[78] Steyn-Ross ML, Steyn-Ross DA, Sleigh JW, Whiting DR (2003) Theoretical predictions for spatial covariance of the electroencephalographic signal during the anesthetic-induced phase transition: Increased correlation length and emergence of spatial self-organization. Physical Review E 68(2):021,902

[79] Steyn-Ross ML, Steyn-Ross DA, Sleigh JW (2004) Modelling general anaesthesia as a first-order phase transition in the cortex. Progress in Biophysics & Molecular Biology 85:369–385

[80] Steyn-Ross ML, Steyn-Ross DA, Wilson MT, Sleigh JW (2009) Modeling brain activation patterns for the default and cognitive states. NeuroImage 45(2):298–311

[81] Steyn-Ross ML, Steyn-Ross DA, Sleigh JW (2012) Gap junctions modulate seizures in a mean-field model of general anesthesia for the cortex. Cognitive Neurodynamics 6(3):215–225

[82] Sun FT, Morrell MJ, Wharen Jr RE (2008) Responsive cortical stimulation for the treatment of epilepsy. Neurotherapeutics 5(1):68–74

[83] Terzano MG, Parrino L (2000) Origin and significance of the cyclic alternating pattern (CAP): Review article. Sleep Medicine Reviews 4(1):101–123

[84] Tsimpiris A, Kugiumtzis D (2010) Measures of analysis of time series (MATS): A MATLAB toolkit for computation of multiple measures on time series data bases. Journal of Statistical Software 33(5)

[85] Varotto G, Tassi L, Franceschetti S, Spreafico R, Panzica F (2012) Epileptogenic networks of type II focal cortical dysplasia: A stereo-EEG study. Neuroimage 61(3):591–598

[86] Vaseghi SV (2008) Advanced digital signal processing and noise reduction. Wiley, Chichester, UK

[87] Wang Y, Goodfellow M, Taylor PN, Baier G (2012) Phase space approach for modeling of epileptic dynamics. Physical Review E 85(6):061,918

[88] Wendling F (2008) Computational models of epileptic activity: a bridge between observation and pathophysiological interpretation. Expert Review of Neurotherapeutics 8(6):889

[89] Wilson MT, Steyn-Ross ML, Steyn-Ross DA, Sleigh JW (2005) Predictions and simulations of cortical dynamics during natural sleep using a continuum approach. Physical Review E 72(5):051,910

[90] Wilson MT, Sleigh JW, Steyn-Ross DA, Steyn-Ross ML (2006) General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. Anesthesiology 104(3):588–593

[91] Wilson MT, Steyn-Ross DA, Sleigh JW, Steyn-Ross ML, Wilcocks LC, Gillies IP (2006) The K-complex and slow oscillation in terms of a mean-field cortical model. Journal of Computational Neuroscience 21(3):243–257

[92] Wilson MT, Barry M, Reynolds JN, Crump WP, Steyn-Ross DA, Steyn-Ross ML, Sleigh JW (2010) An analysis of the transitions between down and up states of the cortical slow oscillation under urethane anaesthesia. Journal of Biological Physics 36(3):245–259

[93] Zhang Z, Koifman J, Shin D, Ye H, Florez C, Zhang L, Valiante T, Carlen P (2012) Transition to seizure: Ictal discharge is preceded by exhausted presynaptic GABA release in the hippocampal CA3 region. The Journal of Neuroscience 32(7):2499–2512

Appendix A

Features for Sleep Analysis

A.1 Initial feature set

- Power in the delta, theta, alpha, beta, and gamma frequency ranges, and total power
- Statistical properties: variance, skewness, and kurtosis
- Composite permutation entropy index (CPEI) [61]
- Spindle index [54]
- Power fractions: high power fraction (beta and gamma fraction of total power), and low power fraction (delta and theta fraction of total power)
- Properties of the log power spectrum curve: slope and offset of a fitted line, maximum alpha power above the linear fit, and maximum delta power above the linear fit [45],[51]
- Delta wave steepness properties: the average steepness of delta waves, fraction of epoch covered by delta waves, and combination delta wave steepness. The code for calculating this was developed for this paper by the authors and is included in Appendix C, containing the code of the m-files deltaWaveSteepnessFxn.m and epochFeatsFromWaveSlope.m. However, the importance and meaning of this feature is highlighted in [68].
- Median frequency [84]
- Autocorrelation functions with delay $\tau = 1$ time-step: Spearman autocorrelation, Pearson autocorrelation, and partial autocorrelation [84]
- Bicorrelation with delay $\tau = 1$ time-step [84]
- Mutual information measures including: equidistant mutual information, equiprobable mutual information, and the first minimums of both types of mutual information. Algorithms were from [84], and for these features, with delay $\tau = 1$ time-step

- Hurst exponent [84]
- Hjorth parameters [84]

Note that several of the more complex feature calculation algorithms were implemented using MATLAB code created by other researchers. In particular, many standard time series analysis features were taken from the MATS toolbox [84], and others from supplemental material to papers including [61; 54; 45; 51]. The code sources are indicated in the above listing.

A.2 Feature selection

A.2.1 Feature pruning

The requirements stated in Section 4.4 provide a baseline for creating heuristic scoring functions to prune features and then select a limited number of feature subsets that are likely to perform well. First, the staged patient data is used to split epochs into SWS or REM sleep. The feature selection algorithm is the only time we make use of labeled sleep stage data in this algorithm other than for visualization purposes. Because both the model and patient data contain labels of areas or epochs that are SWS and REM sleep, this is applicable to both types of data and between the two types of data.

Then, with these labeled epochs, the Spearman rank correlation between SWS and REM of each feature was calculated in both the model and the sleep data. That is to say that SWS and REM were converted into a binary variable, respectively 0 and 1. Then the Spearman rank correlation was calculated between this binary variable and the feature values at a particular feature, giving both a correlation and a p-value. One can think of this as a measure of how likely an REM epoch feature value is to be greater than a SWS epoch feature value. If this correlation is close to 1, that means that most REM feature values are larger than SWS feature values. Conversely, a correlation close to -1 means that most SWS feature values are larger than REM feature values. A correlation close to 0 indicates that SWS and REM features values are well mixed in ranking–a typical feature value of a SWS epoch is equally likely to be greater than or less than a typical feature value of an REM epoch. If all SWS features are greater than all REM features, the correlation would be precisely negative one. The Spearman correlation is a non-linear measure of correlation, and only compares the ranking of values.

To remove features that did not vary significantly between SWS and REM in the patient data, all features with less absolute value Spearman correlation than a threshold of 0.35 in all patients were removed. The features removed due to lack of importance were the low power fraction, the skewness, gamma power, median frequency, bicorrelation, and the mutual information first minimums.

Next, using the same Spearman correlations, all remaining features that had opposite signs of correlation coefficient between the model and sleep were removed from the feature

set if absolute value of the model Spearman correlation was greater than .25. This was implemented because a different sign of correlation coefficient is an indication that the model features vary with an opposite trend when compared to patient data. Opposite trends between SWS and REM may demonstrate that the feature is not modeled well with the model, but only if the model shows a clear trend is this significant. Only few features were of opposite sign, and only the power in the β-band and the maximum alpha band power spectral density above a linear approximation exceeded that threshold.

Finally, to choose subsets of features that reduce redundancy, the Pearson correlation (a linear correlation) of features across the model for each state was calculated. Features that highly correlated to one another in the model space were identified by thresholding correlations and identifying features that had over $\rho = .9$ correlation. From this, a connectivity graph of which features were connected by high correlations could be drawn, for example see Figure 4.10 which shows one of the complete subgraphs, in which all features are redundant with each other and thus all features are fully connected to each other. To remove excess redundancy, the feature selection algorithm allowed only one feature from each complete subgraph to be contained in the final feature set. This also reduced the number of possible subsets.

Naive Bayes models also assume conditional independence of features (conditioned on the state), a 'naive' assumption that gives this model its name. We do not expect the naive Bayes model assumptions to be met exactly, but to avoid high correlations, we do not allow feature sets to contain any two features that given the state have an average Pearson correlation coefficient of 0.4.

A test set of all feature subsets from size one to six were generated after pruning features and limiting redundancy according to these rules.

A.2.2 Feature subset cost function analysis

All subsets generated using the rules above were tested against a cost function designed to separate REM and SWS regions. As a significant goal is to distinguish SWS and REM when the EEG epochs are mapped to the sleep parameter space, the average likelihood of each stage of each subject for each feature subset was calculated. Figure 4.11 shows an example of this type of region.

Next, the states in parameter space that contained 95% of the likelihood for average REM likelihood and SWS sleep were each identified. Areas to the left (lower L) and below (lower Δh_e^{rest}) the cusp of the manifold region were identified as SWS regions, while REM regions were identified as regions above the manifold region (higher Δh_e^{rest}). The manifold itself and the top left corner beyond the cusp of the manifold were allowed to be either REM or SWS - this is because the manifold is thought to be a transition region, and while low L and high Δh_e^{rest} is theorized to contain REM, the results in [51] have indicated that perhaps some SWS may lie in these areas.

From these 95% likelihood regions (composed of a set of states), the percentage overlap of REM into SWS-identified areas and SWS into REM-identified areas of the parameter space

were calculated. A cost function composed of the sum of these percentages was minimized at each feature set size. Under visual inspection, the solution for the feature set with four features was chosen.

Appendix B

Features for Seizure Analysis

The likelihood and HMM calculations described in Chapters 3 and 5 assume a fixed set of appropriately chosen features. To find such a set, we start with a large set of dimensionless or normalized features (e.g. log power ratios, dimensionless quantities, etc.), and then use a feature selection algorithm to reduce this to a smaller set of relevant features. This large set included features in the time domain, frequency domain, wavelet features, measures of coherence, and correlation, among others.

Our assumptions at this point are that features can be classified in one of three categories: irrelevant features that do not reflect meaningful information about seizure dynamics, unmodeled features that are not captured by the mathematical model of the cortical tissue within these parameter planes, and relevant features that can be used in this algorithm for tracking.

B.1 Pruning

As a first step, to reduce the run time of the feature search algorithm, the feature set was pruned using a heuristic pruning technique.

Earlier, we made the assumption that the underlying cortical dynamics of seizures evolve through one of the two parameter planes $\mathcal{H}_j$ hypothesized to have produced the seizures. Given this assumption, values of features from human subject seizures must be contained in the range of values produced by the cortical model under these hypothesized planes. Thus, we calculate for each feature over all human subject data the 5% and 95% quantiles, and compare these values to the minimum and maximum of the same feature calculated from all model simulated data. If the feature range identified by these quantiles in human subject data exceeds the range over all model simulations within a parameter plane, this feature is pruned and considered to be capturing noise or unmodeled dynamics. The 5% and 95% quantiles are used for human subject data rather than the full range to allow for artifacts and noise in the data.

B.2 Cost function

Next, we use a greedy feature selection algorithm to minimize a cost function, shown in Equation B.1 with components explained below. Roughly speaking, we penalize three different things: (1) classification error in the human subject data, to make sure seizure and non-seizure regions are separable in the data; (2) on-average inconsistent localization of seizure and non-seizure likelihoods in parameter planes; (3) number of features included in the feature set. As an additional step, due to the Naive Bayes assumptions of independence, at each greedy step, we only consider features that have an absolute value correlation coefficient less than 0.4 conditioned on the state.

$$J(\mathbf{e}_{sub}^{\mathcal{T}}, \mathcal{L}(\cdot), N_f) = C_{Loc} + \lambda_f N_f^2 \tag{B.1}$$

The first component C_{Loc} function penalizes the amount of normalized average likelihood of human subject seizure epochs that maps into non-seizure regions of the parameter planes and likewise penalizes the fraction of average likelihood of non-seizure epochs that lies in the seizure-regions of the parameter planes. As this criterion involves both simulated data and human subject data, its goal is to pick features that are well-captured by this model and exclude features that are not. It also inherently penalizes classification errors in the human subject data as separability in the likelihood regions is an indication of separability in the human subject data alone.

This C_{Loc} is a criterion based on the average likelihood regions of seizure and non-seizure subject data epochs mapped onto each hypothesis plane. Z_j is a matrix over all states in the $\mathcal{H}_j$ parameter plane that is 1 at states of the model plane that contain only seizures, 0 on states that contain only non-seizures, and the fraction of seizure epochs on states that contained both seizure and non-seizure epochs. $\mathcal{N}_{j,sz}$ is the average likelihood of all seizure points normalized so that it sums to one, shown in Equation B.3. For speed, in these feature calculations, instead of a likelihood calculation based on a kernel estimate of the probability distribution of each feature, we used a simpler probability distribution function of an 8-bin histogram with add-one Laplace smoothing.

$$C_{Loc} = \frac{1}{2} \sum_{\mathcal{H}_j} \sum_{s_k} ((1 - Z_j)\mathcal{N}_{j,sz} + Z_j \mathcal{N}_{j,non-sz}) \tag{B.2}$$

$$\mathcal{N}_{j,sz} = \frac{\sum_{t_{sz}} \mathcal{L}_j(\mathbf{e}^{t_{sz}})}{\sum_{s_k} \sum_{t_{sz}} \mathcal{L}_j(\mathbf{e}^{t_{sz}})} \tag{B.3}$$

Finally, we add a regularization cost on the final number of features via $\lambda_f N_f$, where N_f is the number of features and λ_f=0.001.

Appendix C

MATLAB Code

Here we present two MATLAB m-functions that were used to calculate the delta wave steepness feature used in the sleep analysis of Chapter 4.

The function deltaWaveSteepnessFxn.m is used to process raw EEG data and produce measures of steepness of individual delta waves. The function epochFeatsFromWaveSlope.m is used to parse this delta-wave information into features over epochs.

C.1 deltaWaveSteepnessFxn.m

```
function [ amtUpslopes, amtDownslopes, iUps, iDowns] = ...
    deltaWaveSteepnessFxn(data, sRate, paramIn)
% [ amtUpslopes, amtDownslopes, iUps, iDowns] = ...
% deltaWaveSteepness(data, sRate, varargin)
% Created: 11-13-2012
% Author: Vera Dadok
%
% INPUT:    data: EEG data in row-form
%           sRate: sampling rate of EEG data
%           paramIn: optional structure with fields to set parameters
%                deltaRange - range in Hz of waves to capture
%                nStdDev - number of standard deviations to count wave
%                filterOrder - order of Butterworth filter
%
% OUTPUT:   amtUpslopes:  average steepness of up-slopes of delta waves
%         amtDownslopes: average steepness of down-slopes of delta waves
%                     iUps: indices of up-slopes of delta waves
%                     iDowns: indices of down-slopes of delta waves
%
```

```
%% DETAILS OF METHOD from  this code:
%
% Summary of algorithm:
% (1) Find delta waves by filtering to delta band only
% (2) Identify peaks and valleys of filtered data
% (3) Identify adjacent peaks and valleys with which to calculate slopes.
% (3) Calculate average slope between peaks and valleys within filtered
%     data (Note: in the future you could alter this function to return to
%     the original data and use peaks and valleys on the original data,
%     however that may result in a noisier feature).

%% Parameters Setup:
% Defaults:
par.deltaRange = [.5,3];  %range of frequencies considered a delta wave.
par.filterOrder = 4;     %filter order
par.nStdDev = .5;  %number of standard deviations that
                   %defines the threshold for a 'delta wave'

if nargin>2
    parFields = fieldnames(par);
    for k=1:length(parFields)
        if isfield(paramIn, parFields{k})
            par.(parFields{k}) = paramIn.(parFields{k});
        end
    end
end

%Note: could change this to different fraction of a wavelength.
waveThreshold = 0.75/min(par.deltaRange);
%max length in seconds allowed for a single delta wave.

%% Check inputs for validity:

% Make sure data is a row vector
[a,b] = size(data);
if b==1
    data = data';
end
```

```
%% (1) Filter Signal to delta band:
 %BUTTERWORTH FILTER:
    deltaRange = par.deltaRange;
    deltafreqsFilter = [deltaRange(1)/sRate,deltaRange(2)/sRate]*2;
    [B,A] = butter(par.filterOrder,deltafreqsFilter);
    nStdDev= par.nStdDev;

    %filtered EEG within the  frequency range
    filtData = filtfilt(B,A,data);

%% (2) Find peaks
    minPeakDistance = 1;

    %Find potential peaks/valleys within filtered data:
        [~,iPeaksA] = findpeaks(filtData, ...
            'Minpeakdistance', minPeakDistance);
        [~, iValleysA] = findpeaks(-filtData, ...
            'Minpeakdistance', minPeakDistance);

    %Throw out false peaks that are insignificant

    %OPTION A: use Hilbert amplitude (may be better, but slower)
        %Hilbert transform of filtered data
        hilbData= hilbert(filtData);
        %amplitude of data
        ampData = abs(hilbData);
        stdAmp = std(ampData);
        itmppk = find(ampData(iPeaksA)>nStdDev*stdAmp);
        itmpval = find(ampData(iValleysA)>nStdDev*stdAmp);
        iPeaksB = iPeaksA(itmppk);
        iValleysB = iValleysA(itmpval);

%% Part 1c: find sloped points:
    %Upslopes:
        [iBottoms, iTops] = findUpSlopes(iValleysB, iPeaksB,sRate,...
         waveThreshold);
    %Downslopes:  (these are indices of edges)
        [iBottoms2, iTops2] = findUpSlopes( iPeaksB,iValleysB,...
        sRate, waveThreshold);

%% Part 2: Calculate upslope and downslopes
```

```
%     of everything in delta filtered set.

    %Using sRate converts to mV/s
        amtUpslopes = ...
            calculateSlopeAmounts(iBottoms, iTops, filtData);
        amtUpslopes = amtUpslopes*sRate;
        amtDownslopes = ...
            calculateSlopeAmounts(iBottoms2, iTops2, filtData);
        amtDownslopes = amtDownslopes*sRate;

%% Prep for output:
    iUps = [iBottoms; iTops];
    iDowns=[iBottoms2; iTops2];

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%% ------- HELPER FUNCTIONS ------ %%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function amtSlopes = calculateSlopeAmounts(iStarts, iStops,
    timeSeriesData)
  %[amtSlopes] = ...
  %    calculateSlopeAmounts(iStarts, iStops, timeSeriesData)
  % This function calculates slopes between iStarts and iStops
  % for timeSeriesData, returning a list of slopes.
      amtSlopes= (timeSeriesData(iStops) ...
           -timeSeriesData(iStarts))./(iStops-iStarts);
end

function [iBottoms, iTops] =
     findUpSlopes(iMins, iMaxes, sRate, waveThreshold)
 %[iBottoms, iTops] = findUpSlopes(iMins, iMaxes, sRate, waveThreshold)
 % This function finds adjacent pairs of wave edges that define
 % the boundaries of a min to a max
 % (so edges of a delta wave sloping up or down)
 % waveThreshold default is 2;due to  max 2 seconds per wave.  .5 HZ.

% Merge and sort the two, keeping track of which are mins and maxes. Then
% find the min and max neighbors using a flag.
    flags = [0*ones(1,length(iMins)), ones(1,length(iMaxes))];
    [~, iSortAll] = sort([iMins, iMaxes]);
    sortFlags = flags(iSortAll);
```

```
    iiBottomsTest=find(diff(sortFlags)>0);
    iiTopsTest=iiBottomsTest+1;
    iBottoms = iMins(iSortAll(iiBottomsTest));
    iTops = iMaxes(iSortAll(iiTopsTest)-length(iMins));

  %Now waves that are longer than waveThreshold:
    iWaveLengths = iTops-iBottoms;
    iTooLong = find(iWaveLengths/sRate>waveThreshold);
    iTops(iTooLong) =[];
    iBottoms(iTooLong) =[];
end
```

C.2 epochFeatsFromWaveSlope.m

The feature used in the feature set of Chapter 4 is the third output of this function (avgWaveSlopeWeighted).

```
function [ avgWaveSlopeRaw,amtWaveCoverage, avgWaveSlopeWeighted ] =...
    epochFeatsFromWaveSlope( data, samplingRate, ptsPerEpoch, ...
    ptsInOverlap, paramIn  )
%[ avgWaveSlopeRaw,amtWaveCoverage, avgWaveSlopeWeighted ] =...
%    epochFeatsFromDeltaWaveSteepOut( data, samplingRate, ptsPerEpoch, ...
%    ptsInOverlap, paramIn  )
% This function calculates 3 features of delta-wave steepness on each epoch
% over a segmented time series, using deltaWaveSteepnessFxn.m.
%
% INPUTS:   data -  time series vector
%           samplingRate - sampling rate in Hz
%           ptsPerEpoch - number of points in each Epoch
%           ptsInOverlap - number of points to overlap each window by
%           paramIn - optional input structure that can be passed to
%               deltaWaveSteepnessFxn.m, see that function for details
%                deltaRange - range in Hz of waves to capture
%                nStdDev - number of standard deviations to count wave
%                filterOrder - order of Butterworth filter%
%
% OUTPUTS:  avgWaveSlopeRaw - average slope of waves. If no waves are present
%                   this will be zero.
%           amtWaveCoverage - amount covered (fraction) by  sloping parts
%                   of waves.
%           avgWaveSlopeWeighted - average slope up in wave weighted by
%                   fraction of epoch covered by waves.
%
```

```
% Created: November 2012
% Author:  Vera Dadok
%
% See Also: deltaWaveSteepnessFxn

%% Options:
   avgFunctionHandle = @median;

%% Setup:
    [a,b] = size(data);
    if b==1;
        data=data';
        [a,b] = size(data);
    end
    iEpochsStarts =  1:(ptsPerEpoch-ptsInOverlap):(b-ptsPerEpoch+1);
    nEpochs = length(iEpochsStarts);

    if (nargin<5)
       paramIn=struct();
    end
%% Calculate full data epochs delta waves:
    [ amtUpslopes, amtDownslopes, iUps, iDowns] = ...
        deltaWaveSteepnessFxn(data,samplingRate, paramIn);

%% Get information on each epoch
%Features we'll calculate:
    avgDeltaUp = zeros(nEpochs, 1);
    avgDeltaUpB = zeros(nEpochs, 1);
    amtWaveCoverage = zeros(nEpochs, 1);
    avgDeltaDown = zeros(nEpochs, 1);
    avgWaveSlopeRaw = zeros(nEpochs, 1);
    avgWaveSlopeWeighted = zeros(nEpochs, 1);
%Do the calculations:
    for k=1:nEpochs
        j=iEpochsStarts(k);
        iStart = j;
        iStop = (j+ptsPerEpoch-1);

        %Find waves to consider: for now any wave that begins
        % or ends in the epoch as a wave to use.
        iWaves=find(((iUps(1,:)>iStart) & (iUps(1,:)<iStop))|...
            ((iUps(2,:)>iStart) & (iUps(2,:)<iStop)));
        iWavesDown = find(((iDowns(1,:)>iStart) & (iDowns(1,:)<iStop))|...
            ((iDowns(2,:)>iStart) & (iDowns(2,:)<iStop)));
```

```
        if ~isempty(iWaves)&&~isempty(iWavesDown)
            avgDeltaUp(k) = avgFunctionHandle(amtUpslopes(iWaves));
            avgDeltaDown(k) = avgFunctionHandle(amtDownslopes(iWavesDown));
            avgWaveSlopeRaw(k) = (avgDeltaUp(k)+abs(avgDeltaDown(k)))/2;

            %Get amount of epoch covered by delta-waves.
            %(Future: make sure up and down waves don't ever overlap)
            nCoverUpWaves = sum(min(iUps(2,iWaves), iStop)...
                -max(iUps(1,iWaves), iStart));
            nCoverDownWaves = sum(min(iDowns(2,iWavesDown), iStop)...
                -max(iDowns(1,iWavesDown), iStart));
            amtWaveCoverage(k) = (nCoverUpWaves+ nCoverDownWaves)/...
                ptsPerEpoch;

            %Weight the wave slope by the amount coverage:
            avgWaveSlopeWeighted(k) = amtWaveCoverage(k)*avgWaveSlopeRaw(k);
            avgDeltaUpB(k) = amtWaveCoverage(k)*avgDeltaUp(k);
        else
            if isempty(iWaves)
                avgDeltaUp(k) = 0;
            else
                avgDeltaUp(k) = avgFunctionHandle(amtUpslopes(iWaves));
                avgWaveSlopeRaw(k)  = avgDeltaUp(k);
                nCoverUpWaves = sum(min(iUps(2,iWaves), iStop)...
                    - max(iUps(1,iWaves), iStart));
                amtWaveCoverage(k) = (nCoverUpWaves)/ptsPerEpoch;
                avgWaveSlopeWeighted(k) = amtWaveCoverage(k)*avgWaveSlopeRaw(k);
            end
            if isempty(iWavesDown)
                avgDeltaDown(k) = 0;
            else
                avgDeltaDown(k) = avgFunctionHandle(amtDownslopes(iWavesDown));
                avgWaveSlopeRaw(k)  = abs(avgDeltaDown(k));
                nCoverDownWaves = sum(min(iDowns(2,iWavesDown), iStop)...
                    -max(iDowns(1,iWavesDown), iStart));
                amtWaveCoverage(k) = (nCoverDownWaves)/ptsPerEpoch;
                avgWaveSlopeWeighted(k) = amtWaveCoverage(k)*avgWaveSlopeRaw(k);
            end
        end
    end
end
```